THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

BuRNN: Buffer Region Neural Network Approach for Polarizable-Embedding Neural Network/Molecular Mechanics Simulations

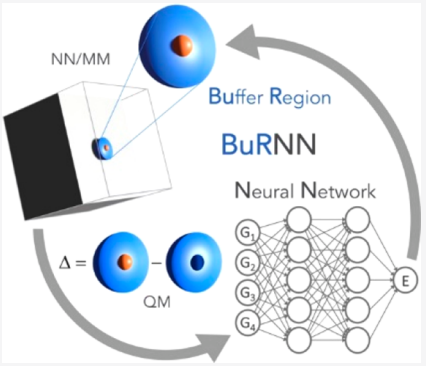
Bettina Lier, Peter Poliak, Philipp Marquetand, Julia Westermayr,* and Chris Oostenbrink*

Cite This: *J. Phys. Chem. Lett.* 2022, 13, 3812−3818

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hybrid quantum mechanics/molecular mechanics (QM/MM) simulations have advanced the field of computational chemistry tremendously. However, they require the partitioning of a system into two different regions that are treated at different levels of theory, which can cause artifacts at the interface. Furthermore, they are still limited by high computational costs of quantum chemical calculations. In this work, we develop the buffer region neural network (BuRNN), an alternative approach to existing QM/MM schemes, which introduces a buffer region that experiences full electronic polarization by the inner QM region to minimize artifacts. The interactions between the QM and the buffer region are described by deep neural networks (NNs), which leads to the high computational efficiency of this hybrid NN/MM scheme while retaining quantum chemical accuracy. We demonstrate the BuRNN approach by performing NN/MM simulations of the hexa-aqua iron complex.



Molecular dynamics (MD) simulations are powerful tools for studying the dynamics of systems consisting of hundreds of thousands of atoms. The energy of the system can be described fully classically by a molecular mechanics (MM) force field, by a quantum mechanical (QM) method, or by a hybrid quantum mechanics/molecular mechanics (QM/MM) technique. The latter approach is very powerful, as it enables an accurate description of a small important part of a system at the appropriate level of quantum chemistry, while the remainder is treated by MM to simulate large system sizes at relevant time scales.[1]

In QM/MM approaches, the electrostatic coupling between the partitioned regions can be treated with different levels of mutual interaction, i.e., embedding schemes.[2−5] Mechanical embedding is the simplest and least accurate approach. Interactions are described via classical point charges only. In contrast, electrostatic embedding is physically better motivated, as the QM system experiences the MM charge distribution being embedded in the QM Hamiltonian. However, QM particles see MM particles as fixed point charges, which neglects polarization in the MM region. To account for polarization effects in the MM region, as well, polarizable force fields can be used.[6,7]

Independently of the scheme, all QM/MM methods are limited by high computational costs of the quantum calculation and issues at the interface, such as overpolarization.[8] Particularly prone to such artifacts are boundaries that cross covalent bonds, although a careful choice of the bond splitting scheme can alleviate them.[4] Furthermore, discrepancies between the forces derived for the QM and MM region can lead to artifical crowding or depletion at the interface, when particles are allowed to change character during a simulation. Several approaches have been proposed to address boundary artifacts either by introducing an intermediate region[9,10] or by restricting the boundary transition.[11]

Alternatively, the whole system can be treated using machine-learned interatomic potentials based on ab initio data.[12] Machine learning (ML) is especially effective for MD simulations as it can learn the relation between a descriptor, i.e., the structure of a system, and a targeted output, i.e., energies and forces, with the accuracy of the reference method, but much lower computational costs. Such ML potentials are available for specific materials at different levels of theory.[13−17] However, universal ML potentials for more complex systems, such as biomolecules, still pose a challenge and are limited by the computational expenses of the reference calculations.[2,12]

Very recently, ML potentials have been combined with QM/MM concepts and were shown to be powerful for, e.g., calculating free energies or transition paths.[18−24] However, these approaches are complicated as ML models need to capture the effects of the environment (MM region) even though only the QM region has to be learned. The introduction of a cutoff, up to which the MM region is included, has emerged as a solution.[21,22,25] One example is FieldSchNet,[20] which circumvents this problem by sampling

Received: March 4, 2022
Accepted: April 18, 2022
Published: April 25, 2022

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818

the environment while keeping the QM region fixed. This model has been shown to be powerful in predicting spectra and chemical reactions with neural networks (NNs) using electrostatic embedding but requires extended sampling. Due to the nature of electrostatics, artifacts at the interfaces are not reduced in the aforementioned approaches.

