Electronic Supplementary Information

Facilitating Ab Initio QM/MM Free Energy Simulations by Gaussian Process Regression with Derivative Observations

Ryan Snyder,[†] Bryant Kim,[†] Xiaoliang Pan,[‡] Yihan Shao,*[,‡] and Jingzhi Pu*[,†]
[†]*Department of Chemistry and Chemical Biology, Indiana University-Purdue University Indianapolis, 402 N. Blackford St., Indianapolis, IN 46202, USA* [‡]*Department of Chemistry and Biochemistry, University of Oklahoma, 101 Stephenson Pkwy, Norman, OK 73019, USA*
*(*Correspondence: yihan.shao@ou.edu and jpu@iupui.edu)*

1. Computational Details

1.1. System Setup

We demonstrate the effectiveness of the dual-level QM-GPR/MM method on the solution-phase S_N2 Menshutkin reaction[1] between ammonia and methyl chloride in water, which has been widely used as a paradigm system for developing QM and QM/MM free energy methods.[2-11] Topologies for the solute molecules NH_3 and CH_3Cl were built based on similar residues available in the standard CHARMM topology files. Specifically, the atom types NH3, HC, CT3, HA, and CLA are used for the nitrogen and hydrogens in NH_3, carbon and hydrogens in CH_3, and chlorine, respectively. The van der Waals (vdW) parameters are assigned based on the standard CHARMM22 force field[12] during the initial setup. The solute molecules are solvated in a 40×40×40 $Å^3$ water box composed of modified TIP3P[13] water molecules employing periodic boundary conditions. The SHAKE[14] algorithm is used to constrain the bond lengths in the water molecules during dynamics.

1.2. Potential Energy Surface

The MNDO97[15]-based combined QM/MM module implemented in CHARMM[16] was used to carry out all molecular dynamics (MD) simulations. The QM region consisting of the solute molecules was treated using the Austin Model 1 (AM1) SE method.[17] AI reference potential energies and forces were obtained using the B3LYP functional[18-20] with the 6-31+G(d,p) basis set[21] through the Gaussian16 package[22] interfaced with CHARMM. A set of pair-specific vdW parameters, optimized for the Menshutkin reaction by Gao *et al.* [6] and implemented using the NBFIx facility in CHARMM, are used to model vdW interactions between the QM and MM regions. Long-range electrostatics for MM/MM and QM/MM interactions are treated by the particle mesh Ewald (PME)[23] and QM/MM-PME[24, 25] methods, respectively. In both PME treatments, the κ parameter that represents the width of the Gaussian screening charge distribution is set to 0.34 $Å^{-1}$, and the reciprocal space summations are performed on a 40×40×40 FFT grid, with maximally up to five *k*-vectors included in each Cartesian direction. For the real-space contribution of the QM/MM-PME electrostatics, a switching function available in CHARMM is applied from 12 to 14 Å to smoothly attenuate the real-space QM/MM electrostatic interactions at a cutoff of 14 Å.

1.3. Collective Variables, Restraints, and Free Energy Simulations

We employed the string method in collective variables[26] (CVs) to calculate free energy and find the minimum free energy path (MFEP). To describe the free energy path for the Menshutkin reaction, we used two CVs: the bond-breaking (C-Cl) and bond-forming (N-C) distances. These CVs are harmonically anchored with a force constant of 1000 kcal/mol/Å^2. Configurational sampling takes place at 20 evenly spaced images along each path. For each string iteration, each CV's fluctuations around its harmonic restraint are measured over 20 ps MD simulations, evolving using a 1 fs timestep. The CVs' fluctuations are used to estimate the free energy mean forces. The path is evolved to minimize the mean forces after their along-the-path components are projected out. Additional reparametrization steps are applied to redistribute the images evenly along the evolved path. Once the MFEP is determined, the mean forces are integrated to obtain the free energy profile.

1.4. Gaussian Process Regression Model Training and Deployment

Gaussian Process Regression (GPR) was performed using a customized model built using the TensorFlow[27, 28] based GPflow program.[29] The trained model was deployed using the USER module in CHARMM.

