CHIMIA

CHIMIA 2020, Volume 74
ISSN 0009-4293
www. chimia.ch
Supplementa to Issue 7-8/2020

SCS
Swiss Chemical
Society

SCS Fall Meeting 2020 (online conference)
Lecture, Short Talk and Poster Abstracts

Session of Computational Chemistry

August 25, 2020
University of Bern (online conference)
http://scg.ch/fallmeeting/2020

Swiss Chemical Society
Haus der Akademien
Postfach
3001 Bern
Switzerland
info@scg.ch
www.scg.ch

Learning how to do chemical reactions from data

A. C. Vaucher[1], A. Cardinale[1], J. Geluykens[1], V. H. Nair[1], O. Schilter[1], P. Schwaller[1], A. Logallo[1], F. Zipoli[1], T. Laino[1]*

[1]IBM Research Europe, Säumerstrasse 4, 8803 Rüschlikon, Switzerland

The synthesis of novel compounds is a complex task requiring 1) an adequate retrosynthetic analysis, 2) the formulation of adequate reaction conditions for each reaction step, and 3) the actual synthesis in the laboratory. The knowledge and experience matured by chemists in decades of practice are key components to characterize a successful design across steps (1) to (3).

With more than two hundred years of reported experiments in the chemical literature, as well as recent advances in automation and machine learning, data-driven approaches provide multiple opportunities to assist chemists and accelerate the material design process. In recent years, different groups worldwide reported several algorithms to support chemists in the retrosynthetic problem [1-3]. Such computational approaches are highly useful to assist chemists in the identification of optimal synthetic pathways of target molecules starting from suitable building blocks. In particular, the approaches relying on machine-learning do not require the manual encoding of reaction rules and are able to learn directly from corpuses of reaction data. The algorithms for retrosynthesis, however, provide only a design of the synthesis of given targets, without specifying the details of the reaction steps to implement that specific transformation.

Here, we present a machine-learning model to predict the sequence of operations to perform any reaction step in the laboratory. To train this model, we used a dataset containing examples of reaction equations and associated laboratory operations, processed from reported experimental procedures in patents. To this aim, we designed a separate machine-learning model trained to extract sequences of operations from experimental procedures in patents [4] and to convert them to a sequence of structured, computer-friendly operations. We used this dataset to design an architecture to predict the exact sequence of laboratory operations given a reaction target.

Finally, we present how the models for the retrosynthetic analysis and for the prediction of laboratory operations can be coupled to commercial chemical robots to provide a platform able to autonomously synthesize molecules.

[1] Law, J.; Zsoldos, Z.; Simon, A.; Reid, D.; Liu, Y.; Khew, S. Y.; Johnson, A. P.; Major, S.; Wade, R. A.; Ando, H. Y., *J. Chem. Inf. Model.* **2009**, *49*, 593–602.
[2] Segler, M. H. S.; Preuss, M.; Waller, M. P., *Nature* **2018**, *555*, 604–610.
[3] Schwaller, P.; Petraglia, R.; Zullo, V.; Nair, V. H.; Haeuselmann, R. A.; Pisoni, R.; Bekas, C.; Iuliano, A.; Laino, T., *Chem. Sci.* **2020**, *11*, 3316–3325.
[4] Vaucher, A. C.; Zipoli, F.; Geluykens, J.; Nair, V. H.; Schwaller, P.; Laino, T., *under review*, **2020**.

Self-Parametrizing System-Focused Atomistic Models

C. Brunken[1,2], M. Reiher[1]*

[1]ETH Zürich, Laboratory for Physical Chemistry, Vladimir-Prelog-Weg 2, 8093 Zurich, [2]
christoph.brunken@phys.chem.ethz.ch

Computational studies of chemical reactions in complex environments such as proteins or metal-organic frameworks require accurate and at the same time efficient atomistic models applicable to the nanometer scale. For arbitrary system classes, an accurate parametrization of the atomistic entities will not be available, but demands a fast automated system-focused parametrization procedure to be quickly applicable, reliable, flexible, and reproducible. We develop and combine [1] an automatically parametrizable quantum chemically derived molecular mechanics model with machine-learned corrections under uncertainty quantification. Our approach first generates an accurate, physically motivated model from a minimum energy structure and its corresponding Hessian matrix by a partial Hessian fitting procedure [2] of the force constants. This model can be applied to generate a large number of configurations (e.g., via molecular dynamics) for which additional reference data can be calculated on the fly. A Δ-machine learning model is trained on these data to provide a correction to energies and forces including uncertainty estimates. The parametrization of large systems is enabled by an autonomous fragmentation approach, which is demonstrated at the example of the copper-containing protein plastocyanin. Our approach may also be employed for the generation of system-focused electrostatic molecular mechanics embedding environments in a quantum-mechanical/molecular-mechanical hybrid model for arbitrary atomistic structures at the nanoscale.

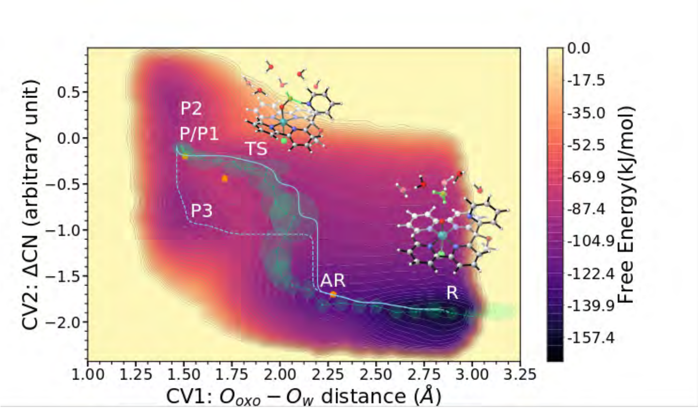
[1] Brunken, C.; Reiher, M., *J. Chem. Theory Comput.* **2020**, *16*, 1646-1665.
[2] Wang, R.; Ozhgibesov, M.; Hirao, H., *J. Comput. Chem*. **2016**, *37*, 2349-2359.

Zooming in on the O-O Bond Formation – Novel Insights from *ab initio* Molecular Dynamics with Enhanced Sampling Techniques

M. Schilling[1], R. A. Cunha[1], S. Luber[1]*

[1]University of Zurich, Department of Chemistry

The process of splitting water into molecular hydrogen and oxygen is among the most promising approaches in the field of renewable energy sources. In order to make this reaction economically viable, in-depth understanding of the crucial O-O bond formation is highly desirable. In the past we have studied a family of Ru-based water oxidation catalysts (WOCs) bearing a Py5 ligand by means of geometry optimization based simulation protocols employing density functional theory (DFT).[1,2] In our current studies we go beyond the static description of the reactive intermediates by utilizing forefront enhanced sampling techniques such as the Bluemoon ensemble and metadynamics together with high-performance DFT-based molecular dynamics simulations.



In doing so, we can more accurately describe solute-solvent interactions as well as the dynamics of the system at ambient temperature. This allows us to get detailed insights into the process of the oxygen-oxygen bond formation event, the associated reaction network, and the flexibility of the product state. Moreover, we demonstrate how crucial the choice of an appropriate collective variable is in order to capture the relevant features of the studied reaction.[3]

Lately, we have applied the developed simulation protocol to investigate the influence of the base-strength onto the O-O bond formation event. To this extent we study a modified version of the Py5 ligand that was proposed earlier in an *in silico* design study, the goal of which was to improve the catalytic performance of those Ru-based WOCs.[2] In short, we find that increasing the basicity of the intramolecular base facilitates a water nucleophilic attack mechanism by altering the solvation shell of the catalysts. This clearly shows how important an adequate description of environmental effects and solute-solvent interactions is.[4]

[1] Gil-Sepulcre *et al.* *ChemSusChem* **2017**, *10*, 4517.
[2] Schilling, Böhler and Luber, *Dalton Trans.* **2018**, *47*, 10480.
[3] Schilling, Cunha and Luber, *JCTC* **2020**, *16*, 2436-2449.
[4] Schilling, Cunha and Luber, *ACS Catal.* **2020** in revision.

Using Ultrafast Spectroscopy and Computer Simulations to Re-examine Our Picture of Photoinduced Electron Transfer

C. Rumble[1], E. Vauthey[1]*

[1]University of Geneva, Department of Physical Chemistry

Spectroscopic experiments, both time-resolved and steady-state, have provided us exquisite insight into the dynamics of photoinduced charge transfer processes. Interpretation of these experiments has been aided by Marcus theory and its numerous derivatives, a linear-response theory based primarily on a picture of spherical reactants immersed in a continuum solvent. Although this model has found much practical success, it completely ignores the molecular nature of the system, and many of the parameters underlying the theory are not directly accessible by experiment. We must therefore combine detailed computer simulations with insight from experiments in order to answer the fundamental question: how good are the predictions of traditional sphere/continuum models when the molecularity of the system is accounted for? To address this, we will present the results of combined experimental and computation studies of electron donor and acceptor systems and will demonstrate the shortcomings of traditional sphere/continuum models when the molecularity of the system is accounted for directly.

The first system we will examine are electron donor/acceptor complexes (EDACs), donor/acceptor pairs whose electronic structure allows for direct excitation from the neutral to ion pair states. Although the thermodynamics and charge recombination kinetics of EDACs have been the subject of intense study, the structural dynamics of the EDACs and their relation to charge recombination are still poorly understood. Using a unique combination of ultrafast spectroscopy and mixed quantum/classical simulations, we find that neither the ground nor excited state of EDACs can be described using a single, well-defined, structure, and that rearrangement from predominately face-to-face π stacks to edge-to-face T-shaped structures dominate the excited state structural dynamics.[1] Given the sensitivity of charge-transfer processes to donor/acceptor orientation, future theoretical descriptions of EDAC recombination dynamics must take into account the molecularity and structural evolution of the excited state in order to properly model the reaction.

Additionally, we will use classical molecular dynamics simulations to explore electron donor/acceptor systems in ways that are impossible with experimental techniques alone. We are able to demonstrate with umbrella sampling molecular dynamics simulations that the hard-sphere model and continuum electrostatics are poor descriptors of the reactant and product potentials of mean force, functions that describe the spatial distribution of donor/acceptor pairs and factor into predictions for the reaction driving force using the Weller equation. We will also show that simulated solvent reorganization energies deviate strongly from the predictions of continuum electrostatics, and that the linear-response approximation can break down in weakly polar solvents. The results of our experiments and simulations demonstrate that the real dynamics underlying photoinduced electron transfer are significantly more complex than traditional theories would predict.

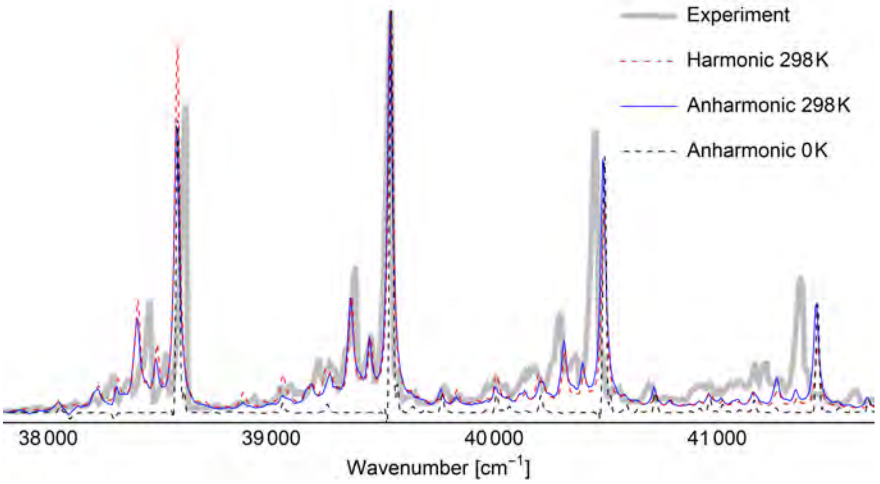
[1] Rumble, C. A.; Vauthey, E. *Phys. Chem. Chem. Phys.,* **2019,** 21, 11797–11809

On-the-fly ab initio semiclassical evaluation of vibronic spectra at finite temperature

T. Begušić[1], J. Vaníček[1]*

[1]Laboratory of Theoretical Physical Chemistry, Institute of Chemical Sciences and Engineering, EPFL

To compute and analyze vibrationally resolved electronic spectra at zero temperature, we have recently implemented the on-the-fly *ab initio* extended thawed Gaussian approximation [1], which accounts for anharmonicity, mode-mode coupling, and Herzberg-Teller effects. Here, we generalize this method in order to evaluate spectra at non-zero temperature [2]. In line with thermo-field dynamics [3, 4], we transform the von Neumann evolution of the density matrix to the Schrödinger evolution of a wavefunction in an augmented space with twice as many degrees of freedom. Due to efficiency of the extended thawed Gaussian approximation, this increase in the number of coordinates results in nearly no additional computational cost. More specifically, compared to the original, zero-temperature approach, the finite-temperature method requires no additional *ab initio* electronic structure calculations. At the same time, the new approach allows for a clear distinction among finite-temperature, anharmonicity, and Herzberg-Teller effects on spectra. We show, on a model Morse system, the advantages of the finite-temperature thawed Gaussian approximation over the commonly used global harmonic methods and apply it to evaluate the symmetry-forbidden absorption spectrum of benzene (see graphic), where all of the aforementioned effects contribute.



