

Materials' Revolution: Computational Design and Discovery of Novel Materials

Progress Report Year 2 February 2015 - January 2016

FNSNF

Fonds national suisse Schweizerischer Nationalfonds Fondo nazionale svizzero Swiss National Science Foundation Les Pôles de recherche nationaux (PRN) sont un instrument d'encouragement du Fonds national suisse Die Nationalen Forschungsschwerpunkte (NFS) sind ein Förderungsinstrument des Schweizerischen Nationalfonds I Poli di ricerca nazionali (PRN) sono un mezzo di sostegno del Fondo Nazionale Svizzero The National Centres of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation

La BiblioTeca de Babel

ozon de van Tilación anal madio,

disolvers an al viento angendrado

By This art you may contemplate The varia. Tion of The 23 let ars ... The Anatomy of Melanchely, part. 2, sect. I mam. IV

bATAN DAD ballidiman

Haber

El universo (que atros llaman la Biblioteca) se compone de un mumero indatinido, y Tal vez infinito, de galerías hexagonales, feor diverses con vastos

By this art you may contemplate the varia tions of the 23 letters... *The Anatomy of Melancholy, part 2, sect. II, mem. IV*

200 RODATAS

El universo (que otros llaman la Biblioteca) se compone de un número indefinido, y tal vez infinito de galerías hexagonales, con vastos pozos de ventilación en el medio, cercados por barandas bajísimas. Desde cualquier hexágono se ven los pisos inferiores y superiores: interminablemente. La distribución de las galerías es invariable. Veinte anaqueles, a cinco largos anaqueles por lado, cubren todos los lados menos dos; su altura, que es la de los pisos, excede apenas la de un bibliotecario normal. Una de las caras libres da a un angosto zaguán, que desemboca en otra galería, idéntica a la primera y a todas. A izquierda y a derecha del zaguán hay dos gabinetes minúsculos. Uno permite dormir de pie; otro, satisfacer las necesidades finales. man a .



www.snf.ch Wildhainweg 3, P.O. Box 8232, CH-3001 Berne

Programmes division National Centres of Competence in Research

NCCR: 2nd Progress Report - Cover Sheet

Title of the NCCR	Materials' Revolution: Computational Design and Discovery of Novel Materials (MARVEL)	
NCCR Director Name, first name Institution	Marzari, Nicola EPFL, Lausanne	
Date of submission	February 23, 2016	

1.	Executive summary		[see explanations: #1]		
2.	. Reaction to the recommendations of the review panel			[2]	
3.	Management				
	3.1 S	Structure and organisation of	the NCCR	[3]	
	3.2 N	lanagement activities and sta	atus of collaboration / integration	[4]	
4.	Research				
	4.1	Results since the last progre	ess report	[5]	
	4.2	New Projects		[6]	
5.	. Knowledge and technology transfer				
6.	Education and training [8				
7.	. Equal opportunities / Advancement of women			[9]	
8.	. Communication			[10]	
9.	. Structural aspects				
10.	Lists				
	10.1	Overview of personnel and f	unctions	[12]	
	10.2	Education and training		[12]	
	10.3	Publications		[12]	
	10.4	Lectures at congresses etc.		[12]	
	10.5 Services, patents, licences, start-up companies etc.			[12]	
	10.6	Cooperation with third parti	ies	[12]	
11.	Statis	tical output data		[12]	
12.	Finan	ce			
	12.1	Finance NCCR Overview	(NIRA Report Nr. 1010)	[13]	
	12.2	Funding sources	(NIRA Report Nr. 1020)	[13]	
	12.3	Allocation to projects	(NIRA Report Nr. 1040)	[13]	
	12.4	Expenditures	(NIRA Report Nr. 1030)	[13]	
	12.5	Comments on finances		[14]	

Contents

1	Executive summary	3	
2	Reaction to the recommendations of the review panel		
3	Management3.1Structure and organisation of the NCCR3.2Management activities and status of collaboration/integration	15 15 19	
4	Research 4.1 Results since the last progress report Vertical Project 1	21 21 38 74 91 119 132 146 167	
5	Knowledge and technology transfer	169	
6	Education and training	177	
7	Equal opportunities/Advancement of women	181	
8	Communication	189	
9	Structural aspects		
10	Lists 10.3 Publications	199 199	

1 Executive summary

NCCR MARVEL has an official start date of May 2014; the contract with the SNSF was signed in July 2014, and groups and PIs started ramping up efforts and hiring personnel in the fall of 2014. So, as we write this, the project has been ongoing for 18 months, with its structure solidly in place, and groups participating to an overall effort dedicated to the different aspects of materials design and discovery.

During year 2 of MARVEL

- a number of high-profile papers directly driven from the Centre have started appearing, most notably in *Nature* on Weyl semimetals, and in *Nature Materials* on topological insulators the full list is available in chapter 10.3;
- numerous recognitions have been bestowed upon the PIs, including the 2015 Körber European Science prize to Nicola Spaldin, the 2015 EuCheMS Lecture Award to Ursula Röthlisberger, the 2016 APS Aneesur Rahman Prize for Computational Physics to Matthias Troyer, the 2016 APS Frank Isakson Prize for Optical Effects in Solids to Dirk van der Marel, a 2014 ERC Advanced Grant to Berend Smit, and a 2015 ERC Starting Grant to Michele Ceriotti; Alessandro Curioni has been named vice president and director of the IBM Research Lab in Zurich, Switzerland (first time that the position goes to a computational person) and Ursula Röthlisberger has been inducted to the International Academy of Quantum Molecular Sciences;
- the "Experiments" Platform Project PP7 was launched, and has seen a very dedicated effort to engage the experimental community, with two 1-day workshops at PSI and at Empa, two rounds of calls for experimental projects, and its first project meeting to review ongoing and starting projects;
- the "Informatics" Platform Project PP6 has

established close working relations between the MARVEL team and the CSCS personnel, with monthly meetings to work on the different challenges of managing high-throughput calculations on a supercomputer centre, on acting as a public data repository, and on federating data across centres; the effort on data has also been singled-out and praised by the Scientific Advisory Board at the Review and Retreat meeting of September 2015;

• a number of verification or validation efforts have started — one, part of a broader collaboration involving also MARVEL members, slated to appear in *Science* in the spring, and one centered on the electronic-structure of challenging transition-metal complexes.

Given the very active coordination efforts in the two PP6 and PP7 Platform Projects, the brainstorming and planning for the NCCR needs to focus onto the five Vertical and Horizontal Projects. The risk is that these become comfortable containers of appropriately themed research, but do not really push the field with new ideas, tools, or collaborations - and the entire NCCR needs and will always benefit from research that pushes the PIs to take up new challenges besides those of their already established and very successful independent research directions. In this respect, the establishment of clear metrics for success, discussed in the response to the recommendations of the Review Panel, is an important step, but it is essential to keep a vigilant eye and find ways to reward projects that take up the challenge and the metrics discussed, while streamlining other efforts.

Highlights from the seven projects are listed below, as contributed by the individual project leaders, followed by a summary of the efforts taking place in the four structural areas, that are then listed in more detail in Sections 5 to 8.

Research

Vertical Project 1: Novel Materials Physics. In this project we focus on the design of materials in which the fundamental physics and resulting functionalities are determined by strong interactions between individual electrons, or by new and emerging phenomena, such as topological properties of their band structures. The challenge here is the development of methods that can later be used in high-throughput material design. In the second year we have continued developing the Wannier-based Z2PACK software package, used it to identify a new type of Weyl semimetals, and started a high-throughput search for more topological materials that turned up first candidates. This approach has already resulted in a theoretical prediction followed by experimental validation of a novel topological insulator phase in Bi₄I₄, while several Weyl semimetals candidates have been identified. We have continued the integration of DMFT methods (for accurate description of strong correlations) with DFT (for description of crystal structure and chemistry) and applied the combined methodology in understanding the trends in properties across the rare-earth nickelate series, and in designing a new material which is described by a single-band Hubbard Hamiltonian.

Vertical Project 2: Novel Materials Applications. During the second year of MARVEL, we have streamlined the internal structure of VP2 to focus its research activities. In particular, we have defined three research themes that should guarantee a good balance between targeted computational design and broader scope discovery. Among the important achievements of the second year are a joint experimental/theoretical study that identified the molecular origin of J - V hysteresis effects in halide perovskites (Röthlisberger), the investigations of water splitting catalysts via a mechanistic characterization of the reaction path for cobaltcubane based oxidation catalysts (Hutter) and the establishment of linear scaling relationships and volcano plots in homogenous catalysis (Corminboeuf). Work in progress involves high-throughput screening of candidate materials for next-generation solid-state lithium conductors, in a joint collaboration between the Marzari group and IBM, perovskites for photocatalytic water splitting (Marzari), and exfoliation of novel 2D materials (Marzari). Important technical milestones have been the development of a standard solid-state pseudopotential library (Marzari) and the development of an efficient scheme for accurate GW calculations of band gaps (Pasquarello).

Horizontal Project 3: Advanced Quantum Simulations. We developed efficient continuous-time quantum Monte Carlo methods for correlated lattice fermions and found ways to mitigate fermion sign problem in quantum cluster simulations. We implemented an extended DMFT method and combined it with a finite temperature Matsubara axis GW code in order to be able to perform full self-consistent GW+EDMFT calculations. We created a library for electrochemical simulations based on Poisson-(Boltzmann) solvers and implemented analytical gradients for RI-MP2 methods and applied them in molecular dynamics studies of water and solvated electrons. We assembled a library of molecules based on Co water oxidation catalysts for the assessment and verification of DFT and DMFT methods.

Horizontal Project 4: Advanced Sampling Methods. Significant progress has been made in predicting new materials, characterizing them by properties of their energy landscape, machine learning for predicting their properties and the understanding of the growth and nucleation of materials. All these topics are closely related and require powerful methods to sampling the potential and free energy surface. All these methods profit from certain descriptors that are also called fingerprints or collective variables that can quantity the similarity or dissimilarity of different structures.

Horizontal Project 5: Materials Informatics. The work in Horizontal Project 5 focused on the design of big-data system, merging the expertise on cognitive, neural networks and machine learning, to predict novel materials connected to the MARVEL activities and on the development of domain-specific language frameworks that will allow the easy development of high-performance applications in the context of MARVEL. In particular, we aim at complementing in a synergistic way recent developments on machine learning and neural networks with cognitive computing to support the activity of the vertical projects and help in either predicting unknown properties of materials or even construct new ones based on all the properties of similar class of materials known to us at the present day. In particular, in the second year we were able to develop the framework for a material discovery system that is built around a knowledge graph that can support complex queries such as semantic search. In addition, we developed advanced



machine learning and statistical inference algorithms that are crucial to populate the knowledge graph and to allow prediction of properties. We stress that in year 2 we initiated a deep collaboration with VP2 to showcase our approach in the context of novel materials for batteries. Finally, work towards (materials) domain specific languages has quite progressed at the foundational level to now allow practical instantiations.

Platform Project 6: Informatics. The Informatics Platform is focused on providing the hardware and software infrastructure for the activities of MARVEL, and on populating and disseminating materials data to the scientific community at large. On the hardware side, we have completed the procurement and purchase of high-performance computing to support the numerical efforts of all the groups, with 180 dual-socket nodes of a Cray XC40. We have put in place an agreement with CSCS to ramp-up Petabyte-class backed-up storage to cover both the internal research activities in data-mining and machine learning, and to act as a public repository of data and workflows. On the software side, sustained development of the materials' informatics platform AiiDA (www.aiida.net) has continued, to support high-throughput calculations, data preservation and storage, scientific workflows, and public dissemination. AiiDA is slated for 1.0 release in the summer of 2016, and several training workshops have already taken place (4) or are planned (4). Work on the Materials Cloud dissemination platform, its integration with AiiDA, and data service federation has also started, as has work on the SIRIUS domain-specific library, for hybrid CPU-GPU implementations of all-electron (Exciting) and pseudopotential codes (Quantum-ESPRESSO).

Platform Project 7: Experiments. Turning the results of the computational approach into real materials requires the intense collaboration between theory and experimental groups. The PP7 is the linking point and provides funding for experimental verification projects. Two calls for proposals have been launched. In the first call, which was launched at the end of 2014, eight proposals have been granted and six of them have started. The second call was lunched at the end of 2015 and six proposals have been granted. These projects cover different topics of VP1 and VP2 and have all an experimental and a theoretical group leaders. First results include the successful highpressure synthesis of powder samples of nickelates and the comparison of high-throughput theoretical calculations with measured crystalline structure of an oxynitride photocatalysts leading to an input for the refinement of the model.

Other aspects

A sustained effort in the four structural areas has taken place, with

- in KTT, direct support or organization of 10+ conferences in the field, very active interactions with the experimental community at PSI and Empa, efforts in open-source codes such as CP2K, BIGDFT, Quantum-ESPRESSO, PLUMED, and AiiDA, data dissemination via the Materials Cloud, pseudopotential verification, and a focused tech transfer activity led by Pascale Van Landuyt (from EPFL's industrial liaison office), that dedicates 20% of her time to MARVEL to proactively work on providing to interested companies a connection to the different research portfolios within the project;
- in education, the first MARVEL Junior Retreat has been organized, four MAR-VEL Distinguished Lectures were held and recorded, MARVEL students have been supported to conferences abroad, and our long-term engagement for science in developing countries though our involvement in the African School series on Electronic Structure Methods and Applications (ASESMA);
- in equal opportunities, many activities have taken place in collaboration with the relevant offices at EPFL, including workshops and summer camps for girls (7-10 years old, and 11-13 years old) and mixed camps for girls and boys; three "professional skills development" workshops for women, led by Nancy Houfek, where held at ETHZ and EPFL, open to all NCCRs, and saw more than 70 participants from the PhD to the professorial level;
- in communication, our website was launched, the second Review and Retreat took place in September, together with the meeting of the Scientific Advisory Board, and several project meetings started; we actively participated in several open days efforts, at EPFL, USI, and CSCS, and organized a visualization contest at EPFL that will become now a regular occurrence.

2 Reaction to the recommendations of the Review Panel

The Review Panel met at the first site visit, on April 22 and 23, 2015, and prepared a report that was approved and communicated on June 10, 2015. This is discussed in detail in the following. In addition, the official letter accompanying the report of the Review Panel mentioned that the Division 4 of the SNSF Research Council endorses the establishment of a focussing strategy for the future and is also very interested to see how the KTT activities will develop in the coming years. The Research Council is pleased to see that the structural development of the NCCR is progressing as planned. With respect to the infrastructure, most of the promised equipment is already fully functional. Two new assistant professors are already in place; the recruitment of the third PATT will follow.

The report from the Panel was divided into six sections: General impression and status of integration, Comments on the projects, Structure related criteria, Structural aspects and support by the Home Institution, Recommendations, and Next site visit, that will be addressed in the following.

General impression and status of integration

Key statements from this section (for completeness please refer to the original document) mentioned that

- The panel is unanimously impressed by the manifold aspects of the NCCR MARVEL, mainly by its overall scientific strength and the coherent vision of the participating researchers.
- The panel appreciates that real synergy between the different individual projects is visible.
- The panel is interested to see some metrics of success in the coming years. The NCCR should define itself, which achievements will decide on success and failure of the research endeavor (define milestones).
- The panel understands that all PIs of MAR-VEL will meet once a year ("MARVEL Review

and Retreat"). This may be too little to encourage and foster more collaborations and synergies.

- For the future development of the NCCR, focusing will be inevitable. The NCCR should play to the strengths of the Swiss community when setting the priorities. One aspect to be taken into account may be the high number of excellent method developers in Switzerland. Another strategy could be to focus on the projects that are sufficiently integrated into the NCCR.
- The agility grants already prove to be a very interesting tool for the development of the program.

We thank the Review Panel for the supporting statements made. We took at heart the issues of 1) focusing the projects, 2) building, when possible, new synergies, and 3) defining metrics and achievements. The different project leaders have been called to the task, so we present here in order the response to the comments on the individual projects, and later, in the Recommendation section, our response to the overall set of recommendations made.

Comments on the projects

VP1 - Novel Materials Physics: VP1 is built up around the two themes topological insulators and strongly correlated materials. The panel thinks that these two topics may not necessarily need to remain separate; they can come closer and be partly merged over the years, e.g. by exploring the strongly correlated materials that are topological insulators. In VP1 there is some high throughput work on the discovery of topological insulators. It is not entirely clear for the panel how this differs from work done elsewhere in this area. The NCCR must show what is unique in their effort and how they differentiate themselves from work done by others. The panel sees this project as a possible testbed for the machine learning efforts. They see, for example, the plot of the metal-insulator transition as a function of the perovskite tolerance factor.

Reaction to the recommendations of the Review Panel

MARVEL

The Panel asked us to clarify how our work on high throughput screening of topological materials differs from other efforts world wide. Our differentiating point is that through the development of Z2PACK we have a unique tool to accurately and directly calculate all topological invariants from first principles.

VP2 - Novel Materials Applications: The panel discussed aspects like the water-sensitivity of oxynitride catalysts or the CO_2 capture and its viability. Generally, the members think that an early dialogue with industrial partners will inform the PIs on the needs and constraints of industrial applications. The idea is not to fully restrict the activities in VP2 to the requirements of industry, but the panel is convinced that the awareness of the industrial situation right from the beginning will help to make later KTT activities easier.

Following the advice of the Review Panel, the research activities in VP2 have been sharpened. A new organization has been set up around three major themes, namely light harvesting, energy storage, and search of novel materials (Fig. 1). On January 29, 2016, a meeting targeted on VP2 has been organized assembling all researchers involved in VP2, including also some researchers active in horizontal projects showing great affinity with the themes of VP2. The meeting has been structured into three major sessions corresponding to the identified VP2 themes. The participants have been asked to split their research activity according to the themes of these sessions and to give corresponding presentations in each of the concerned sessions. This organization has helped to orient the discussion according to the selected themes rather than on the individual work of every group leader. In this way, common problems could be focussed upon and targeted actions towards their solution could be stimulated. More generally, this VP2 meeting brought a higher coherence between the various VP2 projects. Furthermore, it fostered collaborations and synergies in well defined directions.

As for the individual comments to VP2, the Review Panel questions the viability of some of the materials under investigation for water splitting (e.g. oxynitride catalysts) and CO_2 capture, and recommends an early dialogue with industrial partners to identify constraints for industrial applications. The realization of industrial applications is one of the desirable goals of the present project. However, in this early stage, the choice of materials in most VP2 projects is rather guided by the necessity of



Figure 1: New organization in VP2 showing how the activities of the various PIs distribute among the three principal themes of VP2: light harvesting, energy storage, and search of novel materials.

acquiring fundamental understanding and to set accuracy benchmarks. Nevertheless, early interactions with industry already occur for some projects in VP2, such as those focussing on batteries (Curioni, Marzari) and on carbon nanoribbons (Passerone). For other projects, such as for the project on semiconducting polymers (Corminboeuf), we are in the process of contacting industrial partners.

HP3 - Advanced Quantum Simulations: The different parts still seem relatively independent. They deal with complementary techniques, but there may be opportunities to integrate them. The links to the vertical projects, in particular to VP1, can be strengthened and made more explicit.

We highlight these three actions that we have taken:

- The number of collaborations within VP2 has been increased. A first collaboration between groups with DMFT and DFT has been established (Werner, Hutter).
- The number of collaborations with other MARVEL projects has been increased and intensified. See for example the collaborations with VP2 and HP4. Collaborations with VP1 are ongoing and successful, see for example the Troyer-Yazyev and Troyer-Spaldin collaborations.
- Library of test systems has been established, see below.

In addition, HP3 provided also a reply to the comment from the Scientific Advisory Board, that met on September 3 and 4, 2015 and commented on the methodological efforts by saying that

Methods: the development of this activity is as well under way and focused on outstanding open problems in computational materials science. Enhanced connections and comparisons between the different codes used within the center should be pursued, and verification and validation activities should be perhaps enhanced. One opportunity that we saw is that of actively pursuing comparisons and connections between DFT and DMFT codes, which at present are rather separate efforts. It is perhaps worth defining a challenging problem to which both DFT and DMFT, as well as other methods under development, could contribute in a synergistic fashion.

We highlight these four actions that we have taken:

- Validation and verification of the main DFT codes used in MARVEL (CP2K, Quantum-ESPRESSO, BigDFT) has been started. Validation and assessment of pseudopotential libraries is performed in an automatic way and across codes.
- Exchange and reuse of libraries with new and enhanced functionalities (e.g. nonlocal van der Waals functionals, electrochemical solvent models) has been initiated for the DFT codes.
- General interface format will be supported to connect DFT codes with high-level correlation energy codes (DMFT, DMRG, FCIQMC).
- A library of test molecules based on Co catalysts studied in VP2 has been created. The different systems will be calculated with a series of methods, starting from standard local functionals within DFT. Benchmark results for some systems will be generated using MRCI and other high-level quantum chemistry methods. Additional high-level results will be contributed by the groups of Markus Reiher (ETHZ) using DMRG and Ali Alavi (MPI Stuttgart) using FCIQMC methods. DMFT calculations on specific properties of these systems will be performed by the Werner group.

HP4 - Advanced Sampling Methods: The nucleation work, while scientifically very exciting, seems fairly disconnected from the rest of the activities.

There is general connection between the activity of the Parrinello group and that of many of the other groups in this MARVEL project since sampling and the study of rare events are crucial to many of the other groups such as the groups that study batteries (Alessandro Curioni), catalytic processes (Sandra Luber), CO₂ sequestration (Berend Smit, Wanda Andreoni), perovskite solar cells (Ursula Röthlisberger). The study of nucleation is a challenging testing ground for sampling methods. Furthermore the ability of controlling nucleation and crystal shapes has a direct impact on our ability to make those materials that are predicted to have special properties. A direct collaboration between the Parrinello and the Röthlisberger groups on the nucleation mechanism of some practically relevant perovskite systems will start soon and many more are planned with other groups. The Parrinello group has also joined efforts with the Ceriotti group in order to make structural fingerprints available to the materials science community through the free energy methods plugin PLUMED 2. More frequent meetings of the HP4 members are planned to promote collaborative efforts.

HP5 - Materials Informatics: The panel members think that some aspects of the machine learning look quite disconnected from the rest of the efforts. An example for this is the knowledge graph work. It is not clear to the panel how this is going to contribute the rest of the NCCR. If the machine learning is too detached, it runs the risk of not being helpful. The panel likes to point out a concern which is valid for the whole community, not only for the NCCR: How can the community ensure that everybody is not just generating a lot of new databases which are all separated from each other?

We will implement the recommendation of the Review Panel as follows.

• Knowledge graph for material design: We are creating a cognitive system to help accelerate material discovery. In this context the built-up of a knowledge graph that contains all relevant information is crucial. We would like to emphasize that the specifics of the knowledge graph are directly determined by the characteristics of the problems addressed in the other projects within MARVEL. In particular, the documents sources, the types of nodes, the connections as well as the analytics algorithms are directly determined by the material science problems (queries in our terminology) at hand in MARVEL. For instance, this work within HP5 is done in synergy with the work in VP2. We also stress though that the cognitive approach,

which relies in the analysis of a large body of unstructured information, is quite general and directly applicable to many domains. It is however, the expert knowledge that determines the flavor of the knowledge graph that allows for crucial domain adaptation. Thus, in the next two years we will continue the domain adaptation within VP2 but also extend in other vertical (sub-)projects.

• Databases: We are in close synergy with MARVEL members in order not to duplicate any work that may seem related. In particular, we are in close collaboration with the AiiDA designers in order to come up with a synergistic design. The knowledge graph designers have been regularly participating to AiiDA training sessions. Moreover, we stress that we are creating a system that can interface with any storage/database system either within or outside MARVEL. Thus, we design the system so as to allow API based communication with outside services that make it universally consumable.

PP6 - Informatics: This is a strong project. It is not quite clear to the panel how the code strategy Schulthess was presenting will exactly be made concrete. At the next site visits, the panel likes to hear more on how the scientists use the tools developed by the computer specialists.

This project has regular monthly meetings via videoconferencing involving the AiiDA and Materials Cloud teams at EPFL and the CSCS team, to brainstorm on plans, solutions, and federation strategies. In addition we have also an internal monthly team meeting for both AiiDA and the Materials Cloud here at EPFL. A quarterly Usage Report is regularly issued by CSCS providing an overview of the accounting records, statistics, and a highlight of the CSCS-EPFL collaboration on each period.

The Panel commented on this being a strong project, but wanting to hear more on how the tools/code strategy described by Schulthess will be used. This is now introduced in the first group report, and will be presented at the site visit. In addition, we would like to stress that funding for the Schulthess group has started only in Year 2, due to the challenges of finding the appropriate personnel.

PP7 - Experiments: Most of the participants in this Platform seem to focus on testing predictions that were made by computations which is certainly important. However, there should be a feedback mechanism: The experiments must feed back to the theory. In some cases the experiments will validate the predictions and in other cases they will not confirm the theory. That is also useful because it provides boundaries to the applicability of certain theoretical approaches. Therefore the panel is very interested to see an established feedback mechanism.

The feedback loop between theory and experiment is an important aspect and has been taken into account at the evaluation of the proposals. This ensures that only proposal with theoretical co-proposer and with a clear feedback loop between theory and experiment were funded. A first good example in this respect is the project by Lippert and Marzari on artificial photosynthesis. Based on the results of high-throughput theoretical calculations the oxynitride photocatalysts YTaON₂ was synthesized. However, the synthesized crystalline structure is different than theoretically predicted, and the refinement of previous calculations is ongoing. In order to better establish the synergies between theory and experiment, the PP7 organizes regular workshops where all project partners meet and the progress of the experimental effort and for the theory effort for each project is presented.

The experimental layout follows the pattern: computational screening – experimental validation – industrial application. The reviewers' worry is that the industrial applications put severe constraints on the materials that are potentially applicable. If the material screening is done without these constraints in mind, there is a high risk that many materials will never make it in applications. There should be a good balance between the freedom of basic research and requirements set by applications.

While identifying paths towards industrial applications is an important part of PP7, it is not the goal of PP7 to develop industrial applications. The main focus of PP7 is to provide experimental verification for theoretical predictions in several fundamental research areas, all of which have the potential to be technologically useful sometime in the future. In order to make this clearer, the evaluation criteria for the second call have been changed accordingly, stating that scientific excellence is considered in view of basic research and/or potential for technology transfer.

Last, we mention that the individual subprojects meet more often than the whole PP7 project; that the postdocs of the projects at Empa and PSI can take part in the training and education program of the EC co-financed fellowship programs at PSI and Empa, and that PSI is operating several large scale research facilities which are open access user facilities. The instrumentation is constantly improved and in particular developments for VP1 and VP2 related research projects are in-line with the strategy of the PSI. The significant amount of matching fund provided by PSI and Empa for the experimental verification projects is a clear proof for this.

Structure related criteria

The structure related areas are primarily carried by Nicola Marzari and junior PIs. The panel thinks that the senior PIs should get more involved in these activities. In general, the panel would like to see a small set of concrete and measurable actions for each of the structural areas at the next site visit. Most of the structural activities described in the report and during the presentations are concentrated at EPFL. The reviewers would like to get also some information on similar activities at other institutions of the MARVEL network where appropriate.

This has probably been the most difficult suggestion to implement - on one side, having the support team based at EPFL makes it more difficult to organize activities at different sites. On the other side, it is difficult to impose unto PIs additional activities without a clear reward. The failure in hiring an effective program manager has made organization and planning more difficult, if anything. Still, the activities in the four core areas of Education, Equal Opportunities, Knowledge and Technology Transfer, and Communication, detailed in the later sections, have been focused and we'd say very successful, but more could be done with a Program Manager in place, and more contributions from the senior PIs. The Director thinks that at least one senior PI should be added to each of the 4 structural areas, but the challenge is how to make this an effective development.

• The educational activities, coordinated by Nicola Marzari and Michele Ceriotti (MARVEL junior PI), have seen a very successful first Junior retreat in 2015, that will repeat annually (July 18 to 21, 2016, for this year), supported MARVEL students at 3 international conferences, initiated the recording of all the MARVEL Distinguished Lectures and thematic contributions (http://slideshot.epfl.ch/events/14, to be transferred to the Materials Cloud when launched), broadcasted live in 2-way videoconferencing to ETHZ the MARVEL Distinguished Lecture of Prof. G. Ceder, and contributed both to the Khartoum Workshop for the Advancements in Material Sciences, from February 1 to 10, 2015, in Karthoum (Sudan), and to the 3rd and 4th African School on "Electronic Structure Methods and Applications" — the former meant to be held in Abuja (Nigeria) in the summer of 2014, but moved to Johannesburg (South Africa) in Jan-Feb 2015, and the latter slated for June 2016 in Accra (Ghana). 7 conferences have also been sponsored (described in the KTT part), and AiiDA workshops in Lausanne, Zurich, Berlin, Trieste have taken place.

- The equal opportunities activities, coordinated by Lidia Favre-Quattropani (MAR-VEL management), Clémence Corminboeuf (MARVEL junior PI), Helene Füger and Farnaz Moser (EPFL Equal opportunities and Science Outreach) have been extremely successful, with, as suggested, actions for women working within the NCCR MARVEL, such as the 3 "professional skills development" training workshops for women in science that were an ETHZdriven initiative, with the three schools in 2015 targeted at young girls, and with a soon-to-be-launched program of 6-month Master internships. All these activities are meant to enhance and sustain a pipeline leading to more women in senior positions.
- The KTT activities, coordinated by Nicola Marzari and Pascale Van Landuyt (MAR-VEL KTT officer), with contribution by Nicola Spaldin (MARVEL senior PI) have focused on the KT side on scientific aspects (AiiDA and CP2K tutorials, CECAM conferences, work towards the Materials Cloud deployment, and close interactions with PP7 in the soliciting and reviewing experimental proposals); on the TT side Pascale has met with all the EPFL groups (soon to be extended with all groups offsite) to develop a portfolio of MARVEL research that could be presented to companies, prepared the presentations for this, and initiated several contacts with 13 different companies. Industrial engagement remains a core long-term priority, but it requires educating the PIs on the opportunities, and creating some reward system that would make attractive to the PIs and the companies to invest in this effort.
- The communication activities, coordi-

nated by Lidia Favre-Quattropani and Nicola Marzari, and with contributions by Alfredo Pasquarello (MARVEL Deputy Director), have seen the deployment of the web site shortly after the first site visit (www.nccr-marvel.ch), the internal activities including the Review and Retreat in September, the organization of project meetings for VP2 (together with the DFT part of HP3), HP5, and PP7 (in addition to the monthly meetings for PP6), the MAR-VEL Distinguished lectures, Open Doors events at CSCS, USI, EPFL, the production of a short movie about MARVEL, a twopages leaflet presenting MARVEL activities, and the ACCES Visualization Contest.

Structural aspects and support by the Home Institution

As mentioned in the report, all the structural measures are in place — two PATTs have been hired (Ceriotti and Yazyev), one is being actively sought for, and there have been senior hires (Hesthaven and Smit); Smit has been integrated into the NCCR through an agility grant. The CoViz facility has been completed, and EPFL has committed 120'000 CHF to a distant-education classroom (CoViz2) that will be finished in February 2016, co-located with MARVEL headquarters, that have moved in November 2015 in the new ME D building.

Recommendations

The Review Panel concluded with three overall recommendations:

The NCCR should define what set of metrics it wants to apply to decide on success or failure of the MARVEL activities.

We have asked the Project Leaders to define the metrics to apply to each project — these are reported below. It is the view of the Director that metrics should follow closely the mission statement, and should cover these three core goals and relevant measures:

- 1. **Design and discovery of novel materials.** This is the ultimate goal of the project, and so the identification of materials displaying novel physics or improved properties or performance is the first key metric.
- Conceptual, methodological, and algorithmic developments spanning the three horizontal projects of advanced quantum simulations, advanced sampling methods,

and materials informatics. This second key metric for development efforts is in their dissemination and usability by other groups, different from those of the original developers, either within MARVEL or by the scientific community at large.

3. Dissemination of curated data, calculation workflows, and code-specific plugins to enhance the synergies between codes, calculations, and data that support the entire vision of the project and that are enabled by AiiDA and the Materials Cloud.

These below are the metrics that have been elaborated by each project:

VP1: For this fundamental science project our goals are qualitative rather than quantitative, and so are the corresponding metrics:

- understand something not understood before
- predict new material properties and either confirm them or get an interesting surprise

VP2: We have identified the following metrics for success for VP2:

- *High-quality science*, which can be measured in terms of talks at conferences, publications, and number of citations.
- *Transfer of knowhow*, which can be measured in terms of collaborations and joint publications.
- *New research lines* made possible by the NCCR MARVEL.
- *Identification of new materials* with interesting properties.
- *KTT*, for some VP2 projects.

HP3 metrics:

- Joint publications on method development, assessment of methods and method verifications.
- Collaborations with VP1 and VP2.
- Collaborations with HP4 and HP5.

HP4: As metrics we will consider intensified collaborations as reflected by common publications and interfacing of various software packages developed in different groups.

PP6 metrics:

- CSCS hardware: availability, usage metrics, regular reporting.
- AiiDA: stable and updated releases, uptake by the MARVEL and broader community.
- Materials Cloud: deployment, data content, data delivered, user satisfaction.
- Public release of SIRIUS library with supporting paper.
- Improved time-to-solution for Quantum-ESPRESSO and Exciting codes on hybrid architectures.

PP7: The goal of PP7 is to provide experimental verification of theoretical calculations. The metrics by which we judge the success of PP7 is as follows:

- submission of proposals by experimental and theoretical PIs on a common project as a result of a "Call for Proposals";
- selection of projects with the highest potential for experimental verification;
- execution of projects with established feedback loop, regular meetings;
- definition of policies how collaborations between high-throughput theorists and

experiments can be guided most efficiently; these may serve as a blueprint for other people in the community.

In the course of phase I, the NCCR must develop a strategy how to focus the broad research scope for the future. The panel recommends to play to the strengths of the Swiss scientific community and to take into account the degree of integration of the individual projects when reshaping the research portfolio. This should happen in view of the preparation for the phase II proposal at the latest.

As mentioned, several efforts are in place in focusing the activities of the NCCR, after having nurtured them, and the second Site Visit will provide a valuable checkpoint on all these.

For each of the 4 structure-related areas the NCCR should define a small set of very concrete and measurable actions which can be pursued and adapted from year to year.

These have been detailed in the 4 structure-related sections of this report.

Next site visit

The next site visit, planned for April 19 and 20, 2016, will be organized in a more interactive format, allowing for extensive contact with students through an extended poster session.

3.1 Structure and organisation of the NCCR

3.1.1 Structure of the NCCR

This section provides an up-to-date summary of the organisation of MARVEL, the National Centre of Competence in Research (NCCR) Materials' Revolution: Computational Design and Discovery of Novel Materials.

At the scientific level, the NCCR MARVEL is structured around (Fig. 1) two vertical projects of materials design and discovery, with a focus, respectively, on

VP1 novel materials physics,

VP2 novel materials applications,

supported by three horizontal projects on

- HP3 advanced quantum simulations,
- HP4 advanced sampling methods,



Figure 1: Administrative and scientific structure of MARVEL

HP5 materials informatics,

interfaced to two platform projects on

- **PP6** the informatics and supercomputing infrastructure, and
- **PP7** experiments: synthesis, characterization, and testing.

Its home institution is EPFL and participating scientists are coming from a wide range of Swiss academic and industrial institutions, forming a strong network in the field of computational design and discovery of novel materials. MARVEL is organized around different bodies responsible of its operation, the Executive Committee, the Scientific Committee and the Scientific Advisory Board.

Academic and industrial institutions participating in MARVEL

- Federal Institute of Technology, Lausanne (EPFL), Home institution
- Federal Institute of Technology, Zurich (ETHZ)
- University of Geneva (UniGE),
- University of Fribourg (UniFR)
- University of Basel (UniBas)
- University of Zurich (UniZH)
- Università della Svizzera italiana (USI)
- Swiss National Supercomputing Centre (CSCS)
- Paul Scherrer Institute (PSI)
- Materials Science and Technology Research Institute (Empa)
- IBM Research GmbH, Rüschlikon

Executive Committee

The Executive Committee is composed by the director, the deputy director, and a third member from a different institution.

- Nicola Marzari, EPFL, director
- Alfredo Pasquarello, EPFL, deputy director
- Thomas Schulthess, ETHZ

MARVEL Project Leaders

The Project Leaders are in charge of coordinating and overseeing the activities of the seven projects.

- Matthias Troyer, ETHZ, VP1
- Ursula Röthlisberger, EPFL, VP2
- Jürg Hutter, UniZH, HP3
- Stefan Goedecker, UniBas, HP4
- Alessandro Curioni, IBM, HP5
- Thomas Schulthess, CSCS and ETHZ, PP6
- Frithjof Nolting, PSI, PP7

Scientific Committee

The Scientific Committee is composed by the Executive Committee and the seven Project Leaders.

MARVEL Group Leaders

The Group Leaders receive funding from the NCCR MARVEL to support PhD students, postdoctoral associates or research scientists, whose research they supervise.

Principal Investigators (PIs)

- Alessandro Curioni, IBM
- Antoine Georges, UniGE
- Stefan Goedecker, UniBas
- Pierangelo Gröning, Empa
- Jürg Hutter, UniZH
- Michel Kenzelmann, PSI
- Christoph Koch, EPFL
- Nicola Marzari, EPFL
- Frithjof Nolting, PSI
- Michele Parrinello, USI and ETHZ
- Alfredo Pasquarello, EPFL
- Ursula Röthlisberger, EPFL
- Thomas Schulthess, CSCS and ETHZ
- Nicola Spaldin, ETHZ
- Matthias Troyer, ETHZ
- Philipp Werner, UniFR

The three experimental PIs of PP7 (Frithjof Nolting, project leader, Michel Kenzelmann, and Pierangelo Gröning) do not receive support themselves, but kindly coordinate the PP7 activities involving all of PSI and Empa.

Agility members

- Wanda Andreoni, EPFL
- Michele Ceriotti, EPFL
- Volkan Cevher, EPFL
- Clémence Corminboeuf, EPFL
- Daniele Passerone, Empa
- Berend Smit, EPFL
- Joost VandeVondele, ETHZ
- Anatole von Lilienfeld, UniBas
- Oleg Yazyev, EPFL

Experimental project leaders

- Claudia Cancellieri, Empa
- Thomas Lippert, PSI
- Dirk van der Marel, UniGE
- Marisa Medarde, PSI
- Thomas J. Schmidt, PSI
- Thorsten Schmitt, PSI
- Grigory Smolentsev, PSI
- Urs Staub, PSI

Additional experimental project leaders, from May 2016

- Raffaella Buonsanti, EPFL
- Daniele Pergolesi, PSI
- Marco Ranocchiari, PSI
- Marta D. Rossell, Empa
- Christian Rüegg, PSI and UniGE
- Ming Shi, PSI

Scientific Advisory Board

The Scientific Advisory Board has the formal role of suggesting research lines that need to be strengthened within the NCCR, or that have run their course, and act as an independent reviewer of the activities of the NCCR in addition to the official Review Panel of the SNSF. Its members are:

- Giulia Galli, Liew Family Professor, Institute for Molecular Engineering, University of Chicago IL, USA
- Gabriel Aeppli, Synchrotron Radiation and Nanotechnology, Paul Scherrer Institute, Villigen, CH
- Gian-Luca Bona, Director, Empa, Dübendorf, CH



- Gerbrand Ceder, Chancellor's Professor, Department of Materials Science and Engineering, University of California, Berkeley CA, USA
- Karsten Jacobsen, Center for Atomic-scale Materials Design, Technical University of Denmark, Lyngby, DK
- Boris Kozinsky, Robert Bosch Research and Technology Center, Cambridge MA, USA
- Alexander Lichtenstein, Institute of Theoretical Physics, University of Hamburg, DE
- Sadasivan Shankar, Margaret and Will Hearst Visiting Lecturer in Computational Science and Engineering, Harvard School of Engineering and Applied Sciences, Cambridge MA, USA
- Risto Nieminen, Chairman, Psi-K network and Dean, Department of Applied Physics, Aalto University School of Science, Helsinki, FI
- Dominic Tildesley, Director, CECAM -Centre Européen de Calcul Atomique et Moléculaire, EPF Lausanne, CH

Management

A team of administrative and scientific collaborators has been formed to carry out the management and financial duties of MARVEL, and take care of the other responsibilities, such as communication, knowledge and technology transfer, education and equal opportunities. They work with the executive committee (Nicola Marzari, Alfredo Pasquarello, and Thomas Schulthess) and with the other faculty in charge of the structural activities (Michele Ceriotti, education and training, Clémence Corminboeuf, equal opportunities). The team is composed of

- Lidia Favre-Quattropani, 60%, scientific manager, communication, equal opportunities
- Riccarda Caputo, 100%, programme manager, from July to December 2015. A new person will be recruited to replace her.
- Pascale Van Landuyt, 20%, Industrial liaison and tech transfer officer, knowledge and technology transfer
- Valérie Le Dreau, 80%, financial officer until January 2016
- Ursula Vaucher, 80%, financial officer from February 2016
- Elizabeth Gueniat, 60%, administrative assistant

The equal opportunities and outreach activities are organized together with

- Helene Füger, EPFL Equal Opportunities Office
- Farnaz Moser, EPFL Science Outreach Department

Scientific projects

Vertical projects

VP1 — Novel Materials Physics

Project leader:

• Matthias Troyer, ETHZ

Members:

- Matthias Troyer, ETHZ
- Nicola Spaldin, ETHZ
- Antoine Georges, UniGE
- Oleg Yazyev, EPFL

VP2 — Novel Materials Applications

Project leader:

• Ursula Röthlisberger, EPFL

Members:

- Ursula Röthlisberger, EPFL
- Alfredo Pasquarello, EPFL
- Jürg Hutter, UniZH
- Nicola Marzari, EPFL
- Alessandro Curioni, IBM
- Wanda Andreoni, EPFL
- Clémence Corminboeuf, EPFL
- Daniele Passerone, Empa

Horizontal projects

HP3 — Advanced Quantum Simulations

Project leader:

• Jürg Hutter, UniZH

Members:

- Jürg Hutter, UniZH
- Matthias Troyer, ETHZ
- Philipp Werner, UniFR
- Stefan Goedecker, UniBas
- Nicola Marzari, EPLF
- Joost VandeVondele, ETHZ

HP4 — Advanced Sampling Methods

Project leader:

• Stefan Goedecker, UniBas

Members:

- Stefan Goedecker, UniBas
- Michele Parrinello, USI and ETHZ
- Berend Smit, EPFL
- Michele Ceriotti, EPFL
- Anatole von Lilienfeld, UniBas
- Ursula Röthlisberger, EPFL

HP5 — Materials Informatics

Project leader:

• Alessandro Curioni, IBM

Members:

- Alessandro Curioni, IBM
- Christoph Koch, EPFL
- Volkan Cevher, EPFL
- Anatole von Lilienfeld, UniBas

Platform projects

PP6 — Informatics

Project leader:

• Thomas Schulthess, CSCS and ETHZ

Members:

- Thomas Schulthess, CSCS and ETHZ
- Nicola Marzari, EPFL

PP7 — Experiments

Project leader:

• Frithjof Nolting, PSI

Members:

- Frithjof Nolting, PSI
- Michel Kenzelmann, PSI
- Pierangelo Gröning, Empa
- Claudia Cancellieri, Empa
- Thomas Lippert, PSI
- Dirk van der Marel, UniGE
- Marisa Medarde, PSI
- Thomas J. Schmidt, PSI
- Thorsten Schmitt, PSI
- Grigory Smolentsev, PSI
- Urs Staub, PSI

3.1.2 Organisation of the NCCR

Organisational changes

At the level of the management team, Pascale Van Landuyt was appointed in May 2015 as industrial liaison and tech transfer officer at 20%. Riccarda Caputo was appointed in July 2015 as programme manager. She left end of December 2015 and a new person will be recruited shortly to replace her. Valérie Le Dreau, in charge of the financial aspect of MARVEL at the EPFL Research Office, left end of January 2016 for a new life experience in Singapore. She is replaced, at the Research Office, by Ursula Vaucher, who has extensive experience as a financial officer with other NCCRs, guaranteeing a smooth transition. Elizabeth Gueniat will retire end of May 2016. A new administrative assistant will thus be recruited soon.

Thomas Schulthess replaced Nicola Spaldin in the Executive Committee. As professor at ETHZ and director of CSCS, he is the third member from a different institution than EPFL. Concerning the organisation of the scientific projects of MARVEL, a new agility project has started in May 2015 under the leadership of Volkan Cevher at EPFL. Eight more group leaders - Claudia Cancellieri, Thomas Lippert, Dirk van der Marel, Marisa Medarde, Thomas J. Schmidt, Thorsten Schmitt, Grigory Smolentsev, Urs Staub - have joined in May 2015, following the first round of selection of the experimental projects for the experimental platform PP7, adding up to 33 group leaders. Six more experimental group leaders - Raffaella Buonsanti, Daniele Pergolesi, Marco Ranocchiari, Marta D. Rossell, Christian Rüegg, Ming Shi — will enter in PP7 in May 2016 at the beginning of the third year.

Procedural changes

The NCCR-MARVEL Regulations were signed on April 20, 2015 by Nicola Marzari, director of MARVEL, and Andreas Mortensen, EPFL viceprovost of research.



3.2 Management activities and status of collaboration/integration

3.2.1 Activities and measures

Events organisation

In this second year, the management organized several meetings, lectures and events that are described in more details in the chapters 5 to 8 (knowledge and technology transfer, education and training, equal opportunities and communication).

- The Review Panel meeting took place at EPFL on April 22 and 23, 2015. This was the first meeting between the Review Panel and MARVEL group leaders.
- MARVEL was present at the Platform for Advanced Scientific Computing PASC15 Conference at ETHZ, from June 1 to 3, 2015, with a booth and a poster presenting its activities.
- The first MARVEL Junior Retreat was organized from July 7 to 10, with 48 participants, including 7 invited speaker on different subjects. The organisation was under the responsibility of three postdocs, supervised by Michele Ceriotti. MARVEL management helped for the administrative tasks. The second edition is beginning to be organized with a new team of motivated postdocs and will take place from July 18 to 21, 2016.
- The second MARVEL Review and Retreat, gathering annually all MARVEL members, took place on September 3 and 4 at EPFL, with almost 90 participants, including 8 of the 10 members of the Scientific Advisory Board.
- MARVEL management and the group of Nicola Spaldin organized the three workshops for women in science with Nancy Houfek in ETHZ on October 12 and 14, and in EPFL on October 16, 2015.
- MARVEL management coordinated and participated in the common MARVEL and ACCES booth at *Scientastic*, the new science festival of EPFL, on November 21, 2015.
- A project meeting, VP2-Day, took place at EPFL on January 29, 2016, with all the research groups (students and group leaders) participating in that project.
- Four MARVEL distinguished lectures were organized at EPFL on August

31, 2015 with the venue of Prof. Sidney Yip (MIT, Cambridge, USA), on October 13 with Prof. Leonid Levitov (MIT, Cambridge, USA), on November 11 with Dr. Pierre Villars (MPDS, Vitznau, Switzerland), and on January 25, 2016 with Prof. Gerbrand Ceder (University of California at Berkeley, USA). These lectures were all recorded and made available on the MARVEL website www.nccr-marvel.ch. The last one, with Prof. Gerbrand Ceder, was also broadcasted live at ETHZ.

• 13 seminars were organized at EPFL jointly with the THEOS research group of Nicola Marzari and two jointly with the COSMO research group of Michele Ceriotti.

Other activities

MARVEL management, together with the whole group of Nicola Marzari, moved to the new ME D building at EPFL on November 19, 2015.

On a financial level, MARVEL sponsored the following workshops, tutorials and conferences, through its education and training or knowledge and technology transfer budgets.

- CECAM conference "Frontiers of firstprinciples simulations: materials design and discovery", Berlin, February 1 to 5, 2015.
- "Khartoum Workshop for the Advancements in Material Sciences" (KWAMS-15), Khartoum, February 1 to 10, 2015.
- The international conference "Nothing is Perfect — The Quantum Mechanics of Defects", Ascona, April 26 to 29, 2015.
- The CECAM workshop "Future Technologies in Automated Atomistic Simulations", CECAM-HQ-EPFL, Lausanne, June 8 to 10, 2015.
- The CECAM school "4th CP2K Tutorial", CECAM-ETHZ, Zurich, Switzerland, August 31 to September 4, 2015.
- The international symposium and workshop "Electronic Structure Theory for the Accelerated Design of Structural Materials", Moscow, October 26 to 30, 2015.

• The three AiiDA workshops on February 5, 2015 in Berlin, on November 2 to 6, 2015 in Lausanne, and on December 15, 2015 in Trieste.

MARVEL supported the travel expenses of four MARVEL students participating in the "Simons Many Electron Collaboration" Summer School, at the Simons Center for Geometry and Physics, Stony Brook, June 8 to 12, 2015. Contributions were paid from MARVEL to the Psi-k Network and the Quantum-ESPRESSO Foundation as group memberships.

3.2.2 Status of collaboration/ integration

Numerous new collaborations are taking place within the MARVEL groups, and now all new and ongoing efforts are listed within the research descriptions of each of the seven main projects. The annual Review and Retreat, gathering all MARVEL groups, took place on September 3 and 4, 2015 and several project meetings were organized, or will be, for the individual projects.

Outside MARVEL, it is noteworthy to mention that EPFL and CSCS have become part of the newly established EU H2020 Centre of Excellence MaX on "Materials Design at the Exascale" (max-center.eu), and that EPFL is one of the few academic partners that are part of a proposed EU H2020 CSA (Coordination and Support Action) for the newly established European Materials Modeling Council. Also, EPFL is part of the EU H2020 NFFA "Nanoscience Foundries and Fine Analysis" consortium, where theoretical modeling tools, expertise and dedicated support are offered to the experimental community, but where also novel models for data storage and dissemination for experimental efforts are being investigated (AiiDA being one of its three case studies).

4.1 Results since the last progress report

Vertical Project ____ f L

VP1 — Novel Materials Physics

Project leader: Matthias Troyer (ETHZ)

Participating members: Matthias Troyer (ETHZ), Nicola Spaldin (ETHZ), Antoine Georges (UniGE), Oleg Yazyev (EPFL)

Summary and highlights: In this project we focus on the design of materials in which the fundamental physics and resulting functionalities are determined by strong interactions between individual electrons, or by new and emerging phenomena, such as topological properties of their band structures. The challenge here is the development of methods that can later be used in high-throughput material design. In the second year we have continued developing the Z2PACK software package, used it to identify a new type of Weyl semimetals, and started a high-throughput search for more topological materials that turned up first candidates. This approach has already resulted in a theoretical prediction followed by experimental validation of a novel topological insulator phase in Bi₄I₄, while several Weyl semimetals candidates have been identified. We have continued the integration of DMFT methods (for accurate description of strong correlations) with DFT (for description of crystal structure and chemistry) and applied the combined methodology to the understanding of the trends in properties across the rare-earth nickelate series, and in designing a new material which is described by a single-band Hubbard Hamiltonian.

General view of the project

Major research questions

Strongly correlated materials, and especially transition-metal oxides (TMOs) exhibit a wide range of fascinating and potentially technologically important phenomena, including exotic superconductivity, metal-insulator transitions, unusual magnetic ordering and coupled charge/spin/orbital interactions, as a result of their coupled spin, orbital, electronic and lattice orders. These materials display a number of competing phases, with small energy differences between them. An outstanding challenge is to *control these materials* in order to "guide" them towards a specific phase with a desired functionality, determined by their chemical composition and crystal structure. How to achieve control of these materials by acting on these "knobs" is the big research question which underlies the present project.

The discovery of topological insulators, materials realizing topologically nontrivial band structures, opened a new direction of research in condensed matter physics with high potential for technological applications, not the least for quantum devices. The first "practical" Z_2 topological insulators, bismuth and antimony chalcogenides, were discovered approximately five years ago in the well-known class of thermoelectric materials. Topological materials seem to be common among heavy element compounds featuring strong spin-orbit interactions. The search for new members of the family of Z_2 topological insulators and materials

belonging to other topological classes, such as topological crystalline insulators, Chern insulators, Dirac and Weyl semimetals, is another priority direction in this emerging field of research.

Long-term goals

The field of correlated and topological materials is very wide and we thus focus on a few highly promising topics, which we will quickly summarize here and then develop in more detail in the individual subproject reports.

A particularly promising class of materials on which we focus currently are the rare-earth nickelates, which display a phase transition between a high-temperature metallic phase and low-temperature insulating phases. This transition, which can be close to room temperature, is highly sensitive to chemical composition, as well as structural constraints and strain (tunable by changing the substrate or by heterostructuring). This could find applications to switches or to the recently proposed piezoelectric transistors. Deciphering the basic mechanism of the metal-insulator transition (MIT) in nickelates and clarifying to what extent orbitalengineering and the realization of a singleband material are possible for these materials is one of the short and medium term goals of this project.

Oxide heterostructures often exhibit emerging properties that are different from those of the constituent materials. One example is the observation of metallic behavior in heterostructures containing two otherwise insulating oxides. The complex nature of the insulatorto-metal transitions in such heterostructures mainly stems from the simultaneous appearance of several new features compared to the bulk forms of the constituent materials. As a first step we aim to understand the role of bare epitaxial strain in the insulator-to-metal transitions for some prototypical Mott insulators, will then move to an explicit description of interface effects which take place at the boundary between two different oxides, with a long-term goal to achieve a concise and comprehensive picture of oxide heterostructures composed of Mott insulators.

Multiferroic materials, that exhibit simultaneous ferroelectricity and ferromagnetism, are of fundamental interest because of the chemical contraindication between these phenomena, and of potential technological importance because of the cross-coupling between them. Currently practical materials that are multiferroic at room temperature do not exist. In the first year of the project we have focused on a class of materials — the RBaFeCuO₅ (R^{3+} = Y^{3+} and 4f rare-earth ions) series of bilayered perovskites — that is promising for achieving room-temperature multiferroism and studied the origin of the unusual magnetic order that is believed to be responsible for the polar state. While most subprojects of Vertical Project 1 are concerned with understanding new phenomena in a quantitative manner in realistic material simulations instead of effective models, here also follow the opposite path. Inspired by ultracold atomic gases, we propose to engineer materials with precisely (and only) the bands and interactions contained in model Hamiltonians (such as the Hubbard model). Learning to design "simple" materials will be an important first step and will give us model materials that exhibit physical phenomena known from effective models in simple materials, not encumbered by many complications of more complex materials.

For topologically nontrivial materials we plan to set up a fully automated procedure for calculating topological invariants of band structures starting from first-principles calculations and use it for generic scanning of existing materials databases for identifying materials candidates for quantum spin Hall, quantum anomalous Hall effects, as well as 3D topological insulators and Weyl and Dirac semimetals. In parallel, we search for optimal candidate materials for hybrid systems, such as heterostructures and nanowires that are predicted to realize Majorana fermions.

Achievements

On the topic of nickelates, in the first review period we proposed a consistent theoretical description of the metal-insulator transition (MIT) [1] using only the strongly hybridized O-Ni e_{q} states which form the active degrees of freedom close to the Fermi level. In this review period we have calculated the optical spectra of nickelates based on this theoretical description, allowing for direct comparison to experiments, and have systematically investigated the key energy scales involved in this description over the whole RNiO3 series. We have begun a MARVEL-supported collaboration with the group of Marisa Medarde at the Paul Scherrer Institute to verify experimentally our predictions. The understanding that we have developed in this review period will allow us to work towards our goal of "orbital engineering" in the next.

In the strained films, we have extended our



studies of the role of epitaxial strain in the insulator-to-metal transitions from the two prototypical Mott insulators that we studied in the first reporting period - LaTiO₃ (formal t_{2g} occupation d^1) and LaVO₃ (d^2) — to include also SrVO₃ and CaVO₃. We rationalized the observed trends as a function of strain in the insulator-to-metal transitions based on the modifications in the crystal-field splitting and in the bandwidth of these orbitals with strain. We now begin to develop a simple, general model explaining the strain-induced changes of the metallic properties through a limited number of parameters (e.g. bandwidth, on-site energies, filling). We have begun a MARVEL-supported collaboration with the group of Thorsten Schmitt at the Paul Scherrer Institute on the development and verification of this model and its extension to study explicit interface effects.

Motivated by our studies of multiferroic YBaFeCuO₅ in the first reporting period, we proposed a new mechanism in which frustration resulting from chemical mixing can lead to the formation of a ground state magnetic spiral with a net ferroelectric polarization. In this reporting period we performed simulations of a model describing our proposed mechanism and found that it indeed produces a spiral ground state. A manuscript describing our findings is in preparation. Our next step is to identify other material systems in which this mechanism is likely to occur.

The Z2PACK software package [2] was further developed and can now detect special topological invariants associated with unusual Weyl points. With this tool we predicted a new type of Weyl fermions, dubbed the type-II Weyl fermions and identified WTe₂ as a Weyl semimetal where they may appear [3]. While standard (type-I) Weyl points have a point-like Fermi surface and result in low-energy excitations that are the Weyl fermions of quantum field theory, the type-II Weyl points occur at the boundary between electron and hole pockets and realize excitations that are not present in relativistic theories due to the violation of Lorentz symmetry in condensed matter systems. These materials are expected to exhibit a range of novel physical effects, especially in magnetotransport. In high-throughput screening of over 2'000 materials for topological properties we then identified MoP₂ and WP₂ as robust type-II Weyl semimetal candidate materials with relatively simple band structures [4]. Following our recent discovery of the novel quasi-one-dimensional topological insulator Bi_4I_4 [5], we pursued an investigation of a whole family of compounds with formula $Bi_4I_{4-x}Br_x$, and found alternating trivial and topological insulators as x is varied. We have initiated an experimental collaboration to test both of these predictions.

An application of nontrivial topological materials appears in quantum computing with the realization of topological quantum bits in semiconductor/superconductor nanowires. These experiments pose a demand for theoretical/computational guidance of such experiments and optimization of the materials. We have developed realistic effective models of InSb and other zincblende semiconducting wires and heterostructures, to guide the design of new quantum devices. These models were used to identify directions for the wire growth with maximal spin-splittings [6], necessary for inducing Majorana fermions, and to study the quantum spin Hall effect in InAs/GaSb quantum wells.

The "standard model" of condensed matter physics is widely regarded to be the Hubbard model which augments independent-electron band theory through a single parameter to account for electron-electron correlations. However, beyond one dimension, the problem is intractable and so much current research aims at finding appropriate approximations for obtaining the Hubbard model phase diagram. Here we take a new approach — we used *ab initio* methods to design a material whose Hamiltonian is exactly that of the Hubbard model so that its solution can be found by measuring the material properties. After identifying a crystal class and several appropriate chemistries, we used density functional theory to screen for the desired electronic band structure, and dynamical mean field theory to study the Mott transition. Following this, we took the most promising candidate and addressed its structural stability, possibilities for doping and exotic superconductivity, and its similarity with the cuprate high-temperature superconductors. A manuscript has been accepted for publication in Physical Review B.

Milestones

April 2016

- Guidelines for the design and search for topological materials (Troyer and Yazyev).
- Finalizing Z2PACK package and publication (Troyer and Yazyev).
- Prediction of several candidate materials with topologically nontrivial band structures (Troyer and Yazyev).

• Advance research aiming at control of orbital polarization in transition-metal heterostructures ("orbital engineering") (Georges and Spaldin).

April 2017

- Evolution of electronic structure of rareearth nickelates across the series calculated using LDA+DMFT (Georges and Spaldin).
- TRIQS to WANNIER90 interface (Georges and Spaldin).
- New model for spin spirals elucidated and implications explored (Spaldin and Troyer).
- Experimental realization of some candidate materials with topologically nontrivial band structures (Troyer and Yazyev).
- Derivation of effective tight-binding models for topological materials relevant for quantum computing applications (Troyer and Yazyev).
- Propose a material with a large value of orbital polarization (Yazyev and Troyer).

April 2018

- Integration of Z2PACK into WANNIER90 (Troyer, Marzari and Yazyev).
- Strain-dependence of behavior of rareearth nickelates understood (Georges and Spaldin).
- New multiferroics designed based on new spin-spiral mechanism (Troyer and Spaldin).
- Complete characterization of topological properties of the band structure of all materials in the ICSD database (Troyer and Yazyev).

New and ongoing collaborations

Collaborations within MARVEL

- Calculations of topological invariants from first principles, Oleg Yazyev and Matthias Troyer, starting June 2014.
- Designer Hamiltonian materials, Nicola Spaldin, Thomas Schulthess, Matthias Troyer, starting June 2014.

- Molecular Dynamics of PbTe using Tersoff potentials fitted to DFT, Alessandro Curioni and Nicola Spaldin, starting January 2015.
- Rare-earth nickelates using LDA+DMFT, Antoine Georges and Nicola Spaldin (Claude Ederer), starting September 2015.
- Code development in the framework of the TRIQS library, Antoine Georges and Nicola Spaldin (Claude Ederer), starting July 2014.
- A new mechanism for spin-spiral magnetism that can lead to multiferroism, Nicola Spaldin and Matthias Troyer, starting July 2014.

Collaborations with experimentalists

- Identification and implications of the Higgs mode in hexagonal manganites, Nicola Spaldin with Dirk van der Marel, since May 2015.
- Rare-earth nickelates, Nicola Spaldin (Claude Ederer) and Antoine Georges with Marisa Medarde, since May 2015.
- Strained rare-earth titanates and vanadates, Nicola Spaldin (Claude Ederer) with Thorsten Schmitt, since May 2015.
- Ultrafast processes, Nicola Spaldin and Philipp Werner with Urs Staub, since May 2015.
- Materials synthesis and ARPES measurements of Weyl semimetal candidates, Oleg Yazyev and Matthias Troyer with Ming Shi, since September 2015.
- Materials synthesis of bismuth compounds, Oleg Yazyev with Anna Isaeva, starting March 2014.
- ARPES measurements, Oleg Yazyev with Marco Grioni, starting March 2014.
- ARPES measurements at synchrotron facilities, Oleg Yazyev with Luca Moreschini, Eli Rotenberg, Aaron Bostwick, starting March 2014.
- Laser-based ARPES measurements, Oleg Yazyev with Alessandra Lanzara, since September 2015.
- Transport measurements, Oleg Yazyev with László Forró, since March 2015.



1 Discovery and simulation of topological phenomena in materials and devices (Matthias Troyer — ETHZ)

1.1 Research summary

The methods and software were developed to identify topological features in band structures of both insulators and metals. A new topological semimetal phase was found that exhibits a new type of low-energy excitations. Progress towards realistic simulation of topological devices is reported.

1.2 Scientific goals

Band structure topology is a fundamental property of materials, having significant impact on transport properties. In this project we aim at identifying topological structure of existing materials, as well as at predicting novel topological phenomena and materials that host it. Another goal of the project is to design and optimize solid-state devices and heterostructures that realize certain topological phases.

a) *Short-term* The recent discovery [3] by our group of a new type of Weyl fermion in semimetals, dubbed type-II Weyl semimetals, suggests a possibility to experimentally observe a wide range of novel physical phenomena. The materials predicted so far as candidates for this new phase have complicated Fermi surface structure. To study the properties of the type-II Weyl fermions, identification of more material candidates with simpler Fermi surfaces becomes crucial, and this is one of our short term goals. We have already developed the methodology for an automated high-throughput search for this kind of topological semimetals.

One of the physical effects predicted to exist in the type-II Weyl is their behavior in external magnetic fields. In particular, for certain field directions the material becomes insulating, while for the others it is a good conductor. The study of this effect effect is the goal of major importance.

b) *Long-term* Simulation of hybrid superconductor-semiconductor devices for optimizing conditions to realize Majorana fermions in quantum wires to facilitate experimentally feasible schemes of topological quantum computation [19, 20]. By developing a set of highly accurate tight-binding models, we aim at reliable calculation of *g*-factors, optimizing spinorbit coupling and understanding of superconducting proximity effect. This would allow for numerical simulations of realistic experimental

setups used to produce and manipulate topological qubits.

1.3 Results obtained since the last report

Several important results were obtained in the past year.

a) New fermions realized in known materials The search for novel topological phases in materials led to the discovery [3] of the type-II Weyl fermions and their host materials topological type-II Weyl semimetals. Topologically protected crossings of two bands, called Weyl points, occur in these materials near the Fermi level, producing low-energy excitations with linear dispersion. While standard (type-I) Weyl points have a point-like Fermi surface [21, 22, 23, 24, 25] and result in low-energy excitations that are the Weyl fermions of quantum field theory, the type-II Weyl points occur at the boundary between electron and hole pockets (Fig. 1) and realize excitations that are not present in relativistic theories due to the violation of Lorentz symmetry in condensed matter systems. These materials are expected to exhibit a range of novel physical effects, especially in magnetotransport.

Special topological invariants were identified that describe the topology of type-II Weyl semimetals, and the corresponding methods to compute them were implemented in Z2PACK software package [2]. The phase was predicted to exist in orthorhombic WTe₂ [3] and MoTe₂ (Fig. 2) [7].

Following a general methodology developed to search for Weyl semimetals, in collaboration with other members of MARVEL (Gabriel Autès and Oleg Yazyev at EPFL), the type-II Weyl semimetal was also found in materials of a different structure: MoP₂ and WP₂. A publication is in preparation [4].



Figure 1: Possible types of Weyl points. Left panel: Type I Weyl point with point-like Fermi surface. Right panel: Type II Weyl point is the touching point between electron and hole pockets.



Figure 2: *Type-II Weyl points in the Brillouin zone of MoTe*₂*. The red and blue colors indicate Chern numbers* +1 and -1 *respectively.*

b) Numerical methods for topological materials Calculation of topological contributions to orbital magnetoelectric response for real materials remained an unsolved problem for a long time. We developed a numerical solution to this problem [8], illustrating it with real materials applications. We expect that the developed method will be used not only for topological materials, since it greatly facilitates the search for good magnetoelectric materials — a task of immediate technological importance.

c) *Simulation of Majorana wires* A set of highly accurate symmetric tight-binding models was derived for several binary zincblende and wurzite compounds currently used in experiments on realizing Majorana fermions in quantum wires. These models were used to identify directions for the wire growth with maximal spin-splittings (Fig. 3) [6], necessary for inducing Majorana fermions, and to study the quantum spin Hall effect [26] in InAs/GaSb quantum wells.

Besides, a combination of tight-binding modeling with fist-principles calculations allowed us to show that spin-splittings can be greatly enhanced in disordered compounds $InAs_xSb_{1-x}$ relative to the clean InAs and InSb. The same work identified a possible ordered phase of this compound for x = 0.5 as a material realizing several possible topological phases: topological insulator, topological semimetal of a new kind and a semiconductor with spin-orbital



Figure 3: Some of the wire vs external electric field configurations that maximize Rashba spin-splittings in the conduction band of InSb. (a) A [110]-wire in the $[1\overline{10}]$ field. (b) A [111]-wire in the $[1\overline{10}]$ field.

splittings an order of magnitude larger than those in materials used in the current Majorana experiments. A publication is in preparation.

- 1.4 Planned research for next year
 - Identification of more simple type-II Weyl semimetal compounds.
 - Understanding of unusual response of type-II Weyl semimetals to external electric and magnetic fields.
 - Collaboration with the Yazyev group in EPFL on using Z2PACK for identifying novel topological insulators and semimetals.
 - Understanding the possibilities of reliable calculation of materials *g*-factors from specially designed tight-binding models.
 - Realistic tight-binding modeling of metalsemiconductor interfaces as a first step towards numerical simulations of superconducting proximity effects.
- 1.5 Synergies with other computational and experimental efforts

a) Within MARVEL Continued collaboration with Gabriel Autès and Oleg Yazyev at EPFL. Collaboration with the group of Nicola Marzari at EPFL on interfacing Z2PACK with WANNIER90 software package [27].

b) *Outside MARVEL* Collaboration with the group of Prof. N. P. Ong at Princeton University (USA) on experimental realization of the predicted novel response (chiral anomaly of a new kind) of type-II Weyl semimetals and collaboration with the group of Prof. C. M. Marcus at the Niels Bohr Institute (Copenhagen) on optimizing spin-orbit splitting in semiconducting nanowires.

1.6 Personnel

NCCR personnel

- Alexey Soluyanov, postdoc, 100%, from September 2014.
- Quansheng Wu, postdoc, 100%, from May 2015.
- Georg Winkler, PhD student, 100%, from September 2014.

Matching personnel

• Dominik Gresch, PhD student, 100%, from July 2015.



2 High-throughput search of novel topological materials (Oleg Yazyev — EPFL)

2.1 Research summary

Our project aims at the discovery of novel materials realizing topologically nontrivial electronic phases by means of high-throughput database screening.

2.2 Scientific goals

a) *Short-term* Development of a computational methodology for high-throughput screening of databases of known compounds in order to identify candidate topological materials.

b) *Long-term* Discovery of topological materials realizing novel topological phases or displaying novel physical properties via in-depth theoretical studies and experiments performed on promising candidate materials.

2.3 Results obtained since the last report

a) Methodological developments Our methodology for studying new classes of topological materials beyond conventional Z₂ topological insulators has been extended. Using Z2PACK, we are now able to identify Dirac and Weyl semimetals, topological crystalline insulators. Over the course of the past year, Weyl semimetals have attracted considerable attention following the theoretical prediction [22] and experimental confirmation [23, 24] of a Weyl semimetal phase in TaAs and related transition-metal monopnictide compounds (TaP, NbAs and NbP). Weyl semimetals are materials which contain symmetry protected crossing points near the Fermi level forming singly degenerate 3D Dirac cones carrying a non-zero chiral charge. The presence of Weyl points near the Fermi level gives rise to new exotic phenomena such as chiral anomalies, large negative magnetoresistance and the Fermi arcs surface states. By tracking the position of hybrid Wannier centers on the surface of a sphere enclosing a crossing point, we are able to confirm the existence of a Weyl node and determine its chirality.

A high-throughput screening of over 2'000 binary candidates materials from the ICSD database was performed. A new class of Weyl semimetals was identified in the family of transition-metal disphosphide compounds MoP_2 and WP_2 (Fig. 4). We discovered that these materials possess 8 Weyl nodes in their 3D Brillouin zone which come in pair of nodes with the same chiral charge. Importantly, in

these materials the Weyl nodes with opposite chirality are well separated in momentum space. As a consequence, the Weyl phase in MoP₂ and WP₂ is expected to be very stable and should give rise to long Fermi arcs on their surface. In this respect, MoP₂ and WP₂ possess a strong advantage over TaAs and related compounds where the oppositely charged Weyl nodes are located close to each other in momentum space.

The manuscript reporting these results is currently in preparation in collaboration with Alexey Soluyanov and Dominik Gresch (group of Matthias Troyer, ETHZ) [4]. We initiated a collaboration with the group of Ming Shi (PSI) aiming at experimental confirmation of our predictions.

b) Quasi-1D bismuth halide topological insulators Following our discovery of the novel topological insulator Bi_4I_4 [5], we pursued our study of



Figure 4: (a) Crystal structure of MoP₂. (b) Brillouin zone of MoP₂ with dots showing the locations of positively (red) and negatively (blue) charged pairs of Weyl nodes. (c) Calculated band structure of MoP₂ along the high-symmetry lines in the Brillouin zone. (d) Constant energy contours at the energy of Weyl node W_1 and $W_{1'}$, plotted in the neighborhood of the location of W_1 and $W_{1'}$ nodes in the $k_z = 0$ plane. The Weyl nodes are the points of contact between the electron pocket (green) and the hole pocket (red). (e) Band structure of MoP₂ along a path joining the W_1 and $W_{1'}$ Weyl nodes.



Figure 5: Atomic and electronic structure of the topological insulator $ZrTe_5$. (a) Crystal structure of $ZrTe_5$ showing the unit cell dimensions. (b) Atomic structure of individual layers of $ZrTe_5$. (c) ARPES measurements at two different photon energies (hv = 22 eV and hv = 23.5 eV) clearly showing both the bulk valence band dispersing in energy and the surface states. Image courtesy of A. Crepaldi (Sincrotrone Elettra, Trieste). (d) Predicted momentum-resolved local density of states at the (010) surface (red) and in the bulk (blue) along k_x calculated for the strong topological phase of $ZrTe_5$.

other quasi-one-dimensional bismuth halides. By substituting I atoms with Br, it is possible to synthesize a whole family of compounds with formula $Bi_4I_{4-x}Br_x$, most of which have never been reported previously. While all these materials are constructed from the same onedimensional blocks, their packing in the crystal structure is different. This stacking order has important consequences on the topological phase that can arise in the bulk crystals. By using first-principles calculations and symmetrybased arguments, we were able to predict that compounds which contains an even number of non-equivalent chains in the unit cell (Bi₄Br₄ and Bi₄I₂Br₂) are expected to be topologically trivial while materials with an odd numbers of non-equivalent chains (Bi₄I₄ and Bi₄I₃Br) are expected to be Z₂ topological insulators. The compound Bi₄I₃Br which contains three inequivalent chains is of particular interest. In known topological insulators, the topological phase arises from the inversion of a single band and a single topological phase transition is expected. Our first-principles calculations predict three inverted bands for Bi₄I₃Br originating from the three inequivalent chains. This gives rise to a very rich topological phase diagram for this compound whose topology can be controlled by various stimuli, such as applied pressure.

A first series of angle-resolved photoemission spectroscopy (ARPES) measurements on Bi_4Br_4 confirm our prediction that this material is a trivial insulator. Ongoing measurements on the topologically nontrivial compound Bi_4I_3Br are currently in progress.

c) Topological phase in transition-metal pentatelluride ZrTe₅ Having completed a joint experiment-theory project on $(Sb_2)_m - (Sb_2Te_3)$ natural superlattice topological materials [9], we proceeded towards another collaboration with the group of Marco Grioni (EPFL). The transition-metal pentatellurides ZrTe₅ and HfTe₅ are layered materials characterized by a non-symmorphic orthorhombic space symmetry group giving rise to intriguing topological properties (Fig. 5a, b). While the monolayers of these materials have been predicted to be large gap quantum spin Hall insulators [28], a debate still remains over the nature of the topological phase of the bulk crystals. We find that the electronic structure of these materials is very sensitive to the interlayer distance. In particular, when the experimental crystal structure is used, ZrTe₅ is predicted to be a strong topological insulator with Z_2 invariants (1;110) while a full relaxation of the structure leads to the prediction of a weak topological phase with Z_2 invariants (0;110).

These two phases can be differentiated by examining the (010) surface, which is predicted to exhibit topologically protected surface states only in the strong phase. Experimentally, the small size of the band gap makes the observation of the topological states a difficult task. Nevertheless, by carefully comparing our predictions with ARPES measurements, we were able to identify the correct phase of the material. ARPES scans performed at different photon energies reveal the presence of two bands at $\overline{\Gamma}$ with different band dispersion. One of the bands disperses strongly with photon energy changing from a parabolic shape to an M-like shape (Fig. 5c), while the second band is linear and dispersionless. This is in good agreement with our calculations on the strong topological phase, shown in Fig. 5d. We predict both the presence of a parabolic bulk band at Γ , which transforms into an M-shaped band at the *Y* point, as well as Dirac-like surface states. The manuscript reporting these results is currently in preparation.

2.4 Planned research for next year

- Extend the high-throughput search for Z₂ topological insulators and Weyl semimetals to ternary compounds.
- In-depth study of candidate materials identified through the high-throughput of binary materials.
- Collaborative efforts with the experimental group of Ming Shi (PSI) aiming at the investigation of predicted Weyl semimetals MoP₂ and WP₂.

2.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* The described research program implies close collaboration with Alexey Soluyanov and Dominik Gresch (group of Matthias Troyer, ETHZ) on developing methodologies for calculating topological invariants and efficient high-throughput search of novel topological materials.

b) *Outside MARVEL* Our project involves a growing number of groups with expertise in experimental techniques that can be used for synthesis, characterization and investigation of candidate topological materials.

- Materials synthesis and ARPES measurements of Weyl semimetal candidates: group of Ming Shi (PSI).
- Materials synthesis with emphasis on bismuth compounds: Anna Isaeva (TU Dresden).
- Angle-resolved photoemission spectroscopy (ARPES): group of Marco Grioni (EPFL).
- ARPES investigations at synchrotron facilities: Luca Moreschini (groups of Eli Rotenberg and Aaron Bostwick, Lawrence Berkeley National Laboratory).
- Laser-based ARPES measurements: group of Alessandra Lanzara (University of California, Berkeley).
- Transport measurements: Andrea Pisoni, László Forró (EPFL).

2.6 Personnel

NCCR personnel

• Gabriel Autès, postdoc, 100%, from July 2014.

Matching personnel

• Hyungjun Lee, postdoc, 50%, from May 2014.

3 DFT+DMFT study of strain and interface effects in d^1 and $d^2 t_{2g}$ -perovskites (Nicola Spaldin — ETHZ)

3.1 Research summary

The effect of epitaxial strain on the metalinsulator transition (MIT) was evaluated for several correlated oxides, such as $LaTiO_3$, $LaVO_3$, $SrVO_3$, and $CaVO_3$, with the aim of building a comprehensive and concise picture. We are currently investigating other important factors influencing the MIT in oxide heterostructures, such as the electron confinement and, next, the electronic reconstruction at the interfaces.