2. Effect of Training Set Size on the Computational Expense of Molecular Dynamics

In principle, the computational complexity of Gaussian Process Regression predictions scales as $O(N)$[30] or $O(N(M+1))$ when including derivative information.[31] When additional internal coordinate transformations are included in kernel constructions for predictions during dynamics, this is not exactly the case (Figure S1). Considering the limited improvements to the prediction errors beyond 80 training points (Figure 1 in the main text) as well as the growing computational expense for making the predictions during dynamics, we decided to use 80 points as the training-set size in our production runs for a balanced combination of prediction accuracy and computational efficiency.

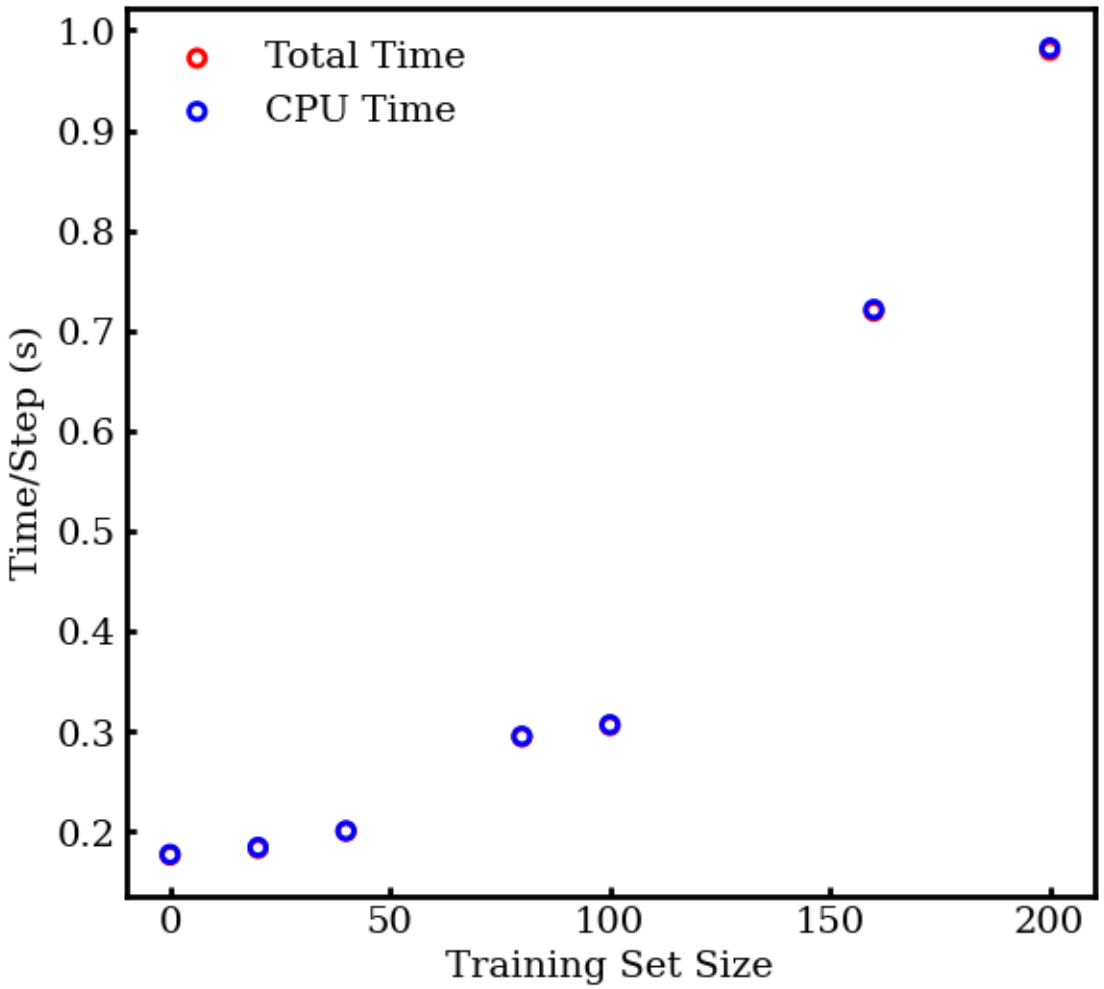

Figure S1. Predictions made during molecular dynamics can make MD steps computationally expensive.

