Uncertainty Quantification of Quantum Chemical Methods

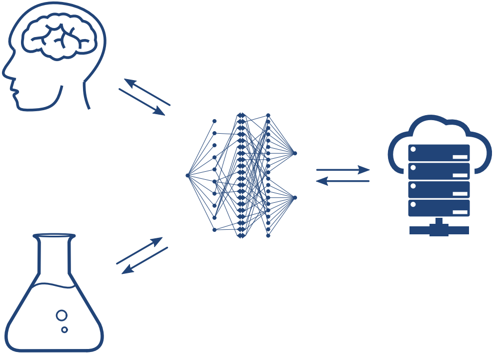
Markus Reiher

Outline

- Introducing context: high-throughput calculations
- General considerations

  M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732
- Selection of a few specific examples

Imagine Unsupervised High-Throughput Calculations such as Reaction Networks



J. Unsleber, M. Reiher, Annu. Rev. Phys. Chem. 71, **2020**, 121

G. N. Simm, A. Vaucher, M. Reiher, J. Phys. Chem. A 123, **2019**, 385

Fully Automated Explorations

- **Invest (cheap) computer time** by massively searching for new intermediates, transition states, elementary steps, side reactions, degradation reactions ...

- Benefits from our new developments: fast semi-empirical methods, new transition-state search engines, stable orbital optimization, ...

  A. C. Vaucher, M. Reiher, *J. Chem. Theory Comput.*, **2018**, 14, 3091–3099.

  A. C. Vaucher, M. Reiher, *J. Chem. Theory Comput.*, **2017**, 13, 1219.

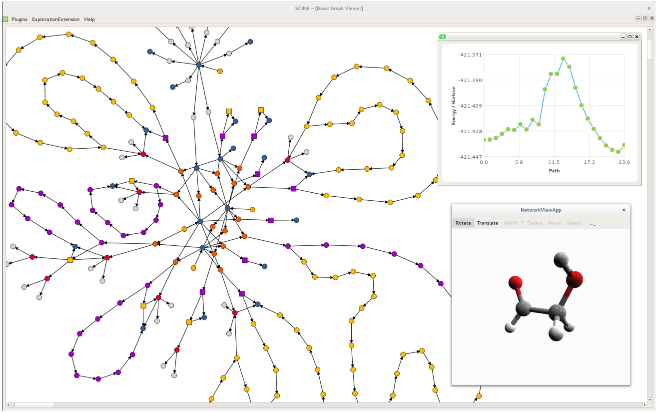
  A. H. Mühlbach, A. C. Vaucher, M. Reiher, *J. Chem. Theory Comput.*, **2016**, 12, 1228.

  T. Husch, M. Reiher, *J. Chem. Theory Comput.*, **2018**, 14, 5169.

  T. Husch, A. C. Vaucher, M. Reiher, *Int. J. Quantum Chem.*, **2018**, e25799

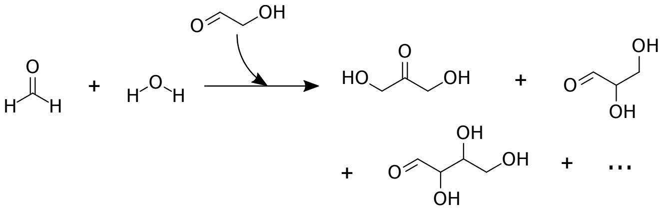
- Big data problem (no manual inspection possible!)

