Impact of explicit long-range interactions on the accuracy of machine learning interatomic potentials in modeling molecular liquids

Olga Chalykh [1] **Dmitry Korogod** [1 2] **Mikhail Polovinkin** [1 2] **Nikita Rybin** [1 2] **Ivan S. Novikov** [3] **Max Hodapp** [4] **Alexander V. Shapeev** [1 2]

[1]*Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Bolshoy boulevard 30, Moscow, 143026, Russian Federation* [2]*Digital Materials LLC, Odintsovo, Kutuzovskaya str. 4A Moscow region, 143001, Russian Federation* [3]*HSE University, Faculty of Computer Science, Pokrovsky boulevard 11, Moscow, 109028, Russian Federation* [4]*Materials Center Leoben Forschung GmbH (MCL), Austria*. Correspondence to: Olga Chalykh olga.chalykh@skoltech.ru, Alexander V. Shapeev a.shapeev@skoltech.ru.

1. Introduction

Machine learning interatomic potentials (MLIPs) have emerged as a powerful approach for atomistic simulations, offering accuracy comparable to *ab initio* molecular dynamics at a computational cost close to that of (semi-)empirical force fields. However, most contemporary MLIPs rely on a locality approximation, in which interatomic interactions are restricted to local neighborhoods within a finite cutoff radius. This approximation limits their ability to model systems whose structure and dynamics are governed by long-range interactions [1].

To address this limitation, explicit long-range interaction terms can be incorporated into the total energy expression [1, 2]. Prior work has primarily focused on electrostatic and dispersion interactions: electrostatics is typically incorporated via a Coulomb interaction term, while dispersion is often included using correction schemes originally developed for density functional theory (DFT) [3].

When long-range interactions are effectively screened beyond the cutoff radius, local MLIPs may still provide an accurate description even for systems where such interactions significantly influence system behavior [4]. However, the quantitative impact of explicitly incorporating long-range interactions remains under investigation.

Building on our recent publication [5] and preprint [6], we report Moment Tensor Potentials (MTPs) [7] augmented with: (i) D2-style dispersion correction [8] (MTP+D2) for carbon tetrachloride (CCl_4) and toluene, and (ii) explicit electrostatics modeled via fixed point charges (MTP-QRd) [9] for ethylene carbonate (EC) / ethyl methyl carbonate (EMC) binary solvent and $LiPF_6$ solutions in 3EC/7EMC (molar ratio). We provide a systematic assessment of how explicit long-range terms affect the accuracy of MLIPs in modeling dispersion- and electrostatics-dominated molecular liquids.

2. Related work, Methods, and Results

2.1 Related work

To incorporate electrostatics, the simplest approaches assign element-specific fixed point charges, which are applicable to systems without chemically distinct atoms of the same element [9]. More transferable models require learning environment-dependent charges, either from reference charges, dipole moments, or directly from energies and forces [1, 2]. However, non-local charge transfer (e.g., in conjugated π-systems) requires MLIPs with globally dependent charges [10].

For dispersion, Lennard-Jones potentials can be applied on top of local MLIPs [11, 12]. By comparison, utilizing dispersion corrections [3] such as D2 or D3 can provide higher accuracy and can be directly added to the energy expression as they employ precomputed parameters [13, 14, 15, 16]. Dispersion corrections dependent on electronic-structure calculations, such as Tkatchenko–Scheffler (TS) or exchange-dipole moment (XDM), require either machine learning [17, 18] or precomputation followed by element-wise averaging of their parameters [19, 20, 21].

In this work, we use the MTP-QRd and MTP+D2 models as minimal implementations of explicit electrostatics and dispersion corrections, respectively, to systematically assess their impact on the accuracy of MLIPs for EC/EMC mixtures and nonpolar liquids.

2.2 Methodology

This section summarizes the methodology introduced in our recent publication [5] and preprint [6]. To quantify the role of explicit long-range interactions, we compare local MTPs to MTPs augmented with explicit electrostatics (MTP-QRd) or dispersion interactions (MTP+D2). Specifically, we use implementations of MTP+D2 in the MLIP-2 code [22, 5] and MTP-QRd in the MLIP-4 code [23, 9].