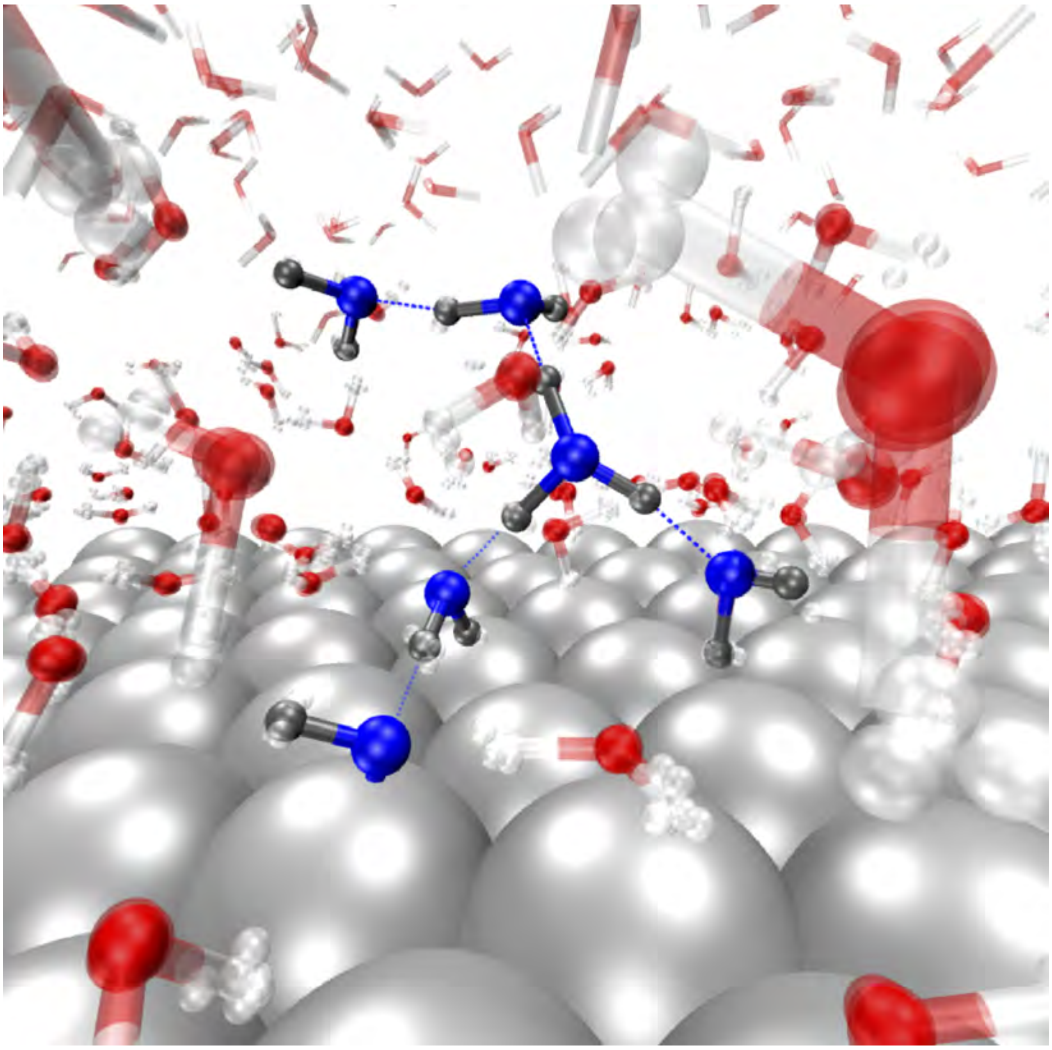
[1] Aurélien Patoz, Tomislav Begušić, Jiří Vaníček, *Journal of Physical Chemistry Letters,* **2018,** 9, 2367-2372.
[2] Tomislav Begušić, Jiří Vaníček, *Submitted* (available at arXiv:2005.09126 [physics.chem-ph]).
[3] Masuo Suzuki, *Journal of the Physical Society of Japan*, **1985**, 54, 4483.
[4] Raffaele Borrelli, Maxim F. Gelin, *Scientific Reports*, **2017**, 7, 1-9.

Ionization of Water as an Effect of Quantum Delocalization at Aqueous Electrode Interfaces

J. Lan[1], V. V. Rybkin[1], M. Iannuzzi[1]

[1]University of Zurich, Department of Chemistry

The enhanced probability of water dissociation at the aqueous electrode interfaces is predicted by path-integral ab initio molecular dynamics. The ionization process is observed at the aqueous platinum interface when nuclear quantum effects are introduced in the statistical sampling, while minor effects have been observed at the gold interface. We characterize the dissociation mechanism of the formed water ions. In spite of the fact that the concentration and lifetime of the ions might be challenging to be experimentally detectable, they may serve as a guide to future experiments. Our observation might have a significant impact on the understanding of electrochemical processes occurring at the metal electrode surface.



[1] Lan J, Rybkin V V, Iannuzzi M. I, *J. Phys. Chem. Lett.* 2020, 11, 9, 3724–3730

Enabling Drug Discovery with an Integrated Computational Approach. A Case Study in Schrödinger’s Molecular Design Platform

T. Knehans[1]

[1]Schrödinger GmbH, Thierschstraße 27, 80538 München, Germany

In structure-based small molecule drug discovery a central challenge, once a protein target has been determined, is to design compounds that bind to the target with sufficient potency and specificity. Furthermore, molecules need to exhibit the right physico-chemical and ADMET property profile to become a drug candidate. Computational chemistry methods have always been an important component in this difficult endeavor and their contribution has increased dramatically over the last years.

The Schrödinger drug discovery platform integrates solutions for predictive modeling, machine learning, data analytics, and collaboration to enable more rapid discovery of such multi-parameter optimal novel chemical matter. Multiple examples of enterprise-wide deployment of these technologies have been demonstrated to increase the probability of success and shorten the length of time required for preclinical drug discovery.

In this presentation, we will examine a case study highlighting the different ways the Schrödinger Technology Platform may be utilized in a discovery project. We will follow the development of a previously described ACC inhibitor and show how predictive methods can help in targeted water displacement in binding sites to drive potency, enumerate compound analogues to optimize physico-chemical properties or address chemical stability issues.

PNAS March 29, 2016 113 (13) E1796-E1805

Kinetic reweighting made easy – An integrator independent dynamic reweighting algorithm for biased simulations

S. M. Linker[1], R. G. Weiss[1], S. Riniker[1]*

[1]Laboratory of Physical Chemistry, ETH Zürich

Kinetics of molecular systems occur often on microsecond or higher timescales. Thus, investigation and sufficient sampling of such systems with computer simulations like molecular dynamics (MD) is challenging. External biasing potentials can accelerate such slow kinetics, but at the expense of altering the kinetic and thermodynamic properties of the system. Only recently, dynamic reweighting algorithms have been proposed to recover the unbiased kinetics (and populations) from biased systems [1-3]. However, especially the highly accurate path reweighting algorithms suffer from various implementation problems, i.e., the reweighting has to be performed on-the-fly, is integrator dependent, and only applicable to stochastic dynamics (SD).

We introduce a new, user-friendly path reweighting algorithm that can be used with any numerical scheme. Hence, our approach extends the scope of path reweighting methods to deterministic MD and is independent of the integrator used. Additionally, the new formulation leads to an increased parameter stability. Our path reweighting algorithm brings us thus closer to the goal of an easily applicable and robust reweighting scheme, which will be crucial to study and understand the dynamics of increasingly complex systems.

[1] Hao Wu, Fabian Paul, Christoph Wehmeyer and Frank Noé, *Proc. Natl. Acad. Sci. U.S.A.,* **2016**, 113, 3221-3230
[2] Lukas S. Stelzl, Adam Kells, Edina Rosta, and Gerhard Hummer, *J. Chem. Theory Comput.,* **2017**, 13, 6328–6342
[3] Luca Donati and Bettina Keller, *J. Chem. Phys.*, **2017,** 149, 072335

Atomistic simulations of catalytic confined environments from supramolecular systems to biomass conversion.

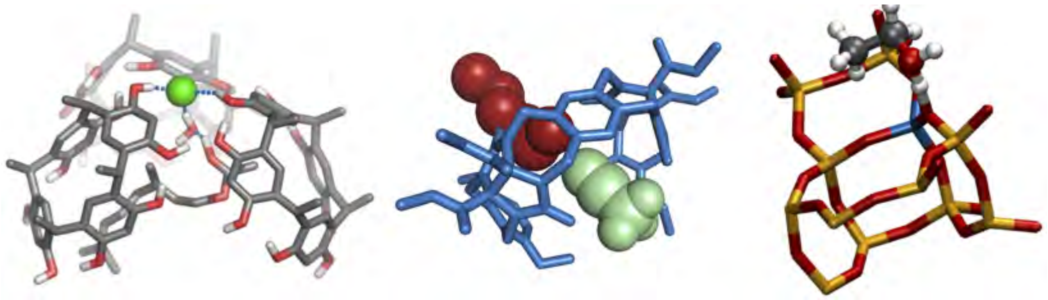
G. Piccini[1]

[1]Department of Chemistry and Applied Biosciences, ETH Zurich, c/o USI Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

In recent years, new materials and processes opened the way to novel catalytic pathways. The main feature of these systems is the simultaneous presence of a liquid substrate/solvent solution confined in nanosized catalytic environments. These processes are highly relevant because they can potentially revolutionize industrial production for a more sustainable chemical industry, ranging from fine chemistry[1] up to biomass conversion[2].Enhanced sampling molecular dynamics simulations with realistic models of condensed phase systems are in able to describe accurately the behavior of complex catalytic interfaces and reactions.

Here, we report the study of two types of catalysis characterized by well-defined cavities hosting a limited number of solvent and substrate molecules: 1) organic reactions in hexameric calixarene nanocapsules and molecular tweezers and 2) liquid-phase biomass conversion over zeolites.

For the first project we characterized the acid-base properties of an hexameric rescorcinarene capsule catalyzing terpene cyclization[3], and studied the conformational changes in the binding mode of an alkylammonium cation within a glycouril based molecular tweezers[4] leading to highly specific catalytic activity. Whereas, for the second project we aimed at understanding the process of ethanol dehydration in the acidic zeolites chabazite. For both research topics, metadynamics simulations have been used in combination with recently developed methods for studying activated processes relevant to catalysis[5-7]. This allowed us to understand the complex dynamics of different catalytic reactions and shed light on the atomistic behavior of these complex molecular systems.