Chemoton: Autonomous Exploration Framework



G. N. Simm, M. Reiher, J. Chem. Theory Comput., *2017 13*, 6108-6119.

Demonstration Example – Formose Reaction



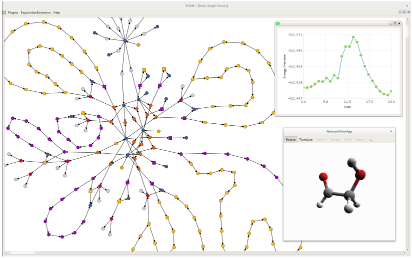
- Prebiotic polymerization reaction
- Initiator (e.g., glycolaldehyde) required
- Autocatalytic mechanism likely

A. Butlerow, *Justus Liebigs Ann. Chem.* 120 (1861) 295; R. Breslow, *Tetrahedron Lett.* 1 (1959) 22.

has been subject of exploration studies: e.g., A. Aspuru-Guzik and co-workers

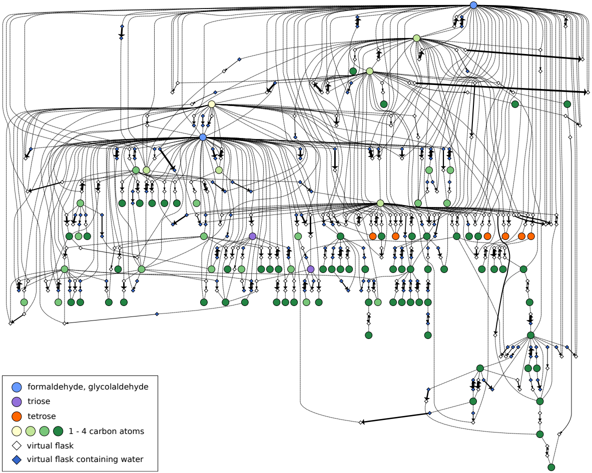
Formose Reaction – Computational Details

- DFT: PBE / double-ζ basis
- Exploration statistics
  - 150 000 calculations
  - 1000 unique molecular configurations
  - 10 000 transition states
- Largest network obtained so far for this reaction — but still small and chemically incomplete



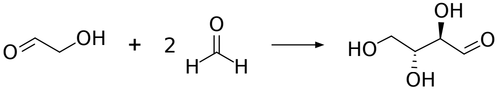
G. N. Simm, M. Reiher, J. Chem. Theory Comput., *2017* *13*, 6108-6119.

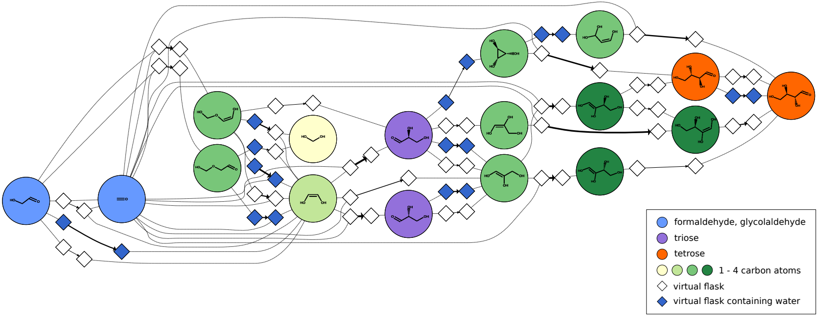
Automated Mechanism Generation – Subnetwork:



G. N. Simm, M. Reiher, J. Chem. Theory Comput., *2017 13*, 6108-6119.

Formation of D-Erythrose





G. N. Simm, M. Reiher, J. Chem. Theory Comput., *2017 13*, 6108-6119.

Reaction Network for Propylene and Dioxygen

Stable and general explorations but suffering from combinatorial explosion

Brute-force exploration with approximative DFT methods (tight binding)

starting from propylene + oxygen

Settings:

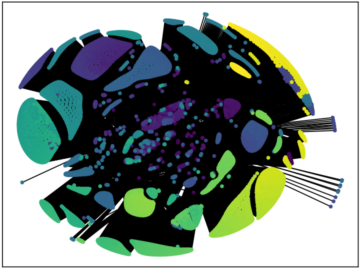
- No conformers, barrier <200 kJ/mol, size limit $C_{10}H_{22}O_7$

Resulting Network

- 4,200 compounds connected by 6,300 reactions

Computational Cost:

- 10^9 individual energy calculations
- 5,000 days of CPU time



Compounds as nodes colored according to their order of discovery from purple to yellow

Steiner M., Reiher M., Top. Catal. **2022**, 65, 6-39

Complex network generated by automated exploration

$\Rightarrow$ now dock to kinetic modeling

... but what about accuracy?