3.2 Scientific goals

The overall aim of this project is to address the most important aspects at the origin of metalinsulator-transition (MIT) phenomena within oxide heterostructures showing strong electronic correlation. Through atomistic simulations based on density functional theory (DFT) and dynamical mean field theory (DMFT), these aspects can be addressed one by one separately, or treated together in a more sophisticated model. We focus our efforts on correlated oxides with perovskite-type crystal structure and valence electrons of $3d-t_{2g}$ type with

 d^1 or d^2 fillings, such as LaTiO₃, SrVO₃, CaVO₃, LaVO₃, etc. In this way we plan to build a comprehensive picture of the MIT in this class of systems and at the same time identify the key parameters that could be tuned to obtain oxide heterostructures with tailored properties.

a) *Short-term* As a first step, we are currently exploring those aspects that do not depend explicitly on the details of the oxide-oxide interfaces, namely, the effect of epitaxial strain and electron confinement. The former occurs ubiquitously in heterostructures due to the lattice mismatch of the different constituents, while the latter is expected to play an important role when the film thickness or the superlattice period are particularly small. In some cases, one of these factors could be decisive in the onset of a MIT inside a thin film or other heterostructure, while in other cases the simultaneous occurrence of several factors is necessary in order to observe the MIT. The results and trends from these initial simulations will also serve as a reference to interpret the latter cases, which will be modeled within a more sophisticated framework that is able to catch their complexity (see here below).

b) Long-term Next we will address more complex models that include explicit interface effects, as well as the other aspects previously addressed, thus approaching the full complexity of a realistic heterostructure. These simulations will allow us to understand the role of electronic reconstruction at the interfaces, and also the interplay among all these factors in driving the system through a MIT. Different oxide material combinations can be employed to build up a heterostructure, but the insight from experiments and from the trends studied in first part will be used as a guide in the choice. In the long run, we plan to obtain a comprehensive picture that could be useful in the interpretation of experiments, but also in the prediction of novel MITs that might occur, in particular material combinations not yet explored.

3.3 Results obtained since the last report

We evaluated the effect of epitaxial strain in the MIT of several t_{2g} -perovskite oxides, namely LaTiO₃, LaVO₃, SrVO₃, and CaVO₃, a list that includes both metallic and insulating materials with d^1 or d^2 filling of the t_{2g} orbitals. The observed trends as a function of strain in the MIT were rationalized based on the modifications in the crystal-field splitting and in the bandwidth of these orbitals with strain. Our simulations



Figure 6: Representation of the two growth orientations for LaVO₃ thin films investigated in this work.

show that a compressive epitaxial strain comparable to that obtain in thin films coherently grown on SrTiO₃ can induce an insulator-tometal transition in LaTiO₃, but not in LaVO₃, consistently with experimental observations. The different t_{2g} -band filling, d^1 and d^2 , respectively, was found to explain the different behavior of these two materials in presence of a compressive strain.

For one of these materials (LaVO₃), we also characterized the effect of the crystallographic direction of growth on the energetic stability of the thin film and on the changes in the paramagnetic Mott insulating state due to strain [10]. The two growth orientations that were compared (Fig. 6) did not result into a significant energetic preference towards one or the other, unless a very strong compressive strain (above -3%) could be achieved throughout the film. The orientation that is favored at large compressive strains shows a complete suppression of the out-of-plane octahedral tilts, which is not observed for the moderate strain (s $\sim -0.5\%$) imposed by the SrTiO₃ substrate (neither in experiments nor in our simulations). If the tilt suppression takes place, the Mott insulating state of LaVO3 bulk is more significantly perturbed towards the metallic regime, but not enough to achieve a good metallic state. Owing to the effective half-filling of the t_{2g} orbitals, an important suppression of the orbital fluctuations at large compressive strain was evinced from our calculations, explaining why the t_{2g} bandwidth increase does not lead to a MIT (Fig. 7).

3.4 Planned research for next year

In the coming year, we plan to complete the ongoing calculations of the electronic confinement effect in $CaVO_3$, a material on the verge of a metal-insulator transition that could


Figure 7: Spectral density $A(\omega)$ of paramagnetic LaVO₃ at $T \simeq 300$ K from DMFT for the bulk geometry (a) and epitaxially-strained geometries with compressive strain (b) and tensile strain (c). The orbital-resolved contributions to $A(\omega)$ are shown with different line styles.

be triggered by the combined effect of tensile epitaxial strain and confinement effects within thin film geometries. The simulation of $LaVO_3/SrTiO_3$ heterostructures will likely be resumed, focusing on the electronic reconstruction effects of this system, which has not yet been addressed by simulations in the literature. Our final aim is to assess whether the metallicity observed at this correlated-insulator/bandinsulator interface could be driven by pure electronic effects, without the need of extrinsic electron doping from O-vacancies.

The DFT+DMFT calculations needed to obtain an accurate description of these systems are not yet out-of-the-box simulation tools and thus require some software consolidation activity that will be part of our efforts next year.

3.5 Synergies with other computational and experimental efforts

a) Within MARVEL In collaboration with the group of Antoine Georges, we are currently developing and testing a software tool that links two popular simulations packages (WAN-NIER90 and TRIQS). It will be then made publicly available to the MARVEL community and anyone else wishing to perform DFT+DMFT calculations with this kind of setup.

We are also in contact with the experimental group of Thorsten Schmitt at PSI in order to start an experimental verification project about vanadate thin films. This collaboration will be successively detailed in a separate proposal.

3.6 Personnel

NCCR personnel

- Gabriele Sclauzero, postdoc, 100%, from January 2015.
- Andrea Scaramucci, postdoc, 50% (PSI), from September 2014.

Matching personnel

- Claude Ederer, senior research, 10%, from July 2014.
- Krzysztof Dymkowski, PhD student, 23%, From February to June 2015.
- Michael Fechner, postdoc, 40%, from May 2015.
- Natalya Fedorova, PhD student, 25%, from May 2015.
- Alexander Hampel, PhD student, 20%, from August 2015.
- Andrea Scaramucci, postdoc, 50% (PSI), from September 2014.

4 High-temperature multiferroic spiral state by local frustration (Nicola Spaldin — ETHZ, Michel Kenzelmann — PSI)

4.1 Research summary

We explore the possibility of stabilizing a multiferroic spin-spiral state at high temperature by a small concentration of strong nearestneighbor frustrating couplings embedded in an otherwise unfrustrated spin lattice. Particular focus is given to the high-temperature spiral state observed in YBaCuFeO₅.

4.2 Scientific goals

a) *Short-term* Analysis of the conditions to stabilize a magnetic spiral state in an Heisenberg model on lattices with unfrustrated couplings when a small concentration of frustrating bonds are inserted. *Ab initio* calculations of magnetic couplings in compounds where local frustration is linked to the appearance of a magnetic spiral state.

b) *Long-term* Identify and propose for experimental investigation novel systems where disorder or small concentrations of impurities are present and can provide local magnetic frustration with features (e.g. direction of the frustrating bonds) that allow for stabilizing a multiferroic magnetic spiral state at high temperature.

4.3 Results obtained since the last report

The *ab initio* (LSDA+U) analysis of the structure and exchange couplings between magnetic ions in YBaCuFeO₅ [11] (which is not geometrically frustrated) reveals that the multiferroic magnetic spiral state occurring at high temperatures is not stabilized by the competition of nearest-neighbor and further neighbor exchange. Moreover, both ab initio results and experimental investigation of the structure indicate that a certain degree of disorder in the ordering of the magnetic Cu^{2+} and Fe^{3+} is present. According to our ab initio results the disorder in the magnetic ions is such that the sign of nearest-neighbor magnetic couplings is generally uniform and, in the *c* direction, is alternating ferromagnetic and antiferromagnetic. However, a small concentration of strong antiferromagnetic Fe³⁺–Fe³⁺ bonds along the *c* direction might be present and create local frustration in the otherwise ferromagnetically coupled *ab* layers.

In collaboration with Hiroshi Shinaoka a) (Troyer's group), Monte Carlo simulations were used to study magnetic properties of a Heisenberg model on the spin lattice of YBaCuFeO₅ where a small concentration of strong frustrating couplings (impurity bonds) in the *c* direction are embedded in the unfrustrated network of nearest-neighbor exchanges. Our study shows that, if certain constraint on the relative distance between the impurity bonds (distance between an impurity bond and neighboring layers above a threshold) are met, a spiral magnetic state can be stabilized. In agreement with experiments, the appearance of the magnetic spiral is preceded by the stability of a collinear commensurate magnetic state at higher temperatures.

Moreover, the found temperature dependence of the wave vector of the magnetic ordering (Fig. 8) is qualitatively in agreement with the observed one.

b) Encouraged by the Monte Carlo results we put efforts to understand the reasons and the conditions for which local frustrating bonds embedded in an unfrustrated spin lattice can stabilize a spiral state instead of other



Figure 8: Temperature dependence of the structure factor at $q = (1/2, 1/2, q_z)$ obtained by Monte Carlo simulations performed by H. Shinaoka for an effective Heisenberg model of YBaCuFeO₅. A concentration of x_{imp} strong antiferromagnetic impurity bonds is embedded in the weakly coupled ferromagnetically *ab* planes.

states (e.g. spin glass). To do this we considered the general case of a three dimensional Heisenberg model in a cubic lattice (with different nearest-neighbor couplings in the three directions) where a few of the magnetic couplings occurring in one direction are substituted by couplings with opposite sign.

We found that, assuming coplanar spin states (XY spins) and under the approximation used by Villain [29, 30, 31], the system can be mapped into Ising variables coupled through a dipole-dipole type of interaction. If strong enough, an impurity coupling induces locally a canting of the spins from the collinear state favored by the unfrustrated network of exchanges. At each impurity, the rotation direction of the canted magnetic moments (clockwise or counterclockwise) can be represented as an Ising variable. The stiffness of the remaining unfrustrating bonds induces a dipoledipole like interaction between such Ising degrees of freedom. When the position of the impurity bonds is such that a ferromagnetic state





Figure 9: Lowest energy state for impurity antiferromagnetic nearest-neighbor bonds (orange dashed lines) forming a superlattice embedded in an otherwise unfrustrated network of ferromagnetic nearest neighbor couplings. The left panel shows the state obtained minimizing the Villain model while the right panel shows numerical results obtained for the Heisenberg model. Arrows in the discs represent the magnetization of each layer while the arrows in the planes the direction of each spin (colors are related to the deviation of each spin from the average magnetization of its layer).

of the Ising variables is stable, i.e. the cantings happen in the same direction at all the impurity bonds, a spiral state (i.e. net rotation between all the layers perpendicular to the impurity bonds) is stabilized. The dependence of the effective coupling between the cantings on the relative position of the impurity bonds sheds light on the minimal distance constraint necessary to stabilize a spiral state in the Monte Carlo simulations.

The ground state of the Villain model can be found analytically for impurity bonds forming a periodic superlattices and the obtained results (stable spiral state or not) are qualitatively in agreement with numerical minimization (Fig. 9).

4.4 Planned research for next year

The research planned for next year aims to investigate the possibility that the mechanism found might play a role in other insulators (e.g. solid solutions of Fe_2O_3 and Cr_2O_3) where incommensurate states with concentration-dependent wave vector was found to appear [32]. It comprises the *ab initio* study of magnetic couplings for such systems and the verification for possible local frustration appearing along the trigonal axes of the corundum structure. Moreover, we will investigate the way in which magnetocrystalline anisotropy affects the mechanism for stabilizing a spiral state.

4.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Hiroshi Shinaoka (group of Matthias Troyer) performed Monte Carlo simulations.

b) *Outside MARVEL* M. Morin and M. Medarde at Paul Scherrer Institute are currently investigating changes of the feature of the spiral state by changing the preparation conditions of YBaCuFeO₅. M. Mostovoy at University of Groningen and C. Mudry at Paul Scherrer Institute are involved in the part concerning the understanding of the mechanism stabilizing the spiral state.

4.6 Personnel

NCCR personnel

• Andrea Scaramucci, postdoc, 50%, from September 2014.

Matching personnel

• Andrea Scaramucci, postdoc, 50% (PSI), from September 2014.

5 Controlling the electronic state of transition-metal oxides (Antoine Georges — UniGE)

5.1 Research summary

- Control of metal-insulator transition in transition-metal oxides.
- Identification of the nature of the insulating state in rare-earth nickelates *R*NiO₃ in relation to optical spectroscopy.

5.2 Research question

Transition-metal oxides (TMOs) present a large diversity of electronic properties, which offer potential promises for electronics or energyrelated applications. Examples are: the colossal magnetoresistance of manganites, the large thermopower of cobaltates, the metal-insulator transition of vanadates or nickelates, the superconductivity of cuprates, etc. These properties originate from extreme responses of a material to external perturbations, which result from a strong coupling between various degrees of freedom and multiple competing phases. The behavior of strongly correlated materials depends on the details of electronic structure, such as bandwidth, orbital population, effective interaction parameters, which renders their functionality relatively easy to control by temperature, chemical composition, crystal structure, electromagnetic field, etc. One of the key research questions is how to achieve new phases with desired properties in a predictable manner by tuning these control parameters.

5.3 Scientific goals

The primary research goal is to investigate the control parameters of the metal-insulator transition (MIT) in rare-earth nickelates and explore the prospects of designing new materials by "orbital engineering".

Rare-earth nickelates, *R*NiO₃, is a very good example of a class of materials in which the MIT can be controlled by a variety of ways, e.g. by chemical composition (choice of the rare-earth cation), pressure, epitaxial strain, gating, light, etc. In all these cases external forces induce changes in a number of crystal structure parameters, such as volume, bond angles, octahedral distortions. The key question is to understand how these changes are coupled to the electronic structure, affecting thus the MIT, and a great deal of progress has been recently achieved about this issue [33, 34, 35], partly as a direct result of MARVEL [12, 13].

a) *Short-term* We have developed [12] a description of the nature of the insulating phase of nickelates which provides both qualitative understanding in terms of a limited number of parameters characteristic of the electronic structure, and predictive power based on LDA+DMFT calculations. Our immediate short-term goals will make use of this framework and address the following points.

• Understand the trends over the rare-earth series. The temperature of the MIT in *R*NiO₃ has a strong variation with the rare-earth cation size, which reveals a strong effect of crystal structure on the electronic structure. As one follows the series from PrNiO₃ to LuNiO₃ changes in the Ni-O-Ni bond angles result in both the band narrowing and the modification of the Peierls pseudogap in the band structure induced by the bond disproportionation. We are planning to consider the two

extreme cases, $PrNiO_3$ and $LuNiO_3$, with the aim to understand how the structural changes affect the MIT phase boundaries.

- Unlike other members of the nickelate series LaNiO₃ is the only one that remains metallic at all temperatures in the bulk. However, it undergoes a robust MIT in ultrathin films (1 - 4 unit cells (uc)) and in superlattices of 1 - 2 uc of LaNiO₃ combined with insulating layers (such as LaAlO₃). The nature of the insulating phases in such geometries is still unknown. Experimental characterizations of ultrathin films show no evidence of bond disproportionation, which suggests that the insulating phase might be different from that of other nickelates. Our goal is to investigate the nature of the insulating phase of LaNiO₃ in ultrathin films (1 - 2 uc). Crystal structure determination will be crucial here, because the MIT appears to be related to the local symmetry breaking associated with distortions at the interface. This will be a collaborative work with the group of Sohrab Ismail-Beigi, Yale University.
- We shall also address the nature of the magnetic ordering at low temperatures. Whether the low-energy description that we have elaborated can account for the observed $\mathbf{Q} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ will be a key test.
- b) Long-term
 - A study of systematic trends in the crystal structure of nickelates and of their influence on the MIT is going to be performed within MARVEL in collaboration with Claude Ederer and coworkers (Nicola Spaldin group, ETHZ) on the theory side, and with the group of Marisa Medarde (PSI) on the materials and experimental side (approved MARVEL "experimental validation" joint project).
 - "Orbital engineering" in oxides.

Nickelates (LaNiO₃) were also proposed as a candidate for material design by means of "orbital engineering". Specifically, since in the bulk Ni ions nominally have a single electron occupying the doubly degenerate e_g shell (d^7 state), one can imagine designing a superlattice in which the degeneracy is lifted and the nickelate becomes a single-band half-filled system analogous to cuprates with much stronger correlation effects including a potential

for high- T_c superconductivity [36]. However, more detailed studies, both theoretical and experimental [37, 38][14], have shown that in reality it is difficult to achieve a sufficient amount of orbital polarization in this material, due to the combined effect of strong covalency with oxygens and Hund's coupling, resulting in a large contribution of the d^8 orbital compensated configuration. In future studies, materials where these effects are less pronounced will be considered, such as early transition-metal oxides. There, the valence electrons occupy less covalent t_{2g} orbitals and a potential transformation of a triply degenerate shell into a half-filled single orbital might have a dramatic effect on the properties of the system. In this respect, heterostructures based on vanadates and titanates look very promising. Collaborations with Claude Ederer et al. (ETHZ) on theory and Jean-Marc Triscone et al. (UniGE) is being considered on these materials.

5.4 Results obtained since the last report

a) *Nickelates* In the previous research period we developed [12] a low-energy description of rare-earth nickelates based only on hybridized states around the Fermi level possessing the e_g symmetry of Ni ions. The main conclusion was that it is indeed possible to describe the MIT in RNiO₃ provided that the effective interaction parameters, which include Coulomb repulsion *U* and Hund's coupling *J*, satisfy a condition $U - 3J \lesssim \Delta_s$, where Δ_s is the splitting between on-site energies of long-bond (LB) and short-bond (SB) sites in the low-Tlow-symmetry structure characterized by bond disproportionation. In this regime we could identify a large range of parameters *U* and *J* in which the high-T (orthorhombic) phase was a correlated metal, while the low-*T* (monoclinic) was insulating, with the insulating state characterized by a large disproportionation of occupancy between LB and SB sites. The question remained whether this description is relevant for real nickelates and if yes, what is the exact parameter range corresponding to particular RNiO₃.

To answer both questions we have calculated optical spectra at energies corresponding to the intraband e_g - e_g transitions (infrared region). The motivation is that optical response allows one to probe exactly the energy scale relevant for our low-energy framework and theoretical results can be compared directly to experiment.



Figure 10: Optical spectra of SmNiO₃ in the insulating phase calculated using DFT+DMFT using a low-energy (e_g -only) description. The double-peak structure of the spectra, also observed in experiment, is signature of the bond-disproportionated insulating (BDI) phase.

The calculated optical spectra for SmNiO₃ are presented in Fig. 10 displaying results for several values of U and J corresponding to the bond-disproportionated insulating phase (partial phase diagram is shown in the inset). The main feature of these spectra is a two-peak structure present in all the considered cases. Detailed analysis of the electronic structure has revealed that this two-peak structure is a hallmark of the interplay between the Mott phenomenon responsible for the gap opening and the Peierls splitting of bands which creates a pseudogap in the unoccupied part of the single-particle spectrum.

Importantly, infrared ellipsometry performed in the group of Dirk van der Marel (UniGE) on three different nickelate systems (prepared in the group of Jean-Marc Triscone) precisely revealed this two-peak structure in the same energy range in the insulating phase. At higher temperatures, in the metallic phase, the spectra are very different and show only one peak at a different position.

Hence a clear signature of the nature of the MIT in those materials was identified, calculated theoretically and observed experimentally, and recently presented in a joint publication [13] with the mentioned experimental groups at UniGE-DQMP.

b) *Electronic and thermoelectric properties of ruthenates* A theoretical study of the electronic structure of the ruthenates $CaRuO_3$ and $SrRuO_3$ was conducted in collaboration with Jernej Mravlje (IJS, Ljubljana) and the group of Andrew J. Millis (Columbia) [15, 16]. It was shown in particular [15] that the puzzling deviations from Drude behavior reported ex-

perimentally might largely be due to interband optical transitions extending down to the Terahertz scale. Thermoelectric properties of Sr_2RuO_4 were considered in another submitted publication [17] in which the *ab*- and *c*-axis Seebeck coefficient of this material were computed from DFT+DMFT for the first time, and a prediction made for future experiments.

c) Participation in the development of opensource software Oleg Peil contributed [18] to the elaboration of an updated version of the electronic structure package of the TRIQS library [39], which provides a set of open-source tools for electronic structure calculations combining DFT and DMFT.

5.5 Planned research for next year

As summarized above, in short-term and long-term goals:

- Systematics of the *RNiO*₃ series (collaboration with ETHZ, PSI and UniGE).
- MIT in ultrathin films of LaNiO₃ (collaboration with Yale).
- Magnetic structure of *R*NiO₃.
- Heterostructures based on *t*_{2g}-oxides: titanates, vanadates (collaboration with ETHZ).
- Finalization of the implementation of a user-friendly interface between WAN-NIER90. and the electronic structure part of the TRIQS library [39][18] (collaboration with ETHZ).

5.6 Synergies with other computational and experimental efforts

- a) Within MARVEL
 - ETHZ-group of Nicola Spaldin: Claude Ederer and coworkers.
 - PSI: Marisa Medarde within the experimental verification project.
- b) Outside MARVEL
 - Synergies with the groups of Jean-Marc Triscone and Dirk van der Marel (UniGE-DQMP).
 - Synergy in methods and software development with the group of Olivier Parcollet, IPhT-CEA-Saclay.
 - Collaboration with Jernej Mravlje (IJS, Ljubljana, Slovenia) on thermoelectric properties of transition-metal oxides.

- Collaboration with the group of Andrew J. Millis (Columbia University) on the electronic properties of ruthenates.
- Starting collaboration with the group of Sohrab Ismail-Beigi, Yale University.

5.7 Personnel

NCCR personnel

• Oleg Peil, postdoc, 100%, from September 2014.

Matching personnel

• Edward Perepelitsky (based in Paris), postdoc, from September 2014, on Ecole Polytechnique (France) funds.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- [1] A. Subedi, O. E. Peil, and A. Georges, *A low energy description of the metal-insulator transition in the rare-earth nickelates*, arXiv:1410.2830 (2014).
- [2] D. Gresch, A. A. Soluyanov, G. Autès, O. Yazyev, B. A. Bernevig, D. Vanderbilt, and M. Troyer, Universal framework for identifying topological materials and its numerical implementation in Z2Pack software package (2015), http://z2pack.ethz.ch.
- [3] A. A. Soluyanov, D. Gresch, Z. Wang, Q. Wu, M. Troyer, X. Dai, and B. A. Bernevig, *Type-II Weyl* semimetals, Nature 527, 495 (2015).
- [4] G. Autès, A. A. Soluyanov, D. Gresch, M. Troyer, and O. V. Yazyev, *Robust Type-II Semimetal Phase in Transition Metal Diphosphides XP*₂ (X = Mo, W), *in prepara tion* (2015).
- [5] G. Autès, A. Isaeva, L. Moreschini, J. C. Johannsen, A. Pisoni, R. Mori, W. Zhang, T. G. Filatova, A. N. Kuznetsov, L. Forró, W. Van den Broek, Y. Kim, K. S. Kim, A. Lanzara, J. D. Denlinger, E. Rotenberg, A. Bostwick, M. Grioni, and O. V. Yazyev, A novel quasi-one-dimensional topological insulator in bismuth iodide β-Bi₄I₄, Nature Materials 15, 154 (2016).
- [6] A. A. Soluyanov, D. Gresch, M. Troyer, R. M. Lutchyn, B. Bauer, and C. Nayak, *Optimizing spin-orbit splittings in InSb Majorana nanowires, to be published in* Physical Review B (2015), arXiv:1510.02025.
- [7] Z. Wang, D. Gresch, A. A. Soluyanov, W. Xie, S. Kushwaha, X. Dai, M. Troyer, R. J. Cava, and B. A. Bernevig, *MoTe₂: Weyl and Line Node Topological Metal*, arXiv:1511.07440 (2015).
- [8] G. W. Winkler, A. A. Soluyanov, and M. Troyer, Smooth gauge and Wannier functions for topological band structures in arbitrary dimensions, Physical Review B 93, 035453 (2016).
- [9] J. C. Johannsen, G. Autès, A. Crepaldi, S. Moser, B. Casarin, F. Cilento, M. Zacchiagna, H. Berger, A. Magrez, P. Bugnon, J. Avila, M. C. Asensio, F. Parmigiani, O. V. Yazyev, and M. Grioni, Engineering the topological surface states in the (Sb₂)_m-Sb₂Te₃ (m = 0 - 3) superlattice series, Physical Review B **91**, 201101 (2015).



- [10] G. Sclauzero and C. Ederer, Structural and electronic properties of epitaxially strained LaVO₃ from density functional theory and dynamical mean-field theory, Physical Review B 92, 235112 (2015).
- [11] M. Morin, A. Scaramucci, M. Bartkowiak, E. Pomjakushina, G. Deng, D. Sheptyakov, L. Keller, J. Rodriguez-Carvajal, N. A. Spaldin, M. Kenzelmann, K. Conder, and M. Medarde, *Incommensurate magnetic* structure, *Fe/Cu chemical disorder, and magnetic interactions in the high-temperature multiferroic YBaCuFeO*₅, Physical Review B **91**, 064408 (2015).
- [12] A. Subedi, O. E. Peil, and A. Georges, Low-energy description of the metal-insulator transition in the rare-earth nickelates, Physical Review B 91, 075128 (2015).
- [13] J. Ruppen, J. Teyssier, O. E. Peil, S. Catalano, M. Gibert, J. Mravlje, J.-M. Triscone, A. Georges, and D. van der Marel, *Optical spectroscopy and the nature of the insulating state of rare-earth nickelates*, Physical Review B 92, 155145 (2015).
 - [14] O. E. Peil, M. Ferrero, and A. Georges, Orbital polarization in strained LaNiO₃: Structural distortions and correlation effects, Physical Review B 90, 045128 (2014).
 - [15] H. T. Dang, J. Mravlje, A. Georges, and A. J. Millis, Band Structure and Terahertz Optical Conductivity of Transition Metal Oxides: Theory and Application to CaRuO₃, Physical Review Letters 115, 107003 (2015).
 - [16] H. T. Dang, J. Mravlje, A. Georges, and A. J. Millis, Electronic correlations, magnetism, and Hund's rule coupling in the ruthenium perovskites SrRuO₃ and CaRuO₃, Physical Review B 91, 195149 (2015).
 - [17] J. Mravlje and A. Georges, *Thermopower and Entropy:* lessons from Sr₂RuO₄, arXiv:1504.03860 (2015).
- [18] M. Aichhorn, L. Pourovskii, P. Seth, V. Vildosola, M. Zingl, O. E. Peil, X. Deng, J. Mravlje, G. J. Kraberger, C. Martins, M. Ferrero, and O. Parcollet, *TRIQS/DFTTools: A TRIQS application for* ab initio calculations of correlated materials, arXiv:1511.01302 (2015).

Other references

- [19] R. M. Lutchyn, J. D. Sau, and S. D. Sarma, Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures, Physical Review Letters 105, 077001 (2010).
- [20] Y. Oreg, G. Refael, and F. von Oppen, *Helical liquids and Majorana bound states in quantum wires*, Physical Review Letters 105, 177002 (2010).
- [21] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, *Topological semimetal and Fermi-arc surface* states in the electronic structure of pyrochlore iridates, Physical Review B 83, 205101 (2011).
- [22] H. Weng, C. Fang, Z. Fang, B. A. Bernevig, and X. Dai, Weyl Semimetal Phase in Noncentrosymmetric Transition-Metal Monophosphides, Physical Review X 5, 011029 (2015).
- [23] B. Q. Lv, H. M. Weng, B. B. Fu, X. P. Wang, H. Miao, J. Ma, P. Richard, X. C. Huang, L. X. Zhao, G. F. Chen, Z. Fang, X. Dai, T. Qian, and H. Ding, *Experimental Discovery of Weyl Semimetal TaAs*, Physical Review X 5, 031013 (2015).
- [24] S.-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C.-C. Lee, S.-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, *Discovery of a Weyl fermion semimetal and topological Fermi arcs*, Science **349**, 613 (2015).

- [25] S.-M. Huang, S.-Y. Xu, I. Belopolski, C.-C. Lee, G. Chang, B. Wang, N. Alidoust, G. Bian, M. Neupane, C. Zhang, S. Jia, A. Bansil, H. Lin, and M. Z. Hasan, A Weyl Fermion semimetal with surface Fermi arcs in the transition metal monopnictide TaAs class, Nature Communications 6, 7373 (2015).
- [26] C. L. Kane and E. J. Mele, Z₂ topological order and the quantum spin Hall effect, Physical Review Letters 95, 146802 (2005).
- [27] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, *wannier90: A tool for obtaining maximally-localised Wannier functions.*, Computer Physics Communications **178**, 685 (2008).
- [28] H. Weng, X. Dai, and Z. Fang, Transition-Metal Pentatelluride ZrTe₅ and HfTe₅: A Paradigm for Large-Gap Quantum Spin Hall Insulators, Physical Review X 4, 011002 (2014).
- [29] J. Villain, Theory of one- and two-dimensional magnets with an easy magnetization plane. II. The planar, classical, two-dimensional magnet, Journal de Physique (Paris) 36, 581 (1975).
- [30] J. Villain, Two-level systems in a spin-glass model. I. General formalism and two-dimensional model, Journal of Physics C 10, 4793 (1977).
- [31] J. Villain, Two-level systems in a spin-glass model: II. Three-dimensional model and effect of a magnetic field, Journal of Physics C 11, 745 (1978).
- [32] D. E. Cox, W. J. Takei, and G. Shirane, A magnetic and neutron diffraction study of the Cr₂O₃-Fe₂O₃ system, Journal of the Physics and Chemistry of Solids 24, 405 (1963).
- [33] I. I. Mazin, D. I. Khomskii, R. Lengsdorf, J. A. Alonso, W. G. Marshall, R. M. Ibberson, A. Podlesnyak, M. J. Martínez-Lope, and M. M. Abd-Elmeguid, *Charge Ordering as Alternative to Jahn-Teller Distortion*, Physical Review Letters **98**, 176406 (2007).
- [34] H. Park, A. J. Millis, and C. A. Marianetti, Site-Selective Mott Transition in Rare-Earth-Element Nickelates, Physical Review Letters 109, 156402 (2012).
- [35] S. Johnston, A. Mukherjee, I. Elfimov, M. Berciu, and G. A. Sawatzky, *Charge Disproportionation without Charge Transfer in the Rare-Earth-Element Nickelates as* a Possible Mechanism for the Metal-Insulator Transition, Physical Review Letters 112, 106404 (2014).
- [36] J. Chaloupka and G. Khaliullin, Orbital Order and Possible Superconductivity in LaNiO₃/LaMO₃ Superlattices, Physical Review Letters 100, 016404 (2008).
- [37] M. Han, C. Marianetti, and A. Millis, Chemical control of orbital polarization in artificially structured transitionmetal oxides: La₂NiXO₆ (X=B,Al,Ga,In) from first principles, Physical Review B 82, 134408 (2010).
- [38] M. Wu, E. Benckiser, M. W. Haverkort, A. Frano, Y. Lu, U. Nwankwo, S. Brück, P. Audehm, E. Goering, S. Macke, V. Hinkov, P. Wochner, G. Christiani, S. Heinze, G. Logvenov, H.-U. Habermeier, and B. Keimer, *Strain and composition dependence of orbital polarization in nickel oxide superlattices*, Physical Review B 88, 125124 (2013).
- [39] O. Parcollet, M. Ferrero, T. Ayral, H. Hafermann, I. Krivenko, L. Messio, and P. Seth, *TRIQS: A toolbox for research on interacting quantum systems*, Computer Physics Communications **196**, 398 (2015).

Vertical Project 2 VP2 — Novel Materials Applications

Project leader: Ursula Röthlisberger (EPFL)

Participating members: Ursula Röthlisberger (EPFL), Alfredo Pasquarello (EPFL), Jürg Hutter (UZH), Nicola Marzari (EPFL), Alessandro Curioni (IBM), Wanda Andreoni (EPFL), Clémence Corminboeuf (EPFL), Daniele Passerone (Empa)

Summary and highlights: During the second year of MARVEL, we have streamlined the internal structure of VP2 to focus its research activities. In particular, we have defined three research themes that should guarantee a good balance between targeted computational design and broader scope discovery.

Among the important achievements of the second year are a joint experimental/theoretical study that identified the molecular origin of J - V hysteresis effects in halide perovskites (Röthlisberger), the investigations of water splitting catalysts via a mechanistic characterization of the reaction path for cobalt-cubane based oxidation catalysts (Hutter) and the establishment of linear scaling relationships and volcano plots in homogenous catalysis (Corminboeuf). Work in progress involves high-throughput screening of candidate materials for next-generation solid-state lithium conductors, in a joint collaboration between the Marzari group and IBM, perovskites for photocatalytic water splitting (Marzari), and exfoliation of novel 2D materials (Marzari). Important technical milestones have been the development of a standard solid-state pseudopotential library (Marzari) and the development of an efficient scheme for accurate GW calculations of band gaps (Pasquarello).

General view of the project

VP2 is one of the main application pillars of MARVEL and is thus key for the overall mission of MARVEL, i.e. the computational design and discovery of novel materials. VP2 also incorporates an important translational function between the methodological developments in electronic structure (HP3) and sampling techniques (HP4) and their applications to problems of practical interest. This close interplay between theory and application serves for the testing and benchmarking of new methodologies and also enables VP2 to make new advances in materials design based on innovative leading-edge developments in computational methods. VP2 also produces and/or relies on the ready availability of big amounts of data which creates intimate links between VP2 and and materials informatics in HP5. VP2 also fosters close collaborations with experiments (with the experimental MARVEL platform PP7 and with external groups) for a continuous validation of the theoretical predictions and cultivates contacts with industrial partners in view of possible future commercializations.

Major research questions

VP2 aims at the fundamental atomistic/electronic understanding of matter and the computational design and discovery of novel materials. With 8 group leaders and over 10 subprojects, VP2 is one of the largest projects within MARVEL. A well focussed structure that creates optimal cohesion and a suitable degree of coherence is an important issue for the success of VP2. Following the recommendations of the scientific panel, we have therefore re-organized and re-focussed the research in VP2 into a new horizontal/vertical structure. The horizontal structure is defined by three principal research themes:

- 1. Light harvesting
- 2. Energy storage
- 3. Novel materials

The major part of applications in VP2 (themes 1 and 2) deals with the characterization and design of materials for sustainable energy solutions, in particular these activities are targeted towards novel materials for light capture and energy storage.

The *light harvesting* theme encompasses projects on photovoltaics, namely solar cells based on carbon nanostructures (Passerone), dye-sensitized solar cells (Röthlisberger), organic/inorganic lead halide perovskites as light harvesting and charge transport materials (Röthlisberger) and rare-earth perovskites



Figure 1: New organization in VP2 showing how the activities of the various group leaders distribute among the three principal themes of VP2: light harvesting, energy storage, and search of novel materials.

for photocatalytic water splitting (Marzari). Theme 2 on *energy storage* includes the characterization and design of water oxidation catalysts (Hutter and Corminboeuf), and the study and design of materials for battery applications: metal/air batteries (Curioni), as well as lithium ion and proton conductors (Marzari).

To reach a balance between highly focussed projects targeted at the search for materials with specific predefined properties and the materials discovery aspect of MARVEL, we defined the third research theme of VP2 as novel materials to leave also room for spontaneously emerging opportunities and materials with new functionalities. This third activity embraces projects with a wider scope and a broader variety of envisaged properties. To this category belong the high-throughput screening of 2D materials (Marzari), the design of organic molecules for applications in electronics and spintronics (Corminboeuf), the investigation of novel mixed perovskites (Röthlisberger), the investigation of garnet type lithium conductors (Curioni), the screening of carbon nanostructures for electronics applications (Passerone) and the studies of amine solutions for CO₂ capture (Andreoni). An overview of the distribution of the different subprojects over the three research themes is shown in Fig. 1.

On top of this horizontal project organization, we defined a vertical classification (a-e) of all VP2 application subprojects as a function of their position along the translation from fundamental aspects to a marketable product: essentially all subprojects within VP2, start with a characterization of the fundamental properties of a material (a); followed by a definition of relevant key features or descriptors (b); that can be used for high-throughput screening (c); leading to the identification of candidate materials that can be synthesized and tested (d); and might ultimately lead to commercialization in collaboration with an industrial partner (e). As mentioned before, the subprojects within VP2 also share several unifying concepts that link them tightly to other activities within MARVEL such as the testing and application of new computational methodologies (HP3 and HP4), the need for a common database to share and exchange results (HP5) and the collaboration with experiments for constant validation (PP7 and external experimental groups).

Long-term goals

The long-term goal of VP2 is on the one hand the gain of a fundamental understanding of the electronic and atomistic structure of materials with interesting properties, in particular for applications in sustainable energy strategies. Part of this activity is also combined with the testing and validation of the new computational methodologies developed in MARVEL. In addition, for several subprojects within VP2, the long-term aim is the computational design and/or discovery of new materials that can be synthesized, experimentally characterized and might in some cases lead to a new marketable product.

Achievements

Below, we list some of the main achievements in VP2 during the second year of MARVEL following the newly defined structure of horizontal research themes (1-3) and vertical classification (a-e) according to the type of pursued activity.

1) Light harvesting The Röthlisberger group has provided the theoretical basis for improving the phase stability of halide perovskites via mixing of the organic cation (1a). They have also designed a panchromatic ullazine based dye sensitizer (1a) and were able to rationalize the first *in situ* atomic force microscopy (AFM) measurements of a dye-sensitized solar cell under working conditions by predicting the arrangements and packing of the dyes as a function of coverage (1a). They also explained the origin of the charge transfer excitations that occur in mixed binary and ternary mixtures of the organic photovoltaic systems formed by C_{60} and fluorinated zinc phtalocyanines (1a). For all four projects, the theoretical predictions were validated by synthesis and experimental

characterization of the systems in collaboration with the Graetzel group at EPFL [1, 2, 3, 4]

2) Energy storage The Hutter group was able to characterize the mechanistic pathway for water oxidation catalyzed by substituted cobalt-cubane systems (2a).

The Corminboeuf group rationalized the activity of transition metal free graphite for water oxidation (2a) and established relationships between thermodynamic descriptors and activation barriers (2b).

The Curioni group studied conductivity in solid-state electrolytes with classical potentials (2a) and screened the ICSD database for intrinsic properties (2c).

3) Novel materials The Mazari group was able to extract 6'000 layered 2D materials from crystallographic databases and computated their binding energies (3c). They also performed a stability analysis of SrRuO₃ and LaRuO₃ perovskites as candidate materials for the oxygen evolution reaction (3a) and identified perovskite materials as promising candidates for water splitting devices (3c).

Testing and benchmarking of new computational methodologies Several subprojects within VP2 are dedicated to the development and testing of computational methods needed for VP2 applications.

The Passerone group identified heterojunction systems as test cases for Koopman's schemes (in collaboration with the Marzari group).

The Pasquarello group performed GW calculations for water taking into account nuclear quantum effects. They also tested the performance of hybrid functionals in combination with nonlocal van der Waals corrections and assessed the Quantum-ESPRESSO and the CP2K codes for the same system.

The Corminboeuf group tested the reliability of computational methodologies for quasiclose-shell molecules.

The Marzari group tested new semiclassical approaches to model ionic conductivity.

Data management and automated workflows The building of a common platform for data sharing/exchange and the establishment of automated workflows are important components for enhancing work efficiency in VP2. Several subprojects have made important steps in this direction.

The Passerone group released a beta version of an AiiDA plugin for CP2K (in collaboration with the Mazari group).

The Marzari group assessed and implemented high-throughput methods to calculate diffusion coefficients from molecular dynamics trajectories.

Experimental validation/industrial collaborations Constant experimental validation studies are of high importance for all VP2 projects. Several groups have established contacts with experimental groups (see also list of collaborations) and some joint experimental/computational studies have already led to publications during the second year of MARVEL.

The Röthlisberger group was able to identify ionic polarization as the phenomenon at the origin of the observed hysteresis in J - Vcurves of perovskite solar cells. The theoretical predictions were verified qualitatively as well as quantitatively via temperature dependent measurements [5].

The Corminboeuf and Röthlisberger groups have initiated new contacts with industrial partners.

Milestones

April 2016

- Water oxidation pathways for substituted Co-cubane systems (Hutter).
- Release of a beta version of the AiiDA plugin for CP2K (Passerone).
- Identification of heterojunction geometries to be simulated with Koopman's schemes (Passerone).
- Defects in amorphous Al₂O₃ causing structural degradation (Pasquarello).
- Hybrid functional for water including nonlocal van der Waals (vdW) interactions (Pasquarello).
- Comparison between Quantum-ESPRESSO and CP2K for liquid water simulation (Pasquarello).
- GW description of the electronic structure of liquid water including quantum nuclei effects (Pasquarello).
- Completion of the framework based on kinetic Monte Carlo (KMC) and classical potential to study conductivity in solid-state electrolyte (Curioni).
- Screening of ICSD database on intrinsic properties (Curioni).



- Manipulating and rationalizing the activity of transition metal free graphite for water oxidation (Corminboeuf).
- Establishing scaling relationships between a thermodynamic descriptor and activation barriers for the elaboration of "kinetic volcano plots" (Corminboeuf).
- Gauging the reliability of the computational methodology for recognizing the quasiclosed-shell (QCS) molecules among known experimental examples (Corminboeuf).
- Assessment and implementation of methods to compute diffusion coefficients from molecular dynamics trajectories in a highthroughput approach (Marzari).
- Testing of novel semiclassical approaches to modeling ionic conductivity (Marzari).
- Extraction of 6'000 layered materials from crystallographic databases and computations of binding energies (Marzari).
- Stability analysis (phase and Pourbaix diagrams) of SrRuO₃ and LaRuO₃ perovskites as candidate materials for the oxygen evolution reaction (Marzari).
- Identification of few perovskite materials with high potential to be used in a water splitting device (Marzari).
- Through a joint theoretical/experimental effort, ionic polarization was identified as the microscopic phenomena at the origin of the hysteresis in J V curves of mixed organic/inorganic metal halide perovskites (Röthlisberger).
- The phase stability of halide perovskites was improved via mixing of the organic cation (Röthlisberger).
- Design of a panchromatic ullazine dye sensitizers with the help of density functional theory (DFT) calculations (Röthlisberger).
- The first AFM measurements of dyesensitized solar cell devices under operating conditions were rationalized with classical molecular dynamics simulations (Röthlisberger).
- Identification of the nature of charge transfer states in binary and ternary mixtures of C₆₀ and fluorinated zinc phtalocyanines (Röthlisberger).

April 2017

- Benchmark results on the MARVEL cobalt water oxidation complexes (Hutter).
- Elucidation of mechanisms of Co-based water oxidation catalysts (Hutter).
- Elucidation of catalytic water reduction pathways for metal pyrphyrin complexes (Hutter).
- DFT study of contaminants (O and H) in heterojunctions and their effect on the electronic properties (Passerone).
- Application of Koopman's schemes to heterojunctions containing porphyrin molecules (Passerone).
- GW schemes applied to heterojunctions (Passerone).
- Band alignment at photocatalytic semiconductor/water interfaces (Pasquarello).
- Benchmark calculations for redox levels validating of the new hybrid-vdW functional (Pasquarello).
- Application of constant Fermi-level simulations to redox levels in water (Pasquarello).
- Application of the framework to several (~100) Li-ion screened material from ICSD (Curioni).
- Extension to Na-ion screened material from ICSD (Curioni).
- Application and validation of molecular volcano plots for hydrogen evolution and oxygen evolution reactions (Corminboeuf).
- Prediction of novel cheap and abundant catalysts for water splitting based on molecular volcano plots (Corminboeuf).
- *In silico* design of diverse new neutral and charged quasi-closed-shell molecules, formulation of the design principles (Corminboeuf).
- Finding descriptors for ionic diffusivity (Marzari).
- Computation of energy relaxation and band structures with four different functionals for 100'000 crystal structures, and of phonon dispersions for 20'000 of them (Marzari).

- High-throughput screening of pyrochlore materials for the oxygen evolution reaction (Ru and Ir based) and for water splitting applications (Marzari).
- Identification of rare-earth perovskites for water splitting (Marzari).
- Verification of normconserving pseudopotentials (Marzari).
- Theoretical rationalization of the temperature dependent photoluminescence behavior of lead halide perovskites (Röthlisberger).
- Development of a polarizable force field for mixed organic/inorganic halide perovskites (Röthlisberger).
- Development of a genetic algorithm optimization protocol for mixed perovskites (Röthlisberger).

April 2018

- Design of new Co-cubane derived water oxidation catalysts (Hutter).
- Photophysical properties of selected heterojunction with TDDFT based non-adiabatic quantum-classical molecular dynamics (Passerone).
- Assessment of vdW interaction schemes validated by experimental results: high-throughput screening (Passerone).
- Study of surface related processes for water splitting at photocatalytic semiconductor/water interfaces (Pasquarello).
- Embedding and adaptation of the framework within the cognitive engine developed in HP5 (Curioni).
- Delivery of a cognitive platform for material design based on the cognitive engine developed in collaboration with HP5 (Curioni).
- Elaboration of tools for catalyst design (Corminboeuf).
- Application of molecular volcano plots to key reactions in homogeneous catalysis and molecular electrochemistry for prediction of novel cheap and abundant catalysts (Corminboeuf).
- Computational evaluation of the designed characteristics of molecules, relevant to potential molecular electronics and spintronics applications, e.g. conductance, redox switching, pH-switching, etc. (Corminboeuf).

- Application of descriptors to a large-scale screening effort within the AiiDA frame-work (Marzari).
- Computation of thermodynamic properties at the quasiharmonic level for 10'000 compounds (Marzari).
- Calculations of phase diagrams for organic/inorganic halide perovskites via minima hopping based on polarizable force fields and DFT calculations (Röthlisberger in collaboration with the Goedecker group, HP4).
- Simulation of crystal growth of halide perovskites via variationally enhanced sampling and polarizable force fields (Röthlisberger in collaboration with the Parrinello group, HP4).

New and ongoing collaborations

- Hutter group (UZH) and Pasquarello group (EPFL), 2014, *Application of hybrid functionals for liquids and interfaces*.
- Hutter group (UZH) and Corminboeuf group (EPFL), 2015, *Application of* CP2K *in materials simulations*.
- Hutter group (UZH), Patzke group (UZH) and Smolentsev group (PSI), 2015, *Time*resolved X-ray absorption spectroscopy to investigate mechanisms of photochemical water splitting reactions with molecular catalysts.
- Passerone group (Empa) and Marzari group (EPFL), 2015, CP2K *plug-in for AiiDA*.
- Passerone group (Empa) and Curioni group (IBM), 2015, *Photophysics of carbon nanoribbons*.
- Passerone group (Empa) and Hutter group (UZH), 2015, *Development of GW schemes*.
- Passerone group (Empa) and Marzari group (EPFL), 2015, *Koopman schemes*.
- Pasquarello group (EPFL) and Cancellieri group (Empa), 2014, *Degradation of Al*₂O₃ *in photocatalytic cells*.
- Curioni group (IBM) and von Lilienfeld group (UniBas), 2015, Machine learning methodologies for screening fast ion conducting materials.
- Curioni group (IBM) with HP5, 2015, *Application of the cognitive platform to solid-state electrolytes*.

- Smolentsev group (PSI) and Corminboeuf group (EPFL), 2015, *Development of novel molecular hydrogen evolution catalysts*.
- Xile Hu (EPFL) and Corminboeuf group (EPFL), 2014, *Development of novel cheap and abundant oxygen evolution catalysts.*
- Peter Staar, Matthieu Mottet, Ivano Tavernelli (HP5), and Marzari group, 2015, *Machine learning approaches to ionic conductivity*.
- Troyer group (ETHZ), Yazyev group (EPFL), and Marzari group (EPFL), 2016, *High-throughput screening of topological materials*.
- Goedecker group (UniBas, HP4) and Marzari group (EPFL), 2015, *Pseudopotential verification*.
- Schmidt group (PSI) and Marzari group (EPFL), 2015, *Investigation of materials for the oxygen evolution reaction*.

- Lippert group (PSI) and Marzari group (EPFL), 2015, Novel materials for water splitting.
- Röthlisberger group (EPFL) and von Lilienfeld group (UniBas, HP4), 2014, *Genetic algorithm optimization of machine learning training sets*.
- Röthlisberger group (EPFL) and Goedecker group (UniBas, HP4), 2014, *Minima hopping of CsPbI*₃.
- Röthlisberger group (EPFL) and Graetzel group (EPFL), 2014, *Experimental validation of theoretical predictions for organic/inorganic halide perovskites*.
- Röthlisberger group (EPFL) and Parrinello group (ETHZ/USI, HP4), 2015, *Simulations of crystal growth of halide perovskites via variationally enhanced sampling*.

1 Novel materials for solar cells (Ursula Röthlisberger — EPFL))

1.1 Research summary

The objective of this project is to design novel materials for solar cell applications following both an evolutionary (significantly improving existing materials) and revolutionary (design of novel materials) approach. During this second year of MARVEL, we focused on the improvement of the properties of several classes of solar cells: (halide) perovskites solar cells, dye sensitizers solar cells (DSSCs), and organic photovoltaics (OPVs). Concerning halide perovskite solar cells, we tackled two problems that strongly limit the use of these materials in prototypical and commercial systems: (i) the hysteresis of the J - V characteristics, and (ii) the instability of formamidinium leadiodide perovskites (perhaps the most promising current perovskite material) at operative conditions. Concerning DSSCs we continued our search for dye sensitizers with enhanced light-harvesting properties. Furthermore, we carried out large-scale classical molecular dynamics simulations to study dye packing effects and their influence on device performance in more details. For the OPV class of solar cells, we studied the nature of the charge transfer state in binary and ternary blends of C₆₀ and pure and fluorinated zinc phtalocyanines and the effect on the overal quantum yield in OPVs. As next steps, we intend to investigate more fundamental optical and electronic properties of perovskites, so as to be able to design materials with tailored properties. Moreover, we intend to exploit the knowledge acquired during this year to design perovskite materials with superior stability and photovoltaic efficiency using genetic algorithms (Sec. 5 of HP4 project) and machine learning models (Sec. 2 of HP5 project).

1.2 Scientific goals

a) Short-term

- Hybrid organic/inorganic halide perovskites Despite of the fact that perovskites of many chemical elements have been extensively investigated, how their properties can be tuned to ideally match the requirements for solar cell applications is still an open question. Thus, our short-term goal is to understand how their chemical, mechanical and optoelectronic properties depend on the parameters that can be controlled in an experimental setting, like chemical composition, size of crystallites, crystal phase, and temperature.
- Dye sensitized solar cells (DSSCs) DSSCs are among the most promising photovoltaic technologies due to their low fabrication cost and thin-flexible structure. The main component of DSSCs are dye

molecules adsorbed on a TiO₂ substrate. The overall efficiency of a DSSC is highly dependent on the light harvesting properties of the dye molecule. Recently the highest efficiency DSSC (13%) was achieved employing a porphyrin-based dye (SM315) [6]. Porphyrin-based dyes have a typical two-peaked absorption spectrum (with the so-called Sored band in the blue and the so-called Q-band in the red region). While in the pioneer SM315 dye, the Q-band was red-shifted significantly, resulting in high photocurrents, the sunlight in the region between the Soret and the Q-band was mostly lost. In this part of the project, we tried to recapture the light in between the Soret and the Qband of the spectrum (between 500-600 nm). Incorporating for the first time the ullazine heterocycle as a donor moiety that is covalently attached to a porphyrin within a molecular dyad, denoted SM63, engineered to exhibit a panchromatic effect [2].

In another work, classical molecular dynamics simulations of ruthenium-based Z907 dyes adsorbed on TiO₂ substrate were performed to study the effects of dye-density coverage on dye arrangement and packing. The results of these simulations were then used for the interpretation of atomic force microscopy (AFM) images of DSSCs taken for the first time under working conditions [3].

- Organic photovoltaics (OPVs) Efficient organic solar cells are based on (electron) donor-acceptor heterojunctions. An optically generated excited molecular state (exciton) is dissociated at this junction, forming a charge-transfer (CT) state in an intermediate step before the electron and hole are completely separated. In this project, we characterized the nature of the charge transfer state in binary and ternary blends of C₆₀ and pure and fluorinated zinc phtalocyanines [4].
- b) Long-term
 - Hybrid organic/inorganic halide perovskites In a longer-term perspective, we will build on the knowledge gained in pursuing the short-term goals to develop improved materials. In particular, we have two major objectives: (i) to design perovskite materials with increased chemical and thermal stability and tunable optoelectronic properties for producing highly efficient and non-degrading so-

lar cells, (ii) to develop lead-free or lowlead-content solar cells, for producing environmentally friendly perovskite based solar cells.

The interest for halide perovskites reaches beyond solar cell applications, with possible additional uses for lasing or photodetectors. Thus, our research might also have possible side effects on these other applicative domains.

1.3 Results obtained since the last report

Hybrid organic/inorganic halide perovskites In this second year, we focused on the identification of the microscopic mechanism at the basis of two phenomena limiting the usability of lead halide perovskites: (i) the hysteresis of the J - V characteristics of perovskite solar cells, and (ii) the limited thermodynamic stability of formamidinium lead-iodide perovskites [1]. Concerning the hysteresis, the problem is that the J - V curve, the (experimental) curve relating the photocurrent and the external applied voltage, is different when the voltage is scanned *forward*, from V = 0 to V_{sc} (V_{sc} is the short-circuit voltage, i.e. the voltage at which the current is zero), than when it is scanned backward (Fig. 2, top). Hysteresis is an important phenomenon for several reasons. First, it makes cells look good which, instead, are not so good. In fact, a quick voltage backward scanning, so that the system does not reach stationary conditions, results in a sizable higher photocurrent. This question has been discussed in detail in [24]. Another related issue is that preventing the phenomenon underlying hysteresis would allow to increase the efficiency of halide solar cells. Finally, it has been hypothesized that the chemical and mechanical instability observed at the interface between the perovskite layer and the contacts [25] is due to the same phenomenon responsible for hysteresis.

By combining first principles molecular dynamics and rare event techniques with temperature dependent experiments, we were able to prove that the hysteresis is due to the ionic polarization of the perovskite layer under the action of the external and internal electric bias [5]. In practice, the perovskite layer acts as a solid electrolyte, and a significant polarization is generated through the transport and accumulation of positive halide vacancies at the interface with the hole conductor (Fig. 2, middle and bottom). The phenomenon is apparently less pronounced in bromide perovskites because of the lower migration barrier of V_{Br}° ,





Figure 2: (Top) J - V curve along the backward and forward scanning path. The difference between the two branches of the curve represents the current-voltage hysteresis discussed in the main text. Vacancy driven halide migration on the equatorial (middle) and axial (bottom) plane of the tetragonal methylamonium lead halide perovskite.

approximately half that of V_I° . Accumulation of halide vacancies at a contact brings the local stoichiometry closer to that of PbI₂, favoring the formation of this latter crystal that, due to the incommensurability of the lattices, results in its detachment from the light harvesting layer.

Our findings suggest that hysteresis can be reduced by replacing halides with other less mobile anions. Obviously, the perovskite of these anionic species must show a photovoltaic efficiency similar to that of halide perovskite.

The second question investigated in this year

was how to improve the **stability** of formamidinium lead iodide perovskite (FAPI). FAPI has a small bandgap ($\sim 1.5 \text{ eV}$), which makes it a very promising material for the light absorbing layer. However, at room temperature its stable phase is the yellow δ phase, which being almost transparent is unsuitable for solar cells applications. Thus, it is of crucial importance to find a way to stabilize the lower band gap perovskite phase at room temperature.

A possible idea on how to improve stability is mixing different kinds of perovskites, e.g. perovskites with different monovalent cations. The idea is that a modest mixing does not significantly affect the properties of FAPI but improves its stability. Concerning the stability, the key observation is that all perovskites have a similar structure, with the quasi-cubic inorganic framework hosting the monovalent cations in its cavities. On the contrary, depending on the type of cation, the δ phase of the compound may be structurally significantly different. This is the case, for example, of the δ phase of FAPI and CsPI (CsPI₃, Fig. 3). δ FAPI and CsPI have also a very different volume per stoichiometric unit, the second being smaller by $\sim 10\%$, while the difference of volume in the perovskite phase is much smaller, in fact less than 1%. This suggests that mixing materials with different cations is energetically more convenient in the perovskite phase than in the corresponding δ phase. Following this idea, we computed the free energy balance of mixing cations in the δ and persovskite phases of $Cs_x FA_{1-x} PI$. It was found that in the δ phase the energetic cost of mixing is not balanced by the mixing entropy. On the contrary, the perovskite system is energetically neutral with respect to the mixing, and the mixing entropy leads to a free energy gain of the order of $\sim k_B T$. In practice, cation mixing shifts the



Figure 3: Atomic structures of the δ (nonperovskite), α and β FAPI (top row), and δ (nonperovskite) and cubic CsPI.



Figure 4: The probability of finding any dye atom along the *z*-axis (normal to indicate that the amphiphilic *Z*907 dye covers most of the TiO_2 surface) for 20 ns of MD simulation in NVE ensemble. Blue clouds show the most probable spaces populated by alkyl chains.

transition temperature from the δ to the perovskite phase by ~ 300 K.

This idea has been proven to work. In collaboration with the group of Michael Graetzel we have prepared mixed Cs/FAPI cation solar cells and shown that they are stable under operative conditions, and have a very promising efficiency of $\sim 18\%$ even without any special optimization.

Dye sensitized solar cells The panchromatic SM63 dye was designed via a density functional theory (DFT) study to gain insight into the changes of the electronic structure that can be affected through introduction of the ullazine heterocycle. An impressive improvement of panchromatic light harvesting ability, up to 800 nm, was achieved in photoelectrochemical devices employing SM63 [2].

The question of how the dye sensitizers arrange on the TiO₂ surface and how alkyl substituents can help in preventing direct contact between the electron conducting substrate and the electrolyte was investigated in a combined experimental/computational study. The AFM image of Z907-based DSSC was simulated using classical molecular dynamics (MD) (Fig. 4). Home-made force fields were derived from quantum calculations and were used for setting up a DSSC composed of TiO₂, solvent and dye molecules. The simulation results were used in turn for the interpretation of the AFM images [3].

Organic photovoltaics (OPV) We show that the observed high quantum yields of charge carrier generation and CT state dissociation are due to



Figure 5: Optimized structure and electron density difference of ternary blends. Electrons (yellow) and hole (blue) are shown for the first excitation ($\rho 1 - \rho 0$).

delocalized CT states visible in the absorption spectra. (Fig. 5) [4].

1.4 Planned research for next year

In the next year, we plan to explore some fundamental characteristics of the excited state properties of hybrid organic-inorganic perovskites. So far, perovskite materials for solar cells have been developed on the basis of empirical rules, without understanding in detail how their properties depend on the structure and composition of the system. For example, recent results have shown that the photoluminescent spectrum of halide perovskite significantly changes as a function of temperature and the nature of the monovalent cation, but there is no theory describing this behavior, and no model predicting the optical properties of this family of materials. For the next year we have planned a joint experimental and theoretical activity in collaboration with the group of Michael Graetzel aimed at addressing these questions. Novel materials and solar cells will be fabricated exploiting the results of this research.

Dye sensitized solar cells Concerning DSSCs, we plan to use our data bank on different sensitizer molecules and collect them in a comprehensive review on the light-harvesting properties of these molecules. In addition, using our home-made force fields for different ruthenium-based and porphyrin-based sensitizers, we plan to investigate larger scale phenomena in different DSSC devices (e.g. dye arrangement, electrolyte accessibility)



a) *Within MARVEL* Concerning our investigations of halide perovskites we work in synergy with several MARVEL groups, e.g. the groups of Stefan Goedecker, Anatole von Lilienfeld and Michele Parrinello. This year we mostly worked in the development of methods and the corresponding software tools. The progress of this collaborative work is described in other sections (Secs. 5 and 6 of HP4 project). For the next year we plan to use these techniques to search for novel materials for solar cells applications.

b) *Outside MARVEL* Already during the present and previous reporting periods we worked in synergy with the experimental group of Michael Graetzel (EPFL). During the

2 Photocatalytic water splitting (Alfredo Pasquarello — EPFL)

2.1 Research summary

- We calculated redox levels in aqueous solutions finding good agreement with experimental data.
- We considered vertex corrections in selfconsistent GW calculations through the use of the bootstrap kernel resulting in very accurate band gap predictions for a large set of materials.
- We studied intrinsic and extrinsic defects in amorphous Al₂O₃ in view of understanding the degradation of this material when used in electrochemical cells.

2.2 Scientific goals

a) *Short-term* We are interested in aligning the band edges of a semiconductor and the water redox levels with respect to a common scale through the explicit modeling of the solid-water interface. We will first establish benchmark results using advanced techniques such as hybrid functionals and GW.

b) *Long-term* In a further stage, the detailed charge transfer processes occurring at the solid-water interface will be addressed with the goal of establishing additional criteria for the efficiency of the photochemical splitting of water.

next year we intend to further intensify this collaboration with the objective of developing complete devices designed on the basis of principles derived from our simulations.

1.6 Personnel

NCCR personnel

• Simone Meloni, senior researcher, 100%, from May 2014.

Matching personnel

- Negar Ashari Astani, PhD student, 100%, from May 2014.
- Ariadni Boziki, PhD student, 100%, from December 2014.

2.3 Results obtained since the last report

a) Redox levels in aqueous solution: effect of van der Waals interactions and hybrid functionals The accurate prediction of the electronic properties of liquid water and of redox levels in aqueous solution is fundamental in the study of photocatalytic water splitting. In order to fulfill this goal, we studied electronic properties of liquid water, such as band gap, band edges, and redox levels, using DFT-MD simulations and thermodynamic integration methods [7]. In particular, we considered two different DFT functionals, namely the PBE and rVV10, with the scope of evaluating the effect of explicitly considering van der Waals (vdW) interactions on the electronic properties. In our investigation, we started with the band gap and the band edges, considering different computational schemes for their determination. We first considered schemes proposed in the literature in which the band edges are evaluated either from HOMO/LUMO Kohn-Sham energy levels or from total-energy differences. However, both these schemes correspond to selecting states in the tail of the density of states. Through simulations with different system sizes, we demonstrated the occurrence of finite-size effects in both schemes. Hence, we considered an additional scheme in which the band edges are linearly extrapolated from the density of states. This scheme shows two advantages. First, the identified band edges depend on features of the density of states which are not affected by the disor-

der in the liquid and thus do not depend on the system size. Second, the proposed scheme corresponds to the procedure used to extract experimental estimates for the band-edge positions. Our study shows that the structural modifications of the liquid induced by vdW interactions do not change the band gap and the band-edge positions within 0.1 eV.

For the redox levels in aqueous solution, we adopted a theoretical formulation, based on the grand-canonical approach [26], which allowed us to establish an analogy with charge transition levels of defects in crystalline materials [27]. In the treatment of the redox halfreaction, free-energy differences are considered and an electron chemical potential is introduced. The free-energy differences are then evaluated by thermodynamic integration and the redox levels result from the condition of equal free energies in the reduced and oxidized systems. This formulation offers the advantage of benefitting from the know-how developed for defects in solids, most noticeably the treatment of electrostatic finite-size effects arising from the simulation of charged species subject to periodic boundary conditions. Our study involves the calculation of the redox levels associated to the standard hydrogen electrode, the oxidation of the hydroxyl ion, and the dehydrogenation of water. We showed that the solvation shells of the ionic species are only moderately affected by vdW interactions, leading to small variations in the thermodynamic integrals. Since the electronic properties are found to weakly depend on the structure of the liquid, we adopted the structural configurations encountered in our DFT-MD simulations to calculate the electronic properties at the hybrid functional level. This scheme allows us to study the electronic properties as a function of the fraction α of Fock exchange included in the hybrid functional in a computationally affordable way. The band gap and the band edges evolve linearly with α . At variance, the redox levels are found to remain essentially constant with α , in analogy to the behavior of charge transition levels of atomically localized defect states. When referred to the potential of the standard hydrogen electrode, the calculated redox levels associated to the oxidation of the hydroxyl ion and the dehydrogenation of water agree with experimental values within 0.2 eV for a wide range of α . For a value of $\alpha \cong 0.4$ we find a good agreement with recent experimental data for all the calculated electronic properties of liquid water, including the band gap, the band edges, and the redox levels (Fig. 6).



Figure 6: Energy levels vs the fraction *α* of Fock exchange introduced in the PBE0 functional, referred to the SHE reference level.

b) Band gaps through self-consistent GW We have proposed a highly accurate GW scheme to describe the band gap of extended systems via the inclusion of vertex corrections in the quasiparticle self-consistent GW method [8]. The vertex corrections in the screened interaction are accounted for through the use of an exchange-correlation kernel based on the bootstrap approximation [28]. We show that this method yields accurate band gaps for a variety of materials, including *sp* semiconductors, insulators, and transition metal compounds with either open or closed *d* shells (Fig. 7). The evaluation of the vertex corrections adds very little overhead to the standard random phase



Figure 7: Theoretical vs experimental band gaps for a variety of materials. G_0W_0 corresponds to oneshot GW on top of a PBE starting point. QSGW is the quasiparticle self-consistent GW in the random phase approximation. Our new scheme is denoted as QSG \tilde{W} where the screened interaction \tilde{W} is vertex corrected via the bootstrap kernel.



approximation, whereas the accuracy of our scheme is comparable to that of the much more demanding Bethe-Salpeter kernel. This makes our scheme well fitted for the band-gap predictions of large-scale systems.

The use of hybrid functionals and GW for the determination of defect levels and band edges has been reviewed, with particular focus on technical aspects emerging in GW calculations, such as finite size corrections, path dependence, and delocalization error [9].

c) First-principles calculations of extrinsic and intrinsic defects in am-Al₂O₃ Amorphous (am) Al₂O₃ is an important ceramic and dielectric material, which has widely been used in many technological applications, such as micro-electronics, lasers, and nuclear reactors. More recently, it has drawn considerable attention for its use in photocatalytic water splitting devices. In particular, its deposition on α -Fe₂O₃ has proved to lower the overpotential associated to the water oxidation reaction [29]. However, in electrochemical cells, the structure of the semiconductor/oxide degrades with use. As the negative charge is collected across the semiconductor/oxide, hydroxylation occurs which modifies the electronic properties of the material. In order to address the modifications induced by hydroxvlation, we built a model of am-Al₂O₃ and inserted an hydroxyl ion inside its bulk structure. Hence, we studied the structural and electronic properties of this system through a combination of ab initio molecular dynamics simulations and electronic structure calculations. The former were performed at the PBE level of theory, while the latter were performed at both the PBE and HSE levels of theory, in order to examine effects associated to the bandgap opening. We found that the introduction of the hydroxyl ion in the am-Al₂O₃ bulk structure resulted in significant rearrangement of the bonding. As for the electronic structure, we found that the band gaps of various phases of Al₂O₃ are well reproduced compared to experiment, provided a value of $\alpha = 0.45$ is used for the fraction of Fock exchange in the HSE functional. We then studied the electronic properties of the hydroxyl ion in am-Al₂O₃. We observed a charge transition level positioned at \sim 2.2 eV above the valence band maximum of am-Al₂O₃, a signature of a deep-level trap inside the band gap. Other possible impurities or defects have also been studied, including the interstitial hydrogen and oxygen atoms. Moreover, we also analyzed the oxygen vacancy defect in am-Al₂O₃, a common intrinsic defect in this material. Charge transition levels have been calculated for all these defects.

2.4 Planned research for next year

The study of the redox levels in aqueous solution is the first step towards the accurate description of redox reactions at watersemiconductor interfaces. We now plan to build water-semiconductor interfaces and to determine the corresponding level alignment in order to establish the accuracy of our calculations with respect to benchmark experimental results. In particular, we also intend to focus on the various intermediate steps towards water splitting occurring in proximity of the interface.

In addition, we are interested in bringing the accuracy of GW calculations to the study of the electronic structure and the associated redox levels of liquid water. So far, two important aspects are less explored, namely the quantum nuclear effect and the many-body effect. We will address the quantum nuclear effect through path-integral molecular dynamics [30]. The many-body effect will be taken into account by our recent vertex-corrected self-consistent GW scheme [8]. Once a reliable electronic structure of liquid water is obtained, the electronic excitation will also be probed by the calculation of optical spectra using either time-dependent density functional theory or the Bethe-Salpeter equation. In addition, given the close analogy between redox levels in liquid water and defect levels in semiconductors, we will assess the accuracy by which the energy of localized electronic states can be determined through a benchmark study involving a series of well characterized point defects in semiconductors.

We will continue our study on the role of am-Al₂O₃ overlayers in electrochemical cells by focusing directly on the α -Fe₂O₃/am-Al₂O₃/H₂O system. In particular, we would like to focus on the reasons which lead to the reduction of the overpotentials in this type of electrochemical cell.

2.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* We are collaborating with Jürg Hutter as far as the use of the code CP2K is concerned and plan to carry out benchmark comparisons between CP2K and Quantum-ESPRESSO for systems like liquid water. We are interacting with an experimental group at Empa (Lars Jeurgens, Claudia Cancellieri, and Patrick Schmutz) addressing the stability of oxides in photochemical cells. We have established preliminary contacts with Michele Ceriotti concerning the use of path-integral molecular dynamics simulations.

b) *Outside MARVEL* We are collaborating with C. Massobrio and M. Boero (CNRS-Strasbourg) in view of studying how our modeling approach for liquid water applies to the description of aqua ions.

2.6 Personnel

NCCR personnel

• Francesco Ambrosio, postdoc, 100%, from January 2016.

• Zhendong Guo, PhD student, 100%, from January 2015.

Matching personnel

- Francesco Ambrosio, postdoc, 100%, from May to December 2015.
- Giacomo Miceli, postdoc, 100%, from May 2014.
- Wei Chen, postdoc, 100%, from May 2014.

3 Photo-catalyst design for water oxidation and reduction (Jürg Hutter - UZH)

3.1 Research summary

We study Co(II)-based cubane molecules in solution as possible water oxidation catalysts. Different theoretical approaches, model systems, and simulation protocols are explored and results compared to experiments. Pyridine-based molecules and their interaction with Au(111) and a rutile(110) surface are studied.

3.2 Scientific goals

We are investigating the basic mechanisms in water oxidation and reduction catalysis. Insight gained from detailed atomistic studies should allow to device design strategies for optimized molecular structures for catalysis. The structure and dynamics of a family of Co(II)based cubane molecules as water oxidation catalysts (WOC) and pyridine-based molecules on TiO_2 and Au(111) as water reduction photocatalysts are investigated. The calculations will help, together with experiments, to deepen the understanding of the mechanisms involved in the reactions and to identify the crucial structural and electronic motifs that determine the catalytic efficiency.

a) *Short-term* Investigate Co(II)-based cubane molecules in solution, understand the geometric and electronic structure and make connection to experimental observation for these systems. Further, establish simulation protocols for the investigation of water reduction catalysts based on ligand molecules with a cobalt atom center adsorbed on a rutile surface.

b) *Long-term* Understand the details and energetics of complete reaction mechanisms, de-

termine the most important factors influencing water oxidation and reduction catalysts activity and, with this knowledge at hand, finally design efficient and robust catalysts.

3.3 Results obtained since the last report

a) *Water oxidation catalyst* Artificial photosynthesis is a promising approach to use solar light to split water into oxygen and hydrogen, which is a powerful and environmentally friendly carrier for energy storage and conversion. We recently brought forward the first homogeneous Co(II)-based cubane water oxidation catalysts, $[Co_4^{II}(hmp)_4(\mu-OAc)_2(\mu_2-OAc)_2(H_2O)_2]$ (hmp = 2-(hydroxymethyl)pyridine) (1) and $[Co_3^{II}Er(hmp)_4(OAc)_5H_2O]$ (2) (collaboration with group of G. R. Patzke, [31][10]).

Elucidation of the detailed structure and water oxidation mechanism of theses WOCs is an indispensable step towards the development of structure-activity relationships and artificial high-performance WOCs. For the study of their water oxidation behavior, we thus focused first on the determination of the catalytic ground state structure. We investigated various ligand exchange reactions of (1) and (2) applying both static and dynamic computational methods. Moreover, we looked at the dependency of the ligand exchange energies on the applied solvation model. To this end, we explored not only different levels of solvation ranging from a periodic box of hundreds of water molecules (explored with Born-Oppenheimer molecular dynamics), first solvation shell, implicit solvent to in vacuo calculations, but also investigated barriers and free energies using the nudged elastic band



Figure 8: Three NEB-frames of the O-O bond formation and oxygen release at the end of the oxo-oxo coupling pathway of (1). The electronic energies relative to the first image are marked with a red dot in the graphs above.

method (NEB) and metadynamics. We found a very strong variation of energies with the type and degree of solvation employed. Even the presence or absence of one water molecule of the first solvation shell had a large impact. From this we concluded that, for such ligand exchange reactions, the use of density functional theory (DFT)-based molecular dynamics or other sampling techniques is advisable. Moreover, we found that cubane (2) containing the redox-inert lanthanide cation has a much more flexible ligand shell than (1), which may also be one reason for its higher water oxidation activity [10].

As a next step, possible water oxidation mechanisms of (1) have been investigated. First, a single-site pathway involving two proton coupled electron transfer (PCET) steps followed by a nucleophilic attack of a water molecule forming the O-O bond; next, another 2 PCET steps take place and finally, O₂ is released. Second, another pathway using two adjacent Co centers, converting both Co(II)-OH₂ to Co(IV)-O and then coupling the two oxo ligands, finally evolving O₂. We compared both mechanisms computing not only free energy differences between all catalytic states, but also calculating barriers and reaction paths using NEB (Fig. 8). We analyzed and characterized all states in terms of properties such as frontier orbitals involved, spin multiplicity, distribution of spin density, and structure. Moreover, we compared the water oxidation behavior of (1) to the one of an *ideal* catalyst as often carried out in heterogeneous catalysis. In doing so, we were able to deduce possibilities for tuning the water oxidation catalyst.

b) Water reduction catalyst Co-Pyrphyrin (CoPy) has been recently synthesized and successfully used as homogeneous water reduction catalyst. We have investigated the adsorption of CoPy on the rutile(110) surface using density functional theory [11]. We observe electronic structure rearrangements occurring upon adsorption, mainly due to the interaction of the cyano groups of CoPy with the five-coordinated Ti atoms of rutile(110). The projected densities of states on the different involved species reveal that the molecular orbitals on the $N_{\mbox{CN}}$ atoms undergo a re-hybridization, the highest occupation molecular orbital (HOMO) is located at the molecule, and the lowest unoccupied states of CoPy are shifted to higher energies. The first band of unoccupied states is mainly on rutile and extends from 1.1 to 2.0 eV, with a maximum at 1.5 eV. The lowest unoccupied states localized on CoPy are instead above an energy of 2.5 eV, i.e. at higher energy with respect to the LUMO of the gas phase molecule. We do not observe strong modifications of the electronic structure at the Co(II) center. From these results we expect that the catalytic properties of CoPy for water reduction are preserved upon adsorption.

Apart from CoPy/oxide interactions, CoPy interactions with a metal surface are also investigated by adsorbing CoPy on a reconstructed Au(111) surface. Calculations show that the interaction is maximized when both the cyano N (N_{CN}) atoms and Co are close to surface Au



Figure 9: Optimized structure of CoPy on reconstructed Au(111) surface. One side view (a) and the top view (b) of the model are shown. For clarity only two Au layers and only the surface layer are shown in (a) and (b), respectively. Color code: yellow: Au, blue: N, dark blue: Co, brown: C, and pink: H.

atoms. This condition is satisfied when the molecular cyano axis is 60 degrees tilted with respect to the (110) Au lattice orientation, and N_{CN} and Co are at top sites, as depicted in Fig. 9. Adsorption of Pyrphyrin (Py), without a metal center, results in the same orientation, being slightly bent to favor the N_{CN}-Au interaction. However, the adsorption strength is more than half eV weaker, indicating that the presence of Co has a clear effect.

3.4 Planned research for next year

a) Water oxidation catalyst We will explore the cause for the difference in turnover frequency and catalytic activity in general between catalysts (1), (2), and a $[Co_3^{II}TmO_4]$ based cubane (3), an analog of (2), using an extended approach to the one used up to now. We will further calculate the redox potential for a simpler mononuclear Co(II)-based water oxidation catalyst using a periodic box of water and DFT-based molecular dynamics and compare the results to the ones obtained from a static approach.

In collaboration with the group of Clémence Corminboeuf, we will investigate different ligands and central transition metal atoms based on a cobalt chloride catalyst studied recently [32].

b) *Water reduction catalyst* We will further investigate the effects of vacant sites on preferential adsorption configurations of CoPy on the rutile(110) surface. Creating defect sites on the surface affects the electronic structure of the complex and also the adsorption strength of the molecule. As our experimental partners do not observe molecular monolayer formation on the rutile(110) surface, we plan to in-

vestigate the effects of defects sites on monolayer formation. In addition, we will investigate the metalation of Pyrphyrin on the reconstructed Au(111) surface. The related metalation mechanism of porphyrin-derived molecules on metal surfaces is not well understood, as is pointed out in several experimental and theoretical studies in the literature. We will then focus on the mechanism of water reduction, the occurrence of possible intermediate products, and the energy barriers for hydrogen production.

3.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* There are currently two established collaboration within MARVEL for this project. A collaboration with Alfredo Pasquarello is on the reliability of simulation protocols for molecular liquids. We started a new collaboration with Clémence Corminboeuf on water oxidation catalysts. We have a joint project with Grigory Smolentsev (PSI) from PP7 on the experimental investigation of the Co-cubanes using time-resolved X-ray absorption spectroscopy. We are also setting up a series of test systems based on the Cobalt catalysts. These systems should be used by several groups within and outside MARVEL to test their computational methods.

b) Outside MARVEL Our efforts in this project are coordinated with the University of Zurich Research Priority Program (URPP) LightChEC (http://www.lightchec.uzh.ch/). The final goal of this program is to create an architecture of catalysts and antennas to mimic the natural photosynthesis model. The individual steps: light harvesting, oxidation and reduction of water, are combined into one single functional system.

3.6 Personnel

Matching personnel

- Sandra Luber, postdoc, 100%, from May 2014.
- Yeliz Gürdal, PhD student, 100%, from May 2014.
- Florian Hodel, PhD student, 100%, from May 2014.



