Linearization of Moment Tensor Potentials

Andre Lomaka[a] and Toomas Tamm[a]

[a] *Department of Chemistry and Biotechnology, Tallinn University of Technology, Tallinn, Estonia*
toomas.tamm@taltech.ee

Moment Tensor Potentials (MTP) are a method of approximation of molecular potential energy surfaces, with application in machine learning models [1]. In the published formulation, the model is non-linear and requires nonlinear optimization techniques for parametrization.
We propose an alternative relaxed formulation of the original potential energy function where parameter optimization is expressed as a linear sparse approximation problem. This change leads to very large number of highly multi-collinear variables (e.g. 49156 for the water dimer) which brings about the need to introduce heuristic approaches to reduce the set. We used importance ranking of variables and Bayesian information criterion to reduce the number of variables by more than an order of magnitude (e.g. to 4200 for the water dimer).
The proposed formulation was tested on water dimer and water trimer with MTP expansion including up to 4-body terms, with the training set comprising approximately 38000 configurations for water dimer and 11000 configurations for water trimer and their corresponding energies obtained from CCSD(T) calculations extrapolated to basis set limit.
The models provide approximations to the short-range quantum-mechanical supplement of a classical polarizable potential [2] with a RMSE of 0.0393 kcal/mol for the water dimer for a test set of 4235 data points not used in the fitting. For the water trimer, the corresponding figures are 0.1179 kcal/mol and 1235 points in the test set.

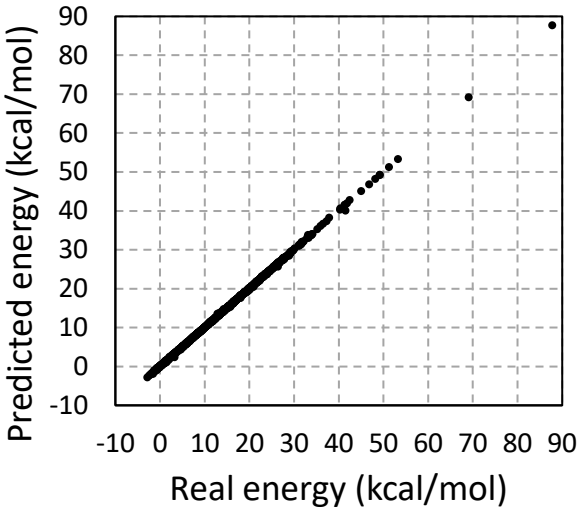

Figure 1: Parity chart for comparing the actual 3B interaction energy values of the test set with those obtained from the model.

References

1. A. V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016).
2. T. T. Nguyen, E. Székely, G. Imbalzano, J. Behler, G. Csányi, M. Ceriotti, A. W. Götz, and F. Paesani, J. Chem. Phys. **148**, 241725 (2018).

This work was supported by Estonian Research Council grant PUT1683.