(Recall: energy differences enter an exponential in a rate expression)

Uncertainty Quantification: Overview

- Quantum chemical methods rely on a range of approximations
- The effect of these approximations is usually difficult to assess
- Therefore, the uncertainty of a quantum chemical result is often not known
- However, knowledge of this (systematic) error is important to assess the reliability of any given method

Traditional Approaches to Determine Uncertainty

- Compare to very accurate reference calculations (only affordable for comparatively small systems)

  G. N. Simm, J. Proppe, M. Reiher, *Chimia*, **2017**, *71*, 202.

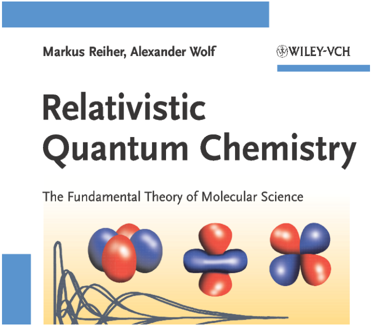
- Rely on benchmark studies
  (not transferable across chemical space)

  T. Weymuth, M. Reiher, *arXiv:2204.06659*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Statement 1: *The theoretical foundations of electronic structure theory are very well established and understood.*


Markus Reiher, Alexander Wolf

WILEY-VCH

Relativistic Quantum Chemistry

The Fundamental Theory of Molecular Science

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Statement 2: *The quantum mechanical Coulomb interaction of two electrons is the crucial challenge for accurate approximations.*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Statement 3: *While the discretization error introduced by orbital and geminal basis sets – as well as any other error that results from the technical implementation of solution procedures – may be systematically reduced, its value in an actual calculation is usually not precisely known.*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Statement 4: *The individual absolute error of a specific quantum chemical result is usually very hard, if not impossible, to assess accurately.*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Observation: *Electronic structure models are affected by some error that is usually not assessed in a specific application.*

Instead one often relies on experience and intuition gained with some approach.

(Most approximations rely on error compensation that leads to more reliable results for relative quantities, but their precise accuracy for a specific case under study is typically not known either.)

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Conclusion 1: *While we have a very good understanding of what approximations in electronic structure theory are feasible and efficient, they are, in all cases, affected by an error that will be unknown for a specific molecular structure under consideration, even if this error is expected to be small for certain approaches.*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Corollary: *Transferability of benchmark results to a specific case under investigation cannot be guaranteed in a rigorous manner.*

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Conclusion 2: *Bayesian uncertainty quantification is a way out of the error-assessment problem, but requires continuous benchmarking – ideally with error assignment for the reference data themselves. This benchmarking needs to be adjusted to the specific systems under study.*

Reference data point calculations must be selected on the basis of confidence intervals taken, for instance, from an underlying machine learning model.

Its accuracy will depend on the measure with which one strides across the parameter space.

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Conclusion 3: *One needs to know how accurate a computed result will be for a specific application.*

Bayesian error estimation can be the key to provide this information, also weeding models of similar type and accuracy, hence reducing the number of models that will be required.

General Remarks on Quantum Chemical Methods

M. Reiher, *Isr. J. Chem.* 62 **2021**, e202100101; arXiv: 2109.03732

Conclusion 4: *Generalist electronic structure models cannot be accurate and fast at the same time (for large molecules or molecular aggregates).*

System-focused models can be made fast and accurate, but their transferability, i.e., their accuracy for related structures, must be monitored because they have not been tested on some arbitrary reference data.