4.1 Research summary

We have developed tools and design guidelines for molecular materials. We plan to further develop these tools and apply them to predict new molecular water splitting catalysts and new materials for molecular electronics.

4.2 Scientific goals

a) Short-term

- Molecular volcano plots will be used to design novel water splitting catalysts.
- The applicability of the molecular volcano plots [12] will be expanded to include kinetic effects.
- The potential of metal free catalysts and exfoliated transition metal oxides to catalyze water oxidation will be explored.
- The requirements for the non-Aufbau orbital configuration will be identified and applied to design new molecules for organic electronics and spintronics applications.

b) *Long-term* The long-term goal of our research is the development of innovative guidelines and tools to rank and predict molecular materials for catalysis and molecular electronics applications. Based on this conceptual understanding, novel materials will be predicted.

4.3 Results obtained since the last report

a) *Molecular volcano plots* Volcano plots and linear scaling relationships are important concepts in heterogeneous catalysis. Based on the shape of the volcano plots it is possible to visualize and rank large datasets obtained through computations or experiments. This analysis allows for the efficient prediction of new catalysts, yet despite their usefulness volcano plots have so far been restricted to heterogeneous catalysis.

Recently, we generalized and applied volcano plots for the first time (Fig. 10) to a homogeneous catalytic reaction, the Suzuki-Miyaura C-C cross coupling of olefins. This choice was guided by its importance in organic chemistry and the significant amount of experimental and computational information available, which aided validation of the obtained results. This proof-of-principle study will serve as a basis for the prediction of novel molecular water splitting catalysts through molecular volcano plots.



Volcano Plots for Homogeneous Catalysis

Figure 10: A schematic representation of a volcano plot is depicted.

b) Water oxidation at exfoliated oxides Recent experimental studies suggest that catalytic activity is significantly increased upon exfoliation of layered transition metal oxides [33]. We investigated their activity towards water oxidation for a CoO_2 model system. Based on this study we have excluded electronic and mechanistic effects. Rather, an increase in conductivity must be at the origin of the observed improvements [33].

c) Molecular electronics We assembled an exhaustive library of quasi-closed-shell molecules, reported in the literature to date [13]. An in-depth analysis of their electronic structure suggests that an orbital energy level conversion requires a combination of a resonance-stabilized radical and a fragment providing high-energy doubly-occupied orbital(s) (HOMO(s), e.g. an unsaturated heterocycle or an anion, Fig. 11). We also established a computational methodology based on DFT to probe such orbital configuration by comparing the electronic energies of the products of one-electron removal from the molecule in question. Our computations successfully predict orbital energy-level conversion in molecules, in which it has already been demonstrated experimentally.



Figure 11: Example of a quasi-closed-shell molecules, a method to detect it, and its materials' applications.

4.4 Planned research for next year

a) *Molecular volcano plots* Until now molecular volcano plots are only capable of capturing system thermodynamic. We aim to expand their applicability by explicitly considering scaling relations between activation barriers and descriptor intermediates. Based on these scaling relationships, volcano plots that describe system kinetics will be constructed and used to optimize catalysts with respect to their kinetic limitations.

b) *Water splitting* We plan to apply our novel molecular volcano plots to design improved molecular water splitting catalysts. In collaboration with the experimental efforts of Grigory Smolentsev (PSI), molecular materials for H_2 evolution will be studied. Together with Sandra Luber and Jürg Hutter (UZH) we also aim to predict new molecular materials for O_2 evolution. Additionally, we will continue our research efforts at layered water oxidation catalysts by considering exfoliated transition metal oxides and transition metal free graphene in collaboration with Xile Hu (EPFL).

c) *Molecular electronics* We will use the established methodology to develop electronic structure trends, observed in a limited number of known non-Aufbau molecules, into general design principles. With those in hand, we will propose a range of new molecules, for which we could confirm the quasi-closed-shell orbital configuration *in silico*, on the basis on the relative energies on their one-electron removal products. We will specifically search for substrates with promising charge-transport properties (through high-energy HOMO(s)), since they would also be likely to possess appreciable magnetoresistance and hence be of interest in molecular electronic devices. Additionally, we will screen for species, whose orbital configuration can be manipulated by external stimuli (chemical bonding, external electric field), since the ability to switch between open- and quasi-closed states is valuable in organic spintronics.

- 4.5 Synergies with other computational and experimental efforts
- a) Within MARVEL
 - Sandra Luber and Jürg Hutter, UZH, VP2: combined efforts in the computational study of molecular water oxidation catalysts.
 - Grigory Smolentsev, PSI, PP7: combined experimental and computational efforts in understanding and predicting improved molecular water reduction catalysts.
 - Michele Ceriotti, EPFL, HP4: REMD@DFTB [14].
- b) Outside MARVEL
 - Xile Hu, EPFL: combined experimental and computational efforts to develop novel water oxidation catalysts.

4.6 Personnel

NCCR personnel

- Michael Busch, postdoc, 100%, from October 2014.
- Ganna Gryn'ova, postdoc, 50%, from August 2015.

5 A standard solid-state pseudopotential library (SSSP) and its application to novel rare-earth perovskites for visible light harvesting applications (Nicola Marzari — EPFL)

5.1 Research summary

One of the requirements for high-throughput simulations is an accurately verified pseudopotential library that covers the entire periodic table. In this project, we have curated one with the goal of identifying novel rare-earth perovskites with potential for light harvesting and water splitting applications using a densityfunctional theory framework combined with a screening approach based on fundamental descriptors (phase stability, band gaps, and position of the band edges vs the redox levels of water). In addition, collaborations with experimentalists have been established for the synthesis and characterization of previously identified Ta- and Nb-based perovskites.

5.2 Scientific goals

a) *Short-term* One of the preliminary steps of the project is connected with the investigation and identification of optimal pseudopotentials, both in terms of accuracy and efficiency, among the ones available in literature or, in some important case, generated by ourselves. This set will be used in the project. This is a part of a big effort that the scientific community



is making for a broad validation and verification of density-functional theory (DFT) simulations [15]. In addition, due to the well-known problem of estimating the band gap in DFT, we will try different approaches, like G_0W_0 or the GLLB-SC functional, to obtain reliable electronic properties of materials. We have recently started the calculations to relax the rareearth perovskites in the simplest 5-atom unit cell, using a standard functional in the generalized gradient approximation (GGA-PBEsol). The band gap calculations will be performed starting from these optimized structures.

b) Long-term One of the long-term objective is to define a standard procedure, based on quantities like convergence of phonons, stresses, and heats of formation, as well as structural properties like the equations of state, to test new pseudopotentials in an automatic way using the AiiDA platform. The first results of the pseudopotential testing are currently available online at the address http://materialscloud.org/sssp which has been visited more than 3500 times in the last 5 months. Another objective is to refine the results of potentially interesting materials for light harvesting using more expensive and reliable methods. In addition, the catalytic properties of the most promising compounds will be investigated in a realistic electrochemical environment.

5.3 Results obtained since the last report

The project is based on density-functional theory calculations and a systematic verification of the method is the starting point. In the last report, we have described our preliminary results of the tests of various pseudopotential libraries. Those tests were based on the accurate reproducibility of all-electron results following the procedure proposed by K. Lejaeghere *et al.* [34] in which a quality factor, called Δ , gives the measure of the discrepancies between equations of state (calculated by an all-electron code and within the framework of the pseudopotential approximation) for elemental compounds. Smaller the Δ , better the pseudopotential can reproduce the all-electron results. The delta-measure is the first attempt, done in a systematic way, to benchmark various codes and approximations. Recently our efforts, together with the ones of around 70 other scientists, have been collected in a manuscript under review in Science in which dozens of codes and methods have been benchmarked in a rigorous way using the Δ -measure [15]. In addition, all the produced data are available in the supplementary material for further tests.

Although the delta gives an indication of the quality of a pseudopotential for structural properties, it gives no information about other quantities. We have then expanded the properties to test to include the convergence of the phonons, heats of formation, and pressures, i.e. trace of the stress tensor, as well as the computational cost. We have now systematically tested 10 pseudopotential libraries: pslibrary 1.0.0 US, PAW, US low accuracy, and PAW low accuracy [35], pslibrary 0.3.1 US and PAW [36], the sets from Garrity, Bennett, Rabe, and Vanderbilt, GBRV versions 1.2, 1.4, and 1.5 [37] and from Schlipf and Gygi, SG15 versions 1.0 and 1.1 [38] and from Topsakal and Wentzcovitch for the rare-earth elements [39]. We are currently expanding these sets by including the PseudoDojo (http://www.pseudodojo.org/) and norm-conserving pseudopotential recently produced by Saha and Goedecker (MARVEL, UniBas).

We select optimal pseudopotentials using criteria like small delta with respect to allelectron calculations and, with respect to pseudopotential calculations at very high and converged cutoffs, convergence of the phonons within 2%, heats of formation, δH_f , within 3 meV, pressures, δP , within 0.5 Kbar. Other information, like computational cost and presence of semi-core states, has been taken into consideration. The selected pseudopotentials and the suggested wave functions cutoffs are shown in the top part of Fig. 12. The color of each element indicates the library from which the pseudopotential is taken. For some cases, we propose two different pseudopotentials, one with the focus on efficiency, and thus with a lower required cutoff (top left triangle), and one with the focus on accuracy (lower right triangle). The two proposed library, called Standard Solid State Pseudopotentials (SSSP) efficiency and accuracy, are among the pseudopotential libraries with the smallest delta (https://molmod.ugent.be/deltacodesdft).

We are currently working on improving some pseudopotentials which have a particularly large delta or slow convergence patterns, like oxygen, and copper. The produced data, together with the scripts to reproduce our results and expand the tests, are available online on http://materialscloud.org/sssp. In the bottom part of Fig. 12, we show the convergence pattern for solid oxygen. So far, we have investigated 7 libraries (2 PAW, 3 US, and 2 NC). For each library, the convergence of the phonons, of the heats of formation, and of the

pslibrary.1.0.0 US 🔳 pslibrary.1.0.0 PAW 🔳 pslibrary.0.3.1 US 📕 pslibrary.0.3.1 PAW 🔳 GBRV-1.2 (US) $\Delta_{eff} = 0.42 \text{ meV}$ $\Delta_{acc} = 0.31 \text{ meV}$ ■ GBRV-1.4 (US) ■ GBRV-1.5 (US) ■ SG15 (NC) ■ SG15-1.1 (NC) ■ Goedecker ■ THEOS H 55 He 55 RE Wentzcovitch ⁴⁰ Be 55 35 😋 Mg 35 Al 30 Si 30 P 30 ⁵ Ga 7(Ge 40 Br 30 K 50 As 30 Rh 45 Rb 50 Sr 35 Mo 3 Tc 30 Ru Pd 55 Ag 35 Cd ₄ Sb 4 Ba 40 Hf₁₂ Au as Yb 🚛 Lu 45 Tb 40 DV AC Nd ac Euse Pr 50 m 40 о [Kbar] = 1344.8 cm Z = 60.06 0.0 SG15 Z = 6 A = 0.174 = 1331.0 cm -1.18 0.0 [meV] 50 0.0 GBRV-1.2 Z = 6 0.0 200 0.08 0.06 [meV] -0.5 0.12 0.0 [Kbar] = 1352.3 031US Z = 6 ∆ = 1.161 0.0 0.0 = 1353.3 cr 031PAW Z = 6 ∆ = 0.634 0.0 [meV] -0.01 0.0 100US Z = 6 0.02 0.0 0.03 0.05 -0.0 0.02 [meV] -1.51 -0.66 -0.05 -3.55 -3.38 100PAW Z = 6 Δ = 4,401 0.14 -0.05 -0.01 0.25 0.25 0.1 0.02 0.02 0.0 90 100 110 120 130 140 150 Energy cutoff [Ry]; dual = 8 (PAW/US), dual = 4 (NC); q-point = [0.5, 0.5, 0.5]

Figure 12: Top panel: the standard solid-state pseudopotential library. Each squares indicates the original library from which the optimal pseudopotentials are taken. For some selected elements, two pseudopotentials have been identified with specific focus on efficiency and accuracy. The wave functions cutoffs (in Ry) and the duals, when different from 8, are reported. The delta values for the SSSP efficiency and accuracy are also indicated. Bottom panel: the convergence pattern for bulk oxygen. The libraries investigated are, from the bottom to the top, pslibrary.1.0.0 PAW (in green) and US (in red), pslibrary.0.3.1 PAW (in violet) and US (in blue), GBRV 1.2 (in cyan), SG15 (in black) and from Goedecker (in yellow). Each line represents the convergence of the phonons as a function of the wave functions cutoff with respect to the value at fully converged cutoff (200 Ry). The error in the pressure δP and the convergence of the heat of formation δH_f as a function of the cutoff are shown. The values of the Δ -measure, the number of valence electrons Z, and the highest vibrational frequency ω_{max} are also indicated. The selected pseudopotentials and suggested cutoffs for the SSSP efficiency and accuracy are indicated with silver and gold circles, respectively.

pressures with respect to the wave functions cutoff is shown. Using the criteria mentioned above, we have identified the pseudopotential from the GBRV-1.2 library with a cutoff of 45 Ry as optimal for the SSSP efficiency. Due to the relative large delta of around 2 meV, we have selected the pseudopotential from the pslibrary.0.3.1 (with cutoff of 70 Ry) for the SSSP accuracy. All the convergence patterns for the 80 investigated elements are available online. This part of the project is powered by AiiDA: workflows have been implemented for fast calculation and analysis of each pseudopotential. All the workflows are available for the NCCR MARVEL and for the broader scientific community - more than 3'500 people have visited the website in the last 5 months.

MARVEL

The project is now continuing with the structure optimization of the lanthanide perovskites in 5-atom unit cell. We apply some simple structural and chemical based rules, i.e. size of the cations, total number of electrons and the global valence of the compound equal to zero [40], to reduce the number of calculations to run to around 20%. It has been recently shown that, although oxo-perovskites have a high stability, they have poor absorption properties mainly due to the indirect transition at the band gap and weak transition-matrix elements [41]. Oxynitrides, despite being less stable, especially in a water environment, show higher absorption with a better placed band edges with respect to the redox levels of water. Next to the oxo-perovskites (with gen-

Research

eral formula ABO₃), we are also investigating oxynitride materials (*ABO*₂N and *ABON*₂) for a total of around 400 possible compounds. At today, we have the optimized structures of around half of these compounds. The stability of the optimized compounds is evaluated with respect to around 2'000 solid and dissolved phases (for the corrosion in water) in which the material can separate [42]. We are currently working, together with experts in Prof. Nørskov group (Stanford, USA), to implement and test the GLLB-SC functional in the Quantum-ESPRESSO code, which will be used to evaluate the band gaps. In fact, it has been shown [43] that the GLLB-SC functional gives a very good approximation of more expensive GW methods with an error of less than 15%. Together with Prof. Jacobsen group (DTU, DK), we are working on refining the descriptors with a focus on both stability and light absorption properties.

5.4 Planned research for next year

This project will follow two main directions: on the pseudopotential side, the verification of pseudopotentials and the identification of the optimal ones can be improved. The scientific community is discussing about new protocols and new properties to test and we are part of this. On the screening side, our goal is to identify few rare-earth materials for light harvesting and to have a better understanding of both the oxygen evolution reaction and the optical properties of tantalate and niobate perovskites.

5.5 Synergies with other computational and experimental efforts

a) Within MARVEL The project has two active collaborations with experimental groups and several contributions from the theoretical community. On the experimental side, we are investigating the catalytic properties of perovskites like $La_{1-x}Sr_XCOO_3$ for the oxygen evolution reaction [16] together with Thomas J. Schmidt's group at PSI. Together with Thomas Lippert's group (PSI), we are comparing the band gaps of different stoichiometries of YTaON materials with experiments. In addition we are investigating the structural and electronic properties of tantalate and niobate oxynitride perovskites previously suggested [44, 45].

On the theoretical side, several collaborations have been established for the testing of new pseudopotentials. The group of Stefan Goedecker, for example, has provided novel norm-conserving pseudopotentials for testing.



b) Outside MARVEL Due to the fact that felectrons elements have not been much studied within the pseudopotential approximation and only recently low-cost band gap calculations have been possible, additional efforts have to be made in the production of band gap calculations. This will be done in collaboration with the group of Prof. Jens K. Nørskov (Stanford University, USA) which is working in the implementation of the GLLB-SC functional in the Quantum-ESPRESSO code. We will work in close contact with the group of Prof. Karsten W. Jacobsen (DTU, DK) where the descriptors for the identification of novel light harvesting materials have been first discovered and with the developers of the GPAW code (DTU, DK), in which the GLLB-SC functional has been originally implemented.

In addition, we have a broad collaboration with around 70 scientists for the verification and validation of pseudopotential libraries.

5.6 Personnel

NCCR personnel

- Ivano E. Castelli, postdoc, 50%, from September 2014 to February 2015.
- Nicolas Mounet, postdoc, 100%, from February 2015.

Matching personnel

- Ivano E. Castelli, postdoc, CE-COFUND fellow 40%, from March to August 2015.
- Gianluca Prandini, PhD student, 100%, from January 2015.

6 The industrial way for post-combustion capture of CO₂: optimizing the solvent (Wanda Andreoni — EPFL)

6.1 Research summary

Post-combustion carbon capture mainly employs solvent wet scrubbing, primarily using alkanolamine aqueous solutions as chemical absorbents. The crucial need for more efficient solvents has promoted a widespread empirical effort for the optimization of the solvent and of process-related parameters. For an independent, although partial, screening of amine solvents we use computer simulations that enable us to characterize the main chemical reactions involved in both the uptake of CO₂ via absorption and its release, to gain an understanding of the corresponding mechanisms and the role of different physico-chemical factors. The first step of this research was the comparison of two primary amines.

6.2 Scientific goals

The heart of our research is a detailed comparison of the behavior of different amine aqueous solutions in the uptake and release of CO_2 . In particular, we estimate the free-energy barriers associated with activated reactions leading to the formation of different products. In analogy with the empirical search for optimum absorbents, we use as reference aqueous monoethanolamine (MEA) at 30 wt.%, namely the most frequently used amine solution in the removal of CO_2 from combustion gases. In particular, we refer to our own results which have led to a detailed characterization of several reaction paths [46, 47].

a) *Short-term* The first choice for our comparison is on primary amines (like MEA) and is motivated by suggestions from empirical search and the availability of experimental data for verification. We have focused on 2-amino-2-methyl-1,3-propanediol (AMPD) and ammonia solutions.

b) *Long-term* The long-term goal of this research is to establish the physico-chemical factors that can be expected to affect — and possibly improve — the dynamics and relative importance of the reactions accompanying the uptake of CO_2 and its release with simultaneous amine regeneration. Attempts to-date have almost exclusively taken into account intrinsic properties of the amine molecules, such as structure and pK_a. Our simulations should provide additional and important information on the influence of the characteristics of the specific solution on the reaction parameters

and the role of water and concentration in different cases. In particular, it is essential to obtain robust arguments for the prediction of carbamate formation because carbamates are considered to provide fast kinetics, but also lower absorption capacity and too high regeneration energy.

6.3 Results obtained since the last report

AMPD is the only primary amine emerging from the outstanding few candidates identified as promising alternatives to MEA [48], on the basis of a higher absorption capacity (0.7 vs 0.5). In analogy with the case of MEA, measurements of the kinetics of the absorption process — via stopped-flow technique — have been interpreted as involving a (metastable) zwitterion, leading to the formation of a carbamate. However, the stability of the carbamate is lower than in MEA: in some cases it was observed at early stages of CO₂ loading [49] and was not observed in others [50].

Our simulation protocol [47] is based on DFTmolecular dynamics (Car-Parrinello MD) [51] and uses metadynamics [52, 53] as samplingenhancement technique with simple geometric collective variables. Free-energy profiles are reconstructed from MTD and also refined with umbrella sampling [54] in the neighborhood of transition states. For each reaction, we have simulated both direct and reverse processes. In this way we can also estimate the relative thermodynamic stability of reactants and products in solution.



Figure 13: AMPD and MEA isomers.



Figure 14: Zwitterion formation.

The molecular structure of two AMPD isomers is shown in Fig. 13 and compared with MEA. The difference with the latter is the addition of one more hydrophilic hydroxyl functional group and of a hydrophobic methyl group.

Snapshots from our simulations (Figs. 14 and 15) illustrate key steps of the CO₂ uptake from an AMPD solution. The formation of the zwitterion (Fig. 14) is the rate limiting step — as in the case of MEA — and also corresponds to a free-energy barrier of the order of 10 kcal/mol. Specific values do not differ significantly for the two amines, within the accuracy of our calculations. Deprotonation of the zwitterion (Fig. 15) corresponds to a lower free-energy barrier (by 5-6 kcal/mol) and is the only activated process in the formation of both carbamic acid and carbamate, which is facilitated by water via barrierless formation of a hydronium as mediating agent or of a proton wire. This scenario is analogous to the one found in the case of MEA. However, in the latter the formation of carbamic acid had to be "guided" and was never observed to take place in MTD simulations using only the N-H distance as reaction coordinate, in contrast to the case of AMPD. This discrepancy led us to an interesting finding in our models, namely from examination of the paths available to the proton in the solution, we observed a much reduced accessibility to other amines in the case



Figure 15: *AMPD-CO*₂: from zwitterion (left) to carbamic acid (top right) or to carbamate (bottom right).

of AMPD. Therefore, in our view, the difference between the two amines, that indeed renders carbamates less available for AMPD, is due to the different interaction of the amine-derived zwitterion with water. This discrepancy must be mainly related to the presence of the hydrophobic methyl group in AMPD that renders the aqueous environment much less isotropic than in MEA.

On passing from MEA to AMPD, the free energy barrier for the direct release of CO₂ from carbamate decreases (from 50 to 35 kcal/mol), whereas it is unaltered for the release from the zwitterion (7-10 kcal/moll). The deprotonation barriers of the protonated amines are within 1 kcal/mol, therefore irrelevant within the accuracy of the calculations, but consistent with expectations from the difference in pKa (from 9.5 to 8.84). In our previous work on MEA [47]., we have suggested that the most probable mechanism of the release of CO₂ involves again the zwitterion and that the amine deprotonation is the rate-limiting step. Our findings for AMPD are also consistent with this proposal. No improvement of the amine regeneration process has indeed been reported.

Further inspection of direct and reverse reactions indicate that, contrary to MEA, in the AMPD solution the systems carbamate + protonated amine are less stable relative to the amine solution with solvated CO_2 . Although this finding is in agreement with observations, more accurate calculations are necessary to confirm this estimate.

Partial results for aqueous ammonia — that we will complete soon — indicate easier formation of the zwitterion relative to MEA and AMPD, which also triggers easier formation of the carbamate. The system carbamate + protonated ammonium is also found to be stable as in aqueous MEA.

6.4 Planned research for next year

In the second phase of our research we intend to explore ammonia solutions [55] and

ring amines — like benzylamine [56] — which have more recently emerged as effective CO_2 absorbents, and also to adopt novel more advanced metadynamics strategies (e.g. [57]) and explore the outcome of more sophisticated DFT functionals, so as to verify and possibly increase the reliability of our predictions.

6.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Thus far, we have had no opportunity of interaction with other MARVEL

groups. We intend to collaborate with Alfredo Pasquarello and Michele Parrinello.

b) *Outside MARVEL* For method development: we are collaborating with Fabio Pietrucci at the Pierre et Marie Curie, Paris 6.

6.6 Personnel

NCCR personnel

• Changru Ma, postdoc, 100%, from March 2015.

7 A search for solid-state lithium conductors (Nicola Marzari — EPFL)

7.1 Research summary

Liquid salts and organic solvents, used in commercial batteries as electrolytes, are flammable and subject to the formation of passivating films. Solid-state electrolytes avoid these problems by construction, but at the moment they do not show a comparable transport performance combined with an acceptable electrochemical stability. In order to guide the engineering of next generation lithium ion batteries, we will exploit accurate first-principles techniques in a DFT framework and select optimal materials in a high-throughput scheme. Our project runs in collaboration with the project in the IBM group, part of MARVEL.

7.2 Scientific goals

As a first step of this project we aim at a detailed understanding of the microscopic processes related to lithium diffusion. The second step consists in using the acquired knowledge to find descriptors correlated to high ionic conductivity. This would enable a search for candidate electrolyte materials in a high-throughput scheme.

a) *Short-term* The short-term goal is divided into two parts, proceeding in parallel. In the first part we will analyze selected and known high-performing solid-state electrolytes using molecular dynamics (MD) simulations, a precise but expensive technique. The results of this analysis will be used to construct kinetic Monte Carlo (KMC) models. The KMC will be based on a novel implementation of automated site recognition using algorithms commonly employed in computational geometry, and fed with transition rates observed in MD simulations.

The goal is to find descriptors for occupation densities and transition rates in comparable ap-

proach as taken by Wang *et al.* [58]. The IBM group, following a different strategy for computing transition rates and sites, will provide valuable and mutual cross-verification. The second part consists in the implementation and automation of the analysis tools for site recognition. This set of tools will be developed under the AiiDA framework [17].

b) *Long-term* Guided by the experience acquired on selected materials and using automated tools adapted to a high-throughput approach, we plan to repeat the previous analysis on a large number of systems. In close collaboration with the IBM group, we will first include structures of the garnet family. The possibility to vary the Li-composition and the garnets' well-studied geometric structure make it an ideal choice for our site-prediction algorithms. The second step would be to include all Li-containing oxides. We will benefit from the expertise of the IBM group on this topic.

7.3 Results

We produced several long runs at different temperatures for the fast ionic conductor $L_{10}GeP_2S_{12}$ [59], with temperatures from 100 K to 1000 K, using Car-Parrinello MD, and observed the typical unidimensional diffusion characterizing transport in this material, as depicted in Fig. 16. We understood that long trajectories, of the order of 100 ps, are necessary for having a good statistical sampling of phase space, suggesting that MD can only be used to find descriptors and not for material screening. By freezing all ions other than lithium in an equilibrated configuration we observed a large drop in the diffusivity. This means that the motion of the non-diffusing species can be correlated to the lithium diffusivity, thus motivating a phonon analysis of the rigid sublattice

Results — VP2



Figure 16: (*Top*) *Typical plots of lithium mean* square displacements at various temperatures and (bottom) lithium density plot for L_{10} GeP₂S₁₂.

at a later point. The importance of the rigid sublattice has also recently been highlighted by Wang *et al.* We also studied the difference between ionic conductivity, related to collective behavior, and the tracer diffusion coefficient, related to single particle motion, coming to the conclusion that collective behavior is fundamental in this material. Collective modes seem to favor transport and lead to a failure of the Nernst-Einstein equation, which is often applied without corrective factors.

Regarding the automatic site recognition, we developed and implemented an algorithm to tessellate the volume of an arbitrary unit cell into meaningful convex hulls, as depicted in Fig. 17. This was done together with Boris Kozinsky from Bosch RTC. The algorithm combines information from the Voronoi decomposition and Delaunay triangulation on the atomic positions with geometric pattern recognition. The implementation was tested on over 1'000 structures, and found to be robust and accurate. We developed tools to track ions during an MD simulation and detect which site each ion occupies. This allows us to project the continuous trajectory of an ion to a discrete sequence of visited sites. We can therefore detect jumps with a high spatial and temporal accuracy. As a by-product of the cell subdivision, we discovered an automated way to detect melting, enabled by monitoring the volumes of sites throughout the trajectory. The publication of all these findings is in prepara-



Figure 17: Tessellation of a supercell of $L_{10}GeP_2S_{12}$ on a random snapshot during the simulation. Nondiffusive species (*Ge*, *P*, *S*) define the convex hulls of interstitial sites.

tion [18, 19].

7.4 Planned research for next year

We are currently testing the site definition tool on L₁₀GeP₂S₁₂, which is a fundamental ingredient to fit a continuous time Markov chain to each long trajectory. In a first stage we will estimate the transition rates directly from observed jumps in the MD simulation. We will investigate how detailed balance and microscopic reversibility manifest in a coarse grained dynamics. Recomputing the diffusion coefficient from the KMC and comparing it with the one obtained from molecular dynamics will indicate whether transition rates can be used as a descriptor for lithium conductivity. In case of a positive result we will search for descriptors for transition rates on a wide range of materials. The results obtained in the IBM group will be important to guide this research.

7.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* the group of Teodoro Laino, Ivano Tavernelli and Matthieu Mottet at IBM is working on the same problem. They use simulation strategies based on energetics and accelerated sampling methods which will provide a complementary piece of information. Combining the knowledge acquired will favor

the search for useful descriptors of ionic diffusivity.

b) On the experimental side we get insight from Prof. Bettina Lotsch, within the joint collaboration provided by the Max-Planck-EPFL Center for Molecular Nanoscience & Technology.

c) *The set of tools* will be implemented inside the AiiDA framework, developed by the MAR-VEL group. One of the members of the project is an active AiiDA developer.

7.6 Personnel

NCCR personnel

• Aris Marcolongo, postdoc, 100%, from January 2015.

Matching personnel

• Leonid Kahle, PhD student, 100%, from September 2015.

Collaborators

- Boris Kozinsky, Robert Bosch LLC, Research and Technology Center, Cambridge, MA, USA.
- Bettina V. Lotsch, Max-Planck-Institut für Festkörperforschung, Stuttgart, D.

8 Material screening for metal/air batteries (Alessandro Curioni — IBM)

8.1 Research summary

Next generation lithium batteries still faces challenges in the domain of safety and stability. To overcome these difficulties, the combined use of simulation and data-analytics will prove to be a game-changing methodology in the forthcoming years. In this project, we investigate the conductivity of zirconium containing lithium-lanthanum-oxides as a proofof-concept in the development of a framework for fast and reliable evaluation of the ionic conductivity of novel solid-states electrolytes.

8.2 Scientific goals

a) Short-term Kinetic Monte Carlo simulations are a fast way to estimate the ionic conductivity of a material. During this first year, we examined different approaches for the determination of kinetic parameters, such as energy and free energy barriers, and kinetic preexponential factors. Parameters obtained from both static computations and statistical analysis of dynamical simulations are used and tested to improve our understanding of the conductivity phenomenon in LLZO. Machine learning algorithms are also investigated to infer parameter sets or conductivity of new potential electrolyte transporters based on experimental and theoretical data. Finally, we aim at bridging the theoretical effort developed so far with the big-data analytics and cognitive tools designed in the project HP5.

b) *Long-term* Building on the approach developed for the generation of reliable parameters for LLZO, we aim to transfer the acquired

expertise to a large set of potential solid-state electrolytes. Such methods will allow to build a strong framework to investigate a variety of candidate structures in a fast, automated and reliable fashion.

8.3 Results obtained since the last report

a) First principles and kinetic Monte Carlo simulations A kinetic Monte Carlo software was developed and tested within LLZO solidelectrolyte family. The extraction of parameters from static and dynamical computations using different approaches (DFT dynamics, constraint dynamics, nudged elastic band transition state search, and metadynamics) and their testing was performed. Fig. 18 shows a recent estimate of the Lithium ion conductivity in LLZO using current dataset. Convergence of the conductivity parameter σ is obtained in the ns time scale.

b) *Investigation of machine learning descriptors* Machine learning algorithms need to rely on good descriptors able to provide valuable information about the conductivity process. Finding such quantities is as important as building the machine learning engine used for the inferences. During this first part of the project, we developed a geometrical based approach for the rapid detection of ions channels in periodic structures. This approach allows a first quick selection of potential solid-state electrolyte candidates within and outside the Garnet family (Fig. 19).

Furthermore, in collaboration with Horizontal Project 5 (HP5) at IBM and UniBas, we Results - VP2



Figure 18: Conductivity estimated using the kinetic Monte Carlo technique at 300 K. The barriers were estimated by single points calculations and the pre-exponential factor from frequency analysis of molecular dynamics.

are working on the generation of a strong platform for the collection, storage and crossexamination of large amount of published scientific data. We plan to apply this technology to explore the already existing structural information for the identification of potential new solid-state electrolyte transporters.

8.4 Planned research for next year

a) First principles and kinetic Monte Carlo simulations Further investigation of the LLZO family will lead to methods for the generation of more reliable parameters for the prediction of more accurate values of the conductivity. In particular, we are interested in using statistical



Figure 19: Superposition of the lithium scaffold in the LLZO cell and of candidate channels detected by our algorithm.

analysis of trajectories obtained from molecular dynamics simulation to estimate kinetic parameters to use in our KMC scheme. Using this improved parametrization of our KMC model we will investigate the changes in conductivity observed experimentally for a series of doped and metal-substituted LLZO materials. This is a fundamental step for the validation of our approach and it will give us important insights about the level of confidence and predictivity we can expect from this simulation approach. A transfer to other class of solid-state electrolyte transporters will also be considered to see if our current procedure, despite being approximate, holds universally valuable information.

b) *Investigation of machine learning features* With the further development of the material platform (within the HP5 project) and thanks to the design of new descriptors, we work on the design of a material discovery oriented machine learning algorithm that we will train on the database extracted using the cognitive data analytics tools.

8.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* The VP2 team at IBM is working in close contact with the HP5 group in the development of the structural database and related data analytics tools for applications to solid-state electrolytes materials.

We are also planning to collaborate with the group of Nicola Marzari on the refinement of the KMC parameters.

8.6 Personnel

NCCR personnel

- Matthieu Mottet, PhD student, 100%, from June 2015.
- Ivano Tavernelli, senior researcher, 100%, from January 2015.

Matching personnel

• Teodoro Laino, senior researcher, 75%, from January 2015.

9 High-throughput vibrational, dielectric and thermomechanical properties of novel materials (Nicola Marzari — EPFL)

9.1 Research summary

This project aims at providing a database of structural, electronic, magnetic and thermodynamical properties of crystals computed from first principles. Density-functional theory calculations are systematically performed on known bulk crystals, as well as common or novel two-dimensional materials that can be exfoliated from existing three-dimensional ones. As a first step, we are constructing a systematic database of bulk layered materials, their associated 2D mono- and multilayers, and their vibrational, dielectric, spectral and thermomechanical properties.

9.2 Scientific goals

a) Short-term The first phase of the project has two objectives. On the one hand, it aims at constructing a set of tools enabling the high-throughput computation of structural relaxation and phonon dispersions of arbitrary structures from density-functional theory, in a systematic way. These tools are being developed within the AiiDA computational platform [17] in synergy with ongoing work within PP6. To this purpose AiiDA is interfaced with the Quantum-ESPRESSO distribution [60] through automated workflows. On the other hand, a core short-term goal of the project is the construction of a large database of two-dimensional materials and their bulk layered counterparts, automatically screened from crystallographic databases. The AiiDA workflows developed are then applied to the structures of this database, thus allowing predictions of manyfold properties of these novel materials on a large scale.

b) Long-term In a second phase, the set of tools developed and tested on 2D and layered materials will be applied more generally to three dimensional crystals, as available in structural databases such as the Inorganic Crystal Structure Database (ICSD) [61] and the Crystallography Open Database (COD) [62]. Phonon frequencies and the vibrational thermomechanical properties will be computed and made available. In addition, quantities that can be derived easily from the same linearresponse calculations will be added, including dielectric tensors, effective charges, Raman intensities and infrared spectra, powder diffraction spectra, Debye-Waller factors and mode Grüneisen parameters.

9.3 Results obtained since the last report

The two main axes of the short-term goals were addressed in parallel: on one side, automated and robust workflows were developed to allow seamless high-throughput calculations using Quantum-ESPRESSO, and on the other side, a first-principles database of layered and 2D compounds was constructed. We will describe both these aspects in the following.

a) AiiDA workflows Scientific workflows are developed to automatize DFT calculations performed on a given crystal structure, reducing human intervention to the minimum. First of all, we apply this philosophy to the initial "learning phase" that one has typically to go through when beginning computations on an unknown structure. In many past and current studies, this phase is performed manually for its most part, i.e. trying different calculation setups and convergence parameters before finally getting a sense of the numerical configuration to be used to obtain reliable results. Here we encode this effort by designing a workflow, called the CHRONOS workflow, that will ultimately automatize the whole process.

The goal of this workflow is to distinguish between magnetic and nonmagnetic materials, as well as between metals and insulators, and then to produce a fully relaxed energy calculation together with electronic band structure. The main steps performed in the workflow are represented in Fig. 20. From a given structure, a first set of structures is created where a different magnetization is assigned individually to each atom, possibly also using supercells. This allows to launch magnetic tests in which the spin value imposed on the valence electrons of each atom is set randomly. A non



Figure 20: The CHRONOS workflow.

spin-polarized calculation is also launched in parallel, as well as a simple magnetic configuration on a single cell. From all these computations the workflow chooses the lowest energy configuration, and among those with same or similar energy, the simplest one.

In order to assess the metallic or semiconducting character of this lowest energy configuration, a first preliminary calculation of the electronic bands along a path in the Brillouin zone is carried out. This path follows high-symmetry lines that are determined automatically, depending on the Bravais lattice of the structure studied. From the energy bands, the direct/indirect band gap is calculated, enabling the identification of metallic compounds.

Finally, a calculation of the relaxed structure and electronic bands is performed, using the magnetic setup used to obtain the lowest energy magnetic configuration; smearing is adapted to this kind of compound (metallic or insulating), and typically tight convergence parameters are chosen, since we want this final calculation to be as accurate as possible.

Note that each of the calculations (energy and band structures) mentioned above are actually performed thanks to a sub-workflow, the ENERGY workflow, which deals with all the machinery needed to launch actual Quantum-ESPRESSO calculations on the cluster, restarting them if needed, and handling the most common failure cases in an automatic way.

The fully relaxed calculation performed at the end of the CHRONOS workflow, is then used as input for the PHONON workflow, shown in Fig. 21. First, for each \mathbf{q} point on a grid of the Brillouin zone, the dynamical matrix is computed self-consistently from density-functional perturbation theory; these computations are typically launched in parallel to maximize the calculation efficiency. Then, the dynamical matrices are collected and Fourier-interpolated to obtain the phonon frequencies on a fine mesh of \mathbf{q} vectors.



Figure 21: The PHONON workflow.



Figure 22: Set of band gaps and magnetizations computed with the CHRONOS workflow, using Quantum-ESPRESSO and PBE exchange-correlation.

As for the energy computations, each individual phonon calculation at a given **q** point is actually performed by a sub-workflow that manages again both the restarting of calculations and several failure cases (e.g. too slow selfconsistent convergence).

In the end, these workflows allow the highthroughput computation of ground state energies, structure relaxations, magnetic properties, electronic band structures and vibrational dispersions, with manual intervention limited to the minimum. A simple example of results from the CHRONOS workflow is given in Fig. 22 where we show the total and absolute magnetizations vs band gap for a set of 81 two-dimensional materials (also found in [63]). From this plot one can distinguish four kinds of materials: metals, insulators, magnetic metals, and magnetic insulators, the later category being relevant for electronic applications.

Finally, as an example of the application of the PHONON workflow, we show in Fig. 23 the phonon dispersion curves of the two-dimensional material $PtTe_2$.



Figure 23: *Phonon dispersion of two-dimensional PtTe*₂*. Its crystal structure is also shown.*

b) Layered and 2D materials database To build a database of layered materials, our starting point consists in a set of bulk 3D structures extracted from the ICSD [61] and COD [62] databases. The structures downloaded from these databases are not always readily usable, and a certain number of initial steps are required. To begin with, all compounds with partial occupancies, or where some critical information is missing (such as atomic positions), are filtered out. Then, using COD-TOOLS [62] the CIF files [64] are searched and repaired, and the structures refined using the SPGLIB software [65] to avoid loss of symmetry due to round-off errors in the definition of the atomic positions or cell parameters. The primitive cell is also obtained for each crystal. Finally, redundant structures are found using the relevant PYMATGEN tools [66], and removed.

From the set of structures obtained, a systematic screening of potential two-dimensional structures is launched. As a first step, all interatomic distances are evaluated on a supercell of size $3 \times 3 \times 3$, and chemical bonds are identified as those for which

$$d_{i,j} < r_i^{vdw} + r_j^{vdw} - \Delta, \tag{1}$$

where $d_{i,j}$ is the distance between two atoms *i* and *j*, r_i^{vdw} is the van der Waals radius of atom *i* as obtained from [67], and $\Delta = 1.3 \pm 0.2$ Å. Two-dimensional substructures are then found from connected groups which are periodic in two independent directions. Note that the algorithm is general and does not assume any specific orientation of the 2D plane with respect to the crystal or Cartesian axes, contrary to a previous systematic screening of the ICSD database [63]. Moreover, the algorithm also finds lower dimensionality substructures, such as 1D chains or 0D clusters. Examples of layered and 2D structures found using this algorithm are shown in Fig. 24.

The point groups of the 3D layered structures



Figure 24: *Examples of layered structures extracted from the ICSD database.*



Figure 25: Distribution of point groups in the layered materials found.

found are very diverse, as shown in Fig. 25. Some low-symmetry point groups, such as $\overline{1}$, 2/m or mmm are very common, as well as others with more symmetries, such as 4/mmm or $\bar{3}$ m. No cubic point group is present, since the equivalence of the three spatial directions is incompatible with a layered geometry. From the knowledge of the bulk point group, one can in principle find the Raman and infrared activity fan diagrams of any multi-layer slab of such materials. In order to achieve this goal we first take advantage of the weak nature of van der Waals coupling between the layers, so that we can consider only nearest-neighbor interactions. In addition, we are interested in lowlying vibrational modes when the layers move rigidly with respect to one another. This allows us to compute the normal modes of the multilayer structure by mapping the problem exactly onto a linear chain model, where the spring constants are related to the elastic properties of the material. In order to determine if these normal modes are Raman and/or infrared active, we need to know the point group of the multilayer, which in general will be different from the one in the bulk and will depend on the number N of layers and on the stacking order. By using group-theoretical arguments, we identify the evolution of the point group from the bulk to any possible multilayer structure for all 530 Hall space-group settings of the 230 space groups. For example, we consider MoS₂ for which the space group number is 194, and Hall space-group number is 488. Using this method we obtain that the point group evolves from 6/mmm in the bulk to $\overline{6}m2$ when *N* is even and $\bar{3}m$ when *N* is odd. In Fig. 26 we report the fan diagram of MoS₂ showing the frequencies and infrared/Raman activity of the so-called layer-breathing modes as a function of the number of layers obtained through the above procedure.

From the set of layered structures and their two-dimensional substructures found by the


Figure 26: Fan diagram of MoS₂ showing the frequencies and infrared/Raman activity of the socalled layer-breathing modes as a function of the number of layers N obtained through the universal automated procedure described in the main text.

screening algorithm described above, we perform first DFT computations to check the binding energy between layers. For this we use the Quantum-ESPRESSO software [60] with the SSSP (version 0.7) [20] pseudopotentials and the optimal cutoffs. We shall insist here on the importance of the SSSP library, resulting from a parallel project within MARVEL VP2: it provides converged and carefully checked pseudopotentials for all the elements in the periodic table, which is essential for our calculations.

To compute the binding energy we first relax the cell and atomic positions of the 3D layered structure, before extracting the 2D layers from the structure obtained and calculating the self-consistent energy (this time without relaxation). The binding energy is then simply the difference in energy between the 3D bulk structures (after relaxation) and the sum of the energies for each individual layer. This procedure is performed with two different van der Waals functionals, namely the vdW-DF2-C09 [68] and the revised VV10 [69].

In Fig. 27, the binding energies between the



Figure 27: Distribution of the binding energies of the bulk 3D vs individual 2D layers. Note that the distribution extends also out of the range.

	ICSD	COD	
Database entries	172′370	132'695	
downloaded	$(\leq 6 \text{ species})$	$(\leq 4 \text{ species})$	
Structures selected	50′584	52'314	
Distinct layered	1/868	608 (+1'100 also	
materials	4 000	in the ICSD)	
Distinct 2D	4/057	684 (+1'306 also	
materials	4 937	in the ICSD)	
Relaxed 3D	1'387 (vdW-DF2-C09)		
structures	660 (<i>rVV10</i>)		
Binding energies	359 (vď	W-DF2-C09)	
computed	346 (rV	V10)	

Table 1: Database statistics. The structures selected

 (2nd row) are all the exploitable 3D non-redundant

 structures.

layers is exhibited for a subset (~ 350 entries) of the layered compounds found. A peak is visible for both functionals around 20 meV/Å², which is consistent with other studies on van der Waals bonded compounds [70]. More specifically, calculations using the random-phase approximation done for graphite and boron nitride resulted in binding energies of respectively 18 and 14 meV/Å² [70].

Finally, Table 1 summarizes the statistics of the database in its current state. From the 305'065 entries looked at in the ICSD and COD databases, 5'476 three-dimensional layered structures are found, translating into 5'641 two-dimensional materials. This represents the largest database of layered materials reported so far.

9.4 Planned research for next year

On the AiiDA tools side, the workflows developed will continue to be improved. In particular, we should improve the magnetic screening algorithm in the CHRONOS workflow, as the current implementation with purely random magnetizations in supercells might be both inefficient and rather time-consuming. In the same workflow, a systematic convergence check of the k points grid should also be performed (currently conservative estimates are used). Finally, the stresses during a variable cell relaxation should be corrected by taking into account the Pulay stress in a systematic manner. These improvements are necessary to obtain later accurate phonon dispersions curves, in particular for the ZA branch close to the Gamma point in 2D materials.

Then, these workflows will be extensively applied to the database of 3D layered and 2D materials that we have built. The ultimate goal will be to get the phonon dispersion curves of all these materials and for various lattice parameters, in order to be able to

compute thermomechanical properties (as well as Raman intensities, infrared spectra, powder diffraction spectra, Debye–Waller factors, mode Grüneisen parameters, dielectric tensors and effective charges).

We will also screen materials with certain characteristics, e.g. ferroelectrics, piezoelectrics, and superconductors, among others.

9.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Synergies with PP6 are naturally taking place through the improvement of the workflows within the AiiDA platform. Moreover, the AiiDA core code itself is improved continuously following the progress of this work.

This project is also in close interaction with the ongoing efforts on the testing of pseudopotentials, which takes place within VP2.

Finally, the usage of this database platform will be broadened to all MARVEL.

b) *Outside MARVEL* A collaboration with Prof. Andrea Ferrari from the University of Cambridge is taking place to compare theoretical predictions on layered materials with ex-

periments, in particular for the Raman and infrared spectra and the elastic constants.

9.6 Personnel

NCCR personnel

- Nicolas Mounet, postdoc, 100%, from February 2015.
- Giovanni Pizzi, postdoc, 100% (on PP6), from September 2014 to July 2015 and from January 2016.

Matching personnel

- Giovanni Pizzi, postdoc, 100%, from August to December 2015.
- Andrea Cepellotti, PhD student, 100%, from May 2014.
- Philippe Schwaller, master student, 30%, from September 2014.
- Marco Gibertini, postdoc, 100%, from May 2014.

10 Systematic doping of molecular precursors: toward a high-throughput scanning of carbonbased nano-material properties (Daniele Passerone — Empa)

10.1 Research summary

In this project we plan to apply to carbon-based nanostructures methods developed in MAR-VEL both for the high-level calculation of electronic properties and for the high-throughput screening of structure and composition. The Koopman's schemes developed within MAR-VEL are particularly suitable for experimentally accessible heterojunctions containing porphyrins. Based on the previous successful results, we will then focus on heterojunctions and their optical properties, their doping with O and H, and in collaboration with IBM, we will perform the first steps toward the automatic screening of the photophysical properties of a well-defined class of carbon nanosystems.

10.2 Scientific goals

a) *Short-term* The short-term goal of this project is the calculation of the electronic properties of selected carbon based heterojunctions, with particular emphasis on graphene nanoribbons (GNRs) using methods beyond DFT.

We will rely on the collaboration with HP3 (Marzari) to apply Koopman's schemes [21] to our graphene-based nanostructures. In addition, we will investigate the effect of contaminants on such properties. An example of the structures that we could investigate is provided in Fig. 28, a summary of some molecular precursors used at Empa and the corresponding possible resulting heterostructures that can



Figure 28: Examples of molecular precursors (a) and heterojunctions (b-f) that could result from the bottom-up joint experimental and theoretical strategy deployed at Empa. Figure courtesy of *P.* Ruffieux (Empa).

be obtained with bottom-up synthesis.

b) *Long-term* The long term view of this project is the ability to design, by automatic parameter screening, carbon based heterostructures whose properties are of interest for photovoltaic applications and/or for IT technologies.

This objective will cover several years of activity within MARVEL. The ongoing first step of the long term strategy concerns the identification of appropriate indicators to discriminate between "useful" and "useless" nanomaterials. Consequently, simple computational protocols to associate a particular geometry and composition to a set of values for such indicators must be defined. At this point we will be ready for high-throughput screening by means of the AiiDA platform.

10.3 Results obtained since the last report

We focused part of our research activity on a recent experimental/theoretical controversy. A recent theoretical work by Mazzarello and coworkers [71] investigated the spin-polarized states at the edges of adsorbed zigzag nanoribbons (ZGNRs) on Au(111). According to that work, such states should not be quenched on the metallic substrate. Experiments performed in our laboratory showed that such edge states are not visible by scanning tunneling spectroscopy (STS) when the ribbons are directly adsorbed on Au(111). When intercalated with an insulating interlayer, instead, the measurements match the theoretical prediction. We could show theoretically that when considering edges modified with a phenyl unit the states should be well visible also on the metal; this prediction has been confirmed experimentally, together with the particular spatial distribution of such states along the edge [22]. The disagreement between theory and experiment for the pristine ribbon case remains to be solved.

An important achievement of our research activity of last year, obtained jointly within MAR-VEL and the SNSF project 153661, is related to Fourier transform (FT-) scanning tunneling spectroscopy (STS). In an *ab initio* study of the band structure of GNRs, we conclude that the interpretation of the experimental STS spectra requires great care. Our theoretical analysis [23] explains the high STS intensity observed at the edges [72] by using simple symmetry arguments. The *ab initio* STS simulations reveal the validity of the Fourier-transform approach, that can be applied to single adsorbed nanostructures, at variance with usual angle-



Figure 29: Chemical sketch of a 7-AGNR (armchair graphene nanoribbon) consisting of 5 biantryl units (a) together with QPlus atomic force microscopy (AFM) image (b) and scanning tunneling microscopy (STM) image (constant current) (c). The variation of the band gap as a function of length and termination is plotted in (d). Experimental data (dashed lines) are shifted by $\Delta_0 = 0.73$ eV for a better comparison with DFT results (continuous lines) The black dotted line at 1.57 eV represents the bulk value.

resolved photoemission spectroscopy (ARPES) approaches that require a periodic distribution of objects on the surface.

Another achievement in collaboration with experimental teams concerns the study of GNRs of different length and termination. We focused on the case of the N = 7 armchair GNR (Fig. 29) and could explain why short CH-terminated ribbons exhibit a larger band gap when compared to CH₂-terminated ribbons of the same geometric length. We demonstrated how the π -electronic framework of GNRs is sensitive even to single atomic modifications, and that the exact dependence of the gap on the length can be rationalized by introducing an "effective length" of the ribbon that keeps into account the termination.

10.4 Planned research for next year

We list here the planned milestones for next year.

- In collaboration with PP6 we will develop an AiiDA plugin for CP2K that will be useful also for a project to be conducted outside MARVEL (SNSF proposal 165526, under evaluation) where we plan the investigation of electronic properties of heterojunctions by means of the GW schemes that are currently under development in HP3 (Hutter).
- We plan to apply the method developed by Marzari to study the electronic and optical properties of heterojunctions such as depicted in Fig. 28. The method has been

reported to be very efficient and accurate and would be extremely useful for heterojunctions containing porphyrin molecules (Fig. 28f).

- Still related to the heterojunctions plotted in Fig. 28, we plan here to do an extensive study at the DFT level of the effect of contaminants (such as O and H) on the electronic properties of the heterojunctions.
- Although the characterization of the photophysical properties of selected heterojunctions (selected by the outcome of preliminary screening done within and outside MARVEL) is a long term project that relies on the collaboration with IBM, we plan already this year to perform preliminary calculations on model systems to assess the validity of the methodology and to define the possibility for automatic screening that will then be used in a more mature phase of the project.

10.5 Synergies with other computational and experimental efforts

a) Within MARVEL The long term view of this project is to screen a large set of graphene based heterojunctions for their electronic and optical properties, in order to identify appealing nanomaterials for photovoltaics applications. The interplay with activities in the other MARVEL groups will be essential for defining a roadmap for the accurate computation of the photophysical properties of this class of systems. We hereby describe the collaboration within VP2 with IBM with the focus on non-adiabatic molecular dynamics simulation schemes.

Recently, a TDDFT-based [73] non-adiabatic mixed quantum-classical molecular dynamics in the CPMD code was implemented. This approach is based on the trajectory surface hopping (TSH) scheme of Tully [74]. The approach exploits functionals of the occupied and virtual Kohn-Sham orbitals [75]. Linear response TDDFT (LR-TDDFT) [76, 77] is exploited in this scheme for computing excited state energies, forces and non-adiabatic coupling vectors which govern the conical intersections topological properties. An interesting recent development is the description of intersystem crossings (transitions between potential energy surfaces with different spin multiplicity), a promising perspective for a correct description of the fluorescence and phosphorescence phenomena. We forecast the following milestones for our collaboration with the IBM team.

- 1. Computing electronic properties of onedimensional graphene-related devices, possibly with intercalation of porphyrin units.
- 2. Studying the relationship between chemical modifications at the graphene edges and the size of the optical gap and oscillator strength. Tuning the excitation energy is fundamental for designing efficient light-absorbing dyes in the field of solar cell devices. Application to the valence and conduction band position in the nanostructure with respect to the corresponding bands in standard semiconductor junctions (e.g. TiO₂ nanoparticles), with the goal of an efficient electron transfer.
- 3. Estimate of excited state lifetimes by means of photoexcited dynamics (population dynamics in the singlet and triplet state, non-adiabatic coupling, intersystem crossing).
- 4. Design of novel nanostructures with potential in the field of efficient solar cell devices.

b) *Outside MARVEL* The issue of properly including surface effects in the electronic properties of graphene nanomaterials is being the object of a long-lasting collaboration with the Modena group of Prof. Molinari. We already applied a simple scheme of approximation in the past [78], and we plan to test novel approaches currently developed in the Molinari group and the surface embedding strategy exploited in the Louie group [72, 79].

Another important issue concerns the modeling of the van der Waals (vdW) interactions, in particular for molecules adsorbed on metal systems. In the case of the adsorption geometry and electronic properties of GNRs on noble metals, we can use the results from experiment as a validation for the different theoretical schemes. Topographic features (such as the corrugation profile of the adsorbed nanostructures) obtained with picometric precision using the Qplus atomic force microscope (AFM) [80] can be compared with simulated images calculated with tools such as the code realized in the Jelinek group in Prague [81].

10.6 Personnel

NCCR personnel

• Carlo A. Pignedoli, senior researcher, 60%, from October 2014.

Matching personnel

• Aliaksandr Yakutovich, PhD student, 15% from May 2015.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- [1] C. Yi, J. Luo, S. Meloni, A. Boziki, N. Ashari-Astani, C. Grätzel, S. M. Zakeeruddin, U. Röthlisberger, and M. Grätzel, Entropic stabilization of mixed A-cation ABX₃ metal halide perovskites for high performance perovskite solar cells, Energy & Environmental Science 9, 656 (2016).
- [2] S. Mathew, N. Ashari Astani, B. Curchod, J. Delcamp, J. Frey, U. Rothlisberger, M. K. Nazeeruddin, and M. Grätzel, Synthesis, characterization and ab initio investigation of a panchromatic ullazine-porphyrin photosensitizer for dyesensitized solar cells, to be published in Journal of Materials Chemistry A (2016).
- [3] K. Voïtchovsky, N. Ashari Astani, I. Tavernelli, N. Tétreault, U. Rothlisberger, F. Stellacci, M. Grätzel, and H. A. Harms, *In Situ Mapping of the Molecular Ar*rangement of Amphiphilic Dye Molecules at the TiO₂ Surface of Dye-Sensitized Solar Cells, ACS Applied Materials & Interfaces 7, 10834 (2015).
- [4] W. Tress, B. Beyer, N. Ashari Astani, M. Schwarze, F. Gao, R. Scholz, S. Meloni, U. Rothlisberger, and K. Leo, *Charge transfer exciton delocalization: the key to highly efficient organic solar cells, submitted* (2016).
- [5] S. Meloni, T. Moehl, W. Tress, M. Franckevicius, M. Saliba, Y. Lee, P. Gao, M. K. Nazeeruddin, S. M. Zakeeruddin, U. Rothlisberger, and M. Graetzel, Ionic Polarization Induced Current-Voltage Hysteresis in CH₃NH₃PbX₃ Perovskite Solar Cells, to be published in Nature Communications (2016).
- [6] S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B. F. E. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M. K. Nazeeruddin, and M. Grätzel, Dye-sensitized solar cells with 13% efficiency achieved through the molecular engineering of porphyrin sensitizers, Nature Chemistry 6, 242 (2014).
- [7] F. Ambrosio, G. Miceli, and A. Pasquarello, *Redox levels in aqueous solution: Effect of van der Waals interactions and hybrid functionals*, The Journal of Chemical Physics 143, 244508 (2015).
 - [8] W. Chen and A. Pasquarello, Accurate band gaps of extended systems via efficient vertex corrections in GW, Physical Review B 92, 041115 (2015).
 - [9] W. Chen and A. Pasquarello, First-principles determination of defect energy levels through hybrid density functionals and GW, Journal of Physics: Condensed Matter 27, 133202 (2015).
 - [10] F. Evangelisti, R. Moré, F. Hodel, S. Luber, and G. R. Patzke, 3d–4f Co₃¹¹Ln(OR)₄ Cubanes as Bio-Inspired Water Oxidation Catalysts, Journal of the American Chemical Society 137, 11076 (2015).
- [11] Y. Gurdal, S. Luber, J. Hutter, and M. Iannuzzi, Noninnocent adsorption of Co-pyrphyrin on rutile(110), Physical Chemistry Chemical Physics 17, 22846 (2015).
- [12] M. Busch, M. Wodrich, and C. Corminboeuf, Linear Scaling Relationships and Volcano Plots in Homogeneous Catalysis? Revisiting the Suzuki Reaction, Chemical Science 6, 6754 (2015).
- [13] G. Gryn'ova, M. Coote, and C. Corminboeuf, *Theory* and practice of uncommon molecular electronic configurations, WIREs Computational Molecular Science 5, 440 (2015).

- [14] R. Petraglia, A. Nicolaï, M. D. Wodrich, M. Ceriotti, and C. Corminboeuf, *Beyond Static Structures: Putting Forth REMD as a Tool to Solve Problems in Computational Organic Chemistry*, Journal of Computational Chemistry **37**, 83 (2016).
- [15] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Dulak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Granas, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Haann, P. J. Hasnip, N. A. W. Holzwarth, D. Iusan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Kucukbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordstrom, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, The Kohn-Sham equation of state for elemental solids: a solved problem, submitted (2015).
- [16] X. Cheng, E. Fabbri, M. Nachtegaal, I. E. Castelli, M. El Kazzi, R. Haumont, N. Marzari, and T. J. Schmidt, Oxygen Evolution Reaction on La_{1-x}Sr_xCoO₃ Perovskites: A Combined Experimental and Theoretical Study of Their Structural, Electronic, and Electrochemical Properties, Chemistry of Materials 27, 7662 (2015).
- [17] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, *AiiDA: automated interactive infrastructure* and database for computational science, Computational Materials Science 111, 218 (2016).
 - [18] Prateek Mehta, Leonid Kahle, and Boris Kozinsky, Descriptors of Lithium-Ionic Conductivity from ab initio Molecular Dynamics, in preparation (2016).
- [19] Aris Marcolongo, Leonid Kahle, and Nicola Marzari, Correlated motion in fast ionic conductors, in preparation (2016).
- [20] I. E. Castelli, Standard Solid State Pseudopotentials (SSSP), http://www.materialscloud.ch/sssp/.
- [21] N. L. Nguyen, G. Borghi, A. Ferretti, I. Dabo, and N. Marzari, First-Principles Photoemission Spectroscopy and Orbital Tomography in Molecules from Koopmans-Compliant Functionals, Physical Review Letters 114, 166405 (2015).
- [22] P. Ruffieux, S. Wang, B. Yang, C. Sanchez, J. Liu, T. Dienel, L. Talirz, P. Shinde, C. A. Pignedoli, D. Passerone, T. Dumslaff, X. Feng, K. Müllen, and R. Fasel, On-surface synthesis of graphene nanoribbons with zigzag edge topology, arXiv:1511.05037 (2015).
- [23] H. Söde, L. Talirz, O. Gröning, C. A. Pignedoli, R. Berger, X. Feng, K. Müllen, R. Fasel, and P. Ruffieux, *Electronic band dispersion of graphene* nanoribbons via Fourier-transformed scanning tunneling spectroscopy, Physical Review B 91, 045429 (2015).

Other references

- [24] J. A. Christians, J. S. Manser, and P. V. Kamat, Best Practices in Perovskite Solar Cell Efficiency Measurements. Avoiding the Error of Making Bad Cells Look Good, The Journal of Physical Chemistry Letters 6, 852 (2015).
- [25] Z. Xiao, Y. Yuan, Y. Shao, Q. Wang, Q. Dong, C. Bi, P. Sharma, A. Gruverman, and J. Huang, *Giant Switch-able Photovoltaic Effect in Organometal Trihalide Perovskite Devices*, Nature Materials 14, 193 (2015).

- [26] M. Todorova and J. Neugebauer, Extending the Concept of Defect Chemistry from Semiconductor Physics to Electrochemistry, Physical Review Applied 1, 014001 (2014).
- [27] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, *Finite-size supercell correction schemes for charged defect calculations*, Physical Review B 86, 045112 (2012).
- [28] S. Sharma, J. K. Dewhurst, A. Sanna, and E. K. U. Gross, Bootstrap Approximation for the Exchange-Correlation Kernel of Time-Dependent Density-Functional Theory, Physical Review Letters 107, 186401 (2011).
- [29] F. Le Formal, N. Tétreault, M. Cornuz, T. Moehl, M. Grätzel, and K. Sivula, *Passivating surface states on water splitting hematite photoanodes with alumina over-layers*, Chemical Science 2, 737 (2011).
- [30] M. Ceriotti, G. Bussi, and M. Parrinello, Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat, Physical Review Letters 103, 030603 (2009).
- [31] F. Evangelisti, R. Güttinger, R. Moré, S. Luber, and G. R. Patzke, *Closer to Photosystem II: A Co*₄O₄ *Cubane Catalyst with Flexible Ligand Architecture*, Journal of the American Chemical Society **135**, 18734 (2013).
- [32] H. Liu, M. Schilling, M. Yulikov, S. Luber, and G. R. Patzke, Homogeneous Photochemical Water Oxidation with Cobalt Chloride in Acidic Media, ACS Catalysis 5, 4994 (2015).
- [33] F. Song and X. Hu, Exfoliation of layered double hydroxides for enhanced oxygen evolution catalysis, Nature Communications 5, 4477 (2014).
- [34] K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier, Error Estimates for Solid-State Density-Functional Theory Predictions: An Overview by Means of the Ground-State Elemental Crystals, Critical Reviews in Solid State and Materials Sciences 39, 1 (2014).
- [35] A. Dal Corso, Pseudopotentials periodic table: From H to Pu, Computational Materials Science 95, 337 (2014).
- [36] A. Dal Corso, Projector augmented wave method with spin-orbit coupling: Applications to simple solids and zincblende-type semiconductors, Physical Review B 86, 085135 (2012).
- [37] K. F. Garrity, J. W. Bennett, K. M. Rabe, and D. Vanderbilt, *Pseudopotentials for high-throughput DFT calculations*, Computational Materials Science 81, 446 (2014).
- [38] M. Schlipf and F. Gygi, Optimization algorithm for the generation of ONCV pseudopotentials, Computer Physics Communications 196, 36 (2015).
- [39] M. Topsakal and R. Wentzcovitch, Accurate projected augmented wave (PAW) datasets for rare-earth elements (RE = La - Lu), Computational Materials Science **95**, 263 (2014).
- [40] I. E. Castelli and K. W. Jacobsen, *Designing rules and probabilistic weighting for fast materials discovery in the Perovskite structure*, Modelling and Simulation in Materials Science and Engineering 22, 055007 (2014).
- [41] I. E. Castelli, K. S. Thygesen, and K. W. Jacobsen, Calculated optical absorption of different perovskite phases, Journal of Materials Chemistry A 3, 12343 (2015).
- [42] I. E. Castelli, K. S. Thygesen, and K. W. Jacobsen, Calculated Pourbaix Diagrams of Cubic Perovskites for Water Splitting: Stability Against Corrosion, Topics in Catalysis 57, 265 (2014).
- [43] I. E. Castelli, F. Hüser, M. Pandey, H. Li, K. S. Thygesen, B. Seger, A. Jain, K. A. Persson, G. Ceder, and K. W. Jacobsen, *New Light-Harvesting Materials Using Accurate and Efficient Bandgap Calculations*, Advanced Energy Materials 5, 1400915 (2015).

- [44] I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, and K. W. Jacobsen, *Computational* screening of perovskite metal oxides for optimal solar light capture, Energy & Environmental Science 5, 5814 (2012).
- [45] I. E. Castelli, D. D. Landis, K. S. Thygesen, S. Dahl, I. Chorkendorff, T. F. Jaramillo, and K. W. Jacobsen, *New cubic perovskites for one- and two-photon water splitting using the computational materials repository*, Energy & Environmental Science 5, 9034 (2012).
- [46] C. A. Guido, F. Pietrucci, G. A. Gallet, and W. Andreoni, *The Fate of a Zwitterion in Water from* ab Initio *Molecular Dynamics: Monoethanolamine (MEA)-CO₂*, Journal of Chemical Theory and Computation 9, 28 (2013).
- [47] C. Ma, F. Pietrucci, and W. Andreoni, Capture and Release of CO₂ in Monoethanolamine Aqueous Solutions: New Insights from First-Principles Reaction Dynamics, Journal of Chemical Theory and Computation 11, 3189 (2015).
- [48] G. Puxty, R. Rowland, A. Allport, Q. Yang, M. Bown, R. Burns, M. Maeder, and M. Attalla, *Carbon Dioxide Postcombustion Capture: a Novel Screening Study of the Carbon Dioxide Absorption Performance of 76 Amines*, Environmental Science & Technology 43, 6427 (2009).
- [49] Q. Yang, M. Bown, A. Ali, D. Winkler, G. Puxty, and M. Attalla, A carbon-13 NMR study of carbon dioxide absorption and desorption with aqueous amine solutions, Energy Procedia 1, 955 (2009).
- [50] D. Fernandes, W. Conway, X. Wang, R. Burns, G. Lawrance, M. Maeder, and G. Puxty, *Protonation constants and thermodynamic properties of amines for post-combustion capture of CO*₂, The Journal of Chemical Thermodynamics 54, 183 (2012).
- [51] R. Car and M. Parrinello, Unified Approach for Molecular Dynamics and Density-Functional Theory, Physical Review Letters 55, 2471 (1985).
- [52] A. Laio and M. Parrinello, *Escaping free-energy minima*, Proceedings of the National Academy of Science of the USA 99, 12562 (2002).
- [53] A. Barducci, M. Bonomi, and M. Parrinello, *Metady-namics*, WIREs Computational Molecular Science 1, 826 (2011).
- [54] G. Torrie and J. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation - Umbrella sampling, Journal of Computational Physics 23, 187 (1977).
- [55] X. Wang, W. Conway, D. Fernandes, G. Lawrance, R. Burns, G. Puxty, and M. Maeder, *Kinetics of the Reversible Reaction of CO₂(aq) with Ammonia in Aqueous Solution*, The Journal of Physical Chemistry A **115**, 6405 (2011).
- [56] G. Richner, G. Puxty, A. Carnal, W. Conway, M. Maeder, and P. Pearson, *Thermokinetic properties* and performance evaluation of benzylamine-based solvents for CO₂ capture, Chemical Engineering Journal 264, 230 (2015).
- [57] F. Pietrucci and A. Saitta, Formamide reaction network in gas phase and solution via a unified theoretical approach: towards a reconciliation of different prebiotic scenarios, Proceedings of the National Academy of Science of the USA 112, 15030 (2015).
- [58] Y. Wang, W. D. Richards, S. P. Ong, L. J. Miara, J. C. Kim, Y. Mo, and G. Ceder, *Design principles for solid-state lithium superionic conductors*, Nature Materials 14, 1026 (2015).



- [59] A. Kuhn, J. Koehler, and B. V. Lotsch, *Single-crystal X-ray structure analysis of the superionic conductor Li*₁₀*GeP*₂*S*₁₂, Physical Chemistry Chemical Physics 28, 11620 (2013).
- [60] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials, Journal of Physics: Condensed Matter 21*, 395502 (2009).
- [61] http://www.fiz-karlsruhe.com/icsd.html.
- [62] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quirós, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration, Nucleic Acids Research 40, D420 (2012).
- [63] S. Lebegue, T. Bjorkman, M. Klintenberg, R. M. Nieminen, and O. Eriksson, *Two-Dimensional Materials from Data Filtering and Ab Initio Calculations*, Physical Review X 3, 031002 (2013).
- [64] S. Hall and B. McMahon, eds., International Tables for Crystallography, Definition and exchange of crystallographic data (Springer, The Netherlands, 2005), vol. G of International Tables for Crystallography.
- [65] A. Togo, http://spglib.sourceforge.net.
- [66] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, *Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis,* Computational Materials Science 68, 314 (2013).
- [67] S. Alvarez, A cartography of the van der Waals territories, Dalton Transactions 42, 8617 (2013).
- [68] I. Hamada and M. Otani, Comparative van der Waals density-functional study of graphene on metal surfaces, Physical Review B 82, 153412 (2010).
- [69] R. Sabatini, T. Gorni, and S. de Gironcoli, Nonlocal van der Waals density functional made simple and efficient, Physical Review B 87, 041108 (2013).
- [70] T. Björkman, A. Gulans, A. V. Krasheninnikov, and R. M. Nieminen, van der Waals bonding in layered compounds from advanced density-functional first-principles

calculations, Physical Review Letters **108**, 235502 (2012).

- [71] Y. Li, W. Zhang, M. Morgenstern, and R. Mazzarello, Electronic and Magnetic Properties of Zigzag Graphene Nanoribbons on the (111) Surface of Cu, Ag, and Au, Physical Review Letters 110, 216804 (2013).
- [72] Y.-C. Chen, T. Cao, C. Chen, Z. Pedramrazi, D. Haberer, D. G. de Oteyza, F. R. Fischer, S. G. Louie, and M. F. Crommie, *Molecular bandgap engineering of bottom-up synthesized graphene nanoribbon heterojunctions*, Nature Nanotechnology **10**, 156 (2015).
- [73] E. Runge and E. K. U. Gross, Density-Functional Theory for Time-Dependent Systems, Physical Review Letters 52, 997 (1984).
- [74] E. Tapavicza, I. Tavernelli, and U. Rothlisberger, *Trajectory Surface Hopping within Linear Response Time-Dependent Density-Functional Theory*, Physical Review Letters 98, 023001 (2007).
- [75] I. Tavernelli, B. F. E. Curchod, and U. Rothlisberger, On nonadiabatic coupling vectors in time-dependent density functional theory, The Journal of Chemical Physics 131, 196101 (2009).
- [76] J. Hutter, Excited state nuclear forces from the Tamm-Dancoff approximation to time-dependent density functional theory within the plane wave basis set framework, The Journal of Chemical Physics 118, 3928 (2003).
- [77] B. F. E. Curchod, U. Rothlisberger, and I. Tavernelli, Trajectory-Based Nonadiabatic Dynamics with Time-Dependent Density Functional Theory, ChemPhysChem 14, 1314 (2013).
- [78] P. Ruffieux, J. Cai, N. C. Plumb, L. Patthey, D. Prezzi, A. Ferretti, E. Molinari, X. Feng, K. Müllen, C. A. Pignedoli, and R. Fasel, *Electronic structure of atomically precise graphene nanoribbons*, ACS Nano 6, 6930 (2012).
- [79] M. M. Ugeda, A. J. Bradley, S.-F. Shi, F. H. da Jornada, Y. Zhang, D. Y. Qiu, W. Ruan, S.-K. Mo, Z. Hussain, Z.-X. Shen, F. Wang, S. G. Louie, and M. F. Crommie, *Giant bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor*, Nature Materials 13, 1091 (2014).
- [80] F. J. Giessibl, Advances in atomic force microscopy, Reviews of Modern Physics 75, 949 (2003).
- [81] P. Hapala, G. Kichin, C. Wagner, F. S. Tautz, R. Temirov, and P. Jelínek, *Mechanism of high-resolution STM/AFM imaging with functionalized tips*, Physical Review B **90**, 085421 (2014).

		~
I la uima ustal	Ductor	
norizontai	Project	`

HP3 — Advanced Quantum Simulations

Project leader: Jürg Hutter (UZH)

Participating members: Jürg Hutter (UZH), Matthias Troyer (ETHZ), Philipp Werner (UniFR), Stefan Goedecker (UniBas), Nicola Marzari (EPFL), Joost VandeVondele (ETHZ)

Summary and highlights: We developed efficient continuous time quantum Monte Carlo methods for correlated lattice fermions and found ways to mitigate fermion sign problem in quantum cluster simulations. We implemented an extended dynamical mean field theory (DMFT) method and combined it with a finite temperature Matsubara-axis GW code in order to be able to perform full self-consistent GW+EDMFT calculations. We created a library for electrochemical simulations based on Poisson-(Boltzmann) solvers and implemented analytical gradients for RI-MP2 methods and applied them in molecular dynamics studies of water and solvated electrons. We assembled a library of molecules based on Co water oxidation catalysts for the assessment and verification of density functional theory (DFT) and DMFT methods.

General view of the project

Horizontal Project 3 (HP3) aims at developing new improved methods and algorithms for the description of strongly correlated systems and electrochemical devices. These methods are especially targeted towards the needs in the Vertical Projects 1 and 2.

Major research questions

For a class of strongly correlated systems, that give raise to extremely interesting but scarcely known states of matter, the most commonly used methods for materials based on density functional theory (DFT) fail. Recently, a quantum Monte Carlo (QMC) algorithm that is at the same time free from systematic errors and optimally performing has become available. However, this continuous time approach (CTQMC) has to be further developed to allow to study realistic systems. We aim to predict observables adequate for experimental verification and measures of interest in the field of quantum information and computation using this methodology. The long-term goal is the realization of a scheme for studying strongly correlated states of matter with satisfactorily predictive power.

Another method to address strongly correlated systems is the combination of density functional based *ab initio* calculations with dynamical mean field theory (LDA+DMFT). Replacing the LDA input by GW allows to put the theory (GW+DMFT) onto a firm diagrammatic basis and avoids conceptual problems related to LDA+DMFT. This approach has the potential to become a true *ab initio* method for strongly correlated compounds. At present, self-consistent GW+DMFT calculations have been implemented only for very simple models. The goal is to implement the GW+DMFT approach for real multi-band materials.

The useful work that can be extracted from an electrochemical system is described by the free energy of electrons. Being able to compute, rationalize, and manipulate this quantity in complex chemical systems is important for a sustainable energy future. The capability to accurately compute this quantity in realistic systems is still limited. The challenge is formidable, as it requires to describe finite temperature systems, interfaces between ordered solids and disordered aqueous systems, and the balance between localized and delocalized electrons. It is essential to obtain a qualitatively correct picture for such systems related to research planed in VP2 of this NCCR. The search for density functionals that yield reliable results for a wide range of such systems has lead to methods that are based on an explicit treatment of correlation. Functionals based on the random phase approximation (RPA) have the potential to provide both, high accuracy and general applicability. We are looking for algorithms to calculate the energy at these levels of theory while still providing a computational efficiency that allows application to large and complex systems.

Other approaches to study complex interface systems are related to orbital-dependent functionals and continuum solvation models. Both of these approaches are actively followed in our project.



Long-term goals

Extend CTQMC for long-range Coulomb forces, enabling its use in the simulation of realistic materials, and develop a general scheme to calculate fidelity in CTQMC. Apply the CTQMC scheme with long range forces for the *ab initio* simulation of large scale molecules and realistic materials. Implement the extraction of phenomenological parameters from CTQMC to describe the dynamical responses of strongly correlated and low-temperature phases and apply quantum entanglement and fidelity measures to quantify the correlation and to characterize novel ground state phase diagrams.

Using an imaginary frequency GW code, develop a GW+DMFT self-consistency loop for multi-orbital systems. Extend the selfconsistent GW+DMFT calculations to a larger subspace of orbitals, covering an energy window of ~ 20 eV, while restricting the impurity calculation to a few strongly correlated bands. Implement more advanced forms of RPA with higher order corrections that improve the accuracy of the electronic structure method. Increase the performance of the methods to enable simulations of large systems, and to enable sampling. Add features to RPA and MP2 required for the description of electrochemical systems, such as MD capability and unrestricted calculations.

Apply the computational framework of the resolution-of-the-identity (RI) methods to GW-based methods. Develop a solver for quasi-particle equations defined by orbital-dependent and nonlocal potentials. Transform this into a general scheme for property calculations. Develop a solver for quasi-particle equations with frequency-dependent self-energies. Explore a wide range of approximations for the self-energy (G_0W_0 to Coupled Cluster methods).

The core objective of the continuum models and algorithms developed will be to describe complex electrostatic environments where, due to an implicit solvent, the dielectric constant varies in space, or where mobile ions can shield the charge of the system. Development and implementation of an electrochemical library to solve the generalized Poisson and Poisson-Boltzmann equations.

Achievements

 $SrVO_3$ is a prototypical example of a strongly correlated metal. This widely studied material provides an interesting test case for the more accurate *ab initio* methods. Over the course

of the last year, the Werner group made significant advances in our effort to enable calculations in a fully *ab initio* and fully selfconsistent manner [1]. The finite temperature Matsubara-axis GW code developed by collaborators at Lund university [2] and the multiband EDMFT framework developed within the MARVEL project have been integrated and are ready for production use. A first application focused on SrVO₃ and achieved noteworthy results.

