Transfer Learning Lithium and Electrolyte Potential Energy Surfaces from Pure and Hybrid DFT

Anonymous Author(s)
Affiliation
Address
email

Abstract

One of the most important problems in rational design of batteries is predicting the properties of the Solid Electrolyte Interphase, which (for a metallic anode) is the part of the battery where metallic and non-metallic components come into contact. However, there is a fundamental problem with predicting the properties of such a mixed material: the two components are best simulated with incompatible levels of density functional theory. Pure functionals perform well for metallic properties, while hybrid or long-range-corrected density functionals perform better for molecular properties and reaction barriers. We demonstrate a simple method to obviate this conflict by training a machine learning potential energy surface using both levels of theory via transfer learning. We further show that the resulting model is more accurate than models trained individually to these levels of theory, allowing more accurate property prediction and potentially faster materials discovery.

1 Introduction

The Solid Electrolyte Interphase (SEI) is a key part of lithium-ion battery design which has proved resistant both to experimental characterization and to accurate simulation[1]. It is difficult to measure the properties of the SEI, such as electronic and ionic conductivity, and therefore difficult to gather enough data to predict how its properties will change with respect to electrode or electrolyte composition.

Quantum mechanics simulations can allow us to calculate materials properties that are difficult to characterize experimentally. For example, Kuai et al[2] simulate the boundary between lithium metal and an electrolyte, as in a battery with a lithium metal anode, over short timescales using Density Functional Theory (DFT). DFT is the method of choice for quantum mechanics simulation of materials, providing excellent coverage of chemical space, reasonable accuracy, and tolerable cost (at least for small systems). But the SEI is too large, and forms over too long a timescale, to be convenient to describe with DFT. More fundamentally, the SEI is also plagued by a discrepancy between the types of DFT appropriate for accurate simulation of its different components.

A large number of DFT methods, known as "functionals", have been developed based on different approximations and parameterization (for further details, see Goerigk et al[3]). DFT functionals without other quantum mechanical calculations added are known as "pure" DFT. Pure DFT functionals such as the popular Perdew–Burke-Ernzerhof (PBE)[4] functional are computationally efficient and are known to accurately predict the properties of many metals, including lithium[5]. However, pure DFT produces less accurate results for organic systems, and in particular, can dramatically underestimate reaction barriers[3].

Hybrid DFT improves on pure DFT by adding an additional quantum mechanical property, Hartree-Fock exchange, thus reducing errors for many classes of systems. Long-range corrected (LRC) hybrid

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

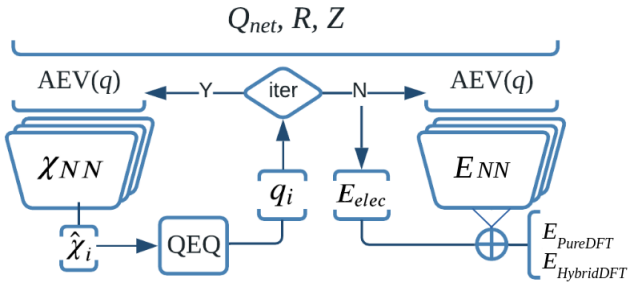

Figure 1: Diagram of transfer learning architecture.

functionals such as ωB97X-D3(BJ)[6] achieve particularly high accuracy[3]. However, hybrid DFT is much more costly for periodic systems than pure DFT, making periodic hybrid DFT rare aside from computationally cheaper "screened" variants[7]. Even more problematic for SEI simulation, Hartree-Fock theory introduces poor convergence and non-physical effects when applied to metals[8]. Because of the limitations of both pure and hybrid DFT, there is no broadly satisfactory method to study organic reactions on a metallic surface with high accuracy.

Machine-learning models for potential energy functions [9] have shown promise in simulating electrolytes[10] and metals[11, 12] at much lower cost than either pure or hybrid DFT. Furthermore, machine learned potential energy surfaces have been trained to output multiple target tasks including different types of DFT[13]. Therefore, it is natural to implement a combined machine learning potential energy surface which can cover both Li metal and common electrolytes, and also the interactions between these components which produce the SEI.

2 Methods

We begin from the dataset of Dajnowicz et al, which is composed of non-periodic organic electrolyte clusters labeled using the LRC hybrid functional (ωB97X-D3(BJ)/def2-TZVPD).[10] We supplement this dataset with calculations at the same level of theory including metallic lithium clusters, organic reactions of common electrolytes, and lithium-organic complexes. In addition, we add periodic bulk lithium metal data labeled using the pure DFT functional PBE-D3[4, 14].