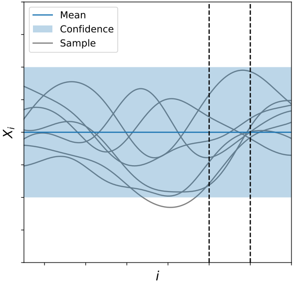
Having system-focused models requires autonomy regarding parametrization and uncertainty quantification, and therefore, efficient automated procedures are needed in order to be practical.

Examples

- Error-controlled reaction network exploration
- Improved semi-classical dispersion interactions

Gaussian Processes (GPs)

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution (a distribution over functions). It is defined by a mean function and a covariance function (kernel).



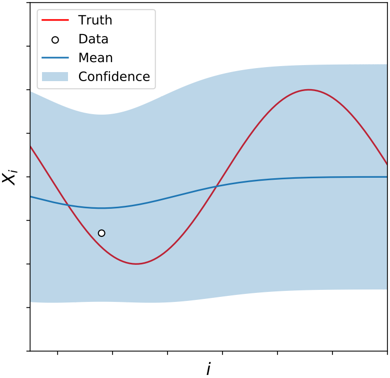
$$E[X_i] = 0$$

$$\mathrm{cov}(X_i, X_j) = \exp\left(-\frac{1}{2\ell^2}|i-j|^2\right) + \delta_{i,j}\sigma^2$$

$$\sigma^2 = 1, \ell^2 = 1$$

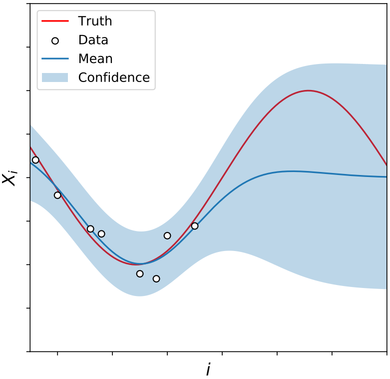
Error Estimation with Gaussian Processes

Training: Optimization of hyperparameters (σ^2 and ℓ^2) with (training) data.



Error Estimation with Gaussian Processes

Training: Optimization of hyperparameters (σ^2 and ℓ^2) with (training) data.



Part 1:

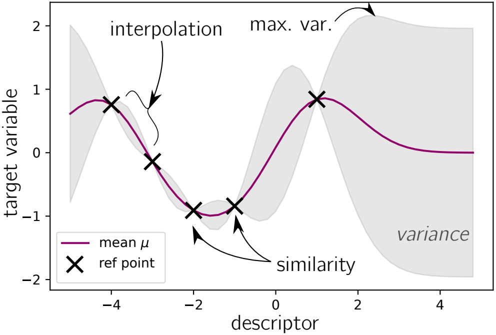
- Error-controlled reaction network exploration

Error-Controlled Exploration of Reaction Networks

- Obtain the observable of interest (*e.g.*, energies) from a statistical model which is fast to evaluate (*e.g.*, via machine-learning using Gaussian processes)
- Obtain error estimate from this statistical model
- If error is above a given threshold, add the corresponding point to the training set and retrain the model

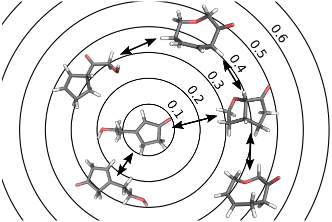
G. N. Simm, M. Reiher, *J. Chem. Theory Comput.*, **2018**, *14*, 5238.

Gaussian Process (GP) regression



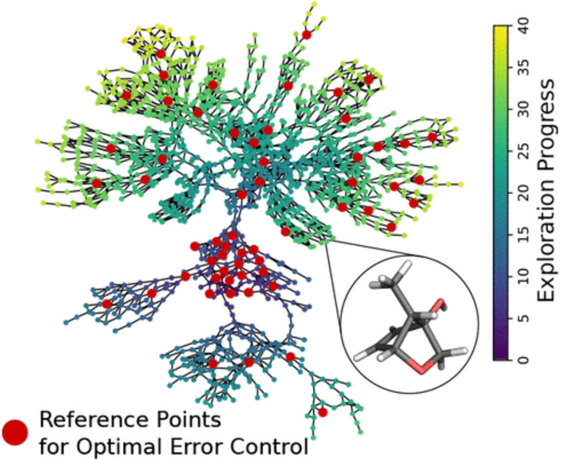
- mean of GP interpolates between reference data
- variance is constructed from similarity measure
- get new reference point where variance is high