For electrostatics-dominated systems, we assess densities and intra- and intermolecular interactions in EC/EMC binary solvents, as well as ionic conductivity in 1M $LiPF_6$ solutions in 3EC/7EMC, using ensembles of three MLIPs to estimate uncertainty arising from random parameter initialization. For dispersion-dominated systems, we examine dimer binding curves, liquid densities, and radial distribution functions (RDFs) of CCl_4 and toluene, using ensembles of five MLIPs.

To determine when explicit long-range terms are required and when increasing the cutoff radius (R_{cut}) of a local MTP is sufficient, we perform comparisons in two regimes: (i) when R_{cut} is insufficient to capture long-range interactions, and (ii) when the R_{cut} is large enough to include the dominant contribution of long-range interactions.

2.3 Results

Our results show that explicit long-range interaction terms are required when the local MTP cutoff is too short to capture long-range effects, whereas the impact at larger cutoffs depends on the system.

For CCl_4, a short cutoff radius R_{cut}=5.5Å leads to unphysical dimer binding curves, which are corrected either by extending R_{cut} to 7.5Å or by incorporating explicit dispersion (Fig. 1 (a)). Furthermore, MTPs with a short cutoff failed to sustain a stable liquid phase in MD simulations, leading to unphysical expansion of the simulation cell. By contrast, both extended-cutoff MTP and MTP+D2 yield stable MD, with density mean absolute errors (MAE) of 16% and 20%, respectively, compared to experiment [24], and nearly identical RDFs (Fig. 1 (c)).

Unlike CCl_4, for toluene the inclusion of explicit dispersion improves model performance compared to MTP with R_{cut}=7.5Å. For MTP+D2, unphysical oscillations in the dimer binding curves are significantly reduced compared to extended-cutoff MTP (Fig. 1 (b)). While liquid density MAE remain comparable (12% and 10% for MTP and MTP+D2, respectively, relative to experiment [25]), the impact of explicit dispersion is evident in the more accurate C7-C7 RDF obtained with MTP+D2 (Fig. 1 (d)).

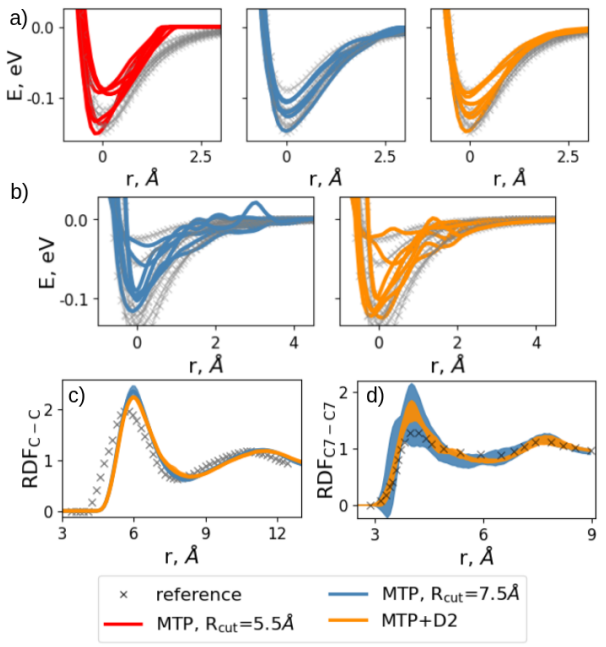

Fig. 1: (a) CCl_4 dimer binding curves calculated using a local MTP with short cutoff R_{cut}, an extended-cutoff MTP, and MTP+D2, compared to PBE-D3 reference data (binding curves minima were shifted to zero); (b) toluene dimer binding curves obtained with extended-cutoff MTP and MTP+D2, compared to PBE-TS; (c) C-C RDFs in liquid CCl_4 obtained from MD simulations in extended-cutoff MTP and MTP+D2, compared to experiment [26]; (d) C7-C7 RDFs in liquid toluene from MD simulations using extended-cutoff MTP and MTP+D2, compared to experiment [27].

In electrostatics-dominated systems, both local MTP and MTP-QRd reproduce densities of EC/EMC binary solvents when R_{cut} is set to 5Å (Fig. 2 (a)). However, MTP-QRd achieves comparable accuracy with fewer machine-learned parameters (Fig. 2 (b)). Both models predict ionic conductivity in $LiPF_6$ solution in 3EC/7EMC with MAE of 11% and 6%, respectively (Fig. 2 (c)), indicating that both local and long-ranged MLIPs can reproduce transport properties with good accuracy in the present simulations.

