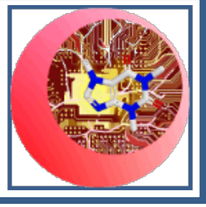
P33

TMCS Theory and Modelling in Chemical Sciences

Efficient Construction of Accurate Interatomic Potentials using Machine Learning[1]

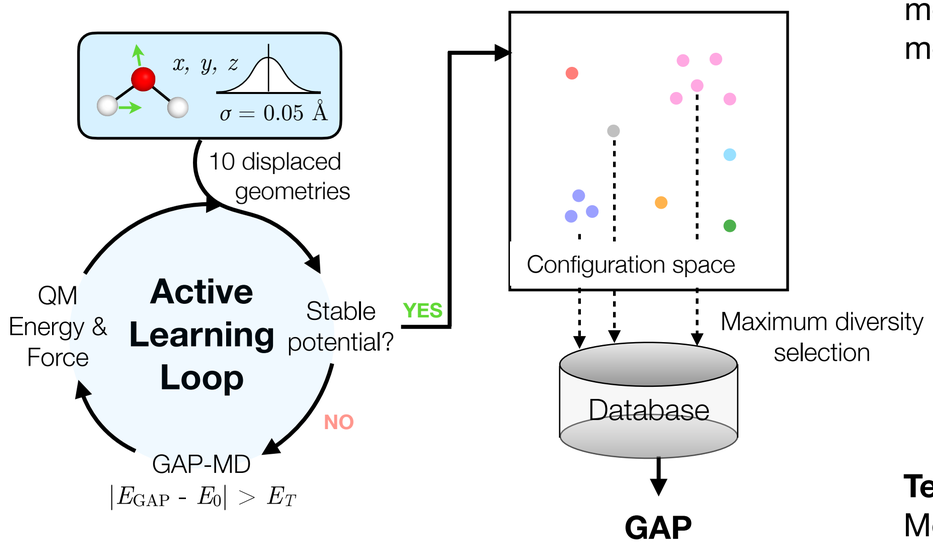
Tom A. Young, Tristan Johnston-Wood, Volker L. Deringer and Fernanda Duarte

UNIVERSITY OF OXFORD



Overview. Accurate molecular simulation requires fast and reactive interatomic potentials. Gaussian Approximation Potentials (GAPs) provide such a framework, and can now be trained quickly and autonomously for condensed and phase molecular systems. By leveraging active learning with backtracking, inter-intramolecular decomposition, hierarchical QM reference methods and maximum-diversity selection, GAPs can be generated efficiently using just 100s of *total* QM calculations.

Method

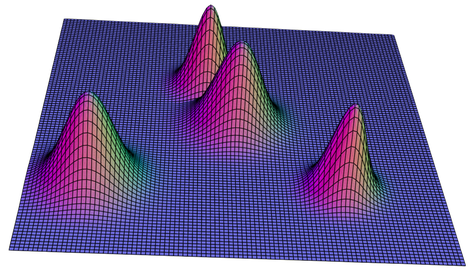


Highlights

- **Systematically improvable**
- **Fully reactive**
- **Automated construction**

`github.com/t-young31/gap-train`

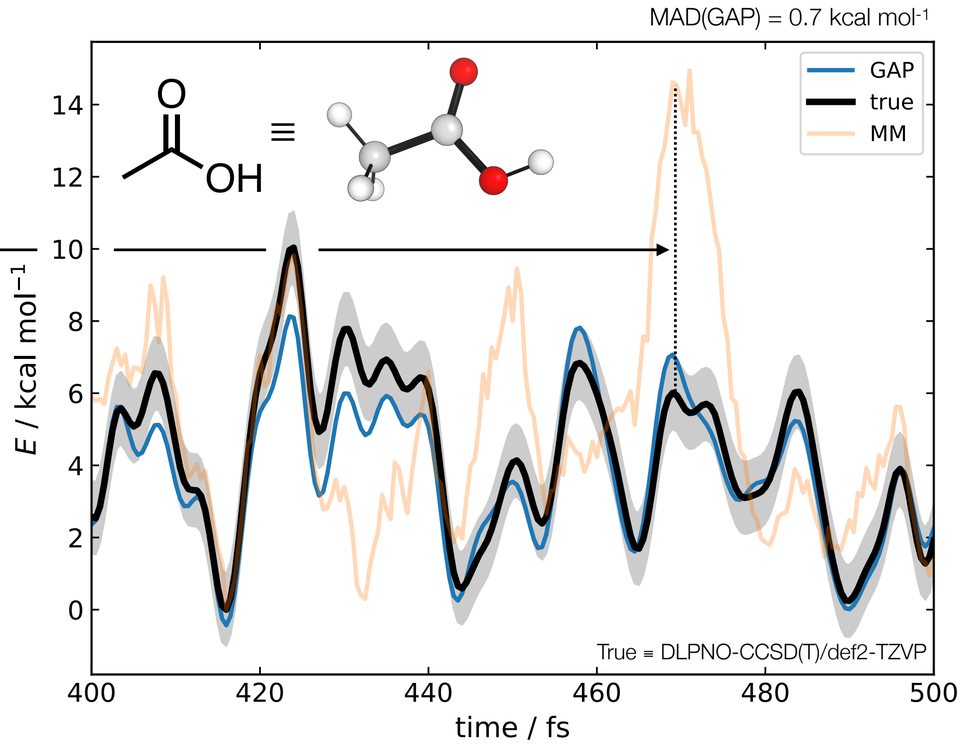
To calculate accurate free energies requires accurate underlying potential energies, however using conventional molecular mechanics, even for simple molecules, errors exceed: **8 kcal mol^{-1}**



SOAP atomic environment descriptor used in a GAP

Tested systems. $H_2O_{(l)}$, $AcOH_{(g)}$, $C_nH_{2n+2(g)}$, $MeCN_{(l)}$, CH_3Cl+Cl, C_2H_4+C_4H_6…

Background. Gaussian Approximation Potentials (GAP) are a machine learning (ML) approach to building interatomic potentials (f_{GAP}: $\mathbf{R} \rightarrow E$). Using reference energies (E) and forces ($-\mathrm{d}E/\mathrm{d}R_x$) a Gaussian Process is *trained* to predict energies at new nuclear configurations ($\mathbf{R}^*$).[2]



[1] T. Young, et. al, *Chem. Sci.*, 2021,**12**, 10944.
https://github.com/t-young31/gap-train

GitHub ← Open source implementation

[2] A. Bartok, et. al, *Phys. Rev. Lett.*, **104**, 136403.

RSC INTEREST GROUP CHEMICAL INFORMATION AND COMPUTER APPLICATIONS

RSC INTEREST GROUP BIOLOGICAL AND MEDICINAL CHEMISTRY SECTOR