The Troyer group developed efficient continuous time quantum Monte Carlo methods (LCT-QMC) [3, 4] suitable for studying lowtemperature phases of strongly correlated lattice fermions. Combined with a newly discovered design principle of sign-problem free fermionic QMC methods [5], the phase diagram of asymmetric Hubbard model [6] was calculated. A simple, efficient and generic approach to compute an important quantum information concept, fidelity susceptibility using quantum Monte Carlo methods [7] was developed. Based on this approach, they laid out the conceptual and technical foundation of investigating impurity quantum phase transitions and crossovers using fidelity susceptibility [8]. A collaboration of the Werner group with the Troyer group has produced insights into the accuracy of the cRPA downfolding procedure [9], and the basis dependence of the sign problem in cluster-DMFT calculations [10].

The Marzari group has over the last few years developed and tested different Koopmancompliant (KC) functionals. They have demonstrated that KC functionals can be reliable and accurate theoretical tools for the prediction of UPS spectra and can also be exploited successfully in constructing orbital tomography momentum maps, in close agreement with ARPES measurements [11].

In a joint effort, the Goedecker and Marzari groups have worked on self-consistent continuum solvation (SCCS) models. The SCCS model was coupled with linear response time dependent density function perturbation theory (TDDFpT), in order to allow the calculation of electronic excitations and optical spectra of solvated systems [12]. Extensions of the generalized Poisson solver to partially periodic and aperiodic systems, in order to include the effects of the surrounding embedding [13], were developed. They developed and extensively tested a novel solver for both the generalized Poisson and the Poisson-Boltzmann equation. Tests on accuracy and performance of DFT calculations in the presence of an implicit solvent have been carried out [14].

In the VandeVondele group, nuclear gradients for the unrestricted MP2 method (RI-UMP2) have been implemented. The new functionality has been applied in a study of the solvated electron in bulk water. This development allows for the first time to perform finite temperature simulations of redox active species in the condensed phase at this level of theory.

In a joint collaboration of the VandeVondele and Hutter groups, improved performance of post-Hartree-Fock methods was achieved [15]. Analytic gradients and stress tensor have been implemented for MP2 [16]. Applications of the MP2 and RPA codes have been done for several systems. They investigated properties of liquid water using these methods and compared them with a series of local and hybrid density functionals [17].

Milestones

April 2016

- MP2 LSD forces: implementation and first applications (VandeVondele).
- First electrochemistry results based on RPA (energies only) (VandeVondele).
- G₀W₀: implementation and first applications (Hutter).
- Release of stand-alone library for environment effect in first principles simulations (Goedecker, Marzari).
- First applications of Koopmans-compliant functionals to solids (Marzari).
- Implementation of *ab initio* GW+DMFT simulations of correlated materials. Benchmark results for SrVO₃ (Werner).

April 2017

- RPAX energies implemented (VandeVondele).
- Electrochemical system (e.g. solvated electron) consistent MP2 geometries + energies (VandeVondele).
- Benchmark on MARVEL system (cobalt water oxidation complexes) (VandeVondele).
- G₀W₀: reduced scaling and local embedding implementation (Hutter).
- Local self interaction correction: implementation and assessment (Hutter).
- CP2K: assessment of the solid-state system features (Hutter).

- Simulations of electrified interfaces using modified Poisson-Boltzmann approaches; integration of electrochemical library in DFT codes (Goedecker, Marzari).
- Screening in Koopmans' compliant functionals from linear-response theory; band gaps and band structures of solids (Marzari).
- *Ab initio* GW+DMFT results for charge transfer insulators. Inhomogeneous DMFT results for molecules (Werner).

April 2018

- Self interaction correction: general orbital dependent solvers (Hutter).
- Self-consistent GW and Bethe-Salpeter functionality (Hutter).
- CP2K: complete tested and verified pseudopotential/basis set setup (Hutter).
- Band-gaps in solution from solvation model coupled with Koopmans-compliant functionals (Marzari).
- *Ab initio* GW+DMFT for molecules (Werner).

New and ongoing collaborations

This is a list of ongoing collaborations between MARVEL groups where at least one partner is from HP3.

- Marzari group and Pasquarello group (VP2), 2015, *Bandgap of water*.
- VandeVondele group and Hutter group, pre-MARVEL, *Wave function methods in* ab initio *MD*.
- Hutter group and Marzari group, 2015, Orbital-dependent functionals and selfinteraction corrections.
- VandeVondele group and Passerone group (VP2), 2015, *Application of MP2 in material science*.
- VandeVondele group and Ceriotti group (HP4), 2015, *Path-integral sampling methods for MP2 MD*.
- Hutter group, VandeVondele group and Marzari group, 2015, CP2K *binding to AiiDA framework*.
- Hutter group and Pasquarello group (VP2), 2014, *Application of hybrid function- als for liquids and interfaces*.



- Hutter group and Corminboeuf group (VP2), 2015, *Application of* CP2K *in materials simulations*.
- Marzari group and Goedecker group (HP4), 2015, *Continuum solvent models*.
- Werner group and Hutter group, 2015, *Co-Cubane systems as benchmarks for DMFT*.

1 Algorithm development for correlated fermion simulations (Matthias Troyer — ETHZ)

1.1 Research summary

We developed efficient continuous time quantum Monte Carlo methods for correlated lattice fermions; found out ways to mitigate fermion sign problem in quantum cluster simulations; and designed a simple and efficient approach to compute fidelity susceptibility for correlated materials.

1.2 Scientific goals

a) *Short-term* Improve fermionic simulation methods with focus on negative sign problem and suboptimal scalings of the algorithms; explore usage of quantum information measures in material science research.

b) *Long-term* Design and develop efficient and powerful fermionic simulation algorithms to facilitate material discovery.

1.3 Results obtained since the last report

We developed efficient continuous time quantum Monte Carlo methods (LCT-QMC) [3, 4] suitable for studying low-temperature phases of strongly correlated lattice fermions. These methods are free from time-discretization error and enjoy many flexibilities of the continuous time QMC framework [19]. Combined with a newly discovered design principle of signproblem free fermionic QMC methods [5], we mapped out the phase diagram of asymmetric Hubbard model [6] using LCT-QMC method. These results bridge the known limiting cases of SU(2) symmetric Hubbard model and the Falicov-Kimball model that is relevant to rareearth compounds.

We systematically investigated the dependence of average sign on the choice of single-particle basis in the context of quantum cluster problems. We found out the sign problem can be substantially reduced by using a nontrivial single-particle basis which diagonalizes a subset of the intracluster hoppings [10]. This discovery has broad implication for cluster dynamical mean field theory (DMFT) studies of multi-orbital systems and low-dimensional Hubbard models.

Werner group and Troyer group, pre-

• Werner group and Staub group (PSI, PP7),

2015, Charge and spin dynamics in photo-

MARVEL, Solvers for DMFT.

doped antiferromagnetic materials.

We developed a simple, efficient and generic approach to compute an important quantum information concept, fidelity susceptibility using quantum Monte Carlo methods [7]. It provides a fresh look at quantum phases and phase transitions going beyond the conventional Ginzburg-Landau paradigm. Based on this approach, we laid out the conceptual and technical foundation of investigating impurity quantum phase transitions and crossovers using fidelity susceptibility [8].

Constrained random-phase approximation (cRPA) is one of the popular downfolding schemes to derive low-energy effective Hamiltonian of multiorbital systems. We examined its reliability in [9] and found that the violation of the Pauli principle in the cRPA leads to overscreening effects when the interorbital interaction is small. This problem can be overcome by using a variant of the cRPA method which restores the Pauli principle.

1.4 Planned research for next year

Find the optimal single-particle basis in cluster DMFT calculations of low-dimensional Hubbard models Thanks to the improved sign in the Monte Carlo sampling, we may reach lower temperature and larger cluster sizes than the previous study [20]. Furthermore, combine the basis rotation idea [10] with the LCT-QMC methods [3, 4] may provide a quantum impurity solver that performs better than the currently adopted methods [19]. These new technical breakthroughs may shed new light on the pairing mechanism and pseudogap puzzle of the cuprates and related organic superconducting compounds.

Implement fidelity susceptibility calculation in DMFT studies of multi-orbital Hubbard models This quantity provides a unified and numerically inexpensive tool to discover and diagnose various phase transitions and crossovers in realistic materials, such as (orbital-selective) Mott transitions, high-spin to low-spin transitions, Fermi-liquid to non-Fermi-liquid crossovers, spin-freezing crossovers, etc.

1.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* With the group of Philipp Werner at UniFR.

b) *Outside MARVEL* With the group of Silke Biermann at École Polytechnique, Paris.

1.6 Personnel

Matching personnel

- Hiroshi Shinaoka, postdoc, 50%, from May 2014.
- Lei Wang, postdoc, 50%, from May 2014.
- Ye-Hua Liu, postdoc, 50%, from January 2015.
- Mauro Iazzi, postdoc, 50%, from May 2014 to September 2015.

2 Dynamical mean field based methods - GW+DMFT (Philipp Werner - UniFR)

2.1 Research summary

- Multi-orbital impurity solvers;
- frequency dependent interactions;
- extended DMFT;
- self-consistent GW+DMFT.

2.2 Scientific goals

Calculating the electronic structure of strongly correlated materials from first principles is a challenging problem for which reliable and accurate methods still need to be developed. The GW+DMFT scheme [21] should allow a fully self-consistent treatment of screening and correlation effects in the solid. Up to now, self-consistent GW+DMFT calculations have been implemented only for simple one-band models [22, 23][1]. The goal of this project is to implement and test the self-consistent GW+DMFT approach for realistic multi-band materials.

a) Short-term We have merged our recently implemented multi-obital GW+DMFT code with the Matsubara-axis ab initio GW code developed by our collaborators in Lund, and obtained first self-consistent GW+DMFT results for SrVO₃. As will be discussed in the following section, these preliminary results raise questions about the proper interpretation of satellite features in previous ab initio studies. It appears that what was previously interpreted as Hubbard bands in the SrVO₃ spectral function [24] are in fact plasmon satellites resulting from strong nonlocal screening. In the short term, we will extend our analysis to related compounds, such as YTiO₃ and SrMoO₃ to make sure that our *ab initio* GW+DMFT scheme correctly captures known material trends.

b) Long-term The long-term goal is to extend the self-consistent GW+DMFT calculations, which are currently performed only within the strongly correlated low-energy space, to a larger subspace of orbitals, covering an energy window of $\sim 10 - 20$ eV, while restricting the impurity calculation to a few strongly correlated bands. Once the optimal scheme, in terms of accuracy and computational effort, has been identified, and potential double counting issues have been resolved, the method can be applied to a broad range of materials. The hope is that these ab initio GW+DMFT simulations will allow quantitative predictions for the electronic structure of three dimensional strongly correlated compounds, and a proper physical interpretation of satellite features.

2.3 Results obtained since the last report

SrVO₃ has been considered a prototypical example of a strongly correlated metal ever since photoemission and inverse photoemission experiments twenty years ago [25] showed features well outside the low-energy quasiparticle band. These features have been conventionally explained as Hubbard bands, since they can easily be generated in DMFT [26] calculations for Hubbard-type models, when using an appropriate ad hoc value for the local Coulomb repulsion (e.g. [24, 27]).

This widely studied material provides an interesting test case for the more accurate *ab initio* methods that have been developed in recent years. The important methodological innovations are (i) the constrained random phase approximation (cRPA) [28] that removes the ambiguity in choosing the interaction parameters, (ii) extended DMFT (EDMFT) [29] to include the dynamical screening effect due to nonlocal interactions in the low-energy effective model, (iii) the combination with GW [21, 22][1] to take



Figure 1: Spectral function of SrVO₃ calculated for different approximations: DMFT with the static interaction taken from cRPA, the pure GW and the self-consistent GW+EDMFT result.

into account nonlocal self-energy effects, and (iv) impurity solvers for the auxiliary DMFT impurity problem capable of dealing with dynamically screened interactions [30].

Over the course of the last year, we made significant advances in our effort to enable these kinds of calculations in a fully *ab initio* and fully self-consistent manner. The finite temperature Matsubara-axis GW code developed by our collaborator Fredrik Nilsson in the Ferdi Aryasetiawans group at Lund university and the multi-band EDMFT framework we developed within the MARVEL project have been integrated and are ready for production use. As a first application, we have revisited SrVO₃, with noteworthy results.

A simple DMFT description based on the cRPA static interaction reproduces the experimentally observed sidebands and supports their interpretation as Hubbard bands. A DMFT treatment with the proper cRPA retarded local interaction on the other hand, albeit considered a less severe approximation, underestimates the quasiparticle bandwidth [31]. In Fig. 1 we show how the more elaborate GW+EDMFT calculation for this material again vields spectral functions consistent with experimental findings. The additional effects considered in this scheme are a band-widening due to the momentum dependent self-energy, and a reduction of the local interaction due to nonlocal screening. Interestingly, the local interaction is so strongly reduced that Hubbard physics can be excluded as the origin of the sidebands. The new picture which emerges is that of a weakly correlated metal with prominent plasmonic satellites whose positions are related to structures in the fully screened interaction.

For the explanation of both, the reduced static interaction and the formation of satellite features in the spectral functions, we can draw on



Figure 2: Fully screened interaction for SrVO₃ calculated in GW and GW+EDMFT.

previous work [22][1]. Both are a consequence of nonlocal interactions, which are reflected in the EDMFT auxiliary impurity problem by a selfconsistently determined retardation of the local interaction. Marked in Fig. 1 is the static value of this interaction, U(0), which is clearly too small to account for the separation between the peaks at ~ -1.3 eV and ~ 2.7 eV in a Hubbard picture. Instead, those can be explained as plasmonic satellites of the quasiparticle structure, subject to a retardation channel which is evident in the fully screened interaction at ω_p (Fig. 2). The quasiparticle structure is hardly modified compared to GW by the rather small effective on-site interaction.

2.4 Planned research for next year

In the coming months, we will extend the *ab* initio GW+DMFT scheme to low-energy models with weakly and strongly correlated bands. A first target material will be cuprates in a *d*p model description, with one strongly correlated *d*-band (treated within GW+DMFT) and two less correlated *p*-bands (treated only within GW). In such calculations, which treat correlation effects differently in different orbitals, a double-counting issue may appear and will have to be addressed. Once this scheme has been implemented and tested, it will enable the study of a wide range of materials. In particular, we will investigate if the classification of materials such as cuprates or NiO as charge transfer insulators is appropriate. We also plan to address the challenging correlation effects in iron within a *s*-*d* model description.

2.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Within HP3, we are collaborating with the group of Jürg Hutter on the simulation of Co-based water oxydation catalysts. The goal is to apply the GW+DMFT method to these molecular systems. At this

point, we have received the *ab initio* derived input data for these calculations. The goal for the next year is to set up the GW+DMFT framework and test it on these molecular benchmark systems.

There has also been an active collaboration with Hiroshi Shinaoka in the group of Matthias Troyer. This has produced insights into the accuracy and limitations of the cRPA downfolding procedure [9], and the basis dependence of the sign problem in cluster-DMFT calculations [10]. We have also published an *ab initio* study of pyrochlore iridates [18]

We have furthermore started a collaboration with the experimental group of Urs Staub at PSI. The goal of this MARVEL-funded project is to understand the interplay of charge and spin dynamics in photo-doped antiferromagnetic materials. b) *Outside MARVEL* We are collaborating with Denis Golez on an ERC funded effort related to the non-equilibrium extension of DFMT. Here, the goal is to implement a non-equilibrium version of GW+DMFT for simple model systems.

2.6 Personnel

NCCR personnel

• Lewin Boehnke, postdoc, 100%, from December 2014.

Matching personnel

- Denis Golez, postdoc, 100%, from May 2014 to May 2015.
- Hugo Strand, postdoc, 100%, from June 2015.

3 Koopmans' compliant functionals for molecules and solids (Nicola Marzari - EPFL)

3.1 Research summary

Improve the performance and widen the range of applicability of Koopmans-compliant functionals in order to efficiently compute theoretical spectral properties of complex molecules and band structures of extended systems.

3.2 Scientific goals

The interpretation of experimental spectra, such as those obtained with ultraviolet photoemission spectroscopy (UPS) or angular-resolved photoemission spectroscopy (ARPES), often requires theoretical support, due to the complexity of the data. In fact, theoretical prediction can help resolve spectral contribution coming from quasi-degenerate excitations, or to find the correspondence between photoemission peaks and the probability density of the states from which the electrons were emitted. From a theoretical point of view, photoemission spectra have been studied with many-body perturbation theory, time-dependent extension of density functional theory (DFT), density-matrix functional theory or with wave function methods of quantum chemistry. However due to the significant computational requirements of these approaches, and their own limits in term of ultimate accuracy, applications are limited in systems' size and complexity. Recently, in our group, we have introduced Koopmanscompliant (KC) functionals [32, 33] to enforce a generalized criterion of piecewise linearity

with respect to the fractional removal or addition of an electron from any orbital — and not only the highest occupation molecular orbital (HOMO) — in approximate DFT functionals, and to extend to the entire electronic manifold the self-interaction linearization imposed by DFT+Hubbard U [34] methods. The condition of Koopmans' compliance is naturally akin to that of enforcing a correct description of charged excitations, and thus leads to orbital energies that can be interpreted as the quasiparticle excitations measured in photoemission experiments [35].

a) Short-term Our aim is to develop a generalization of the KC functionals to accurately deal with orbital relaxation effects that naturally happen when a photoemission process is considered. In the KC functionals definition, the Koopmans correction is made up by the product of two important terms: (i) an orbitaldensity dependent corrections aiming at describe the addition/removal of an electron in a frozen-orbitals picture, and (ii) a screening factor α which takes into account orbital relaxation effects [33]. In previous calculations, we showed that a unique (identical for all the orbitals) screening factor α , chosen so that the ionization potential (IP) of a neutral molecule exactly matches the electron affinity (EA) of the molecular cation (i.e. enforcing Koopmans' condition), is sufficient to accurately predict not only the IP and EA energies but also molecular photoemission spectra for a variety of molecular systems (Figs. 3 and 4). However, in



Figure 3: *IPs and EAs of 23 organic photo-*voltaic molecules, calculated using the standard *DFT* method (*PBE exchange correlation functional*), *Perdew-Zunger (PZ) self interaction correction and Koopmans-compliant functionals: KI and KIPZ* (from [11]).

order to deal with complex systems, the screening coefficients should be calculated by averaging the nonlocal permittivity in some suitable system — and orbital-dependent fashion. In this respect we are currently working on the development and implementation of an efficient scheme to compute an orbital-dependent screening parameter α , based on the linear response technique of density functional perturbation theory [36].

b) *Long-term* On a long term perspective, we plan to use the KC functionals to study the spectroscopic properties of complex materials, such as organic/semiconductor or semiconductor/semiconductor interfaces, that are beyond the capability of the state-of-the-art approaches usually employed to deal with charged excitations, e.g. the many-body perturbation theory techniques. Moreover, an extension of the KC functionals into the time-dependent domain of DFT will be considered.

3.3 Results obtained since the last report

In the last few years we have developed and implemented different flavors of KC functionals [33] and we have assessed their performance in calculating ionization energies, geometries and atomization energy of standard sets of molecules, showing the accuracy of the method against experimental results and their favorable price/performance ratio compared with correlated wave function based methods and many-body perturbation theory techniques. We have demonstrated that KC functionals can be reliable and accurate theoretical tools for the prediction of UPS spectra and can also be exploited successfully in constructing orbital tomography momentum maps, in close



Figure 4: Ultra-violet photoemission spectrum for fullerene C_{60} calculated using standard DFT method (PBE exchange correlation functional), Perdew-Zunger (PZ) self interaction correction and Koopmans-compliant functionals: KI and KIPZ (from [11]).

agreement with ARPES measurements [11].

3.4 Planned research for next year

We are currently working on the development and implementation of an efficient approach to evaluate an orbital-dependent screening parameter α . A physically sound definition of this parameter has been derived within the framework of the linear response theory. In the next year, we will first test the accuracy of this approach by applying it to the study of selected systems such as organic/inorganic semiconductors, and by comparing the results to those obtained by other approaches such as the GW method. Once validated, the improved version of KC functionals will be applied to study the band structures of extended systems and the photoemission spectra of heterogeneous systems such as organic molecule-semiconductor interfaces.

3.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Collaboration with the group of Alfredo Pasquarello on GW and Koopmans calculations of the band gap of water.

b) *Outside MARVEL* Collaboration with the ARPES experimental community at PSI.

3.6 Personnel

Matching personnel

- Ngoc Linh Nguyen, postdoc, 100%, from May 2014.
- 4 Development of models, algorithms, and codes to include the effects of an aqueous environment in electronic-structure calculations (Stefan Goedecker — UniBas, Nicola Marzari — EPFL)

4.1 Research summary

The study of quantum simulations in the presence of wet environments plays a key role in many scientific fields. If quantum simulation methods are to applied to these processes and systems it becomes imperative to have a qualitatively and quantitatively correct modeling of the effects of the environment. The present research will considerably extend the range of applications of such methods.

4.2 Research question

The computational study of chemical reactions in complex, wet environments is critical for applications in many fields, and of crossdisciplinary interest to physics, chemistry, materials science, chemical engineering, and biology. It is often essential to study chemical reactions in the presence of an applied electrochemical potential, establishing the correct relation between charge and potential, and taking into account the complex electrostatic screening coming from the solvent and the electrolytes. Simulations of various systems in an aqueous, or a generic wet environment are of great interest, but pose serious problems to standard fully atomistic approaches, due to their intrinsic multiscale nature. For this reasons, new models and algorithms are needed, which allow to reduce the complexity of the environment, whitout affecting its effects on the quantum-mechanical system of interest. Continuum approaches can provide an answer to the problem, especially when combined with efficient and robust numerical algoritms into a computational library able to handle the resulting generalized Poisson or Poisson-Boltzmann equations, without affecting the performances of electronic structures codes.

4.3 Scientific goals

a) *Short-term* Integration of the developed electrochemical library in the existing electronic structures codes used within MARVEL. Test and validation of overall performances on

atomistic simulations from vacuum to wet environments.

b) Long-term The core objective of the continuum models and algorithms developed in the electrochemical library will be to describe complex electrostatic environments where, due to an implicit solvent, the dielectric constant varies in space, or where mobile ions can shield the charge or multipoles of the system of interest; it will allow to embed the quantum simulation engines into a robust and efficient Poisson-Boltzmann solver that has been extensively verified and validated. The Poisson Boltzmann equation is more difficult to solve than the generalized Poisson equation due to the fact that it leads to a nonlinear problem and due to the possibly strong variation of the ion concentration with respect to the electric potential. Hence in contrast to the standard or generalized Poisson equation there is no general consensus on what flavor of the Poisson Boltzmann equation is most appropriate for various problems. As a consequence the Poisson Boltzmann library should allow to solve various flavors of this equation. It should also be possible for users to add easily new variants to the library.

4.4 Results obtained since the last report

This is a new project started in late 2014 and supported by PASC, the Swiss Platform for Advanced Scientific Computing. It is reported here because it's a significant matching effort to the goals of MARVEL, and because it will underpin validation efforts with other MARVEL projects (e.g. in the groups of Pasquarello and VandeVondele) and design efforts in the field of photo-electrocatalysis.

The starting framework of the project is a novel continuum model, based on a reformulation of the variational approach of Fattebert and Gygi [37] and extended to model nonelectrostatic effects in terms of self-consistent contributions. The resulting self-consistent continuum solvation (SCCS) model [38] was carefully parameterized and tested on neutral and charged organic molecules in aqueResults — HP3



Figure 5: Comparison of the absorption spectra of the 4-AP molecule in vacuum, water, and carbon-tetrachloride, using the Gaussian 09 code and the PCM model on the one hand and the turboTDDFT code and the SCCS model on the other hand. The hybrid B3LYP density functional was used for both calculations.

ous solutions, showing results in perfect agreement with similar approaches (as the polarizable continuum model (PCM) of Tomasi and coworkers [39]) and in good agreement with experimental data.

As part of the efforts of this project, the SCCS model was coupled with linear-response timedependent density function perturbation theory (TDDFpT), in order to allow the calculation of electronic excitations and optical spectra of solvated systems. The resulting framework [12] provides results in very good agreement with the PCM approach, as shown for example in Fig. 5 for the case of a small organic molecule in polar and apolar solutions.

As the original formulation of the generalized Poisson solver is based on the use of fast Fourier transforms and assumes a periodically repeated system, we extended some of the most common approaches to correct for periodic boundary conditions in partially periodic and aperiodic system, in order to include the effects of the surrounding embedding [13].

As a parallel and coordinated effort, we developed and extensively tested a novel solver for both the generalized Poisson and the Poisson-Boltzmann equations. A preconditioned conjugate gradient (PCG) method has been implemented for the generalized Poisson equation and the linear regime of the Poisson-Boltzmann, allowing to solve iteratively the minimization problem with some ten iterations of a ordinary Poisson solver. Being challenging to define an action integral for all flavors of the nonlinear Poisson-Boltzmann equation and then apply a minimization scheme, a self-consistent procedure has been imple-



Figure 6: Euclidean norm of the residual vector (top graphs) and accuracy of the numerical solution (lower graphs) for SC (blue circles) and PCG (black squares) solvers with free, surface and periodic boundary conditions.

mented for the nonlinear Poisson-Boltzmann problem. Both solvers exhibit very high accuracy and parallel efficiency, and intrinsically allow for the treatment of different boundary conditions, as for example surface systems. The solver has been integrated into the BIGDFT and Quantum-ESPRESSO electronicstructure packages and will be released as an independent program, suitable for integration in other codes.

Fig. 6 shows solver performances. The top graphs report the residual norm as function of the iteration number, whilst lower panels present the output accuracy. Analytic test functions have been used, in particular a Gaussian potential and an error function for the dielectric cavity which mimics an atomistic cavity where the dielectric function takes a value of one in the neighborhood of the atoms and typically a higher value in the region of the implicit solvent. The PCG solver (black squares) exhibits a fast convergence reaching an accuracy of ~ 10^{-10} with some ten iterations. The figure also reports results from a self-consistent scheme (SC, blue circles).

To further validate the electrochemical library once integrated in *ab initio* packages, electrostatic solvation energies and structure optimizations have been investigated for a small set of neutral molecules and a TiO₂ slab. Table 1 reports electrostatic solvation energies ΔG^{el} obtained both with the rigid and the charge-dependent cavity under free boundary conditions. As reference, in Table 1 self-consistent continuum solvation (sccs) model calculations from [38] together with PCM calculations have been reported. The table shows a good agreement for both approaches, validating the whole electrochemical library.



Figure 7: *TiO*₂ *surface in contact with water: isosurfaces of the polarization charge in the implicit dielectric medium.*

	PCM	$sccs_{QE}^{P}$	$\operatorname{sccs}^F_{\operatorname{BigDFT}}$
NH ₃	-6.65	-5.39	-5.35
H_2O	-8.98	-8.21	-8.23
CH_4	-0.61	-0.68	-0.63
CH ₃ OH	-6.78	-5.89	-5.83
CH ₃ NH ₂	-4.51	-4.53	-4.45
CH ₃ CONH ₂	-12.53	-11.87	-11.87

Table 1: Electrostatic solvation energies ΔG^{el} (in *kcal/mol*).

Numerous processes of practical interest involve surfaces in contact with neutral or ionic solvents, leading to an induced polarization charge of the dielectric medium or an electric double layer. The BIGDFT package allows to use exact surface boundary conditions avoiding spurious interactions in the direction orthogonal to the surface exposed to the wet environments. To show a further application of the solvation library in BIGDFT, a TiO₂ surface in contact with pure water has been simulated. The full DFT simulation in presence of the solvent has been initialized starting from its relaxed state in vacuum. Fig. 7 shows the TiO₂ wet surface as well as isosurfaces of the polarization charge density.

Furthermore extensive tests on accuracy and performances of DFT calculations in the pres-

ence of an implicit solvent have been carried out. All results have been submitted and accepted as publication in *The Journal of Chemical Physics* [14]. This work comes from the close and intense collaboration between the two groups.

4.5 Planned research for next year

The integration of these various software developments will allow to tackle new challenging physical problems. The electrochemical library will be highly optimized for massively parallel computer architectures using both traditional CPUs as well as GPUs. Extensive testing, validation and verification of all the component of the new library will be performed. Further development, validation and testing of the Poisson-Boltzmann solver for different electrostatic environments will be done.

4.6 Synergies with other computational and experimental efforts

a) *Within MARVEL* The study of liquid-solid interfaces is a key theme of VP2 and HP3, and (electrified) liquid-solid interfaces play a crucial role in batteries, fuel cells, dye-sensitized solar cells, and photo-electro-catalysis; thus we believe this project will support manyfold future activities.

b) *Outside MARVEL* the model and the optimized algorithms are delivered as an opensource library that can be easily integrated into other codes.

4.7 Personnel

Matching personnel

- Giuseppe Fisicaro, postdoc, 100%, from January 2015.
- Olivero Andreussi, postdoc, 100%, from January 2015.

5 Electrochemistry beyond GGA DFT (Joost VandeVondele — ETHZ)

5.1 Research summary

- Massively parallel implementation of spin-unrestricted MP2 forces.
- Solvated electron.
- Advanced Random Phase Approximations.
- Electrochemical properties of novel materials.

5.2 Scientific goals

- a) Short-term
 - Application of the novel spin-unrestricted MP2 forces implementation to the prototypical and challenging electronchemical species — the hydrated electron in the bulk water — via molecular dynamics

simulation.

• Parallel implementation and benchmarking of the advanced random phase approximation (RPA) methods in CP2K to treat systems with complicated electronic structure, including those with small or zero band gap.

b) *Long-term* Electrochemical properties of novel materials obtained from accurate manybody electronic structure methods beyond DFT and molecular dynamics.

5.3 Results obtained since the last report

The first milestone of the group was the generalization of the implementation of resolutionof-identity (RI) MP2 forces [16] for the spinunrestricted (RI-UMP2) calculations in the condensed phase in CP2 κ [40]. This objective has been fully achieved, allowing for the first time to perform finite temperature simulations of redox active species in the condensed phase. This allows us to extend the leading position our group has in the simulation of aqueous systems from the first simulation of bulk liquid water at the MP2 level [41][17] to electrochemical systems.

To summarize the approach followed in the past ten months, we base our formalism on the geometric derivative of the unrestricted RI-MP2 energy as:

$$E^{(2)^{(x)}} = 4 \sum_{Q}^{AUX} \sum_{\mu\nu}^{AO} \Gamma^{Q}_{\mu\nu} (\mu\nu|Q)^{x} - 2 \sum_{PQ}^{AUX} \Gamma^{PQ} (P|Q)^{x} + (1) + 2 \sum_{pq}^{SO} [P^{(2)}_{pq} F^{x}_{pq} - W^{(2)}_{pq} S^{x}_{pq}]$$

where μ , ν are basis functions, p, q are spin orbitals, and

$$\sum_{p}^{SO} = \sum_{p^{\alpha}}^{SO^{\alpha}} + \sum_{p^{\beta}}^{SO^{\mu}}$$
(2)

 F_{pq}^{x} and S_{pq}^{x} are the derivatives of the Fock and overlap matrixes, respectively. Specific RI-MP2 density matrices $P_{p^{\alpha}q^{\alpha}}^{(2)}$, $P_{p^{\beta}q^{\beta}}^{(2)}$ and $W_{p^{\alpha}q^{\alpha}}^{(2)}$, $W_{p^{\beta}q^{\beta}}^{(2)}$ are constructed from electron-repulsion integrals (ERI) between spin-orbitals of similar and different spin, which makes the spinunrestricted generalization nontrivial.

Within RI-approximation 4-center ERI in the MP2 theory are approximated by 2- and 3-ones

7000 6000 5000 wall time, 4000 Computational 3000 2000 1000 0 512 1024 1536 1792 2304 0 256 769 1280 2048 Num

Figure 8: Parallel performance of RI-UMP2: computational time for one molecular dynamics step for solvated electron in 31-water (blue) and 63-water (red) periodic boxes. One compute node (of the CSCS Piz Daint supercomputer) consists of 8 CPUs and 1 GPU. These curves demonstrate the good scalability of the code, to top supercomputer sizes, while also demonstrating that 63-water molecules is not fully in the $O(N^5)$ regime, as could could be expected for MP2.

using auxiliary functions [42]:

$$(i^{\sigma_1}a^{\sigma_1}|j^{\sigma_2}b^{\sigma_2}) \approx \sum_{PQ} (i^{\sigma_1}a^{\sigma_1}|P)(P|Q)^{-1}(Q|j^{\sigma_2}b^{\sigma_2})$$
(3)

where $\sigma_1 = \sigma_2 = \alpha$, $\sigma_1 = \sigma_2 = \beta$, or $\sigma_1 = \alpha$, $\sigma_2 = \beta$; *i*, *j* are occupied and *a*, *b* are virtual indices. The 3-center ERI are computed numerically in the efficient Gaussian and plane waves approximation (GPW) employing the reciprocal space [43, 44, 45, 41, 46][15].

The implementation is based on the loop over pairs of occupied spin-orbitals, in which ERI and their contributions to the terms in Eq. (1)are calculated on the fly. Parallelization strategy is based on dividing the range of auxiliary basis functions P, Q, ... and virtual orbitals *a*, *b*, ... between processors. The formal scaling on the algorithms is N^5 although for many important systems this is not achieved. The program implementation involves both MPI and OMP parallelization as well as GPUacceleration for the big matrix multiplications. Parallel performance for the solvated electron in periodic boxes containing 31 and 63 water molecules are illustrated in Fig. 8. Thus, using thousands of processors, it is possible to perform one molecular dynamics step for these relevant systems in a few minutes, or dozens of minutes, respectively.

One-particle density matrices $P_{p^{\alpha}q^{\alpha}}^{(2)}$, $P_{p^{\beta}q^{\beta}}^{(2)}$ needed to calculate the forces allow for performing population analysis and calculating the spin density. These are important properties for the system at hand. In Fig. 9 spin densities for the solvated electron in a periodic box containing 63 water molecules are plotted at the spin-unrestricted GGA DFT (BLYP) and





Figure 9: Isosurfaces of spin densities of the solvated electron (isovalue = 0.001 Bohr). Top: BLYP, bottom: MP2, for the same configuration of bulk liquid water.

RI-MP2 levels. It can be clearly seen that with DFT the electron is much more delocalized, exhibiting the infamous delocalization error [47], while with MP2 the spin density is mostly confined to the cavity. This illustrates the ability of MP2 to give a qualitatively accurate description of anionic states and radicals,

necessary for the reliable computation of electrochemical properties, and thus supporting our research plan.

5.4 Planned research for next year

- Perform MD study of the bulk hydrated electron at RI-UMP2 level.
- Implement spin-restricted and unrestricted AXK-RPA [48] (and possibly other variants of RPA) in CP2K.
- Benchmark their performance on cobaltcontaining water-oxidation complexes with complex electronic structure as part of the recently proposed MARVEL benchmark.
- 5.5 Synergies with other computational and experimental efforts
- a) Within MARVEL
 - MP2 and RPA functionality is made available to all MARVEL partners as part of CP2K, users introduced to the functionality (Passerone Group).
 - Development of MP2 and advanced RPA, close collaboration on CP2K development (Hutter group).
 - Development of advanced sampling algorithms combining DFT and MP2 for NQE (Ceriotti group).
 - CP2K developments to facilitate interfacing with AiiDA and other community projects (Marzari Group).

b) *Outside MARVEL* Collaborations on aqueous electrochemistry are ongoing with M. Sprik (Cambridge) and J. Cheng (Aberdeen).

5.6 Personnel

NCCR personnel

• Vladimir Rybkin, postdoc, 100%, from December 2014.

6 Density functional theory and beyond (Jürg Hutter — UZH)

6.1 Research summary

We develop methods and algorithms for electronic structure calculations that go beyond local density functional theory. Based on the Gaussian and plane waves framework we have implemented MP2, RPA and GW methods. We will in the future explore reduced scaling algorithms and methods for enhanced accuracy.

6.2 Scientific goals

The search for density functionals that yield reliable results for a wide range of systems, from molecules to solids, metals, and semiconductors with strong and weak interactions has led to methods that include an explicit treatment of electron correlation. Functionals based on these principles have the potential to provide both, high accuracy and general applicability. We are looking for algorithms to calculate the energy at these levels of theory while still providing a computational efficiency that allows applications to large systems and a sampling of many configurations. The goal is to construct software programs that implement highlevel correlation methods into CP2K [40] and that can use modern massively parallel computer architectures.

a) *Short-term* We are working towards reduced scaling algorithms for RPA correlation energy calculations. The goal is to make it possible to calculate an order of magnitude larger systems using this method. Similar methods will also be applied to GW methods and we want to make it possible to sample spectroscopic signals for larger systems.

b) *Long-term* Development and implementation of high-level methods for electron correlation using efficient algorithms for large systems. The methods should be reliable and predictive for complex systems, like electrochemical interfaces and defect structures. We are especially seeking methods that still allow for an extensive sampling of phase space. The computer implementations will take advantage of modern hardware, massively parallel computers with heterogenous node structures.

6.3 Results obtained since the last report

We have worked on auxiliary density matrix methods (ADMMs) [49] to speed up Hartree-Fock exchange (HFX) calculations. We have implemented three ADMMs with charge constrains in CP2K. The implementation is based on previous work in our group and can handle HFX in total energy calculations as well as molecular dynamics simulations with use of analytical forces and the analytical stress tensor. These new ADMMs can also be used together with the post-Hartree-Fock methods 2nd-order Moller–Plesset (MP2) and randomphase approximation (RPA) [15].

Analytic gradients and stress tensor have been implemented for MP2. Applications of the MP2 and RPA codes have been done for several



Figure 10: Execution time for a G_0W_0 calculation of water molecules on Piz Daint in a cc-TZVP basis.

systems. We specifically investigated properties of liquid water using these methods and compared them with a series of local and hybrid density functionals [17, 16].

The main project has been the implementation of G_0W_0 and eigenvalue-self-consistent GW (evGW) in the CP2K code. These methods are based on the previous implementation of the resolution-of-the-identity (RI) approach together with the Gaussian and plane-waves method for computing the correlation energy in RPA [45]. The G_0W_0 and evGW implementation displays high efficiency and good parallel scalability which is the key for the study of large systems containing hundreds of atoms using accurate basis sets (Fig. 10).

We applied the implementation to study the HOMO-LUMO gap of organic semiconductors with evGW. Specifically, we investigated acenes (Fig. 11) and in collaboration with L. Talirz and C. A. Pignedoli (Empa), anthenes containing hundreds of atoms. The anthenes have been synthesized recently on a substrate and are promising candidates for future electronic devices.

6.4 Planned research for next year

We have started to implement a method to compute the RPA correlation energy in $\mathcal{O}(N^3)$ operations compared to the $\mathcal{O}(N^4)$ computational cost in the current state-of-the-art implementation. The $\mathcal{O}(N^3)$ -RPA method relies on the RI with a local metric, imaginary time and imaginary frequency grids and sparse matrix multiplications employing the DBCSR library [50]. We already found agreement between the RPA correlation energy in the $\mathcal{O}(N^3)$ method and the state-of-the-art implementation. Currently, we are investigating the performance of the parallel $\mathcal{O}(N^3)$ -RPA implementation.

The primary focus is on the implementation of the $O(N^3)$ G_0W_0 and evGW method for large-scale applications. As basis, we plan to



Figure 11: evGW HOMO-LUMO gap of oligoacenes for the PBE0 (blue) and the tuned CAM-B3LYP starting point (green) with open-shell (broken symmetry, circles) and closed-shell configuration (triangulars). We find that the evGW gaps are only weakly dependent on the starting DFT functional, but strongly dependent on restricting the DFT calculation to closed shell or not: the deviation of the evGW HOMO-LUMO gaps between open-shell and closed-shell can be as large as 2 eV (for 10-acene and 11-acene).

make use of the $\mathcal{O}(N^3)$ -RPA method, which is for itself already a first milestone. We will then focus on quasiparticle-self-consistent GW (QSGW), also at $\mathcal{O}(N^3)$ numerical cost. This will allow us to study systems by means of QSGW without the dependence on the underlying exchange-correlation functional in DFT. An open question that we will tackle is the issue of converging GW calculations with respect to the basis set size. We will explore dual basis set methods where a specific additional basis set only for virtual orbital is used.

6.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* We do have close collaborations on the development of electron correlation methods with Joost VandeVondele (ETHZ). We anticipate collaborations on orbital-dependent functionals in general with the group of Nicola Marzari (EPFL) and on the combination of GW with DMFT with Philipp Werner (UniFR). There are synergies from the applications of our developments within the VP2 project of MARVEL.

b) *Outside MARVEL* Our developments depend strongly on the basic high-performance

computing modules available within CP2K. There is a co-design effort, supported by the PASC initiative, to improve these basic libraries. Besides the groups mentioned above, this efforts include contributions from CSCS, Cray, Intel, and Nvidia.

6.6 Personnel

NCCR personnel

- Jan Wilhelm, PhD student, 100%, from November 2014.
- Tiziano Müller, PhD student, 100%, from April 2015.

Matching personnel

- Mauro DelBen, postdoc, 100%, from January 2015.
- Patrick Seewald, PhD student, 100%, from May 2015.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- L. Huang, T. Ayral, S. Biermann, and P. Werner, *Extended dynamical mean-field study of the Hubbard model with long-range interactions*, Physical Review B **90**, 195114 (2014).
- [2] P. Werner, R. Sakuma, F. Nilsson, and F. Aryasetiawan, *Dynamical screening in La₂CuO₄*, Physical Review B **91**, 125142 (2015).
- [3] M. Iazzi and M. Troyer, Efficient continuous-time quantum Monte Carlo algorithm for fermionic lattice models, Physical Review B 91, 241118 (2015).
- [4] L. Wang, M. Iazzi, P. Corboz, and M. Troyer, Efficient continuous-time quantum Monte Carlo method for the ground state of correlated fermions, Physical Review B 91, 235151 (2015).
- [5] L. Wang, Y.-H. Liu, M. Iazzi, M. Troyer, and G. Harcos, *Split Orthogonal Group: A Guiding Principle for Sign-Problem-Free Fermionic Simulations*, Physical Review Letters **115**, 250601 (2015).
- [6] Y.-H. Liu and L. Wang, Quantum Monte Carlo study of mass-imbalanced Hubbard models, Physical Review B 92, 235129 (2015).
- [7] L. Wang, Y.-H. Liu, J. Imriška, P. N. Ma, and M. Troyer, Fidelity Susceptibility Made Simple: A Unified Quantum Monte Carlo Approach, Physical Review X 5, 031007 (2015).
- [8] L. Wang, H. Shinaoka, and M. Troyer, Fidelity Susceptibility Perspective on the Kondo Effect and Impurity Quantum Phase Transitions, Physical Review Letters 115, 236601 (2015).
- [9] H. Shinaoka, M. Troyer, and P. Werner, Accuracy of downfolding based on the constrained random-phase approximation, Physical Review B 91, 245156 (2015).
- [10] H. Shinaoka, Y. Nomura, S. Biermann, M. Troyer, and P. Werner, Negative sign problem in continuous-time quantum Monte Carlo: optimal choice of single-particle basis for impurity problems, Physical Review B 92, 195126 (2015).



- [11] N. L. Nguyen, G. Borghi, A. Ferretti, I. Dabo, and N. Marzari, First-Principles Photoemission Spectroscopy and Orbital Tomography in Molecules from Koopmans-Compliant Functionals, Physical Review Letters 114, 166405 (2015).
- [12] I. Timrov, O. Andreussi, A. Biancardi, N. Marzari, and S. Baroni, Self-consistent continuum solvation for optical absorption of complex molecular systems in solution, The Journal of Chemical Physics 142, 034111 (2015).
 - [13] O. Andreussi and N. Marzari, Electrostatics of solvated systems in periodic boundary conditions, Physical Review B 90, 245101 (2014).
- [14] G. Fisicaro, L. Genovese, O. Andreussi, N. Marzari, and S. Goedecker, A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments, The Journal of Chemical Physics 144, 014103 (2016).
 - [15] M. Del Ben, O. Schütt, T. Wentz, P. Messmer, J. Hutter, and J. VandeVondele, *Enabling simulation at the fifth rung of DFT: Large scale RPA calculations with excellent time to solution*, Computer Physics Communications 187, 120 (2015).
- [16] M. Del Ben, J. Hutter, and J. VandeVondele, Forces and stress in second order Møller-Plesset perturbation theory for condensed phase systems within the resolution-ofidentity Gaussian and plane waves approach, The Journal of Chemical Physics 143, 102803 (2015).
- [17] M. Del Ben, J. Hutter, and J. VandeVondele, Probing the structural and dynamical properties of liquid water with models including non-local electron correlation, The Journal of Chemical Physics 143, 054506 (2015).
- [18] H. Shinaoka, S. Hoshino, M. Troyer, and P. Werner, *Phase Diagram of Pyrochlore Iridates: All-in–All-out Magnetic Ordering and Non-Fermi-Liquid Properties*, Physical Review Letters 115, 156401 (2015).

Other references

- [19] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Continuous-time Monte Carlo methods for quantum impurity models*, Reviews of Modern Physics 83, 349 (2011).
- [20] P. Werner, E. Gull, O. Parcollet, and A. J. Millis, Momentum-selective metal-insulator transition in the two-dimensional Hubbard model: An 8-site dynamical cluster approximation study, Physical Review B 80, 045120 (2009).
- [21] S. Biermann, F. Aryasetiawan, and A. Georges, First-Principles Approach to the Electronic Structure of Strongly Correlated Systems: Combining the GW Approximation and Dynamical Mean-Field Theory, Physical Review Letters 90, 086402 (2003).
- [22] T. Ayral, S. Biermann, and P. Werner, Screening and nonlocal correlations in the extended Hubbard model from self-consistent combined GW and dynamical mean field theory, Physical Review B 87, 125149 (2013).
- [23] P. Hansmann, T. Ayral, L. Vaugier, P. Werner, and S. Biermann, Long-Range Coulomb Interactions in Surface Systems: A First-Principles Description within Self-Consistently Combined GW and Dynamical Mean-Field Theory, Physical Review Letters 110, 166401 (2013).
- [24] E. Pavarini, S. Biermann, A. Poteryaev, A. I. Lichtenstein, A. Georges, and O. K. Andersen, *Mott Transition* and Suppression of Orbital Fluctuations in Orthorhombic 3d¹ Perovskites, Physical Review Letters **92**, 176403 (2004).
- [25] K. Morikawa, T. Mizokawa, K. Kobayashi, A. Fujimori, H. Eisaki, S. Uchida, F. Iga, and Y. Nishihara, Spectral weight transfer and mass renormalization

in Mott-Hubbard systems SrVO₃ and CaVO₃: Influence of long-range Coulomb interaction, Physical Review B **52**, 13711 (1995).

- [26] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Reviews of Modern Physics 68, 13 (1996).
- [27] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and O. K. Andersen, Dynamical mean-field theory using Wannier functions: A flexible route to electronic structure calculations of strongly correlated materials, Physical Review B 74, 125120 (2006).
- [28] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, *Frequency*dependent local interactions and low-energy effective models from electronic structure calculations, Physical Review B **70**, 195104 (2004).
- [29] P. Sun and G. Kotliar, Extended dynamical mean-field theory and GW method, Physical Review B 66, 085120 (2002).
- [30] P. Werner and A. J. Millis, *Dynamical Screening in Correlated Electron Materials*, Physical Review Letters 104, 146401 (2010).
- [31] R. Sakuma, P. Werner, and F. Aryasetiawan, *Electronic structure of SrVO₃ within GW+DMFT*, Physical Review B 88, 235110 (2013).
- [32] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, *Koopmans' condition for densityfunctional theory*, Physical Review B 82, 115121 (2010).
- [33] G. Borghi, A. Ferretti, N. L. Nguyen, I. Dabo, and N. Marzari, Koopmans-compliant functionals and their performance against reference molecular data, Physical Review B 90, 075135 (2014).
- [34] M. Cococcioni and S. de Gironcoli, Linear response approach to the calculation of the effective interaction parameters in the LDA + U method, Physical Review B 71, 035105 (2005).
- [35] A. Ferretti, I. Dabo, M. Cococcioni, and N. Marzari, Bridging density-functional and many-body perturbation theory: Orbital-density dependence in electronic-structure functionals, Physical Review B 89, 195134 (2014).
- [36] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, *Phonons and related crystal properties from density-functional perturbation theory*, Reviews of Modern Physics 73, 515 (2001).
- [37] J.-L. Fattebert and F. Gygi, Density Functional Theory for Efficient Ab Initio Molecular Dynamics Simulations in Solution, Journal of Computational Chemistry 23, 662 (2002).
- [38] O. Andreussi, I. Dabo, and N. Marzari, *Revised self-consistent continuum solvation in electronic-structure cal-culations*, The Journal of Chemical Physics 136, 064102 (2012).
- [39] J. Tomasi, B. Mennucci, and R. Cammi, *Quantum mechanical continuum solvation models*, Chemical Reviews 105, 2999 (2005).
- [40] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. Vande-Vondele, CP2K: atomistic simulations of condensed matter systems, WIREs Computational Molecular Science 4, 15 (2014).
- [41] M. Del Ben, M. Schönherr, J. Hutter, and J. Vande-Vondele, Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory, The Journal of Physical Chemistry Letters 4, 3753 (2013).
- [42] O. Vahtras, J. Almlöf, and M. W. Feyereisen, *Integral approximations for LCAO-SCF calculations*, Physical Review Letters 213, 514 (1993).

- [43] G. Lippert, J. Hutter, and M. Parrinello, A hybrid Gaussian and plane wave density functional scheme, Molecular Physics 92, 477 (1997).
- [44] M. Del Ben, J. Hutter, and J. VandeVondele, Second-Order Moller-Plesset Perturbation Theory in the Condensed Phase: An Efficient and Massively Parallel Gaussian and Plane Waves Approach, Journal of Chemical Theory and Computation 8, 4177 (2012).
- [45] M. Del Ben, J. Hutter, and J. VandeVondele, *Electron Correlation in the Condensed Phase from a Resolution of Identity Approach Based on the Gaussian and Plane Waves Scheme, Journal of Chemical Theory and Computation* 9, 2654 (2013).
- [46] M. Del Ben, M. Schönherr, J. Hutter, and J. Vande-Vondele, Correction to "Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory" (4, 3753 (2013)), The Journal of Physical Chemistry Letters 5, 3066 (2014).

- [47] P. Mori-Sanchez, A. J. Cohen, and W. Yang, *Localiza*tion and Delocalization Errors in Density Functional Theory and Implications for Band-Gap Prediction, Physical Review Letters **100**, 146401 (2008).
- [48] F. Furche, Developing the random phase approximation into a practical post-Kohn-Sham correlation model, The Journal of Chemical Physics 129, 114105 (2008).
- [49] M. Guidon, J. Hutter, and J. VandeVondele, Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations, Journal of Chemical Theory and Computation 6, 2348 (2010).
- [50] U. Borštnik, J. VandeVondele, V. Weber, and J. Hutter, Sparse matrix multiplication: The distributed blockcompressed sparse row library, Parallel Computing 40, 47 (2014).

4 Horizontal Project



Project leader: Stefan Goedecker (UniBas)

Participating members: Stefan Goedecker (UniBas), Michele Parrinello (USI and ETHZ), Berend Smit (EPFL), Michele Ceriotti (EPFL), Anatole von Lilienfeld (UniBas), Ursula Röthlisberger (EPFL)

Summary and highlights: Significant progress has been made in predicting new materials, characterizing them by properties of their energy landscape, machine learning for predicting their properties and the understanding of the growth and nucleation of materials. All these topics are closely related and require powerful methods to sampling the potential and free energy surface. All these methods profit from certain descriptors that are also called fingerprints or collective variables that can quantity the similarity or dissimilarity of different structures.

General view of the project

Crystallographic data banks such as ICSD, OQMD, Materials Project, or AFLOW contain several hundred thousand structures. The majority of them are of experimental origin, but an increasing number is now also originating from theoretical predictions. To exploit such large databases one has to be able to navigate in them to establish correlations between certain descriptors of these materials and their properties. As a first requirement a descriptor has to allow to determine whether two structures are identical or not. This seems to be a simple problem, but it actually turns out to be rather difficult one for which, at least until very recently, no satisfactory solution has been found. In a second step, a descriptor can be used as an input into machine learning schemes to predict directly properties without the costly numerical solution of the electronic Schrödinger equation. Such machine learning based schemes could open up entirely new possibilities allowing to simulate very large systems at a small fraction of the cost of a density functional calculation. Another very important question is whether a theoretically predicted structure can be synthesized. Since one can find in a computer simulation typically a very large number of structures, presumably only a small fraction of these materials can be synthesized. In addition, varying environments during nucleation and growth can lead to different structures. Understanding better the nucleation and growth process is therefore of a very important link between theoretical structure prediction and technological applications.

Major research questions

The potential energy surface, which gives the energy of a system as a function of its atomic coordinates, is the central quantity to understand the structure and dynamics of any condensed matter system. The global minimum gives the structural ground state, other local minima metastable states and the properties of the saddle points determine the speed of chemical reactions as well as other dynamical processes. For a system containing N atoms the potential energy surface is a 3N dimensional function. This high dimensionality makes the exploration of the potential energy surface extremely difficult. The main goal of the Horizontal Project 4 (HP4) is to develop methods to sample and to understand this potential energy surface for various systems.

The group of Goedecker focuses on methods for finding low-energy local minima as well as reaction pathways on this potential energy surface. In addition it works on the development and use of fingerprints. The group of Parrinello focuses on the free energy which is typically not given as a function of all the 3N coordinates of the system but as a function of a smaller subset of order parameters. The group of Parrinello is also interested in kinetic processes occurring on the free energy surface. In particular it works on methods to accelerate processes occurring on the free energy surface. The group of von Lilienfeld focuses on the chemical compound space, i.e. is interested in how the potential energy surfaces and resulting properties change if the identity of atoms changes (this is in strong synergy with HP5). The group of Ceriotti is also interested in characterizing the potential energy surface and in combining fingerprints with machine learning methods.

Long-term goals

Project HP4 is at the heart of materials discovery. Sampling the potential energy, free energy and chemical compound spaces is the core ingredient for finding new materials. The hope is to find among all the materials that can now be predicted by simulation methods those materials that have useful properties for technological applications such as sustainable energy generation and storage. In addition a longterm goal is to give theoretical guidance to experimental synthesis efforts for predicted materials. Another goal is to replace expensive density functional calculations by much faster methods based on machine learning. Such methods could of course also be used for structure prediction and the study of kinetic events.

Achievements

The Parrinello group has further strengthened its leadership in the development of methods for studying nucleation events. Among other things, a method has been developed that allows to simulate growth at constant chemical potential. The Goedecker group has developed a fingerprint that allows to detect similar structures in large structural datasets in an efficient and reliable way. It also did further structure predictions for superconducting materials. It also embarked on structure prediction work that is of industrial interest by studying the SiC-SiO₂ interface which is of great interest to power electronics. In a collaboration between the Röthlisberger and von Lilienfeld groups a software for chemical compound and materials design (EVOLVE) has been developed. The range of properties that can be predicted by machine learning methods has been increased considerably. In addition to the atomization energy some ten other properties can now be predicted. The Smit group has extended its topological data analysis for nano-porous media to screen for a large number of properties such as carbon capture, methane storage, Krypton/Xenon separation and methane/carbon separation. The Parrinello group will continue its work on developing novel methods for sampling free energy landscapes and obtaining kinetics of rare events. Part of this work will focus on developing a general and modular software framework for variationally-enhanced sampling within the PLUMED 2 plugin that will allow for a quicker implementation and testing of new ideas and methods. This framework will be made publicly available to the wider scientific community. Almost all methods developed by

the Parrinello group are generally applicable but special attention will be made on employing them to study nucleation and growth of crystalline phases from the melt and solution in technologically relevant systems. Among other things, this will include the calculation of nucleation rates via variationally-enhanced sampling simulations. The Parrinello group will also work on developing methods for investigating growth at constant chemical potential in complex crystals.

Milestones

- Simulation of crystal growth from the melt and solution (Parrinello).
- A machine learning model for perovskites (von Lilienfeld).
- Further structure predictions (Goedecker).
- Application of structure prediction methodology to industrially relevant problems (Goedecker).
- A demonstration of an improved framework to construct similarity kernels between structures and molecules and its application to machine learning (Ceriotti).
- A web API for representing structural landscapes (Ceriotti).
- Integration of structural fingerprints with AiiDA and PLUMED (Ceriotti).
- Extend topological data analysis to other classes of nano-porous materials (Smit).
- Novel methods for sampling free energy landscapes and obtaining kinetics of rare events (Parrinello).
- Calculation of nucleation rates (Parrinello).

New and ongoing collaborations

From the very beginning of MARVEL a substantial number of collaborations existed within project HP4. All of them are continued. The i-PI interface, developed by the group of Ceriotti, is used by the groups of Goedecker to couple different minima hopping with Quantum-ESPRESSO and has now also been ported to the CP2K group of Joost VandeVondele. The collaboration of the groups of Goedecker and Röthlisberger on the structure of perovskite materials for photovoltaic applications is also continued and will be extended



ing recent results on alanates of the Goedecker group with their machine learning tools and is collaborating with the Röthlisberger group on optimizing the composition of training sets for machine learning models.

1 Variationally-enhanced sampling (Michele Parrinello — USI and ETHZ)

1.1 Research summary

This project aims at developing variationallyenhanced sampling, a new generally applicable enhanced sampling method that is based on a rigorous variational principle. In particular we will focus on how the variational property of the method can be used to extend the method in novel and innovative ways. So far we have demonstrated how variationallyenhanced sampling can be used to obtain kinetic information from atomistic simulation in an accurate and efficient way. We have furthermore demonstrated how the method can be extended to effectively handle sampling of highdimensional free energy landscapes. A major ongoing effort is to develop a general and modular software framework for variationallyenhanced sampling that allows us to more quickly implement and test new ideas.

1.2 Scientific goals

a) *Short-term* In the short-term the main work is on developing a general software framework for variationally-enhanced sampling. This framework is designed to be very modular such that new ideas can be quickly implemented and tested. It will be integrated in the PLUMED 2 [15] enhanced sampling plugin which gives access to a wide range of molecular dynamics codes. The software framework will be made publicly available in the future to allow the wider scientific community to both use the method and also contribute to its development.

b) *Long-term* In the long-term we plan to explore various ideas how the variational property of the method can be utilized in novel ways to develop effective enhanced sampling methods. One possible avenue for this is to use empirical and phenomenological free energy models to construct bespoke bias potentials. Many other ideas will also be considered. All developments resulting from this work will be made available to other researchers in the MARVEL project and the wider scientific community.

1.3 Results obtained since the last report

To avoid confusion we note that we have renamed the variational approach presented in [1]. In the previous report it was called "variational approach to enhanced sampling" but will hereafter be called "variationallyenhanced sampling".

Since the last report we have demonstrated how variationally-enhanced sampling can be extended to effectively tackle two major issues in enhanced sampling. First, how to obtain kinetic information from biased simulations in an accurate and efficient manner. Second, how to effectively sample high-dimensional free energy landscapes. As explained in detail below, both of these developments make use of the flexibility of variationally-enhanced sampling regarding the choice of the bias potential and target distribution. Both these works are clear examples of how the variational property of the method can be employed in novel and innovative ways.

a) *Kinetic information from variationallyenhanced sampling* Obtaining kinetic rates of rare transitions between metastable states from atomistic simulation is a rather difficult problem as the time scales involved generally reach far beyond the times achievable in conventional molecular dynamics simulations. Many different approaches have thus been introduced to solve this important problem.

An appealing strategy to tackle this problem is to shift the focus onto the free energy surface and construct a bias potential $V(\mathbf{s})$ that enhances the fluctuations of a set of collective variables $\mathbf{s} = (s_1, \dots, s_d)$ chosen such that they properly distinguish the relevant metastable states. If this bias potential is constructed in such a way that the transition state region remains bias-free can the kinetic information of the physical (unbiased) system can be obtained from a simulation of the biased system. This is achieved by making use of the relation [16, 17]

$$t^* = t_{MD} \langle e^{\beta V(\mathbf{s})} \rangle_V \tag{1}$$

where all quantities on the right hand side are obtained in a biased molecular dynamics simu-

lation under the influence of the bias potential $V(\mathbf{s})$. In this way one can relate the time t_{MD} measured in a biased simulation to the physical time t^* one would obtain in an unbiased simulation. Here the average $\langle e^{\beta V(\mathbf{s})} \rangle_V$ is the so-called acceleration factor which measures the speed up or boost of a biased simulation as compared to an unbiased simulation.

The issue that remains is to construct a bias potential that leaves the transition state bias-free and yet results in a significant boost. In the socalled infrequent metadynamics method [17], previously introduced by our group, this is achieved by making the deposition rate of the metadynamics bias slower than the time spent in the transition state. This method has been quite successful in many fields, like for instance in nucleation [2] (see below, section 7). However, the need to reduce the rate of bias deposition can be a severe computational handicap and force one to rather lengthy calculations.

In this work we showed how the inherent flexibility of variationally-enhanced sampling [1] in the choice of the bias potential and target distribution can be used to construct a bias potential that effectively accelerates transitions between metastable states while ensuring a biasfree transition state region. This is achieved by constructing a bias potential that only fills the underlying free energy surface up to a certain cutoff value.

We consider within variationally-enhanced sampling a bias potential of the form

$$V(\mathbf{s}; \boldsymbol{\alpha}) = v(\mathbf{s}; \boldsymbol{\alpha}) S(-v(\mathbf{s}; \boldsymbol{\alpha}) - F_c)$$
(2)

where S(x) is a sigmoidal switching function that goes smoothly from 1 to 0 around x = 0, F_c is a cutoff parameter, and $v(\mathbf{s}; \boldsymbol{\alpha})$ is some given functional form that represents the underlying free energy surface. Generally it is most convenient to expand $v(\mathbf{s}; \boldsymbol{\alpha})$ in a set of basis functions, $v(\mathbf{s}; \boldsymbol{\alpha}) = \sum \alpha_k f_k(\mathbf{s})$, as described in the previous report and done in our previous publications [1, 3]. We furthermore consider a target distribution of the form

$$p(\mathbf{s}) = \frac{S(F^*(\mathbf{s}) - F_c)}{\int d\mathbf{s} S(F^*(\mathbf{s}) - F_c)}$$
(3)

where $F^*(\mathbf{s})$ is an estimate of the underlying free energy surface. This choice targets the sampling of the collective variables towards the region in \mathbf{s} space in which $F^*(\mathbf{s}) \leq F_c$. The underlying free energy surface is of course unknown a priori so this target distribution is iteratively updated in a similar manner as done in [3] by using the current best estimate of the



Figure 1: A hypothetical one dimensional free energy surface F(s) (black solid line) where a bias potential $V(s; \alpha)$ (red solid line) given by Eq. 2 fills F(s) up to the cutoff value F_c (green solid line). The switching function, $S(-v(s) - F_c)$ (blue dashed line) is also shown.

free energy surface $F^*(\mathbf{s}) \approx -v(\mathbf{s}; \boldsymbol{\alpha})$. The optimization of the bias potential in Eq. 2 (with respect to the variational parameters $\boldsymbol{\alpha}$) results in a bias that only fills the free energy surface up to the given cutoff value F_c , as shown for a hypothetical one-dimensional free energy surface in Fig. 1.

In order to establish the efficiency and accuracy of the method we considered a few benchmark systems where accurate reference results can be obtained. For instance, we considered the classical example of the $C_{7eq} \rightarrow C_{7ax}$ conformational transition of alanine dipeptide in vacuum, which is characterized by a barrier of around 16 k_BT (\approx 33 kJ/mol) and a mean first transition time around 26 μ s. Results obtained with different values of the cutoff F_c are shown in Fig. 2 where we can observe that for



Figure 2: Mean first passage times for the $C_{7eq} \rightarrow C_{7ax}$ transition in alanine dipeptide obtained using different F_c values (left axis). The shaded gray region depicts the standard error in the reference results. For each F_c value we performed 60 independent runs with a static bias potential. Also shown on a logarithmic scale is the corresponding acceleration factor $\langle e^{\beta V(s)} \rangle_V$ (blue, right axis) which measures the speed up as compared to unbiased simulations.

all F_c values we obtain transition times that are in very good agreement with the reference results. This is even the case for F_c values that are just a few $k_B T$ below the apparent free energy barrier. In Fig. 2 we can also observe that the method results in a significant speed up or boost of up to $10^4 - 10^5$ as compared to unbiased simulations. We furthermore found the method to be superior to infrequent metadynamics in terms of efficiency (result not shown).

We underline that this approach requires no a priori knowledge of the underlying free energy surface and is fully self-exploratory. Furthermore, it is completely general regarding the choice of the collective variables. We expect it to find wide usage in many problems where one is interested in obtaining kinetic information from atomistic simulations, both within materials science and also in chemistry, physics, and biology. A manuscript describing the method has been published in *Physical Review Letters* [4].

b) Sampling of high-dimensional free energy landscapes Most enhanced sampling approaches are rather limited in the number of the collective variables that can be biased at the same time, for example in conventional metadynamics one generally can at most use 3 to 4 collective variables simultaneously. In this work we have showed how we can use the flexibility of variationally-enhanced sampling to overcome this limitation and approximate yet very effective multidimensional biases.

The general idea is to represent the multidimensional bias potential $V(\mathbf{s})$ with an approximate but computationally tractable expression. For instance, one can think of writing $V(\mathbf{s})$ as a sum of two body biases

$$V(s_1, s_2, \dots, s_d) = \sum_{i,j} V(s_i, s_j)$$
(4)

where each two body term is expanded in some set of basis functions as done in [1, 3]. The computational advantages of such an approximate expression are evident as now the number of variational parameters is a manageable $\left(\frac{d(d-1)}{2}\right)L^2$, where *L* is the number of basis functions per collective variable, compared to L^d for the exact representation. Of course, any such approximate form will be lacking in variational flexibility when trying to represent the true high-dimensional free energy landscape. However, the deficiencies of the approximate form can be largely corrected by employing a cleaver choice of the target distribution. To

understand this point let us consider the relationship between the free energy surface $F(\mathbf{s})$, the target distribution $p(\mathbf{s})$, and the bias potential $V(\mathbf{s})$ that minimizes $\Omega[V]$ in variationally-enhanced sampling [1],

$$V(\mathbf{s}) = -F(\mathbf{s}) - \frac{1}{\beta} \log p(\mathbf{s})$$
 (5)

where we ignore unimportant constants. If the target distribution is chosen as the equilibrium distribution, $p(\mathbf{s}) \propto \exp[-\beta F(\mathbf{s})]$, the minimum of $\Omega[V]$ is trivially given by $V(\mathbf{s}) = 0$. Of course $F(\mathbf{s})$ is the very quantity we want to obtain and is not available a priori. However, we can imagine having some initial estimate $F_o(\mathbf{s})$ of the underlying free energy landscape and use that to define the target distribution as $p(\mathbf{s}) \propto \exp[-\beta F_o(\mathbf{s})]$. In this case can we rewrite Eq. 5 as

$$F(\mathbf{s}) = -V(\mathbf{s}) + F_o(\mathbf{s}) \tag{6}$$

where we can view the bias $V(\mathbf{s})$ as a correction to the initial estimate $F_o(\mathbf{s})$. There is a conceptual as well as a technical consequence of this observation. From a conceptual point of view, sampling can be greatly accelerated with an adroit choice of $p(\mathbf{s})$ that captures as much as possible any a priori information we have. From a technical point of view the requirements on the variational flexibility of the bias potential are much less stringent since $V(\mathbf{s})$ only needs to correct the errors in $F_o(\mathbf{s})$, especially those errors around the minimum. Thus, even an approximate form of $V(\mathbf{s})$ can still be capable of correcting $F_o(\mathbf{s})$.

A manuscript describing the idea presented here and it application to protein folding will be published [5]. In future work we will explore how it can be appropriately adapted to handle other problems where one is interested in exploring high-dimensional free energy landscapes.

c) *Well-tempered sampling* In the last report we described an iterative scheme for achieving a well-tempered target distribution in variationally-enhanced sampling. We furthermore showed that well-tempered sampling generally results in a significant improvement in convergence. This is mainly due to the fact that the well-tempered target distribution focuses the sampling towards the relevant low free-energy regions. A manuscript describing the iterative scheme has been published in the *Journal of Chemical Theory and Computations* [3].

1.4 Planned research for next year

One of the major ongoing efforts that will continue in the next year is to develop a general software framework for variationallyenhanced sampling within the PLUMED 2 [15] plug-in. This framework is designed to be very modular such as one can easily implement new bias potential representations, basis set expansions, target distributions, and optimization algorithm. This will allow us to very quickly implement and test new ideas. This framework will be made publicly available in the future to allow the wider scientific community to contribute to the development of the method.

A rather generally applicable idea that we will be focusing on in the next year is to use empirical and phenomenological free energy models to construct bespoke bias potentials whose parameters can then be optimized within variationally-enhanced sampling. One area where this has shown great promise is for nucleation phenomena where we can employ a bias potential based on classical nucleation theory that only depends on two variational parameters (see below, section 7). Another interesting idea based on this which we are starting to explore is to study second order phase transitions by employing a bias potential based on Landau's theory of phase transitions.

We also plan to explore various ways how the variationally-enhanced sampling method can be generally improved. For example if the optimization of the variational parameters can be improved by employing other types of stochastic optimization algorithm. We also plan to implement other types of basis sets like splines and wavelets and test their performance.

1.5 Personnel

NCCR personnel

• Omar Valsson, postdoc, 100%, from October 2014.

Matching personnel

• James McCarty, postdoc, 50%, from September 2015.

2 A fingerprint based metric for measuring similarities of crystalline structures (Stefan Goedecker — UniBas)

2.1 Research summary

Based on a characterization of the local environment of all atoms in a cell, we introduce crystal fingerprints that can be calculated easily and allow to define configurational distances between crystalline structures that satisfy the mathematical properties of a metric. This distance between two configurations is a measure of their similarity/dissimilarity and it allows in particular to distinguish structures. The new method is a useful tool within various energy landscape exploration schemes, such as minima hopping, random search, swarm intelligence algorithms and high-throughput screenings.

2.2 Scientific goals

a) *Short-term* Develop a fingerprint for crystalline structures which allows to measure configurational distances and local structure in a reliable and efficient way.

b) *Long-term* We will implement our fingerprint method in various energy landscape exploration schemes, such as minima hopping, random search, swarm intelligence algorithms and high-throughput screenings.

2.3 Results obtained since the last report

Based on the configurational fingerprint for clusters [18], which we developed before, we introduced a new fingerprint for crystalline structures. Since there is no unique representation of a crystal by a group of atoms (e.g. the atoms in some unit cell) we will use atomic fingerprints instead of global fingerprints in the crystalline case. However, this atomic fingerprint is closely related to our global fingerprint for non-periodic systems. For each atom k in a crystal located at \mathbf{R}_k we obtain a cluster of atoms by considering only those contained in a sphere centered at \mathbf{R}_k . For this cluster we calculate the overlap matrix elements $S_{i,i}^k$ as described in reference for a non-periodic system. In this first step, the amplitudes of the Gaussians c_{norm} are chosen such that the Gaussians are normalized to one. To avoid that the eigenvalues have discontinuities when an atom enters into or leaves the sphere we construct in a second step another matrix T^k such that

$$T_{i,j}^{k} = f_{c}(|\mathbf{R}_{w(i)} - \mathbf{R}_{k}|)S_{i,j}^{k}f_{c}(|\mathbf{R}_{w(j)} - \mathbf{R}_{k}|)$$
(7)

The eigenvalues of this matrix T^k are sorted in descending order and form the atomic

fingerprint vector \mathbf{V}_k .

The Euclidean norm $|\mathbf{V}_k - \mathbf{V}_l|$ measures the dissimilarity between the atomic environments of atoms *k* and *l*. The atomic fingerprints \mathbf{V}_k^p and \mathbf{V}_k^q of all the N_{at} atoms in two crystalline configurations *p* and *q* can now be used to define a configurational distance d(p,q) between the two crystals:

$$d(p,q) = \min_{P} \left(\sum_{k}^{N_{at}} |\mathbf{V}_{k}^{p} - \mathbf{V}_{P(k)}^{q}|^{2} \right)^{1/2}$$
(8)

where *P* is a permutation function which matches a certain atom *k* in crystal *p* with atom P(k) in crystal *q*. The optimal permutation function which minimizes d(p,q) can be found with the Hungarian algorithm in polynomial time. If the two crystals *p* and *q* are identical the Hungarian algorithm will in this way assign corresponding atoms to each other.

Furthermore, we introduced contraction schemes that allow to considerably reduce the amount of data necessary to characterize a crystalline structure. These schemes are based on a function $\tau(i)$ that designates a certain property of the Gaussian orbital *i* and encodes it in form of a contiguous integer index. In case of a multicomponent crystal it can indicate on which kind of chemical element the Gaussians are centered and whether the orbital is of *s* or *p* type. The principal vector is thus chopped into pieces whose elements all carry the same value $\tau(i)$. In the following presentation of numerical results we have always considered the central atom to be special, independent of its true chemical type. Having m atomic species in the unit cell and using atomic Gaussian orbitals with a maximum angular momentum l_{\max} , $\tau(i)$ runs from 1 to $(m + 1)(l_{\max} + 1)$. Now we can construct a contracted matrix t^k

$$t_{\nu,\mu}^{k} = \sum_{i,j} \delta_{\nu,\tau(i)} u_{i}^{k} T_{i,j}^{k} u_{j}^{k} \delta_{\mu,\tau(j)}$$

together with its metric tensor s^k

$$s_{\nu,\mu}^k = \sum_i \delta_{\nu,\tau(i)} u_i^k u_j^k \delta_{\mu,\tau(j)}$$

where \mathbf{u}^k is the principal vector of the matrix T^k of Eq. 7. The eigenvalues λ of the generalized eigenvalue problem

$$t^k \mathbf{v} = \lambda s^k \mathbf{v}$$

form again an atomic fingerprint of length $(m + 1)(l_{max} + 1)$ which is much shorter than

the non-contracted fingerprint \mathbf{V}_k .

Fig. 3 shows all possible pairwise configurational distances obtained with several fingerprints for various datasets. Different fingerprints are plotted along the *x*- and *y*-axis. LFP stands for the uncontracted long fingerprint and in square parenthesis it is indicated whether only *s* or both *s* and *p* orbitals were used to set up the overlap matrix, SFP[s] stands for the short contracted fingerprint with s orbitals only where the properties used for the contraction are central atom and the element type of the neighboring atoms in the sphere. For materials that have only one type of element (Si in our case) the atomic fingerprint has only length two and the coincidence theorem is not satisfied. Even though there are hyperplanes in the configurational space where different configurations have identical fingerprints, it is very unlikely that different local minima lie on such hyperplanes and the fingerprint can therefore nevertheless well distinguish between identical and distinct structures. If both *s* and *p* orbitals are used (SFP[sp]) the atomic fingerprint has at least length 4 and no problem with the coincidence theorem arise. In addition we also show the configurational distances arising from the Oganov [19] and BCM [20] fingerprints as well as from a fingerprint based on the amplitudes of symmetry functions [21]. All our datasets contain both the global minimum (geometric ground state) as well as local minima (metastable) structures, obtained from minima hopping runs [22]. Energies and forces were calculated with the DFTB+ [23] method for SiC and the molecular crystals, and the Lenosky tight-binding scheme was used for Si [24]. For the CsPbI₃ perovskite and the transparent conductive oxide Zn₂SnO₄ plane wave density functional theory (DFT) calculations were used as implemented in the Quantum-ESPRESSO code [25].

Our scheme allows to characterize crystalline structures by rather short fingerprint vectors and to decide more reliably whether structures are identical or not than previously proposed methods. Furthermore, the new fingerprint can be used to accurately explore local environments to create atomic and structural attributes for machine learning techniques.

2.4 Personnel

NCCR personnel

• Li Zhu, postdoc, 100%, from August 2014 to November 2015.





Figure 3: Correlation between different fingerprints for all the 8 test sets obtained during structure prediction runs. SFP[x] and LFP[x] indicate short and long fingerprints with x orbitals, respectively. "opt high" and "opt low" indicate the quality of the geometry relaxation.

Matching personnel

• Migle Grauzinyte, PhD student, 50%, from January 2015.

3 Mapping the landscape of materials and molecules (Michele Ceriotti — EPFL)

3.1 Research summary

This project aims at providing a comprehensive framework to build fingerprints and metrics to compare condensed-phases and molecular structures. Based on such descriptors a nonlinear dimensionality reduction representation of complex chemical and structural landscapes is constructed. Particular attention is dedicated towards providing community-accessible tools for applying these techniques, so as to maximise the impact on the platform, the NCCR MARVEL, and the materials science community at large.