3. Cross Validations of Force Prediction on the Atoms Involved in the Collective Variables (CVs)

Table S1. The root-mean-square errors (RMSEs) of the AM1-GPR/MM models in force prediction on the CV atoms

	RMSE in Force Prediction on N, C, and Cl (in kcal/mol/Å)[a]				
MFEP-level[b]	AM1/MM	AM1-GPR1/MM	AM1-GPR2/MM	AM1-GPR3/MM	AM1-GPR4/MM
AM1/MM	19.4 (19.6)	5.1±6.1 (4.3±2.6)	5.9±15.4	6.1±18.3	6.4±18.4
AM1-GPR1/MM	19.6 (19.6)	6.0±6.8	4.9±7.8 (4.0±2.8)	5.2±9.1	5.3±9.5
AM1-GPR2/MM	19.7 (19.4)	6.3±7.3	5.0±7.7	5.1±7.5 (4.2±3.1)	5.2±8.9
AM1-GPR3/MM	19.9 (20.0)	6.6±7.3	5.2±8.8	5.3±9.7	5.1±8.6 (4.1±3.3)
AM1-GPR4/MM	19.7	6.6±7.0	5.0±7.8	5.2±7.9	5.3±8.2

[a]RMSEs in force prediction are compared against the B3LYP/6-31+G(d,p)/MM benchmark; results for both the testing set and the training set (data in parenthesis) are given when applicable; values following the ± correspond to the average variances of force predictions.
[b]The level for determining the MFEP, along which the samples are collected

4. Additional tests on the noise-hyperparameter treatments

In the extended-kernel GPR formalism used by Meyer and Hauser,[32] a single noise hyperparameter is added to the diagonal elements of the extended covariance kernel matrix $\mathbf{K}_{\text{ext}}(\mathbf{X},\mathbf{X})$ to capture the stochastic feature of the data. In this work, we followed their practice for treating the noise hyperparameter. However, there are also other works that suggest the use of separate noise hyperparameters for describing different levels of noise of an underlying function and its derivatives.[33, 34]

While the predictive variance can be used as a measure of overlap between the new and training samples in their input space, it can also be used to measure the expected mean-square error (EMSE) of the predictive distribution. In the case of noisy observations, this involves adding the noise hyperparameter to the variance:[35, 36]

$$\mathrm{E}[(\mathrm{E}[f(\mathbf{x}^*)]-y)^2] = \mathrm{Var}[f(\mathbf{x}^*)]+\sigma_{\mathrm{n}}^2 \qquad \text{(S1)}$$

When this principle is used to estimate the error for gradient predictions, it would make sense to apply a different noise hyperparameter to the diagonal elements of the second derivative terms in the extended kernel to account for differences in units. For the Menshutkin reaction, we also tested the performance of the QM-GPR/MM model with two separate noise hyperparameters. As shown in Table S2, the inclusion of a second noise hyperparameter offers little change to the errors estimated for force predictions; however, the addition of a second noise hyperparameter increases the estimated errors on energy predictions.

Table S2. The root-mean-square errors (RMSEs) and the square root of the expected mean-square errors [SR-EMSEs, defined as the square root of Eq. (S1)] of the AM1-GPR/MM models using a single noise hyperparameter and separate noise hyperparameters for energy and force predictions. The RMSEs, variances (values following ±), and SR-EMSEs of predictions for both the testing and training (data in parenthesis) sets are given for comparison.