To circumvent boundary problems and with the aim of avoiding extensive force field parametrizations, we propose an alternative approach. We introduce an additional buffer region that experiences full electronic polarization by the inner QM region. The buffer region is described at the QM and MM levels. Effectively, the interactions with the QM region are calculated entirely at the QM level while the interactions with the MM region are described at the MM level. Within the buffer region, the interactions are a combination of MM interactions and the effect of the QM region on the electronic degrees of freedom of the buffer region. While this approach minimizes the artifacts that arise from mixing two levels of theory, it comes at considerable computational costs as two QM calculations are required. By using ML to describe QM-derived energy surfaces, an elegant solution emerges. In this work, we train deep NNs to directly predict the difference between the two required QM calculations. Thus, this scheme automatically includes the mutual influence of the QM and the buffer region, without the need for additional external forces in the ML setup.[20] The calculation of all relevant energies in a simulation can efficiently be done with a single evaluation of the NN. Due to the fact that interaction energies are often easier to learn with NNs than potential energies, outstanding accuracy can be achieved with mean absolute errors in the range of a few kilojoules per mole. This range is well below the often-desired chemical accuracy of machine learning models and is independent of the size of the inner region. We refer to this NN/MM approach as a buffer region neural network (BuRNN). Although schemes like ONIOM with different regions exist, this approach is, to the best of our knowledge, novel.

The BuRNN approach partitions a system into an inner region $\mathbb{I}$, a buffer region $\mathbb{B}$, and an outer region $\mathbb{O}$ (Figure 1). The total potential energy, V_{tot}, contains the QM energy of the inner and buffer region, $V^{\text{QM}}_{\mathbb{I}+\mathbb{B}}$, and the MM energy coupling of the outer region to all other regions, $V^{\text{MM}}_{\mathbb{O}(\mathbb{I}+\mathbb{B}+\mathbb{O})}$. The buffer region is calculated at both levels of theory. The difference of the two buffer terms, $V^{\text{MM}}_{\mathbb{B}} - V^{\text{QM}}_{\mathbb{B}}$, is included in V_{tot} and helps to smooth the transition between the QM energy of the inner region and the MM energy of the outer region. In addition, artifacts in the electronic degrees of freedom at the outer edge of the buffer region will largely cancel in the difference $V^{\text{QM}}_{\mathbb{I}+\mathbb{B}} - V^{\text{QM}}_{\mathbb{B}}$. Further details are discussed below. Adding all terms for the total potential energy together leads to

$$V_{\text{tot}} = V^{\text{QM}}_{\mathbb{I}+\mathbb{B}} - V^{\text{QM}}_{\mathbb{B}} + V^{\text{MM}}_{\mathbb{B}} + V^{\text{MM}}_{\mathbb{O}(\mathbb{I}+\mathbb{B}+\mathbb{O})} \quad (1)$$

The subscripts for the potential energy denote the calculated region, and the superscripts the method. Even though interactions in QM are not pairwise additive, it is instructive to consider them as hypothetical pairwise interactions within or between regions. They are indicated with a comma-separated subscript. $V^{\text{QM}}_{\mathbb{I}+\mathbb{B}}$ can then be separated into three terms, i.e., the energy that results from interactions within the inner region, $V^{\text{QM}}_{\mathbb{I},\mathbb{I}}$, between the inner and buffer region, $V^{\text{QM}}_{\mathbb{I},\mathbb{B}}$, and within the buffer region, $V^{\text{QM}}_{\mathbb{B},\mathbb{B}}$:

$$V^{\text{QM}}_{\mathbb{I}+\mathbb{B}} = V^{\text{QM}}_{\mathbb{I},\mathbb{I}} + V^{\text{QM}}_{\mathbb{I},\mathbb{B}} + V^{\text{QM}}_{\mathbb{B},\mathbb{B}} \quad (2)$$

The potential energy of the buffer region, $V^{\text{QM}}_{\mathbb{B}}$, is separately calculated at the QM level, as well, and is denoted as $V^{\text{QM(isolated)}}_{\mathbb{B},\mathbb{B}}$. This should emphasize that the inner region is not part of this particular calculation and term. $V^{\text{QM}}_{\mathbb{B},\mathbb{B}}$ as a hypothetical pairwise interaction within the buffer region as used in eq 2, though, includes the influence of the inner region on the buffer. The difference of the two terms can be seen as the polarization of the buffer, which is denoted as $V^{\text{QMpol}}_{\mathbb{B},\mathbb{B}}$:

$$V^{\text{QMpol}}_{\mathbb{B},\mathbb{B}} = V^{\text{QM}}_{\mathbb{B},\mathbb{B}} - V^{\text{QM(isolated)}}_{\mathbb{B},\mathbb{B}} \quad (3)$$

The energy of the buffer is also described at the MM level, $V^{\text{MM}}_{\mathbb{B}} = V^{\text{MM}}_{\mathbb{B},\mathbb{B}}$. If this term agrees with QM exactly, it will cancel with $V^{\text{QM(isolated)}}_{\mathbb{B},\mathbb{B}}$. The outer region with all involved interactions $V^{\text{MM}}_{\mathbb{O},\mathbb{O}}$, $V^{\text{MM}}_{\mathbb{I},\mathbb{O}}$, and $V^{\text{MM}}_{\mathbb{B},\mathbb{O}}$ is treated at the MM level, but with partial charges of the inner and buffer regions derived from the QM calculation of the inner and buffer region together. Hence, it also includes a representation of the polarization of the inner and buffer regions:

$$V^{\text{MM}}_{\mathbb{O}(\mathbb{I}+\mathbb{B}+\mathbb{O})} = V^{\text{MM}}_{\mathbb{I},\mathbb{O}} + V^{\text{MM}}_{\mathbb{B},\mathbb{O}} + V^{\text{MM}}_{\mathbb{O},\mathbb{O}} \quad (4)$$