3.2 Scientific goals

In broad terms, the objective of this project is to develop and apply fingerprints to assess the similarity of materials and molecules, based on the combination of descriptors of *local* atomic environments [26], pattern recognition algorithms [6], and dimensionality reduction methods [27, 28]. This framework will be used to assist classification and searching of structures in large databases of materials, to rationalise structure-property relations, and to drive accelerated sampling of structural transitions.

a) *Short-term* On a short term, the main objective is to lay the mathematical foundations for a hierarchical structure-matching fingerprint that can be used both in nonlinear dimensionality reduction and in machine learning techniques. The idea is to start from a kernel function $k(\mathcal{X}, \mathcal{X}')$ that compares two atomic *environments* \mathcal{X} and \mathcal{X}' , and then, given two structures A and B, build an environment kernel matrix $C_{ij}^{AB} = k(\mathcal{X}_i^A, \mathcal{X}_j^B)$, that compares all the environments in the two structures. Based on this environment covariance matrix, one can build a global kernel, e.g. by taking $K(A, B) = \sum_{ij} C_{ij}^{AB}$. Such kernel should satisfy the following requisites

- be applicable to periodic and non-periodic structures, including multi-species systems, and be able to compare structures with different numbers and kinds of atoms;
- be smooth and differentiable, so it can be used as the basis of accelerated sampling

schemes;

- be positive-definite, so that it can be readily applied to both nonlinear dimensionality reduction and to kernel-ridge regression machine learning;
- be relatively inexpensive, and be scalable to large systems and/or large datasets.

b) *Long-term* The key idea behind this project is to provide a unified infrastructure to compare atomic and molecular structures, in a way that can be applied both to the representation (sketch-map) of a large database of atomic structures, as well as to accelerated-sampling and search of materials space. This initiative is therefore complementary to the construction and exploration of structural databases, which is one of the central objectives of the NCCR MARVEL, and an increasingly important tool for the materials modeling community.

We aim to provide a way of visualizing the whole dataset (for example containing hundred thousand structures) at once in a so called sketch-map. Such a map shall provide a bird-eye view of materials space and shall also guide its navigation, searching for outliers and for materials with the desired properties. By comparing structural descriptors and computed properties, these representations will also assist the determination of structure-property relations.

Finally, an effective way of comparing structures is a crucial ingredient to accelerate the sampling of structural transition [29, 30, 31, 32] and to drive structure-search methods [22, 33], that are both key aspects of the research program of HP4. Apart from the algorithmic development, a significant effort will be devoted to the technical part of the implementation to facilitate usage by third parties, and integration with other initiatives within the MARVEL community.

3.3 Results obtained since the last report

Since the last report we focused our efforts on the development of structural fingerprints based on local environment descriptors, and on preparing the software infrastructure that will enable more widespread application of nonlin-



Figure 4: Sketch-map of various silicon crystal structures of different phases and polymorphs [7][34]. The color and size of the points vary according to their atomic energy and atomic volumes respectively. Different region of the plots in which structures from the same parent phase are grouped have been indicated with dotted lines. The \mathcal{C} rosses" on the upper left side of the map represent the projected low-density silicon polymorphs reported in [7]. The β luses" represent the structures projected from another database containing structures obtained from AIRSS [34].

ear dimensionality reduction techniques in the analysis of materials simulations.

a) Comparing molecules and solids across structural and composition space Evaluating the (dis)similarity of crystalline, disordered and molecular structures is a critical step in the development of algorithms to classify structures, search chemical space for better compounds and materials, and drive the next generation of machine learning algorithms for predicting the stability and properties of atomic systems. In the last few years several strategies have been designed to compare atomic coordination environments. In particular, the Smooth Overlap of Atomic Positions has emerged as a natural framework to obtain translation, rotation and permutation-invariant descriptors of atomic environments, driven by the design of various classes of machine-learned interatomic potentials.

In this project we aim to develop an algorithm

in collaboration with Prof. Gábor Csányi of University of Cambridge, to build on top of these local descriptors to compare both molecular and bulk structures by introducing powerful metrics that can compare within a unified framework chemical and structural complexity. One example result from this study is shown in Fig. 4. Here we use a database of 1'274 silicon structures containing various ideal and distorted configurations belonging to different silicon phases, e.g. diamond like, simple hexagonal, β -Sn, amorphous and liquid phase. The sketch-map in Fig. 4 is constructed by first building a map of 100 landmark configurations (chosen using farthest point sampling method), and then projecting the rest of the configurations on the map. A match kernel and 5 Å cutoff distance was used in this case to obtain the (dis)similarity matrix. The points in the sketch-map are colored according to the energy per atom of the configura-



Figure 5: A screenshot of the interactive sketch-map web interface. One can zoom/pan in the left panel while the right panel shows the overall map and highlights the active area. A pointer tool-tip can provide any relevant information, e.g. the energy, properties, or a snapshot of the configuration, etc. Further details could be accessed in a separate panel.

tions and the point sizes varies according to the volume/atom. As seen in the Fig. 4 the map is extremely well correlated with the energies and density of the configurations. Furthermore, all the structures have been automatically grouped according to their similarity to different silicon phases. For demonstration purpose, rough outlines has been drawn on the map to indicate different phases.

The "crosses" on the upper left side of the map represent few low-density polymorphs reported in [7]. The "pluses" represent the structures projected from another database containing structures obtained from AIRSS [34]. These structures were not part of our starting database and as a result none of the landmarks belong to this class. Hence, the fact that all of these closely-related structures have been automatically projected together demonstrates the transferability of sketch-maps. Once a reliable reference map is constructed for a given dataset, any new data points can be projected easily to get a meaningful representation. We are already in the final stage of this project and the manuscript is being finalized for submission.

b) *Sketch-map implementation in* PLUMED One of the main objectives during 2015 has been to obtain an efficient and easy to use implementation of the sketch-map algorithm in PLUMED, a widely used software for enhanced sampling and free-energy methods [15]. In collaboration with Gareth Tribello (Queen Mary University Belfast) and with the help of Sandip De and a master project student (Anurag Sharma), we already have delivered a working implementation of sketch-map in PLUMED, which can be obtained from the public PLUMED repository. Small improvements and new features will be added periodically as we develop new algorithms and understandings, together with a more thorough documentation and examples.

3.4 Planned research for next year

a) Developing a web interface for sketch-map Work is in progress on developing an interactive web interface for generating and displaying sketch-maps. This includes both a backend to process structural data and generate a sketch-map representation, and a front-end to display interactively the map, which can then be linked to a more thorough characterization of each structure. The interface will include a fully interactive map which can be zoomed and panned, in a way which is not possible in conventional static documents. A first prototype is already available, and we will use it to generate supporting information for our upcoming publications. During next year we will polish this infrastructure, and define an API to include it in the data processing and visualization pipeline of materials databases, such as those that will be generated by MARVEL. A screenshot of the current version of the interactive interface is shown in Fig. 5.

b) *Oligopeptide database* During 2016 we plan to invest some time in studying a pub-

licly available database of oligopeptides (http://aminoaciddb.rz-berlin.mpg.de/) in collaboration with Carsten Baldauf from FHI, Berlin. Preliminary work highlighted the potential of sketch-map for identifying patterns in the dataset, and for locating outliers corresponding to unforeseen chemical reactions during the automated structural search.

c) *Implementation of a fingerprint module in* PLUMED The construction of a sketch-map relies on the calculation of a (dis)-similarity matrix for the given set of structures. Given the importance of fingerprints for structure search, accelerated sampling of phase transitions, and generation of machine learning potentials [35, 36, 32, 37], we will participate in a collaborative effort within HP4 to obtain an efficient and well-documented library to compute such fingerprints. In particular, we are exploring the possibility of including such library as part of the PLUMED package, so as to make it easy to access by a broad community.

3.5 Synergies with other computational and experimental efforts

The team involved in this project has also contributed to other computational efforts that are not directly related to the goals discussed here, but that are connected with the general objectives of the NCCR MARVEL.

a) *Within MARVEL* Collaborations with other members of MARVEL mainly revolve around the development and use of i-PI [38], a Python interface that was initially developed for path integral molecular dynamics, and is now being extended to other advanced molecular dynamics techniques.

• In collaboration with Clémence Corminboeuf (EPFL) we have used i-PI in combination with the DFTB+ software [23] to perform replica-exchange simulations of complex conformational and chemical rearrangements in organic molecules [8].

• In collaboration with Joost Vandevondele (ETHZ), we are using i-PI in combination with CP2K to include nuclear quantum effects on top of state-of-the-art wave function methods, by combining multiple time step and ring polymer contraction techniques. We are currently finalizing a publication in which we demonstrate such implementation, and further collaboration is foreseen into the modeling of static and dynamic quantum nuclear effects in liquid water treated at the MP2 level of electronic structure.

b) *Outside MARVEL* Obtaining instructive representations of sets of molecular conformations is an urgent need also outside the MAR-VEL collaboration. In the framework of the MPG-EPFL center for molecular nanoscience, we are collaborating with Carsten Baldauf (FHI-Berlin) to map the structural and chemical landscape of proteinogenic amino acids and dipeptides, and rationalise their conformational changes in the presence of divalent cations.

3.6 Personnel

NCCR personnel

• Sandip De, postdoc, 100%, from March 2015.

Matching personnel

• Piero Gasparotto, PhD student, 30%, from March 2015.

4 Topological data analysis for nano-porous materials (Berend Smit — EPFL)

4.1 Research summary

The first steps have been made to apply topological data analysis methods to screen the best performing materials and identify similarities among them.

4.2 Scientific goals

a) *Short-term* The short term aims are to develop topological data analysis methods adapted to the study of nano-porous materials.

b) *Long-term* We intend to apply the developed topological data analysis methods to different screening studies and to extend the methodology to other systems.

4.3 Results obtained since the last report

We often use the term Nano-porous Materials Genome with regard to the development of advanced nano-porous materials because analogous with the human genome, a very large number of materials (infinite, in principle) is possible by combining different molec-


Figure 6: A workflow of a new approach to characterize and compare zeolites based on their topologies. The first step is identification of pore structures represented with a finite set of data points using the open-source software Zeo⁺⁺. The second step is characterization of pore shape with persistent homology analysis and encoding information into barcode form. Last, we use the outputs from the persistent homology analyses in a form of barcode as a descriptor (i.e. fingerprint) for identifying and comparing zeolites.

ular building blocks. The present computational approach uses brute-force molecular simulations to generate the thermodynamic data needed to predict the performance of these materials in different applications, but this approach is limited to tens of thousands of structures due to computing time constraints. So, we are seeking to tackle this issue by transitioning away from the brute-force characterization to high-throughput screening approaches based on big-data analysis. For a successful screening, it is ultimately required to use a precise descriptor to establish relationship between properties and performances of materials. While in recent years, researchers have tried to develop simple material descriptors such as diameter of largest included sphere and pore volume to characterize porous materials, these descriptors unfortunately remain insufficient to find the best materials. Recognizing that these failures are due to an insufficient amount of information with simple descriptors, in this work, we aim to develop a new approach that characterizes and compares materials with more complete description of materials (high-dimensional data about pore topology, iso-surface of energy, charge map, etc.). In conjunction with big-data analysis techniques, this approach identifies materials based on shapes like face recognition tech-



Figure 7: (a) Two-dimensional histogram (red dots) of deliverable capacity (D.C.) of zeolites in the initial set against that of most similar ones out of the entire set by TDA based descriptor, in comparison with the results (green dots) by convention descriptors in a form of vector $\{D_i, D_f, \rho, A.S.A., A.V.\}$, where D_i and *D_f* are a maximum diameter of largest included sphere and free sphere, respectively, ρ is the density (g/cm³) of a zeolite, A.S.A. and A.V. indicate accessible surface area (m^2/cm^3) and volume fraction to a gas probe molecule. The blue, diagonal line is a perfect match. (b) Examples of zeolites and their pore topologies (blue) in the initial set (left), selected set by conventional descriptors (center), and by TDA (right). K_H is the Henry coefficient (mol/kg/Pa) of a zeolite for gas probe molecule, and Q_{ad} denotes heat of adsorption (kJ/mol).

nique.

We started this project with the development of a new approach to compare and characterize zeolites using topological data analysis. There are two important reasons why we chose zeolites as a starting material in our research: (i) we already possess over a hundred thousand simulation data regarding the zeolites for several important applications such as carbon capture and methane storage, which can serve as a reliable reference set; (ii) as these materials are in the all-silica from of the material, from a chemical point of view all zeolites have the same chemical composition and only differ in

their pore topology. This allows us to initially focus on how to describe pore topology, and we can easily extend our approach to other systems by simply adding information about energy or charge. Fig. 6 shows a workflow of our new approach for comparing pore topologies. This approach has been successfully applied to screen top-performing zeolites for several applications such as carbon capture [39], methane storage [40], Krypton/Xenon separation [41] and methane/carbon separation. Fig. 7 shows an example of this approach to methane storage application.

4.4 Planned research for next year

In the next period, we plan to extend our methodology to other classes of nano-porous materials (metal organic frameworks, zeolitic imidazolate frameworks, porous polymer networks, etc.) and other applications. Because there is an additional dimension of chemical composition along with the pore topology for these classes of porous materials, we will improve our descriptor to use an image of energy landscape and charge map.

4.5 Personnel

NCCR personnel

• Yongjin Lee, postdoc, 100%, from November 2014.

Matching personnel

- Senja Bartel, postdoc, 100%, from September 2015.
- Davide Tiana, postdoc, 100%, from June 2015.

5 Development of an evolutionary algorithm tool set for chemical compound and materials design (Ursula Röthlisberger — EPFL, Anatole von Lilienfeld — UniBas)

5.1 Research summary

We have developed and implemented a range of evolutionary algorithms for efficient search of chemical space, enabling accelerated material and chemical compound design. We have created a flexible and modular package of genetic algorithm (GA) implementations, which we call EVOLVE. During this year, we have further developed EVOLVE to include additional evolutionary computation (EC) models, multiobjective optimisation, and support for cluster parallelization schemes. In collaboration with the group of Anatole von Lilienfeld, we are extending the optimisation of machine learning models of molecular properties with EVOLVE to the investigation of a priori selection criteria for training set selection [9]. We now look to apply EVOLVE to the compositional and structural optimisation of perovskite crystals to increase stability against undesirable phase transitions.

5.2 Scientific goals

a) Short-term

• GA-QM/ML Models of Molecular Properties. In collaboration with the group of von Lilienfeld, in the first year, we have used EVOLVE to optimize training set composition for machine learning models of molecular properties, thereby facilitating a reduction in predictive error. During this year, the work has been extended

significantly. The database of molecules used for learning and optimisation has been increased from 6095 constitutional isomers of $C_7H_{10}O_2$ to 21800 stoichiometrically diverse molecules, while maintaining a significant reduction in predictive error through GA optimisation. The number of properties we have analyzed has also increased from solely atomization enthalpies to 10 thermochemical and electronic properties. Ensembles of GA optimisations have been performed, which we will now use to investigate the possibility of generating a priori selection criteria.

• Compositional Optimisation of Perovskites. Perovskite crystals of the form ABX_3 where $A = MA^+$ (CH₃NH₃⁺) or FA^+ ((NH₂)₂CH⁺), $B = Pb^{2+}$ and $X = I^-$, have recently emerged as promising materials for solar cell applications. One of the most promising perovskites, FAPbI₃, has been found to exhibit superior performance with respect to light harvesting and optimal open-circuit voltage properties. Unfortunately, the perovskite phase of FAPbI₃ is not stable at room temperature and quickly converts to the unsuitable δ -phase. Recent studies have shown that the stability can be improved by mixing cations, e.g. FA⁺ or MA⁺ [42], and simultaneous mixing of I⁻ and Br⁻ [43]. For each cationic/anionic stoichiometry, many different structural isomers exist within the PbI₃ framework, thus a systematically exhaustive investigation of all possible configurations is not possible. We will determine the (near) optimal set of configurations using EVOLVE with respect to maximizing stability against conversion to the δ -phase.

- b) Long-term
 - EVOLVE development. In the past year, we have extended EVOLVE to include particle swarm optimization (PSO) and Bat algorithm (BA), and additionally included more EC variation, selection and reproduction operators. We now plan to implement more EC algorithms such as learnable evolution, island models via migrator operators, as well as to further develop the interoperability between several simultaneous EC computations for a given optimisation problem. Our long-term goal is to develop and polish EVOLVE into a software which can be released to the wider scientific community interested in materials design, including detailed API, analysis tools and extensive tutorials.

In a continuation of the collaboration with the group of von Lilienfeld, we plan to develop a machine learning model for the prediction of DFT total energies of mixed perovskites systems which will allow for more extensive GA optimizations with the predictive accuracy of an electronic structure method. The initial training set will be generated by performing GA runs on minimal cell perovskite samples for which the DFT total energy evaluated with Quantum-ESPRESSO is used as fitness function.

5.3 Results obtained since the last report

We have applied EVOLVE to optimize the training set composition for molecular models based on 10 thermochemical and electronic properties. We made use of a subset of 21'800 stoichiometrically diverse molecules from a larger database [10] containing chemical properties and relaxed geometries of 133'885 small organic molecules, computed at the DFT/B3LYP/6-31(2df,p) level of theory. Fig. 8 displays the learning curves of randomly generated and GA-optimized training sets for enthalpy of atomization, using both the direct-learning and the Δ -learning method [11]. There is a substantial and systematic reduction in RMAE across all properties after GA optimiza-



Figure 8: GA optimized learning curves for ML models trained directly on enthalpy of atomization, H (green, yellow). Significant reductions in predictive error are obtained. applying GA to Δ -learning methods (blue, red) additionally shows this approach is stable with respect to changes in the underlying ML model.

tion. The error scatterplot and distribution insets show that the GA optimized training sets approach an ideal mapping from Δ^{GA} to Δ^{ref} .

5.4 Planned research for next year

Next year we will investigate the selection bias in machine learning of molecular properties through EVOLVE. Here, we will attempt to construct simple criteria for generating training sets to be used in such models. We will also apply EVOLVE to the compositional and structural optimisation of perovskite crystals with respect to improving the stability of perovskite against conversion to the δ -phase and develop a machine learning model for rapid fitness evaluation.

- 5.5 Synergies with other computational and experimental efforts
- a) Within MARVEL
 - GA-QM/ML Models of Molecular Properties. This project is a collaboration with the group of Anatole von Lilienfeld at Uni-Bas.
- b) Outside MARVEL
 - Optimisation of Perovskite Unit Cells. This project is a collaboration with the experimental group of Michael Graetzel at EPFL.

5.6 Personnel

NCCR personnel

• Nicholas J. Browning, PhD student, 100%, from May 2014 to August 2015.

Matching personnel

- Marta Andreia Da Silva Perez, postdoc, 50%, from May 2015.
- Nicholas J. Browning, PhD student, 100%, from September 2015.

6 Phase discovery and phase diagrams of halide perovskites (Ursula Röthlisberger — EPFL, Stefan Goedecker — UniBas, Michele Parrinello — USI and ETHZ)

6.1 Research summary

The chemical and crystal phase stability is a crucial issue in the search for novel materials for solar cells. Also important is the possible formation of undesired phases which are kinetically favored. Another aspect to consider is how *dopants* or *additives* can affect the morphology, size and structure of crystal grains. Investigating all these aspects is the objective of this project which is an intra-MARVEL collaboration between VP2 (Novel Materials Applications) and HP4 (Advanced Sampling Methods) involving the three groups of Röthlisberger, Goedecker and Parrinello.

The achievement of the objectives summarized above requires the testing and tuning of existing computational techniques as well as the development of novel approaches. The exhaustive explorations of the phase space of crystals, and simulations of the nucleation mechanisms are computationally highly demanding tasks even when using enhanced sampling methods. We therefore adopt a hierarchical multiscale approach for the application of these methods to systems of practical interest. This hierarchical multiscale approach consists in using a simpler model of the atomistic system for a preliminary screening. At a second stage, this model is then replaced by a more accurate quantum mechanical description of the interatomic forces. Thus, during the second year of MARVEL, we followed a double path. On the one hand we continued exploring the phase space of a prototypical perovskite system, CsPbI₃, to test and tune the minima hopping (MH) method for application to halide perovskites. On the other hand, we worked at the development of a classical polarizable force field, which will speedup the MH exploration. and enable simulations of the nucleation process.

6.2 Scientific goals

a) *Short-term* Due to the liquid processing of hybrid organic-inorganic lead halide perovk-

sites, there is only limited control of the crystal phase that is formed. This is a problem for halide perovskite systems, like CH₃NH₃PbI₃ (MAPI) and [HC(NH₂)₂]PbI₃ (FAPI), that exist in several phases [44]. In particular, since the crystalline structure crucially affects the optical and transport properties of this class of materials, a complete knowledge of the "phase space", and the stability ranges of the various crystalline phases, is crucial for the *in silico* design of suitable photovoltaic materials.

Investigating the thermodynamics and kinetics of crystal formation is a very challenging task from the point of view of simulations. Only few methods exist to explore the phase space of such systems that are often characterized by energy barriers of hundreds of k_BT . Here we have employed the minima hopping method (MH, [33]) to identify low-energy crystalline structures of CsPbI₃, together with a large number of defect structures. In a next step, we will study the effect of the crystalline phase and defects on the optical and transport properties of these materials. To perform this search the Quantum-ESPRESSO code was coupled with the minima hopping algorithm using a communication scheme based on sockets.

b) Long-term The application of MH, or other enhanced sampling techniques, to identify the phases of a crystalline system requires an accurate model of the atomic forces. Density functional theory (DFT) has proven to work quite well for lead halide perovskites, with an error in the lattice parameter of only $\sim 1\%$. However, DFT is computationally rather expensive, especially if one considers the large atomistic samples necessary to allow the formation of some of the phases. For example, the unit cell of β -FAPI contains 144 atoms. This makes the use of DFT unsuitable for an extensive screening of the phase space of the materials we are interested in. Thus, we are currently developing a classical force field which is at the same time accurate enough to reproduce the main features of the system (at a semi-quantitative



Figure 9: Polyhedral views of the experimental δ -phase (a) and cubic (b) structures of CsPbI₃.

level) and computationally expedient.

We plan to use this classical force field in MH simulations to explore the phase space of more complex perovskite systems identified as good candidate materials for solar cells in our MAR-VEL project on novel materials for solar cells. Another important problem we intend to investigate is the formation of perovskite crystal grains. It has been noticed that the morphology and structure of the perovskite crystallites can be significantly affected by the thermodynamics and chemical conditions under which the deposition process is carried out, and the performances of the materials are largely affected by these factors [45]. Thus, in order to optimize also this part of the process, we intend to study the crystallization mechanism of some prototypical perovskite systems. For this we will perform variationally-enhanced sampling simulations [1] in combination with the force field mentioned above.

6.3 Results obtained since the last report

Several MH simulations were performed with cells containing 4 and 8 formula units of CsPbI₃ starting from different input configurations. In addition to the experimental δ -phase (Fig. 9a) and cubic perovskite structure (Fig. 9b), many new structures were discovered. Several of these metastable structures are shown in Fig. 10. All the structures have in common that they contain PbI₆ octahedra



Figure 10: *Polyhedral views of various predicted structures of CsPbI*₃*.*

that share corners, edges or faces. In the (non-perovskite) δ -phase each PbI₆ octahedron shares an edge and has three neighboring PbI₆ octahedrons, whereas in the cubic phase the

PbI₆ octahedra share all their six corner I atoms with one of the six neighboring PbI₆ octahedron. The structures shown in Fig. 10a, b and c can be considered as distortions of the cubic structure, resulting from the rotation of the octahedra around their three I-Pb-I axes. In structure d of Fig. 10, the PbI₆ polyhedra share both edges and corner I ions. Due to their smaller bandgap, the structures with highly connected octahedra are more appealing for photovoltaics than the δ -CsPbI₃ phase which has a too large band gap.

Since the use of small unit cells led to a significant number of dynamically unstable structures, we did all the runs with 40 atoms. This large unit cell makes the structure prediction runs quite slow. However, we expect that the resulting large database will allow a much better understanding of the crystalline structures and their defects in these perovskite materials. This improved understanding can help in the design of better materials for photovoltaics.

6.4 Planned research for next year

During the next year, we will focus on the exploration of the phases of some practically relevant perovskite systems, such as MAPI. As explained above, this requires the development of an *classical* force field. A fixed-point charge force field for MAPI has been recently introduced [46]. While this force field is able to correctly describe some of the features of MAPI, e.g. the transition from the *orthorhombic* to the tetragonal phase at a temperature between 100 and 200 K, it fails to catch some of the relevant details of these two phases, e.g. the break of symmetry of the Pb-I bonds when the system passes from the orthorhombic to the tetragonal phase. To capture these details, which can be crucial for the identification of new phases, we intend to extend this model by introducing

a polarization term which can make the two classes of ions distinguishable.

A second research direction we intend to pursue is the nucleation mechanism of MAPI. This project will be carried out in collaboration with the group of Michele Parrinello. The results of the simulations will be compared with the data obtained from the group of Michael Graetzel, who is also working on this subject from the experimental point of view.

6.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* This project is a joint effort between VP2 and HP4. The exploration of the phase space of perovskite systems involves the groups of Röthlisberger and Goedecker while the study of the nucleation mechanism of MAPI will be performed as a collaboration between Röthlisberger and Parrinello.

b) *Outside MARVEL* The theoretical predictions will be validated by experiments in the group of Michael Graetzel at EPFL.

6.6 Personnel

NCCR personnel

- Li Zhu, postdoc, 100%, from August 2014 to November 2015.
- Simone Meloni, senior researcher, 100%, from May 2014.

Matching personnel

- Negar Ashari Astani, PhD student, 100%, from May 2014.
- Ariadni Boziki, PhD student, 100%, from December 2014.

7 Development of advanced sampling methods for nucleation and growth (Michele Parrinello — USI and ETHZ)

7.1 Research summary

This project aims at overcoming the sampling and time scale limitations that hinder simulations of two processes that lie at the heart of materials science, namely, nucleation and growth. Previously we have shown how to recover free energy surfaces (FES) associated to nucleation phenomena taking into account finite-size effects, and have developed and implemented a novel algorithm for simulating growth at constant chemical potential. Building upon these achieved goals, we now demonstrate how to calculate nucleation rates for vapor-liquid phase transitions. Moreover, we describe a method to calculate the chemical potential in dense systems. In the near future we plan to refine our methods and tackle the important problem of nucleation of crystals from the melt and solution, both for the calculation of the FES and nucleation rates, using variationally-enhanced sampling.

7.2 Scientific goals

a) *Short-term* In the last year we have calculated nucleation rates for the condensation of liquid droplets from a supersaturated vapor using metadynamics, developed a technique to study growth at constant chemical potential (C μ MD), and proposed a novel methodology to compute the chemical potential in dense systems. In the short term we plan to:

- validate the CµMD method [12] on more complex systems and processes. So far it has been applied with success on urea crystal growth from solution;
- continue the development and testing of the technique devised to enhance the sampling of test particle insertion in dense systems; this is expected to be completed in the immediate future;
- calculate nucleation rates employing the variationally-optimized free energy flooding [4] in combination with the functional form provided by classical nucleation theory to construct the bias potential.

b) *Long-term* In the long term we aim at combining the features of the different methods we developed, in order to devise new schemes for the molecular simulation of materials. This includes the application of the novel variational approach [1] to the sampling of test particle insertion aiming at a full grand-canonical sampling method. Furthermore we plan to study nucleation of crystalline clusters from the melt and solution taking advantage of the array of tools developed for gas-liquid nucleation.

7.3 Results obtained since the last report

a) Nucleation rates from metadynamics Recently it has been shown that accurate transition times of thermally activated processes can be obtained from well-tempered metadynamics (WTmetaD) simulations [17, 47]. We have employed this method to obtain nucleation rates for the condensation of liquid droplets from a supersaturated vapor. For the first time nucleation rates are calculated in a regime of moderate to low supersaturations where experiments are usually performed.

As discussed in detail in [17], the rate calculation does not require a converged estimate of the free energy barrier associated with nucleation, which is hard to sample. Instead, this approach calculates rates by quantifying the acceleration achieved in metadynamics. Crucial to this procedure is the hypothesis of negligible bias deposition at the transition state. Such



Figure 11: Average nucleation times in finite size (in blue) and macroscopic (in red) cases as a function of supersaturation ratio *S*. While conventional molecular dynamics simulations are limited to times of the order of 100 ns, WTmetaD allows extending the time scale well into the hour regime.

a bias is constructed through *infrequent deposition* in the space of one or more order parameters that are able to describe the stable intermediates and the product states. The fulfillment of such conditions can be checked a posteriori owing to the fact that escape times from long-lived metastable states obey Poisson statistics [48]. When this is the case, the whole transition time distribution can be recovered from a set of WTmetaD simulations. Nucleation times in a macroscopic system were obtained by rescaling the calculated times using a finite-size correction developed in a previous publication under MARVEL [13].

Nucleation rates computed with WTmetaD for argon are reported in Fig. 11 as a function of the supersaturation ratio *S*, together with literature data computed from brute-force unbi-



Figure 12: Water nucleation rates as a function of the scaled supersaturation. With WTmetaD simulations we are able to bridge the gap between simulations and experiments. For reference, we show the results of very expensive brute-force simulations.



Figure 13: Geometry arrangement of $C\mu MD$ for the study of planar crystal growth. The vertical lines indicate where the force F^{μ} is applied. The gray-shaded areas indicate the CR volume.

ased molecular dynamics (MD) simulations. In Fig. 11 it can be seen that WTmetaD allows extending the time scale well into the hour regime. This implies a speed up of up to 10 orders of magnitude relative to typical standard MD simulations [49].

Water nucleation rates are shown in Fig. 12 as a function of a scaling relation that allows to include in the same plot results obtained at different temperatures. It can be seen that there is a gap between nucleation rates measured in experiments and those calculated from computer simulations. With WTmetaD simulations we are able to obtain nucleation rates within this gap and the results are in good agreement with very expensive brute-force simulations [50].

b) Constant chemical potential MD MD study of chemical processes in solution, such as crystallization, self-assembly, adsorption or electrochemical reactions is of primary importance in material science. However, the presently available computational resources impose a limit to the size of the systems that can be investigated. In the framework of the aforementioned processes such a limited size usually affects the dynamics, preventing MD from reproducing the behavior of realistic macroscopic systems. For example, in the study of crystallization from solution, the latter is depleted as the phase transition proceeds, so that the chemical potential of the crystal environment changes, affecting the process itself [31]. To overcome this problem, we have developed and tested the constant chemical potential MD (C μ MD) technique [12], a method that mimics constant chemical potential conditions in the environment solution of the considered system. This method avoids to appeal for particle insertion, which in dense environments such as liquids is

highly inefficient. In $C\mu MD$ part of the simulation volume is used as a finite reservoir for the rest of the system, where the process of interest takes place. An external adaptive force is then applied to regulate the exchange of molecules between these two regions, in order to maintain the relevant region of the simulation at constant chemical potential. In Fig. 13 we show the C μ MD geometry for planar crystal growth of urea in aqueous solution. The force F^{μ} is applied at both sides of the crystal, maintaining a constant solution composition in the highlighted control region (CR), which surrounds the slab. Since the previous report we have completed the validation of $C\mu MD$ in the case of urea crystal growth in aqueous solution, testing both the crystallization in the $\{001\}$ and {110} direction. The method has allowed to evaluate urea growth rates at constant chemical potential, confirming the results of our previous works on urea crystallization [31, 30].

c) Test particle insertion using metadynamics During last year we have been working on a new method to enhance the sampling of the test particle insertion technique. Test particle insertion is a wide-spread approach for the calculation of chemical potential of liquid systems. In this technique, also referred to as the Widom method [51], the excess chemical potential is calculated by sampling the insertion energy of an extra test particle over the canonical trajectory of a molecular system, according to the following formula

$$\mu^{\text{ex}} = -\beta^{-1} \ln \left\langle \frac{1}{V} \int \exp\left(-\beta \Delta U(r)\right) d^3 r \right\rangle$$
(9)

where *V* is the system volume, *N* is the number of particles, $\Delta U(r)$ is the energy cost of inser-



Figure 14: Sketch showing the most salient features of the variationally-enhanced sampling method using a model as a functional form for the bias potential. It can be seen that in a standard unbiased simulation (shown in blue) the system is constrained to move either in the vapor basin (below) or the liquid basin (above) due to the presence of high free energy barriers. By means of a bias potential constructed employing the functional form of classical nucleation theory (CNT) and with the parameters ($\Delta\mu$, σ) variationally optimized, the system is able to explore the whole CV space sampling the well-tempered distribution (shown in green).

tion of a particle at position r, $\beta = (k_B T)^{-1}$, and the average is performed in the canonical ensemble. It is clear that the largest contributions to μ^{ex} come from those configurations characterized by regions with small $\Delta U(r)$. In dense systems such configurations are very rare, and ordinary sampling is not sufficient to obtain an accurate estimate for μ^{ex} [52]. To overcome this problem we have designed a technique that uses well-tempered metadynamics (WTmetaD) to enhance the density fluctuations of the system, so that the relevant configurations are sampled and the chemical potential is calculated with higher accuracy. Up to now the method has been tested with high-density Lennard-Jones fluids, but it can be generalized to applications of higher interest in the framework of material studies.

d) Variationally-enhanced sampling of nucleation phenomena We have explored the interesting possibility of using a model to construct a bias potential to retrieve the free energy surface of nucleation phenomena. We employ the variationally-enhanced sampling (VES) method [1] developed in the context of MARVEL to optimize the parameters of the model. Once the parameters are fully optimized, the system samples the well-tempered distribution [3] in the collective variables (CV) of choice. In the particular case of nucleation of liquid droplets from the gas phase, a natural choice of collective variable is the number of molecules with a coordination number larger than a certain cutoff.

The simplest yet effective model for nucleation is the well-known classical nucleation theory (CNT). In the framework of CNT corrected for finite size effects [13], the free energy as a function of the number of molecules in a cluster of the new phase, *n*, can be approximated as:

$$F_{CNT}(n;\Delta\mu,\sigma) = -\Delta\mu n + \sigma n^{2/3} + nk_BT + (n_{tot} - n)k_BT \ln\left(\frac{n_{tot} - n}{n_{tot}}\right)$$
(10)

where $\Delta \mu$ is the difference in chemical potential between the parent and child phase, σ is the effective surface energy of the interface, and n_{tot} is the total number of molecules in the system. In many enhanced-sampling methods the bias potential V(s) is constructed as to compensate totally or partially the free energy surface of the system. For instance, in metadynamics V(s) is constructed as a sum of kernels, typically chosen to be gaussians, and in variationally-enhanced sampling V(s) is generated with basis set expansion (e.g. Fourier series or Chebyshev polynomials) and the coefficients of this expansion are optimized. In this work we employ the functional form of CNT (Eq. 10) to construct the bias potential, i.e. $V(n; \Delta \mu, \sigma) = -(1 - \gamma^{-1}) F_{CNT}(n; \Delta \mu, \sigma)$ where the factor $1 - \gamma^{-1}$ allows the sampling of the well-tempered distribution in n. The parameters $(\Delta \mu, \sigma)$ are optimized until V(s) minimizes a functional $\Omega[V]$. In the case of the welltempered distribution, the V(s) that minimizes $\Omega[V]$ satisfies $V(s) = -(1 - \gamma^{-1})F(s)$ with F(s) the FES corresponding to the CV s. Details of the VES method can be found in [1, 3]. In Fig. 14 the main features of our method are

sketched. It can be seen that in a standard un-

biased simulation the system remains either in the vapor basin or the liquid basin due to the presence of high free energy barriers. By means of a bias potential constructed employing the functional form of CNT and with the parameters ($\Delta\mu, \sigma$) variationally optimized, the system is able to explore the whole CV space sampling the well-tempered distribution (shown in green).

7.4 Planned research for next year

During the last year we have developed and validated an array of tools for the study of nucleation and growth phenomena. We have shown how to retrieve both free-energy surfaces and nucleation rates from enhancedsampling simulations, and to study growth at constant chemical potential. However at this stage we have limited our research to relatively simple systems such as gases and organic molecules with reduced conformational flexibility. Hence, for the next year we have defined the following goals:

a) Sampling at constant chemical potential In the framework of constant chemical potential sampling we plan to further explore the possibilities of $C\mu$ MD method testing it with more complex crystals (e.g. acetylsalicylic acid and para-amino benzoic acid), possibly coupling its action with enhanced sampling methods such as metadynamics or variational approach. We also plan to complete the validation studies for the aforementioned enhanced sampling for test particle insertion technique. The coupling of this method with the variational technique will be also explored.

b) *Calculation of nucleation rates via variationally-enhanced sampling simulations* In order to speed up the infrequent deposition approach to the calculation of rates, we will obtain nucleation rates for the gas-liquid phase transition employing the variationally-optimized free energy flooding in combination with the functional form provided by clas-

sical nucleation theory to construct the bias potential.

c) Nucleation of crystalline phases from the melt and solution During the processing stage, nucleation plays an important role in controlling the microstructure and thus the properties of many materials. Molecular simulations of real systems aided with powerful enhanced sampling methods can provide an insight into phenomena occurring at a length scale inaccessible to experiments. We plan to tackle technologically relevant systems such as those that exhibit a crystalline phase that grows either from the melt or solution. Our main objectives are the retrieval of free energy surfaces and the calculation of rates.

7.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Together with the group of Michele Ceriotti we will make structural fingerprints, collective variables and machine learning methods developed in the context of MARVEL widely available to the materials science community. To this end, we will include them in the analysis and free energy methods plugin PLUMED 2. We expect it to be used both as a method to navigate and query databases, and to accelerate sampling to study structural transformations.

7.6 Personnel

NCCR personnel

• Pablo Piaggi, PhD student, 100%, from November 2014.

Matching personnel

- Claudio Perego, postdoc, 40%, from October 2014.
- Matteo Salvalaglio, postdoc, 50%, from May 2014 to August 2015.

8 Superconductivity in highly metastable phases of phosphorus-hydride compounds under high pressure (Stefan Goedecker — UniBas)

8.1 Research summary

The phase stability and superconducting properties of the recently reported phase in PH_3 under pressure have been investigated with *ab initio* calculations. Our extensive structural search show that the hydrogen-rich phosphorus phases are thermodynamically unstable. Although including vibrational zero-point effects to our calculations improves the stability of PH_1 , PH_2 and PH_3 , the hybrid HSE06 functional favors decomposition, rendering all binary phases metastable. Nevertheless, the low-enthalpy ground states at fixed composition may have been synthesized through non-

equilibrium effects and stabilized by strain. Several phases were predicted to be good phonon-mediated superconductors and could thus in principle account for the measured high- T_c in experiments.

8.2 Scientific goals

Recently, compressed phosphine was reported to metallize at pressures above 45 GPa, reaching a superconducting transition temperature (*T_c*) of 100 K at 200 GPa [53, 54]. However, neither the exact composition nor the crystal structure of the superconducting phase have been conclusively determined. In this work the phase diagram of PH_n (n = 1, 2, 3, 4, 5, 6) was extensively explored by means of ab initio crystal structure prediction using the minima hopping (MH) method. The results do not support the existence of thermodynamically stable PH_n compounds, which exhibit a tendency for elemental decomposition at high pressure. Although the inclusion of vibrational entropy for the calculation of the phase diagram improves the stability of $PH_{1,2,3}$ it remains questionable if the experimentally measured high value of T_c can be fully attributed to a phase-pure compound of PH_x .

a) Short-term Exploration of the composition and configurational space of PH_x at high pressure.

b) *Long-term* The core objective of this research is two fold, (i) to find novel materials with high T_c , (ii) to reduce the tremendous pressure at which the chemically precompressed compound attains metallicity and superconductivity. Hence a systematic, automate and smart way of exploration is necessary in order to explore the large family of hydrogen rich materials, i.e. all molecules containing hydrogen.

8.3 Results obtained since the last report

All scientific goals planned during this first stage have been reached. The formation enthalpy for different compositions studied at 120 GPa of pressure is shown in Fig. 15 with respect to elemental decomposition. Clearly, none of the mixed phosphorus-hydrogen compositions considered in this work are stable towards elemental decomposition in the whole pressure range studied. We also investigated the influence of vibrational entropy and zeropoint energy (ZPE) on the phase stability. For PH the hull distance is less than 5 meV/atom at 0 K, whereas for PH₂ and PH₃ it is about



Figure 15: Predicted formation enthalpies of PH_n with respect to decomposition into P and H at 120 GPa. The solid red line denotes the convex hull of stability. Black triangles represent PBE and circles hybrid functional values. The color scale dots indicate the free energy within the harmonic approximation at temperatures between 0 and 400 K in steps of 100 K on top of PBE energies.

7 meV/atom and 30 meV/atom, respectively. Similar to our results obtained with the HSE06 functional, there is a clear distinction in the vibrational effect on the stability for PH_n-phases with hydrogen content lower or higher than n = 3: PH₁, PH₂ and PH₃ increase in stability, whereas PH₄, PH₅ and PH₆ decrease. This behavior can be attributed to the existence of H₂ molecules in the crystal lattice of the hydrogenrich phases which induce high-energy vibrational modes responsible for an increased ZPE. The structures of the ground states of PH, PH₂ and PH₃ are shown in Fig. 16, indicating the dominating tendency of P atoms to form polymers with hydrogen saturating the dangling bonds, resulting in 1D chains or 2D layers. This picture is confirmed when looking at the isosurfaces of the electron localization function (ELF), clearly showing that electrons are local-



Figure 16: Low-lying enthalpy structures found for PH, PH₂ and PH₃ compositions at pressure of 120 GPa. The large and small spheres denote the P and H atoms, respectively. The ELF at a fixed value of 0.8 is shown in the upper part of each structure.

ized to form P-P and P-H bonds.

To assess whether any of these three metastable phases could be a reasonable candidate for superconductivity, we calculated the Eliashberg spectral functions and determined the electronphonon (ep) interaction. Our calculated behavior of T_c with respect to pressure shows a fair agreement with experiments for all three structures (PH, PH₂ and PH₃); the best agreement is found for PH₂, which has a T_c of 40 K at 100 GPa that increases under pressure and reaches a maximum value of 70 K at 200 GPa. The similar superconducting behavior of the three compounds, however, masks significant differences in the Eliashberg functions, which are reflected in a different behavior of $ep(\lambda)$ and ω_{log} (coupling frequency), this due to the different vibrational character of H atom.

In conclusion, the phase stability and superconducting properties of the recently reported phase in PH₃ under pressure have been investigated with ab initio calculations. Our extensive structural searches show that the hydrogenrich phosphorous phases are thermodynamically unstable. While including vibrational zero-point effects to our calculations improves the stability of PH1, PH2 and PH3, the hybrid HSE06 functional favors decomposition, rendering all binary phases metastable. Nevertheless, the low-enthalpy ground states at fixed composition may have been synthesized through non-equilibrium effects and stabilized by strain. Several phases were predicted to be good phonon-mediated superconductors and could thus in principle account for the measured high- T_c , in experiments. In our opinion, the phase that yields the best agreement with experiments is PH₂, which is both close to the convex hull and reproduces the evolution of the experimentally measured T_c . Since it is unclear, however, if the experimentally observed high- T_c can be fully attributed to a phase of a binary phosphorus-hydride, further experiments are called for to validate both our predictions and the results of Drozdov *et al.* Our results have been submitted to *Physical Re*-

view B, Rapid Communication [14].

8.4 Planned research for next year

The implementation of an automate determination of superconducting properties for materials "on flight" while being explored by crystal prediction methods will allow us to correlate structure, dynamical properties (phonons) and electronic structure in order to better search for patterns (high- T_c) in the large space of materials.

8.5 Synergies with other computational and experimental efforts

a) *Outside MARVEL* This work is done in collaboration with the group of Prof. Gross (Max-Planck Institute in Halle, Germany), Prof. Boeri (Institute of Theoretical and Computational Physics, Graz University of Technology, Austria), Prof. Wolverton (Department of Materials Science and Engineering, Northwestern University, USA) and Prof. Profetta (Dipartimento di Fisica Università degli Studi di L'Aquila, Italy).

8.6 Personnel

Matching personnel

 José A. Flores-Livas, postdoc, 50%, from May to July 2015.

9 Atomic-scale analysis of SiC-oxide interface for improved high-power MOSFETs (Stefan Goedecker — UniBas)

9.1 Research summary

Conformational sampling algorithms on the *ab initio* and more approximate levels were used to study SiC-SiO₂ interfaces and different interface defects, e.g. dangling bonds, carbon clusters and near-interface oxide traps. We found that the carbon clusters at the interface can create states both near the conduction band and the valence band. We also worked towards a better understanding of the thermal oxidation process of 4H-SiC crystals by simulating the early phases of oxidation and the creation of the first oxide layers as well as different carbon removal pathways in the later stages of oxidation.

9.2 Scientific goals

a) *Short-term* Exploration of the SiC-SiO₂ interface structures, and understanding the origin of interface states.

b) *Long-term* Based on the understanding of the atomistic processes, new passivation techniques for the SiC-SiO₂ interface may be developed.

9.3 Results obtained since the last report

a) Verifications of different density functionals The band gap is a critical property for studying the SiC/SiO₂ interface. The PBE functional fails in reproducing the experimental band gaps for 4H-SiC as it does for most materials. One way to deal with this problem is to use hybrid density functionals, which contain an explicit contribution of Hartree-Fock exchange. However, the computational cost for hybrid functionals is very high. Another way is to use meta-GGAs, which can yield band gaps with an accuracy similar to hybrid functional or the GW method, but computationally less expensive. We compared the SiC band gaps obtained from PBE [55], PEB0 [56], HSE [57], and meta-GGA method [58] (modified Becke-Johnson exchange potential in combination with L(S)DA-correlation (MBJ)). The SiC band gap obtained with HSE (3.2 eV) and meta-GGA (3.1 eV) is significantly closer to the experimental value (3.3 eV) than the one obtained with PBE (2.2 eV) and PBE0 (4.0 eV). The meta-GGA can give a reasonable band gap value and the calculation is cheaper. Therefore, the electronic properties in this study are mainly obtained with the meta-GGA method.

b) The pure interface To investigate the SiC/SiO_2 interface, we first performed the simulations of the pure interfaces. We could find a clear band gap there, but no dangling bonds in the interface (Fig. 17).

c) *SiC/SiO*₂ *defects:* (*A*) *dangling bonds,* (*B*) *ox-ide traps,* (*C*) *carbon clusters* Dangling bonds, oxide traps and carbon clusters are believed to be the most dominant contributions to interface states in the SiC/SiO₂ system.

(A) **Dangling bonds:** Here, we have started with a structure where a few dangling bonds



Figure 17: The structure (inset) and DOS (right) of pure *SiC/SiO*₂ interface with no dangling bonds.



Figure 18: (a) The structure (inset) and DOS of SiC/SiO₂ interface with oxygen trap. (b) The structure (inset) and DOS of SiC/SiO₂ interface (with oxygen trap) passivated by nitrogen. The interface states are shown in (a) by a yellow arrow.

were manually constructed at the interface. The resulting defect states can be found in the band gap. These defect states can be passivated easily by adding hydrogen atoms. After adding the hydrogen atoms, the defect states disappeared.

(B) Oxide traps: For the oxygen traps (missing oxygen atoms in SiO_2 , the calculations results revealed that the resulting states in the gap cannot be eliminated by hydrogen. However, they disappear after the additional nitrogen atoms are added in the interface region (Fig. 18). This means that nitrogen can be used to passivate the interface states.

(C) Carbon clusters: To simulate the behavior of carbon clusters at the SiC/SiO₂ interfaces, we added C_4 , C_6 , C_8 , and C_{10} clusters close to the interfaces. After the minima hopping runs, we got the stable interface structures with carbon clusters of different sizes. For the interface with C_4 cluster, we used a unit cell which contains 160 atoms. And we used much larger unit cells to simulate the interface structures



Figure 19: The DOS of SiC/SiO₂ interface structures with C_4 (a), C_6 (b), C_8 (c) and C_{10} (d) clusters, showing interface states.

with C_6 (382 atoms/cell), C_8 (384 atoms/cell) and C_{10} (386 atoms/cell) clusters. From the results of the simulations, it can be found that all the structures have some states in the band gap (Fig. 19). The carbon clusters may be the strong candidates for the defects in the SiC/SiO₂ interfaces. In the next step, we will simulate the interface after introducing carbon clusters with varying sizes and shapes in order to identify states in the band gap. The gained understanding will help to develop passivating methods that eliminate states in the band gap.

d) *SiC oxidation: different competing processes* The motivation behind this part of our theoretical work was to address long-standing questions related to the oxidation process of SiC and to understand the dynamics during the process. There have been a lot of ef-



Figure 20: Formation of the first layer of SiO_2 on the Si-terminated (0001) surface of 4H-SiC. The snapshots are the DFT-optimized conformations found along a minima hopping search trajectory on DFTB potential energy surface.



Figure 21: Carbon clusters at the interface between SiO₂ and 4H-SiC found during the early phases of SiO₂ layer formation.

forts to understand the SiC oxidation better - these include a few theoretical studies too. Knaup and coworkers investigated the mechanism of SiC oxidation with a combination of SCC-DFTB and DFT methods using a model of the SiC/SiO₂ interface. It has been found that carbon dimers are formed as intermediate complexes of SiC oxidation. Here we attempt to go beyond the energetics of the reactants and would like to investigate into the kinetic factors, e.g. activation barriers along the reaction pathways, and look for novel products and pathways, formation of intermediate layers in the oxidation process which the previous studies or model like Deal-Grove or the variants might have missed.

The formation of the first SiO_2 layer: During thermal oxidation, the penetration of oxygen in SiC is considered to be the rate-limiting step for the dry-oxidation process. At the outset we wanted to understand the first SiO_2 layer formation and subsequently how more oxygen diffuses through this layer and how the boundary of SiC moves towards the SiC bulk as SiO_2 keeps on growing (Fig. 20).

Formation of Carbon clusters in the interface: For the Si-terminated (0001) surface of SiC, the oxygen atoms seem to bind to the first silicon layer quite easily and form a Si-O-Si bridge with the next layer. Close to the interface, carbon atoms are found to align in a chainlike conformation (Fig. 21). It has been observed in different independent runs and these conformations are among the lowest ones in energy.

9.4 Personnel

NCCR personnel

• Li Zhu, postdoc, 100%, from August 2014 to November 2015.



Matching personnel

• Shantanu Roy, postdoc, 50%, from May 2015.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- [1] O. Valsson and M. Parrinello, Variational Approach to Enhanced Sampling and Free Energy Calculations, Physical Review Letters 113, 090601 (2014).
- [2] M. Salvalaglio, P. Tiwary, M. Mazzotti, and M. Parrinello, Overcoming timescale and finite-size limitations to compute condensation rates at physically relevant conditions, arXiv:1508.01642 (2015).
- [3] O. Valsson and M. Parrinello, Well-Tempered Variational Approach to Enhanced Sampling, Journal of Chemical Theory and Computation 11, 1996 (2015).
- [4] J. McCarty, O. Valsson, P. Tiwary, and M. Parrinello, Variationally Optimized Free-Energy Flooding for Rate Calculation, Physical Review Letters 115, 070601 (2015).
- [5] P. Shaffer, O. Valsson, and M. Parrinello, Enhanced, targeted sampling of high dimensional free energy landscapes using variationally-enhanced sampling, with an application to chignolin, to be published in Proceedings of the National Academy of Science of the USA (2016).
- [6] P. Gasparotto and M. Ceriotti, Recognizing molecular patterns by machine learning: An agnostic structural definition of the hydrogen bond, The Journal of Chemical Physics 141, 174110 (2014).
 - [7] M. Amsler, S. Botti, M. A. L. Marques, T. J. Lenosky, and S. Goedecker, *Low-density silicon allotropes for photovoltaic applications*, Physical Review B **92**, 014101 (2015).
 - [8] R. Petraglia, A. Nicolaï, M. D. Wodrich, M. Ceriotti, and C. Corminboeuf, *Beyond Static Structures: Putting Forth REMD as a Tool to Solve Problems in Computational Organic Chemistry*, Journal of Computational Chemistry **37**, 83 (2016).
- [9] N. J. Browning, R. Ramakrishnan, A. O. von Lilienfeld, and U. Rothlisberger, Addressing the selection bias: Genetically optimized molecular data sets for the training of machine learning models of molecular enthalpies of atomization, in preparation (2015).
 - [10] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, *Quantum chemistry structures and properties* of 134 kilo molecules, Scientific Data 1, 140022 (2014).
 - [11] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, *Big Data Meets Quantum Chemistry Approximations: The* Δ-*Machine Learning Approach*, Journal of Chemical Theory and Computation **11**, 2087 (2015).
- [12] C. Perego, M. Salvalaglio, and M. Parrinello, Molecular dynamics simulations of solutions at constant chemical potential, The Journal of Chemical Physics 142, 144113 (2015).
- [13] M. Salvalaglio, C. Perego, F. Giberti, M. Mazzotti, and M. Parrinello, *Molecular-dynamics simulations of urea nucleation from aqueous solution*, Proceedings of the National Academy of Science of the USA 112, E6 (2015).
- [14] J. A. Flores-Livas, M. Amsler, C. Heil, A. Sanna, L. Boeri, G. Profeta, C. Wolverton, S. Goedecker, and E. K. U. Gross, Superconductivity in metastable phases of phosphorus-hydride compounds under high pressure, Physical Review B 93, 020508 (2016).

Other references

- [15] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, *PLUMED 2: New feathers for an old bird*, Computer Physics Communications 185, 604 (2014).
- [16] A. F. Voter, Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events, Physical Review Letters 78, 3908 (1997).
- [17] P. Tiwary and M. Parrinello, From Metadynamics to Dynamics, Physical Review Letters 111, 230602 (2013).
- [18] A. Sadeghi, S. A. Ghasemi, B. Schaefer, S. Mohr, M. A. Lill, and S. Goedecker, *Metrics for measuring distances in configuration spaces*, The Journal of Chemical Physics **139**, 184118 (2013).
- [19] A. R. Oganov and M. Valle, *How to quantify energy landscapes of solids?*, The Journal of Chemical Physics 130, 104504 (2009).
- [20] Y. Wang, J. Lv, L. Zhu, and Y. Ma, CALYPSO: A method for crystal structure prediction, Computer Physics Communications 183, 2063 (2012).
- [21] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, The Journal of Chemical Physics 134, 074106 (2011).
- [22] M. Amsler and S. Goedecker, *Crystal structure prediction using the minima hopping method*, The Journal of Chemical Physics 133, 224104 (2010).
- [23] B. Aradi, B. Hourahine, and T. Frauenheim, DFTB+, a Sparse Matrix-Based Implementation of the DFTB Method, The Journal of Physical Chemistry A 111, 5678 (2007).
- [24] T. J. Lenosky, J. D. Kress, I. Kwon, A. F. Voter, B. Edwards, D. F. Richards, S. Yang, and J. B. Adams, *Highly* optimized tight-binding model of silicon, Physical Review B 55, 1528 (1997).
- [25] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials*, Journal of Physics: Condensed Matter **21**, 395502 (2009).
- [26] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, Physical Review B 87, 184115 (2013).
- [27] M. Ceriotti, G. A. Tribello, and M. Parrinello, From the Cover: Simplifying the representation of complex freeenergy landscapes using sketch-map, Proceedings of the National Academy of Science of the USA 108, 13023 (2011).
- [28] G. A. Tribello, M. Ceriotti, and M. Parrinello, Using sketch-map coordinates to analyze and bias molecular dynamics simulations, Proceedings of the National Academy of Science of the USA 109, 5196 (2012).
- [29] F. Giberti, M. Salvalaglio, M. Mazzotti, and M. Parrinello, *Insight into the nucleation of urea crystals from the melt*, Chemical Engineering Science **121**, 51 (2014).
- [30] M. Salvalaglio, T. Vetter, F. Giberti, M. Mazzotti, and M. Parrinello, Uncovering Molecular Details of Urea Crystal Growth in the Presence of Additives, Journal of the American Chemical Society 134, 17221 (2012).

- [31] M. Salvalaglio, T. Vetter, M. Mazzotti, and M. Parrinello, *Controlling and Predicting Crystal Shapes: The Case of Urea*, Angewandte Chemie International Edition 52, 13369 (2013).
- [32] R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, *Nucleation mechanism for the direct* graphite-to-diamond phase transition, Nature Materials 10, 693 (2011).
- [33] S. Goedecker, Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems, The Journal of Chemical Physics 120, 9911 (2004).
- [34] L. Rapp, B. Haberl, C. J. Pickard, J. E. Bradby, E. G. Gamaly, J. S. Williams, and A. V. Rode, *Experimental* evidence of new tetragonal polymorphs of silicon formed through ultrafast laser-induced confined microexplosion, Nature Communications 6, 7555 (2015).
- [35] S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network, Physical Review B 92, 045131 (2015).
- [36] B. Schaefer, S. Mohr, M. Amsler, and S. Goedecker, Minima hopping guided path search: An efficient method for finding complex chemical reaction pathways, The Journal of Chemical Physics 140, 214102 (2014).
- [37] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, International Journal of Quantum Chemistry 115, 1032 (2015).
- [38] M. Ceriotti, J. More, and D. E. Manolopoulos, *i-PI:* A Python interface for ab initio path integral molecular dynamics simulations, Computer Physics Communications 185, 1019 (2014).
- [39] L.-C. Lin, A. H. Berger, R. L. Martin, J. Kim, J. A. Swisher, K. Jariwala, C. H. Rycroft, A. S. Bhown, M. W. Deem, M. Haranczyk, and B. Smit, *In silico screening of carbon-capture materials*, Nature Materials 11, 633 (2012).
- [40] C. Simon, J. Kim, D. Gomez-Gualdron, J. Camp, Y. G. Chung, R. L. Martin, R. Mercado, M. W. Deem, D. Gunter, M. Haranczyk, D. Sholl, R. Q. Snurr, and B. Smit, *The Materials Genome in Action: Identifying the Performance Limits for Methane Storage*, Energy & Environmental Science 8, 1190 (2015).
- [41] C. M. Simon, R. Mercado, S. K. Schnell, B. Smit, and M. Haranczyk, What Are the Best Materials To Separate a Xenon/Krypton Mixture?, Chemistry of Materials 27, 4459 (2015).
- [42] N. Pellet, P. Gao, G. Gregori, T.-Y. Yang, M. K. Nazeeruddin, J. Maier, and M. Grätzel, *Mixed-Organic-Cation Perovskite Photovoltaics for Enhanced Solar-Light Harvesting*, Angewandte Chemie International Edition 53, 3151 (2014).
- [43] N. J. Jeon, J. H. Noh, W. S. Yang, Y. C. Kim, S. Ryu, J. Seo, and S. I. Seok, *Compositional engineering of per*ovskite materials for high-performance solar cells, Nature 517, 476 (2015).

- [44] C. C. Stoumpos, C. D. Malliakas, and M. G. Kanatzidis, Semiconducting tin and lead iodide perovskites with organic cations: phase transitions, high mobilities, and near-infrared photoluminescent properties, Inorganic Chemistry 52, 9019 (2013).
- [45] S. Mastroianni, F. D. Heinz, J.-H. Im, W. Veurman, M. Padilla, M. C. Schubert, U. Würfel, M. Grätzel, N.-G. Park, and A. Hinsch, *Analysing the effect of crystal size and structure in highly efficient CH*₃NH₃PbI₃ perovskite solar cells by spatially resolved photo- and electroluminescence imaging, Nanoscale 7, 19653 (2015).
- [46] A. Mattoni, A. Filippetti, M. I. Saba, and P. Delugas, Methylammonium Rotational Dynamics in Lead Halide Perovskite by Classical Molecular Dynamics: The Role of Temperature, The Journal of Physical Chemistry C 119, 17421 (2015).
- [47] A. Barducci, G. Bussi, and M. Parrinello, Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method, Physical Review Letters 100, 020603 (2008).
- [48] M. Salvalaglio, P. Tiwary, and M. Parrinello, Assessing the Reliability of the Dynamics Reconstructed from Metadynamics, Journal of Chemical Theory and Computation 10, 1420 (2014).
- [49] G. Chkonia, J. Wölk, R. Strey, J. Wedekind, and D. Reguera, *Evaluating nucleation rates in direct simulations*, The Journal of Chemical Physics **130**, 064505 (2009).
- [50] R. Angélil, J. Diemand, K. K. Tanaka, and H. Tanaka, Homogeneous SPC/E water nucleation in large molecular dynamics simulations, The Journal of Chemical Physics 143, 064507 (2015).
- [51] B. Widom, Some Topics in the Theory of Fluids, The Journal of Chemical Physics 39, 2808 (1963).
- [52] S. G. Moore and D. R. Wheeler, *Chemical potential perturbation: A method to predict chemical potentials in periodic molecular simulations*, The Journal of Chemical Physics **134**, 114514 (2011).
- [53] A. P. Drozdov, M. I. Eremets, and I. A. Troyan, Superconductivity above 100 K in PH₃ at high pressures, arXiv:1508.06224 (2015).
- [54] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin, *Conventional superconductivity* at 203 kelvin at high pressures in the sulfur hydride system, Nature Materials 525, 73 (2015).
- [55] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Physical Review Letters 77, 3865 (1996).
- [56] M. Ernzerhof and G. E. Scuseria, Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional, The Journal of Chemical Physics 110, 5029 (1999).
- [57] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, The Journal of Chemical Physics **118**, 8207 (2003).
- [58] A. D. Becke and E. R. Johnson, A simple effective potential for exchange, The Journal of Chemical Physics 124, 221101 (2006).



5 Horizontal Project



Project leader: Alessandro Curioni (IBM)

Participating members: Alessandro Curioni (IBM), Christoph Koch (EPFL), Volkan Cevher (EPFL), Anatole von Lilienfeld (UniBas)

Summary and highlights: The work in Horizontal Project 5 (HP5) focused on the design of big-data system, merging the expertise on cognitive, neural networks and machine learning, to predict novel materials connected to the MARVEL activities and on the development of domain-specific language frameworks that will allow the easy development of high-performance applications in the context of MARVEL. In particular, we aim at complementing in a synergistic way recent developments on machine learning and neural networks with cognitive computing to support the activity of the Vertical Projects and help in either predicting unknown properties of materials or even construct new ones based on all the properties of similar class of materials known to us at the present day. In particular, in the second year we were able to develop the framework for a material discovery system that is built around a knowledge graph that can support complex queries such as semantic search. In addition, we developed advanced machine learning and statistical inference algorithms that are crucial to populate the knowledge graph and to allow prediction of properties. We stress that in year 2 we initiated a deep collaboration with VP2 to showcase our approach in the context of novel materials for batteries. Finally, work towards (materials) domain specific languages has quite progressed at the foundational level to now allow practical instantiations.

General view of the project

While in the first year, we set the outline and the general objectives of how we can leverage computer science research, and specifically big-data-systems research, to support the efforts of the MARVEL community, in this second year we started to combine the different expertise (cognitive, machine learning and neural networks) into a unique platform to bring forward the design of a platform able to support scientists of the Vertical Projects in the discovery of novel materials. We started to explore possible pipelines with the MARVEL platform AiiDA, with the goal to integrate these pipelines within the material discovery platform and test/validate it on the project of solid-state electrolytes of the VP2. Concurrently, we continued to study the use of domain-specific languages (DSLs) and scripting languages in computational materials science to develop specialized high-level and high-productivity programming languages for the computational materials science community.

Major research questions

The key research questions that we target are as follows.

• The use of cognitive and analytics techniques for boosting materials prediction

within MARVEL.

- The design of specialized high-level programming and scripting languages for the computational materials science community that make MARVEL teams more productive in their coding efforts. How can we make programs developed in such languages run highly efficiently on highperformance computing platforms?
- Accelerate reasoning based on deep neural networks so as to render them a powerful and useful machine learning technique for the MARVEL community.
- The exploitation of property trends across chemical compound space is at the heart of the computational design of new materials. Can we use a machine learning approach to leverage this information in the material discovery process?

Long-term goals

In the next year, we plan to include other rich data-sources such as Wikipedia, PhD theses (e.g. from EPFL and ETHZ) and targeted journals and books, given we get the permission from the publisher. Furthermore, we plan to broaden our annotation tools, such that we

not only extract sum-formulas of new crystallographic structures, but also are able to detect property-names and values. Examples of the latter could be the conductivity of a certain material or its critical temperature of its phase transition.

In terms of analytics, in the long term, we plan to develop and implement advanced inference techniques as well as a direct link between the knowledge graph of materials and simulation software such as CPMD.

Our long-term goal is to introduce the use of domain-specific languages (DSLs) and scripting languages in computational materials science and develop specialized high-level and high-productivity programming languages for the computational materials science community. We will extensively evaluate the potential of domain-specific compilation to automatically generate HPC code that is competitive in performance with the best code achievable by expert human programmers. If another major effort to develop a performance critical quantum simulation or machine learning code base is started during the duration of MARVEL, we will attempt to get involved in this effort to see whether our techniques can lead to more highlevel implementations at higher productivity without a performance penalty.

We would like to take a two-pronged approach in improving our predictive capabilities for materials discovery. The first one relates to the computational scalability for Kernel based methods, which currently outperform the neural network based approaches. In the second approach, we would like to develop active sampling schemes that work in conjunction with our predictive setup. The idea is to leverage the existing materials simulations and use this as guidance in sampling new configurations that have attractive properties.

In terms of the work in predicting material properties, in the long run we plan to simultaneously explore multitudes of compounds, and to iteratively navigate the corresponding materials space in order to converge upon new materials which exhibit desirable properties. Our approach will enable the use of optimization methods with unprecedented efficiency. This will enlarge the materials space which we explore.

Achievements

Regarding the design of a platform for supporting the discovery of new materials, we made sensible improvements with regard to the parsing and annotation of specific datasources. We have successfully parsed the ICSD-database (170 K materials), all arXivpapers (600 K papers) and all US-patents (10 M patents). We have annotated all papers and patents by searching for mentions of sumformulas of materials. In this way, we found a surprising number of materials that were not yet included in the ICSD, allowing us to enrich the ICSD-database. Concurrently, we focused on the theoretical development of cognitive algorithms as well as on a practical implementation with regard to the search of new materials for batteries.

We have been developing a compilation framework for DSLs that allows to plug in very powerful domain-specific compiler optimizations to produce code that is competitive in performance with the most efficient handwritten low-level code developed by HPC experts.

The design and discovery of new materials can be accelerated using machine learning techniques. In this quarter, we have demonstrated this idea by predicting atomization energies using [1].

Our previous short-term goal has been to assess the applicability, transferability, and predictive power of alchemical derivatives [8, 9] of first principles properties in the compositional space of solid materials. To this end, we have first investigated simple materials, such as alkali-halogenide crystals. We have established a good predictive power of energetics using first order Hellmann-Feynman derivatives [10, 8] for alkali halides.

Milestones

The milestones for the next two years are as follows.

April 2016

- Release the first version of the material discovery knowledge graph.
- Release the first version of the DSL compiler framework.
- Finalize study on alchemical predictions of alkali halogenide solids and on machine learning based predictions of main group quaternary crystals.
- Set up a scalable convex optimization framework for new learning formulations in materials discovery. The new formulations will enable a host of new regularization functions, including gradient regularization along with various different kernels and will be robust to outliers in the training data.



April 2017

- Advanced search on the knowledge graph including inference.
- Proof of concept for symbolic computing in the materials discovery space.
- Extend the alchemical study of solids to also include earth-alkali and chalcogenide elements.
- Demonstrate improvements in learning quality with convex optimization. As an alternative, develop new, non-convex neural network approaches for the same class of problems.

April 2018

- Release of materials design cognitive assistant for MARVEL related activities.
- Identify two additional use-cases for domain specific compilation with MARVEL teams.
- Apply resulting alchemical methods to explore phase diagrams of solid mixtures.
- Provide new rotation and shift invariant representations for accelerating materials discovery that exploit the scalable convex or non-convex optimization machinery in the previous milestones.

New and ongoing collaborations

Cognitive material design A collaboration with VP2 is in progress since the beginning of 2015

for the study of electrolytes. This is planned to remain and intensify for the next two years (2015-2018). Outside MARVEL we plan a new collaboration with STFC (Hartree Center in the UK) for cognitive material design acceleration. The emphasis and particular Hartree Center group leader is yet to be determined, however, there is an agreement at the Lab leadership level (2016-2018).

Domain specific languages Ongoing collaborations since the beginning of the project with the AiiDA team (lead by Nicola Marzari) as well as with the infrastructure team (led by Thomas Schulthess) (2014-2018). New planned collaboration with computer science groups in EPFL planned. Group leaders to be determined (2016-2018).

Acceleration of machine learning Ongoing collaborations within MARVEL with the groups of Ceriotti, Curioni, von Lilienfeld, and Marzari. The emphasis is on new and improved machine learning kernels tailored for the topics akin in MARVEL (2015-2018).

Prediction of chemical properties Ongoing collaborations within MARVEL with the groups of Goedecker (alanate crystals), Röthlisberger (tuning of molecular properties prediction) (2015-2018). Ongoing collaborations outside MARVEL with the groups of Andrienko (Max Planck, Mainz) on the design of superior molecular electronic devices, Keith (Univ. of Pittsburgh) on alchemical couplings for the computational screening trough the different alloy combinations and Armiento (Lynkoping Univ.) on prediction of crystal energy. All the above between 2015-2018.

1 Data-driven approach to the prediction of new materials (Alessandro Curioni — IBM)

1.1 Research summary

In this horizontal project, we apply a datadriven approach for the discovery of new materials. In particular, we target to predict unknown properties of materials or even construct new materials based on all the properties of similar class of materials known to us at the present day. Such an approach requires essentially two main components. First, one has to have advanced parsing and annotating algorithms to extract material-related data from a variety of documents, such as books, articles, PhD theses, patents, webpages, etc. Second, one needs to have fast and scalable algorithms, that can handle the large amount of data and make sensible predictions on the properties of materials. Given that we have already good progress and experience with these components, the targeted research questions in this horizontal project are of the kind: *Given properties A, B and C with values* α , β , γ , *find all relevant materials that satisfy these property-values?* or "*Given that material A, B and C have respectively property-values* α , β , γ , *can we predict what the property-value* δ *is for materials D?*" or "*Given a family of materials with similar atomic structure, can we find new members?*".

1.2 Scientific goals

The scientific goals of this horizontal project are essentially to develop and implement the two main components described in the previous subsection, i.e. a parsing and annotating component for technical documents and an an-



Figure 1: Overview of current accomplished goals as well as the future goals of Horizontal Project 5.

alytics component for making sensible predictions. We have made significant progress on both fronts over the last year.

- a) Short-term
 - Parsing and annotation: In Fig. 1, we show our progress and planned accomplishments with regard to the parsing and annotation of specific data-sources. we have successfully parsed the ICSD-database (170 K materials), all arXiv-papers (600 K papers) and all US-patents (10 M patents). We have annotated all papers and patents by searching for mentions of sum-formulas of materials. In this way, we found a surprising number of materials that were not yet included in the ICSD, allowing us to enrich the ICSD-database.
 - Analytics: In this section, we have focused on the theoretical development of cognitive algorithms as well as on a practical implementation with regard to the search of new materials for batteries. The latter was done in close collaboration with VP2, which is concerned with fundamental research on batteries. On the algorithmic side, we have investigated how to do quantitative searches on large knowledge graphs. We came to the conclusion that the subgraph-centrality algorithm was a suitable candidate, since it involves the full knowledge graph in the query (contrary to simple filters) and is therefore a much more robust search-algorithm. We have researched intensively how to implement this new algorithm efficiently on accelerators as well as how to scale it on distributed compute systems. This resulted in a paper, currently submitted to IPDPS'16 [2].

- b) Long-term
 - Parsing and annotation: In the next year, we plan to include other rich data-sources such as Wikipedia, PhD theses (e.g. from EPFL and ETHZ), and targeted journals and books, given we get the permission from the publisher. Furthermore, we plan to broaden our annotation tools, such that we not only extract sum-formulas of new crystallographic structures, but also are able to detect property-names and values. Examples of the latter could be the conductivity of a certain material or its critical temperature of its phase transition.
 - Analytics: In the long term, we plan to develop and implement advanced inference techniques as well as a direct link between the knowledge graph of materials and simulation software such as CPMD. In this way, we can use the inference algorithms to predict new materials, based on desired properties. By having a direct link to simulation software, we can then easily verify the existence of the material as well as its properties. The result of these simulations can then be fed back into the knowledge graph of materials to increase the accuracy of predictions.

1.3 Results obtained since the last report

The plan of the past year was to set up an initial knowledge graph of materials and to have implemented simple algorithms to query



Figure 2: Web-based front-end of the materials knowledge graph. Using a "familiar" search bar, one can search materials as well as any type of literature related to it.

the latter. As discussed in detail in the previous section, we have populated the knowledge graph with 170 K materials from the ICSDdatabase and enriched it with all the papers from the public arXiv-repository and as well as all patents from the US-patent office. To query the graph, we have implemented two functionalities, i.e. an advanced search functionality, which allows us to search through all the literature. This search is based on the subgraph-centrality [2]. Furthermore, we have also implemented an initial correlation module, which allows us to correlate quantitative attributes of all ingested materials (e.g. its density vs the volume of the unit cell). As it is shown in Fig. 1, the graph can be queried using a web-based front-end as well as through a IPython-notebook. The latter is very convenient for rapid prototyping. In Fig. 2, we show a screenshot of the web-based front-end.

1.4 Planned research for next year

In the next period, we plan to reach two goals. First, we want to improve the annotation of the properties of the materials. This will enrich dramatically our materials knowledge graph. Second, we plan to develop and implement advanced inference techniques. The research of these inference techniques will be in close collaboration with the research in VP2.

1.5 Synergies with other computational and experimental efforts

a) Within MARVEL With regard to the use of cognitive algorithms to help the battery research of VP2, we are in the process of identifying characteristics features of the known batteries. Combined with data from the kinetic Monte Carlo simulations, we plan to exploit these characteristics features by searching in the materials database to other materials with similar features. This is work in progress. We refer the reader to the report of VP2 for technical details with regard to the simulations and characteristic features.

1.6 Personnel

NCCR personnel

• Peter Staar, postdoc, 75%, from January 2015.

Matching personnel

• Costas Bekas, senior researcher, 25%, from January 2015.

2 "Alchemical" coupling and machine learning predictions (Anatole von Lilienfeld — UniBas)

2.1 Research summary

The exploitation of property trends across chemical compound space is at the heart of the computational design of new materials. We have carried out numerical calculations for various crystalline systems in order to establish a property trends through use of "alchemical" coupling between materials. For alkali halide crystals we have found that first order derivatives with respect to coupling parameter predict property trends across materials compound space with unprecedented efficiency and accuracy. We also developed machine learning models for the prediction of crystal formation energies.

2.2 Scientific goals

a) *Short-term* Our next short-term goal is the extension of the methods developed to explore phase diagrams of binary and tertiary mixtures of alkali halide crystals (alchemical coupling), as well as other crystals and properties (machine learning).

b) *Long-term* In the long run, we plan to simultaneously explore multitudes of compounds, and to iteratively navigate the corresponding materials space in order to converge upon new materials which exhibit desirable properties. Our approach will enable the use of optimization methods with unprecedented efficiency. This will enlarge the materials space which we explore.

c) Alchemical coupling The potential energy of a material, the most basic property one would like to control, can be Taylor expanded in the space of all possible compounds and materials,

$$E(\lambda) = E_0 + \frac{\partial E_0}{\partial \lambda} \Delta \lambda + \frac{1}{2} \frac{\partial^2 E_0}{\partial \lambda^2} \Delta \lambda^2 + \text{HOT} \quad (1)$$

where we have truncated the higher order terms (HOT). The coupling parameter λ can be chosen to relate any two materials, represented by their respective Hamiltonians. This includes systems with identical number of valence electrons, such as NaCl and KCl. The Hamiltonian could, for example, read,

$$\hat{H}(\lambda) = \hat{H}_{\text{NaCl}} + \lambda (\hat{H}_{\text{KCl}} - \hat{H}_{\text{NaCl}})$$
(2)

and its first order derivative is given analytically by virtue of the Hellmann-Feynman theorem [10, 8],

$$\frac{\partial E}{\partial \lambda}\Big|_{\lambda} = \left\langle \frac{\partial \hat{H}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} = \left\langle \hat{H}_{\text{KCl}} - \hat{H}_{\text{NaCl}} \right\rangle_{\lambda} \quad (3)$$

where $\langle ... \rangle$ corresponds to the Bra-Ket notation of the quantum mechanical observable. In practice, first order derivatives can be calculated by simply interpolating the external potential. Pulay-forces can be avoided, and elements from different periods in the periodic table can be coupled through use of interpolated pseudopotentials in a plane-wave basis set.

d) *Machine learning* Our machine learning (ML) model is based on kernel ridge regression [11] which maps the nonlinear energy difference between a computationally demanding quantum mechanics based energy estimate and an inexpensive approximate baseline model into a linear feature space. More specifically, we construct a ML model of the energy difference to the sum of static, atom-type dependent, atomic energy contributions ϵ_{It} , obtained through fitting of each atom type *t* in all chemical elements. The energy-predicting function of any crystal is then a sum of weighted exponentials in similarity *d* between query and training crystal,

$$E(\mathbf{x}) = \sum_{I}^{N'} \epsilon_{It} + \sum_{i}^{N} \alpha_{i} e^{-d_{i}/\sigma}$$
(4)

where N' is the number of atoms/unit cell, and the second sum runs over all N training instances.

2.3 Results obtained since the last report

Our previous short-term goal has been to assess the applicability, transferability, and predictive power of alchemical derivatives [8, 9] of first principles properties in the compositional space of solid materials. To this end, we have first investigated simple materials, such as alkali-halogenide crystals. We have established a good predictive power of energetics using first order Hellmann-Feynman derivatives [10, 8] for alkali halides.

a) *Coupling alkali halides* In order to calculate all possible alchemical transmutations for alkali halide crystals, we first carried out density functional theory calculations for all rock salt (rs) (64 atoms supercell) and simple cubic (cscl) (54 atoms supercell) phases. We considered all the 16 combinations of Na, K, Rb, Cs, and F, Cl, Br, and I for a range of lattice constants at the Γ -point only. All calculations have been performed using the pseudopotential plane-wave based code CPMD [12], the LDA and PBE functionals [13], and Goedecker-Teter-Hutter pseudopotentials, as parameterized by Krack [14]. All alkali and halide pseudopotentials account respectively for nine and seven valence-electrons. A major part of the calculations has been carried out thanks to access to the Piz Dora at the CSCS computational center.