Distance Metric in Chemical Space



- Application of ML to chemical systems requires an adequate kernel → distance *d* in chemical space as Input (e.g., SOAP by Csanyi et al.)
- Distance metric introduced by the kernel demonstrated at the example of a reaction subnetwork
- Contour lines represent distance $d(\mathbf{x}, \mathbf{x}')$ between reactant in the center ($\mathbf{x}$) and products ($\mathbf{x}'$).

Error-Controlled Exploration of Reaction Networks



G. N. Simm, M. Reiher, *J. Chem. Theory Comput.*, **2018**, *14*, 5238.

How Obtain Reference Data in Rolling Fashion?

- Explicitly correlated coupled cluster, if single-reference

see work by Tenno, Klopper, Werner, Neese, ...

Black-box multi-configuration SCF: DMRG

shown to work also for transition metal compounds:

K. H. Marti, I. Malkin Ondik, G. Moritz, M. Reiher, J. Chem. Phys. 128 (2008) 014104

see also work by Yanai, Kurashige, Chan, Legeza, Wouters, Van Neck, ...

DMRG with Matrix Product States (MPS) and Matrix Product Operators (MPO)

Our new MPO-based DMRG program: QCMaquis

Download: http://www.reiher.ethz.ch/software/maquis.html

$$|\Psi\rangle = \sum_{\boldsymbol{\sigma}} c_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle \quad \rightarrow \quad |\Psi\rangle = \sum_{\boldsymbol{\sigma}} \sum_{a_1,\ldots,a_{n-1}} M^{\sigma_1}_{1a_1} M^{\sigma_2}_{a_1 a_2} \cdots M^{\sigma_n}_{a_{n-1}1} |\boldsymbol{\sigma}\rangle$$

$$\widehat{\mathcal{W}} = \sum_{\boldsymbol{\sigma},\boldsymbol{\sigma'}} w_{\boldsymbol{\sigma}\boldsymbol{\sigma'}} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma'}| \quad \rightarrow$$

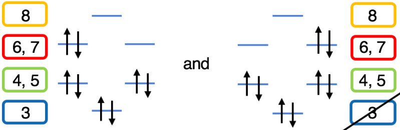
$$\widehat{\mathcal{W}} = \sum_{\boldsymbol{\sigma}\boldsymbol{\sigma'}} \sum_{b_1,\ldots,b_{n-1}} W^{\sigma_1\sigma_1'}_{1b_1} \cdots W^{\sigma_l\sigma_l'}_{b_{l-1}b_l} \cdots W^{\sigma_n\sigma_n'}_{b_{n-1}1} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma'}|$$

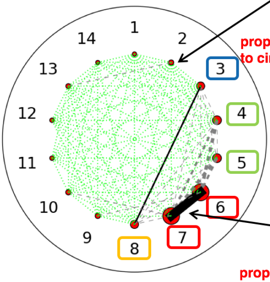
S. Keller, M. Dolfi, M. Troyer, M. Reiher, J. Chem. Phys. 143, 244118 (2015)

How to choose active orbitals in an automated way if multi-configuration calculations must be launched in an automated way?