Model	Energy[a]		Solute Atom Force[b]		CV Atom Force[b]	
	RMSE	SR-EMSE	RMSE	SR-EMSE	RMSE	SR-EMSE
Single Noise Hyperparameter	6.3±0.7 (5.6±0.7)	3.6 (3.6)	3.7±5.1 (3.0±2.6)	4.2 (3.9)	5.1±6.1 (4.3±2.6)	4.3 (3.9)
Separate Noise Hyperparameters	7.6±1.4 (6.9±1.3)	7.0 (7.0)	3.6±5.1 (2.9±2.5)	4.0 (3.6)	5.0±6.0 (4.2±2.5)	4.1 (3.6)

[a] in kcal/mol

[b] in kcal/mol/Å

Notably, Solak *et al.*,[37] who introduced the concept of derivative observations in GPR, suggested using a noise hyperparameter for each individual gradient component. Future investigations are therefore needed to identify the best practice for the noise-hyperparameter treatment when GPR with derivative observations is used for facilitating AI/MM free energy simulations.

5. AM1-GPR/MM Improves the Internal Coordinate Force on the Collective Variables

The string method in collective variables (CVs) constructs the free energy profile and optimizes the minimum free energy path (MFEP) using the free energy gradients estimated from the fluctuations of the CVs around a set of harmonic restraints.[26] In our recent work,[3] we showed that directly correcting the internal forces on the CVs along the MFEP can reproduce high-level (i.e., AI/MM) free energy profiles from low-level (i.e., SE/MM) simulations. Here we examine whether the Cartesian-force-based GPR corrections would also lead to the improved internal forces on the CV degrees of freedom essential for obtaining the target-level free energy profile.

Along the AM1/MM MFEP, the overall root-mean-square deviation (RMSD) between B3LYP/6-31+G(d,p)/MM and AM1/MM internal forces, when averaged over both CV bonds, drops from 29.3 to 8.9 kcal/mol/Å, after the GPR1 correction is applied (Figure S2). Specifically, the RMSD in force on the N-C bond drops from 23.1 to 11.5 kcal/mol/Å and that on the C-Cl bond drops from 34.4 to 5.2 kcal/mol/Å.

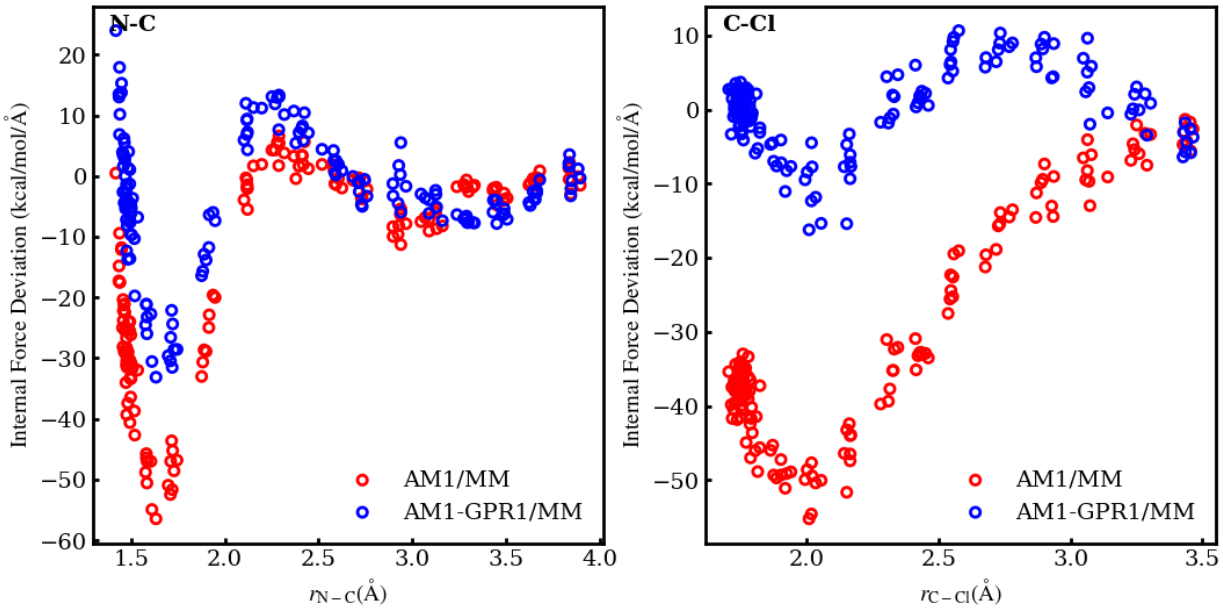

Figure S2. Internal force deviations from the B3LYP/6-31+G(d,p)/MM results before (AM1/MM) and after (AM1-GPR1/MM) the GPR1 corrections are applied along the AM1/MM MFEP. Although offering reasonable descriptions of internal forces in the bond dissociation regions, AM1/MM displays large errors at short bond distances. With the GPR1 corrections, the AM1-GPR1/MM method significantly improves the internal force descriptions on the N-C and C-Cl bonds at short bond distances.