[1] J. Meeuwissen, J. N. H. Reek, *Nat. Chem.* **2010**, *2*, 615–621.
[2] T. Ennaert, J. Van Aelst, J. Dijkmans, R. De Clercq, W. Schutyser, M. Dusselier, D. Verboekend, B. F. Sels, *Chem. Soc. Rev.* **2016**, *45*, 584–611.
[3] S. Merget, L. Catti, G. Piccini, K. Tiefenbacher, *J. Am. Chem. Soc.* **2020** *142*, 9, 4400-4410.
[4] M. Knezevic, M. Heilmann, G. Piccini, Tiefenbacher K. *Angew. Chem. Int. Ed.* **2020** *Accepted article*.
[5] D. Mendels, G. Piccini, M. Parrinello, *J. Phys. Chem. Lett.* **2018**, *9* (11), 2776–2781.
[6] E. Grifoni, G. Piccini, M. Parrinello, *Proc. Natl. Ac. Sci.* **2019**, 116 (10), 4054–4057
[7] G. Piccini, M. Parrinello, *J. Phys. Chem. Lett.* **2019**, *10* (13), 3727–3731.

Simulating the Ghost: Quantum Dynamics of Solvated Electron

V. V. Rybkin[1]

[1]University of Zurich, Department of Chemistry

The nature of bulk hydrated electron has been a challenge for both experiment and theory due to its short life time and high reactivity, and the need for a high-level of electronic structure theory to achieve predictive accuracy. The lack of a classical atomistic structural formula makes it exceedingly difficult to model the solvated electron using conventional empirical force fields, which describe the system in terms of interactions between point particles associated with atomic nuclei.

Here we overcome this problem using a machine-learning model, that is sufficiently flexible to describe the effect of the excess electron on the structure of the surrounding water, without including the electron in the model explicitly. The resulting potential is not onlyable to reproduce the stable cavity structure, but also recovers the correct lo-calization dynamics starting from neat water. The machine-learning model achieves the accuracy of the state-of-the-art correlated wave function method (second-order Møller-Plesset perturbation theory) it is trained on. It is sufficiently inexpensive to afford a full quantum statisticaland dynamical description, and allows us to achieve a highly accurate determination of the structure, diffusion mechanisms and vibrational spectroscopyof the solvated electron.

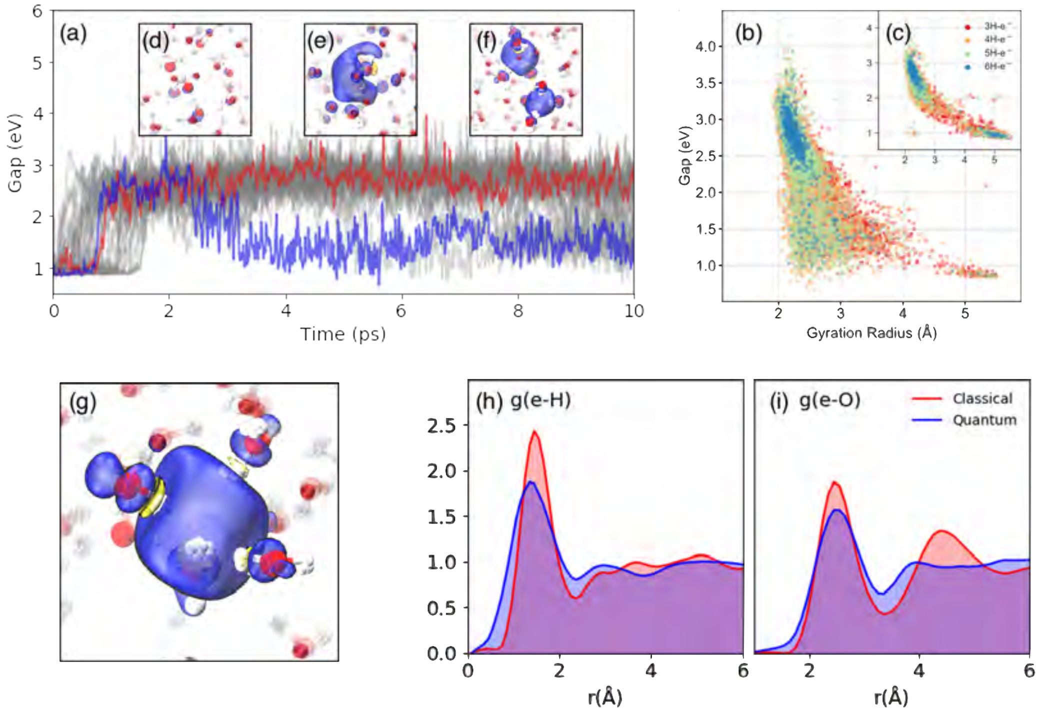


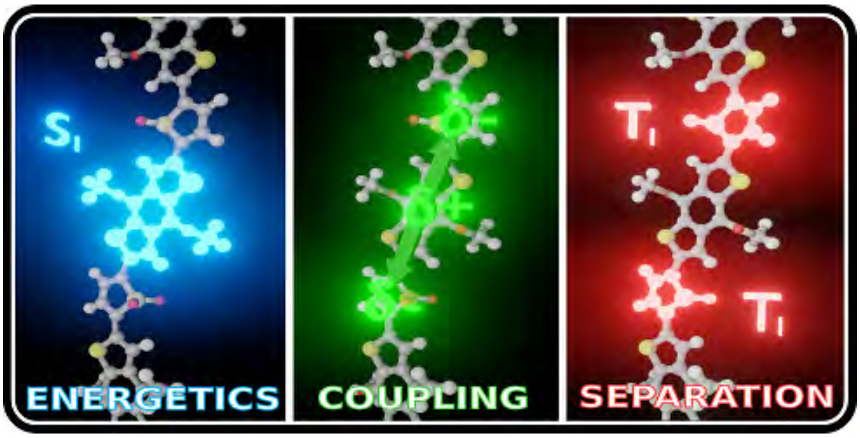
Figure: a) Time evolution of the gap of the excess electron. The representative evolution curves are marked in red (only involved with single-cavity) and blue (also involved twin-cavity). (b) Band gaps, gyration radii of the spin density distribution with different solvated electron coordination, as obtained from the quantum simulations. (c) distribution as obtained from classical simulations. The coordination number of solvated electron with respect to hydrogen are marked in red (3 H atoms), orange (4 H atoms), yellow (5 H atoms) and green (6 H atoms). Subplots (d,e,f): Spin densities of the solvated electron from quantum dynamics showing (d) the delocalized electron (e) the pre-solvated electron (f) a twin-cavity (g) a single cavity. (h,i) RDFs of hydrogen (h) and oxygen (e) atoms wrt to center of the solvated electron from classical (red) and quantum (blue) molecular dynamics.

Designing Singlet Fission Candidates from Donor-Acceptor Copolymers

T. Blaskovits[1], M. Fumanal[1], S. Vela[1], C. Corminboeuf[1]*

[1]Laboratory for Computational Molecular Design, ISIC, SB, EPFL

Singlet fission (SF) has shown great potential to improve the power conversion efficiency in photovoltaic devices by means of promoting singlet-triplet exciton splitting upon the absorption of light.[1] Traditionally, SF is targeted as an intermolecular process, however its dependence on crystal packing makes molecular design difficult. In contrast, intramolecular SF (iSF) enables the exploration of tunable bi-chromophoric systems following well-defined structure-property relationships. However, the number of reported materials that undergo both inter- and intramolecular singlet fission remains remarkably narrow, and clear trends pointing towards new materials have not yet been established. In this context, the building-block approach to conjugated donor-acceptor polymer synthesis provides fertile ground to formulate simple and robust rules for the design of candidates which respond to SF requirements, namely appropriate energies and electronic coupling.[2,3] In this work, we focus our analysis on the $E(S_1)\geq 2E(T_1)$ thermodynamic condition, the appropriate charge transfer (CT) character of the S_1 state and the geometric separation of the local T_1 states in the material.[4] These parameters, computed using time-dependent density functional theory, are initially applied to a small, chemically diverse, curated library of truncated dimers of synthetically feasible donor-acceptor copolymers. The efficiency of the protocol is validated by the correct identification of some of the few materials that have previously exhibited SF behaviour experimentally. Furthermore, we demonstrate that the CT character of the excited state can be mapped to the frontier molecular orbital energies of the constituent monomers, thus providing a cost-effective initial step for an accelerated screening of promising iSF donor-acceptor pairs. This computationally cheap and scalable framework has been applied to a database constructed of 5888 unexplored, though synthetically accessible, donor-acceptor compounds.[5] By means of high-throughput screening, our approach identifies novel copolymers potentially capable of iSF, thereby enriching the palette of singlet fission materials.



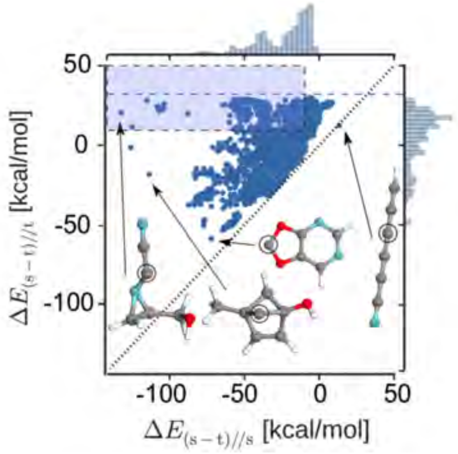
[1] Millicent B. Smith, Josef Michl, *Chem. Rev.*, **2010**, *110*, 6891-6936.
[2] David Casanova, *Chem. Rev.*, **2018**, *118*, 7164-7207.
[3] Eric Busby, Jianlong Xia, Qin Wu, Jonathan Z. Low, Rui Song, John R. Miller, Xiaoyang Zhu, Luis M. Campos, Matthew Y. Sfeir, *Nat. Mater.*, **2015**, *14*, 426-433.
[4] J. Terence Blaskovits, Maria Fumanal, Sergi Vela, Clémence Corminboeuf, **2020**, DOI: 10.26434/chemrxiv.11861817.v2
[5] J. Terence Blaskovits, Maria Fumanal, Sergi Vela, Clémence Corminboeuf, *Manuscript in preparation.*

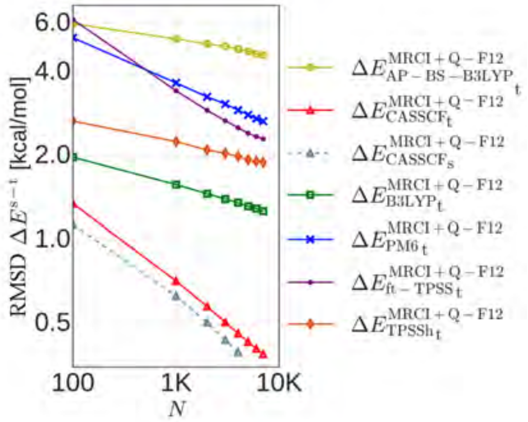
Spin gap learning of carbenes with strongly correlated electrons

M. Schwilk[1], D. N. Tahchieva[1], O. A. von Lilienfeld[1]*

[1]Institute of Physical Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel

Machine learning (ML) of large chemical spaces of molecules with strongly correlated electrons is a mostly unexplored task. We present multi-level ML approaches for the QMspin carbene data set [1] with accurate learning of spin gaps in free carbenes at small computational cost [2], enabling accurate ML predictions for hundreds of thousands of carbene candidates in chemical space. Our recently published QMspin data set of ~13k carbene structures with singlet-triplet gaps computed at MRCI level of theory and verified carbene electronic structure also gives new fundamental insights into general carbene chemistry [1] that are tightly linked to the potential presence of strong electron correlation. Finally, it is shown how orbital entanglement, orbital energies, Coulomb, and exchange interactions of the carbene open shell orbitals can be accurately learned through chemical space [2] and how this constitutes a new approach to multi-state ML.