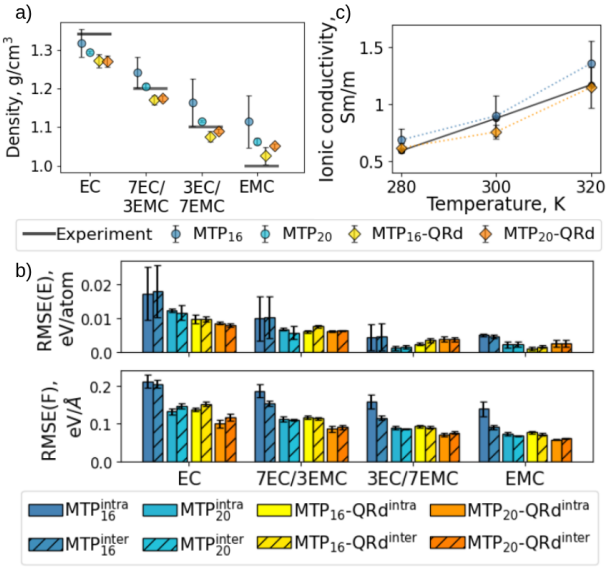

Fig. 2: (a) Densities of EC/EMC mixtures calculated with MTP and MTP-QRd of levels 16 and 20 [7], compared to literature values [4] obtained by interpolating experimental data [28]; (b) root mean squared errors (RMSE) in intra- and intermolecular energies per atom and forces for MTP and MTP-QRd models of levels 16 and 20 relative to PBE-D3 calculations for 7EC/3EMC; (c) temperature dependence of ionic conductivity predicted by MTP and MTP-QRd of level 20 for 1M $LiPF_6$ solution in 3EC/7EMC, compared to experiment [29].

In summary, we benchmarked MTP potentials with and without explicit long-range interactions and showed that explicit dispersion is essential when these interactions are not screened beyond the employed local MLIP cutoff. Moreover, even when the cutoff radius is sufficient to include dominant long-range effects, explicit long-range terms can yield improved accuracy and computational effeciency.

Acknowledgments

This work was supported by the Russian Science Foundation (grant number 23-13-00332, https://rscf.ru/project/23-13-00332/).

References

[1] Dylan M Anstine and Olexandr Isayev. Machine learning interatomic potentials and long-range physics. *The Journal of Physical Chemistry A*, 127(11):2417–2431, 2023.

[2] Jorg Behler. Four generations of high-dimensional neural network potentials. *Chemical Reviews*, 121(16):10037–10072, 2021.

[3] Stefan Grimme, Andreas Hansen, Jan Gerit Brandenburg, and Christoph Bannwarth. Dispersion-corrected mean-field electronic structure methods. *Chemical reviews*, 116(9):5105–5154, 2016.

[4] Ioan-Bogdan Magdău, Daniel J Arismendi-Arrieta, Holly E Smith, Clare P Grey, Kersti Hermansson, and Gábor Csányi. Machine learning force fields for molecular liquids: Ethylene carbonate/ethyl methyl carbonate binary solvent. *npj Computational Materials*, 9(1):146, 2023.

[5] Olga Chalykh, Dmitry Korogod, Ivan S. Novikov, Max Hodapp, Nikita Rybin, and Alexander V. Shapeev. Moment tensor potential and equivariant tensor network potential with explicit dispersion interactions. *The Journal of Chemical Physics*, 163(13), October 2025.

[6] Olga Chalykh, Mikhail Polovinkin, Dmitry Korogod, Nikita Rybin, and Alexander Shapeev. Active learning and explicit electrostatics enable accurate modeling of electrolytes. *arXiv preprint arXiv:2510.03479*, 2025.

[7] Alexander V. Shapeev. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Modeling & Simulation*, 14(3):1153–1173, 2016.

[8] Stefan Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006.

[9] Dmitry Korogod, Olga Chalykh, Max Hodapp, Nikita Rybin, Ivan S. Novikov, and Alexander V. Shapeev. Incorporating coulomb interactions with fixed charges in moment tensor potentials and equivariant tensor network potentials, 2025.