The total energy in terms of hypothetical interaction energies can then be written as

$$V_{\text{tot}} = V^{\text{QM}}_{\mathbb{I},\mathbb{I}} + V^{\text{QM}}_{\mathbb{I},\mathbb{B}} + V^{\text{MM}}_{\mathbb{I},\mathbb{O}} + V^{\text{QMpol}}_{\mathbb{B},\mathbb{B}} + V^{\text{MM}}_{\mathbb{B},\mathbb{B}} + V^{\text{MM}}_{\mathbb{B},\mathbb{O}} + V^{\text{MM}}_{\mathbb{O},\mathbb{O}} \quad (5)$$

Thus, BuRNN ensures that the interactions within and between neighboring regions are computed at the appropriate

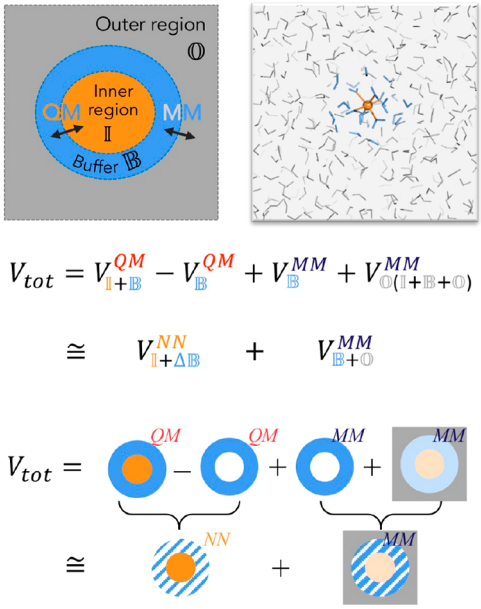


Figure 1. Scheme of the BuRNN approach, which distinguishes three regions: (1) inner region $\mathbb{I}$ (orange), which is described entirely by quantum mechanics (QM), (2) buffer region $\mathbb{B}$ (blue), which is described by both QM and molecular mechanics (MM), and (3) outer region $\mathbb{O}$ (gray), which is described entirely by a classical MM force field.

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818

levels. One of the benefits is that any artifacts in the electronic degrees of freedom will cancel in the difference $V^{QM}_{\mathbb{I}+\mathbb{B}} - V^{QM}_{\mathbb{B}}$, as mentioned above. This is true with the assumption that these artifacts are due to the interface with the outer region and that the relevant polarization of the buffer region predominantly takes place at the interface between the inner and buffer regions. More specifically, artifacts in the QM calculation may arise in the electron density at the outer boundary of the buffer region, which differs from a solvated boundary. However, these artifacts will cancel in the difference as it will be very similar in both QM terms. Any remaining artifacts potentially arising at the interface between the buffer and the outer region are, furthermore, relatively far from the inner region of interest.

The interactions between $\mathbb{I}$ and $\mathbb{O}$ and between $\mathbb{B}$ and $\mathbb{O}$ are computed using a mechanical-embedding scheme with charges assigned from the QM calculation, which is appropriate because of the large distances between the inner and outer regions. The direct electronic influence on the inner region due to the outer region will be relatively small, and the interaction is largely electrostatic. The remaining artifacts may arise from the fact that particles moving from the outer region into the buffer region switch from a mechanically embedded MM interaction to a full QM interaction with the inner region. Also in this case, the interaction is a relatively long distance from the inner region, where the $\mathbb{I}$, $\mathbb{B}$ interaction will be largely electrostatic in nature, such that these artifacts can be expected to be small.

Despite the accuracy and benefit of this scheme, the burden lies in the high computational costs that remain because two computationally expensive QM calculations are required. To overcome this limitation, we describe the first two terms of eq 1 directly using a deep NN:

$$V^{NN}_{\mathbb{I}+\Delta\mathbb{B}} \cong V^{QM}_{\mathbb{I}+\mathbb{B}} - V^{QM}_{\mathbb{B}} \tag{6}$$

which is equal to

$$V^{NN}_{\mathbb{I}+\Delta\mathbb{B}} = V^{QM}_{\mathbb{I},\mathbb{I}} + V^{QM}_{\mathbb{I},\mathbb{B}} + V^{QMpol}_{\mathbb{B},\mathbb{B}} \tag{7}$$