Entanglement Measures

dioxygen $^1\Delta_g$





proportional to circle area

proportional to thickness of connecting lines

C.J. Stein, M. Reiher, *Chimia* **2017**, *4*, 170.
arXiv: 1702.00450

deviation from a pure state α (—, ↑, ↓, ↑↓)
is measured by **single-orbital entropy**

$$s_i(1) = -\sum_{\alpha=1}^{4} w_{\alpha,i} \ln w_{\alpha,i}$$

$w_{\alpha,i}$: eigenvalue of the reduced one-orbital density matrix for orbital *i*

analogously for two orbitals: **two-orbital entropy**

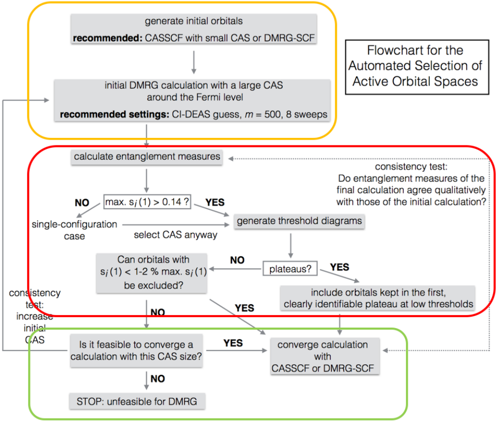
$$s_{ij}(2) = -\sum_{\alpha=1}^{16} w_{\alpha,ij} \ln w_{\alpha,ij}$$

$w_{\alpha,ij}$: eigenvalue of the reduced two-orbital density matrix for orbitals *i, j*

mutual information

$$I_{ij} = \frac{1}{2}\left[s_i(1) + s_j(1) - s_{ij}(2)\right](1 - \delta_{ij})$$

Automated Orbital Selection Algorithm



Flowchart for the Automated Selection of Active Orbital Spaces

generate initial orbitals
recommended: CASSCF with small CAS or DMRG-SCF

initial DMRG calculation with a large CAS around the Fermi level
recommended settings: CI-DEAS guess, *m* = 500, 8 sweeps

calculate entanglement measures

max. $s_i(1) > 0.14$?

NO → single-configuration case → select CAS anyway

YES → generate threshold diagrams

plateaus?

YES → include orbitals kept in the first, clearly identifiable plateau at low thresholds

NO → Can orbitals with $s_i(1) <$ 1-2 % max. $s_i(1)$ be excluded?

YES → converge calculation with CASSCF or DMRG-SCF

NO → Is it feasible to converge a calculation with this CAS size?

YES → converge calculation with CASSCF or DMRG-SCF

NO → STOP: unfeasible for DMRG

consistency test: increase initial CAS

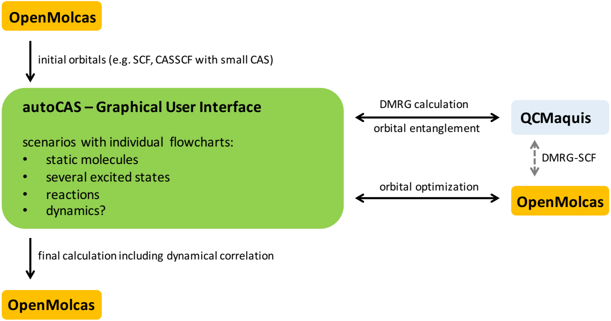
consistency test: Do entanglement measures of the final calculation agree qualitatively with those of the initial calculation?

recipe

1. **partially converged DMRG calculation** with a large number of active orbitals
2. **identify "important" orbitals**
3. **converge the calculation** with only "important" orbitals

Stein, C. J., Reiher, M. *JCTC*, **2016**, *12*, 1760.

Black-Box DMRG Calculations: autoCAS



OpenMolcas

initial orbitals (e.g. SCF, CASSCF with small CAS)

autoCAS – Graphical User Interface

scenarios with individual flowcharts:
- static molecules
- several excited states
- reactions
- dynamics?