[1] Max Schwilk*, Diana N. Tahchieva*, O. Anatole von Lilienfeld, Large yet bounded: spin gap ranges in carbenes 2020, arXiv:2004.10600

[2] Diana N. Tahchieva*, Max Schwilk*, O. Anatole von Lilienfeld, Spin gap learning of carbenes with strongly correlated electrons 2020, in preparation

*equal contributions

Quantum Dynamics with the Time-Dependent Density Matrix Renormalization Group

A. Baiardi[1], M. Reiher[1]

[1]ETH Zürich

Nowadays, ultrafast spectroscopies enable one to probe the dynamics under non-equlibrium conditions of a molecule on the natural time scales of both its electronic and nuclear motions. Ab-initio quantum simulations are often the key to translate the experimental outcomes into structural and electronic molecular properties [1-2].
In this contribution, we show how the Density Matrix Renormalization Group (DMRG) algorithm [3] can be applied to molecular quantum dynamics. We encode the wave function as a matrix product state and obtain the equations of motion by applying the Dirac-Frenkel variational principle. We apply the resulting algorithm, known as tangent space-based time-dependent DMRG (TD-DMRG) [4], to both electronic [5] and vibrational [6-7] Hamiltonians and simulate the nuclear and electronic dynamics, possibly coupled together, of molecules with several dozens of degrees of freedom [8]. We further improve TD-DMRG for long-time propagations by coupling it with different single-particle optimization schemes and with efficient integration algorithms. We also extend it to imaginary-time propagations to derive an alternative optimization algorithm to standard DMRG [8]. When feasible, we assess the accuracy of the simulations by comparison with state-of-the-art Configuration Interaction results.

[1] Meyer H. D., *WIRES Comp. Mol Sci.*, **2011**, 2, 351.
[2] Sato T., Ishikawa K. L., *Phys. Rev. A.*, **2013**, 88, 023402.
[3] Baiardi A., Reiher M., *J. Chem. Phys.*, **2020**, 152, 040903.
[4] Haegeman J., Lubich C., Oseledets I., Vandereycken F., *Phys. Rev. B.*, **2016**, 94, 1.
[5] Keller S., Dolfi M., Troyer M., Reiher M., *J. Chem. Phys.*, **2015**, 143, 244118.
[6] Baiardi A., Stein C.J., Barone V., Reiher M., *J. Chem. Theory Comput.*, **2017**, 13, 3764.
[7] Baiardi A., Stein C.J., Barone V., Reiher M., *J. Chem. Phys.*, **2019**, 150, 094113.
[8] Baiardi A., Reiher M., *J. Chem. Theory Comput.*, **2019**, 15, 3481.

Real-Time Spectroscopy for an Interactive Quantum Chemistry Framework

F. Bosia[1], T. Weymuth[1], S. Polonius[1], M. Reiher[1]*

[1]ETH Zürich, Laboratory of Physical Chemistry

A framework for interactive quantum chemistry established in our group allows for the interactive exploration of chemical reactivity along potential energy surfaces.[1,2] In order to facilitate the efficient spectroscopic identification of compounds, an extension of this framework has been developed.[3] We present a real-time spectroscopy tool that provides infrared and ultraviolet-visible spectroscopic information during the interactive exploration. Characteristic spectral signals may serve as diagnostic probes that do not require ultimate accuracy with respect to peak position and intensity. Such diagnostic bands from the interactive spectroscopy software can provide valuable hints on structures with characteristic spectroscopic properties. Since very accurate ab initio electronic structure methods are too inefficient to be applied in an interactive framework, we rely on our open-source in-house semiempirical quantum chemistry program, Sparrow.[4,5] This allows for fast evaluation of the spectroscopic properties of interest with either NDDO or DFTB semiempirical models that may be later refined with more accurate methods if deemed necessary. The approximations invoked to improve the accuracy vs. computational cost trade-off of the real-time spectroscopy tool are presented and critically reviewed.

[1] M. P. Haag, A. C. Vaucher, M. Bosson, S. Redon, M. Reiher, *ChemPhysChem*, **2014**, 15, 3301-3319.
[2] A. C. Vaucher, M. Reiher, *J. Chem. Inf. Model.,* **2016**, 56, 1470-1478.
[3] F. Bosia, T. Weymuth, S. Polonius, M. Reiher, **2020**, *in preparation*.
[4] T. Husch, A. C. Vaucher, M. Reiher, *Int. J. Quantum Chem.*, **2018**, 118, e25799.
[5] F. Bosia, T. Husch, A. C. Vaucher, M. Reiher, *Zenodo*, **2020**, 10.5281/zenodo.3768527.

A time-reversible integrator for the time-dependent Schrödinger equation on an adaptive grid

S. Choi[1], J. Vaníček[1]*

[1]Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

One of the most accurate methods for solving the time-dependent Schrödinger equation uses a combination of the dynamic Fourier method with the split-operator algorithm on a tensor-product grid. To reduce the number of required grid points, we let the grid move together with the wavepacket but find that the naïve algorithm based on an alternate evolution of the wavefunction and grid destroys the time reversibility of the exact evolution. Yet, we show that the time reversibility is recovered if the wavefunction and grid are evolved simultaneously during each kinetic or potential step; this is achieved by using the Ehrenfest theorem together with the splitting method. The proposed algorithm is conditionally stable, symmetric, and time-reversible and conserves the norm of the wavefunction. The preservation of these geometric properties is shown analytically and demonstrated numerically on a three-dimensional harmonic model and collinear model of He–H_2 scattering. We also show that the proposed algorithm can be symmetrically composed to obtain time-reversible integrators of an arbitrary even order. We observed 10 000-fold speedup by using the tenth-order instead of the second-order method to obtain a solution with a time discretization error below 10^{-9}. Moreover, using the adaptive grid instead of the fixed grid resulted in a 64-fold reduction in the required number of grid points in the harmonic system and made it possible to simulate the He–H_2 scattering for six times longer while maintaining reasonable accuracy. Applicability of the algorithm to high-dimensional quantum dynamics is demonstrated using the strongly anharmonic eight-dimensional Hénon–Heiles model.

[1] Seonghoon Choi, Jiří Vaníček, J. Chem. Phys., **2019**, 151, 234102.

The Role of Substrate Binding in the O_2 Activation Pathway of Naphthalene 1,2-Dioxygenase

K. S. Csizi[1,2], L. Eckert[1], C. Brunken[1], M. Reiher[1]*, T. B. Hofstetter[2]*

[1]ETH Zürich, Laboratory of Physical Chemistry , [2]Eawag, Swiss Federal Institute of Aquatic Science

Mononuclear non-heme iron enzymes constitute a large superfamily of biocatalysts ubiquituous in nature. [1] The scope of catalyzed reactions is broad, yet in many cases they involve oxygenations with molecular O_2 via transition-metal oxygen intermediates. Naphthalene 1,2-dioxygenase (NDO) is a representative that carries out cis-dihydroxylation of naphthalene. It was observed experimentally that binding of O_2 to the metal center is gated by the presence of naphthalene in the active site and reduction of the Rieske cofactor in the proximate enzyme subunit, triggering a change in the coordination environment of the metal center. [2, 3] The presence of naphthalene in the active site without binding to the catalytic iron is plausible biologically to avoid unproductive O_2 activation. Whereas the actual catalytic dihydroxylation step is well understood [4], the mechanism that initiates dioxygen activation remains elusive.

We present a combined DFT and CASSCF/CASPT2 study to elucidate the mechanism that gates the O_2 activation process. Extensive DFT studies of reaction energies and barriers allow for a feasible broad sampling of the reaction network and theoretical validation of experimental observations. [5] As the spin-state problem is an inherent phenomenon in transition metal chemistry and DFT does not systematically predict spin-state energetics, accurate ab-initio reference data are generated from CCSD and CASSCF/CASPT2 calculations, where the active-space selection follows an automated protocol developed within our group. [6] Possible Multi-state-reactivity involving the change of spin state during the course of reaction is investigated with particular focus on the generation of radical intermediates on the substrate that could react in a spin-allowed process with triplet dioxygen.

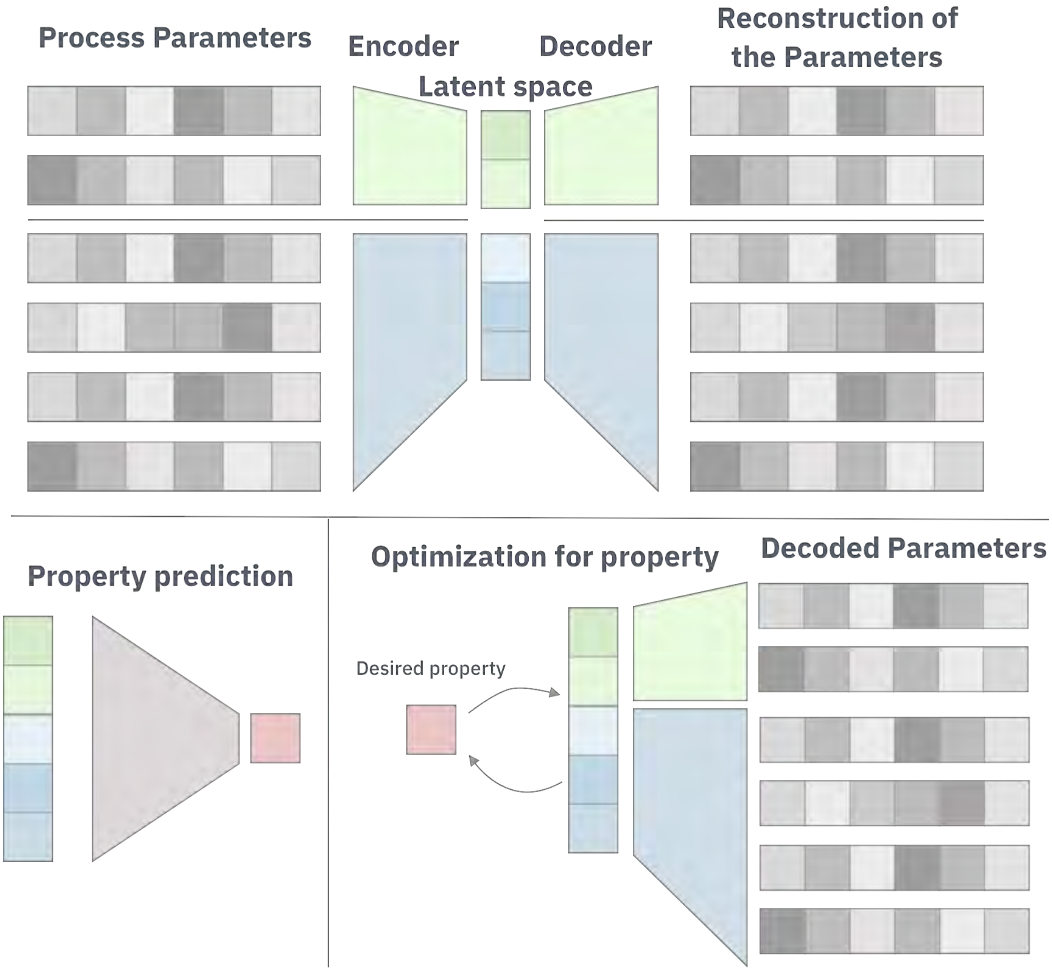
[1] E. I. Solomon, S. Goudarzi, K. D. Sutherlin, *Biochemistry* **2016**, 55(46), 6363–6374.
[2] T. C. Yang, M. D. Wolfe, M. B. Neibergall, Y. Mekmouche, J. D. Lipscomb, B. M. Hoffman, *J. Am.Chem. Soc.* **2003**, 12(8), 2034–2035.
[3] M. D. Wolfe, J. V. Parales, D. T. Gibson, J. D. Lipscomb, *J. Biol. Chem.* **2001**, 276(3), 1945–1953.
[4] A. Bassan, M. R. Blomberg, P. E. Siegbahn, *J. Biol. Inorg. Chem.* **2004**, 9(4), 439–452.
[5] M. Reiher, *Chimia* **2009**, 63(3), 140–145.
[6] C. J. Stein, M. Reiher, *J. Chem. Theory Comput*. **2016**, 12(4), 1760–1771.