It now becomes clear that the introduction of a buffer region into the BuRNN is akin to a polarizable embedding with the polarization described at the full QM level. Thus, the NN represents the full interactions within the inner region, the interactions between the inner and buffer regions, and the polarization of the buffer region due to the inner region in a single term, $V^{NN}_{\mathbb{I}+\Delta\mathbb{B}}$. The Δ sign is used to emphasize that the BuRNN essentially includes a Δ learning,[26] bringing interactions of the buffer region from the MM to the QM level. The total BuRNN energy can finally be rewritten as

$$V_{tot} = V^{NN}_{\mathbb{I}+\Delta\mathbb{B}} + V^{MM}_{\mathbb{I},\mathbb{O}} + V^{MM}_{\mathbb{B},\mathbb{B}} + V^{MM}_{\mathbb{B},\mathbb{O}} + V^{MM}_{\mathbb{O},\mathbb{O}} \tag{8}$$

All MM terms can be computed classically from a single call to the force field. The workflow of an NN/MM BuRNN simulation is illustrated in Figure 2. The training data set is based on QM calculations and can be generated via sampling from MD simulation snapshots of the targeted system and extended using adaptive sampling.[27] The training set generation and sampling of initial data are explained in detail in sections S1.1 and S1.2 of the Supporting Information. We employ NNs to predict interaction energies, interaction forces, and charges to carry out NN/MM simulations. As NN models, we use SchNet,[28,29] a deep convolutional continuous-filter NN,

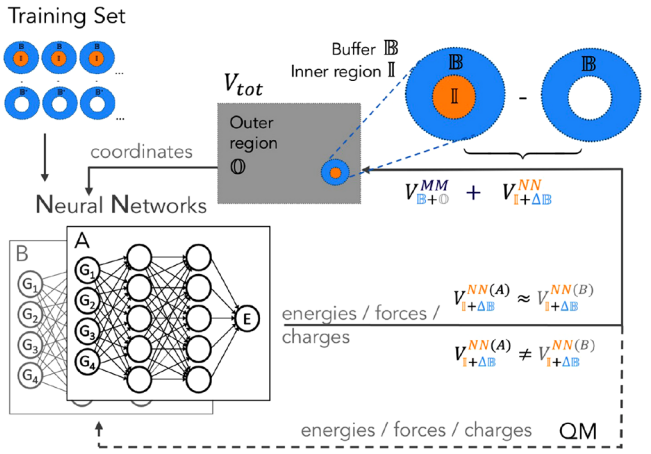


Figure 2. Process of a BuRNN simulation that includes adaptive sampling. At every xth time step during MM, two neural networks (A and B) are compared. When predictions diverge, the training set is expanded by additional QM calculations.

that was adapted to allow for charge fitting. A full description of the NN models, including learning curves, model accuracy, and hyperparameter optimization, can be found in section S1.3. The mean absolute error assessed from five independently trained NN models is 1.7 ± 0.3 kJ/mol for energies, 8.4 ± 0.4 kJ mol^{-1} nm^{-1} for forces, and 0.027 ± 0.001 au for partial charges. Models trained on a larger inner region are comparable in accuracy, as the NN models are always trained on the whole system, i.e., the inner and buffer regions, which have >50 atoms. For all of the outputs, these are very small errors, well below the chemical accuracy defined as 1 kcal/mol in recent (machine learning) studies.[26,30] We have implemented the BuRNN approach in the GROMOS simulation software[31] (see section S1.3). Importantly, we generate the training set and conduct MD simulations by applying two initially trained NNs, A and B. In adaptive sampling, their prediction differences can be used to assess the reliability of NN models during the MD simulation. Whenever the NN predictions for interaction energies are similar, i.e., when $V^{NN(A)}_{\mathbb{I}+\Delta\mathbb{B}} \approx V^{NN(B)}_{\mathbb{I}+\Delta\mathbb{B}}$, predictions are deemed accurate and the simulation is continued. If predictions start to diverge from each other and exceed a predefined threshold, additional reference QM calculations ($V^{QM}_{\mathbb{I}+\mathbb{B}}$and $V^{QM}_{\mathbb{B}}$) are performed for relevant configurations and added to the training set. In this work, we carried out four rounds of adaptive sampling, i.e., until dynamics could be run up to 1 ns without further interruptions. By replacing QM calculations with NNs during MD simulations, the BuRNN reduces the computational costs considerably and enables long time scales while retaining high accuracy.

Here, we demonstrate the use of the BuRNN for the hexa-aqua iron $[Fe(H_2O)_6]^{3+}$ complex in water as a model system. This system has the advantage of being relatively simple for testing our approach, but the classical description of transition metal interactions is notoriously difficult, which makes it a good use case of the BuRNN for simulating this system. Especially challenging is the coordinative bond between Fe and O as it is somewhat between a covalent bond and an ionic bond and is often addressed with specialized force fields.[32,33]

In our test case, the Fe^{3+} ion comprises the inner region, and water molecules up to 0.5 nm are treated as the buffer region,

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818

which roughly accounts for the first two solvation shells, where the first solvation shell is expected to be formed by the hexacoordinated waters. During extensive MD simulations (10 ns), the water molecules are freely diffusing between the buffer and outer regions, smoothly switching interactions between the NN (QM) level and the MM level of theory. BuRNN simulations are validated using experimental data and are compared to QM/MM simulations of the $Fe(H_2O)_6^{3+}$ complex (QM region) in classical SPC water and two distinct fully classical descriptions. In addition, we performed BuRNN simulations using a larger inner region that additionally comprises the first solvation shell.