DMRG calculation

orbital entanglement

QCMaquis

DMRG-SCF

orbital optimization

OpenMolcas

final calculation including dynamical correlation

OpenMolcas

autoCAS: http://www.reiher.ethz.ch/software/autoCAS

OpenMolcas: https://gitlab.com/Molcas/OpenMolcas

Software: C. J. Stein, M. Reiher, J. Comput. Chem. 40, 2019, 2216

Part 2:

- Improved semi-classical dispersion interactions

Physical descriptor for disperion

From the atom pair-wise DFT-D3(BJ) correction[1]

$$E_{IJ}^{\mathrm{D3(BJ)}} = \sum_{n=6,8} s_n \frac{C_n^{IJ}}{R_{IJ}^n + \left(a_1 \sqrt{C_8^{IJ}/C_6^{IJ}} + a_2\right)^n} \quad \forall I \neq J,\ 0 \text{ else}$$

with R_{IJ}: distance I to J, $C_{\{6,8\}}^{IJ}$: dispersion coefficients, and parameters $s_{\{6,8\}}$, $a_{\{1,2\}}$

we took – in analogy to the Coulomb matrix[2] – the sorted eigenvalues of the resulting matrix as descriptor.[3]

---

[1] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[2] M. Rupp, A. Tkatchenko, K.-R. Müller, *Phys. Rev. Lett.* **2012**, *108*, 058301

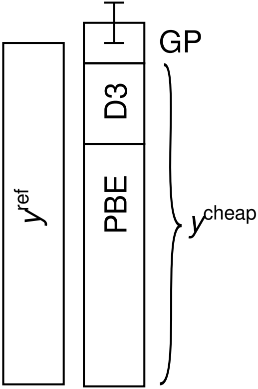
[3] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

The target variable

Our target variable: dispersion energies.[1]

- result of the reference calculation, y^{ref}
- Δ-ML: difference between reference calculation and cheap method, $\Delta y = y^{\text{ref}} - y^{\text{cheap}}$

In our example:
ref $\equiv$ DLPNO-CCSD(T)/CBS
cheap $\equiv$ PBE-D3(BJ)/ma-def2-QZVPP



---

[1] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

Similarity measures

Gaussian kernel:

$$k\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \exp\left(-\frac{1}{2\ell^2}\left(\boldsymbol{x}_i - \boldsymbol{x}_j\right)^2\right) \tag{1}$$

- For two descriptors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ (here, the eigenvalues of the D3 correction above) that are the same, the kernel yields unity.
- Similarity decays exponentially with the tunable length parameter ℓ.

Data sets[1]

Set	#	Description
S13x8	104	dispersion-dominated subset of S66x8
ROTA	1,100	ethyne–pentane dimers; varying relative orientations; centroid distances d_c = 3.5–10 Å
CONF	44	ethyne–pentane dimers; varying relative orientations; pentane conformations; d_c = 5.2 Å
TOTAL	1,248	all molecular reference systems (dimers)

---

[1] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

Electronic-structure calculations[1]

- ORCA 4.0.1. and CP-corrected.
- PBE: ma-def-QZVPP basis set, def2-QZVP auxiliary basis set
- DLPNO-CCSD(T): aug-cc-pVT/QZ basis sets, aug-cc-pVQ/5Z aux. basis sets
- TZ and QZ DLPNO-CCSD(T) energies extrapolated to CBS.
- D3 with Becke–Johnson (BJ) damping.

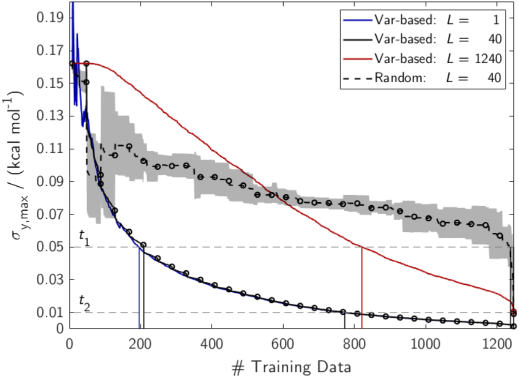
---

[1] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

Batchwise variance based sampling (BVS)

Instead of sampling randomly, *L* new points at maximum variance are added to train set[1]:

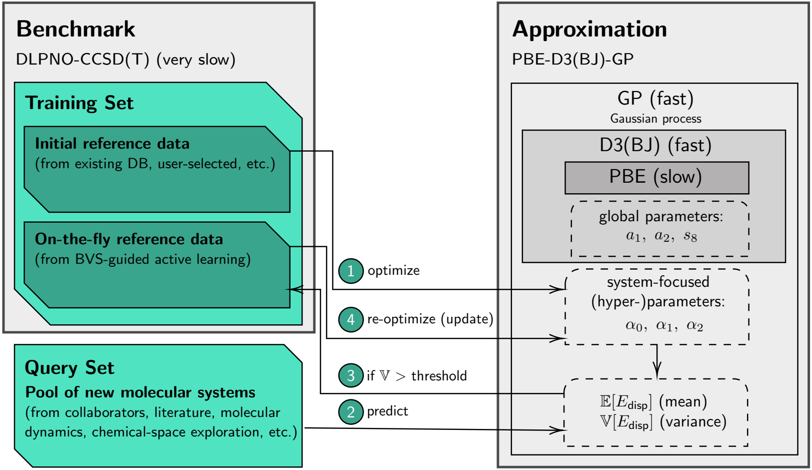
- *L* = 40 as good as *L* = 1 but much less ref. data needed
- Faster than random sampling.



---

[1] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

Workflow overview[1]



Benchmark

DLPNO-CCSD(T) (very slow)

Training Set

Initial reference data
(from existing DB, user-selected, etc.)

On-the-fly reference data
(from BVS-guided active learning)

Query Set

Pool of new molecular systems
(from collaborators, literature, molecular dynamics, chemical-space exploration, etc.)

Approximation

PBE-D3(BJ)-GP

GP (fast)
Gaussian process

D3(BJ) (fast)

PBE (slow)

global parameters:
$a_1,\ a_2,\ s_8$

system-focused (hyper-)parameters:
$\alpha_0,\ \alpha_1,\ \alpha_2$

$\mathbb{E}[E_{\text{disp}}]$ (mean)
$\mathbb{V}[E_{\text{disp}}]$ (variance)

① optimize

④ re-optimize (update)

③ if $\mathbb{V} >$ threshold

② predict

---

[1] J. Proppe, S. Gugler, M. Reiher, *J. Chem. Theory Comput.* **2019**, *15*, 6046.

Related work on uncertainty quantification

- Bootstrapping and jacknifing for physico-chemical models such as those for ^{57}Fe Mössbauer isomer shifts prediction.[1]
- Quantification of parameter dependence of Grimme's semiclassical D3 dispersion correction with bootstrap analysis.[2]
- Δ-machine learning to improve on system-focused self-parametrizing atomistic models. [3], [4]

---

[1] J. Proppe, M. Reiher, *J. Chem. Theory Comput.* **2017**, *13*, 3297.

[2] T. Weymuth, J. Proppe, M. Reiher, *J. Chem. Theory Comput.* **2018**, *14*, 2480.

[3] C. Brunken, M. Reiher, *J. Chem. Theory Comput.* **2020**, *16*, 1646.

[4] C. Brunken, M. Reiher, *J. Chem. Theory Comput.* **2021**, *17*, 3797.

Conclusions

- Traditional ways of benchmarking quantum chemical methods are not necessarily reliable because of a lack of transferability
- GPR delivers error estimates, exploited to identify need for system-specific reference data on the fly
- We obtained **system-focused, self-improving models** equipped with confidence intervals (at examples: reaction network exploration and dispersion interactions)

Acknowledgment

Financial support: ETH Zurich, SNF, FCI, DFG, NCCR Catalysis

- **Current group:** Alberto Baiardi, Moritz Bensberg, Katja Csizi, Robin Feldmann, Nina Glaser, Stefan Gugler, Stephanie Grimmel, Veronika Klasovita, Max Mörchen, Charlotte Müller, Miguel Steiner, Paul Türtscher, Jan Unsleber, Thomas Weymuth