Advanced Data-Driven Manufacturing using Autoencoders

T. Gaudin[1,3], O. Schilter[2], F. Zipoli[2], T. Laino[2]

[1]IBM Research Europe, [2]IBM Research Europe, Säumerstrasse 4, 8803 Rüschli, [3]University of Toronto

In this work we present a data-driven approach to the design and fabrication of novel materials via encoder-decoder based models, which offer a promising approach to compress the data into a reduced latent space to improve the material design task. For instance, a formulation of an alloy or polymeric material with a specific composition can undergo various processes, and one single process can be applied to various formulation compositions. Here, we used autoencoders to generate continuous representations of the independent parameters in a reduced space, which is called latent space [1]. We used an autoencoder for the composition of the formulation and another one to encode the sequence of processing steps. From these autoencoders we constructed a latent representation from which we can predict various properties (in the case of an alloy, it could be some mechanical properties and for polymeric materials, it could be physico-chemical ones) using an extra neural network. It is inspired by the work of Gomes-Bombarelli et al. [2] in the domain of molecular structure prediction, we opted for the construction of latent spaces, enabling the use of Gaussian processes to identify points that have certain desired properties. These points are then decoded back to the input parameter space. This approach exploits that latent space has fewer dimensions than the parameter space, making it easier to search. It could also lead to solutions that were potentially not observed in the dataset yet. Moreover, having this flexible framework with different autoencoders also allows to condition the search on a part of the parameters. For instance, we can look for a specific process yielding the highest tensile strength given an alloy composition. The benefit of the presented framework is its general applicability to a variety of material design problems.



[1] D. P. Kingma and M. Welling, Auto-encoding Variational Bayes, ICRL 2014, https://arxiv.org/abs/1312.6114

[2] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams and A. Aspuru-Guzik, ACS Central Science, 2018, 4, 268–276.

Electric Field Gradient calculation within Frozen-Density Embedding Theory

Y. Gimbal-Zofka[1], M. Fu[1], N. Ricardi[1], C. E. Gonzalez[1], T. A. Wesolowski[1]*

[1]Department of Physical Chemistry, University of Geneva

Frozen-Density Embedding Theory[1],[2] (FDET) provides a system-independent formal framework for multi-level computational methods which describe the effect of a frozen electron density of the environment on the chromophore. Within this framework, it is possible to treat either explicitly or implicitly the effect of the embedding species on the electron distribution in the environment in numerical simulations. In the state specific pre-polarization protocol[3], the response of the environment to the change of the electronic structure of the chromophore can be taken into account in the process of generating . In practice, due to some approximations made for the FDET potential, number of pre-polarization cycles and the choice of , this treatment is not exact.

The goal is to demonstrate the potential of FDET for Electric Field Gradient (EFG) calculation compare to commonly used approximations , and in particular the state specific pre-polarization with ESP-derived point charges[3] (ChelPG[4])

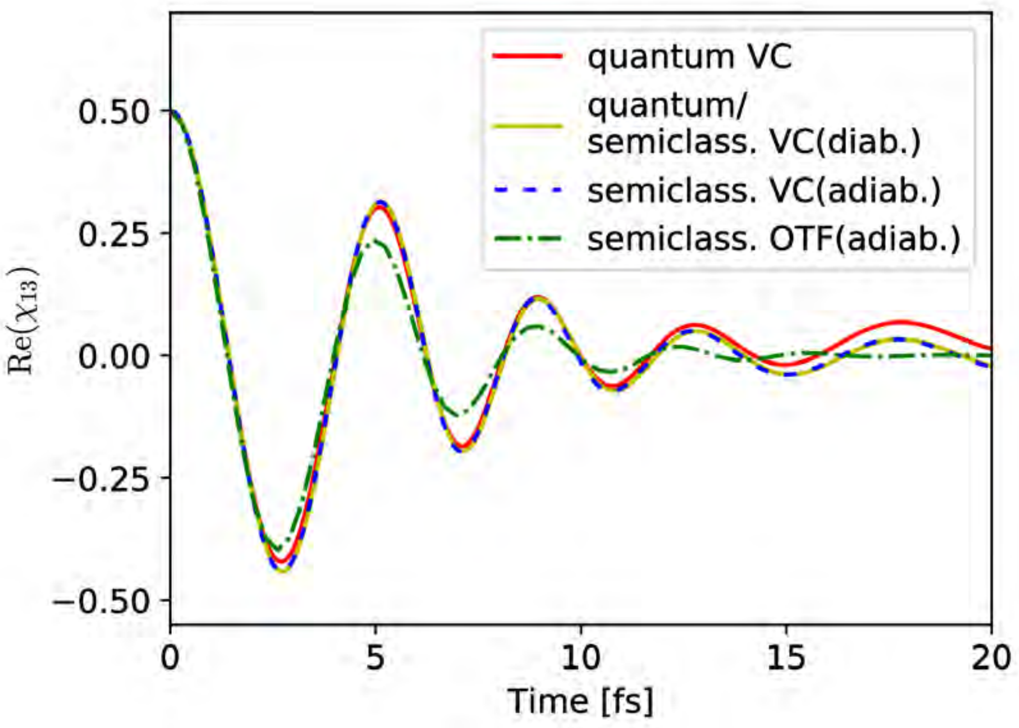
[1] Tomasz Wesolowski, Arieh Warshel, *J. Phys. Chem.*, **1993**, 97, 8050
[2] Tomasz Wesolowski, *Phys. Rev. A*, **2008**, 77, 012504
[3] Niccolò Ricardi, Alexander Zech, Yann Gimbal-Zofka, Tomasz A. Wesolowski, *Phys. Chem. Chem. Phys.*, **2018**, 20, 26053-26062
[4] C. M. Breneman and K. B. Wiberg. *J. Comput. Chem.*, **1990**, Vol. 11, No. 3, 361-373

On-the-fly ab initio semiclassical evaluation of electronic coherences in polyatomic molecules reveals a simple mechanism of decoherence

N. Golubev[1], T. Begušić[1], J. Vaníček[1]

[1] Laboratory of Theoretical Physical Chemistry, EPFL

Irradiation of a molecular system by an intense laser field can trigger dynamics of both electronic and nuclear subsystems. The lighter electrons usually move on much faster, attosecond time scale but the slow nuclear rearrangement damps ultrafast electronic oscillations, leading to the decoherence of the electronic dynamics within a few femtoseconds. We show that a simple, single-trajectory semiclassical scheme can evaluate the electronic coherence time in polyatomic molecules accurately by demonstrating an excellent agreement with full-dimensional quantum calculations. In contrast to numerical quantum methods, the semiclassical one reveals the physical mechanism of decoherence beyond the general blame on nuclear motion. In the propiolic acid, the rate of decoherence and the large deviation from the static frequency of electronic oscillations are quantitatively described with just two semiclassical parameters---the phase space distance and signed area between the trajectories moving on two electronic surfaces. Because it evaluates the electronic structure on the fly, the semiclassical technique avoids the "curse of dimensionality" and should be useful for preselecting molecules for experimental studies.

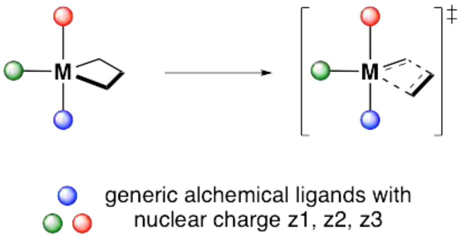


Identifying and Designing Ligand Fields by Computational Alchemy

C. Gordon[1], C. Copéret[1]

[1]ETH Zürich

The identification of ligand fields that enable chemical reactivity is at the core of homogeneous catalysis. Today, this process mainly relies on a wealth of empirical knowledge, gained by organometallic chemists over the last decades, in combination with experimental screening. In this work, we present an alternative approach and tool for the rapid identification of ligand fields that enable chemical processes. The approach relies on the use of generic, alchemical ligands, the properties of which are modified by a computational modification of their nuclear charge.[1,2] These alchemical ligands can emulate a broad range of ligand strengths and can be back-transformed into chemically accessible ligands. Thus, this approach allows us to reduce the chemical space of possible ligands to only a few parameters (i.e. the nuclear charges of the alchemical ligands). This tool can identify guidelines for catalyst design, by providing ranges of ligands and ligand-fields that yield the targeted chemical properties. As a proof of principle, I will illustrate this approach with the investigation of olefin metathesis catalysts.



step 1: minimize $\Delta E^{\ddagger}$ by optimizing nuclear charge z1, z2, z3
step 2: identify real ligands that emulate optimized alchemical ligands

[1] E. B. Wilson, *J. Chem. Phys.* **1962**, *36*, 2232.
[2] O. A. von Lilienfeld, *J. Chem. Phys.* **2006**, *125*, 154104.

Automated Chemical Reaction Space Exploration of a Proton Reduction Mechanism

S. A. Grimmel[1], M. Reiher[1]*

[1]ETH Zürich, Laboratory of Physical Chemistry

For an in-depth understanding of a chemical system, all relevant intermediates as well as the elementary reactions connecting them must be known. However, due to the immense complexity resulting from, *e.g.*, the presence of side reactions and conformational freedom, manual investigations of reaction networks by means of quantum-chemical methods are often unfeasible and error-prone. Therefore, a variety of different approaches for the automated exploration of chemical reaction networks have been proposed in the last decade.[1,2]
In this endeavor, we present our software Chemoton[3] that allows for the automated exploration of reaction networks based on first-principles heuristics.[4] The application of conceptual electronic structure theory enables us to establish algorithms that are agnostic regarding the type of molecules under consideration. As an example, we discuss the possibility of exploiting the molecular electrostatic potential for predicting protonation sites that we applied to the exploration of the mechanism of proton reduction catalyzed by a hydrogenase model complex.[5]

[1] Gregor N. Simm, Alain C. Vaucher, Markus Reiher, *J. Phys. Chem. A,* **2019**, 123, 385-399.
[2] Jan P. Unsleber, Markus Reiher, *Annu. Rev. Phys. Chem.,* **2020**, 71, 121-142.
[3] Gregor N. Simm, Markus Reiher, *J. Chem. Theory Comput.,* **2017**, 13, 6108-6119.
[4] Maike Bergeler, Gregor N. Simm, Jonny Proppe, Markus Reiher, *J. Chem. Theory Comput.,* **2015**, 11, 5712-5722.
[5] Stephanie A. Grimmel, Markus Reiher, *Faraday Discuss.,* **2019**, 220, 443-463.

Complete active space analysis of a reaction pathway: investigation of the oxygen-oxygen bond formation[1]

R. Han[1], S. Luber[1]*

[1]University of Zurich, Department of Chemistry

Water nucleophilic attack is an important step in water oxidation reactions, which have been widely studied using density functional theory (DFT). Nevertheless, a single determinant DFT picture may be insufficient for a deeper insight into the process, in particular during the oxygen-oxygen bond formation. In this work, we use complete active space self-consistent field (CASSCF) calculations and describe an approach for a complete active space analysis along a reaction pathway. This is applied to the water nucleophilic attack at a Ru-based catalyst, which has successfully been used for efficient water oxidation[2] and in silico design of new water oxidation catalysts recently[3].

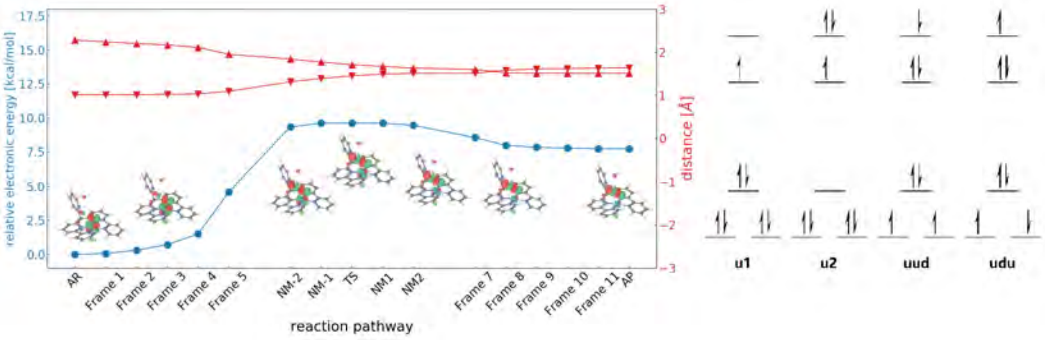


Figure 1. Reaction pathway of water nucleophilic attack at a Ru-based catalyst (left) and short-cut scheme of four major configuration state functions from the CASSCF calculations (right).