To validate the method, we first look at the Fe−O radial distribution functions (RDF) $g(r)$ in Figure 3a. All simulations

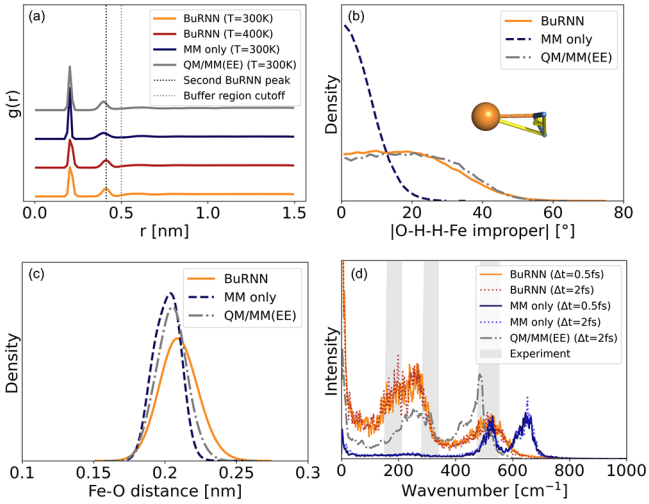


Figure 3. Coordination of Fe^{3+} by water molecules with BuRNN simulations and when using MM only. (a) Radial distribution function for BuRNN at temperatures of 300 and 400 K, with MM only and a QM/MM simulation using electrostatic embedding (EE). The dashed lines indicate the second BuRNN peak and the cutoff used to define the buffer region. (b) Probability distribution of the O−H−H−Fe improper dihedral. (c) Distribution of the Fe−O distance. (d) Power spectrum of the Fe−O coordinative bond for different simulations. Experimental data were taken from refs 50−52.

show a distinct peak at ∼0.20 nm, corresponding to the first coordinative solvation shell of mostly six water molecules, which is slightly narrower and more pronounced in the classical description (blue curve). In contrast, the second solvation shell is more pronounced in the BuRNN (orange curve) and corresponds to an average of 12.7 water molecules. It shows a maximum at ∼0.41 nm, while the MM simulation shows a broader peak with a maximum at 0.40 nm. Simulations with a larger inner region yield almost identical results (Figure S4). In the QM/MM simulation with electrostatic embedding, the second shell also has a maximum at 0.40 nm (gray curve), while a QM/MM simulation with mechanical embedding leads to a maximum at 0.41 nm (Figure S4), as in the BuRNN. Experimentally, it was found at 0.415 nm and comprises 12 water molecules, hence agreeing well with our simulations.[34] We have also performed simulations using the 12−6−4 Lennard-Jones potential[35] and the SPC/E water model[36] and found that the $g(r)$ shows an additional peak at 0.31 nm, representing one additional molecule pushing into the first solvation shell (Figure S4). The transition at 0.5 nm in the RDF obtained with the BuRNN is smooth and does not show any artifacts. This is remarkable as the buffer region ends and the water molecules beyond this distance interact completely according to a pure MM description.

To investigate the robustness of the BuRNN, we performed MD at different temperatures (Figure S5) and show the RDF obtained at 400 K in Figure 3a. As one can see, there is a slight smoothing between 0.5 and 1.0 nm due to the increased level of thermal motion, but the BuRNN simulation remains stable. The two NN models deviate on average by 0.39 ± 0.02 kJ/mol.

In addition, we sought to investigate the propensity of the BuRNN to describe water exchange. Hence, we used umbrella sampling[37] to pull a water molecule away from the complex and observed the spontaneous exchange of this water molecule with another (see Supporting Movie S1). The energy predictions and MD simulations are stable during this process. In the regular simulations, the hexacoordination is stably maintained. Water molecules in the second solvation shell (within the buffer region) readily exchange with water molecules from the outer region. All water molecules (786 molecules) visit the buffer region at least once during the simulation, with an average lifetime of 14.4 ps. This agrees with estimates from NMR experiments that determine a lifetime that is below their resolution limit of 100 ps.[38] We further computed the self-diffusion rate for BuRNN and MM only simulations and found values of 0.98×10^{-5} and 0.92×10^{-5} cm^2/ps, respectively. Both approaches overestimate the diffusion constant compared to experimental estimates of $0.55-0.68 \times 10^{-5}$ cm^2/ps,[38−41] in line with the observation that bulk SPC (simple point charge) water has a diffusion constant that is slightly too large.[42]