[10] Tsz Wai Ko, Jonas A Finkler, Stefan Goedecker, and Jörg Behler. A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer. *Nature communications*, 12(1):398, 2021.

[11] Patrick Rowe, Gábor Csányi, Dario Alfè, and Angelos Michaelides. Development of a machine learning potential for graphene. *Physical Review B*, 97(5):054303, 2018.

[12] Volker L Deringer, Miguel A Caro, and Gábor Csányi. A general-purpose machine-learning force field for bulk and nanostructured phosphorus. *Nature communications*, 11(1):5461, 2020.

[13] Ning Wang and Shiping Huang. Molecular dynamics study on magnesium hydride nanoclusters with machine-learning interatomic potential. *Physical Review B*, 102(9):094111, 2020.

[14] Penghua Ying and Zheyong Fan. Combining the D3 dispersion correction with the neuroevolution machine-learned potential. *Journal of Physics: Condensed Matter*, 36(12):125901, 2023.

[15] Oliver T Unke and Markus Meuwly. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *Journal of chemical theory and computation*, 15(6):3678–3693, 2019.

[16] Kun Yao, John E Herr, David W Toth, Ryker Mckintyre, and John Parkhill. The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics. *Chemical science*, 9(8):2261–2269, 2018.

[17] Nguyen Thien Phuc Tu, Nazanin Rezajooei, Erin R Johnson, and Christopher N Rowley. A neural network potential with rigorous treatment of long-range dispersion. *Digital Discovery*, 2(3):718–727, 2023.

[18] Nguyen Thien Phuc Tu, Siri Williamson, Erin R Johnson, and Christopher N Rowley. Modeling intermolecular interactions with exchange-hole dipole moment dispersion corrections to neural network potentials. *The Journal of Physical Chemistry B*, 128(35):8290–8302, 2024.

[19] Max Veit, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl, and Gábor Csányi. Equation of state of fluid methane from first principles with machine learning potentials. *Journal of chemical theory and computation*, 15(4):2574–2586, 2019.

[20] Heikki Muhli, Xi Chen, Albert P Bartók, Patricia Hernández-León, Gábor Csányi, Tapio Ala-Nissila, and Miguel A Caro. Machine learning force fields based on local parametrization of dispersion interactions: Application to the phase diagram of C 60. *Physical Review B*, 104(5):054106, 2021.

[21] Rina Ibragimova, Mikhail S Kuklin, Tigany Zarrouk, and Miguel A Caro. Unifying the description of hydrocarbons and hydrogenated carbon materials with a chemically reactive machine learning interatomic potential. *Chemistry of Materials*, 37(3):1094–1110, 2025.

[22] Ivan S Novikov, Konstantin Gubaev, Evgeny V Podryabinkin, and Alexander V Shapeev. The mlip package: moment tensor potentials with mpi and active learning. *Machine Learning: Science and Technology*, 2(2):025002, 2020.

[23] Alexander Shapeev. Mlip-4, 2024.

[24] Nguyen Q Hien, Anthony B Ponter, and Walter Peier. Density and viscosity of carbon tetrachloride solutions containing chlorine. *Journal of Chemical and Engineering Data*, 23(1):54–55, 1978.

[25] Kenneth R Harris. Temperature and density dependence of the viscosity of toluene. *Journal of Chemical & Engineering Data*, 45(5):893–897, 2000.

[26] Lawrence J Lowden and David Chandler. Theory of intermolecular pair correlations for molecular liquids. Applications to the liquids carbon tetrachloride, carbon disulfide, carbon diselenide, and benzene. *The Journal of Chemical Physics*, 61(12):5228–5241, 1974.

[27] Marta Falkowska, Daniel T Bowron, Haresh G Manyar, Christopher Hardacre, and Tristan GA Youngs. Neutron scattering of aromatic and aliphatic liquids. *ChemPhysChem*, 17(13):2043–2055, 2016.

[28] Paul H Johnson. The properties of ethylene carbonate and its use in electrochemical applications a literature review. 1985.

[29] MS Ding, K Xu, SS Zhang, K Amine, GL Henriksen, and TR Jow. Change of conductivity with salt content, solvent composition, and temperature for electrolytes of $lipf_6$ in ethylene carbonate-ethyl methyl carbonate. *Journal of the Electrochemical Society*, 148(10):A1196, 2001.