[1] R. Han and S. Luber, *J. Comput. Chem.* **2020**, 1, 12.
[2] M. Gil-Sepulcre, et al., *ChemSusChem* **2017**, 10, 4517-4525.
[3] M. Schilling, M. Böhler, S. Luber, *Dalton Trans.* **2018**, 47, 10480.

Excited-state properties for semi-empirical tight binding

A. Hehn[1], F. Belleflamme[1], J. Hutter[1]*

[1]Universität Zürich

Understanding properties of electronic excited states is fundamental for describing various kinds of spectroscopy and for modeling photochemical phenomena. The therefore required analytical excited-state gradients have been reported for time-dependent density functional theory for Gaussian and plane wave basis sets and have become a standard tool in quantum chemistry program packages [1,2]. Based on these developments, but aiming now for a highly efficient semi-empirical approach suitable for excited-state dynamics and large-scale applications on confined materials, we present excited-state gradients for the simplified Tamm-Dancoff approximation (sTDA) [3] as implemented in the mixed Gaussian and plane waves framework of the CP2K program package [4]. For electronic spectra, sTDA was already proven to be a powerful alternative to conventional hybrid density functionals, enabling to treat larger system sizes of thousands of atoms while still yielding reasonable accuracy with deviations in the range of 0.5 eV. We outline the implementation strategy for corresponding excited-state properties and show first benchmark results in comparison to standard time-dependent density functional theory approaches.

[1] F. Furche, R. Ahlrichs, *Journal of Chemical Physics*, **2002**, 117, 7433-7447.
[2] J. Hutter, *Journal of Chemical Physics*, **2003**, 118, 3928-3934.
[3] S. Grimme, *Journal of Chemical Physics*, **2013**, 138, 244104-14.
[4] The CP2K developers group. CP2K is freely available from http://www.cp2k.org/.

Quantum tunnelling in electron-transfer reactions

E. R. Heller[1], J. O. Richardson[1]*

[1]Lab. for Physical Chemisty, ETH Zürich

Electron transfer rates between weakly coupled states are predominantly calculated using Marcus theory. Due to its inherent classical character, however, experimental reaction rates can deviate from the Marcus result by several orders of magnitude due to the influence of quantum tunnelling and zero-point energy. These effects prove to be of particular importance in the so called “inverted” regime, where reaction rates become smaller with increasing thermodynamic driving force.

In principle the correct quantum mechanical rate can be obtained by applying Fermi’s golden rule. This, in turn, requires knowledge of the nuclear wavefunctions of the two reactive wells, which are not accessible for all but very small systems. Semiclassical instanton theory is able to capture nuclear quantum effects in a numerically efficient way allowing application to complex, anharmonic systems. Recently this method has been extended to describe the inverted regime by locating an optimal tunnelling pathway formed by two trajectories, one of which travels in negative imaginary time giving a picture reminiscent of the scattering of particles and antiparticles [1]. Furthermore, the theory was amended to implicitly account for the effect of a quantum or classical solvent bath [2]. The method thus enables to capture nuclear tunnelling effects in even larger multiscale-type problems not only in electron-transfer reactions but also in light-matter interactions, electrochemistry and other fields.

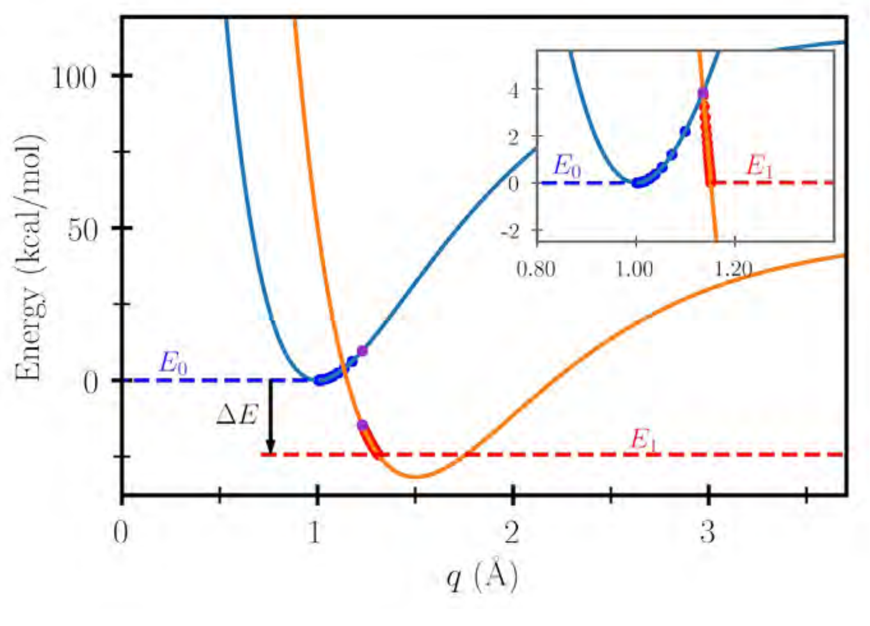


Figure 1: Reactant potential (blue solid lines) and product potential (orange solid lines) of a system of two interacting Morse oscillators. The trajectories forming the instanton are shown at discrete time steps by blue (reactant) and red (product) dots and their energies, E_0 and E_1, are depicted by the blue and red dashed lines. Due to the coupling of this system to an external bath, these energies are separated by the energy gap ΔE. The inset shows the instanton for the same system but without the presence of a bath, where $\Delta E = 0$.

[1] E. R. Heller, J. O. Richardson, *The Journal of Chemical Physics*, **2020**, 152, 034106.
[2] E. R. Heller, J. O. Richardson, *arxiv [physics.chem-ph]*, **2020**, 2005.05860.

Quantum Tunnelling in Isotopically-Substituted Systems with Machine Learning and Ab Initio Instanton Theory

G. Laude[1], E. M. Jahr[1], J. O. Richardson[1]*

[1]Lab. for Physical Chemisty, ETH Zürich

Instanton theory is a semiclassical approximation to the path-integral formulation of quantum mechanics. This theory allows for a description of quantum tunnelling in terms of dominant tunnelling pathways. By applying this theory and combining it with a ring-polymer discretisation of the aforementioned pathways, we obtain a computationally efficient method of computing reaction rates and tunnelling splittings in molecular systems [1, 2]. However, one is still prevented from studying systems with a large number of degrees of freedom due to prohibitive computational costs. We have thus combined instanton theory with a machine-learning method, Gaussian Process Regression (GPR) to resolve this. This allows us to limit the amount of *ab initio* electronic structure calculations required by using a small subset of data to accurately predict the PES around the dominant tunnelling pathway. This has been successfully applied to compute rates for H + CH_4 and H + C_2H_6 [3]. We have then applied it compute the kinetic isotope effect in the reactions of Muonium (μ^+e^-) and H with CH_4, C_2H_6 and C_3H_8 [4], and demonstrate our findings here.

We next consider quantum tunnelling effects in asymmetric double-well systems, wherein the local minima in the two wells have the same energy, but the widths differ slightly. We have successfully derived a semiclassical instanton formulation for the tunnelling splitting in such systems. The new approach was benchmarked using a set of one- and two-dimensional models, for which the results compare favourably with numerically exact quantum mechanics. We now apply this new formulation to compute tunnelling splittings in various isotopically-substituted versions of the proton transfer in malonaldehyde in full dimensionality [5].

[1] Jeremy O. Richardson, *International Reviews in Physical Chemistry*, **2018,** 37, 171-216.
[2] Jeremy O. Richardson, *The Journal of Chemical Physics*, **2018,** 148, 200901.
[3] Gabriel Laude, Danilo Calderini, David P. Tew and Jeremy O. Richardson, *Faraday Discussions*, **2018,** 212, 237-258.
[4] Gabriel Laude, Danilo Calderini, Ralph Welsch and Jeremy O. Richardson, *Physical Chemistry Chemical Physics* (submitted).
[5] Elena M. Jahr, Gabriel Laude and Jeremy O. Richardson, *in preparation*.

Spin-mapping approach to nonadiabatic molecular dynamics

J. E. Runeson[1], J. O. Richardson[1]*

[1]Laboratory for Physical Chemisty, ETH Zurich

Many important chemical reactions are *nonadiabatic*, meaning that they involve transitions between different electronic states. Typical examples include light-induced reactions found in biological processes such as photosynthesis, vision, and DNA repair, as well as light-harvesting compounds relevant for future energy technology. Nonadiabatic processes are particularly challenging to computational chemists, because the coupling between the nuclear and electronic motion prevents us from using the famous Born--Oppenheimer approximation.

It is therefore necessary to develop new approximations in order to simulate realistic large-scale nonadiabatic processes. A common starting point is to divide the system into a subsystem that is treated with quantum dynamics and an environment that is treated with classical dynamics. There is however no unique way to treat the coupling between the two parts, and different approaches may lead to various levels of accuracy.

We have resently developed a method that can robustly describe the nonadiabatic dynamics of general multi-level systems in complex condensed-phase environments [1]. For two-level systems, the method can be thought of as a mapping to a spin in a magnetic field, and for higher-level systems we use its group-theoretical generalization [2]. This leads to a set of classical equations of motion that are able to describe complex quantum problems. As an example, we have simulated electron transfer in photosynthetic green sulphur bacteria with state-of-the-art accuracy (see figure). It outperforms conventional Ehrenfest calculations, scales linearly with system size, and opens the path towards nonadiabatic extensions of path-integral-based methods that include nuclear zero-point energy, which could make it a powerful tool for computing rates in general nonadiabatic reactions.

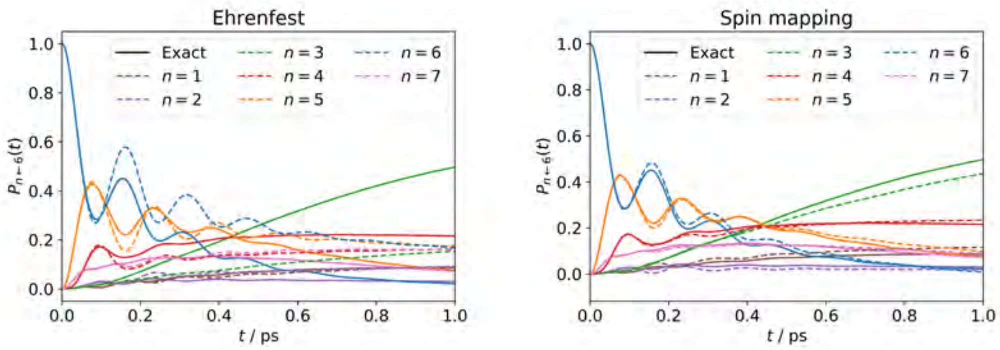


Figure: Electron population dynamics in a seven-state model of the Fenna--Matthews--Olsen complex, where solid lines indicate the exact benchmark. Our spin-mapping dynamics (right) outperforms conventional Ehrenfest dynamics (left) for the same computational effort.

[1] Johan E. Runeson and Jeremy O. Richardson, J. Chem. Phys., **2019**, 151, 044119.
[2] Johan E. Runeson and Jeremy O. Richardson, J. Chem. Phys., **2020**, 152, 084110.

Nonadiabatic effects in electronic spectroscopy: a novel quantum-classical approach

J. R. Mannouch[1], J. O. Richardson[1]*

[1]Lab. for Physical Chemisty, ETH Zurich

Spectroscopy offers a powerful tool for investigating the microscopic dynamics and structure of many systems. The diversity of available methods means that a wide range of dynamical processes can now be probed by these techniques. However such quantities are often challenging to obtain theoretically for complex condensed phase systems, as they usually require the computation of the full quantum dynamics of systems containing a large number of degrees of freedom.

Previously, semiclassical approximations have been successfully employed to calculate the underlying optical response functions for both linear and nonlinear spectroscopy, by describing the dynamics of nuclei on single adiabatic surfaces using ensembles of classical trajectories[1]. However such approximations are much harder to apply to more complicated systems that contain conical intersections or avoided crossings within the excited electronic state manifold; these crossings facilitate ultra-fast population transfer during the excited state dynamics and mean that nonadiabatic couplings can not be neglected.