Along the AM1-GPR4/MM MFEP (the converged MFEP), the GPR4 correction reduces the overall RMSD in CV forces from 28.1 to 8.4 kcal/mol/Å (Figure S3). On the N-C and C-Cl bonds, the force deviations are reduced from 24.8 to 11.0 kcal/mol/Å and from 31.0 to 4.7 kcal/mol/Å, respectively.

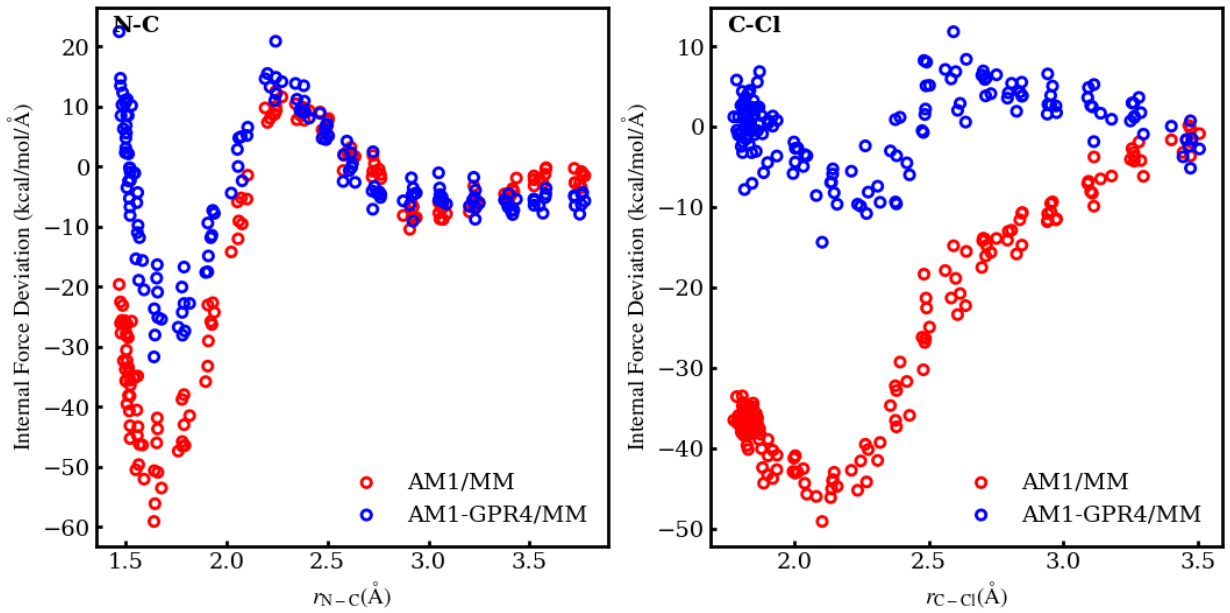

Figure S3. Internal force deviations from the B3LYP/6-31+G(d,p)/MM results before (AM1/MM) and after (AM1-GPR4/MM) the GPR4 corrections are applied along the AM1-GPR4/MM MFEP. Similar to the results based on AM1/MM MFEP shown in Figure S2, the AM1-GPR4/MM method mainly improves the AM1/MM's internal force descriptions on the N-C and C-Cl bonds at short bond distances.

These results show significant improvements on the internal forces. This indicates that explicit predictions of Cartesian atomic forces in our GPR models also improve the description of the CV bonds, which helps rationalize the observed improvements in MFEP and free energy profile.