Figure 3b shows the distribution of the O−H−H−Fe improper dihedral angles defining the co-planarity of the iron and a water molecule. A value of 0°, which is predominant in pure MM simulations, implies that the water molecule and the Fe^{3+} ion are in the same plane. Larger values as observed for the BuRNN with a mean angle of 19.3° and for QM/MM simulations (mean angle of 20.3°) indicate a more tetrahedral arrangement in which the iron interacts with the lone pairs on the oxygen. For comparison, a BP86-D3/def2-TZVP/COSMO estimate for this angle in $[Fe(H_2O)_{18}]^{3+}$ lies at 16°.[43]

MD simulations are further compared by the geometries visited during the simulations. Figures 3c shows radial distances between the Fe and O that agree well with the range of experimental estimates for the Fe−O bond lengths of 0.199−0.210 nm.[34,44−50] O−Fe−O angles are almost identical among the BuRNN, QM/MM, and "MM only" and reflect angles expected for an octahedral arrangement (peaking at around 90° and 175°). Figure 3d shows that there are clear differences for the frequencies at which the Fe−O bonds vibrate, implying that the Fe−O interaction is indeed not captured well by a purely classical description. In the QM/MM simulations and when using the BuRNN, the vibrations take place at lower frequencies and are in better agreement with experimental bands observed at ∼180, ∼310, and ∼500 cm^{-1}.[50−52] The frequencies obtained with quantum chemistry are better aligned with experiment and the BuRNN than with pure MM (Figure S4b), while those obtained from 12−6−4 Lennard-Jones potential simulations are even higher than those observed with the simple MM only approach (Figure S4).

In this work, we have introduced the BuRNN, a buffered region neural network NN/MM scheme, as an alternative to QM/MM simulations that experiences full electronic polar-

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818

ization by the inner QM region. The BuRNN minimizes artifacts at the interface between regions by ensuring that interactions that go over boundaries are treated at a consistent level of theory. Inconsistencies at the edges of the buffer (i) can be expected to cancel in the difference between two QM terms and (ii) are far removed from the inner region. These advantages come at the cost of an additional QM calculation, which is elegantly solved by training NNs directly on the energy difference. A single evaluation of the NN is required to evaluate the energies, and a second NN is used to derive charges for full mutual polarization. The BuRNN allows fast hybrid NN/MM simulations and has the advantage of being applicable to any system and usable with any molecular ML model.

We have demonstrated the use of the BuRNN by realistic simulations of hexa-aqua iron in water. This shows that it can be applied for metal−ligand interactions without the need for additional force field parameters. The good agreement and high stability of BuRNN for long time-scale MD simulations, including external perturbation, such as changing temperature or forces that lead to water exchange, make our method very promising for future application in the simulation of more complex systems.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.2c00654.

Computational details for the quantum calculations, training set generation, training and validation of the neural networks and simulation settings, and further analyses of simulations with alternative settings (QM/MM with mechanical embedding, classical with a 12−6−4 Lennard-Jones potential, BuRNN with alternative charge distributions, BuRNN at higher temperatures, and BuRNN with a larger inner region) (PDF)

Supporting Movie S1 (MP4)

Transparent Peer Review report available (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Julia Westermayr − *Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K.;* orcid.org/0000-0002-6531-0742; Email: julia.westermayr@warwick.ac.uk

Chris Oostenbrink − *Institute for Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, University of Natural Resources and Life Sciences, Vienna, 1190 Vienna, Austria;* orcid.org/0000-0002-4232-2556; Email: chris.oostenbrink@boku.ac.at

Authors

Bettina Lier − *Institute for Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, University of Natural Resources and Life Sciences, Vienna, 1190 Vienna, Austria;* orcid.org/0000-0002-8032-0084

Peter Poliak − *Institute for Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, University of Natural Resources and Life Sciences, Vienna, 1190 Vienna, Austria; Department of Chemical Physics, Institute of Physical Chemistry and Chemical Physics, Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, 812 37 Bratislava, Slovakia*

Philipp Marquetand − *Institute of Theoretical Chemistry, University of Vienna, 1090 Vienna, Austria;* orcid.org/0000-0002-8711-1533

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.2c00654

Author Contributions

B.L. and P.P. contributed equally to this work. B.L.: conceptualization, investigation, data curation, funding acquisition, software, review and editing, and visualization. P.P.: investigation, data curation, methodology, software, writing of the original draft, and review and editing. P.M.: methodology, supervision, and review and editing. J.W.: investigation, data curation, methodology, writing of the original draft, review and editing, and visualization. C.O.: conceptualization, investigation, data curation, methodology, writing of the original draft, review and editing, supervision, and resources.

Funding

Open Access is funded by the Austrian Science Fund (FWF).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Michael Gastegger for discussion regarding the training process of ML models. B.L. is a recipient of a DOC Fellowship of the Austrian Academy of Sciences (ÖAW) at the Institute for Molecular Modeling and Simulation at the University of Natural Resources and Life Sciences, Vienna (Grant 25743). The project was further supported by the Austrian Science Fund, FWF doctoral program BioToP (W1224), and Erwin-Schrödinger Project J 4522-N (J.W.). The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC).

■ REFERENCES

(1) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.