To evaluate dynamical quantities in nonadiabatic systems, similar semiclassical approximations for the full electron-nuclear dynamics can be derived with the use of mapping variables[2,3]. Within these approaches the electronic subsystem is mapped exactly onto a fictitious system, which has a well defined classical limit. The dynamics associated with the mapped system is then described by an ensemble of classical trajectories. Such approaches have already been successfully used to calculate other time-dependent correlation functions associated with coupled electon-nuclear systems[4,5]. In particular, we have recently developed a new classical-trajectory mapping-based technique, called spin-PLDM, which seems ideally suited to obtaining optical response functions.

In this poster, we will show how spin-PLDM can be applied to accurately and efficiently compute linear absorption spectra of nonadiabatic condensed phase systems, such as pyrazine and the Fenna-Matthews-Olsen complex. We will also briefly outline and discuss how our formalism can be easily extended to treat more complex nonlinear spectra, such as 2D electronic spectroscopy.

[1] P. L. McRobbie, E. Geva, *J. Phys. Chem. A,* **2009**, 113, 10425–10434.
[2] G. Stock, M. Thoss, *Phys. Rev. Lett.,* **1997**, 78, 578–581.
[3] H. D. Meyer, W. H. Miller, *J. Chem. Phys.,* **1979**, 70, 3214–3223.
[4] M. A. C. Saller, A. Kelly, J. O. Richardson, *J. Chem. Phys.,* **2019**, 150, 071101.
[5] J. E. Runeson, J. O. Richardson, *J. Chem. Phys.* **2019**, 151, 044119.

Extending spectroscopic applications of real time propagation

J. Mattiat[1], S. Luber[2]*

[1]University of Zurich, Department of Chemistry, [2]University of Zurich, Department of Chemistry

In this contribution simulations of various spectroscopic signatures in the time domain using real time time dependent density functional theory (RT-TDDFT) will be presented. RT-TDDFT allows to directly simulate the optical response of a molecular system with regard to an electro magnetic pulse. The field is treated explicitly and non-perturbatively as opposed to standard energy domain approaches which rely on perturbation theory.

The favorable scaling of RT-TDDFT compared to perturbative approaches and the fact that the whole frequency dependent response functions are obtained in one go, makes it a promising candidate for every day spectroscopic applications for large systems.

In particular recent progress for the calculation of UV/VIS absorption [1], electric circular dichroism (ECD) [2], (resonance) Raman [1] and Raman optical activity (ROA) spectra [3] will be presented.

Emphasis will be on a origin independent description of linear response properties, which is vital for the prediction of ECD and ROA spectra.

In addition a novel method for the calculation of (Resonance) Raman spectra for condensed phase systems will be presented.

[1] J. Mattiat, S. Luber, *J. Chem. Phys.*, 149, 174108, **2018**
[2] J. Mattiat, S. Luber, *Chem. Phys.*, 527, 110464, **2019**
[3] J. Mattiat, S. Luber, *J. Chem. Phys.*, 151, 234110, **2019**

Reliability of Tailored Coupled Cluster in Different Correlation Regimes

M. Mörchen[1], L. Freitag[2], M. Reiher[1]*

[1]ETH Zürich, Laboratory of Physical Chemistry, [2]University of Vienna, Institute for Theoretical Chemistry

The description of strongly correlated electrons within the coupled cluster framework is still an unsolved problem in the theoretical chemistry community. The tailored coupled cluster (TCC) method, originally introduced by Kinoshita et al. [1] is a very promising ansatz, as it keeps the simplicity of a single-reference coupled cluster method, while incorporating the multi-reference wavefunction through amplitudes obtained from a previous active space calculation. We present an investigation of the TCC wavefunction based on model systems [2,3] introduced by Jankowski et al. [4]. Since these systems rely on an accurate description for both static and dynamic correlation, they provide adequate test cases for multi-reference coupled cluster methods. We investigated the reliability of the TCC method with respect to the best available wavefunction provided by the full configuration interaction method. In addition to the energy error and T1 and D1 diagnostics [5-7], we used the overlap of the TCC and full configuration interaction wavefunction as a quality measure for the wavefunction. Based on this, we critically reviewed all advantages and drawbacks of the TCC method.

[1] T. Kinoshita, O. Hino, R. J. Bartlett, *J. Chem. Phys.*, **2005**, 123, 074106.
[2] K. Jankowski, J. Paldus, *Int. J. Quantum Chem.*, **1980**, 18, 1243–1269.
[3] K. Jankowski, L. Meissner, J. Wasilewski, *Int. J. Quantum Chem.*, **1985**, 28, 931–942.
[4] M. Mörchen, L. Freitag, M. Reiher, **2020**, *in preparation*.
[5] T. J. Lee, J. E. Rice, G. E. Scuseria, H. F. Schaefer, *Theor. Chim. Acta*, **1989**, 75, 81–98.
[6] T. J. Lee, P. R. Taylor, *Int. J. Quantum Chem.*, **1989**, 36, 199–207.
[7] C. L. Janssen, I. M. Nielsen, *Chem. Phys. Lett.*, **1998**, 290, 423–430.

Frozen Density Embedding Theory exploiting experiment-derived densities for the environment.

N. Ricardi[1], M. Ernst[2], T. A. Wesolowski[3]*, P. Macchi[4]*

[1]Department of Physical Chemistry, University of Geneva, [2]Department of Physical Chemistry, University of Bern, [3]Department of Physical Chemistry, Geneva, [4]Department of Physical Chemistry, Milano

Frozen Density Embedding Theory[1] is a theoretical framework for multiscale calculations which allows to account for the effect of an environment on a system of interest. The environment is described only by its electron density and the position of the nuclei, and the interaction with the system of interested is modelled by means of orbital-free density functionals. The quantum mechanical treatment of the interaction accounts for non-electrostatic components of the interaction, mitigating issues such as electron spill out, in turn allowing for larger and hence more accurate basis sets than conventional electrostatic embedding. The fact that the environment is modelled via its density allows for several[2] options to obtain it, each presenting its positives and negatives, and consequently a domain of applicability. These mainly differ in the basis set expansion and in the modelling of mutual polarisation . Two peculiar options stick out of this set of variants: time-averaged densities and experimental densities. The former consists in the possibility of averaging the density of the environment over a period of time and perform a single embedding calculation instead of a large set, while the latter refers to the possibility of using densities obtained from X-ray diffraction. Densities from X-ray diffraction are routinely used to deduce the position of the nuclei. Obtaining wavefunctions, albeit appealing, presents several obstacles. One valid approach is the so called X-ray restrained Hartree Fock[3] where a linear combination of the energy of the Slater determinant and of a parameter for the agreement with the experimental data is minimised. The contribution of the latter can be slowly increased, slowly turning on the perturbation due to crystal field and electron correlation. This is possible up to a certain extent, which depends on the quality of the crystallographic data.

We show that for eight H-bonded complexes consisting of a small chromophore (acrolein, acrylic acid, or acetone) coordinated by a glycylglycine molecule, X-ray restrained densities can be successfully used as environment densities in Frozen Density Embedding Theory calculations. We calculated vertical excitation energies and compared them to supermolecular calculations. Specifically, we show that the change is gradual and smooth with the increase of weight for the experimental information, and that the range spun by the experiment-derived densities is similar to that of common quantum mechanical methods (HF, CCSD, MP2, PBE).

Direct comparison of experiment-derived to conventional post Hartree-Fock densities shows some similar features but also specific features, which can be attributed to the crystal field effects.A similar behaviour is found in the densities and properties from embedding calculations, which shows that not only embedding in an experiment-derived density is numerically sound, but indeed captures some physical effect.

This paves the way for embedding calculations of chromophores embedded in crystals, where several challenges for producing environment densities arise.

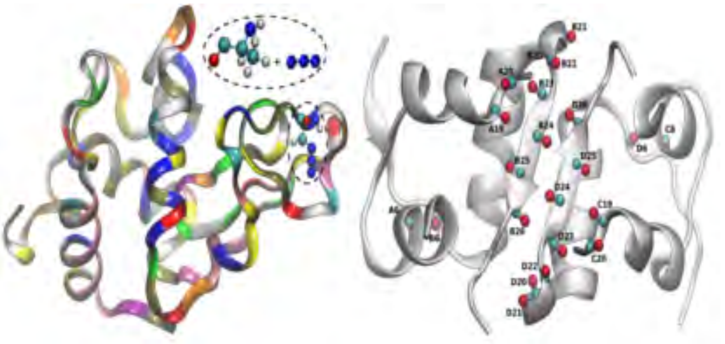
[1] Tomasz Adam Wesolowski, *Physical Review A* , ***2008***, *77*(1), 1–9.
[2] Niccolò Ricardi, Alexander Zech, Yann Gimbal-Zofka, Tomasz Adam Wesolowski, *Physical Chemistry Chemical Physics*
[3] Piero Macchi, Jean-Michel Gillet, Francis Taulelle, Javier Campo, Nicolas Claiser, Claude Lecomte, *IUCrJ*, **2015**, *2*, 441–451.

Towards Spatially Sensitive Infrared Probes

S. M. Salehi[1]

[1]Department of Chemistry, University of Basel

Finding the relationship between local structure and dynamics from IR spectroscopy is a challenging subject and molecular dynamics simulations have become a powerful tool to study the structure, dynamics, and function of complex systems at atomic detail. To achieve this, two different probes were chosen to study the local dynamics: Azide probe (N$_3^-$) [1] and Amide-I probe (CO) [2,3]. The N$_3^-$ asymmetric stretch is well separated from most organic chromophores/protein bands and the CO probe is native to all proteins, sensitive and not strongly influenced by side chains.[4] First, vibrational spectroscopy of N$_3^-$ will be discussed in the gas phase and solvent with its application in Azido-homoalanine (AHA) [5] and protein (Azido-Lysozyme) which demonstrate site-selective dynamics. Second, the amide-I vibration for monomeric and dimeric Wild type (WT) and mutant (B24G and B24A) of insulin is studied using validated energy functions for the CO probe with the aim to quantify the differences in the spectroscopy and dynamics of monomeric with dimeric and WT with mutant insulin.



[1] Salehi et al. J. Phys. Chem. B 2019, 123, 3282-3290
[2] Dhayalan et al. ChemBioChem 2016, 17, 415−420.
[3] Raghunathan et al. J. Phys. Chem. B 2018, 122, 7038−704.
[4] Ganim et al. A. Acc. Chem. Res. 2008, 31, 432–441.
[5] Zanobini et al. J. Phys. Chem. B 2018, 122, 10118-10125

Molecular dynamics simulations to understand conformational behavior of cyclic peptides at polar/apolar interfaces

C. Schellhaas[1], S. M. Linker[1], S. Riniker[1]*

[1]Laboratory of Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland

We use molecular dynamics simulations and Markov state modeling of cyclic peptides in order to rationalize the origin of their membrane permeability. Previous work has shown that an important determinant for this permeability is a peptide's conformational behaviour in the water phase [3]. In this study, we focus on the conformational behaviour of cyclic peptides at the polar/apolar interface of water and chloroform that mimics the cell membrane. Our findings support the hypothesis that only peptides in the so-called 'closed' conformation, where polar groups are shielded from the environment, are able to permeate through the apolar region of a membrane. Moreover, the polar/apolar interface catalyses the interconversion from an open to the closed conformation, a key event for permeation. Based on our simulations, we show how specific amino acids and different local environments influence this closing event. These findings may be helpful in the design of future generations of membrane permeable cyclic peptides.

[1] Celine Adessi, Claudio Soto, *Current medicinal chemistry*, **2002**, 9.9, 963-978.
[2] Andrew T. Bockus, et al. *Journal of medicinal chemistry*, **2015**, 58.11, 4581-4589.
[3] Witek, Jagna, et al. *Journal of chemical information and modeling*, **2018**, 59.1, 294-308.