References

1. N. Menshutkin, *Z. Physik. Chem.*, 1890, **5**, 589-600.
2. B. Kim, Y. Shao and J. Pu, *J. Chem. Theory Comput.*, 2021, **17**, 7682-7695.
3. B. Kim, R. Snyder, M. Nagaraju, Y. Zhou, P. Ojeda-May, S. Keeton, M. Hege, Y. Shao and J. Pu, *J. Chem. Theory Comput.*, 2021, **17**, 4961-4980.
4. J. Wu, L. Shen and W. Yang, *J. Chem. Phys.*, 2017, **147**, 161732.
5. X. Pan, P. Li, J. Ho, J. Pu, Y. Mei and Y. Shao, *Phys. Chem. Chem. Phys.*, 2019, **21**, 20595-20605.
6. J. Gao and X. Xia, *J. Am. Chem. Soc.*, 1993, **115**, 9667-9675.
7. T. N. Truong, T.-T. T. Truong and E. V. Stefanovich, *J. Chem. Phys.*, 1997, **107**, 1881-1889.
8. J. J. Ruiz-Pernia, E. Silla, I. Tunon, S. Marti and V. Moliner, *J. Phys. Chem. B*, 2004, **108**, 8427-8433.
9. O. Acevedo and W. L. Jorgensen, *J. Phys. Chem. B*, 2010, **114**, 8425-8430.
10. H. Nakano and T. Yamamoto, *J. Chem. Theory Comput.*, 2013, **9**, 188-203.
11. X. Pan, J. Yang, R. Van, E. Epifanovsky, J. Ho, J. Huang, J. Pu, Y. Mei, K. Nam and Y. Shao, *J. Chem. Theory Comput.*, 2021, **17**, 5745-5758.
12. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586-3616.
13. E. Neria, S. Fischer and M. Karplus, *J. Chem. Phys.*, 1996, **105**, 1902-1921.
14. J.-P. Ryckaert, G. Ciccotti and H. C. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-337.
15. W. Thiel, *University of Zurich: Zurich, Switzerland*, 1998.
16. B. R. Brooks, C. L. Brooks, III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545-1614.
17. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902-3909.
18. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
19. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
20. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623-11627.
21. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654-3665.
22. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K.

Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, Gaussian, Inc., Wallingford, CT, 2016.

23. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
24. K. Nam, J. Gao and D. M. York, *J. Chem. Theory Comput.*, 2005, **1**, 2-13.
25. K. Nam, *J. Chem. Theory Comput.*, 2014, **10**, 4175-4183.
26. L. Maragliano, A. Fischer, E. Vanden-Eijnden and G. Ciccotti, *J. Chem. Phys.*, 2006, **125**, 024106.
27. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu and X. Zheng, *arXiv*, 2016, arXiv:1603.04467.
28. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, M. Wicke, Y. Yu and X. Zheng, *OSDI'16: Proceedings of 12th USENIX Symposium on Operating Systems Design and Implementation*, 2016, 265-283.
29. A. G. de G. Matthews, M. Van Der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani and J. Hensman, *J. Mach. Learn. Res.*, 2017, **18**, 1299-1304.
30. D.-T. Nguyen, M. Filippone and P. Michiardi, presented in part at the Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, Limassol, Cyprus, 2019.
31. D. Eriksson, K. Dong, E. H. Lee, D. Bindel and A. G. Wilson, *arXiv*, 2018, arXiv:1810.12283.
32. R. Meyer and A. W. Hauser, *J. Chem. Phys.* , 2020, **152**, 084112.
33. D. Born and J. Kästner, *J. Chem. Theory Comput.*, 2021, **17**, 5955-5967.
34. A. Denzel and J. Kästner, *J. Chem. Theory Comput.*, 2020, **16**, 5083-5089.
35. C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, MIT press Cambridge, MA, 2006.
36. A. Lederer, J. Umlauft and S. Hirche, *arXiv*, 2019, arXiv:1906.01404.
37. E. Solak, R. Murray-Smith, W. Leithead, D. Leith and C. Rasmussen, *NIPS*, 2003, **15**, 1033-1040.