(2) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM simulations: What are some of the "burning issues"? *J. Phys. Chem. B* **2021**, *125*, 689−702.

(3) Brunk, E.; Rothlisberger, U. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chem. Rev.* **2015**, *115*, 6217−6263.

(4) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Ang. Chem. Int. Ed.* **2009**, *48*, 1198−1229.

(5) Gao, J.; Truhlar, D. G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.

(6) Ganguly, A.; Boulanger, E.; Thiel, W. Importance of MM polarization in QM/MM studies of enzymatic reactions: Assessment of the QM/MM drude oscillator model. *J. Chem. Theory Comput.* **2017**, *13*, 2954−2961.

(7) Loco, D.; Lagardère, L.; Caprasecca, S.; Lipparini, F.; Mennucci, B.; Piquemal, J.-P. HybridQ QM/MM molecular dynamics with amoeba polarizable embedding. *J. Chem. Theory Comput.* **2017**, *13*, 4025−4033.

(8) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: The future gold standard for complex (bio)systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433−14448.

(9) Bernstein, N.; Várnai, C.; Solt, I.; Winfield, S. A.; Payne, M. C.; Simon, I.; Fuxreiter, M.; Csányi, G. QM/MM simulation of liquid water with an adaptive quantum region. *Phys. Chem. Chem. Phys.* **2012**, *14*, 646−656.
(10) Watanabe, H. C.; Cui, Q. Quantitative analysis of QM/MM boundary artifacts and correction in adaptive QM/MM simulations. *J. Chem. Theory Comput.* **2019**, *15*, 3917−3928.
(11) Kirchhoff, B.; Jónsson, E. Ö.; Dohn, A. O.; Jacob, T.; Jónsson, H. Elastic collision based dynamic partitioning scheme for hybrid simulations. *J. Chem. Theory Comput.* **2021**, *17*, 5863−5875.
(12) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine learning force fields. *Chem. Rev.* **2021**, *121*, 10142.
(13) Deringer, V. L.; Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **2017**, *95*, 094203.
(14) Deringer, V. L.; Bernstein, N.; Csányi, G.; Ben Mahmoud, C.; Ceriotti, M.; Wilson, M.; Drabold, D. A.; Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature* **2021**, *589*, 59−64.
(15) Smith, J. S.; Isayev, O.; Roitberg, A. E. Ani-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, *8*, 3192−3203.
(16) Westermayr, J.; Marquetand, P. Machine learning for electronically excited states of molecules. *Chem. Rev.* **2021**, *121*, 9873−9926.
(17) Westermayr, J.; Gastegger, M.; Marquetand, P. Combining SchNet and SHARC: The SchNarc machine learning approach for excited-state dynamics. *J. Phys. Chem. Lett.* **2020**, *11*, 3828−3834.
(18) Zhang, Y.-J.; Khorshidi, A.; Kastlunger, G.; Peterson, A. A. The potential for machine learning in hybrid QM/MM calculations. *J. Chem. Phys.* **2018**, *148*, 241740.
(19) Wu, J.; Shen, L.; Yang, W. Internal force corrections with machine learning for quantum mechanics/molecular mechanics simulations. *J. Chem. Phys.* **2017**, *147*, 161732.
(20) Gastegger, M.; Schütt, K. T.; Müller, K.-R. Machine learning of solvent effects on molecular spectra and reactions. *Chem. Sci.* **2021**, *12*, 11473−11483.
(21) Pan, X.; Yang, J.; Van, R.; Epifanovsky, E.; Ho, J.; Huang, J.; Pu, J.; Mei, Y.; Nam, K.; Shao, Y. Machine-learning-assisted free energy simulation of solution-phase and enzyme reactions. *J. Chem. Theory Comput.* **2021**, *17*, 5745−5758.
(22) Böselt, L.; Thürlemann, M.; Riniker, S. Machine learning in QM/MM molecular dynamics simulations of condensed-phase systems. *J. Chem. Theory Comput.* **2021**, *17*, 2641−2658.
(23) Shen, L.; Wu, J.; Yang, W. Multiscale quantum mechanics/molecular mechanics simulations with neural networks. *J. Chem. Theory Comput.* **2016**, *12*, 4934−4946.
(24) Shen, L.; Yang, W. Molecular dynamics simulations with quantum mechanics/molecular mechanics and adaptive neural networks. *J. Chem. Theory Comput.* **2018**, *14*, 1442−1455.
(25) Zeng, J.; Giese, T. J.; Ekesan, Ş.; York, D. M. Development of range-corrected deep learning potentials for fast, accurate quantum mechanical/molecular mechanical simulations of chemical reactions in solution. *J. Chem. Theory Comput.* **2021**, *17*, 6993−7009.
(26) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big data meets quantum chemistry approximations: The Δ-machine learning approach. *J. Chem. Theory Comput.* **2015**, *11*, 2087−2096.
(27) Behler, J. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem.* **2015**, *115*, 1032−1050.
(28) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet − a deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148*, 241722.