Using first order derivatives we predicted energies in rs and cscl phases for all possible $(16 \times 15 = 240)$ alkali halide transmutations using LDA and PBE functionals. All resulting potential energy curves are shown in Fig. 3. For all alkali chlorides we also included the zincblende (zb) phase in all these predictions. Corresponding alchemical predictions of lattice constants, energy differences between the phases, and bulk moduli are specified in Table 1. The error of predicting lattice constants and energy differences varies from 1.0 to 13.0%. The bulk moduli predictions are in general less accurate, but remain reasonable: the error ranges from 3.0 to 40.0%.

Using NaCl as a starting compound, alchemical predictions of zb, rs, and cscl crystal structures of KCl are also illustrated in Fig. 4. Here, the bottom row corresponds to the density functional theory (DFT) energies of the reference salt (in this case NaCl). The middle row corresponds to the DFT energies of the targeted salt (KCl). The top row shows the energies estimated using Hellmann-Feynman derivatives. While the alchemically estimated curves of KCl are much higher in total energy than targeted, the location of the energy minimum occurs at a lattice constant which compares relatively well to the actual value for KCl. Furthermore, the predicted relative energies, i.e. the difference between the mutual symmetries (zb-rs, zb-cscl, and rs-cscl), are in very good agreement with the actual energy differences of KCl — for all lattice values. Finally, we stress that the computational cost for obtaining the estimated curves is negligible. They have been obtained by virtue of the analytical Hellmann-Feynman derivatives. No emperical parameters have been used to obtain this result.

While alkali halides typically favor the rs crystal structure, three salts have a lower energy in the cscl structure rather than rs, namely CsCl, CsBr, and CsI. Unfortunately, DFT/PBE fails to correctly account for this energy trend. In [15] it has been pointed out that either the LDA functional, or an empirical London dispersion



Figure 3: Actual and predicted energies [eV] as a function of interatomic distance [Å] for all possible 16 × 16 alkali halide couplings considered. Diagonal plots correspond to actual DFT results, while off-diagonal plots refer to alchemical predictions. There are four lines in each plot: a pink solid line corresponds to rs phase, green solid line to cscl phase, dashed pink and green lines obtained by inclusion of a dispersion correction. Columns and rows indicate the initial and targeted salts, respectively.

correction can results in the correct trend. Using C_6 coefficients from [16, 15], we have reproduced this observation (Table 2).

b) Machine learning models of crystals For ML training and evaluation, we have generated DFT data for elpasolite crystals, a cubic crystal of quaternary composition. We have considered all \sim 2 M crystals containing all maingroup elements up to Bi. The training and testing data consists of 10 k crystal structures, i.e. 0.5% of the total number of 2 M possible crystals. This set has been obtained through random selection of elpasolites ensuring an unbiased composition.

For a training-set size of 10 k, the mean absolute error of the ML model decays to less than 100 meV/atom. Such inaccuracies are comparable to errors one would expect from a typical DFT prediction. The screening of 2 M crystals has enabled us to resolve the Wyckoff-position specific energy contribution each main-group element from the periodic table can make on average. We also identified several crystals with low-lying formation energies despite the fact that elements have unusual oxidation states. This study has recently been submitted for publication [3, 4].

		predicted from				
		NaCl	KC1	RbCl	CsCl	MAE
		latt	ice constar	ıts, Å		
NaCl	rs	5.40	5.90	6.00	6.35	0.68
	cscl	3.35	3.63	3.68	3.85	0.37
	zb	6.05	6.60	6.75	7.15	0.78
KC1	rs	6.10	6.05	6.10	6.35	0.13
	cscl	3.58	3.65	3.70	3.85	0.05
	zb	6.85	6.75	6.95	7.15	0.23
RbC1	rs	6.35	6.40	6.35	6.45	0.05
	cscl	3.80	3.85	3.80	3.90	0.05
	zb	7.05	7.20	7.15	7.20	0.07
CsC1	rs	6.55	6.95	6.75	6.65	0.17
	cscl	3.90	4.18	4.05	3.95	0.13
	zb	7.45	7.80	7.60	7.40	0.27
	rs	0.05	0.28	0.25	0.45	
MAE	cscl	0.04	0.19	0.16	0.27	
	zb	0.08	0.33	0.37	0.52	
		energ	y differenc	es, meV		
	rs-cscl	-617.2	-584.4	-604.7	-588.8	24.6
NaCl	rs-zb	-318.2	-314.7	-329.6	-320.9	5.9
	zb-cscl	-299.0	-269.7	-275.2	-267.9	28.1
	rs-cscl	-542.5	-598.7	-593.1	-588.7	23.9
KC1	rs-zb	-296.2	-314.3	-316.4	-317.8	7.9
	zb-cscl	-246.3	-284.4	-276.7	-270.9	19.8
	rs-cscl	-596.9	-577.4	-575.0	-578.3	9.2
RbC1	rs-zb	-314.5	-304.7	-306.1	-308.5	4.1
	zb-cscl	-282.4	-272.8	-268.9	-269.8	6.1
	rs-cscl	-439.5	-567.4	-569.2	-560.6	45.5
CsCl	rs-zb	-257.0	-295.0	-297.9	-293.3	14.2
	zb-cscl	-182.5	-272.4	-271.2	-267.3	31.3
	rs-cscl	66.4	14.00	8.9	13.9	
MAE	rs-zb	20.9	2.2	6.0	2.9	
	zb-cscl	45.5	12.8	11.8	15.2	
		bulk	moduli, 10	11 dynes		
NaCl	rs	2.317	1.617	1.516	2.068	0.583
	cscl	1.041	0.857	0.775	0.851	0.213
	zb	1.393	1.245	1.285	1.134	0.172
KC1	rs	1.478	1.715	1.726	2.147	0.227
	cscl	1.193	0.884	0.820	0.859	0.133
	zb	0.997	1.136	0.962	1.136	0.129
RbCl	rs	1.368	1.657	1.759	1.988	0.241
	cscl	0.830	0.781	0.828	0.799	0.026
	zb	1.138	0.996	0.944	1.103	0.135
CsCl	rs	1.546	1.291	1.771	1.797	0.261
	cscl	0.937	0.612	0.730	0.852	0.149
	zb	0.956	0.861	0.836	0.919	0.059
	rs	0.293	0.436	0.279	0.303	
MAE	cscl	0.132	0.157	0.123	0.020	
	zb	0.123	0.086	0.122	0.139	

Table 1: Alchemical predictions of alkali chloride properties in rock-salt (rs), CsCl (cscl), and zincblende (zb) crystal structure. Lattice constants (top block), energy differences (middle block), and bulk moduli (bottom block) are shown. The columns indicate the alkali chlorides from which predictions are made, and rows correspond to predicted al-kali chlorides. Diagonal elements correspond to the actual DFT values (PBE) of the row's alkali chloride. Lattice constants and bulk moduli have been obtained by fitting calculated data to the Birch-Murnaghan equation of state. Energy differences refer to the energies at the predicted lattice constants of the corresponding two crystal structure symmetries.

2.4 Planned research for next year

We plan to investigate if the predictive power of first order Hellmann-Feynman derivatives also carries over to systems other than alkali halides.



Figure 4: Potential energy curves for the range of different cell parameters. E_f is the energy of a crystal in rs symmetry, E_b in cscl, and E_{zb} in zincblende. Two curves on the top correspond to the prediction of KCl from NaCl with Hellmann-Feynman derivatives.

- 2.5 Synergies with other computational and experimental efforts
- a) Within MARVEL
 - In collaboration with Stefan Goedecker (UniBas), we have started to perform preliminary test calculations for alanate crystals, recently explored using enhanced sampling methods developed in the Goedecker group [17].
 - In collaboration with Ursula Röthlisberger (EPFL), we started numerical explorative studies of optimizing the composition of training sets for obtaining more efficient machine learning models of molecular properties, extending our preceding work published in [18].
- b) Outside MARVEL
 - In collaboration with Dennis Andrienko at the Max-Planck-Institute for Polymer Research (Mainz, Germany) we have started to perform preliminary test calculations

	CPMD, ΔE , eV/f.u.			VASP [15], ΔE , eV/f.u.		
	PBE	PBE+D2	LDA	PBE	PBE+D2'	LDA
CsF	-0.212	-0.164	-0.128	-0.22	-0.08	-0.11
CsCl	-0.080	0.024	0.078	-0.10	0.16	0.06
CsBr	-0.072	0.026	0.086	-0.08	0.19	0.06
CsI	-0.060	0.016	0.082	-0.06	0.21	0.07

 Table 2:
 Energy difference between rs and cscl phases calculated with LDA, PBE, and PBE+D2

for the application of the alchemical coupling approach towards the design of superior molecular electronic devices.

- In collaboration with John A. Keith at the University of Pittsburgh (PA, USA), we have started an application of alchemical couplings for the computational screening trough the different alloy combinations. The main goal is to find a catalyst with the most favorable hydrogen adsorption energies.
- In collaboration with Rickard Armiento at the University of Lynkoping (SE), we performed the ML crystal energy prediction work [3, 4].

2.6 Personnel

NCCR personnel

- Alisa Solovyeva, postdoc, 100%, from September 2014.
- Zhenwei Li, postdoc, 100%, from January 2015.

Matching personnel

- Diana N. Tahchieva, PhD student, 100%, from January 2015.
- Felix Faber, PhD student, 100%, from November 2014.

3 Domain-specific languages and compilation (Christoph Koch — EPFL)

3.1 Research summary

We are creating a compiler infrastructure for domain-specific languages which allows to easily add domain-specific compiler optimizations for generating highly efficient low-level code.

3.2 Scientific goals

Software libraries for a particular domain are traditionally written as either object-oriented components or subroutine libraries. However, an alternative development methodology, which has recently obtained much attention, is to define a domain-specific programming language (DSL). By using DSLs, programs can be written using the idioms of its particular domain, a fact which improves the productivity of developers as well as the maintainability of software programs themselves.

Like every other programming language, a DSL needs suitable programming tools and compilers that support it. Extensible optimizing compilers are compiler backends that are used to resolve the performance issues of highlevel DSLS. Such compilers provide interfaces in order to specify optimizations that can be applied to a certain DSL. In addition, extensible optimizing compilers provide interfaces for generating code in a different (target) programming language. This way, a program written in a high-level DSL is compiled into a program in a low-level language (such as C). Furthermore, the generated program can reuse the existing libraries written in its own language (for example, if the generated program is in the Python programming language, we can reuse the numerous libraries written in Python for the generated code).

In our experience, existing extensible optimizing compilers are not yet adequate for highperformance computing. The fundamental reason for this is that such systems basically force developers to become compiler experts before developing any software. Thus, in this project, we investigate the infrastructure needed for easy development of DSLs and compilers and programming tools for them. Many people involved in MARVEL develop

code in one way or the other; increasingly it is code in scripting languages that has the purpose of doing pre- and post-processing or providing glue code between several existing code bases. Scripting languages are known to be rather inefficient, and their purpose is often not just to call code in more efficient languages such as C and Fortran but to move substantial amounts of data between these more efficient and tuned code bases, and to transform this data. Such computations should be made efficient in a high-performance computing (HPC) environment. While there exists expertise on optimizing and tuning HPC code in this community, it is a work-intensive and difficult task. General-purpose compilers cannot be relied upon to generate low-level code competitive in performance with expertly hand-optimized code. However, there is recent progress that suggests this can be done well for more restricted languages.

a) *Short-term* We are developing an extensible compiler framework for DSLs.

b) *Long-term* We want to study the use of domain-specific languages (DSLs) and script-

ing languages in computational materials science and develop specialized high-level and high-productivity programming languages for the computational materials science community. This will create real value for the MAR-VEL community in that it will make the various teams more productive in their coding efforts.

Some MARVEL teams such as the AiiDA team are already now using Python as a host language for embedded DSL development.

Some of the code bases are performancecritical, so it is important to have compilers that automatically transform DSL programs into highly efficient code to run on highperformance computing platforms. We will embark on a long-term collaborative effort involving all interested MARVEL teams to codesign DSLs and compiler optimizations for our compiler framework as their needs and challenges develop.

We will extensively evaluate the potential of domain-specific compilation to automatically generate HPC code that is competitive in performance with the best code achievable by expert human programmers. If another major effort to develop a performance critical quantum simulation or machine learning code base is started during the duration of MARVEL, we will attempt to get involved in this effort to see whether our techniques can lead to more highlevel implementations at higher productivity without a performance penalty.

One risk factor to the success of the DSL project is supporting the Python scripting language as a frontend language. Python is an untyped language, which is difficult to automatically produce highly optimized code from. This is a substantial and difficult research challenge. However, it is of wider interest in computer science. For example, JavaScript engines in web browsers and efficient implementations of the JVM (such as HotSpot) face the same problem that great performance benefits could be obtained if unavailable type information could be effectively discovered. (In the case of Java, the type information is eliminated by the compiler and not available in the Java bytecode received by the JVM; in the case of JavaScript, the type information never existed.) For that reason, work on this problem is currently starting at a number of places internationally.

3.3 Results obtained since the last report

Since the start of the MARVEL project, Amir Shaikhha (supported, since 2015, by Lionel Parreaux) has been developing a compilation framework for DSLs that allows to plug in very powerful domain-specific compiler optimizations to produce code that is competitive in performance with the most efficient handwritten low-level code developed by HPC experts. We want to make this framework easy to use and to work with programming languages and scripting languages that members of the MAR-VEL community want to use. The compiler framework is called SC and is developed on github. It will be released relatively early in 2016.

As expected, building a production-quality compiler generation framework is a multiple person-year effort. We have progressed to the point that the first release of our compiler framework is almost ready. We are working on finalizing the key documentation and a tutorial to make the framework usable by outsiders. We have submitted a first technical paper describing our work (for consumption by the data management research community, our group's principal research community) [5].

In the past year, we also worked on cost-based optimization inside SC. The goal is to be able to automatically create cost functions from source code, and using this for search-based code optimization (program synthesis). This work is driven by Amir Shaikhha and Mohammed El Seidy. One exciting aspect of our design is that we cannot just search for faster equivalent code but do transforms such as turning a batch algorithm into an online algorithm or automatic differentiation. The latter transforms code for a mathematical function into the differential of the function; applications of this includes automatically generating gradient code for (gradient-descent-style) machine learning techniques. We note that the scope extends beyond that of computer algebra systems such as Mathematica since the calculi themselves are definable and the functions are defined in convenient programming languages (DSLs).

We have been interacting with members of Marzari's group who seem eager to apply these capabilities in the materials science space as soon as the features are ready.

3.4 Planned research for next year

For this first compiler release, the input language is Scala, and the output is highly optimized C code; supporting further frontend languages such as Python will be future work. In the coming year, we will also start interacting closely with HPC computing experts from CSCS to build domain-specific HPC code optimizations and test our framework against the



code quality of expert HPC programmers.

Furthermore, we will work on completing a functional prototype of the cost-based search and synthesis extensions of our compiler framework, and will experiment further with automatic differentiation.

3.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* The goal of this work is to make other MARVEL teams more productive in their coding efforts while still allowing their codes to be efficient. This will have impact on all computational efforts in MAR-VEL, as long as teams are willing to work with us, and is particularly relevant to (HPC and other) software infrastructure and materials informatics projects, such as AiiDA and our advanced analytics and machine learning efforts. We have been interacting on this with Thomas Schulthess, Nicola Marzari's team, and several of the machine learning teams. These interactions will have to be intensified in the coming

months.

b) *Outside MARVEL* Our compiler infrastructure will be evaluated by other computer science groups at EPFL for developing DSLs and productively building efficient software systems, such as database management systems and software-defined networking systems. An effort to build a database management system in this way is already underway.

3.6 Personnel

NCCR personnel

• Amir Shaikhha, PhD student, 100%, from May 2014.

Matching personnel

- Mohammed El Seidy, PhD student, 100%, from May 2015.
- Lionel Parreaux, PhD student, 100%, from May 2015.

4 Speeding up the training of deep neural networks (Christoph Koch — EPFL, Volkan Cevher — EPFL)

4.1 Research summary

Deep neural networks (DNNs) are currently the best learning machines for many cuttingedge applications. For instance, Google uses DNNs to classify images and videos, to transcribe and recognize speeches, and to achieve "human-level performance." In stark contrast to its abundant success in applications, methods for training DNNs have not been so fruitful over past decades, despite of numerous efforts. The bottleneck lies in the large-scale and complicated nature of the loss functions to be minimized in DNNs. Even nowadays, the goto algorithm for most practitioners remains the classical stochastic gradient descent. In the past year, we demonstrated a principled way of improving stochastic gradient descent. Our algorithm relies on a simple yet deep observation: the gradient descent approximates functions by quadratic functions, while this choice is not the best possible for DNNs. Instead, employing a non-Euclidean norm gives the optimal approximation.

4.2 Scientific goals

How can we make reasoning in deep neural nets fast to make them a powerful and useful machine learning technique for the MARVEL community?

Machine learning bears the promise of opening up new opportunities for computational materials science, and several groups inside MAR-VEL are interested in it, including Alessandro Curioni's (IBM), Anatole von Lilienfeld's (Uni-Bas), and Michele Ceriotti's (EPFL). Our goal is to do fundamental work creating such machine learning tools and techniques for our collaborators.

4.3 Results obtained since the last report

The output layer of DNNs uses the Softmax function. Training DNNs with Softmax outputs requires us to minimize a very special function called the "log-sum-exp" (lse) function. To this end, gradient descent proceeds by approximating lse by a quadratic function, then minimizes the approximated quadratic form; stochastic gradient descent inherits this procedure with stochastic approximation, so as to tame the heavy computation required for exact approximation.

A natural question rises: are quadratic functions good enough for approximating lse? The answer turns out to be NO! To see this in a heuristic way, we refer to Fig. 5. If we draw the contour lines of lse, then obviously it looks much more like squares, rather than ellipsoids.



Figure 5: Geometry of the cost function to be optimized.

However, the contour lines of a quadratic function are always ellipsoids. This simple observation hints that approximating lse by quadratic functions is a poor choice, while approximating it by functions with square-shaped contour lines is the best one. This leads us to consider approximating lse by $|| \cdot ||_{\infty}$, whose contours lines are precisely squares. The corresponding algorithm is shown in Fig. 5: the red arrows represent the trajectory of approximating lse by $|| \cdot ||_{\infty}$, then minimizing the approximated function. This procedure always results in the optimal direction, while gradient descent takes a while to figure the truth out.

Lastly, we note that the above idea can be implemented in a modular fashion: whenever there is a gradient descent-like algorithm, we can replace the quadratic approximation by $|| \cdot ||_{\infty}$, and it always works better than its gradient descent counterpart. We refer readers to our works [6, 7], where we developed various non-Euclidean descent algorithms with provable guarantees. Although the formulas there are quite different, their origin is always the idea we described in this section.

a) *Experiments with real-world data* We tested our algorithms on two of the most popular datasets: MNIST and Cifar (Fig. 6). In this figure, we plot the testing performance vs training times. The *y*-axis corresponds to the log-likelihood, which measures how accurately we predict the data.

The classical gradient descent-based algorithms we chose are stochastic gradient descent (SGD), ADAdgrad, and RMSprop. The counterparts in the $||\cdot||_{\infty}$ case are denoted by SSD, ADAspectral, and RMSspectral, respectively. In all the cases we have implemented, utilizing $||\cdot||_{\infty}$ is no worse, and usually much better, than their gradient descent counterparts.



Figure 6: *State-of-the-art algorithmic performance on real-world datasets.*

b) Applications of DNNs for atomization energy prediction The design and discovery of new materials can be accelerated using machine learning techniques. In this quarter, we have demonstrated this idea by predicting atomization energies using DNNs [1].

Finding good feature representations for molecules is hard, in part because of the graphlike structured geometry of the molecules that need to be represented as high-dimensional vectors. The desired properties of a descriptor are invariance to rotation and translation of the molecule and invariance with respect to the permutation of atom indexes.

To this end, we propose a novel histogram based representation, which satisfies both. While there are other descriptors which overcome the translation and rotation invariance, the main property of the proposed descriptor is robustness to atom indexing. Moreover, unlike other descriptors, its dimensionality is not dependent on the size of the molecules and it varies with the number of unique atom types in the dataset.

We evaluate its performance on two datasets (with approximately 7 k molecules each) and obtain state-of-the-art atomization energy predictions using neural networks.



We would like to take a two-pronged approach in improving our predictive capabilities for materials discovery.

The first one relates to the computational scalability for Kernel based methods, which currently outperform the neural network based approaches as observed in [3]. It turns out that we can improve the flexibility of these methods to go well beyond Kernel-Ridge regression ideas. For this purpose, we will develop a stochastic three-operator splitting method and integrate it with a online optimization algorithm in order to achieve automatic tuning of the prediction method. The progress on the splitting method would be significant even on the convex optimization theory side.

In the second approach, we would like to develop active sampling schemes that work in conjunction with our predictive setup. The idea is to leverage the existing materials simulations and use this as guidance in sampling new configurations that have attractive properties. For this purpose, we will study the application of Bayesian Gaussian process optimization algorithms. For this approach to ultimately succeed, we will also have to study materials representations that are translation and rotation invariant.

4.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* We interact with the other machine learning teams in MARVEL, those of Ceriotti, Curioni, von Lilienfeld, and Marzari.

4.6 Personnel

NCCR personnel

• Edo Collins, PhD student, 100%, from March to September 2015.

Matching personnel

- Radu Ionescu, PhD student, 100%, from May 2015.
- Milos Nikolic, PhD student, 10%, from May 2014.
- Viviana Petrescu, PhD student, 100%, from February to August 2015.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- [1] V. Petrescu, V. Cevher, and C. Koch, Atomization energy prediction using neural networks with bag of bonds histogram descriptor, submitted (2016).
 - [2] P. Staar, P. Barkoutsos, R. Istrate, A. C. I. Malossi, I. Tavernelli, N. Moll, H. Giefers, C. Hagleitner, C. Bekas, and A. Curioni, Stochastic matrix-function estimators: Scalable Big-Data kernels with high performance, in Proceedings of the 30th IEEE International Parallel & Distributed Processing Symposium (IEEE IPDPS 2016), to be published (2016).
 - [3] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, *Crystal Structure Representations for Machine Learning Models of Formation Energies*, International Journal of Quantum Chemistry 115, 1094 (2015).
 - [4] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, *Machine Learning Energies for 2 M Elpa*solite (ABC₂D₆) Crystals, arXiv:1508.05315 (2015).
- [5] A. Shaikhha, Y. Klonatos, L. Parreaux, M. Dashti, and C. Koch, *How to Architect a Query Compiler, submitted* (2015).
- [6] D. E. Carlson, E. Collins, Y.-P. Hsieh, L. Carin, and V. Cevher, Preconditioned Spectral Descent for Deep Learning, in Advances in Neural Information Processing Systems 28 (NIPS 2015), C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds. (2015).
- [7] D. Carlson, Y.-P. Hsieh, E. Collins, L. Carin, and V. Cevher, *Stochastic Spectral Descent for Discrete Graphical Models*, IEEE Journal of Selected Topics in Signal Processing (2016), doi:10.1109/JSTSP.2015.2505684.

Other references

- [8] O. A. von Lilienfeld, Accurate ab initio gradients in chemical compound space, The Journal of Chemical Physics 131, 164102 (2009).
- [9] O. A. von Lilienfeld, First principles view on chemical compound space: Gaining rigorous atomistic control of molecular properties, International Journal of Quantum Chemistry 113, 1676 (2013).
- [10] R. P. Feynman, Forces in Molecules, Physical Review 56, 340 (1939).
- [11] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition* (Springer, New York, 2009).
- [12] J. Hutter et al., Computer code CPMD, v 3.17, © IBM Corp. 1990-2008, © MPI-FKF Stuttgart 1997-2001.
- [13] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Physical Review Letters 77, 3865 (1996).
- [14] M. Krack, Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, Theoretical Chemistry Accounts 114, 145 (2005).
- [15] F. Zhang, J. D. Gale, B. P. Uberuaga, C. R. Stanek, and N. A. Marks, *Importance of dispersion in density functional calculations of cesium chloride and its related halides*, Physical Review B 88, 054112 (2013).
- [16] S. Grimme, Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction, Journal of Combinatorial Chemistry 27, 1787 (2006).
- [17] T. D. Huan, M. Amsler, M. A. L. Marques, S. Botti, A. Willand, and S. Goedecker, *Low-Energy Polymeric Phases of Alanates*, Physical Review Letters **110**, 135502 (2013).
- [18] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *Fast and accurate modeling of molecular atomization energies with machine learning*, Physical Review Letters **108**, 058301 (2012).

	6	
Platform Pro	oject U	PP6 — Informatics

Project leader: Thomas Schulthess (ETHZ and CSCS)

Participating members: Thomas Schulthess (ETHZ and CSCS), Nicola Marzari (EPFL)

Summary and highlights: The Informatics Platform (PP6) is focused on providing the hardware and software infrastructure for the activities of MARVEL, and on populating and disseminating materials data to the scientific community at large. On the hardware side, we have completed the procurement and purchase of high-performance computing to support the numerical efforts of all the groups, with 180 dual-socket nodes of a Cray XC40. We have put in place an agreement with CSCS to ramp-up Petabyte-class backed-up storage to cover both the internal research activities in data-mining and machine learning, and to act as a public repository of data and workflows. On the software side, sustained development of the materials' informatics platform AiiDA (www.aiida.net) has continued, to support high-throughput calculations, data preservation and storage, scientific workflows, and public dissemination. AiiDA is slated for 1.0 release in the summer of 2016, and several training workshops have already taken place (4) or are planned (4). Work on the Materials Cloud dissemination platform, its integration with AiiDA, and data service federation has also started, as has work on the SIRIUS domain-specific library, for hybrid CPU-GPU implementations of all-electron (Exciting) and pseudopotential codes (Quantum-ESPRESSO).

General view of the project

Major research questions

The accuracy and predictive power of materials simulations based on first principles calculations allows nowadays a paradigm shift for computational design and discovery, in which massive HPC (high-performance computing) and HTC (high-throughput computing) efforts can be launched to identify novel materials or materials with improved or designed properties and performance; where behavior of ever-increasing complexity can be addressed in computational experiments; where data and workflows sharing can greatly enhance the synergies between different communities and efforts; and where services can be provided in the form of data, codes, expertise, workflows, turnkey solutions, and a liquid market of computational resources.

The impact of materials modeling on the overall scientific and economic progress will depend on our ability to promptly identify and take up the emerging scientific, technological, and societal challenges and to provide swift solutions to them. Coping with the latter task will require the generation, storage, retrieval, and analysis of ever increasing amounts of data, and the management of ever more complex codes. The complexity of the corresponding workflows calls for the implementation of a middleware infrastructure — what we label here a materials informatics ecosystem — to organise and coordinate systematically thousands to hundreds of thousands of simulations, searching for optimal properties and performance, while acquiring a variety of heterogeneous microscopic data from the *ab initio* calculations.

This infrastructure should ideally allow for an automated design and implementation of complex workflows and task tracking, based on a scripting interface for job creation and submission, and aimed at the creation and feed of heterogeneous databases of structures and properties that will in turn drive further simulations. The data thus generated will be used, e.g., for data-mining/machine learning, or to build classical neural networks to further ramp up the time and length scales accessible to numerical modeling.

Complementing this software infrastructure, it is essential to deploy the hardware infrastructure that supports the activities of all the groups, and that is amenable to optimal performance for high-throughput simulations, and integrating and supporting it through the informatics ecosystem alluded to above.

Last, open-access data dissemination and sharing needed to be supported by appropriate hardware that guarantees the storage, preservation, provenance, and dissemination of the data.



Long-term goals

- Delivery of the AiiDA materials informatics ecosystem for the automation of high-throughput calculations, the automatic storage of data in graph databases, a working environment where workflows and turnkey solutions are enabled, and sharing of data and workflows is made possible, while provenance, storage and preservation, reproducibility, and reuse are guaranteed.
- Delivery of the Materials Cloud repository for tools, properties, and data relevant to electronic-structure simulations.
- Development of domain-specific libraries to sustain the transition of electronicstructure codes to novel and heterogeneous environments based on accelerators, GPU, and multi-core architectures.
- Delivery of the hardware infrastructure that sustains the computational and data/storage needs of the MAR-VEL groups and the long-term storage, preservation, and dissemination needs.

Achievements

Hardware The computing infrastructure was installed at CSCS at the end of 2014, as agreed in the collaboration agreement between EPFL, ETHZ, and CSCS. The hardware includes 180 dual-socket compute nodes (2x Intel Haswell 12-Core, E5-2690 v3 @ 2.6 GHz), with 64 GBytes of memory (DDR4 @ 2133 MHz) and is clustered to the Cray XC40 supercomputer Piz Dora. MARVEL compute nodes have full access to the scratch filesystem of 2.7 PetaBytes shared with Piz Daint. CSCS is responsible for the hardware infrastructure as well provisioning core services, such as computing, storage, identity management and security.

After setting up the required software and user configuration, the available resources were distributed among the different MARVEL projects via node-hours per quarter. Early users were invited in Q4'2014 for software testing while Q1'2015 was dedicated to pre-production usage and fine-tuning. Production workload began on April 1st, 2015 and resources have been utilized steadily since. An initial amount of 100 TB of permanent storage was also distributed among all MARVEL projects. Additionally, members of the AiiDA team, together with CSCS, have defined the requirements for a new storage service that would be one of the pillars for data dissemination. A possible implementation is being tested at the moment and the first proof-of-the-concept is scheduled to be available in the summer of 2016. Given the close collaboration between CSCS and EPFL members regarding AiiDA deployment, infrastructure services, codes, semantics and databases among others, a meeting is scheduled monthly where the different members update the others on the progress on the different areas and possible blockers between them. Additionally, a Quarterly Report is being regularly issued by CSCS providing an overview of the accounting records, statistics and a highlight of the CSCS-EPFL collaboration on each period.

Software The AiiDA infrastructure has gone through its first public release (February 2015), a major new release (December 2015) and it is slated for v1.0 in the summer of 2016; the hire of two computer specialists is allowing to tackle some of the more technical challenges highlighted in the group report. Different efforts are taking place to scale the efficiency of the back end to large datasets, to deploy common queries in a database-agnostic, userfriendly approach, to create pipelines to and from common open-access repositories, to increase the number of plugins available for different codes, and to develop and refine workflows for automatic calculations. Notably, four tutorials have been already taken place in Berlin, Zurich, Lausanne, and Trieste, to full capacity, and four more are planned in Kyoto (\sim 20 attendees, March 2016, supported by University of Kyoto), Lausanne (\sim 30 attendees, June 2016, supported by Psi-k and MAR-VEL), in Trieste (\sim 100 attendees, July 2016, supported by ICTP) and in Trieste (\sim 100 attendees, January 2017, supported by Psi-k, CE-CAM, MARVEL, and MaX).

The Materials Cloud effort has also started, not only defining the key structure of the web portal, but working on the RESTful APIs and on secure federation of data with the CSCS infrastructure. A data management plan has been outlined, also part of the collaboration with the European MaX Centre of Excellence, and a collaboration in the form of meetings and white papers for computational and experimental metadata has taken place in the framework of the EU H2020 NFFA ("Nanoscience Foundries and Fine Analysis") program.

The domain-specific library SIRIUS has been nominated for a best paper award at Supercomputing 2015 for its CPU-GPU distributed eigenvalue solver, and a proof-of-concept for Quantum-ESPRESSO has also been achieved.

Milestones

April 2016

- Release of AiiDA with enhanced backend support (SQLAlchemy, JSONB fields) and improved query functionality including graph traversal (QueryBuilder) toward AiiDA 1.0 (to be released in summer 2016).
- Materials Cloud: proof-of-concept (internal), and data for two curated sections (2D + SSSP). Working at remote communication to CSCS.
- Initial prototype of the AiiDA plugin for the Exciting code, and AiiDA workflow to produce equation of states (EoS) from the Exciting code.

April 2017

- AiiDA 1.0 out with stable workflow engine, and benchmarked towards big data with efficient data management (storing, accessing, querying) with at least > 10 M nodes, > 100 M attributes in the DB. Full integration with Materials Cloud with seamless sharing capabilities (i.e. direct contribution of data to/from any AiiDA instance from/to Materials Cloud). Complex provenance queries across the AiiDA repository can be built graphically.
- Materials Cloud: first public version released, ability to import data directly from AiiDA, capability for contributing to properties section. Advanced usability: extended node visualisation, and graphical query builder.
- SIRIUS: implementation of the PAW method in the SIRIUS library, implementation of spin-collinear and non-collinear magnetism in the SIRIUS library and integration with Quantum-ESPRESSO code, initial implementation of real-space projectors for the pseudopotential part of SIRIUS, initial implementation of iterative solver for the full-potential LAPW.
- Complete AiiDA plugin for the Exciting code.

April 2018

• Materials Cloud: implementation of a "App Store"-like repository of plugins and workflows on Materials Cloud (both developed within MARVEL or contributed). These will be installable directly into any AiiDA instance with a simple command. Have support for contributors to create their own curated properties section.

- Availability in the repository of advanced workflows for materials properties (including at least structure relaxation, total energy and forces, phonon spectra, band structure, molecular dynamics with Quantum-ESPRESSO) to be downloaded and used by anyone having AiiDA.
- SIRIUS: complete full-potential and pseudopotential ground-state implementation in the SIRIUS library, implementation of exact exchange kernel.

For SIRIUS, by "implementation" we assume CPU, Nvidia GPU and (probably) Intel Xeon-Phi implementation of the code.

New and ongoing collaborations

- Automatic pipelines to COD and TCOD repositories, Saulius Grazulis and Andrius Merkys (Vilnius University, Lithuania), October 2014-onwards.
- Yambo plugin, Andrea Marini, Daniele Varsano and Andrea Ferretti (National Research Council, Italy), Gianluca Prandini and Antimo Marrazzo (EPFL), December 2015-onwards.
- FLEUR plugin, Jens Broeder, Gregor Michalicek, Daniel Wortmann, and Stefan Blügel (Jülich, Germany), December 2015-onwards.
- CP2K plugin, Aliaksandr Yakutovich (Empa), Uli Schauer (ETHZ), Tiziano Müller, Andreas Glosse, Patrick Seewald (UZH), Leonid Kahle (EPFL), 2016.
- SIESTA plugin, Victor Garcia Suarez (Uni Oviedo), 2016.
- i-PI plugin, Venkat Kali (EPFL), 2016.
- VASP plugin, Mario Zic (Trinity College Dublin) and Boris Kozinsky (Bosch RTC), 2016.
- Data import/export with the Materials Project, Anubhav Jain and Kristin Persson (LBL, USA), August 2015-onwards.
- Exciting plugin, Anton Kozhevnikov (CSCS), 2016.
- Quantum-ESPRESSO NEB, Marco Gibertini (EPFL), and Quantum-ESPRESSO DOS, PDOS and WANNIER90, Daniel Marchand (EPFL and McGill)), 2016.

1 Domain-specific library for electronic structure calculations (Thomas Schulthess — CSCS and ETHZ)

1.1 Research summary

The development of the SIRIUS library started within work package 8 of the PRACE second implementation phase (PRACE-2IP), with the main goal of finding major performance bottlenecks in the ground-state calculations of the Exciting and Elk codes. The library has now also been extended to the ultrasoft pseudopotential functionality of Quantum-ESPRESSO code, and the current development is focused on extending the SIRIUS functionalities and algorithms to create a domain-specific library with architecture-dependent backends for LAPW/PW/PAW-based materials science community codes.

1.2 Scientific goals

The material science community has developed a large number of electronic structure codes based on Kohn-Sham (KS) density functional theory (DFT). Some codes are used "inhouse" by the original group of developers along with a small group of collaborators. Other codes have grown into community codes with the large user base and a long list of contributors. Nevertheless, most of the codes take a spectral approach to solving a set of second-order differential equations (the Kohn-Sham equations) by expanding wave functions in terms of a fixed number of basis functions and converting the solution of the Kohn-Sham equation into a generalized eigenvalue problem of a square matrix. The eigenvalue problem is solved using either iterative subspace diagonalization techniques (as implemented in the family of plane-wave pseudopotential methods) or a by a straightforward diagonalization (used in the linear combination of atomic orbitals or augmented plane-wave methods).

Thus, we have an odd situation where (i) all electronic structure codes are based on the same DFT foundations, (ii) each of the electronic structure codes uses one of the wellknown basis sets (such as plane waves, augmented plane waves, localized orbitals, and a few others) to expand Kohn-Sham wave functions, but (iii) the particular implementation of each code solely depends on the personal tastes of the developers or the corresponding community, making it hard for external contributors to commit changes.

Moreover, the recent emergence of new architectures in computing, such as Graphic Pro-

cessing Units (GPU) or Intel's Many-Core processors dubbed Xeon Phi, has shown that typical electronic structure codes are hard to port and maintain on these new emerging platforms. Since this new diversity of architectures is a consequence of the approaching end of CMOS scaling, it is likely to stay with us and even increase as we reach the end of Moore's law some times at the beginning of the next decade. Thus, an entirely new approach to the implementation of electronic structure codes will have to be taken [1], in oder to sustain development of performant codes in the future. This approach will have to separate concerns of the scientific community that is interested in developing new methods from those developers who have to make sure the software performs efficiently on rapidly evolving architectures.

Inspired by the successful example of MeteoSwiss which has performed a refactoring of COSMO code and introduced a domainspecific library for stencil calculations, we are aiming at creating a domain specific library for electronic structure calculations. The library will include a support of several most popular diagonalization-based methods, such as plane-waves (PW) pseudopotential and projected augmented-waves (PAW), as well as fullpotential linearized augmented plane waves (LAPW). A prototype of the SIRIUS library is available [2] and includes a hybrid CPU-GPU implementation of a complete DFT groundstate loop (Hamiltonian diagonalization, density construction, generation of effective potential, symmetrization and mixing) for ultrasoft pseudopotential PW and full-potential LAPW methods. We plan on turning this library into supported product that will be extended with the following steps.

a) *Short-term* The SIRIUS library has to be fully interfaced with Quantum-ESPRESSO (QE) and Exciting codes and the library's API has to be revised and improved if necessary. The tests on correctness and performance must be leveraged from the manual to an automated fashion using the AiiDA framework.

b) *Mid to long-term* Extend the functionality of the library. There are several major pieces of the canonical Fortran90 codes that have to be re-implemented inside SIRIUS and ported to GPUs as well as other novel architectures such as Intel's Xeon Phi many-core processors. This includes magnetism and forces in the pseudopotential part of the library and exact ex-

change energy calculation for the hybrid functionals.

1.3 Results obtained since the last report

In the period between December 2014 and November 2015 the work was focused on two major aspects of the SIRIUS library.

a) Simulations of large unit cells with fullpotential LAPW The ultimate goal of this effort was to demonstrate that the full-potential simulations of large unit cells containing more than a thousand atoms are possible with the reasonable amount of resources. The work was focused around development of the new stateof-the-art hybrid CPU-GPU distributed eigenvalue solver and resulted in a paper that has been nominated for the best paper award at the 2015 Supercomputing conference [3].

b) Plane-wave pseudopotential engine The library was extended with the basic plane-wave pseudopotential functionality to run a nonmagnetic ground-state loop. The research was focused on iterative solver methods (block-Davidson, Chebyshev polynomial filtering, direct inversion in the iterative subspace, locally optimal block-preconditioned conjugate gradient), distributed problem-specific FFTs and GPU implementation. Preliminary results of the performance benchmark of Au surface test (112-atom unit cell) are shown on Fig. 1. Production version of QE code runs only on CPUs. The version of QE recompiled with the SIRIUS library takes an advantage of GPUs installed in the Piz Daint nodes. The main effect of including GPUs in the computation is significant reduction of its footprint on the machine. The simulation performs on 18 hybrid nodes as well as on 128 standard multi-core nodes. This



Figure 1: *Performance benchmark of the Quantum*-ESPRESSO code on *Piz Daint supercomputer.*

reduction in footprint is very important for high-throughput calculations, where the number of runs on a given resource needs to be maximized. Furthermore, as a result of this reduction in footprint, the strong-scaling limited is reached much sooner on the hybrid system.

1.4 Planned developments for next year

a) *Iterative solver for LAPW method* Even with the fastest distributed GPU-enable eigenvalue solver, straightforward diagonalization of a Hamiltonian matrix will remain a bottleneck. This problem is highly exposed in the "molecule-in-the-box" calculations when for the few occupied Kohn-Sham wave functions the large eigenvalue problem has to be solved. In order to circumvent this problem and make LAPW fly, iterative subspace diagonalization (based, for example, on a block-Davidson method) has to be introduced, parallelized and ported to GPUs. Attention should be payed to the good preconditioning schemes for the LAPW basis.

b) *Projected augmented-wave method* is considered as a better alternative to the ultrasoft pseudopotential method. VASP code is using it by default and QE also has a PAW implementation. The task of this development would be to understand the PAW implementation and its pitfalls in QE and and re-implement both CPU & GPU versions of PAW in SIRIUS.

c) Second-variational magnetism in pseudopotential Full-potential codes like Exciting and Elk use second-variational approach to solve the magnetic Hamiltonian:

$$\hat{H}_{\alpha\beta} = -\frac{1}{2}\nabla^2 \delta_{\alpha\beta} + V^{eff}_{\alpha\beta}(\mathbf{r})$$
(1)

where $V_{\alpha\beta}^{eff}(\mathbf{r})$ is a 2 × 2 matrix acting on spinor states

$$V_{\alpha\beta}^{eff}(\mathbf{r}) = \left(\mathbf{I} \cdot v^{eff}(\mathbf{r}) + \boldsymbol{\sigma} \cdot \mathbf{B}^{eff}(\mathbf{r})\right)_{\alpha\beta}$$
(2)

Instead of solving the Hamiltonian directly in the basis of spinor states, one can split the problem into two parts. First, the nonmagnetic Hamiltonian is diagonalized:

$$\left(-\frac{1}{2}\nabla^2 + v^{eff}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \qquad (3)$$

Then the first-variational states $\phi_i(\mathbf{r})$ are used to expand the true spinor wave functions

$$\Psi_{i\alpha} = \sum_{j}^{N_{fv}} C^{\alpha}_{ij} \phi_j \tag{4}$$

and set up the second-variational Hamiltonian which includes all remaining operators (coupling with magnetic field, Hubbard U correction, spin-orbit coupling, etc.):

$$\langle \phi_j | \hat{H}_{\alpha\beta} | \phi_{j'} \rangle = \epsilon_j \delta_{jj'} \delta_{\alpha\beta} + \sigma_{\alpha\beta} \langle \phi_j | \mathbf{B}^{eff} | \phi_{j'} \rangle \quad (5)$$

The ultimate purpose of second-variational diagonalization is a gain in performance: nonmagnetic problem will be solved using iterative subspace diagonalization (slow part), but the magnetic Hamiltonian will be solved by direct diagonalization of a small matrix (fast part).

The initial implementation should be straightforward. Use the nonmagnetic wave functions to evaluate

$$\langle \phi_i | \mathbf{B}^{eff} | \phi_{i'} \rangle$$
 (6)

and set up the second-variational Hamiltonian (Eq. 5). This will be sufficient to benchmark the collinear and non-collinear magnetism and help in designing the SIRIUS library.

1.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* The further development of the domain-specific library from a prototype stage to a production-quality software will require a close interaction with the user community. We believe that the MARVEL network is a perfect candidate for this role due to the mutual interest in high-throughput calculations.

2 The AiiDA Materials' Informatics Platform (Nicola Marzari — EPFL)

2.1 Research summary

- Providing the MARVEL community and the scientific community at large with the tools to deal most efficiently with the complex requirements of computational science, including the automation of high-throughput simulations on remote resources;
- the storage, preservation, provenance, and efficient searching of data;
- the development of a working environment where workflows connecting different and complex codes are constructed;
- the sharing and dissemination of data and workflows.

2.2 Scientific goals

The goal of the project is to provide the MAR-VEL users, and the scientific community at b) *Outside MARVEL* Recently ETHZ has become a partner of the Centre of Excellence (CoE) for computing applications MaX — "Materials design at the eXascale". MaX aims at endowing researchers and innovators with powerful new instruments to address the key scientific, industrial and societal challenges that require novel materials.

ETHZ, represented by CSCS and by the group of Thomas Schulthess at the Department of Physics, is contributing to the work package 4 "Transition to the exascale architectures". The activity covers such topics as fast and scalable algorithms for full-potential and pseudopotential density functional theory methods, implementation of the algorithms on modern and future extreme-scale computing architectures, as well as their integration with the scientific community codes officially supported by CeO MaX.

The activities of CeO MaX and MARVEL regarding the HPC side of the scientific community codes are common and a strong synergy between MARVEL and MaX is expected in the forthcoming years.

1.6 Personnel

NCCR personnel

- Pengxiang Xu, senior researcher, 100%, from January 2016.
- Urs Hähner, PhD student, 100%, from August 2015.

large, with a computational tool, AiiDA, to help them deal efficiently with computational simulations; in particular for the automation of high-throughput simulations, but also beyond this for the storage, provenance, analysis and sharing of any type of simulation.

a) *Short-term* After the first release, the main short-term goals are:

- increasing support of scientific codes by the development of additional code plugins, with particular care to codes used inside the MARVEL community;
- support of users (mailing lists, feature requests) and general maintenance of the code (bug fixes, regular releases);
- improvement of those parts of the code that can benefit of further optimisation or





Figure 2: The main components of the AiiDA infrastructure and their interactions. The core AiiDA component is the API, whose ORM represents stored objects as Python classes. The AiiDA daemon is a background process that takes care of most automated operations. It interacts with the remote clusters via different channels using the appropriate scheduler plugins.

of an improved design (see "Planned research for next year" below);

• strengthening of the user base and organisation of tutorials.

b) *Long-term* More generally, in the long term we plan to make AiiDA a platform that fully implements the four pillars of the ADES model [4], that are:

- Automation: Automate job execution, in particular input file creation, copy of files to remote supercomputers, job submission, file retrieval and parsing. Scale up to hundreds of jobs to be able to support high-throughput simulations.
- Data: Store calculations, their inputs and results (extracted from XML or JSON files, or parsed) in a database, in a format suitable for computational materials science. This requires to store attributes with a flexible schema (in AiiDA, we implement entity-attribute-value tables to store any type of attribute). The database structure is designed not only to guarantee the reproducibility of calculations, but also the provenance of the data, by representing calculations in a directed acyclic graph.
- Environment: Much more that the data, the future of computational materials science will rely on preparing, storing and

sharing workflows that generate the data, and that combine different computational engines. Also, the infrastructure should be a research facilitator, so it must be easy to use and provide materials scientists with a natural research environment.

• Sharing: While the databases of each user (or trusted group of users) need typically to be private, import/export features are needed to share selected portions of the databases with other researchers or to automatically upload successful calculations to a public area. Also, common formats and ontologies need to be defined and used to make it possible to share data between different codes and groups, and to search into data efficiently.

2.3 Results obtained since the last report

Since the last report, we have released the first public version of the AiiDA code (v.0.4.0, with an open-source MIT license) on February 27, 2015, followed by a minor release (v.0.4.1) on April 8 and a new major release (v.0.5.0) on December 17, 2015, with all recent updates and new functionalities. The current AiiDA infrastructure is illustrated in Fig. 2. Along with the release of the code, we have also published a paper describing the code and the ADES model behind it; the preprint has been available on
arXiv since April 2015, and the paper is now published in a peer-reviewed journal [4].

At the same time, we have spent significant efforts to build up a solid user base and strengthen the user support: we have published online the full documentation of AiiDA (http://aiida-core.readthedocs.org, Fig. 3), including installation instructions, tutorials for users and developers, extensive code documentation, etc.; we have opened a mailing list to support user requests; we have organized three tutorials addressed not only to users but also to developers.

After the first release, code developments have continued steadily thanks to the combined efforts of the various AiiDA developers, now including also two computer scientists (Spyros Zoupanos and Snehal Waychal) who have joined MARVEL in the summer of 2015. The main developments of the past year include:

- Improvements to the user interface and especially to the command line verdi, including commands to directly list, show and export data objects such as crystal structures, band structures, trajectories; improved commands to manage groups of nodes.
- Support for multiple profiles/databases, switchable at execution time from the command line, also as a preliminary step toward flexible data sharing.
- Support for "Inline Calculations", i.e. calculations that are run in place without the need to go through a job scheduler, but are nevertheless automatically stored in the AiiDA database to keep provenance.
- Improvement of the AiiDA calculation plugins API to allow for multiple executable calls within the same scheduler job. This makes it easier to support codes

🖬 AiiDA	Docs * Welcome to AliDA's documentation!	View page so
archdocs		
er's guide	🖧 🖓 🕹 🕹	7
Databases for AiiDA	Automated Interactive Infrastructure and Database	for Atomistic simulations
nstallation and Deployment of AJIDA	Welcome to AiiDA's documentation	n!
etup of computers and codes		
scripting with AiiDA	AiiDA is a sophisticated framework designed from scratch to be	a flexible and scalable infrastruc
StructureData tutorial	for atomistic simulations. Being able to store the full data proven	ance of each simulation, and ba
Quantum Espresso PWscf user-tutorial	on a tailored database solution built for efficient data mining imp the ability to interact seamlessly with any number of HPC machi	lementations, AiiDA gives the u nes and codes thanks to its flexib
Quantum Espresso Phonon iser-tutorial	plugin interface, together with a powerful workflow engine for th	te automation of simulations.
Quantum Espresso Car-Parrinello Iser-tutorial	The software is available at http://www.aiida.net.	
Setting parsed calculation results	This is the documenation of the AiiDA framework. For the first s	etup, configuration and usage, re
seudopotential families tutorial	to the user's guide below.	
Manually loading pseudopotentials	If, instead, you plan to add new plugins, or you simply want to un	derstand AliDA internals, refer t
AiDA schedulers	the developer's guide.	
Theck the state of calculations		
AiiDA workflows	User's guide	
her guide resources	User's guide	
weloper's guide	 Databases for AliDA 	

Figure 3: The first page of the AiiDA documentation. Different sections are aimed at the end users, and at developers. that require pre- and post-processing steps (such as Yambo, FLEUR, WANNIER90, i-PI).

- A first version of a Python Querytool has been implemented, to support common queries to the AiiDA graph database with a database-agnostic and user-friendly interface.
- Improved support for job schedulers (now also supporting independently both Torque and PbsPro, and with a DirectScheduler implementation to emulate a scheduler on a machine on which none is installed).
- Significant efforts have been focused on developing automated and robust workflows to allow for seamless highthroughput calculations using Quantum-ESPRESSO; more details are described in the VP2 project, section 9 "Highthroughput vibrational, dielectric and thermomechanical properties of novel materials". As a part of the same project, another workflow has been implemented to scan crystal structures and identify twodimensional layered substructures. Also, a workflow for calculation of phonon dispersions in the frozen-phonon approximation using PhonoPy has been developed, as well as a workflow to initialise and execute molecular-dynamics simulations using the Quantum-ESPRESSO Car-Parrinello code.
- A Worker class has been implemented to allow for an automatic scheduling and management of the number of workflows running at a given time.
- A new optimized backup script is now available, performing efficiently also with databases composed of tens of millions of nodes.
- Many new simulation codes are now supported as AiiDA plugins: NWChem, PhonoPy, CODTools, GPAW, ASE and its calculators (over 20, including VASP, LAMMPS, etc.). Also, many of the Quantum-ESPRESSO plugins have been significantly improved (for instance, better support of spin-polarized calculations and automatic computation of the bandgap in the pw.x code, whose input structure in AiiDA is represented in Fig. 4).
- Triggered the development of new calculation plugins, also thanks to the organisation of the AiiDA developers' tutorial



Figure 4: A simple example of how a Quantum-ESPRESSO calculation, the executable code and input/output data are represented as nodes in a graph. In this example, we have the relaxation of a diamond crystal structure.

and coding days in November 2015. In most cases, we also involved the simulation code developers. Native plugins under development include the Quantum-ESPRESSO NEB code, WANNIER90, FLEUR, Yambo, CP2K, VASP, Exciting, SIESTA, i-PI.

- New plugins to support data structures have been developed: band structures with automatic generation of the *k*-point path (following the high-symmetry lines reported in [5] for 3-dimensional Brillouin zones; we also extended the class to 2D Brillouin zones), molecular-dynamics trajectories, CIF-format structures, pseudopotentials in the UPF format. Other data structures are under developments, to support (in a code-independent format): Gaussian basis sets, atomic-orbital basis sets, Wannier functions, force constants calculated from phonons.
- Importers and exporters between different formats have been implemented, such as for instance converters between the AiiDA StructureData format for crystal structures and the ASE Atoms format (https://wiki.fysik.dtu.dk/ase/ [6]), and the Materials Project pymatgen format (http://materialsproject.org [7]). This also allows now for seamless integration of AiiDA with these Python packages for structure manipulation, structure matching, and all other features implemented in these external packages.
- Importers and exporters from a large number external databases has of been implemented improved. or The currently supported databases ICSD, OQMD, NNINC, PCOD, are: MPOD, TCOD. In the case of TCOD (http://crystallography.net/tcod [8]), it

is not only possible to import structures into AiiDA, but also to export the full calculation provenance to and from the TCOD database. In this respect, we are closely collaborating with the TCOD team to develop a code-independent ontology for DFT simulations. In order to support ontologies in AiiDA, we have defined and documented an abstract API for generic exporters. We then provide within AiiDA a specific implementation to export results run in AiiDA using Quantum-ESPRESSO into the TCOD database, correctly assigning calculation results with the flags defined in the TCOD ontology.

• Unit tests for the vast majority of the code to ease the process of code development and validation. Unit tests also for code parsers, that can be automatically generated from calculations that have been already run with AiiDA.

Note that some of the tasks reported in the report of last year (such as the AiiDA web interface, the REST API, and the creation of the web portal) have now been branched out to the Materials Cloud subproject and their description can be found in the next section.

2.4 Planned research for next year

We list here the main development directions for next year.

- Decouple the code from the current ORM (Django) by abstracting the backend interface into an API.
- Implement a second backend using SQLAlchemy as the Python interface, and using PostgreSQL 9.4 or later with JSON fields to store Attributes and Extras. This should provide a significant speed-up in



storing and retrieving big sets of nodes from the database.

- The previous item will also require to deploy an improved Querytool, exposing the same API for common queries, but implementing it for each database backend (Django and SQLAlchemy). We plan to switch all queries appearing in the code to use only this Querytool.
- Investigate at least one graph database (identified as the Titan Distributed Graph Database) as a further AiiDA backend, to assess whether it can provide improved query efficiency and scalability for large databases.
- Improve the current workflow engine to scale up to hundreds of concurrent running workflows, to improve the data model, to facilitate queries, and to make it easier for developers to write and debug new workflows. Preliminary work in this direction has been made with the current developments of the Worker class discussed above.
- Work on the enhancement of AiiDA sharing capabilities to support seamless sharing between different backend types, to remove bottlenecks in the process of import and export so as to make them more efficient, and simplifying the user experience while performing database migration and sharing operations.
- Coordinate plugin development efforts in various groups (the list of plugins under development is listed in the previous section).
- Prepare and release a virtual machine with the most recent code version and make it available to users. Also, release the virtual machine used for the AiiDA tutorial, including the data and the tutorial instructions.

2.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* AiiDA has been one of the project proposed to the students of Christoph Koch's "Big Data" course held at EPFL during the 2015 spring semester. A group of eight students has worked for the full semester and has investigated possible efficient extensions of the AiiDA backend using other SQL and NoSQL databases. The Titan Graph Database has been identified as a promising backend solution for large-size databases. Also, the group has investigated how to port functionality for the identification of two-dimensional structures to the Scala language and the Apache Spark infrastructure in order to distribute the tasks over a large commodity cluster.

Four tutorials of the AiiDA code have been held to date. The first users' tutorial was addressed to the MARVEL community and was held in October 2014 in Zurich (\sim 30 attending). A second users' tutorial was held for the Psik/CECAM conference in Berlin in February 2015 (\sim 40 attending). The third tutorial was held in Lausanne in November 2015 (~40 attending) and mainly addressed to the communities of MARVEL and of the new EU H2020 MaX Centre of Excellence (CoE) on "Materials Design at the Exascale". This last tutorial was addressed both to users and developers of AiiDA, and was followed by three coding days for interested plugin developers, where the developments of at least nine new plugins was initiated. The last one has been held within the Master in HPC (www.mhpc.it) in Trieste, Italy on December 15, 2015, to all the students of the Master.

We are collaborating with Michele Ceriotti to ensure compatibility between AiiDA and the software/library to compute the crystal structure fingerprints he is developing. In particular, we are aiming at a Python interface of the code (that will be implemented either in C/C++ or Fortran for efficiency reasons) in order to make it possible to compute efficiently thousands of fingerprints of crystal structures already stored in the AiiDA database.

We are holding monthly meetings with CSCS to coordinate the efforts on core services, codes support, support for metadata and semantics, and data sharing of large files stored at CSCS to the outside world.

b) Outside MARVEL We are now also part of two new H2020 Centres of Excellence: the einfrastructures centre MaX ("Materials Design at the Exascale") and the integrating infrastructure NFFA ("Nanoscience Foundries & Fine Analysis"), both of which have AiiDA as a core component for the management of computer simulations. Some of the tutorials described above involved also centres outside MARVEL (such as the MaX CoE). Also, we have applied for co-funding for a MARVEL/Psi-k school on high-throughput computations using AiiDA to be held in June 2016. We are collaborating closely with the Crystallography Open Database (http://crystallography.net [8]). In particular, Andrius Merkys visited the group

for a year, and worked on automatic importers and exporters for the TCOD database including the full provenance and using the TCOD ontology to flag result variables; these efforts are now released together with AiiDA. We are also closely interacting with the Materials Project (http://materialsproject.org [7]) and the Open Quantum Materials Database (http://oqmd.org [9]) to maintain data compatibility and interoperability; converters and/or exporters are now available into AiiDA. We are in close interaction with Atsushi Togo (Tokyo University), author of the Spglib and phonopy packages. Both are now already usable from AiiDA, and we plan to improve further the interoperability. To this aim, Giovanni Pizzi will visit Togo's group in Japan in the spring 2016.

2.6 Personnel

NCCR personnel

• Giovanni Pizzi, postdoc, 100%, from September 2014 to July 2015 and from January 2016.

3 The Materials Cloud portal (Nicola Marzari — EPFL)

3.1 Research summary

The goal is to provide the MARVEL community and the scientific community at large with the tools to deal most efficiently with the complexity needed and empowered by computational science, including the automation of high-throughput simulations on remote resources. To this end we are developing a rich web interface to expose not only the AiiDA database but also the ability to carry out complex calculations without domain expertise. In the past year much of the server-side hardware and software infrastructure has been procured and implemented and client-side work is ongoing.

3.2 Scientific goals

A major goal for the informatics platform is to build a portal centered around the materialscloud.org website to enable sharing and collaboration within the NCCR as well as to give easy access to a curated set results to the wider world.

a) *Short-term* To design, validate and put in place the software and hardware infrastructure needed to support the dissemination of results

- Spyros Zoupanos, postdoc, 100%, from June 2015.
- Nicolas Mounet, postdoc, 100% (on VP2), from February 2015.
- Martin Uhrin, postdoc, 100%, from January 2015.
- Snehal Waychal, system specialist, 100%, from August 2015.

Matching personnel

- Giovanni Pizzi, postdoc, 100%, from August to December 2015.
- Andrea Cepellotti, PhD student, 100%, from May 2014.
- Andrius Merkys, visiting PhD student, 100%, from October 2014 to October 2015.
- Daniel Marchand, visiting Master student, 100%, from June 2015 to December 2015.

and data by members of the NCCR to the wider scientific community.

b) *Long-term* To build a portal that enables seamless sharing and collaboration using a centralized AiiDA-based infrastructure where groups can have their own space and give selective permission for others to view and add to the current set of calculations. This should be seen as a replacement to the current way of collaborating where files are emailed or otherwise shared and provenance is difficult or impossible to keep track of.

Furthermore, the portal should enable sharing of workflows and have the ability to launch certain standard workflows such as structure relaxations directly from the Materials Cloud website which would run directly on CSCS resources, possibly as backfill.

3.3 Results obtained since the last report

Fig. 5 shows a schematic of the sharing infrastructure that AiiDA will enable. Shown are AiiDA instances that can be run privately by groups within or outside the NCCR and the Materials Cloud website which acts as a portal for disseminating both public and private data generated by members of the NCCR and our collaborators.



Figure 5: The AiiDA sharing infrastructure.

In the last year the following progress has been made to realizing these goals.

a) *Materials Cloud website* The goals and conceptual layout of the Materials Cloud website have been agreed upon and planned. The site will provide open, web based access to materials data and various tools to analyze it. A part of this website will provide a fully functional interface for AiiDA that will expose a central database of all materials and simulations open to members of the NCCR only. It will be organized into three main sections.

- Learn/Tools: a section containing learning materials such as video lectures, tutorials, and presentations related to material science.
- Explore/Properties: this will contain a curated set of results including structures and their properties as generated by MAR-VEL members. Members will be encouraged to have their own sections related to their projects. In addition, there will be a section to publish AiiDA workflows and plugins to enable others to easily carry out complex calculations designed by domain experts.
- Work/Data: this section will give access to a full AiiDA database instance which groups can use to collaborate by giving others in the consortium read and, optionally, write access to their set of results and calculations. It will be possible to easily make public final, polished, results by publishing in the *Explore* section.

To support the Materials Cloud website a server was purchased which is hosted at EPFL and the website went live in August 2015. The current website, shown in Fig. 6, shows results of pseudopotential testing carried out by Ivano Castelli.



Figure 6: The landing page of materialscloud.org.

b) Data dissemination To support the dissemination of results, infrastructure is being prepared at CSCS which will provide a significant storage capacity to enable the goals of the platform. Over a series of meetings a design for both the back and frontend was agreed upon that meets the security requirements of CSCS and supports the use-case requirements for the platform. It was agreed that iRODS will be used as the storage system, a schematic for which can be seen in Fig. 7. Care was taken to ensure that the proposed solution could be fairly easily replicated at other computing centers, should they choose to run a similar AiiDA service.

CSCS have currently set up a testing server running iRODS and after an initial evaluation of the suitability, functionality will be added into AiiDA to support this usage scenario.

- c) Other
 - A RESTful API was written for AiiDA using the *Tastypie* web service allowing clients to query the database in a standard way.



Figure 7: The backend design for the iRODS service that will serve results from the activities of the NCCR.

• A group of computer science students investigated alternative backends for the AiiDA database, including the Titan and Neo4j database, finding these to offer some benefits but with some drawbacks and an implementation cost. Furthermore the students investigated the benefits of replacing the Django ORM with SQLAlchemy and found this to give significantly greater flexibility on the types of queries that can be executed with a corresponding significant speed improvement. This change which will now be implemented.

3.4 Planned research for next year

The Materials Cloud web interface will be developed using modern web technologies and various open source libraries. High-level client-server architecture of AiiDA web interface is show in Fig. 8. The main components at server side are AiiDA APIs, Data ORM (explained in section TODO:xyz) and RESTful services.

RESTful web services provide light weight, highly scalable and maintainable interface to the client applications. The RESTful API talks with ORM through the attributes and methods provided by the AiiDA ORM API to access or update the corresponding entries in the database tables or repository files. These services are commonly used to create APIs for web based application. Using these services user can access list of data nodes and filter it by applying various properties. In AiiDA, the RESTful services are written using Tastypie web service API framework. The client side architecture is explained in the next section.

The web client interface will be the main development task for next few months. This will be developed using *AngularJS* JavaScript framework and various open source libraries like *D3.js, Bootstrap, JSmol, JQuery UI*, etc.



Figure 8: AiiDA web UI Architecture.

AngularJS is a structural framework for dynamic web apps that extend HTML's syntax to express the application's components clearly and succinctly. Angular's data binding and dependency injection eliminate much of the code that we would otherwise have to write. The goal of this new web interface will be to provide high performance, interactive and rich user experience for exploring AiiDA data. Angular uses the Model, View (or Template), Controller, this isolates the application logic from the user interface layer and supports the separation of concerns. The Model component of JavaScript framework syncs with server side data using RESTful service. Templates get compiled at runtime into views which uses the model data and updates the DOM. The Controller provides the interface between Model and View based on user inputs.

Some of the important features of Materials Cloud website are:

- Authentication: an authentication layer will be implemented allowing members of MARVEL to access private sections of the website not accessible to the public.
- Node and metadata display: this feature will provide an interface to access different types of nodes allowing users to search and filter giving immediate access to materials properties, structures and the full provenance of calculations (Fig. 9 for an example).
- Visualization: this will be one of the important features of Materials Cloud web interface. Using this feature one can visualize nodes in different formats using visualization plugins. To implement visualization plugins different JavaScript libraries will be used. For example, *D3.js* could be used to display band structure or *JS*-*mol* to visualize crystal structure. The plugins will be modularized in such a way

<mark>8</mark> Aii	DA Web Interface	Home Computers	Calculations	Workflows	Data Nodes	1 ailda@localhost -
Ca	lculations	S List				
Ţ	Filters				•	
	Label ~ Float	×	State = FINS	SHED *		
ID	Label	Туре	Computer	State	Creation -	
499	Test float sum >	FloatsumCalculation	theospc	FINISHED	15 Oct 2015 14:01:39 CEST	Action +
494	Test float sum >	FloatsumCalculation	theospc	FINISHED	15 Oct 2015 14:00:11 CEST	Action +
485	Test float sum >	FloatsumCalculation	theospc	FINISHED	15 Oct 2015 12:26:34 CEST	Action +
474	Test float sum >	FloatsumCalculation	theospc	FINISHED	14 Oct 2015 17:19:09 CEST	Action +
418	Test float sum >	FloatsumCalculation	theospo	FINISHED	1 Sep 2015 11:00:42 CEST	Action +

Figure 9: *A filtered list of calculation nodes from the AiiDA web interface.*

that it will be easy to add new visualization views or data type using existing code. Also the same visualization views, for example crystal structures, can be displayed in different ways using different JavaScript libraries.

- **Computation:** authenticated users will be able to access AiiDA features to submit calculations and workflows without having to use a command line interface. This will enable non-domain-experts to carry out complex calculations easily.
- App Store: the website will allow users to share their workflows and plugins to extend the functionality of AiiDA by adding new features and support for new codes.
- **Own space:** groups within MARVEL will be able to have their own space and give selective permission for others to view and share their results.

3.5 Personnel

NCCR personnel

- Martin Uhrin, postdoc, 100%, from January 2015.
- Snehal Waychal, system specialist, 100%, from August 2015.
- Giovanni Pizzi, postdoc, 100%, from September 2014 to July 2015 and from January 2016.

Matching personnel

- Giovanni Pizzi, postdoc, 100%, from August to December 2015.
- Andrius Merkys, visiting PhD student, 100%, from October 2014 to October 2015.
- Andrea Cepellotti, PhD student, 100%, from May 2014.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

- T. C. Schulthess, *Programming revisited*, Nature Physics 11, 369 (2015).
- [2] Domain specific library for electronic structure calculations, https://github.com/electronic-structure/ SIRIUS.
- [3] R. Solcà, A. Kozhevnikov, A. Haidar, S. Tomov, J. Dongarra, and T. C. Schulthess, *Efficient implementation of* quantum materials simulations on distributed CPU-GPU systems, in SC'15, Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (ACM, New York, 2015), article No. 10, doi:10.1145/2807591.2807654.
- [4] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, *AiiDA: automated interactive infrastructure* and database for computational science, Computational Materials Science **111**, 218 (2016).

Other references

- [5] W. Setyawan and S. Curtarolo, *High-throughput electronic band structure calculations: Challenges and tools*, Computational Materials Science **49**, 299 (2010).
- [6] S. R. Bahn and K. W. Jacobsen, An object-oriented scripting interface to a legacy electronic structure code, Computing in Science & Engineering 4, 56 (2002).
- [7] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, *Commentary: The Materials Project: A materials genome approach to accelerating materials innovation*, APL Materials 1, 011002 (2013).
- [8] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quirós, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, *Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration*, Nucleic Acids Research 40, D420 (2012).
- [9] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD)*, The Journal of The Minerals, Metals & Materials Society (TMS) 65, 1501 (2013).

Platform Project 7 PP7 — Experiments

Project leader: Frithjof Nolting (PSI)

Participating members: Frithjof Nolting (PSI), Michel Kenzelmann (PSI), Pierangelo Gröning (Empa), Claudia Cancellieri (Empa), Thomas Lippert (PSI), Dirk van der Marel (UniGE), Marisa Medarde (PSI), Thomas J. Schmidt (PSI), Thorsten Schmitt (PSI), Grigory Smolentsev (PSI), Urs Staub (PSI)

Summary and highlights: Turning the results of the computational approach into real materials requires the intense collaboration between theory and experimental groups. The PP7 is the linking point and provides funding for experimental verification projects. Two calls for proposals have been launched. In the first call, which was launched at the end of 2014, eightproposals have been granted and six of them have started. The second call was lunched at the end of 2015 and six proposals have been granted. These projects cover different topics of VP1 and VP2 and have all an experimental and a theoretical group leaders. First results include the successful high-pressure synthesis of powder samples of nickelates and the comparison of high-throughput theoretical calculations with measured crystalline structure of an oxynitride photocatalysts leading to an input for the refinement of the model.

General view of the project

Major research questions

The experimental verification platform (PP7) on experimental synthesis and characterization within the NCCR harnesses the vast and diverse experimental capabilities in materials' synthesis, analysis and characterization of Swiss laboratories. The goal is to create a tight synergy between experiment and the computational activities of this NCCR, with the objectives of

- providing experimental synthesis, characterization, and measurement of the physical properties with which the performance of the materials' design effort can be evaluated,
- helping directing this effort to support some of the strategic goals of the experimental materials' laboratories.

The major research questions are directly related with the activities of VP1, Novel Material Physics, and VP2, Novel Materials Applications. In particular

- materials with novel multifunctional properties (multiferroics and artificial heterostructures),
- materials with novel physical properties (topological insulators, model Hamiltonian, dyn. excited materials),

• novel materials for energy applications (dye-sensitized solar cells, photocatalytical water splitting, metal/air batteries, thermoelectrics).

The task of the overall PP7 is to coordinate the activities between the individual theoretical and experimental partners. In order to enable close collaboration projects for experimental verifications, a proposal-based system for funding of postdoc or PhD projects was established.

Long-term goals

The long term goal is to demonstrate the power and potential of computational driven materials design with a close feedback system with experimental verification. Therefore the progress of the different subprojects will be monitored and the tight connection and feedback loop between theory and experiment ensured. Besides meeting within the individual subprojects a workshop will be held where the results for each subproject from the theory part and the experiment part will be presented. A further long-term goal is to evaluate the proposal-based system and, if necessary, redefine future directions.

Achievements

The PP7 platform has now, in the second year of MARVEL, started. Eight out of 20 submitted proposals for experimental verification were selected (Table 1). Four of them are in the area of VP1 and four in the area of VP2. Besides the scientific merit, the connection to MARVEL theory topics and feedback loops were criteria for selection. Most of the awarded projects could fill the scientist position, e.g. postdocs, and started the project. First results are achieved which can be found in the detailed reports of the subprojects. Some highlights are:

- In the project about nickelates by Medarde and Spaldin/Georges, a new highpressure synthesis setup was successfully commissioned and enabled the successful growth of large (1 g) powder samples of nickelates with R = La to Dy, as well as DyNiO₃ nanocrystals, which so far has not been possible.
- In the project about artificial photosynthesis by Lippert and Marzari, the oxynitride photocatalysts YTaON₂ was synthesized, based on the results of high-throughput theoretical calculations. However, the obtained crystalline structure is different than expected and the refinement of previous calculations is ongoing.
- In the project about the oxygen evolution in perovskites by Schmidt and Marzari, the bulk electronic structure, surface composition, conductivity and electrochemical activity of LaSrCoO has been theoretical and experimentally investigated suggesting that the oxygen evolution reaction (OER) activity is driven by the electronic structure.

A second call for experimental verification projects was lanced. As the previous call, this was started by a one day workshop hold this time at Empa (Fig. 1). At this workshop, the experimental group leaders of the projects of the



Figure 1: Daniele Passerone giving the introduction talk at the Empa workshop on October 23, 2015.

first call gave presentations about their project. Two invited speakers from industry presented their work about "Realistic material design & simple computational modeling closed-loop in an industrial environment" (BASF) and "Computational design of novel materials for energy storage applications" (IBM). In short talks, new ideas for experimental verification projects were presented. The second call for experimental proposals closed at the end of November and the evaluation was made in December 2015. Based on the experience from the first call and the recommendation from the Review Panel, the selection criteria were slightly modified. They now emphasize more clearly the connection to MARVEL theory and the fact that both, applied and basic research, are at the focus. The criteria are:

- tight connection with computational efforts of MARVEL,
- link to topics of MARVEL,
- connections to theory group in MARVEL,
- scientific excellence (in view of basic research and/or potential for technology transfer),
- feasibility, in particular on a two year time scale.

The evaluation committee consisted of Frithjof Nolting (PSI), Michel Kenzelmann (PSI), Pierangelo Groening (Empa), Maksym Kovalenko (Empa/ETHZ) from the experimental side. From the theory side Nicola Spaldin (ETHZ) and Nicola Marzari (EPFL) were members of the evaluation committee, ensuring that the connection to MARVEL theory is taken into account.

Milestones and metrics

The goal of PP7 is to provide experimental verification of theoretical calculations. The metrics by which we judge the success of PP7 is as follows:

- submission of proposals by experimental and theoretical PIs on a common project as a result of a "Call for Proposals";
- selection of projects with the highest potential for experimental verification;
- execution of projects with established feedback loop, regular meetings;
- definition of policies how collaborations between high-throughput theorists and experiments can be guided most efficiently; these may serve as a blueprint for other people in the community.

Topics	Group leader Experiment	Group leader(s) Theory	Start	End
Structural and <i>in situ</i> electrochemi- cal characterization of oxide phase transformation at oxide-liquid in- terface	Claudia Cancellieri (Empa)	Alfredo Pasquarello (EPFL), Daniele Passerone (Empa)	01/2016	12/2017
Theory and experiment synergy for artificial photosynthesis	Thomas Lippert (PSI)	Nicola Marzari (EPFL)	06/2015	05/2017
Higgs and Goldstone modes in hexagonal manganites	Dirk van der Marel (UniGE)	Nicola Spaldin (ETHZ)	?	?
$RNiO_3$ perovskites: exploring the boundary between localized and itinerant behavior	Marisa Medarde (PSI)	Antoine Georges (UniGE), Nicola Spaldin (ETHZ)	08/2015	07/2017
Development of advanced electro- catalysts for water splitting: corre- lation between electronic structure, surface properties and electrochem- ical activity	Thomas Schmidt (PSI)	Nicola Marzari (EPFL)	08/2015	07/2017
Resonant inelastic X-ray scattering on thin films and oxide heterostruc- tures for future mottronics and or- bitronics	Thorsten Schmitt (PSI)	Nicola Spaldin (ETHZ)	01/2016	12/2017
Time-resolved X-ray absorption spectroscopy to investigate mech- anisms of photochemical water splitting reactions with molecular catalysts	Grigory Smolentsev (PSI)	Jürg Hutter (UZH)	?	?
Testing ultrafast processes in con- densed matter	Urs Staub (PSI)	Philipp Werner (UniFR), Nicola Spaldin (ETHZ)	09/2015	08/2017

Table 1: Experimental verification projects selected in the first call. Start and end are referring to the actual term of employment of the postdocs.

This translates into the following milestones:

October 2014

• Organization of workshop to bring together experimentalist with MARVEL theorist.

December 2014

• First Call for Proposals.

February 2015

• Selection of about 50% of the possible experimental proposals (Table 1).

November 2015

• Second Call for Proposals.

December 2015

• Selection of remainder of possible experimental proposals (Table 2).

Topics	Group leader Experiment	Group leader(s) Theory	Start	End
Experimental realization of novel topological semimetals	Ming Shi (PSI)	Oleg Yazyev (EPFL), M. Troyer (ETHZ)	?	?
Microscopic origin of the magneto- electric properties in strained and doped Aurivillius phases predicted by DFT	Marta D. Rossell (Empa)	Claude Ederer, Nicola Spaldin (ETHZ)	?	?
Using computational chemistry to predict the performance of metal- organic frameworks catalysis in the hydroformylation of olefins	Marco Ranocchiari (PSI)	Berend Smit (EPFL)	?	?
The search for new proton con- ductors: high-throughput screen- ing and experimental synthesis and characterization	Daniele Pergolesi (PSI)	Nicola Marzari (EPFL)	?	?
Single-band Hubbard model in new fluorides	Christian Rüegg (PSI/UniGE)	Nicola Spaldin (ETHZ), Thomas Schulthess (ETHZ), Matthias Troyer (ETHZ)	?	?
Colloidal nanocrystals as model systems to uncover struc- ture/properties relations in CO ₂ electroreduction	Raffaella Buonsanti (EPFL)	Nicola Marzari (EPFL)	?	?

 Table 2: Experimental verification projects selected in the second call.

April 2016

- Organization of PP7 wide workshop with focus on the start of the projects and first results.
- Running of all 8 subprojects from the first call.
- Start of at least 4 subprojects from the second call.

April 2017

- Organization of PP7 wide workshop with focus on status, interim results and next steps of the projects.
- Running of all 14 subprojects.
- Evaluation of proposal based funding system.

- Define future strategy of PP7 within MAR-VEL.
- Define guidelines for fostering collaborations between high-throughput theorists and experimental validation of results.

April 2018

- Organization PP7 wide workshop with focus on the results of the subprojects.
- Finish of most subprojects.