Chemical Reaction Grammar as Seen by Neural Networks: Unsupervised Attention-guided Atom-Mapping

P. Schwaller[1,3], B. Hoover[2], J. L. Reymond[3], H. Strobelt[2], T. Laino[1]*

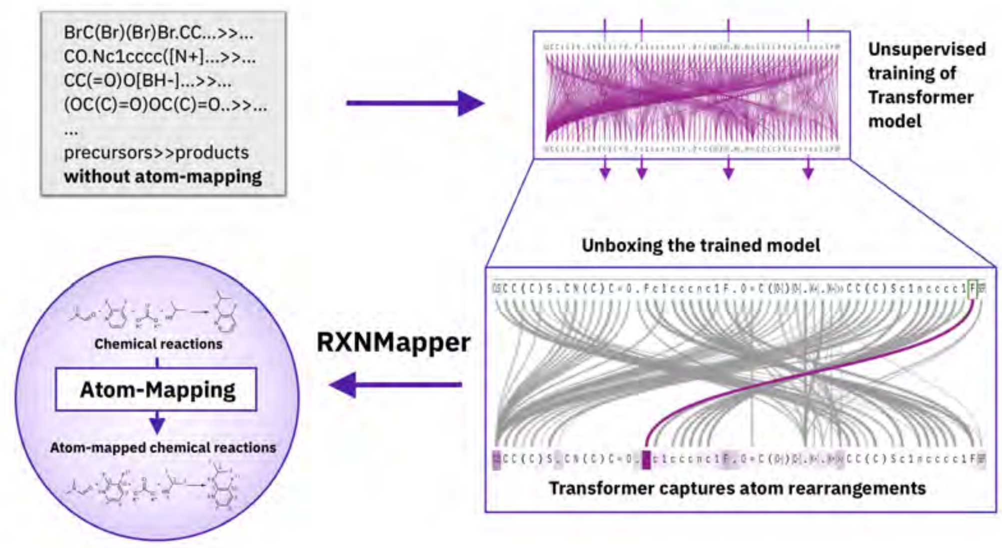
[1]IBM Research Europe, Säumerstrasse 4, 8803 Rüschli, [2]IBM Research – Cambridge / MIT-IBM Lab, United States, [3]Department of Chemistry and Biochemistry, University of Bern

Knowing how atoms rearrange during a chemical transformation is fundamental to numerous applications aiming to accelerate organic synthesis and molecular discovery. This labelling is known as atom-mapping and is an NP-hard problem. Current solutions use a combination of graph-theoretical approaches, heuristics, and rule-based systems. Unfortunately, the existing mappings and algorithms are often prone to errors and quality issues, which limit the effectiveness of supervised approaches. Self-supervised neural networks called Transformers, on the other hand, have recently shown tremendous potential when applied to textual representations of different domain-specific data, such as chemical reactions. Here we demonstrate that attention weights learned by a Transformer, without supervision or human labelling, encode atom rearrangement information between products and reactants. We build a chemically agnostic attention-guided reaction mapper that shows a remarkable performance in terms of accuracy and speed, even for strongly imbalanced reactions. Our work suggests that unannotated collections of chemical reactions contain all the relevant information to construct coherent sets of reaction rules. This finding provides the missing link between data-driven and rule-based approaches and will stimulate machine-assisted discovery in the chemical domain.

Demo: http://rxnmapper.ai

Code: https://github.com/rxn4chemistry/rxnmapper

Reference: 10.26434/chemrxiv.12298559



Benchmarking Tight-Binding Predictions for Metal-Organic Frameworks

B. Sertcan[1], A. Hehn[1], J. Hutter[1]*

[1]Universität Zürich

Metal-organic frameworks (MOFs) are metal clusters linked with organic ligands forming porous and stable networks. Modification of the building blocks enables modular design and controllable synthesis of millions of materials with different chemical composition and functionalities, qualifying MOFs as competitive alternatives to traditional porous materials like zeolites or covalent-organic networks [1]. Due to this versatility, high-throughput screening is an efficient and necessary tool to identify the most suitable candidates from the thousands of hypothetical or experimental structures stored in databases. We used the semi-empirical tight binding method GFN-xTB [2], which was recently implemented in CP2K [3], to screen a test set of 10 diverse MOFs [4] and the CoRE-MOF database [5], presenting benchmark results in comparison to density-functional theory.

[1] O. Farha, J. T. Hupp, Acc. Chem. Res. 43, 1166 (2010).
[2] S. Grimme, C. Bannwarth, P. Shushkov, J. Chem. Theory Comput. 13, 5 (2017).
[3] T. D. Kühne, et al., arXiv:2003.03868 [physics.chem-ph] preprint
[4] D. Nazarian, P. Ganesh, D. S. Sholl, J. Mater. Chem. A 3, 22432 (2015).
[5] D. Nazarian, J. Camp, Y. G. Chung, R. Q. Snurr, D. S. Sholl, Chem. Mater. 29, 2521 (2017).

The effects of van der Waals dispersion corrections on the study of carbon nitrides membranes performance

M. Tohidi Vahdat[1,2], D. Campi[2], N. Marzari[2]*, K. V. Agrawal[1]*

[1]Laboratory of Advanced Separations (LAS), [2]Theory and Simulation of Materials (THEOS)

Simplicity and energy efficiency are two main reasons that have attracted a lot of attention toward membrane-based separations, especially using nanoporous two-dimensional materials. Graphitic carbon nitrides frameworks are highly promising for molecular sieving because they are stable in a wide range of operating conditions. The two most studied frameworks in the family of carbon nitrides are C_2N and poly(triazine imide) (PTI). They have been already synthesized, and recently exfoliated to atom-thick single-layer in our and other laboratories [1][2].

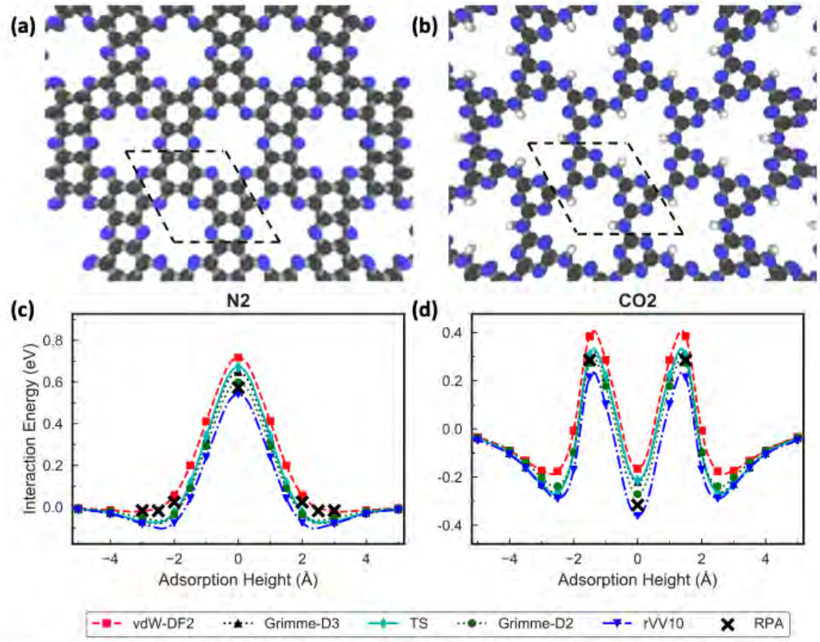


Figure 1. a) The structure of C_2N lattice. b) The structure of PTI lattice. c) The potential energy surface of N_2, d) CO_2 on a C_2N monolayer using van der Waals-corrected DFT in comparison to the points calculated by EXX/RPA method.

The potential energy surfaces of different gases on a C_2N and PTI monolayer have been investigated with van der Waals-corrected density functional theory (DFT). We noticed that different adsorption behaviour for gases are outlined in literature. So in this study, we explore the effects of using different van der Waals approximations on the study of membrane performance of C_2N and PTI. Due to large discrepancy among different van der Waals dispersions, we benchmark our calculations based on more advanced theory technique, which is the exact exchange correlation in conjugation with random phase approximation (RPA). The way that computational study of materials for membrane application could help the community is to address whether a set of materials is promising for a particular separation or not. In this way the community can focus on potentially interesting materials in the way of realizing high performance membranes. Here, we show more emphasis on the accuracy of capturing van der Waals interactions is crucial.

[1] B. V. Lotsch, *JACS*, **2014**, pp. 1–14.
[2] L. F. Villalobos, Sci. Adv., **2020**, 6, eaay9851.

Unassisted Noise-Reduction of Chemical Reactions Data Sets

A. Toniato[1], P. Schwaller[1,2], A. Cardinale[1,3], J. Geluykens[1], T. Laino[1]*

[1]IBM Research Europe, Säumerstrasse 4, 8803 Rüschlikon, Switzerland, [2]Department of Chemistry, University of Bern, Bern, Switzerland, [3]Department of Chemistry, University of Pisa, Pisa, Italy

Existing deep learning models applied to reaction prediction in organic chemistry are able to reach extremely high levels of accuracy (> 90% for NLP- based ones [1]). With no chemical knowledge embedded than the information learnt from reaction data, the quality of the data sets plays a crucial role in the performance of the prediction models. While human curation is prohibitively expensive, the need for unaided approaches to remove chemically incorrect entries from existing data sets is essential to improve the performance of artificial intelligence models in synthetic chemistry tasks. Inspired by a previous work on classification tasks [2], we propose a machine learning-based, unassisted approach to remove chemically wrong entries (noise) from chemical reaction collections. Results show that models trained on cleaned and balanced data sets improve the quality of the predictions without a decrease in performance. For the retrosynthetic models the round-trip accuracy is enhanced by 13% and the value of the cumulative Jensen Shannon metric is lowered down to 70% of its original value, while maintaining high values of coverage (97%) and constant class-diversity (1.6) at inference.

[1] Philippe Schwaller ,Teodoro Laino, Théophile Gaudin, Peter Bolgar, Christopher A. Hunter, Costas Bekas, and Alpha A. Lee*****, *ACS Central Science*, **2019**, 5, 1572–1583.
[2] Mariya Toneva, Alessandro Sordoni, Remi Tachet des Combes, Adam Trischler , Yoshua Bengio, Geoffrey J. Gordon, *arXiv e-prints* **2018**, arXiv:1812.05159

Nonadiabatic Instanton Rate Theory – Quantum Tunnelling beyond the Born-Oppenheimer Approximation

R. A. Zarotiadis[1], J. O. Richardson[1]*

[1]Lab. for Physical Chemisty, ETH Zürich

The accurate prediction of (quantum) rates and rate mechanisms is key to advance our understanding of chemical reactions in a variety of disciplines, ranging from astrophysics to biology. In many cases the Born-Oppenheimer approximation is a valid choice and molecular reactions are well described on a single electronic surface.

Semiclassical instanton theory, which is commonly used to accurately predict rates within the Born-Oppenheimer approximation, is based on locating the optimal tunnelling pathway (the instanton) on the adiabatic surface using the stationary-action principle, and can therefore predict the tunnelling rate as well as its mechanism.

In the nonadiabatic limit, where the Born-Oppenheimer approximation breaks down, Golden-Rule Instanton (GRI) theory is the method of choice. In GRI-theory, the instanton is optimised on the two weakly-coupled diabatic states. A priori however, the distinction into adiabatic and nonadiabatic regimes i.e. strong or weak coupling of the two diabatic states is only valid in limiting cases, and in general a more universal rate theory would be desirable. Towards this aim, we developed the Nonadiabatic Ring-Polymer Instanton (N-RPI) theory.

N-RPI theory bridges the gap between the adiabatic and nonadiabatic limit by constructing the effective electronic surface on which to optimise the tunnelling path from the two coupled diabatic states. It is therefore ensured that the instanton path cannot be dominated by the optimal path residing entirely on either diabatic surface. NRPI-theory hence overcomes pitfalls of other quantum instanton theories that aim for applicability to arbitrary coupling strengths.

In order to demonstrate the predictive power of our newly developed N-RPI theory, we investigate the linear crossing model, which is the most fundamental nonadiabatic curve-crossing system. We show that N-RPI theory successfully predicts the rate for this system in both the adiabatic and the nonadiabatic limit and crucially also for intermediate coupling strengths.