(29) Schütt, K. T.; Kessel, P.; Gastegger, M.; Nicoli, K. A.; Tkatchenko, A.; Müller, K. R. SchNetPack: A deep learning toolbox for atomistic systems. *J. Chem. Theory Comput.* **2019**, *15*, 448−455.
(30) Friesner, R. A. Ab initio quantum chemistry: Methodology and applications. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6648−6653.
(31) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. Architecture, implementation and parallelisation of the gromos software for biomolecular simulation. *Comput. Phys. Commun.* **2012**, *183*, 890−903.
(32) Xia, M.; Chai, Z.; Wang, D. Polarizable and non-polarizable force field representations of ferric cation and validations. *J. Phys. Chem. B* **2017**, *121*, 5718−5729.
(33) Li, P.; Merz, K. M. Metal ion modeling using classical mechanics. *Chem. Rev.* **2017**, *117*, 1564−1686.
(34) Caminiti, R.; Magini, M. An x-ray diffraction study on the first and the second hydration shell of the Fe(III) ion in nitrate solutions. *Chem. Phys. Lett.* **1979**, *61*, 40.
(35) Li, P.; Song, L. F.; Merz, K. M. Parameterization of highly charged metal ions using the 12−6-4 LJ-type nonbonded model in explicit water. *J. Phys. Chem. B* **2015**, *119*, 883−895.
(36) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(37) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(38) Herdman, G. J.; Salmon, P. S. Dynamics of water protons in concentrated gallium(3+), aluminum(3+), iron(3+) and dysprosium-(3+) aqueous solutions: A study using incoherent quasi-elastic neutron scattering. *J. Am. Chem. Soc.* **1991**, *113*, 2930−2939.
(39) Amira, S.; Spångberg, D.; Probst, M.; Hermansson, K. Molecular dynamics simulation of Fe2+(aq) and Fe3+(aq). *J. Phys. Chem. B* **2004**, *108*, 496−502.
(40) Lide, D. R.; Kehiaian, H. V. *CRC handbook of thermophysical and thermochemical data*; CRC Press: Boca Raton, FL, 1994.
(41) Buffle, J.; Zhang, Z.; Startchev, K. Metal flux and dynamic speciation at (bio)interfaces. Part I: Critical evaluation and compilation of physicochemical parameters for complexes with simple ligands and fulvic/humic substances. *Environ. Sci. Technol.* **2007**, *41*, 7609−7620.
(42) Glättli, A.; Daura, X.; van Gunsteren, W. F. Derivation of an improved simple point charge model for liquid water: SPC/A and SPC/L. *J. Chem. Phys.* **2002**, *116*, 9811−9828.
(43) Radoń, M. Benchmarking quantum chemistry methods for spin-state energetics of iron complexes against quantitative experimental data. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4854−4870.
(44) Magini, M.; Caminiti, R. On the structure of highly concentrated iron(III) salt solutions. *J. Inorg. Nucl. Chem.* **1977**, *39*, 91−94.
(45) Sham, T. K.; Hastings, J. B.; Perlman, M. L. Structure and dynamic behavior of transition-metal ions in aqueous solution: An EXAFS study of electron-exchange reactions. *J. Am. Chem. Soc.* **1980**, *102*, 5904−5906.
(46) Piccaluga, G. X-ray diffraction study of the coordination of Fe(III) in a highly hydrolyzed solution of iron(III) chloride. *Z. Naturforsch. A* **1982**, *37*, 154−157.
(47) Apted, M. J.; Waychunas, G. A.; Brown, G. E. Structure and specification of iron complexes in aqueous solutions determined by X-ray absorption spectroscopy. *Geochim. Cosmochim. Acta* **1985**, *49*, 2081−2089.
(48) Asakura, K.; Nomura, M.; Kuroda, H. Fe k-edge XANES and EXAFS of the X-ray absorption spectra of $FeCl_3$ aqueous solutions. A structural study of the solute, iron(III) chloro complexes. *Bull. Chem. Soc. Jpn.* **1985**, *58*, 1543−1550.
(49) Herdman, G. J.; Neilson, G. W. Ferric ion (Fe(III)) coordination in concentrated aqueous electrolyte solutions. *J. Phys.: Cond. Matter* **1992**, *4*, 627−638.
(50) Ohtaki, H.; Radnai, T. Structure and dynamics of hydrated ions. *Chem. Rev.* **1993**, *93*, 1157−1204.
(51) Böhm, F.; Sharma, V.; Schwaab, G.; Havenith, M. The low frequency modes of solvated ions and ion pairs in aqueous electrolyte solutions: Iron(II) and iron(III) chloride. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19582−19591.
(52) Mink, J.; Németh, C.; Hajba, L.; Sandström, M.; Goggin, P. L. Infrared and Raman spectroscopic and theoretical studies of hexaaqua

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818

metal ions in aqueous solution. *J. Mol. Struct.* **2003**, *661−662*, 141−151.

https://doi.org/10.1021/acs.jpclett.2c00654
J. Phys. Chem. Lett. 2022, 13, 3812−3818