New and ongoing collaborations

First call In the first call for experimental verification projects seven postdoc projects and one PhD project were granted. Six of these projects

are now in progress. The leader of one project is in the process to hire suitable staff. The balance between VP1 related and VP2 related projects is about 50/50. Table 1 gives a list of the projects from the first call. In the table the group leader of the experimental verification part is listed and also the group leaders from the corresponding theory effort. As it is obvious from the call regulations, each project has an experimental and a theory part. Worth noting is that several projects involve more than one theory group leader, demonstrating the networking character of these projects. Actually, some of the projects are also collaborations between different experimental groups.

Second call In the second call six postdoc projects have been granted, which now start searching for candidates (Table 2). Again the split between VP1 and VP2 related topics is 50/50 and some projects involve more than one theory group leader and some are collaborations between different experimental groups.

1 In situ structural and electrochemical characterization of phase transformations at oxide-liquid interface (Claudia Cancellieri — Empa)

1.1 Research summary

Our project addresses the microstructureproperty relationships of defective functional oxides in different environmental conditions. As a first step, the structure and properties of amorphous Al₂O₃ will be studied by experiment and theory. To investigate the local chemical state of the Al and O ions in the amorphous state, electrochemically-grown amorphous Al-oxide films prepared under different growth conditions, will be investigated by absorption spectroscopy. Additionally, the electronic properties of the oxide films will be studied by valence band measurements. The thus-established experimental details on the local atomic structures and electronic properties of the amorphous oxide films will be compared to corresponding density functional theory (DFT) model predictions.

1.2 Scientific goals

a) *Short-term* The short-term goal of this project (end of next year) is to understand the local atomic structure and electronic properties of electrochemically-grown amorphous Al_2O_3 layer, from an experimental viewpoint supported by theoretical calculations. Theoretical and experimental results on the local atomic structure and electronic properties of the non-crystalline, defective Al-oxide phase will be compared with the thermodynamically-preferred bulk state of Al_2O_3 , presented by single-crystalline α -Al₂O₃.

b) *Long-term* The final aim of this study is to arrive at a fundamental understanding of the structure and electronic properties of functional oxides (in particular, Al_2O_3) at their surfaces in contact with liquids (in particular, water). It is envisaged that the theoretical predictions will be able to address the defect structure

ture and electronic properties resulting from the incorporation of H into the oxide subsurface, which will be compared with experimental investigations. The thus-obtained fundamental knowledge on the microstructureproperty relationships of defective functional oxides in contact with water can be exploited to tailor functional oxide surface properties towards higher catalytic process efficiencies and long-term stability/corrosion resistance in operating environments.

1.3 Results obtained since the last report

We have grown electrochemically dense Aloxide on Al metal and Al-coated substrates. The observed linear relation between the applied voltage during growth and the oxide thickness allows very accurate control of the oxide layer thickness. Microstructural analysis by SEM, XRD and FIB/TEM evidence an amorphous state of the oxide phase. The surface roughness of the oxide layer strongly depends on the growth conditions. X-ray absorption spectroscopy (XAS) at the Al and O edges was performed at the Swiss Light Source (SLS) at PSI, at the PHOENIX beamline in total electron yield (TEY) mode in order to be surface sensitive. The results, shown in Fig. 2, demonstrate a complex local coordination of Al cations by O nearest-neighbor anions, which cannot be described by a randomly interconnected network of tetrahedrally- and octahedrally-coordinated [AlO₄] and [AlO₆] building blocks. *Ab initio* calculations, within MARVEL, are already ongoing in order to predict the local atom structure of amorphous Al₂O₃. As soon as the structural details has been calculated, the electronic properties can be derived, which will be validated by other experimental techniques.



Figure 2: (a) EXAFS results on TEY mode on Al anodized layer compared with references Al-oxides and hydroxides with known structure. (b) NMR results on anodized Al sample showing the 3 different Al coordinations present: tetragonal (4), pentagonal (5) and hexagonal (6).

1.4 Planned research for next year

The MARVEL postdoc will start officially in our laboratory (Joining Technologies and Corrosion) at Empa in January 2016. During the first stage of his project, we will maintain our focus on the growth of amorphous Aloxides by electrochemical methods and the effect of the growth conditions on the developing oxide-film microstructure (e.g. thickness, composition, morphology, defect structure, electronic properties). The experimental strategy envisaged in this project is to investigate the surface reactivity of defective Al₂O₃ films, considering different degrees of structural disorder. To this end, the microstructure-property relationships of non-crystalline, defective Aloxide films grown electrochemically will be characterized and compared with those of less-defective Al₂O₃ oxide and oxyhydroxides phases grown by, e.g., thermal oxidation, reactive magnetron sputtering, and PVD/ALD. This part represents the core activity and most challenging aspect of the project. Finally, the stability of the grown oxide films under operating conditions will also be investigated, considering reactivity, hydroxide formation, corrosion resistance and influence of defects. During each successive experimental step (addressing crystalline and defective oxide structures, as well as hydroxide formation at reacting oxide surfaces), active interaction with the modeling efforts in the MARVEL framework will be striven for, aiming at comprehensive knowledge and model predictions of the structural and electronic properties of functional oxide surfaces after growth, processing and subsequent operation. The laboratory of Joining Technologies and Corrosion offers the following experimental platform and competences.

• Sample preparation of thin functional ox-

ide layers using different techniques and growth methods, including electrochemical anodizing, PVD, ALD and thermal oxidation.

- Electrochemical characterizations (core competence of the laboratory): in particular, experimental investigations of the electronic properties and surface reactivity of semi-conducting oxides by electrochemical impedance spectroscopy (EIS), photoelectrochemistry (implementation of an *in situ* technique during the anodizing process) [1] as well as specific local electrochemical methods such as environmental AFM operating under controlled atmosphere (liquid, gas, temperature) [5] and local capillary based methods for oxide defect identification.
- Cutting-edge experimental investigations of the oxide-film microstructures by a combinatorial experimental approach using various surface-analytical techniques, both in the laboratory (XPS, AES) [6, 7] and at the synchrotron (see below). Structural analysis will be also performed by Xray diffraction methods and transmission electron microscopy (TEM) at Empa [2].

Atom-specific synchrotron-based techniques like NEXAFS are currently explored and applied to reveal local atomic environments in non-crystalline, defective oxide phases, as well as to reveal oxide-to-hydroxide transformations at functional oxide surfaces in contact with water. Nuclear magnetic resonance (NMR) will also be performed to derive the local chemical environment and different ratios of oxide coordination number for thick oxide films and bulk oxide references (not applicable to thin oxide films).

1.5 Synergies with other computational and experimental efforts

a) Within MARVEL A kick-off meeting with the theoretical group from EPFL (Alfredo Pasquarello) to discuss about tasks and workplanning was held at Empa. Future meetings and discussions will be planned on a regular basis (e.g. every 2 months) once the project will officially start in January 2016 with Fabio Evangelisti, MARVEL postdoc for two years. Among the issues to be tackled by the modelers (e.g. energy-level alignment), the longterm stability of the semiconductor/oxide surface in contact with humidity/water probably is the most-challenging one. As a first step, the effect of hydrogen incorporation in the oxide subsurface on the resulting defect structure and electronic band structure will be addressed. As a next step, simulations will be attempted to investigate the oxide-to-hydroxide transformation at the oxide surface in contact with water. Absorption and photoemission spectroscopy techniques will be used to investigate the microstructure-property relationship of oxide films after growth and controlled water exposure. The thus obtained experimental results can be compared with the model predictions. Very likely, additional novel experiments will have to be designed to validate theory predictions. With the group of Daniele Passerone, Carlo Pignedoli and Daniele Scopece, we are actively collaborating on simulating the XAS spectra obtained on amorphous Al_2O_3 and α - Al_2O_3 and the XAS response of defective alumina structures.

b) *Outside MARVEL* Collaborations with T. Huthwelker and C. Borca from Swiss Light Source (SLS), PSI are ongoing to measure absorption edge of different oxides.

1.6 Personnel

NCCR personnel

• Fabio Evangelisti, postdoc, 50%, from January 2016.

Matching personnel

• Fabio Evangelisti, postdoc, 50%, from January 2016.

2 Theory and experiment synergy for artificial photosynthesis (Thomas Lippert — PSI)

2.1 Research summary

A possible answer to the world's increasing energy demand is the conversion of solar energy into chemical energy by artificial photosynthesis, which converts water into H₂ and O₂ by photocatalytic solar water splitting. Oxynitride perovskites are among the most promising candidates to achieve one-step, efficient water splitting under solar light irradiation. However, to date, a stable semiconductor enabling overall water splitting has not been found. This project is based on very close synergies between high-throughput theoretical calculations on tens of thousands of candidate pervoskites, and the experimental synthesis, thin film growth, characterization and testing of the most promising materials identified.

2.2 Scientific goals

Recent theoretical studies by Castelli and co-workers [8, 9] report a comprehensive high-throughput computational screening of 19'000 oxides and oxynitrides in the cubic perovskite structure to identify semiconductors suitable for photo-electro-chemical water splitting under solar light. Based on criteria for structural stability and for the size and position of the band gap, the screening selected a handful of oxynitride materials for further investigations. Some of the oxynitride materials have already been synthesized showing electronic properties in good agreement with the theoretical prediction.

The screening selected other compounds as potentially highly performing photocatalysts that have not been investigated experimentally so far; these are new N-substituted tantalate and niobate perovskites with the chemical compositions MgTaO₂N and MgNbON₂, as well as YTaON₂ and YNbON₂.

a) *Short-term* LaTaON₂ and CaNbO₂N are well-known Ta- and Nb-based oxynitride photocatalysts. These two materials will be synthesized and used as reference materials for the new compositions under investigation.

The preparation of the new compounds will then be addressed in form of powders using standard ceramic processing routes (solid-state reaction and high-temperature ammonolysis). The fabrication of some of the proposed materials may turn out to be experimentally not possible at all or not possible in the expected perovskite structure. Moreover, the experimentally measured band gap may be different to the theoretically expected.



During this initial phase of the project, the feedback with theoretical calculations focusses on the structural and chemical stability of the new materials, as well as the size of the band gap.

b) *Long-term* Photo-electro-chemical (PEC) characterization will be performed with the materials selected after the first screening described above. The standard 3-electrode configuration will be used with a Xe lamp to simulate the solar spectrum.

Potentiodynamic and potentiostatic measurements of the photocatalytic activity of the different materials will be performed, as well as long-time stability test to probe their chemical stability in operating condition.

The synthesis of new oxynitride materials will also be investigated using pulsed reactive crossed beam laser ablation (PRCLA) [10] to grow thin films with different crystalline and crystallographic properties and probe how these affect the photocatalytic activity.

Thin films "as grown" and after PEC measurement will be used to analyze by X-ray photoemission spectroscopy the surface modifications induced by the photo-electro-chemical processes.

Finally, combining X-ray absorption and emission spectroscopy, resonant inelastic X-ray scattering (RIXS) maps will be acquired at the Swiss Light Source (SLS) at PSI. From the RIXS maps, the density of states can be calculated [11, 12], thus achieving more insights on the mechanism leading to a reduction of the band gap as a consequence of the N substitution into the O sites.

2.3 Results obtained since the last report

This is our first report of the first 6 months of this project which started in June 2015.

Among the new materials potentially attractive for solar water splitting selected by theoretical calculations, the system Y-Ta-N-O was first investigated. The first step we addressed was the synthesis of the perovskite structure YTaON₂. As a reference compound the wellknown LaTaON₂ oxynitride photocatalyst was used. The fabrication, photo-electrode preparation, PEC measurement of the reference material were successfully achieved.

The new Y-Ta-O-N system was then investigated. The starting oxide YTaO₄ was synthesized either by solid-state reaction of Y_2O_3 and Ta_2O_5 , or polymerized complex method. Through thermal ammonolysis of YTaO₄, the pyrochlore $Y_2Ta_2O_5N_2$ and the fluorite YTa(O,N, \Box)₄ have been successfully fabricated at various temperatures, ammonia flows and treatment time.

Perovskite YTaON₂ also appeared as a small trace along with the main fluorite YTa(O,N, \Box)₄ phase after very long ammonolysis time, but it was impossible to isolate the perovskite structure YTaON₂, indicating the perovskite structure is less stable than the other two structures for this system.

This first result on the stability of the predicted perovskite structure for this system is currently the subject of further investigation from the theoretical point of view. The refinement of previous calculation is ongoing as well as the development of new models to describe the optical properties of the two crystalline structures experimentally synthesized.

UV-Vis diffuse reflectance showed that the absorption edges of the two structures are both within the visible light range. The band gaps are determined to be 2.48 eV and 2.14 eV for pyrochlore and fluorite structures, respectively. The oxynitride powders have been deposited on FTO glass by electrophoretic deposition. Preliminary PEC measurements have been successfully performed and the two structures showed responses in both alkaline and neutral electrolytes under a 404 nm monochromatic light illumination.

We consider this an important result since, even though with a different than expected crystalline structures, new oxynitride materials have been synthesized and tested, showing potentials for solar-driven water splitting.

We also started the investigation of the system Y-Nb-O-N for which theoretical simulations foresaw potential applicability for solar water splitting for the perovskite structure YNbON₂. The results are still preliminary but it seems that also for this system the perovskite structure is unstable, while the fluorite oxynitride can be synthesized.

2.4 Planned research for next year

As mentioned above, in the case of the Y-Ta oxynitride system the predicted perovskite structure turned out to be unstable while different structures were successfully synthetized. This may apply also for some of the other selected compounds. In such a case, new theoretical calculations will be performed to model structural stability and optical properties of the new materials.

The synthesis and characterization of the new oxynitride materials will continue to test all the originally selected compounds but also other materials may be included.

The synthesis of the new oxynitride materials in form of thin films will be addressed. PRCLA is the selected deposition method. Our previous studies show that PRCLA is an effective route, alternative to conventional techniques, for N incorporation in oxides [10].

Oxinitride thin films fabrication and test have rarely been reported in the literature. Thin films are ideal model systems to investigate the effect of crystalline quality and crystallographic surface orientation on the photoelectro-chemical properties. This investigation is not possible with commonly used powder samples.

Among the new oxynitride materials, those with good chemical and structural stability and appropriate band gap will be selected.

PEC characterization of the new powder materials will be performed and compared with the reference compounds. For the photocatalytic characterization of thin films, TiN will be used as current collector following a wellestablished experimental approach developed in our laboratory.

On the base of the experimental results, a refinement of the theoretical simulation may be needed to match potential differences in size and position of the band gap.

If different photocatalytic properties will be observed in the case of thin films as a function of different crystallographic orientations of the surface, theoretical simulation will be particularly important to rationalize this effect.

The more promising materials will be used for two SLS campaigns planned for the next year to identify the modifications induced to the electronic structure by the N substitution [11, 12].

2.5 Synergies with other computational and experimental efforts

a) *Within Marvel* Since the project starts (June 2015), the research team met during the scheduled MARVEL meetings at EPFL and at Empa, at the PP7 MARVEL meeting at PSI, as well as several times through Skype meetings to discuss the specific issues of the experiment. The next important meeting is scheduled for the beginning of January 2016 to discuss recent developments that require additional feedback from the computational side to the experimental activity.

b) *Outside MARVEL* Regular proposals will be submitted in collaboration with J. Szlachetko for the allocation of suitable beamtime at the SLS for high-energy X-ray spectroscopy on the new oxynitride materials.

Within this MARVEL project a close collaboration on the computational side is planned for the next year with the group of Prof. K. W. Jacobsen at the Technical University of Denmark, author of the two theoretical studies [8, 9] from which this project started.

2.6 Personnel

NCCR personnel

• Wenping Si, postdoc, 50%, from July 2015.

Matching personnel

• Wenping Si, postdoc, 50%, from July 2015.

3 Quest for the Higgs-mode in hexagonal manganites (Dirk van der Marel — UniGE)

3.1 Research summary

Recently a new phenomenon has been added to the already rich physics of hexagonal manganites: a hidden form of antiferroelectric order was observed - and described theoretically - in addition to the well-known improper ferroelectricity in this material [13]. The duality of these two states of matter, and the fact that the energy cost for moving from one to the other is rather small, opens novel perspectives for engineering topological objects, in particular ferroelectric and antiferroelectric vortices. Another consequence, which is very interesting from a fundamental perspective, comes from the peculiar energy-landscape for the collective coordinates describing aforementioned order: when improper antiferromagnetism and ferromagnetism are degenerate, this implies that the energy as a function of order collective coordinate takes the shape of a sombrero-shaped potential with an approximate U(1) symmetry [14, 15]. The lowtemperature phase will have the order parameter falling inside the rim of the sombrero, however, as a result of the U(1) symmetry the direction around the hat will be a random one. Spontaneous symmetry breaking occurs also in superconductors, ferromagnets, and the Higgs-field (Fig. 3), to mention just a few examples, and is revealed by the appearance of new collective modes. The aim of the present study is to demonstrate these modes experimentally in the hexagonal manganites.





Figure 3: Potential landscape of the energy as a function of tilting of the MnO₅ octahedra. The amplitude mode is equivalent to a Higgs mode. The rotation of the tilt angle gives rise to the Goldstone mode in the limit vanishing six-fold corrugation of the potential landscape.

3.2 Scientific goals

a) Short-term In the hexagonal manganites, the spontaneous U(1)-symmetry breaking associated with either ferro- or antiferro states brings the lattice coordinates (in the present case a tilting of MnO₅ octahedra) in a state whereby the azimuthal angle of the tilt is free to rotate (Fig. 4). Note, that obtaining the perfect U(1) symmetry requires appropriate materials engineering: h-InMnO₃ has a six-fold denting of the rim of the hat, with 4 meV potential barriers separating neighboring minima around the perimeter of the hat. In rapid quenched h-InMnO₃ the denting pattern shifts by 30 degrees, and the potential barriers are reduced to 1 meV. Further fine-tuning by partial substitution of Ga on the Mn-sites make it possible to suppress the potential barriers entirely, allowing the realization of a completely flat rim around the Mexican hat [13]. The short-term goal is to demonstrate the first experimental realization of the Higgs-mode in the context



Figure 4: Crystal structure of InMnO₃, figure adopted from [16].

of ferroelectricity, and to determine its modefrequency.

b) *Long-term* According to the Goldstone theorem, a new type of "sound" should emerge, corresponding to the collective motion of the coordinate describing the azimuthal rotation [17]. The long-term goal of this study is to demonstrate experimentally the Goldstone mode in the context of U(1) symmetry in an improper (anti-)ferroelectric system, and to determine the corresponding sound-velocity. We aim to ultimately measure the effect on the Goldstone mode dispersion of the discrete sixfold rotation symmetry away from optimal tuning.

3.3 Results obtained since the last report

We have performed infrared and polarization resolved Raman spectroscopy of ErMnO₃ and two types of InMnO₃ (para- and ferroelectric). We performed polarization resolved Raman spectroscopy on the three samples in a range between 50 and 850 cm⁻¹. We used two differents polarization configuration (z,(x,x)z) and The characterization of the polarz(x,y)z).ization has been done by analysing the wellknown extinction of the 514 cm^{-1} phonon peaks in Si. Each sample has been measured from 10 K to 350 K for each polarization. Plots of experimental data are shown in Fig. 5. For sake of visibility curves have been shifted on the vertical axis proportionally to their temperatures. Interestingly, the energy and tempera-



Figure 5: Raman spectra of ErMnO₃ (sample displayed in top right corner) for temperatures from 10 K (bottom curve) to 350 K (top curve). Polarization is $z(x, \bar{x})\bar{z}$. The conspicuous broadening of the mode at 250 cm⁻¹ at elevated temperatures is a first indication of Higgs-related physics, and arise from thermal activation of rotational motion of the MnO₅ octahedra due to the flat potential landscape shown in Fig. 3.

ture dependence of one of the observed modes follows the expected behavior, making it a potential candidate for the Higgs mode.

3.4 Planned research for next year

Additional Raman spectroscopy at low frequencies will be carried out on ErMnO₃, InMnO₃, HoMnO₃ and YMnO₃ to study in detail the temperature dependence of the observed soft modes. New measurements of HoMnO₃ and YMnO₃ will increase our window of investigation on hexagonal manganites. Vibrational spectra will be calculated using the frozen-phonon method as implemented in the PHONOPY software [18, 19]. Further experiments will be carried out using electricfield gating of the samples. The instrumentation used will be THz time-domain spectroscopy for the lowest frequencies (1 cm^{-1} to 100 cm^{-1}).

3.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* Our experimental work has been suggested by the group of Nicola Spaldin (ETHZ).

3.6 Personnel

NCCR personnel

• PhD student position not yet filled.

Matching personnel

• Adrien Stucky, PhD student, 80%, from May 2015.

4 RNiO₃ perovskites: exploring the boundary between localized and itinerant behavior (Marisa Medarde — PSI)

4.1 Research summary

The first three months of the project were devoted to identify the experimental conditions necessary to synthesize the $RNiO_3$ perovskite family (R = rare earth). Large (~1 g) powder samples of the nickelates with R = La to Dy, as well as DyNiO₃ nanocrystals could be prepared under high O₂ pressures (P_{max} = 2'000 bar). Neutron diffraction measurements, scheduled for mid-November 2015, should provide novel insight into the charge and magnetic order that will be compared with theoretical predictions. Ceramic samples of the remaining nickelates and larger single crystals of YNiO₃ are expected to be available in 2016.

4.2 Scientific goals

a) *Short-term* The main experimental goal for the first three months was the identification of the experimental conditions necessary to synthesize all members of the $RNiO_3$ perovskite family (R = rare earth). For the theory partners, the main objective was the coordination of the efforts of the two teams in order to exploit optimally the experimental data, some of which should already be available at the end of 2015.

b) *Long-term* (i) Improve the theoretical understanding of the metal-insulator transition (MIT) in $RNiO_3$ perovskites. (ii) Test state-of-the-art theoretical tools on a model system particularly clean. (iii) Advance in the under-

standing of the physics relevant to the region at the limit between localized and itinerant behavior.

4.3 Results obtained since the last report

a) Sample synthesis The origin of the MIT in $RNiO_3$ perovskites has been the main open question in this nickelate family since the first report of the transition in 1991 [20]. Although significant progress has been done during the last 25 years [3], in particular due to the recent availability of good quality thin films, the challenging high-pressure synthesis procedure remains a serious stopper. The O₂ pressure required to stabilize the Ni³⁺ formal valence increases by decreasing the size of the lanthanide ion, requiring the use of highly specialized equipment. Moreover, the mass of the samples synthesized under very high pressures is usually very small, typically below 300 mg.

Since sample amounts of ~1 g are necessary for many experimental techniques, the main goal of the first three months of the project has been to investigate the synthesis conditions necessary to obtain such an amount using the oxygen-gas high-pressure synthesis setup recently installed in our laboratory (Fig. 6). Being the maximal available oxygen pressure relatively modest (2'000 bar, to be compared with pressures up to 60'000 bars used by some authors [21]), our efforts have been focussed on the possible ways to increase the reaction kinetics.



Figure 6: High oxygen-gas pressure setup at the Laboratory for Scientific Developments and Novel Materials (PSI).

For the nickelates with R = La, Nd, Sm, Eu and Gd, this could be achieved using precursors obtained by soft chemistry methods. As shown in Fig. 7, pure powder samples could be obtained following this procedure. For DyNiO₃, which has a significantly smaller ionic radius, this approach failed and an alternative method had to be used. In order to increase the reactivity of the precursors, they were mixed with LiCl and KCl in 1:5 weight ratio. These low-melting-point ionic salts (605 and 770°C, respectively) provided an ionic liquid flux that enabled the formation the DyNiO₃ phase (Fig. 7). Interestingly, DyNiO₃ nanocrystallites were obtained, raising reasonable hope for the growth of larger single crystals of the nickelates with smaller rare-earth ions.

b) *Structural studies* Understanding the lattice distortions and its relationship with the charge redistribution at the MIT requires a precise determination of the oxygen positions and hence the use of neutron diffraction. The first experiments are scheduled for the second half of November 2015 at the SINQ (PSI). In order to investigate also the nickelates with R = Sm, Eu and Gd, which present an extremely large neutron capture cross section, the less absorbing ¹⁵⁴Sm, ¹⁵³Eu and ¹⁶⁰Gd isotopes have been purchased.



Figure 7: *X*-ray powder diffraction patterns of the $RNiO_3$ nickelates with R = La to Dy.

c) *Transport properties* Investigating the existence of multiferroicity in the nickelates, in particular for the most insulating ones, is one of the main goals of the project. Ferroelectricity arising from combined magnetic and charge order has been theoretically predicted to exist below T_N [22], but to our best knowledge never reported to date. For this purpose, we have designed and built an high-voltage stick compatible with our PPMS device, and we have tested it during the last two weeks of October 2015. Electric polarization can now be measured between 5 and 400 K under applied voltages of \pm 400 V and magnetic fields up to 9 T.

d) *Spectroscopic studies* X-ray absorption measurements aimed to investigate the evolution of the Ni 3d - O 2p hybridization have been performed at the SLS in collaboration with Marco Grioni (EPFL) and Bernard Delley (PSI). NMR measurements have been carried out by Joël Mesot's group (PSI and ETHZ). The analysis of the obtained data is in progress.

4.4 Planned research for next year

The priority for next year is to complete the synthesis of the full nickelate series in powder form and to investigate the transport properties. Time permitting, this will be extended to some doped and/or $RNiO_3$ solid solutions, to be used for thin film growth by Ming Shi's group (SLS). Temperature-dependent studies of the crystal and magnetic structures will be carried out at the SINQ, giving preference to the ¹⁵⁴Sm, ¹⁵³Eu and ¹⁶⁰Gd nickelates. We will also try to grow YNiO₃ single crystals large enough to do X-ray diffraction and extract electron density maps.

4.5 Synergies with other computational and experimental efforts

a) Within MARVEL We had two meetings with the theory groups of Antoine Georges and Nicola Spaldin (one separately, one all together). The collaboration will be based on continuous experiment-theory feedback. On one side, the new experimental data obtained on RNiO₃ perovskites will be used as input for the description of the different aspects of the electronic structure of these materials. On the other side, the theoretical predictions on aspects such as the decomposition of the lattice distortions as superposition of "frozen" modes, the macroscopic magnetic properties or the details of the charge redistribution across the metal-insulator transition will be experimentally investigated and the results contrasted with the predictions. Further collaboration is foreseen with Nicola Marzari, who expressed his interest on joining the theory team.

His main interest is the study of the charge redistribution in nickelates within the broader framework of charge order phenomena in oxide perovskites.

b) *Outside MARVEL* Collaboration with Marco Grioni (EPFL) and Bernard Delley (PSI) for the investigation of the Ni 3d - O 2p hybridization using X-ray absorption spectroscopy.

4.6 Personnel

NCCR personnel

• Dariusz Gawryluk, postdoc, 50%, from August 2015.

Matching personnel

• Dariusz Gawryluk, postdoc, 50%, from August 2015.

5 Oxygen evolution reaction on $La_{1-x}Sr_xCoO_3$ perovskites: a combined experimental and theoretical study (Thomas J. Schmidt — PSI)

5.1 Research summary

Experimental and theoretical studies on $La_{1-x}Sr_xCoO_3$ perovskite series have suggested that the oxygen evolution reaction (OER) activity is driven by the perovskite electronic structure and benefits by an increased overlap between the occupied O 2*p* valence bands and the unoccupied Co 3*d* conduction bands. In the next project year, metal-like conducting perovskites will be investigated with particular focus in correlating band structure and electrochemical activity towards the oxygen evolution reaction.

5.2 Scientific goals

The main goal of the present project is to provide a fundamental understanding of the oxygen evolution reaction (OER) mechanism and of the design principles governing the OER of highly active perovskite catalysts. The achievements of this ambitious project could represent an important breakthrough in the development of cost-effective, efficient and reliable energy storage devices. This goal will be gained by integrating electronic-structure calculations and electrochemical interface properties into a single model.

a) *Short-term* In the short term, we want to identify potential OER catalysts, which not only display high activity towards the OER but

that can also help in a better understanding of the design principles governing the OER. The selected perovskite is $La_{1-x}Sr_xRuO_3$ (LSRO) with $0 \leq x \leq 1$. Varying the Sr content in the LSRO, the conduction mechanism remains metallic-like (which our previous investigations has shown to to be beneficial for the OER, as extensively explained in paragraph 5.3), but the oxidation state of Ru ions should change for charge balancing, being La in 3+ oxidation state and Sr in 2+ oxidation state. This means that varying the A-site from La to Sr the number of d electrons, and thus, the band structure also changes. Therefore, this perovskite series represents, both from a theoretical and experimental point of view, a particular suitable model system for understanding the influence of the electronic structure on the oxide electrochemical interface and, thus, on the electrochemical activity.

b) *Long-term* The novelty of the present research project lies in the emphasis pointed towards the fundamental importance of understanding the electrochemical interface level in operando conditions. This will be achieved via several physico-chemical and electrochemical investigations, and particularly with *in situ* X-ray absorption spectroscopy (XAS) measurements of nanocatalyst systems which will provide snapshots of the electronic states under operational conditions. Being nanopowders

dominated by the surface, changes in the electronic and geometric structure at the surface level can be detected by operando XAS. The experimental study will be coupled and validated by density functional theory (DFT) simulations of the electronic structure of the selected catalysts. DFT calculations will be carried out not only in the bulk, but also at the surfaces under real operative conditions, where the solvent, the electrolyte, and the applied potential is included in the calculation thanks to the multiscale model developed by the Marzari group at EPFL.

5.3 Results obtained since the last report

We have investigated from an experimental and theoretical point of view the bulk electronic structure, surface composition, conductivity and electrochemical activity towards the oxygen evolution reaction (OER) for the $La_{1-x}Sr_xCoO_3$ perovskites (with x = 0, 0.2, 0.4, 0.6, 0.8, 1), which are promising candidates as OER catalysts. It is found that Sr substitutions have the effect of aligning atoms along the Co-O-Co axis and increasing the average oxidation state of the Co cations. The increase of Sr content also improves the ex situ electronic conductivity as well as the activity towards the OER. According to DFT calculations, the alignment of the Co-O-Co bonds and the oxidation of the Co cations enhance the overlap between the occupied O 2*p* valence bands and the unoccupied Co 3d conduction bands (Fig. 8), thus also increasing the ex situ electronic conductivity as a function of the Sr faction.



Figure 8: Projected density of states for the O 2*p* and Co 3*d* states for $La_{1-x}Sr_xCoO_3$, as obtained in the DFT (PBEsol) calculations. The calculations are left fully free to relax, and a hexa-rombohedral unit (6 × ABO₃) is used for $x \leq 0.4$, followed by a supercell of the 2 × 2 × 2 cubic primitive cell for $x \geq 0.6$, in this latter case the composition steps for the A-cations can only be multiples of 1/8 = 0.125). The DFT calculations have been carried out by the Marzari group at EPFL.



Figure 9: Current density (μ A cmoxide-2) as a function of the ex situ electronic conductivity of $La_{1-x}Sr_xCoO_3$ series; the red circle represents the SrCoO_{2.5}.

Since as the electronic conductivity increases the OER activity of $La_{1-x}Sr_xCoO_3$ is enhanced (Fig. 9), we suggest that the OER activity is driven by its electronic structure and benefits by an increased overlap between the occupied O 2*p* valence bands and the unoccupied Co 3*d* conduction bands. This outcome can significantly help in the search for more active OER catalysts. The experimental and theoretical results on the $La_{1-x}Sr_xCoO_3$ perovskite catalysts is currently under consideration in *Chemistry of Materials* [4].

5.4 Planned research for next year

In the first part of the next year, the synthesis and characterization of $La_{1-x}Sr_{x}RuO_{3}$ (LSRO) perovskites will be carried out. LSRO will be produced as nanopowders by the flames pray method. The as-synthesized powder will be characterized by X-ray diffraction (XRD) in order to establish the material The nanopowders will be characphases. terized also by (i) X-ray photoelectron spectroscopy (XPS) for the surface chemical composition, (ii) Brunauer-Emmett-Teller (BET) for the surface area estimation, (iii) scanning electron microscopy (SEM) and transmission electron microscopy (TEM) for the microstructural investigation, and (iv) rotating ring electrode (RDE) technique for the measurement of the activity of the perovskite catalysts towards the OER.

In the second part of the year, *in situ* XAS measurements will be carried out. The *in situ* XAS measurements will be performed at the SuperXas beamline at the Swiss Light Source (SLS) of PSI in close collaboration with the group of Maarten Nachtegaal. Time-resolved

XAS (TR-XAS) will provide relevant information about *in situ* changes of the catalyst electronic and geometric structure. Particularly, due to the SuperXAS beamline ability of acquiring XAS spectra within 500 ms, the potential will be scanned in the 0 - 1.5 V RHE range and XAS spectra recorded during the anodic and cathodic potential sweep.

5.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* An active collaboration with the group of Nicola Marzari at EPFL has been already established and it has resulted in a joined publication now under consideration in *Chemistry of Materials* [4]. The results ob-

tained from the planned experiments for the next year (see paragraph 5.4) will be also correlated to DFT simulations of the perovskite band structures carried out by the group of Nicola Marzari at EPFL.

5.6 Personnel

NCCR personnel

• Daniel Abbott, postdoc, 50%, from August 2015.

Matching personnel

• Daniel Abbott, postdoc, 50%, from August 2015.

6 Resonant inelastic X-ray scattering on thin films and oxide heterostructures for future mottronics and orbitronics (Thorsten Schmitt — PSI)

6.1 Research summary

Our goal is to study LaTiO₃ and LaVO₃ thin films strained via different substrates as a function of temperature with X-ray absorption spectroscopy (XAS) and resonant inelastic X-ray scattering (RIXS). We plan to reveal changes in the orbital levels and the degree of localization of the valence band across the metal-insulator transition (MIT) of these prototypical Mott systems. Understanding such strained films in close interplay with DFT+DMFT calculations is important to predict properties and a prerequisite for orbital engineering of oxide heterostructures like superlattices with confined slabs of such transition metal oxides.

6.2 Scientific goals

We aim to study LaTiO₃ and LaVO₃ films and superlattices as a function of strain and temperature combining XAS and RIXS. With this we aim at understanding the electronic structure changes in terms of Ti/V 3*d* orbital reconstruction, Ti/V 3*d* – O 2*p* covalence and degree of electronic localization when driving these films into a metallic state.

a) *Short-term* We expect to obtain information on the Ti/V 3*d* levels and their orbital occupation in LaTiO₃ and LaVO₃ thin films and heterostructures through measurements of the onsite *dd*-excitations with RIXS and by comparison with state-of-the-art DFT+DMFT calculations from our MARVEL collaborators. In particular, we will be able to track changes in the splitting among the t_{2g} and e_g levels across the

MIT, giving information on the orbital reconstruction as a function of substrate strain and temperature. From this we will gain deeper knowledge of the MIT tune-ability in the prototypical Mott-materials LaTiO₃ and LaVO₃. We will correlate the changes in the electronic structure seen by XAS and RIXS to the corresponding changes in the conductivity. This will allow studying how orbital reconstruction phenomena induced by strain, corresponding changes in the Ti/V 3d - O 2p hybridization and changes of the degree of electron localization/delocalization are connected to the MIT mechanism. Furthermore, in combination with the DFT+DMFT calculations, this will lead to a better understanding of how various factors such as strain, confinement effects, or genuine interface effects determine the properties of such complex oxide heterostructures.

b) Long-term Comparison of the experimental results with DFT+DMFT calculations within MARVEL will ease better understanding of the MIT in such thin films and heterostructures, a prerequisite for predicting properties as well as for orbital engineering of oxide heterostructures for future mottronics and orbitronics applications in multifunctional electronic devices switching their state based on undergoing a MIT or an orbital reconstruction. Our experimental results will provide important benchmarks for the accuracy and predictive power of the advanced quantum simulations developed in HP3, in particular the DFT+DMFT approach, for such inhomogeneous materials systems with distinct properties due to electronic correlation effects.



6.3 Results obtained since the last report

We have searched for a suitable person for carrying out this project as a postdoctoral project. The selected scientist will start working on this project in January 2016.

6.4 Planned research for next year

We plan to grow by pulsed laser deposition (PLD) epitaxial films of LaVO₃, 25 nm thick, on various substrates. The films will be characterized with X-ray diffraction to verify the strain conditions and by resistivity measurements to determine the temperature for which the electronic behavior changes from insulating to metallic. We will deposit LaVO₃ thin films on at least three different substrates exerting compressive strain. From the experience of investigating strained LaVO₃ films, we will then extend our studies towards confined LaVO₃ tuned to different thickness by varying n, the number of LaVO₃ unit cells, in superlattices. All films and superlattices will be investigated by XAS and RIXS to reveal changes in the orbital levels and the degree of localization of the valence band across the corresponding MIT. These experimental results will create feedback to the DFT-DMFT calculations of the theory partners.

6.5 Synergies with other computational and experimental efforts

a) *Within MARVEL* The thin oxide films we propose to study are ideal for benchmarking

the advanced quantum simulation methods developed within the MARVEL project HP3 (in particular the DFT+DMFT approach) and to apply them to the study of novel materials physics found within artificial oxide heterostructures (VP1). Specifically, there are ongoing activities in the Spaldin group (research team led by Claude Ederer), applying DFT+DMFT-based methods to study the effect of epitaxial strain on the MIT in heterostructures of LaTiO₃ and other early transition metal oxides with d^1 and d^2 electron configuration (e.g. SrVO₃, LaVO₃, and CaVO₃). Our expected results for LaVO₃ films grown on different substrates with varying lattice mismatch will allow verifying these theoretical results and a detailed comparison between theory and experiment will lead to a better understanding of the role of strain in such oxide heterostructures. We had already a kick-off meeting between the experimental and theory partners by telephone and have regularly exchange by e-mail to refine the objectives of the starting project.

6.6 Personnel

NCCR personnel

• Daniel McNally, postdoc, 50%, from January 2016.

Matching personnel

• Daniel McNally, postdoc, 50%, from January 2016.

7 Time-resolved X-ray absorption spectroscopy to investigate mechanisms of photochemical water splitting reactions with molecular catalysts (Grigory Smolentsev — PSI)

7.1 Research summary

The goal of the project is to determine the molecular mechanism of photocatalytic water splitting by a cobalt cubane catalyst. Computational predictions about the structure of intermediates and oxidation state of metal centers will be verified using time-resolved X-ray absorption spectroscopy in the microsecond time range.

7.2 Scientific goals

A technology to harness solar energy to produce fuels is highly desirable, as it can potentially replace fossil fuels by a renewable energy source without a carbon footprint. The rational computational development of molecular systems that catalyze such reactions requires detailed knowledge of the reaction mechanism. We focus on the development of promising systems consisting of earth abundant metals. Accordingly, the goals of the project are the following.

a) *Short-term* Investigate reaction intermediates formed in Co-based photocatalytic water splitting systems. The first target for our investigation will be a promising cobalt cubane catalyst as shown in Fig. 10. It reveals unprecedented parallels to the oxygen-evolving center of photosystem II, which is the most important model system for biomimetic water oxidation catalysts. We aim at clarifying the reaction mechanism by determination of the local electronic and atomic structure of all rele-



Figure 10: *Structure of* Co₄O₄ *cubane catalyst*

vant intermediates formed during the catalytic cycle based on the time-resolved X-ray absorption data.

b) *Long-term* Development of more efficient earth abundant molecular catalysts for water splitting on the basis of computational predictions and experimental verification of catalytic mechanisms using time-resolved X-ray spectroscopy.

7.3 Results obtained since the last report

The project is still in the preparation phase, the contract of the MARVEL postdoc has not started yet. As a preliminary step we optimized the experimental setup for the planned experiments. In particular, the detection system has been optimized for measurements with a time resolution of 500 ns and for samples with a concentration of the metal of interest below 1 mM. In house beamtime for experiments on the Co cubane system at the SuperXAS beamline of Swiss Light Source (SLS) has been reserved for the next experimental cycle (January-June 2016).

7.4 Planned research for next year

Next year will be focused on deciphering the mechanism of water splitting catalyzed by a Co cubane catalyst (Fig. 10). More precisely, we would like to (i) determine the dominant oxidation state of Co atoms and the structure of the complex in the state that can be accumulated in solution upon visible light irradiation; (ii) with 500 ns time resolution we would like to track intermediate states formed within 50 microseconds after excitation by a short laser pulse; (iii) with 30 microseconds resolution we

would like to track intermediates formed during the first milliseconds of the reaction; (iv) finally we would like to determine if the decomposition of the catalyst limits the long-term operation of the system by measuring the system after the photocatalytic reaction (when it does not produce oxygen any more). For this purpose we will perform time-resolved X-ray absorption experiment using pump-flow-probe and pump-sequential-probe methods that we recently developed at the SuperXAS beamline. The synergy of efforts of theoretical and experimental groups will allow to enhance the understanding of the mechanisms of water splitting by molecular Co-based catalysts and design better catalysts by rational design. Timeresolved XAS spectroscopy alone provides qualitative information, for example about the oxidation state of the metal center. In order to extract quantitative information about the structure of intermediates, modeling of spectra for different structural models and comparison with the experiment is required. Such models can be obtained on the basis of quantum chemical calculations. On the other hand, DFT calculations often require some constrains from experiments, for example knowledge about the oxidation state. A few alternative models that have similar total energies often exist and additional arguments coming from the experiment are required. In all such cases comparison and verification of theoretical models with experimental data coming from XAS will be very fruitful.

7.5 Synergies with other computational and experimental efforts

a) Within MARVEL The contract of MARVEL postdoc has not started yet, but we already collaborate with the group of Jürg Hutter. On December 16, 2015 we had a meeting at UZH at which we have discussed plans for next few months, details of the next experiments at SuperXAS beamline and preliminary calculations. Before that we wrote together the beamtime proposal. The first experiment at SuperXAS beamline of SLS is scheduled for May 7 to 10, 2016. As another preliminary step, we are estimating the time-resolved X-ray absorption signal on the basis of DFT models that have been obtained by the group of Jürg Hutter. Next meeting is planned for the end of January 2016.

b) *Outside MARVEL* Synthesis of samples and characterization of their catalytic activity will be performed by the group of Prof. G. R. Patzke (UZH).

7.6 Personnel

NCCR personnel

• 50% postdoc position not yet filled.

8 Testing ultrafast processes in condensed matter (Urs Staub — PSI)

8.1 Research summary

Simulations of Hubbard-type models within the framework of dynamical mean field theory have predicted a series of novel transient states in antiferromagnetic Mott-insulating systems. These calculations also predict that the relaxation time scale of excited states critically depends on the spin background. Using ultrafast pump-probe techniques, we aim for first experimental tests of key predictions and to provide feedback for an improved and material specific modeling.

8.2 Scientific goals

a) Short-term The aim of this project is to test recent theoretical predictions for nonequilibrium phenomena in strongly correlated electron systems. Transient (but potentially long-lived) states may exhibit new and technologically useful properties [23, 24, 25]. Such states can be induced by non-adiabatic tuning of external driving forces or by other perturbations such as photo-excitation. Using recently developed theoretical tools Philipp Werner and collaborators have predicted a number of nontrivial transient states in antiferromagnetic Mott-insulating systems [26, 27]. Using ultrafast pump-probe techniques (optical and X-ray probe) we plan to test some of these results in order to validate the theoretical framework and to provide experimental input for an improved and material specific modeling.

b) Long-term A combined experimental and theoretical effort aimed at clarifying the nature of excited states in condensed matter is essential for further advances in our understanding of non-equilibrium phenomena and the eventual systematic development of new functional materials with specific transient properties. This project will push the development of both, cutting-edge theoretical tools as well as the development of suitable experimental methods. In particular, the large facility built at PSI (SwissFEL) opens up new possibilities to check theoretical predictions on the time evolution of electronic and magnetic states after a field-induced perturbation. The experimental verification of the theoretical results will not only advance our understanding on how to create and control new transient materials, but also push the theoretical tools towards an *ab initio* version of dynamical mean field theory, which will hopefully enable quantitative predictions for the non-equilibrium properties of specific materials. Such a tool will be essential for the long-term goal of designing functional materials, whose driven or otherwise perturbed transient states may possibly exhibit technologically relevant properties.

8.3 Results obtained since the last report

This activity has started in September 2015 when the associated postdoc position has been filled (Michael Porer). Optical pump-probe experiments and static X-ray characterization studies are planned for 2016. The first material to be tested is NiO, a prototypical Mottinsulating antiferromagnet.

8.4 Planned research for next year

a) *Background* Driving condensed matter systems out of equilibrium is a promising route for creating new functional materials, with short-lived (picoseconds) [23, 28] or longlived (days) [29] transient properties, which are distinct from those of known equilibrium phases.

The theoretical description of excited states and non-equilibrium dynamics in condensed matter systems is challenging, but there has been substantial progress in recent years, in particular thanks to the development of nonequilibrium dynamical mean field theory [26]. Simulations of Hubbard type models have provided insights into phenomena ranging from photo-doping and magnetization dynamics to non-equilibrium lattice effects. While some of these studies were aimed at explaining experimental observations, others have uncovered unexpected phenomena, which have yet to be systematically explored in real materials. Particularly relevant for the present project are several results obtained by Philipp Werner and collaborators on the properties of perturbed (e.g. photo-doped) antiferromagnetically ordered systems. For example, they found that



Matching personnel

• 50% postdoc position not yet filled.

long-range order in photo-excited antiferromagnets, described by the Hubbard model with strong coupling, survives for potentially long times after the excitation, which might be interpreted as a "dark state" [27], a state not existing in equilibrium and possibly observed recently in CuO [30]. Beyond a critical photo-doping concentration, this trapped state however is expected to disappear, and the relaxation dynamics is characterized by nonthermal critical behavior in the corresponding doping range. Another set of recent theoretical predictions concerns the nonlinear carrier transport in antiferromagnetic systems and the cooling of photo-doped carriers in such materials. In a strong field, as it exists for example in heterostructures with polar layers, the efficient dissipation of excess kinetic energy to the spin background in antiferromagnets dramatically enhances the mobility. This property may be essential for example in Mott insulating solar cells [31], and more generally for the design of functionalities in (light controlled) devices.

b) Experimental techniques Photon-based techniques have been adapted and employed to access the ultrashort time scales of electron-, spin-, and lattice dynamics required to investigate excited states in solid matter. Dynamic information using these photon pulses is typically obtained using the pump-probe technique. A first pulse excites the sample and a second pulse with a variable time delay "probes" the temporal evolution of the excited system until it relaxes back to the ground state. In particular, the advent of X-ray free electron lasers (XFEL) capable of producing very short and intense X-ray pulses offers new and exciting possibilities to study non-equilibrium transient states of materials. In recent years we have used ultrashort X-ray pulses to study structural, electronic and magnetic dynamics on ultrafast time scales. Structural dynamics has been studied using femtosecond hard X-rays at the SLS slicing (FEMTO) beamline [32, 33, 34], electronic and structural dynamics with XFEL radiation [35]. Investigations on excited, transient magnetic states have been performed by directly studying the magnetic subsystem with soft X-rays [36, 30].

c) *Experimental schedule* We start by identifying suitable materials, for which the model Hamiltonian calculations [27] approximately applies, or for which model calculations are technically doable. To guide the experimental verification and provide parameters for at least semi-quantitative model calculations, we will have additional input from DFT calculations provided by the group of Nicola Spaldin. The focus of the experiments will first be set on testing specific theoretical results by the group of Philipp Werner: the calculations predict a dramatic change in the relaxation time scale of the photo-excited state, when crossing from a paramagnetic phase into an antiferromagnetic phase. Experimentally, this can be achieved by lowering the temperature. The electronic relaxation dynamics of an above band-gap photo-excited system can be studied by an optical pump-optical probe experiment. This can be done with the laser system at FEMTO for example by collecting the pumpinduced changes of the optical reflectance of the material. To study the effect of photoexcitation on the magnetic lattice, ultrafast Xray experiments at a XFEL are ideally suited. As prerequisite, we will perform regular static X-ray characterization of select materials to be able to propose a XFEL experiment (LCLS or SwissFEL). In this second step, we will look for the predicted non-thermal magnetically ordered "dark states" and the doping dependence of this trapping phenomenon.

8.5 Synergies with other computational and experimental efforts

a) Within MARVEL This project is focused on the experimental validation of simulation results obtained in the group of Philipp Werner. The experiments will both test existing model predictions and provide a feedback for material-specific adjustments to the model system. For realistic model parameters, and general guidance on suitable materials, we collaborate with the group of Nicola Spaldin.

8.6 Personnel

NCCR personnel

• Michael Porer, postdoc, 50%, from September 2015.

Matching personnel

• Michael Porer, postdoc, 50%, from September 2015.

MARVEL-related publications

List of publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR.

 A. Beni, C. Cancellieri, C. Borca, T. Huthwelker, L. P. H. Jeurgens, and P. Schmutz, EXAFS and NMR investigations of amorphous Al-oxides: from experimental to theoretical validation, in preparation.

- [2] R. Gaspari, R. Erni, Y. Arroyo, M. Parlinska-Wojtan, J. Dshemuchadse, C. A. Pignedoli, D. Passerone, P. Schmutz, and A. Beni, *Real space crystallography of* a complex metallic alloy: high-angle annular dark-field scanning transmission electron microscopy of o-Al₄(Cr,Fe), Journal of Applied Crystallography 47, 1026 (2014).
- [3] A. Subedi, O. E. Peil, and A. Georges, Low-energy description of the metal-insulator transition in the rare-earth nickelates, Physical Review B 91, 075128 (2015).
- [4] X. Cheng, E. Fabbri, M. Nachtegaal, I. E. Castelli, M. El Kazzi, R. Haumont, N. Marzari, and T. J. Schmidt, Oxygen Evolution Reaction on La_{1-x}Sr_xCoO₃ Perovskites: A Combined Experimental and Theoretical Study of Their Structural, Electronic, and Electrochemical Properties, Chemistry of Materials 27, 7662 (2015).

Other references

- [5] K. Gałazka, S. Populoh, L. Sagarna, L. Karvonen, W. Xie, A. Beni, P. Schmutz, J. Hulliger, and A. Weidenkaff, *Phase formation, stability, and oxidation in (Ti, Zr, Hf)NiSn half-Heusler compounds, Physica Status So*lidi (a) **211**, 1259 (2014).
- [6] E. Ura-Binczyk, A. Beni, M. Lewandowska, and P. Schmutz, Passive oxide film characterisation on Al-Cr-Fe and Al-Cu-Fe-Cr complex metallic alloys in neutral to alkaline electrolytes by photo- and electrochemical methods, Electrochimica Acta 139, 289 (2014).
- [7] A. Beni, N. Ott, M. Pawelkiewicz, M. Wardé, K. Young, B. Bauer, P. Rajput, B. Detlefs, J. Zegenhagen, R. McGrath, M.-G. Barthés-Labrousse, L. P. H. Jeurgens, and P. Schmutz, *Hard X-ray Photoelectron* Spectroscopy (HAXPES) characterisation of electrochemical passivation oxide layers on Al-Cr-Fe complex metallic alloys (CMAs), Electrochemistry Communications 46, 13 (2014).
- [8] I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, and K. W. Jacobsen, *Computational* screening of perovskite metal oxides for optimal solar light capture, Energy & Environmental Science 5, 5814 (2012).
- [9] I. E. Castelli, D. D. Landis, K. S. Thygesen, S. Dahl, I. Chorkendorff, T. F. Jaramillo, and K. W. Jacobsen, *New cubic perovskites for one- and two-photon water splitting using the computational materials repository*, Energy & Environmental Science 5, 9034 (2012).
- [10] I. Marozau, A. Shkabko, M. Döbeli, T. Lippert, M. Mallepell, C. W. Schneider, A. Weidenkaff, and A. Wokaun, *Pulsed laser deposition and characterisation* of perovskite-type LaTiO_{3-x}N_x thin films, Acta Materialia 59, 7145 (2011).
- [11] J. Szlachetko and J. Sa, Rational design of oxynitride materials: From theory to experiment, CrystEngComm 15, 2583 (2013).
- [12] J. Szlachetko, M. Pichler, D. Pergolesi, J. Sa, and T. Lippert, Determination of conduction and valence band electronic structure of La₂Ti₂O₇ thin film, RSC Advances 4, 11420 (2014).
- [13] F.-T. Huang, X. Wang, S. M. Griffin, Y. Kumagai, O. Gindele, M.-W. Chu, Y. Horibe, N. A. Spaldin, and S.-W. Cheong, *Duality of Topological Defects in Hexagonal Manganites*, Physical Review Letters **113**, 267602 (2014).
- [14] B. B. Van Aken, T. T. M. Palstra, and N. A. Spaldin, *The origin of ferroelectricity in magnetoelectric YMnO*₃, Nature Materials **3**, 164 (2004).
- [15] S. Artyukhin, K. T. Delaney, N. A. Spaldin, and M. Mostovoy, *Landau theory of topological defects in multiferroic hexagonal manganites*, Nature Materials 13, 42 (2014).

- [16] T. Yu, T. A. Tyson, P. Gao, T. Wu, X. Hong, S. Ghose, and Y.-S. Chen, *Structural changes related to the magnetic transitions in hexagonal* InMnO₃, Physical Review B 90, 174106 (2014).
- [17] J. Goldstone, Field Theories with "Superconductor" Solutions, Nuovo Cimento 19, 154 (1961).
- [18] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Physical Review B 59, 1758 (1999).
- [19] A. Togo, F. Oba, and I. Tanaka, First-principles calculations of the ferroelastic transition between rutile-type and CaCl₂-type SiO₂ at high pressures, Physical Review B 78, 134106 (2008).
- [20] P. Lacorre, J. B. Torrance, J. Pannetier, A. I. Nazzal, P. W. Wang, and T. C. Huang, *Synthesis, crystal structure, and properties of metallic PrNiO*₃: *Comparison with metallic NdNiO*₃ and *semiconducting SmNiO*₃, Journal of Solid State Chemistry **91**, 225 (1991).
- [21] G. Demazeau, A. Marbeuf, M. Pouchard, and P. Hagenmuller, Series of oxidized trivalent nickel compounds derived from perovskite, Journal of Solid State Chemistry 3, 582 (1971).
- [22] G. Giovanetti, S. Kumar, D. Khmoskii, S. Picozzi, and J. van den Brink, *Multiferroicity in rare-earth nickelates RNiO*₃, Physical Review Letters **103**, 156401 (2009).
- [23] D. Fausti, R. I. Tobey, N. Dean, S. Kaiser, A. Dienst, M. C. Hoffmann, S. Pyon, T. Takayama, H. Takagi, and A. Cavalleri, *Light-Induced Superconductivity in a Stripe-Ordered Cuprate*, Science **331**, 189 (2011).
- [24] V. R. Morrison, R. P. Chatelain, K. L. Tiwari, A. Hendaoui, A. Bruhács, M. Chaker, and B. J. Siwick, A photoinduced metal-like phase of monoclinic VO₂ revealed by ultrafast electron diffraction, Science 346, 445 (2014).
- [25] R. Mankowsky, A. Subedi, M. Först, S. O. Mariager, M. Chollet, H. T. Lemke, J. S. Robinson, J. M. Glownia, M. P. Minitti, A. Frano, M. Fechner, N. A. Spaldin, T. Loew, B. Keimer, A. Georges, and A. Cavalleri, *Nonlinear lattice dynamics as a basis for enhanced superconductivity in* YBa₂Cu₃O_{6.5}, Nature **516**, 71 (2014).
- [26] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, *Nonequilibrium dynamical mean-field theory* and its applications, Reviews of Modern Physics 86, 779 (2014).
- [27] P. Werner, N. Tsuji, and M. Eckstein, Nonthermal symmetry-broken states in the strongly interacting Hubbard model, Physical Review B 86, 205101 (2012).
- [28] M. Porer, U. Leierseder, J.-M. Ménard, H. Dachraoui, L. Mouchliadis, I. E. Perakis, U. Heinzmann, J. Demsar, K. Rossnagel, and R. Huber, *Non-thermal sepa*ration of electronic and structural orders in a persisting charge density wave, Nature Materials 13, 857 (2014).
- [29] L. Stojchevska, I. Vaskivskyi, T. Mertelj, P. Kusar, D. Svetin, S. Brazovskii, and D. Mihailovic, Ultrafast Switching to a Stable Hidden Quantum State in an Electronic Crystal, Science 344, 177 (2014).
- [30] U. Staub, R. A. de Souza, P. Beaud, E. Möhr-Vorobeva, G. Ingold, A. Caviezel, V. Scagnoli, B. Delley, W. F. Schlotter, J. J. Turner, O. Krupin, W.-S. Lee, Y.-D. Chuang, L. Patthey, R. G. Moore, D. Lu, M. Yi, P. S. Kirchmann, M. Trigo, P. Denes, D. Doering, Z. Hussain, Z. X. Shen, D. Prabhakaran, A. T. Boothroyd, and S. L. Johnson, *Persistence of magnetic order in a highly excited* Cu²⁺ *state in CuO*, Physical Review B **89**, 220401 (2014).
- [31] M. Eckstein and P. Werner, Ultrafast Separation of Photodoped Carriers in Mott Antiferromagnets, Physical Review Letters 113, 076405 (2014).

- [32] P. Beaud, S. L. Johnson, A. Streun, R. Abela, D. Abramsohn, D. Grolimund, F. Krasniqi, T. Schmidt, V. Schlott, and G. Ingold, Spatiotemporal Stability of a Femtosecond Hard X-Ray Undulator Source Studied by Control of Coherent Optical Phonons, Physical Review Letters 99, 174801 (2007).
- [33] P. Beaud, S. L. Johnson, E. Vorobeva, U. Staub, R. A. D. Souza, C. J. Milne, Q. X. Jia, and G. Ingold, Ultrafast Structural Phase Transition Driven by Photoinduced Melting of Charge and Orbital Order, Physical Review Letters 103, 155702 (2009).
- [34] L. Rettig, S. O. Mariager, A. Ferrer, S. Grübel, J. A. Johnson, J. Rittmann, T. Wolf, S. L. Johnson, G. Ingold, P. Beaud, and U. Staub, *Ultrafast Structural Dynamics* of the Fe-Pnictide Parent Compound BaFe₂As₂, Physical Review Letters **114**, 067402 (2015).
- [35] P. Beaud, A. Caviezel, S. O. Mariager, L. Rettig,

G. Ingold, C. Dornes, S.-W. Huang, J. A. Johnson, M. Radovic, T. Huber, T. Kubacka, A. Ferrer, H. T. Lemke, M. Chollet, D. Zhu, J. M. Glownia, M. Sikorski, A. Robert, H. Wadati, M. Nakamura, M. Kawasaki, Y. Tokura, S. L. Johnson, and U. Staub, *A time-dependent order parameter for ultrafast photoinduced phase transitions*, Nature Materials **13**, 923 (2014).

[36] S. L. Johnson, R. A. de Souza, U. Staub, P. Beaud, E. Möhr-Vorobeva, G. Ingold, A. Caviezel, V. Scagnoli, W. F. Schlotter, J. J. Turner, O. Krupin, W.-S. Lee, Y.-D. Chuang, L. Patthey, R. G. Moore, D. Lu, M. Yi, P. S. Kirchmann, M. Trigo, P. Denes, D. Doering, Z. Hussain, Z.-X. Shen, D. Prabhakaran, and A. T. Boothroyd, *Femtosecond Dynamics of the Collinear-to-Spiral Antiferromagnetic Phase Transition in CuO*, Physical Review Letters **108**, 037203 (2012).



Year 2 saw, as planned, the start of the first eight experimental projects planned for the first call of PP7; a second call has also now been completed, and six experimental projects will start in May 2016.

As mentioned in the year 1 report, a new computational group leader, Prof. Volkan Cevher (EPFL) has been involved in the activities of HP5, with funding drawn from the allocation to Prof. Christoph Koch, a long-time collaborator.

The agility grants for years 1-2 will conclude, and a renewal call will be issued to the current holders (with the exception of Prof. Wanda Andreoni, that has since retired from EPFL).

Knowledge and technology transfer in MARVEL happens at different levels, from the dissemination of open-source materials' simulation codes, to the training in the use of those codes and of the newly developed materials informatics framework, to the sharing of all results from materials simulations in a network of open-access servers, to the verification and validation of calculations. The technology transfer activities this year aimed mainly to identify tech transfer opportunities and to make MARVEL's activities known in the industrial world, with the focus on materials development with specific properties.

Software

AiiDA

AiiDA (Automated Interactive Infrastructure and Database for Computational Science, www.aiida.net) is the infrastructure associated to the PP6 Informatics Platform [1] and is described in more details in the scientific report on PP6.

Last year, the first public version of the AiiDA code (v.0.4.0, with an open-source MIT license) was released on February 27, 2015. A minor release (v.0.4.1) followed on April 8 and a new major release (v.0.5.0) on December 17, 2015, with all recent updates and new functionalities.

In order to make sure that the different groups in MARVEL have the tools and knowledge to use the software infrastructure AiiDA, we have invested significant energy and efforts to organize activities of knowledge transfer.

In particular, we have organized three tutorials in the past year.

- On February 5, 2015 in Berlin, for the whole community, held after the CECAM/Psi-k research conference, gathering about 40 participants. This was a user tutorial meant to explain how to use AiiDA for daily research.
- On November 2 to 6, 2015 in Lausanne, for the whole community, and specifically targeted to MARVEL and MaX members, gathering over 40 participants (Fig. 1). This tutorial was composed of three parts: a user tutorial, similar to the tutorial in February; a developer tutorial, addressing people interested in using advanced

AiiDA features, such as workflows, and in developing new AiiDA plugins for external codes; and finally three coding days, where six members of the AiiDA team assisted more than ten researchers for the development of new plugins.

- A tutorial for users has been held within the Master in HPC (www.mhpc.it) in Trieste, Italy on December 15, 2015, to all the students of the Master.
- A tutorial for users will be held in Kyoto, Japan, on March 1st, 2016.
- A new tutorial (supported jointly by MARVEL, MaX and Psi-k), both for users and developers, is going to be held in Lausanne on June 22 to 24, 2016.

A lot of effort was also put in increasing the number of simulation codes supported, or the interaction with existing databases. The three coding days mentioned above were very effective in gathering people together and pushing the development of plugins. Also, regular skype calls are being held between members of the AiiDA team and researchers who are developing plugins, in order to assist them in the development process and ease the integration of the plugins into AiiDA. Skype calls or direct meetings are being held in particular for the development of the CP2K, Exciting, FLEUR, i-PI, VASP, Yambo plugins.

Furthermore, code development has taken place directly in the group of Nicola Marzari,



Figure 1: AiiDA tutorial on November 2 to 6, 2015 at EPFL, in Lausanne.

also by group members that are not directly involved in the AiiDA team (Marco Gibertini, Gianluca Prandini), or from visiting students (Daniel Marchand, May to December 2015) or visiting researchers (Victor Garcia Suarez, September to November 2015) that have led to the development of plugins for the Quantum-ESPRESSO NEB, DOS, PDOS, PROJWFC codes, for WANNIER90, and SIESTA, Yambo.

The AiiDA team is in many cases (CP2K, FLEUR, Yambo, i-PI, Quantum-ESPRESSO) in direct contact with the code developers to assist them in changing the codes so that they conform to good guidelines for extraction of code results by automatic codes. This is leading to improvements to the quantum codes that are of benefit of the whole Materials Science community and not only to those using AiiDA.

We are also holding regular interactions with external collaborators to make sure that AiiDA can be integrated or easily interfaced with external codes, libraries or databases.

- Materials Project (www.materialsproject.org): we had a few skype calls with the developers of the Materials Project, and we implemented a converter from crystal structures between AiiDA and pymatgen (part of the Materials Project). The conversion is implemented now both in AiiDA and in pymatgen.
- COD/TCOD (www.crystallography.net): Andrius Merkys visited the group of Nicola Marzari for one year (until October 2015) and developed automated converters to export calculations in the AiiDA database to the format and ontology required by TCOD, and upload them to the TCOD online database.
- Spglib (spglib.sourceforge.net): AiiDA has various routines and workflows using the SPGlib library for spacegroup recognition and cell reduction. Giovanni Pizzi will

visit the author of the library, Atsushi Togo, from February 24 to March 8, 2016 to develop advanced AiiDA workflows that exploit the capabilities of the SPGlib library to power research and discovery of novel materials.

AiiDA, the ADES model that lies behind its design, and the provenance model that it adopts [1] have also been presented in various workshops and conferences, including the Psi-k conference in San Sebastian (Spain), that is held every 5 years, and different CECAM workshops.

Also, members of the AiiDA team have supervised two projects involving students from the computer science department, in collaboration with Christoph Koch. The first involved four students attending his "Big Data" course and focused on investigating the performance and feasibility of new backends for AiiDA, including new graph-oriented databases such as TitanDB and Neo4j. The second one involved a computer science student, Jocelyn Boullier, who added the support in AiiDA for the SQLAlchemy library and the use of JSONB fields provided by recent versions of PostgreSQL, and he improved the database schema to optimize the query performance.

Finally, a non-academic article was published in the internal newspaper of EPFL "Flash", titled "AiiDA, un nouveau logiciel pour aider les chercheurs à gérer les simulations de matériaux". The article appeared in the issue number 10 (on December 16, 2015) and described to the whole EPFL community of researchers the efforts done at EPFL for the development of the code and the outcome of the AiiDA tutorial held in early November.

Materials Cloud

The Materials Cloud is a platform that will be used as a central hub for sharing scientific knowledge and expertise both internally amongst the MARVEL community and with the national and international community at large. The goal is not only to provide direct access to curated and raw results of computational studies but also to provide the means for non domain experts to carry out such investigations themselves. To this end Materials Cloud is a website that is underpinned by AiiDA which provides both the database to enable data sharing and a comprehensive workflow engine that enables complex calculations to be carried out in an automated fashion.

The Materials Cloud platform has been designed to scale to petabytes of data over its lifetime. The data must not only be easily accessible by means of database queries but is expected to be usable by workflows to do analysis and further computations.

Hardware

In February 2015 a 2×2.4 GHz, 64 GB RAM server was purchased with 24 TB of storage that is hosted at EPFL and serves the Materials Cloud website front-end. This is to be supported by a comprehensive storage backend hosted at and managed by CSCS. This flexible architecture allows storage needs to grow over time to support the activities of the MARVEL community.

Software

To enable remote access to data at CSCS, a solution has been designed and tested that uses iRODS for data management allowing secure access to both public and private parts of the shared AiiDA repository. AiiDA has a comprehensive workflow engine which is currently being extended in close collaboration with Robert Bosch to enable experts to easily encode their knowledge as a series of steps with heuristics to determine what action to take next which can then be used with minimal intervention by non domain experts.

Steps have been taken to enable AiiDA to seamlessly share data between local instances and the Materials Cloud as illustrated in Fig. 2. The goal is to make it as easy as possible to share data and workflows within a federated model whereby groups can work privately and retain full control until they are ready to go public.

Website

In April 2015 the public website went live hosting the Standard Solid State Pseudopotentials (SSSP) library, a collection of thoroughly tested



Figure 2: The AiiDA sharing infrastructure.

Quantum-ESPRESSO pseudopotentials. These represent an important step that enables community to perform reliable and efficient quantum mechanical simulations (Cf. scientific report on VP2).

In August 2015, Snehal Waychal joined the Marzari group as a systems specialist in charge of developing both the front and backend of Materials Cloud to achieve the goals previously mentioned. A milestone of March 2016 has been planned for proof of concept that brings together all the required elements including remote communication with backend storage at CSCS, interactive visualisation of results and the ability to query the database.

Related activities

A Horizon 2020 grant has been awarded for on *MAterials design at the eXascale (MAX)*, federating the activities of the supercomputing centers CINECA (Italy), CSCS, BSC (Spain), Jülich (German) and KTH (Sweden). It is envisaged that this project will reuse much of the technology behind the Materials Cloud.

Environ

Environ is a computational library developed with the Quantum-ESPRESSO package and aimed at introducing environment effects into atomistic first-principles quantum-mechanical simulations, in particular for applications in surface science and materials design.

The Goedecker and Marzari groups created a website to advertise and document the Environ interface, www.quantum-environment.org. Two release of the library took place in the last year, Environ-0.1 in March 2015 and Environ-0.2 in January 2016.

CP2K

A new official CP2K version (v.3.0) has been released on December 22, 2015, including Marvel related developments to provide a simulation

package to the community. CP2K is a quantum chemistry and solid state physics software package that can perform atomistic simulations of solid state, liquid, molecular, periodic, material, crystal, and biological systems.

Hardware

The computing infrastructure was installed at CSCS at the end of 2014, as agreed in the collaboration agreement between EPFL, ETHZ, and CSCS. The hardware includes 180 dual-socket compute nodes (2x Intel Haswell 12-Core, E5-2690 v3 @ 2.6 GHz), with 64 GBytes of memory (DDR4 @ 2133 MHz) and is clustered to the Cray XC40 supercomputer Piz Dora. MAR-VEL compute nodes have full access to the scratch filesystem of 2.7 PetaBytes shared with Piz Daint. CSCS is responsible for the hardware infrastructure as well provisioning core services, such as computing, storage, identity management and security.

Conferences and collaborations

MARVEL members organize or participate to the organization of a lot of conferences, in particular in the framework of their collaboration with the CECAM (Centre Européen de Calcul Atomique et Moléculaire), either at its headquarters in Lausanne, at EPFL or in its nodes across Europe. They were part of the organizations of the following ones and some of these workshops also received some financial support from MARVEL.

- CECAM workshop "Future Technologies in Automated Atomistic Simulations", CECAM-HQ-EPFL, Lausanne, June 8 to 10, 2015.
- CECAM workshop "Molecular Quantum Dynamics Methods: Benchmarks and State of the Art", CECAM-HQ-EPFL, Lausanne, June 15 to 19, 2015.
- CECAM Workshop "Next generation quantum based molecular dynamics: challenges and perspectives, CECAM-DE-MM1P, University of Bremen (Germany), July 13 to 17, 2015.
- CECAM school "Hybrid Quantum Mechanics / Molecular Mechanics (QM/MM) Approaches to Biochemistry and Beyond", CECAM-HQ-EPFL, Lausanne, February 9 to 13, 2015.

• CECAM school "4th CP2K Tutorial", CECAM-ETHZ, Zurich, Switzerland, August 31 to September 4, 2015.

Nicola Marzari was in the organization committee of the CECAM conference "Frontiers of first-principles simulations: materials design and discovery" in Berlin from February 1 to 5, 2015.

Nicola Marzari is also co-chair of the next Platform for Advances Scientific Computing PASC16 conference that will take place from June 8 to 10, 2016 at EPFL in Lausanne. Other MARVEL members are part of the minisymposia, contributed talks and posters committee for the "emerging domains (special topic: HPC & big data)" and "materials" sessions. The MARVEL management will contribute to the organization, at a local level.

Technology transfer

The initial aim of tech transfer (TT) activities is to make MARVEL's activities known in the industrial world, with the focus on materials development with specific properties. By increasing awareness and creating a list of interested companies, we will be able to generate synergies in the form of collaborations. We can explain the process as follows:

- Increase the awareness about first principle modeling potential within the industrial world.
- Establish direct contact (focused on defined market segments).
- Start discussion about a possible collaboration, i.e. the company has a specific request.
- non disclosure agreement (NDA) signature/further discussion/negotiation.
- Signature of industrial contract.

Since May 2015 (when Pascale Van Landuyt has been hired for TT and industrial liaison), and prior to start contacting companies, it was necessary to establish an inventory of tech transfer potential and of MARVEL group leaders' interests.

Assessment of potential TT opportunities within the MARVEL's labs

All the group leaders of EPFL labs have been visited for an "interview" about their current research work in the frame of MARVEL and also generally speaking. What is their interest



• Energy harvesting, conversion and storage:

MARVEL activities: solid-state electrolytes for Li ions batteries or fuel cells, metal-air batteries, perovskites for solar cells, photocatalytic water splitting, thermoelectrics.

Topics for possible collaboration: how is it possible to improve electrolytes conductivity? Rational design of dye-sensitized and perovskite solar cells with higher efficiencies. Prediction of the effect of doping.

• Information-and-communication technologies:

MARVEL activities: new electronics, optimization of semiconducting polymers, Si/SiO₂ interface.

Topics for possible collaboration: Understanding the charge transportation in conducting polymers. Study of oxidesemiconductor interfaces for microelectronic technology. Behavior of disordered materials which find applications in glass manufacturing.

• Organic crystals/pharmaceuticals compounds:

MARVEL activities: prediction of nucleation and growth of organic crystals from solutions, polymorphism and phase stability, peptides/proteins.

Topics for possible collaboration: Prediction of specific spatial conformation of peptides/proteins. Interaction between two membrane proteins. Understanding the affinity for specific ligands.

• Catalysis and chemical synthesis:

MARVEL activities: carbon capture, methane storage, separation process.

Topics for possible collaboration: Identification of zeolite structures with optimal properties for carbon dioxide or methane capture. Modeling of small chemical molecules. Quantification of chemical concepts like atomic charges, hyperconjugation, chemical bonds.

• Watch making industry:

MARVEL activities: optimization of optical, mechanical, thermal properties of metals or ceramics.

On this basis, we will be able to start to target specific companies in these sectors.

Increase the awareness

AiiDA is certainly a platform which is interesting to promote. Three companies participate to the second AiiDA tutorial in November 2015 at EPFL. One of them was interested by a license. MARVEL management (Lidia Favre-Quattropani and Pascale Van Landuyt) has been present with a poster and leaflets presenting the NCCR and its fields of activity at the PASC15 conference in Zurich (June 1 to 3, 2015).

A database of companies in contact with MAR-VEL is currently set up with the aim to send by emails specific news or to inform about activities.

Direct contacts and ongoing negotiations

Since May 2015, 13 companies have been in contact with MARVEL at different levels.

Sectors of activities:

- 3 watch-making/materials;
- 2 energy;
- 4 chemistry/pharma;
- 2 electronic;
- 2 others (small companies active in modeling).

Countries:

- Switzerland: 5;
- Europe: 5;
- US: 1;
- Japan: 2.

Level of the interactions:

- 1 company is on the point to sign for an AiiDA licence with 2 year of software support.
- 3 companies signed a non disclosure agreement (NDA), leading to 2 ongoing discussions and 1 closed discussion.
- 2 companies are in discussion without NDA.
- 5 companies have received a first email of awareness raising.



Figure 3: Synergy between MARVEL and other TT actors to approach companies..

- 1 start-up/entrepreneur has expressed an interest for MARVEL's activities.
- 1 company has given a conference for MARVEL community at EPFL.

Strategy to contact companies in synergy with other actors in TT

First contact with targeted companies, especially when located outside Switzerland, is a key issue. Thanks to the synergy with Alliance, EPFL's Corporate Relations, and the EPFL's Centers, we can benefit of entries in companies approaching EPFL with an interest for collaboration.

Watch-making sector is contacted thanks to Alliance network. Large companies in Europe, Japan and US have been approached thanks to EPFL's Corporate Relations team and we are now starting discussion with three EPFL thematic Centers (Energy Center, Transportation Center and Cloud Computing Center) to identify common target-companies (Fig. 3).

In December 2015, MARVEL direction and management did a first meeting with Nathalie Jongen and Carey Sargent from the Competence Centre for Materials Science and Technology (CCMX) management team based in the Institute of Materials at EPFL. The plans are to develop synergies, participate to some of their activities (annual meeting, technology aperitif), and maybe to plan common actions.

Education and support to MARVEL members for industrial contacts and entrepreneurship

Support has been given to MARVEL members to discuss intellectual property (IP) issues, or to explain the potential and eligibility criteria of CTI projects.

One group of two MARVEL students was selected to participate and present their project in connection with the materials informatics platform AiiDA to the Pre-seed Workshop "specifically targeted towards high-technology entrepreneurial ideas within Switzerland" organized and supported by the NCCR Chemical Biology in March 2015.

Information about CTI entrepreneurship courses has been forwarded to MARVEL PhD students and postdocs and this will continue in the future.

Following the agreement between the SNSF and the Swiss technology transfer association (swiTT) allowing NCCRs to be member of swiTT, Pascale Van Landuyt and Riccarda Caputo followed, as continuing education on tech transfer, the swiTT training day "Evaluating invention disclosures" on September 15, 2015, in Berne.

Action plan for 2016

Assessment of potential TT opportunities within the MARVEL's labs

We plan to visit the MARVEL group leaders in other institutions, to complete the roadmap and broaden the TT actions at Swiss and international levels. Current contacts of the MARVEL members will be also added in our database of industrial contacts.

Increase the awareness

Awareness about MARVEL's activities will be increased in the industrial world thanks to :

- Thematic workshops (with CCMX or EPFL's Centers), as, e.g., technology aperitif (16:00 to 18:00) upon invitation with presentation from companies and MAR-VEL's lab.
- Participation to events with a poster or a MARVEL booth. We will target events where companies are predominantly present. For example, the next CCMX annual meeting (May 11, 2016) will open its poster session to non-CCMX research scientists. We have planned to be present with several posters from MAR-VEL scientists.
- Emails to MARVEL industrial contacts to inform about, e.g., AiiDA tutorials, or scientific news published to a large audience (e.g. Vulcano effect, actu.epfl.ch/news/volcano-plots-usedfor-first-time-in-homogeneous-2/).

Direct contacts and ongoing negotiations

• In the next months, we will follow-up the contacts and negotiations initiated in 2015.


- New targeted contacts will be created in synergy with Alliance, EPFL's Corporate Relations and EPFL Centers. Similar strategy will be explored with the other academic partners of MARVEL, after group leaders' visits outside EPFL.
- The roadmap will be regularly updated with the input of the newly visited MAR-VEL members, as well as thanks to discussions with companies.

Education and support to MARVEL members for industrial contacts and entrepreneurship

Case by case, we will provide support to the MARVEL members regarding the IP issues, contracts and CTI projects. We will share information to PhD students and postdocs regarding CTI entrepreneurship and possible hiring by companies active in modeling.

MARVEL-related publications

 [1] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, *AiiDA: automated interactive infrastructure* and database for computational science, Computational Materials Science 111, 218 (2016).

The education and training activities of MARVEL focus on efforts within the community, dedicated to its current PhD and postdoctoral researchers (e.g. through the MARVEL Junior Retreat, or the support to attend notable international conferences), or to the community at large, through training schools and workshops of broad, fundamental interest dedicated to researchers worldwide. A sustained effort is also present to support activities in electronic-structure simulation and modeling in sub-saharan Africa. At the institutional level, Nicola Marzari and Michele Ceriotti have been working together to provide an integrated offering in the field, with new Master classes in the EPFL Institute of Materials and also with some high-profile effort involving former MIT colleague Prof. Craig Carter. We have been experimenting with distant-education technologies (e.g. 2-way broadcast to ETHZ of Prof. Gerd Ceder's Distinguished Lecture); situation would improve once the new classroom will be completed in the spring 2016 (see "Structural aspects", chapter 9).

PhD students and postdoctoral researchers

In the MARVEL community

Junior Retreat

The first Marvel Junior Retreat was organized in Männedorf near Zurich from July 7 to 10, 2015, with 48 participants, with intention of gathering young researchers working in the NCCR MARVEL, from physics, chemistry, or material science perspectives to know each other better in a friendly environment. The organisation was under the responsibility of three postdocs, Ivano Castelli in the group of Nicola Marzari, Sandip De in the group of Michele Ceriotti, and Shinaoka Hiroshi in the group of Matthias Troyer, supervised by Michele Ceriotti. Besides presentations from junior theoreticians, a few distinguished experimentalists were also invited to give us inspiration on the most relevant The program included one inchallenges. troduction talk about MARVEL, three invited speakers from experimental groups, one invited speaker from theoretical group and two talks about career development and funding possibilities inside Switzerland and SNSF (sites.google.com/site/marvelretreat2015/). From the participants, there were sixteen oral

presentations and eleven poster presentations. The retreat was a great opportunity to know other people working on the similar or complementary projects. An informal atmosphere stimulated many academic and non-academic discussions among the participants, which we believe have led to many new collaborations. Few pictures from the event are presented in Fig. 1. The organizers strongly feel that the first Marvel junior retreat was very successful in achieving its target and this tradition should continue in coming years.

A second edition is beginning to be organized with a new team of motivated postdocs and will take place from July 18 to 21, 2016 in Les Diablerets.

Participations to schools and tutorials

As already announced last year, MARVEL students were encouraged to participate to schools and workshop, in particular through some financial support from MARVEL.

• Seven postdocs and PhD students had the opportunity to participate in the international conference "Nothing is Perfect — The Quantum Mechanics of Defects", in Ascona, from April 26 to 29, 2015. The full registration fees were payed to five of them by MARVEL and half payed to the last two. The conference was co-organized by Alfredo Pasquarello, MARVEL deputy director, and was a great success by judging on the quality of the scientific presentations, on the number of participants (60, with 30 oral presentations and 28 posters),



Figure 1: Few selected moments from Marvel Junior Retreat 2015. Photography © 2015 www.sandipde.com

and the intensity of the informal discussions taking place during the breaks.

• MARVEL supported the travel expenses of four MARVEL students participating in the "Simons Many Electron Collaboration" Summer School, focusing on electronic structure and its many body extensions, at the Simons Center for Geometry and Physics, Stony Brook, June 8 to 12, 2015. On its side, the Simons Collaboration supported all their local expenses. This was a first action of a strong contact established between MARVEL and the "Simons Collaboration on the Many Electron Problem" (www.simonsfoundation.org/mathematics-and-physical-science/many-electroncollaboration/) directed by Andrew Millis.

As mentioned in chapter 5, p. 172, as activities at the frontier between knowledge transfer and education, MARVEL members are part of the organization of several CECAM (Centre Européen de Calcul Atomique et Moléculaire) workshops and schools and MARVEL also support financially some of them. In particular, Marcella Iannuzzi (in the group of Jürg Hutter) and Teodoro Laino (in the group of Alessandro Curioni) organized the 4th CP2K Tutorial, from August 31, 2015 to September 4, 2015 at CECAM-ETHZ, in Zurich. The tutorial was supported by CECAM and MARVEL. The program included 22 lectures by international experts (four members of MARVEL: Teodoro Laino, Jürg Hutter, Joost VandeVondele, Jan Wilhelm) and hands-on sessions. There were 41 participants in total, 10 students from MARVEL. In the future, we are thinking of recording these CP2K tutorials and offering them on the Education platform (see below, p. 179).

The first test of live broadcast from EPFL to ETHZ of the last distinguished lecture on January 25, 2016 was a success (chapter 8, p. 191). On another side, the planned distant education room in the new ME D building will be available soon (chapter 9, p. 197). This will be a chance to develop distant teaching through the MARVEL participating institutions and as well as to make local seminars distantly available.

On slightly different issue, MARVEL students were invited by the NCCR QSIT to participate to their QSIT Career Building Event about "De-





Figure 2: Khartoum Workshop for the Advancements in Material Sciences.

siring to understand, to think autonomously, and to create — case studies from my lives as scientist and administrator" with Olaf Kübler (professor of electrical engineering at ETHZ and former president of ETHZ) on Monday December 7, 2015 at ETHZ.

Education platform

In August 2015, Prof. Tony Paxton from King's College in London gave two lectures at EPFL on "Introduction to the tight binding approximation. Lecture I: theory. Lecture II: applications". Prof. Paxton was one of the first developers of the tight binding model — probably the simplest quantum mechanical scheme that can account for bond breaking. His visit received the support of MARVEL. The two lectures were recorded and are available through the MARVEL website on the same platform as the recorded distinguished lectures (slideshot.epfl.ch/events/14, chapter 8, p. 191).

The idea is to build an "Education platform" collecting a library of video lectures and associated material on advances topics in electronic structure simulations, to act as permanent repository. It should benefit from the Materials Cloud platform (chapter 5, p. 170) for hosting and dissemination.

Outside MARVEL

Two core activities will take place in collaboration with the International Centre for Theoretical Physics (ICTP) in Trieste, that has a 50-year track record of training activities, and offers affordable, excellent facilities: a 2-week workshop on "High-Performance and High-Throughput Materials Simulations" will be organized in January 2017, slated for 100+ students, in collaboration with the EU H2020 CoE MaX, and a tutorial on high-throughput computing, also for 100+ students, will be organized in July 2016 as part of the 2-week "College on Multiscale Computational Modeling of Materials for Energy Applications". Also, we continue our engagement with science in developing countries, particularly sub-Saharan Africa, sending lecturers and tutors both to the "Khartoum Workshop for the Advancements in Material Sciences", that took place in Sudan in February 2015 (KWAMS-15), with a special focus on computational material science and the calculated physical and chemical properties of a wide range of materials (Fig. 2), and to the 4th African School on "Electronic Structure Methods and Applications" (ASESMA-2016) that will take place in Accra (Ghana) in June 2016. Remarkably, KWAMS-15 was organized by Dr. Maali M. Hussien, that was a student and the only female in attendance at the 2008 ICTP/NSF East African School in Computational and Materials Sciences, a 2-week workshop held by Nicola Marzari in Addis Ababa in September 2008, and a forerunner of ASESMA.

At the level of Master students

Courses

The Institute of Materials at EPFL rethought its teaching offerings at the Master level starting with the 2015-2016 academic years, and two MARVEL professors will teach new classes in a synergetic approach particularly useful to computational materials science: Nicola Marzari will teach in the fall the core class "Fundamentals of solid-state materials" on the fundamentals of quantum mechanics as applied to atoms, molecules, and solids and to explain the electronic, optical, and magnetic properties of solids, while Michele Ceriotti will teach in the spring the core class "Statistical mechanics", dedicated to expound the microscopic picture and statistical interpretation of thermodynamics. The core classes will be complemented by the spring modeling class on "Atomistic and Quantum Simulations of Materials" by Nicola Marzari, and by "Modeling problem solving, computing and visualisation II" by Craig Carter, professor at MIT, and Michele Ceriotti. Part of this class will be of-

fered via videoconference and will be simultaneously taught with an MIT subject.

For the younger generation

Simulation camp

It is our ambition, if still alive by the end of this report, to start planning an activity for lateyear high-school students dedicated to teaching fundamentals of simulation techniques to entice new generations to the field of computational science. In year 2, MARVEL has continued the implementation of its equal opportunities strategy through the organisation of specific measures in the field of MARVEL and the support of shared actions organized in the framework of this strategy together with other NCCRs, the EPFL Equal Opportunities Office and, for its activities addressed to girls, the newly established EPFL Science Outreach Department.

Strategy and recommendations from the site visit

The equal opportunities strategy adopted by MARVEL is focused on:

- contributing to a better awareness concerning gender issues,
- promoting MARVEL related fields among girls to encourage them to choose scientific and engineering fields,
- recruiting women scientists and offering them an adequate environment to develop their career and to access to higher positions,
- tending to a gender balance at all levels of MARVEL.

Based on the site visit in 2015, the expert panel recommended to focus on concrete steps to promote the equal opportunity strategy within MARVEL. Several measures that integrate the panel's recommendations have been undertaken in year 2 or are currently discussed.

Numbers

Table 7.1 gives an overview of the number and share of women involved in MARVEL. Compared to year 1, with 5 female PhD students and 5 female postdocs, new women were hired at these two levels, with respectively similar shares than in year 1.

At the outset of year 2, eight experimental projects which have been attributed on a competitive basis have started within project PP7 (out of 20 proposals). Two of these projects are led by women. At the beginning of year 3, six new experimental projects will start (out of 13 proposals), among which two led by women, bringing to 6 out of 33 (i.e. 18%) the MARVEL

female group leaders in year 2 and to 8 out of 39 (i.e. 20%) in year 3. In these two rounds, all projects presented by women were financed.

Raising gender awareness

MARVEL website

The MARVEL website (www.nccr-marvel.ch) is online since June 2015. The homepage gives a direct access to the *Equal Opportunities* page and subpages, to show that this theme is an important issue for MARVEL. These pages share the current situation and the goals for equal opportunities in the framework of MARVEL. They also gather the actions and measures for female scientists available in the different institutions, as well as those organized directly by MARVEL, and present the actions concerning the young girls offered at EPFL. The *Equal Opportunities* pages have to be updated regularly, as for the whole MARVEL website.

	Women	Men
Scientific committee	2 (20%)	8 (80%)
Group leaders	6 (18%)	27 (82%)
Postdocs	9 (13%)	59 (87%)
PhD	8 (22%)	29 (78%)
Scientific advisory boards	1 (10%)	9 (90%)

Table 7.1: Number and share of women involvedin MARVEL in year 2



Figure 1: Lunch events organized by the EPFL Equality Office. Left: Accenture lunch event. Center: Google lunch. Right: lunch event with Prof. Clémence Corminboeuf and Prof. Stéphanie Lacour.

Equal opportunities meeting

At the *MARVEL Review and Retreat* meeting of September 3–4, 2015, a one-hour equal opportunities session was organized, with a focus on gender and young researchers. Eight participants took part (five women and three men). The aim of the session was to provide a time to reflect and exchange on career goals and coming steps. For MARVEL, it was an chance to receive an input for future activities for equal opportunities and career development.

Working conditions to combine a career with family and dual career issues were main topics brought up by participants. Participants however also stressed that the family issue was just one issue in gender equality. Other topics they wished to see addressed were sensitivity to gender and diversity issues, awareness of men on gender issues, equal opportunity measures to support women's careers, and measures to provide role models for kids and combat stereotypes.

NCCR gender platform

Clémence Corminboeuf, the MARVEL group leader in charge of equal opportunity at the scientific level, and Lidia Favre-Quattropani, from the MARVEL management team, attended the NCCR gender platform organized by the SNSF in Bern on November 4, 2015. Helene Füger, from the EPFL Equal Opportunities Office, and Farnaz Moser, from the EPFL Science Outreach Department, with whom they work closely, were also present. This meeting was a useful chance to meet other NCCRs' equal opportunities actors to share good practices and learn about new well-working actions.

Support work/life balance

The topic of work/life balance was discussed at the equal opportunities session organized at

the *MARVEL Review and Retreat* meeting. As a next step, it is planned to draft a document of good practices and recommendations to support work/life balance, to be submitted to the MARVEL executive committee.

Advance women scientists

Lunch events

The EPFL Equal Opportunities Office (EEO) organizes annually a series of lunch events which are advertised by the NCCRs linked to EPFL, and open to all women PhD students and postdocs from these NCCRs. The aim of these lunches is to provide women PhD students and postdocs insights in different careers and the chance to interact with role models (Fig. 1). In 2015, the lunch events reported in Table 7.2 were organized.

Training workshops

Workshops with Nancy Houfek

Through the initiative of Nicola Spaldin, MAR-VEL Professor at ETHZ, MARVEL organized in October 2015 three "professional skills development" workshops for women led by Nancy Houfek. Nancy Houfek is a stage director, actor and theatre educator. From 1997 to 2014, she was Head of Voice and Speech for the Tony Award winning American Repertory Theater at Harvard, teaching vocal production, coaching the professional acting company, and administering the master (MFA) program in voice training pedagogy. She has also created a series of workshops designed to help women with negotiation and leadership challenges, which has been presented to more than one hundred institutions and professional groups throughout the USA and Canada (www.nancyhoufek.com).

The organisation of the workshops has been carried out by Nicola Spaldin's group together with the MARVEL management and the EPFL



Date	Event	Partner	Total par- ticipation
22.05.2015	Accenture lunch event Participants got to know more about careers in an international management consulting and technology services company. Three women in different positions talked about their job, their career and answered partici- pants' questions.	Accenture	76
23.09.2015	Lunch meeting with Professor Jill M. Far- rant, laureate of the 10th Erna Hamburger Prize Participants got the opportunity to interact with a leading researcher in her field, to learn more about her career, research, leadership and attitude towards being a role-model.	WISH Foundation	55
09.10.2015	Google lunch Two women working at Google in manage- ment and programming positions presented their work. A doctoral student form EPFL, and Anita Borg Award laureate, presented the award and her involvement in promot- ing equal opportunities in IT careers and out- reach activities for girls.	Google	85
12.11.2015	Lunch event with Prof. Clémence Cormin- boeuf and Prof. Stéphanie Lacour PhD students and postdocs participating had the opportunity to learn more about the ca- reers of the professors, their motivations, choices, obstacles they faced and how they addressed them, as well as how they organize their career and family life.	NCCRs MARVEL and Synapsy	25
27.11.2015	CISCO industry visit Four female and two male staff members pre- sented their careers and positions in this IT network software company. Participants had the opportunity to interact with these profes- sionals at their company's offices at the EPFL innovation park.	Cisco	24

Table 7.2: Lunch events organized in order to support female PhD students and postdocs to continue building their career.

EEO. Among these three one day workshops, one was specifically for professors and senior scientists with a long term position, on October 12, 2015 at ETHZ, and two were for PhD students and postdocs, on October 14 at ETHZ and on October 16 at EPFL. In order to allow a maximum number of women scientists to ben-

efit from these workshops, MARVEL invited other NCCRs to support the participation of female researchers within their network. The NCCRs QSIT — Quantum Science and Technology, Digital Fabrication, Synapsy, Kidney — control of homeostasis, MUST, RNA & Disease, SwissMAP, Bio-Inspired Materials, Plan-



Figure 2: Enthusiastic participants in the Nancy Houfek workshops for women in science at EPFL (left) and at ETHZ (center and right). Nancy listing negotiation tactics (right).

etS, and the SCCER-FURIES supported these workshops by covering the registration fees for their members (Table 7.3). MARVEL supported the participation of 8 women, 1 professor, 1 senior scientist, 4 postdocs, 1 PhD student and the scientific manager.

These workshops entitled "strong women/strategic performance: achieving success in meetings & negotiations" combined theater, negotiation and leadership techniques and the following topics were covered:

- communication & physical presence,
- theories & preparation for negotiation,
- vocal presence,
- negotiation case study.

The workshop evaluations after participation were all extremely positive and enthusiastic. All the participants (Fig. 2) rated the workshop either *extremely valuable* (76%) or *valuable* (24%). They reported about numerous points they learnt that would enable them to operate effectively in difficult situations in their day-to-day professional life. These cover, e.g.,

- the awareness of body language, physical presence, and management of the voice;
- the need to prepare a negotiation (do researches, practice, rehearse, ...), to take time, and sometimes to ask for time and to postpone;
- the multitude of negotiation tactics (Fig. 2 right) that can be used (and that may also be used by others towards oneself).

The choice of Nancy Houfek as a trainer was very positive, not only for her professional skills but also for her great awareness of the academic environment, particularly the scientific academic environment.

The length of the workshops was adequate for most of the participants, could be longer for some students, but should not be longer for some professors (five of them could only participate to the beginning of the day, due to teaching schedule). 30 participants on October 14 were maybe a little numerous and could be

	Professors and senior scientists	PhD students and postdocs	Total
NCCR Bio-Inspired Materials	2	3	5
NCCR Digital Fabrication	1	4	5
SCCER FURIES	0	4	4
NCCR Kidney	0	4	4
NCCR MARVEL	2	6	8
NCCR MUST	0	1	1
NCCR PlanetS	0	3	3
NCCR QSIT	0	4	4
NCCR RNA & Disease	1	11	12
NCCR SwissMAP	1	0	1
NCCR Synapsy	0	2	2
Individual participants	16	8	24
Total	23	50	73

Table 7.3:Participants in the workshops forwomen in science and their support.

reduce to 20-25.

Participants were also interested by other kind of workshops, mainly on speech/voice training and presentation skills, leadership, project and time management, conflict management, career planning and work-life balance. As the financial balance for the organisation of the workshop is slightly positive, MARVEL will reinvest in one or two years the remaining money (about CHF 3'500) in a new workshop, either on the same topic, or in one of those suggested above. This should also allow to reduce the registration fees, even if they did not seem to be prohibitive (one third of the participants were individual, with their fees directly covered by their lab (Table 7.3)).

Support for international workshop participation

MARVEL supported the participation of Nongnuch Artrith, visiting fellow from MIT in the group of Nicola Marzari, in the special workshop for *Women in high performance computing (HPC)* at ISC 2015 in Frankfurt am Main on July 16, 2015, targeted to early career women working in HPC and anyone interested in gender equality in the field of HPC. She was accepted for a poster together with a 2 min. talk. At the end of the workshop, there was a panel discussion about women in science, technology, engineering, and mathematics (STEM) and HPC. It was a good experience for Nongnuch Artrith, with interesting discussions.

REGARD workshop program

The REGARD workshop program (www.unifr.ch/regard) is targeted to women PhD students, postdocs and senior researchers. Around 20 workshops are organized each year, on issues related to career planning, communication, leadership, etc. Organized by the equal opportunities offices from the universities in the French speaking part of Switzerland, the program is advertised on the NCCRs webpages. Participation is free for women researchers of partner institutions.

Mentoring and coaching

MARVEL encourages its female PhD students and postdocs to take part in a number of academic mentoring programs targeted to women researchers.

• *StartingDoc* group mentoring (www.unil.ch/mentoring), targeted to

PhD students at the outset of their doctoral thesis, is open for women PhD students and postdocs employed by UniGE, UniFR, and EPFL.

- The *Réseau romand de mentoring pour femmes* (www.unifr.ch/f-mentoring), targeted to women at the end of their PhD and postdocs aiming at a career in academia, is open for researchers at EPFL, UniGE, UniFR, and USI.
- "Fix the Leaky Pipeline" (www.fix-theleaky-pipeline.ch) coaching groups for PhD students and postdocs is open to women researchers from EPFL, ETHZ, PSI, and Empa
- *Mentoring Deutschschweiz* (www.academic-mentoring.ch), targeted to women at the end of their PhD and postdoc, is open to researchers from UniBas, UZH, UniFR, and USI.

These mentoring programs are advertised on the MARVEL website. Female PhD students and postdocs are also encouraged via email to participate in one of these mentoring and coaching programmes. MARVEL professors (men and women) are encouraged to volunteer as mentors for one of these programs.

Advance women scientists — Next steps

While the activities proposed to women from MARVEL organized by the EPFL EEO will continue, MARVEL intends to implement a set of specific measures aimed at enhancing the number of women and their career prospects in MARVEL.

- A list of academic prizes in MARVEL related fields and/or for women researchers is currently set up, with the idea to support women in MARVEL in applying for these prizes.
- MARVEL is presently planning to award a number of excellence fellowships for master students starting their master thesis. The aim of these fellowships is to attract excellent female master students or to reward excellent female master students studying in MARVEL institutions, and to encourage them possibly to continue with a PhD. This measure is developed together with the NCCR QSIT — Quantum Science and Technology, with an active collaboration with Anna Fontcuberta i Morral, QSIT professor at EPFL.



Figure 3: Polythème workshop Diamant, alu, caoutchouc, ils sont fous ces matériaux ! in September 2015.

- MARVEL is committed to increase the number of women invited for its distinguished lectures. Young female researchers shall then be given the opportunity to exchange with these role models, for example in the context of a lunch event organized by the EPFL EEO, which will be open for other female PhD students and postdocs.
- After the success of the workshops for women in science with Nancy Houfek, MARVEL is considering to renew the experience in two years.

Actions concerning young girls

The Equal Opportunities Office and the Science Outreach Department of EPFL work together with MARVEL to enhance interest of young girls for STEM fields. With the support of the NCCRs hosted at EPFL, specific activities related to the field of each NCCR are developed. Through the organization of these activities, MARVEL acts for behavior changes, decreasing stereotypes and obstacles related to gender, supports the development of knowledge and interest of young girls for MARVEL related fields, and increases their confidence in their capacities for STEM fields.

Polythème workshop on materials (7 to 10 years old girls)

Within *Polythèque*, a place for children to discover and share ideas on science through adapted books, videos, CDs and games, children can register to participate in scientific workshops *Polythèmes*.

A special workshop related to MARVEL fields, entitled *Diamant, alu, caoutchouc, ils sont fous ces matériaux !*, has been developed in 2014. A first edition of the workshop took place in January 2015. This workshop, with the goal to let young girls discover the world of material sciences having fun, was very successful. A second edition of this workshop took place on September 2, 9 and 16, 2015 with the same success (Fig. 3) and a third edition is planned for March 2, 9 and 16, 2016 and is already fully booked.

The first afternoon, different types of materials and their properties are presented to the young girls; they learn about these materials by performing simple hands-on experiments. They discover first the material. What is it made of? Are there any differences between this or that material? The girls put their new knowledge into practice in a game where one has to place objects belonging to different families of materials (ceramics, polymers, metals, composites) on a giant diagram. During the second afternoon, the girls work in groups to detect the properties of the different families of materials. They test the thermal conductivity, electrical conductivity, magnetism, and the density of different materials. They then build a game called "magnetic maze". Through previous experiments, participants are able to deduce which way should take a magnetized ball in their maze. The third afternoon, they experiment non-newtonian fluids. Preparation of the slime[®] give the opportunity to discuss once again the concepts of atoms, molecules and chemical bonds.

A woman scientist conducts the workshops and acts as a role model for these very young girls.

Scientific summer camp Matériaux super géniaux (11 to 13 years old girls)

As announced last year, a new scientific week for 11 to 13 years old girls, entitled *Matériaux super géniaux*, has been prepared around different fields related to MARVEL. Clémence Corminboeuf participated in the elaboration of its content, as well as the one of the *Polythème* workshop *Diamant*, *alu*, *caoutchouc*, *ils sont fous ces matériaux* !, together with Marion Albertini, a scientist of the EPFL Science Outreach Department. The first edition took place from August 17 to 21, 2015. Besides hands-on experiments and theoretical explanations, young



Figure 4: Scientific summer camp "Matériaux super géniaux" in August 2015.

girls could discover the research running in different research labs of MARVEL. Indeed, Clémence Corminboeuf hosted a group of participants and guided them through an interactive tour of her laboratory. Nicola Mounet, in the group of Nicola Marzari, presented to another group the theoretical studies taking place in this other MARVEL lab. During the week, the young girls also discovered different families of materials, grew crystals, imagined and performed experiments to measure different properties such as thermal and electrical conductivity, and density of materials, discovered, experimented, and discussed different states of matter, performed tests on different polymers to find what they were made of, produced a medal with tin, could walk on a mixture of cornstarch and water, or make a piezoelectric loudspeaker (Fig. 4). They prepared a presentation given in front of their friends and families on the subjects they learnt about in the visited labs. They also received a certificate at this closing ceremony in presence of their parents. This summer camp was a big success and will be renewed next summer, from August 15 to 19, 2016, as an already fully booked edition.

Chemistry summer camp (11 to 13 years old girls and boys with a 50% quota for girls)

A chemistry summer camp was organized from August 17 to 21, 2015 for 11 to 13 years old girls and boys. A fixed quota of 50% girls was imposed at registration. The purpose of this camp was to increase the interest of children, and especially young girls, for science and engineering. The usefulness of these branches in daily life was also demonstrated. Children are captivated by chemistry experiments. As many youngsters, and especially young girls, want to help people and make a difference in the world, it is important to highlight the role of chemistry in society.

The week was spent in a real chemistry laboratory for students. The first safety rules were introduced through the wearing of lab coats, gloves and protective glasses. Wearing the same clothes as "real chemists" is also part of the dream of becoming a chemist!

The participants, divided into several groups, visited labs carrying out research in chemistryrelated domains. Researchers acted as role models, explaining their research to the youngsters during one hour using appropriate language. During the camp, participants prepared a photosensitive solution to make cyanotypes. They made slime[®]; materials science was thus introduced thanks to this amusing and unusual substance. Notions of ions energy and color were introduced. They extracted DNA from buccal cells or onion cells. To discover the principle of pH, they prepared their cabbage solution and observed color variations depending on the tested solutions. They experimented the emulsion technique, prepared an orangescented moisturizing hand cream, made a hair gel thanks to a polymerizing agent, and performed electrochemical experiments (Fig. 5).

At the end of the camp, parents were invited to a ceremony during which the children presented the work done during the week and the subjects learned about during their lab visit, and received a participation certificate. This was also an opportunity to talk to parents about science and chemistry and ask them to encourage their children, and especially their



Figure 5: Chemistry summer camp for girls and boys in August 2015.

daughter, to go further in these fields.

This camp will be repeated next summer, from August 15 to 19, 2016 and is already fully booked.

Mathematics workshops Maths en jeu (11 to 13 years old girls)

Studies show that in many industrial countries girls don't have enough confidence in their capacities concerning mathematics. In order to overcome this problem, targeted actions are necessary.

The main purpose of the mathematic workshops *Maths en jeu* is to discover mathematics from a common sense point of view. Mathematics is not presented as a hindrance but as a useful tool for a future scientific career. The important goal of these workshops is to give to children, and in particular to girls, confidence in their mathematical skills and raise their interest for this field which constitutes the basis for the scientific and engineering branches. Beginning with one course, following the enthusiasm of participants to continue such workshops, advanced mathematical workshops have also been developed. These workshops are organized within the framework of the collaboration of the NCCRs Synapsy and MARVEL and the EPFL Science Outreach Department. At least three one-hour per week workshops for 7 to 10 years old and 10 to 13 years old children are organized from March to June and from September to December. Advanced workshops are also organized as a function of the demand. At least one of the three workshops is dedicated just to 11 to 13 years old girls and the other are mixte with a 50% quota for girls. Last workshops were successfully organized from March to June 2015 and in September and October 2015. They will be repeated similarly in 2016.

Activities organized for young girls to increase their interest in STEM fields are very successful and there is a demand for organizing more activities. MARVEL is thinking of developing a course or a camp for older girls, from 15-17 years. This could be part of the education and training strategy, in, e.g., a mixed session with a 50% quota for girls dedicated to material simulation and computer science (see also education and training chapter, page 180).

Cooperation with the Equal Opportunities Office and Science Outreach Department of EPFL

To enhance the sustainability and momentum of its equal opportunities measures, MARVEL has a two-pronged strategy. On one hand, it works directly with EPFL Equal Opportunities Office (EEO), and actively encourages its members to participate in the activities organized by the EEO. MARVEL also coordinates with the EEO its support measures for women scientists, and benefits from EEO's expertise and support. For all these activities EEO offers its services pro-bono, unless tasked with a specific mandate.

On the other hand, MARVEL has a long-term structural plan of offering targeted activities for young girls and women, spanning ages going from primary school to high school. These activities take the form of dedicated workshops offered throughout the year, either in the summer or during the school breaks, and arranged around age-appropriate groups (7-10; 11-13; 15-17). These activities are tasked by MARVEL to the EPFL Science Outreach Department (a new entity at EPFL — the promotion of science among the youngsters, especially the young girls, previously part of the EEO, is now part of this new entity), and the costs (salaries, material, secretariat and supervision) are billed to MARVEL on a yearly basis. The breakdown of these will be presented in the final MARVEL accounting report due July 2016.

Last, we are also starting to cooperated with the equal opportunities offices of other MAR-VEL partners, and are proposing joint activities with NCCRs based at other partner institutions. In this second year, MARVEL developed its communication strategy in many ways. The website is online and actively updated. Meetings of the members were organized to strengthen the internal communication, inside each projects and between different projects, and to actively build the MAR-VEL community and network. MARVEL took the opportunity of various external events to open its communication to the outside, to scientists in the domain of material simulation at the national and international level, industrial partners, media and the general public.

Internal Communication

Website

The MARVEL website www.nccr-marvel.ch is online since June 1, 2015. It serves the two main axes of MARVEL communication, internal and external communications. For the internal communication, it is a common place where resides the relevant information on projects, people, events and news. It is regularly updated. A newsletter tool is planned to be implemented in order to foster the transversal communication between the different research groups in MARVEL, as well as to open communication to outside, including industry.

Meetings

To pursue the development of internal communication, to boost the dialogue within each of the seven projects and between the projects, meetings are regularly planned.

Review and Retreat

The second *Review and Retreat* took place at EPFL, during one day and a half, on September 3 and 4, 2015, with a similar format as the first one. This edition gathered almost 80 MARVEL participants, group leaders, senior researchers, postdocs and PhD students, representing 30 of the 33 research groups (Fig. 1). The director of MARVEL took this opportunity to invite the Scientific Advisory Board to meet MAR-VEL members for the first time, and eight of the ten members were present at the *Review and Retreat*. The programme alternated with plenary sessions in the mornings, with overviews and highlights on the seven projects, and par-

allel sessions in the first afternoon, for a critical review of the activities in each project and for an analysis of possibilities and opportunities for group collaborations. A poster session was organized on the Thursday evening with food and wine to give the chance to the students of presenting their work to the whole MAR-VEL community and to everybody of further discussions around 26 posters of high quality (Fig. 2). All the female members of the NCCR were invited, together with the interested male members, to a one-hour session with a focus on equal opportunities and career perspectives (see chapter 7, p. 182 for more details). The Scientific Advisory Board met and prepared a report presented to the director with comments and recommendations. A meeting of the Scientific Committee took also place during the Review and Retreat.

Experimental cooperation

Together with the launching of the second call for experimental projects (Project PP7), a second "MARVEL - Experimental verification" workshop was organized at Empa on October 23, 2015. This year the focus was mainly given to the question "What is the meaning of 'experimental validation'? The simulation-driven discovery of novel materials requires the guide of the experimental groups playing an important role not only for an 'a posteriori' verification of the theoretical prediction, but also for suggesting important steps to be considered in the theoretical design: synthesis, contacts, experimental conditions, roadmap to application." More details can be found in the PP7 project description.



Figure 1: Second MARVEL Review and Retreat, on September 3 and 4, 2015, at EPFL: group picture.



Figure 2: Second MARVEL Review and Retreat, on September 3 and 4, 2015, at EPFL: oral and poster presentations on ongoing research and a lot of discussions.

A meeting for all the participants in Platform Project PP7 — group leaders, postdocs, and PhD students — is planned on March 3, 2016. At this meeting the individual projects will present their status and future plans, with a presentation or a poster from the experimental partner and from the theory partner.

Projects meetings

Individual projects meetings are also planned for the other projects. The format of these meetings aims at fostering collaboration and synergy between the different groups contributing to the same project. They will be structured in a way that all participants, PhD students and postdocs together with group leaders, will be able to present and discuss their own results and get a constructive critical feedback during the open discussion session. The periodicity of the project meetings is set for the moment to once a year, then it is left to the spontaneous initiative of each group leader to strengthen the already set collaborations or initiate any new ones. Depending on the feedback and the general consensus, the project meetings might become more frequent or turn into topical project meetings to foster the transversal synergy between the seven projects.

VP2-Day On January 29, 2016, the VP2-Day, a meeting for all the participants in Vertical Project VP2, took place at EPFL. This meeting also gathered some researchers from Horizontal Project HP3, actively involved in VP2 through the theories and algorithms developed and implemented in HP3.

The meeting gave the opportunity to the 33 participants to share their ongoing research, discuss the open questions, reinforce ongoing cooperations and foresee new ones (Fig. 3). Pascale Van Landuyt, the person in charge of tech transfer, was also attending as observer.

HP5 meeting The members of Horizontal Project HP5 will meet at IBM in Zurich on February 29, 2016.



Figure 3: VP2 Day at EPFL on January 29, 2016.

PP6 meetings The Informatics Platform project PP6 is regularly discussing the advancement of the project through bimonthly internal AiiDA meetings and monthly Materials Cloud Skype meetings between EPFL and CSCS.

Other project meetings A project meeting gathering members from Vertical Project 1 and members of Horizontal Project 3 involved in VP1 should take place at ETHZ during the spring.

Horizontal Project HP4 should organize a project meeting soon.

NCCR MARVEL distinguished lectures

MARVEL regularly invites worldwide recognized experts in computational materials science to give a "NCCR MARVEL Distinguished Lecture". This year four lectures were organized at EPFL, gathering every time a public of about 50 to 70 people.

- On Monday August 31, 2015, Prof. Sidney Yip from MIT gave a lecture "On the mesoscale science frontier in materials theory and simulation".
- On Tuesday October 13, 2015, Prof. Leonid Levitov, also from MIT, presented the "Atomic collapse in graphene".
- On November 11, 2015, Dr. Pierre Villars, from MPDS in Vitznau, talked about "The MARVEL initiative and the integration of the fifth paradigm of science".
- On Monday January 25, 2016, Prof. Gerbrand Ceder from the University of California at Berkeley presented "The materials genome and the transformation of materials science and engineering" (Fig. 4).

These four lectures were all recorded and made available through the MARVEL website on slideshot.epfl.ch/events/14. The idea is to build gradually a repository of lectures and associated material on advances topics in electronic structure simulations, as education platform (chapter 6, p. 179).

For the moment, these lectures are taking place at EPFL, but could be organized in other participating institutions. As a first successful trial, the last lecture, by Prof. Gerbrand Ceder, was broadcasted live at ETHZ, with a distant audience of about ten people.



Figure 4: Sixth "NCCR MARVEL Distinguished Lecture" by Prof. Gerbrand Ceder from the University of California at Berkeley on January 25, 2016. This lecture was recorded to be available on the website and broadcasted live at ETHZ.

External communication

Website

Mainly used for internal communication until now, the website has to be developed for external communication. The different audiences, scientists, industries, media and the general public, have to be addressed according to their interests. Social medias, such as LinkedIn or Twitter have to be studied as possible entries to reach certain target audiences.

MARVEL leaflet

As a material support can be a useful alternative to an immaterial support such a website, we prepared a two-page leaflet to present MARVEL activities and also to advertise its website. It was realized in a similar graphical line as the website (Fig. 5) and targeted an informed public. While the first version is in English, we also plan to translate it in French and German.

It was first used as a complement to the poster presenting MARVEL activities on our booth at the PASC'15 conference in June 2015 (see chapter 5, p. 173). It was also made available at SC15, the international conference for high performance computing, networking, storage and analysis, in Austin, TX, in November 2015. It is always available as a support to present MAR-VEL to interested people.

RTS Courts du jour

The *Radio Télévision Suisse* (RTS) and the EPFL Institute of Materials (IMX) worked together to produce a series of short films presenting the research projects of several professors at IMX with a link with environmental questions. The



Figure 5: *Two-page leaflet presenting MARVEL ac- tivities.*

series is entitled *Imaginez un monde*, Season 2 (Season 1 is about the research carried out in the Institute of Electrical Engineering at EPFL, broadcasted in January and February 2015).

MARVEL took the opportunity to promote its activities through such a film. Together with presentations from Nicola Marzari at EPFL and Thomas Schulthess at CSCS, the film shows the way MARVEL is working to design and discover new materials thanks to the power of supercomputers.

The film is entitled *Un bol d'oxygène: nouveaux* matériaux on ARTE and *Imaginez un monde. Les matériaux de demain* on RTS and was shot at EPFL (Fig. 6) and CSCS in October 2015. It was broadcasted on ARTE's website mid-November just before the 2015 Conference on Climate Change in Paris (future.arte.tv/fr/bold-oxygene) and on the RTS 1 and 2 TV channels on February 5, 2016 and on the RTS



Figure 6: Shooting of Nicola Marzari's presentation for the Court du jour in the hall of the new ME D building where the MARVEL management moved in November 2015.

website (www.rts.ch/emissions/court-dujour/7342318-les-materiaux-de-demain.html) since then.

As ARTE was a partner, a French and a German version of the film already exists. We plan to translate it into English, as the two professors' presentations are already in this language, and into Italian, and to use it as promotional material to advertise MARVEL.

Open doors

Scientastic

EPFL launched in 2015 a new annual event aimed at demonstrating to the general public research activities undertaken within EPFL. The 2015 edition, entitled *Tout se transforme*, was organized by the School of Engineering (STI), and was dedicated to energy research, demonstrating a variety of activities related to reducing the consommation, inventing new storage methods and developing alternative energy solutions. It took place on Saturday November 21 and gathered more than 6'000 visitors.

MARVEL was present and combined its efforts with the "Application-Centered Computational Engineering Science" (ACCES) collaboration at EPFL to present how numerical simulation and scientific visualization are contributing to energy research. Four main stands presented selected research areas of two MARVEL groups of STI (the Michele Ceriotti (COSMO) and the Nicola Marzari (THEOS) groups) and of two other groups active in simulation (Laboratory for Multiscale Mechanics Modeling (LAMMM) from William Curtin and the Laboratory for Hydraulic Machines (LMH) from François Avellan). A large screen presented stereoscopic movies of examples of these and other groups' activities, raising a sustained interest from the public and particularly from the children (Fig. 7, bottom center and right). All the projected movies can be seen on the ACCES website (acces.epfl.ch/cms/site/acces/lang/en/outreach). The THEOS group wrote an interactive software to illustrate the role of simulation in material science, in particular showing lithium diffusion in safe solid-state electrolytes used for Li-ions batteries. It also prepared 3D movies showing this lithium diffusion, as well as atomic models of various of the used components (Fig. 7, top right). The COSMO group developed a prototype of a web interface for sketch-map representation of materials' landscapes and movies showing the movement of water molecules (Fig. 7, bottom left).





Figure 7: Common MARVEL – ACCES booth at Scientastic, EPFL, on November 21, 2015 (Photography credits: top left and top right: © EPFL — Alain Herzog; top center: © EPFL — Murielle Gerber; bottom pictures: © www.sandipde.com)

The LAMMM group explained the problem of crack propagations in various materials, the importance to design alloys with a ductile behavior, in particular in the idea to use magnesium, one of the lightest metals, to improve the energetic efficiency, e.g., in transportation. Together with movies showing crack propagations, they presented samples of magnesium to be compared to other metals (iron, aluminum) and of metallic foams (Fig. 7, top center). The LMH group presented a simulation of the freesurface flow in a Pelton water turbine (first prize of the 2015 ACCES Visualization Contest, see below) together with a scale model of such turbine to show to the public how it works (Fig. 7, top left).

Mark Sawley, from ACCES, did a great job of collecting all the stereoscopic movies, helping the groups to make a 3D version of their 2D movies, and preparing all the practicalities for the projection. MARVEL management coordinated all the preparation and was present on the booth, which had a great success.

CSCS open doors

On Saturday October 17, 2015, in a day called *I cervelli che contano*, the Swiss National Supercomputing Centre (CSCS) invited the general public to discover the mysteries of science revealed by its supercomputers. Nicola Marzari, as a user of CSCS, told the public how to see a world in a grain of sand in a conference entitled *Vedere un mondo in un grano di sabbia*, to explain why the supercomputers are so important in cutting-edge research. A poster presenting MARVEL's activities and goals was prepared for the occasion.

Other open doors

As stated in the communication strategy, the institutions involved in MARVEL organise or participate in a lot of events dedicated to the presentation of science to the public at large. In this respect, Frithjof Nolting took part in the PSI open day, on October 18, 2015.

In this same respect, MARVEL will be actively present at the whole EPFL open doors that will take place from November 3 to 6, 2016, with several events, some dedicated to the general public, with about 65'000 people expected during the four days.

Events

Public conference at the Interdisciplinary Institute of Data Science

In the framework of the inauguration of the Interdisciplinary Institute of Data Science at the *Università della Svizzera italiana* (USI), Nicola Marzari was invited to close a public conference cycle on "data and society: opportunities and fears" (*Dati e Società: opportunità e paure*) on June 16, 2015 with a presentation entitled *Scienza, dati, e società: dal sistema periodico a i sommersi e i salvati.* This led to an interview in the radio broadcast *Il giardino di Albert, puntate radio* of May 9, 2015 on *RSI ReteDue*, entitled *La rivoluzione dei Big Data? Dall'uomo ai materiali, una nuova strada per la ricerca*, by Fabio Meliciani e Clara Caverzasio.

Ig Nobel Award Tour Show

EPFL will welcome the Ig Nobel Award Tour Show 2015 on March 23, 2016 through the sponsorship of MARVEL. The show will be presented by Marc Abrahams, organiser of the Ig Nobel prizes, editor of the Annals of Improbable Research and Guardian columnist, and will feature a whole host of Ig Nobel Prize winners and other "improbable" researchers. The stated aim of the Ig Nobel prizes is to "honor achievements that first make people laugh, and then make them think".

Artistic work

ACCES Visualization Contest

MARVEL is working in close collaboration with the "Application-Centered Computational Engineering Science" (ACCES) collaboration at EPFL. At the beginning of 2015, they organized together a Visualization Contest, aimed at fostering interest in the use of visualization by encouraging the insightful and visually-engaging presentation of scientific and engineering data. The contest was open for contributions from students and postdocs within the School of Engineering (STI) and the School of Architecture, Civil & Environmental Engineering (ENAC) of EPFL. A total of 55 entries were submitted to the Visualization Contest, 35 in the Image category and 20 in the Animation category. 12 were submitted by MARVEL students, one of them receiving the first prize in the Image category (Fig. 8). The prizes were sponsored by MARVEL with a jury of six experts composed of EPFL scientists and artists. These were Marilyne Andersen (architect, ENAC dean), Jamani Caillet (graphic designer and civil engineer, Mediacom), Christian Gonzenbach (artist on the EPFL campus), Nicolas Henchoz (technology design, director of EPFL+ECAL Lab), Demetri Psaltis (biologic imaging, STI dean), Sabine Süsstrunk (computational imaging, IC professor). All the entries were exposed from May 28 to June 5 at EFFL and can be found at all times on the AC-CES website (acces.epfl.ch/viz_images and acces.epfl.ch/viz_animations).

The ACCES Visualization Contest will be renewed every year and the next edition is under preparation with a new 3D Animation category as a response to the success met with stereoscopic movies at Scientastic.

Photo exhibition

The contacts initiated by Nicola Spaldin with Julie Birenbaum, artist working with light and



Figure 8: The first prize in the Image category by Andrius Merkys, MARVEL doctoral student (top) and Animation category by Christian Vessaz (bottom) of the 2015 Visualization Contest.

images, were pursued. Julie Birenbaum came to EPFL and met MARVEL direction. While no appropriate place for an exhibition on the EPFL campus could be found in association with MARVEL events in 2015, the moving of MARVEL in the new ME D building opens new opportunities for an exhibition. Julie visited the new building in January 2016 and is now thinking at exhibiting her work on the ETHZ and EPFL MARVEL groups.

Other communication-related actions

Lidia Favre-Quattropani continues to participate in the monthly meeting of all EPFL Communication officers; she was accompanied by Riccarda Caputo to some of them.

Both also attended the ScienceComm'15 conference in Solothurn on September 24 and 25, 2015. This national conference organized by the *Science et Cité* foundation brings together Swiss science communication experts and offers them a platform for exchanging views and information. This edition was the fifth one and the main topics were Crisis of science and Open science/Citizen science.

MARVEL in web news and in the press

MARVEL was mentioned several times in the electronic news of EPFL, with the promotion of hot research topics. It was also present several times in the EPFL internal *Flash* newspaper. Outside EPFL, some news in relationship with MARVEL were published on the ETHZ website. These news were usually relayed on the MARVEL website and should be part of the newsletter when available. All these entries are listed in the NIRA database as "Other forms of outputs", in a subsection tagged as "Non academic article in newspapers, magazines or on the web". Some examples are given hereafter.

• 11.09.2015, "Volcano plots used for first time in homogeneous catalysis" by Nik Papageorgiou, EPFL (actu.epfl.ch/news/volcano-plots-usedfor-first-time-in-homogeneous-c/) advertising an article published by the group of Clémence Corminboeuf in *Chemical Science*.

- 26.11.2015, "Surprising discovery of a new particle" by Oliver Morsch, ETHZ (www.ethz.ch/en/news-and-events/eth-news/news/2015/11/surprising-discovery-of-a-new-particle%20) presenting an article published by the group of Matthias Troyer in *Nature*.
- 18.12.2015, "Spintronics, lowenergy electricity take a step closer" by Nik Papageorgiou, EPFL (actu.epfl.ch/news/spintronics-lowenergy-electricity-take-a-step-clo/) highlighting an article published by the group of Oleg Yazyev in *Nature Materials*.

9 Structural aspects

Most of the contractual structural measures have already been implemented during year 1, including the tenure-track positions to Michele Ceriotti in Materials and to Oleg Yazyev in Physics, and with the IC School actively searching for a tenure-track professor in data mining. Also we completed the set up of the visualization room "CoViz1" in MXC 320, and the hardware co-located at CSCS.



Figure 1: New ME D building at EPFL. Top: external view. Center: central hall. Bottom: some of the MARVEL offices.

New ME D Building

NCCR MARVEL moved to the new ME D building (Fig. 1) in November 2015, with ample space for the headquarters and personnel. NCCR Robotics has also moved in the same building, allowing closer interactions and synergies between the two centers.

Interior shots of the common hall and of the new MARVEL offices are shown in Fig. 1 center and bottom; dedicated spaces include also a common room and a small conference room.

New distant education classroom "CoViz2"

Audiovisual services at EPFL have agreed to fund a new installation, co-located at MAR-VEL headquarters, for distant-eduction teaching and for webinar streaming. The costs (~120'000 CHF) have been taken up completely by EPFL, and resulted in a classroom with 38 seats with addressable microphones, three pan-tilt-zoom ceiling-mounted cameras, four 75" LED displays, a 27" drawing tablet, and the electronics and controllers to run all the equipment. The room is slated to be finished in March 2016; an interior shot is shown in Fig. 2.



Figure 2: Distant education room "CoViz2" in the new ME D building.

10 Lists

The contents of chapters 10 and 11 are directly provided by the SNSF from the NIRA database. For convenience we report here the publication list over the last period.

10.3 Publications over the last period

All publications have been entered in NIRA, and are listed below, sorted by group leaders. We list publications either resulting directly from the NCCR (marked with a red hexagon) or with minor contributions from the NCCR. The following lists cover the period from February 1st, 2015 to January 31st, 2016.

- 1. Scientific articles in journals with peer review
- 2. Scientific articles in journals without peer review
- 3. Publications from lists 1 and 2 involving several groups

10.3.1 Scientific articles in journals with peer review

Group of Wanda Andreoni

 C. MA, F. PIETRUCCI, AND W. ANDREONI Reaction Dynamics of CO₂ in aqueous amines from ab initio molecular dynamics: 2-amino-2methyl-1,3-propanediol (AMPD) compared to monoethanolamine (MEA)

to be published in Theoretical Chemistry Accounts (2015).

Group(s): Andreoni / Project(s): VP2

Group of Michele Ceriotti

R. Petraglia, A. Nicolaï, M. D. Wodrich, M. Ceriotti, and C. Corminboeuf

Beyond Static Structures: Putting Forth REMD as a Tool to Solve Problems in Computational Organic Chemistry

Journal of Computational Chemistry 37, 83 (2016).

Group(s): Ceriotti, Corminboeuf / Project(s): HP4

Group of Volkan Cevher

 D. E. CARLSON, E. COLLINS, Y.-P. HSIEH, L. CARIN, AND V. CEVHER Preconditioned Spectral Descent for Deep Learning in Advances in Neural Information Processing Systems 28 (NIPS 2015), C. CORTES, N. D. LAWRENCE, D. D. LEE, M. SUGIYAMA, AND R. GARNETT, eds. (2015).

Group(s): Cevher, Koch / Project(s): HP5

 D. CARLSON, Y.-P. HSIEH, E. COLLINS, L. CARIN, AND V. CEVHER Stochastic Spectral Descent for Discrete Graphical Models

IEEE	Jo	ournal	of	Selected	Top-
ics	in	Signal	P	rocessing	(2016),
doi:10	.1109	/JSTSP.2	015.25	05684.	

Group(s): Cevher, Koch / Project(s): HP5

Group of Clémence Corminboeuf

R. Petraglia, A. Nicolaï, M. D. Wodrich, M. Ceriotti, and C. Corminboeuf

Beyond Static Structures: Putting Forth REMD as a Tool to Solve Problems in Computational Organic Chemistry

Journal of Computational Chemistry **37**, 83 (2016).

Group(s): Ceriotti, Corminboeuf / Project(s): HP4

G. GRYN'OVA, M. COOTE, AND C. CORMIN-BOEUF

Theory and practice of uncommon molecular electronic configurations

WIREs Computational Molecular Science 5, 440 (2015).

Group(s): Corminboeuf / Project(s): VP2

M. BUSCH, M. WODRICH, AND C. CORMIN-BOEUF

Linear Scaling Relationships and Volcano Plots in Homogeneous Catalysis ? Revisiting the Suzuki Reaction

Chemical Science 6, 6754 (2015).

 $Group(s): \ Corminboeuf \ / \ Project(s): \ VP2$

Group of Alessandro Curioni

- P. STAAR, P. BARKOUTSOS, R. ISTRATE,
- A. C. I. MALOSSI, I. TAVERNELLI, N. MOLL,
- H. GIEFERS, C. HAGLEITNER, C. BEKAS, AND A. CURIONI
- Stochastic matrix-function estimators: Scalable Big-Data kernels with high performance

in Proceedings of the 30th IEEE International Parallel & Distributed Processing Symposium (IEEE IPDPS 2016), to be published (2016).

Group(s): Curioni / Project(s): HP5

F. FRANCO DE CARVALHO AND I. TAVERNELLI Nonadiabatic dynamics with intersystem crossings: A time-dependent density functional theory implementation

The Journal of Chemical Physics **143**, 224105 (2015).

 $Group(s): \ Curioni \ / \ Project(s): \ VP2$

Group of Antoine Georges

H. T. DANG, J. MRAVLJE, A. GEORGES, AND A. J. MILLIS

Band Structure and Terahertz Optical Conductivity of Transition Metal Oxides: Theory and Application to CaRuO₃

Physical Review Letters 115, 107003 (2015). Group(s): Georges / Project(s): VP1

 J. RUPPEN, J. TEYSSIER, O. E. PEIL, S. CATA-LANO, M. GIBERT, J. MRAVLJE, J.-M. TRISCONE, A. GEORGES, AND D. VAN DER MAREL

Optical spectroscopy and the nature of the insulating state of rare-earth nickelates

Physical Review B 92, 155145 (2015). Group(s): Georges / Project(s): VP1

- H. T. DANG, J. MRAVLJE, A. GEORGES, AND A. J. MILLIS
- *Electronic correlations, magnetism, and Hund's rule coupling in the ruthenium perovskites SrRuO*₃ *and CaRuO*₃

Physical Review B 91, 195149 (2015). Group(s): Georges / Project(s): VP1

- A. SUBEDI, O. E. PEIL, AND A. GEORGES Low-energy description of the metal-insulator transition in the rare-earth nickelates
- Physical Review B **91**, 075128 (2015). Group(s): Georges / Project(s): VP1

Group of Stefan Goedecker

 L. ZHU, M. AMSLER, T. FUHRER, B. SCHAE-FER, S. FARAJI, S. ROSTAMI, S. A. GHASEMI, A. SADEGHI, M. GRAUZINYTE, C. WOLVER-TON, AND S. GOEDECKER

> *A fingerprint based metric for measuring similarities of crystalline structures*

The Journal of Chemical Physics **144**, 034203 (2016).

Group(s): Goedecker / Project(s): HP4

M. AMSLER, S. BOTTI, M. A. L. MARQUES, T. J. LENOSKY, AND S. GOEDECKER Low-density silicon allotropes for photovoltaic

applications Physical Review B **92**, 014101 (2015).

Group(s): Goedecker / Project(s): HP4

J. A. FLORES-LIVAS, M. AMSLER, C. HEIL, A. SANNA, L. BOERI, G. PROFETA, C. WOLVERTON, S. GOEDECKER, AND E. K. U. GROSS

Superconductivity in metastable phases of phosphorus-hydride compounds under high pressure

Physical Review B 93, 020508 (2016).

Group(s): Goedecker / Project(s): HP4

 G. FISICARO, L. GENOVESE, O. ANDREUSSI, N. MARZARI, AND S. GOEDECKER A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments

The Journal of Chemical Physics **144**, 014103 (2016).

Group(s): Goedecker, Marzari / Project(s): HP3

Group of Jürg Hutter

Y. GURDAL, S. LUBER, J. HUTTER, AND M. IANNUZZI Non-innocent adsorption of Co-pyrphyrin on rutile(110)



Physical Chemistry Chemical Physics **17**, 22846 (2015).

 $Group(s): Hutter \ / \ Project(s): \ VP2$

F. Evangelisti, R. Moré, F. Hodel, S. Luber, and G. R. Patzke

3d–4f Co₃^{II}Ln(OR)₄ Cubanes as Bio-Inspired Water Oxidation Catalysts

Journal of the American Chemical Society **137**, 11076 (2015).

Group(s): Hutter / Project(s): VP2

 M. DEL BEN, J. HUTTER, AND J. VANDEVON-DELE

Forces and stress in second order Møller-Plesset perturbation theory for condensed phase systems within the resolution-of-identity Gaussian and plane waves approach

The Journal of Chemical Physics **143**, 102803 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

• M. DEL BEN, J. HUTTER, AND J. VANDEVON-DELE

Probing the structural and dynamical properties of liquid water with models including nonlocal electron correlation

The Journal of Chemical Physics **143**, 054506 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

Group of Christoph Koch

 D. E. CARLSON, E. COLLINS, Y.-P. HSIEH, L. CARIN, AND V. CEVHER Preconditioned Spectral Descent for Deep Learning

in Advances in Neural Information Processing Systems 28 (NIPS 2015), C. CORTES, N. D. LAWRENCE, D. D. LEE, M. SUGIYAMA, AND R. GARNETT, eds. (2015).

Group(s): Cevher, Koch / Project(s): HP5

 D. CARLSON, Y.-P. HSIEH, E. COLLINS, L. CARIN, AND V. CEVHER Stochastic Spectral Descent for Discrete Graphical Models

IEEE Journal of Selected Topics in Signal Processing (2016), doi:10.1109/JSTSP.2015.2505684.

Group(s): Cevher, Koch / Project(s): HP5

I. TRUMMER AND C. KOCH Multiple Query Optimization on the D-Wave 2X Adiabatic Quantum Computer

Proceedings of the VLDB Endowment (2016), arXiv:1510.06437.

Group(s): Koch / Project(s): HP5

Group of Nicola Marzari

- G. FISICARO, L. GENOVESE, O. ANDREUSSI, N. MARZARI, AND S. GOEDECKER
 - *A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments*

The Journal of Chemical Physics **144**, 014103 (2016).

Group(s): Goedecker, Marzari / Project(s): HP3

I. TIMROV, O. ANDREUSSI, A. BIANCARDI, N. MARZARI, AND S. BARONI Self-consistent continuum solvation for optical absorption of complex molecular systems in solution

The Journal of Chemical Physics **142**, 034111 (2015).

Group(s): Marzari / Project(s): HP3

N. L. NGUYEN, G. BORGHI, A. FERRETTI, I. DABO, AND N. MARZARI First-Principles Photoemission Spectroscopy and Orbital Tomography in Molecules from Koopmans-Compliant Functionals

- Physical Review Letters **114**, 166405 (2015). Group(s): Marzari / Project(s): HP3
- A. CEPELLOTTI, G. FUGALLO, L. PAULATTO,
- M. LAZZERI, F. MAURI, AND N. MARZARI Phonon hydrodynamics in two-dimensional materials
- Nature Communications 6, 6400 (2015). Group(s): Marzari / Project(s): VP2
- M. GIBERTINI AND N. MARZARI Emergence of One-Dimensional Wires of Free Carriers in Transition-Metal-Dichalcogenide Nanostructures
- Nano Letters 15, 6229 (2015).

Group(s): Marzari / Project(s): VP2, VP1

G. PIZZI, A. CEPELLOTTI, R. SABATINI, N. MARZARI, AND B. KOZINSKY AiiDA: automated interactive infrastructure and database for computational science

Computational Materials Science **111**, 218 (2016).

Group(s): Marzari / Project(s): PP6

X. CHENG, E. FABBRI, M. NACHTEGAAL, I. E. CASTELLI, M. EL KAZZI, R. HAUMONT, N. MARZARI, AND T. J. SCHMIDT

Oxygen Evolution Reaction on $La_{1-x}Sr_xCoO_3$ Perovskites: A Combined Experimental and Theoretical Study of Their Structural, Electronic, and Electrochemical Properties

Chemistry of Materials 27, 7662 (2015).

Group(s): Marzari, Schmidt / Project(s): PP7, VP2

Group of Michele Parrinello

• O. VALSSON AND M. PARRINELLO Well-Tempered Variational Approach to Enhanced Sampling

Journal of Chemical Theory and Computation **11**, 1996 (2015).

Group(s): Parrinello / Project(s): HP4

 J. MCCARTY, O. VALSSON, P. TIWARY, AND M. PARRINELLO Variationally Optimized Free-Energy Flooding for Rate Calculation

Physical Review Letters **115**, 070601 (2015). Group(s): Parrinello / Project(s): HP4

 C. PEREGO, M. SALVALAGLIO, AND M. PAR-RINELLO

Molecular dynamics simulations of solutions at constant chemical potential

The Journal of Chemical Physics **142**, 144113 (2015).

Group(s): Parrinello / Project(s): HP4

 P. Shaffer, O. Valsson, and M. Parrinello

> Enhanced, targeted sampling of high dimensional free energy landscapes using variationally-enhanced sampling, with an application to chignolin

to be published in Proceedings of the National Academy of Science of the USA (2016).

Group(s): Parrinello / Project(s): HP4

Group of Alfredo Pasquarello

W. CHEN AND A. PASQUARELLO First-principles determination of defect energy levels through hybrid density functionals and GW

Journal of Physics: Condensed Matter 27, 133202 (2015).

Group(s): Pasquarello / Project(s): VP2

W. CHEN AND A. PASQUARELLO Accurate band gaps of extended systems via efficient vertex corrections in GW

Physical Review B **92**, 041115 (2015). Group(s): Pasquarello / Project(s): VP2

F. Ambrosio, G. Miceli, and A. Pasquarello

Redox levels in aqueous solution: Effect of van der Waals interactions and hybrid functionals

The Journal of Chemical Physics **143**, 244508 (2015).

Group(s): Pasquarello / Project(s): VP2

Group of Daniele Passerone

M. D. ROSSELL, P. AGRAWAL, A. Borgschulte, C. Hébert, D. Passerone, and R. Erni

Direct Evidence of Surface Reduction in Monoclinic BiVO₄

Chemistry of Materials 27, 3593 (2015). Group(s): Passerone / Project(s): VP2

A. BASAGNI, F. SEDONA, C. A. PIGNEDOLI, M. CATTELAN, L. NICOLAS, M. CASARIN, AND M. SAMBI

Molecules-Oligomers-Nanowires-Graphene Nanoribbons: A Bottom-Up Stepwise On-Surface Covalent Synthesis Preserving Long-Range Order

Journal of the American Chemical Society **137**, 1802 (2015).

Group(s): Passerone / Project(s): VP2

C. KARAGEORGAKI, D. PASSERONE, AND K.-H. ERNST

Chiral reconstruction of Cu(110) after adsorption of fumaric acid

Surface Science **629**, 75 (2014).

Group(s): Passerone / Project(s): VP2

Group of Ursula Röthlisberger

K. VOÏTCHOVSKY, N. ASHARI ASTANI, I. TAV-ERNELLI, N. TÉTREAULT, U. ROTHLISBERGER, F. STELLACCI, M. GRÄTZEL, AND H. A. HARMS

> In Situ Mapping of the Molecular Arrangement of Amphiphilic Dye Molecules at the TiO₂ Surface of Dye-Sensitized Solar Cells

ACS Applied Materials & Interfaces 7, 10834 (2015).

Group(s): Röthlisberger / Project(s): VP2

C. YI, J. LUO, S. MELONI, A. BOZIKI, N. Ashari-Astani, C. Grätzel, S. M. Zakeeruddin, U. Röthlisberger, and M. Grätzel

*Entropic stabilization of mixed A-cation ABX*₃ *metal halide perovskites for high performance perovskite solar cells*

Energy & Environmental Science 9, 656 (2016).

Group(s): Röthlisberger / Project(s): VP2

S. MATHEW, N. ASHARI ASTANI, B. CUR-CHOD, J. DELCAMP, J. FREY, U. ROTH-LISBERGER, M. K. NAZEERUDDIN, AND M. GRÄTZEL

Synthesis, characterization and ab initio investigation of a panchromatic ullazine-porphyrin photosensitizer for dyesensitized solar cells



to be published in Journal of Materials Chemistry A (2016).

Group(s): Röthlisberger / Project(s): VP2

 S. Meloni, T. Moehl, W. Tress, M. Franckevicius, M. Saliba, Y. Lee, P. Gao, M. K. Nazeeruddin, S. M. Zakeeruddin, U. Rothlisberger, and M. Graetzel

> Ionic Polarization Induced Current-Voltage Hysteresis in CH₃NH₃PbX₃ Perovskite Solar Cells

to be published in Nature Communications (2016).

Group(s): Röthlisberger / Project(s): VP2

Group of Thomas J. Schmidt

- X. CHENG, E. FABBRI, M. NACHTEGAAL, I. E. CASTELLI, M. EL KAZZI, R. HAUMONT, N. MARZARI, AND T. J. SCHMIDT
 - Oxygen Evolution Reaction on $La_{1-x}Sr_xCoO_3$ Perovskites: A Combined Experimental and Theoretical Study of Their Structural, Electronic, and Electrochemical Properties

Chemistry of Materials 27, 7662 (2015).

Group(s): Marzari, Schmidt / Project(s): PP7, VP2

Group of Thomas Schulthess

T. C. SCHULTHESS

Programming revisited

Nature Physics **11**, 369 (2015).

Group(s): Schulthess / Project(s): PP6

R. Solcà, A. Kozhevnikov, A. Haidar, S. Tomov, J. Dongarra, and T. C. Schulthess

Efficient implementation of quantum materials simulations on distributed CPU-GPU systems

in *SC'15*, *Proceedings of the International Conference for High Performance Computing*, *Networking*, *Storage and Analysis* (ACM, New York, 2015), article No. 10, doi:10.1145/2807591.2807654.

Group(s): Schulthess / Project(s): PP6

Group of Nicola Spaldin

G. SCLAUZERO AND C. EDERER

Structural and electronic properties of epitaxially strained LaVO₃ from density functional theory and dynamical mean-field theory

Physical Review B **92**, 235112 (2015). Group(s): Spaldin / Project(s): VP1 N. S. FEDOROVA, C. EDERER, N. A. SPALDIN, AND A. SCARAMUCCI Biquadratic and ring exchange interactions in

orthorhombic perovskite manganites Physical Review B **91**, 165122 (2015).

Group(s): Spaldin / Project(s): VP1

J. A. HEUVER, A. SCARAMUCCI, Y. BLICKEN-STORFER, S. MATZEN, N. A. SPALDIN, C. ED-ERER, AND B. NOHEDA

Strain-induced magnetic anisotropy in epitaxial thin films of the spinel $CoCr_2O_4$

Physical Review B **92**, 214429 (2015).

Group(s): Spaldin / Project(s): VP1

Group of Matthias Troyer

A. A. SOLUYANOV, D. GRESCH, Z. WANG, Q. WU, M. TROYER, X. DAI, AND B. A. BERNEVIG

Type-II Weyl semimetals Nature **527**, 495 (2015).

Group(s): Troyer / Project(s): VP1

G. W. WINKLER, A. A. SOLUYANOV, AND M. TROYER Smooth gauge and Wannier functions for topo-

logical band structures in arbitrary dimensions

Physical Review B **93**, 035453 (2016).

Group(s): Troyer / Project(s): VP1

M. IAZZI AND M. TROYER Efficient continuous-time quantum Monte Carlo algorithm for fermionic lattice models

Physical Review B **91**, 241118 (2015). Group(s): Troyer / Project(s): HP3, VP1

 L. WANG, M. IAZZI, P. CORBOZ, AND M. TROYER
 Efficient continuous-time quantum Monte Carlo method for the ground state of correlated fermions

Physical Review B **91**, 235151 (2015). Group(s): Troyer / Project(s): HP3, VP1

Y.-H. LIU AND L. WANG Quantum Monte Carlo study of massimbalanced Hubbard models

Physical Review B **92**, 235129 (2015). Group(s): Troyer / Project(s): HP3, VP1

- Gloup(s). Hoyer / Hoject(s). Hi 3, V
- H. SHINAOKA, M. TROYER, AND P. WERNER Accuracy of downfolding based on the constrained random-phase approximation

Physical Review B 91, 245156 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3

H. SHINAOKA, Y. NOMURA, S. BIERMANN, M. TROYER, AND P. WERNER Negative sign problem in continuous-time

quantum Monte Carlo: optimal choice of singleparticle basis for impurity problems **Physical Review B 92, 195126 (2015).**

Group(s): Troyer, Werner / Project(s): HP3, VP1

L. WANG, H. SHINAOKA, AND M. TROYER Fidelity Susceptibility Perspective on the Kondo Effect and Impurity Quantum Phase Transitions

Physical Review Letters 115, 236601 (2015). Group(s): Troyer / Project(s): HP3, VP1

L. WANG, Y.-H. LIU, M. IAZZI, M. TROYER, AND G. HARCOS

Split Orthogonal Group: A Guiding Principle for Sign-Problem-Free Fermionic Simulations

Physical Review Letters **115**, 250601 (2015). Group(s): Troyer / Project(s): HP3, VP1

 H. Shinaoka, S. Hoshino, M. Troyer, and P. Werner

Phase Diagram of Pyrochlore Iridates: All-in– All-out Magnetic Ordering and Non-Fermi-Liquid Properties

Physical Review Letters **115**, 156401 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3

 L. WANG, Y.-H. LIU, J. IMRIŠKA, P. N. MA, AND M. TROYER Fidelity Susceptibility Made Simple: A Unified Quantum Monte Carlo Approach

Physical Review X 5, 031007 (2015).

Group(s): Troyer / Project(s): HP3, VP1

A. A. SOLUYANOV, D. GRESCH, M. TROYER, R. M. LUTCHYN, B. BAUER, AND C. NAYAK Optimizing spin-orbit splittings in InSb Majorana nanowires

to be published in Physical Review B (2015), arXiv:1510.02025.

 $Group(s): Troyer \ / \ Project(s): \ VP1$

Group of Joost VandeVondele

 M. Del Ben, J. Hutter, and J. VandeVondele

Forces and stress in second order Møller-Plesset perturbation theory for condensed phase systems within the resolution-of-identity Gaussian and plane waves approach

The Journal of Chemical Physics **143**, 102803 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

M. DEL BEN, J. HUTTER, AND J. VANDEVON-DELE

Probing the structural and dynamical properties of liquid water with models including nonlocal electron correlation

The Journal of Chemical Physics **143**, 054506 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

Group of Anatole von Lilienfeld

F. FABER, A. LINDMAA, O. A. VON LILIEN-FELD, AND R. ARMIENTO

Crystal Structure Representations for Machine Learning Models of Formation Energies

International Journal of Quantum Chemistry **115**, 1094 (2015).

Group(s): von Lilienfeld / Project(s): HP5

R. RAMAKRISHNAN, P. O. DRAL, M. RUPP, AND O. A. VON LILIENFELD Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach

Journal of Chemical Theory and Computation **11**, 2087 (2015).

Group(s): von Lilienfeld / Project(s): HP4

R. RAMAKRISHNAN, P. O. DRAL, M. RUPP, AND O. A. VON LILIENFELD

Quantum chemistry structures and properties of 134 *kilo molecules*

Scientific Data 1, 140022 (2014). Group(s): von Lilienfeld / Project(s): HP4

Group of Philipp Werner

- H. SHINAOKA, M. TROYER, AND P. WERNER Accuracy of downfolding based on the constrained random-phase approximation
- Physical Review B **91**, 245156 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3
- H. SHINAOKA, Y. NOMURA, S. BIERMANN, M. TROYER, AND P. WERNER Negative sign problem in continuous-time quantum Monte Carlo: optimal choice of singleparticle basis for impurity problems
- Physical Review B 92, 195126 (2015). Group(s): Troyer, Werner / Project(s): HP3, VP1
- P. WERNER, R. SAKUMA, F. NILSSON, AND F. ARYASETIAWAN

Dynamical screening in La_2CuO_4

Physical Review B 91, 125142 (2015).

Group(s): Werner / Project(s): HP3





All-out Magnetic Ordering and Non-Fermi-Liquid Properties

Physical Review Letters **115**, 156401 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3

Group of Oleg Yazyev

G. Autès, A. Isaeva, L. Moreschini, J. C. Johannsen, A. Pisoni, R. Mori, W. Zhang, T. G. Filatova, A. N. Kuznetsov, L. Forró, W. Van den Broek, Y. Kim, K. S. Kim, A. Lanzara, J. D. Denlinger, E. Rotenberg, A. Bostwick, M. Grioni, and O. V. Yazyev

A novel quasi-one-dimensional topological insulator in bismuth iodide β -Bi₄I₄

Nature Materials 15, 154 (2016).

Group(s): Yazyev / Project(s): VP1

J. C. JOHANNSEN, G. AUTÈS, A. CREPALDI, S. MOSER, B. CASARIN, F. CILENTO, M. ZA-CCHIAGNA, H. BERGER, A. MAGREZ, P. BUGNON, J. AVILA, M. C. ASENSIO, F. PARMIGIANI, O. V. YAZYEV, AND M. GRI-ONI

Engineering the topological surface states in the $(Sb_2)_m$ -Sb₂Te₃ (m = 0 - 3) superlattice series

Physical Review B **91**, 201101 (2015).

Group(s): Yazyev / Project(s): VP1

10.3.2 Scientific articles in journals without peer review

Group of Antoine Georges

J. MRAVLJE AND A. GEORGES

Thermopower and Entropy: lessons from Sr_2RuO_4

arXiv:1504.03860 (2015).

Group(s): Georges / Project(s): VP1

M. AICHHORN, L. POUROVSKII, P. SETH, V. VILDOSOLA, M. ZINGL, O. E. PEIL, X. DENG, J. MRAVLJE, G. J. KRABERGER, C. MARTINS, M. FERRERO, AND O. PARCOL-LET

TRIQS/DFTTools: A TRIQS application for ab initio calculations of correlated materials arXiv:1511.01302 (2015).

Group(s): Georges / Project(s): HP3

Group of Michele Parrinello

• M. SALVALAGLIO, P. TIWARY, M. MAZZOTTI, AND M. PARRINELLO Overcoming timescale and finite-size limitations to compute condensation rates at physically relevant conditions

arXiv:1508.01642 (2015).

Group(s): Parrinello / Project(s): HP4

Group of Daniele Passerone

 P. Ruffieux, S. Wang, B. Yang, C. Sanchez, J. Liu, T. Dienel, L. Talirz, P. Shinde, C. A. Pignedoli, D. Passerone,

T. DUMSLAFF, X. FENG, K. MÜLLEN, AND R. FASEL

On-surface synthesis of graphene nanoribbons with zigzag edge topology

arXiv:1511.05037 (2015).

Group(s): Passerone / Project(s): VP2

 L. TALIRZ, P. SHINDE, D. PASSERONE, AND C. A. PIGNEDOLI

Synthesis of Atomically Precise Graphene-Based Nanostructures: A Simulation Point of View

in On-Surface Synthesis. Proceedings of the International Workshop On-Surface Synthesis, Ecole des Houches, Les Houches 25-30 May 2014, A. GOUR-DON, ed. (Springer, 2016), Advances in Atom and Single Molecule Machines.

Group(s): Passerone / Project(s): VP2

Group of Matthias Troyer

Z. WANG, D. GRESCH, A. A. SOLUYANOV, W. XIE, S. KUSHWAHA, X. DAI, M. TROYER, R. J. CAVA, AND B. A. BERNEVIG

*MoTe*₂: Weyl and Line Node Topological Metal arXiv:1511.07440 (2015).

Group(s): Troyer / Project(s): VP1

Group of Anatole von Lilienfeld

F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento

Machine Learning Energies for 2 M Elpasolite (*ABC*₂*D*₆) *Crystals*

arXiv:1508.05315 (2015).

Group(s): von Lilienfeld / Project(s): HP5

10.3.3 Publications involving several groups

R. PETRAGLIA, A. NICOLAÏ, M. D. WODRICH, M. CERIOTTI, AND C. CORMINBOEUF

Beyond Static Structures: Putting Forth REMD as a Tool to Solve Problems in Computational Organic Chemistry

Journal of Computational Chemistry **37**, 83 (2016).

Group(s): Ceriotti, Corminboeuf / Project(s): HP4

- D. E. CARLSON, E. COLLINS, Y.-P. HSIEH, L. CARIN, AND V. CEVHER Preconditioned Spectral Descent for Deep
 - Learning

in Advances in Neural Information Processing Systems 28 (NIPS 2015), C. CORTES, N. D. LAWRENCE, D. D. LEE, M. SUGIYAMA, AND R. GARNETT, eds. (2015).

Group(s): Cevher, Koch / Project(s): HP5

 D. CARLSON, Y.-P. HSIEH, E. COLLINS, L. CARIN, AND V. CEVHER Stochastic Spectral Descent for Discrete Graphical Models

IEEE Journal of Selected Topics in Signal Processing (2016), doi:10.1109/JSTSP.2015.2505684.

Group(s): Cevher, Koch / Project(s): HP5

 G. FISICARO, L. GENOVESE, O. ANDREUSSI, N. MARZARI, AND S. GOEDECKER A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments

The Journal of Chemical Physics **144**, 014103 (2016).

Group(s): Goedecker, Marzari / Project(s): HP3

 M. DEL BEN, J. HUTTER, AND J. VANDEVON-DELE

> Forces and stress in second order Møller-Plesset perturbation theory for condensed phase systems within the resolution-of-identity Gaussian and plane waves approach

The Journal of Chemical Physics **143**, 102803 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

 M. Del Ben, J. Hutter, and J. VandeVondele

Probing the structural and dynamical properties of liquid water with models including nonlocal electron correlation

The Journal of Chemical Physics **143**, 054506 (2015).

Group(s): Hutter, VandeVondele / Project(s): HP3

X. CHENG, E. FABBRI, M. NACHTEGAAL, I. E. CASTELLI, M. EL KAZZI, R. HAUMONT,

N. MARZARI, AND T. J. SCHMIDT Oxygen Evolution Reaction on $La_{1-x}Sr_xCoO_3$ Perovskites: A Combined Experimental and Theoretical Study of Their Structural, Electronic, and Electrochemical Properties

Chemistry of Materials 27, 7662 (2015). Group(s): Marzari, Schmidt / Project(s): PP7, VP2

H. SHINAOKA, S. HOSHINO, M. TROYER, AND P. WERNER

Phase Diagram of Pyrochlore Iridates: All-in– All-out Magnetic Ordering and Non-Fermi-Liquid Properties

Physical Review Letters **115**, 156401 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3

• H. SHINAOKA, M. TROYER, AND P. WERNER Accuracy of downfolding based on the constrained random-phase approximation

Physical Review B **91**, 245156 (2015). Group(s): Troyer, Werner / Project(s): VP1, HP3

 H. SHINAOKA, Y. NOMURA, S. BIERMANN, M. TROYER, AND P. WERNER Negative sign problem in continuous-time quantum Monte Carlo: optimal choice of singleparticle basis for impurity problems

Physical Review B **92**, 195126 (2015).

Group(s): Troyer, Werner / Project(s): HP3, VP1

Cover picture

Computational discovery and experimental verification of the new topological insulator bismuth iodide β -Bi₄I₄ (from G. Autès, A. Isaeva, L. Moreschini, J. C. Johannsen, A. Pisoni, R. Mori, W. Zhang, T. G. Filatova, A. N. Kuznetsov, L. Forró, W. Van den Broek, Y. Kim, K. S. Kim, A. Lanzara, J. D. Denlinger, E. Rotenberg, A. Bostwick, M. Grioni, and O. V. Yazyev, *A novel quasi-one-dimensional topological insulator in bismuth iodide* β -Bi₄I₄, Nature Materials **15**, 154 (2016), Vertical Project 1).

Acknowledgements

We would like to thank MaNEP, Materials with Novel Electronic Properties, a NCCR of the first series, and in particular Christophe Berthod at UniGE, for the MaNEP LATEX style that served as a base to the present MARVEL style.