The model architecture we use is QRNN (charge recursive neural network)[15] modified to allow transfer learning in a hard parameter sharing paradigm, as shown in Fig. 1. Each energy network E_{NN} shares parameters for the first N-1 layers, creating a shared latent space, followed by a linear output layer which decodes the latent space into separate predictions. Therefore, the output of each energy network E_{NN} is not one energy value, but one per level of theory to be predicted, as shown in Eqn 1.

$$E_j = \sum_{i}^{N_{atoms}} \mathrm{NN}_{Z_i j}^{QRNN}(G_i^{AEV}, q_i, G_i^{qR}) + E_{disp} + E_{coul}(q) \tag{1}$$

where E_j is the energy for label j, pure or hybrid DFT in this case, $NN_{Z_i j}^{QRNN}$ is output head j of the energy network for chemical element i, E_{disp} is the empirical dispersion correction of the ωB97X-D functional[16, 17] and E_{coul} is a truncated Coulomb energy which decays smoothly at short range[18]. Input G_i^{AEV} is the Atomic Environment Vector (AEV), a rotationally-invariant set of descriptors of the environment of each atom within a radius of 5.2 Angstrom[18], q_i is the atomic partial charge predicted as in QRNN (no transfer learning is used for charge prediction), and G_i^{qR} is the charge-weighted radial AEV of Jacobson et al[15].

For each training batch, the loss function is simply the sum of the energy, atomic force, and charge loss terms over all training points which have the corresponding labels. In any given batch, the loss terms are concentrated towards the hybrid DFT, since the hybrid DFT labeled points are 99% of the dataset and the batches are drawn uniformly at random from the remaining points in each epoch.

The QRNN charge model neural networks χ_{NN} are trained using only the cluster datapoints, which in this dataset means only the hybrid DFT data. Rather than using DFT dipoles as a reference, as

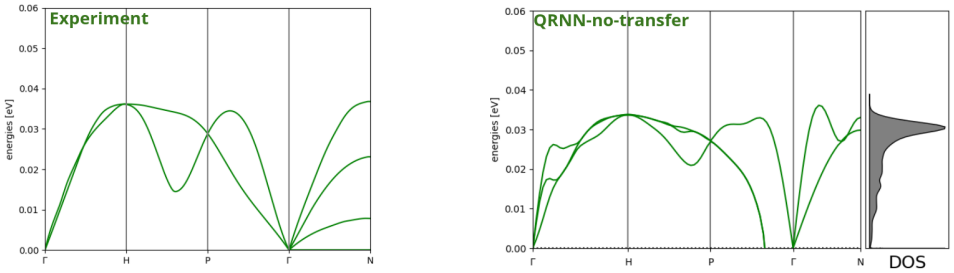

Figure 2: Experimental phonon spectrum from Smith et al, 1968[22] (left) vs QRNN trained with hybrid DFT without transfer learning (right). The QRNN phonon spectrum without transfer learning deviates very significantly from the experimental result. DOS indicates the predicted density of states.

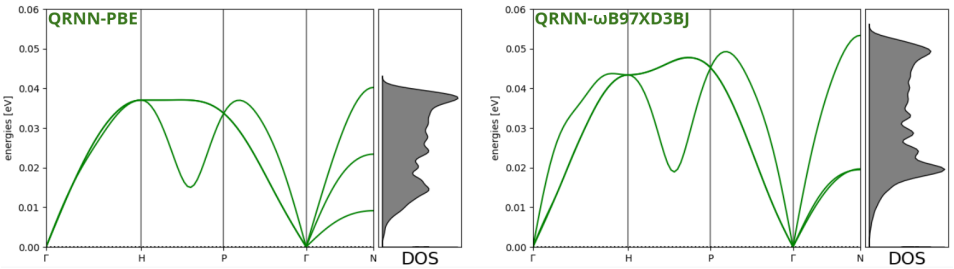

Figure 3: Phonon spectrum with pure DFT head (left) and hybrid DFT head (right) of a transfer-learned QRNN model. The pure DFT head is nearly perfect as compared to the experimental result in Fig. 2; the hybrid DFT head, while not as accurate, is far better than that without transfer learning.

in Dajnowicz et al[10], we used reference atomic partial charges, calculating them with the method Geometries, Frequencies, and Non-covalent 2 Extended Tight Binding (GFN2-xTB)[19]. For large systems, especially systems with high symmetry which have a very small dipole moment, atomic partial charge labels give much more information than the system dipole. Training the charge model neural networks χ_{NN} using periodic data introduces difficulties, since the dipole moment of a periodic system is difficult to define[20], and no current GFN2-xTB implementation provides fully consistent support for periodic systems[21]. The subject of generating consistent charge training using both periodic and non-periodic data merits further study. However, since in our dataset the periodic data is only 1% of the full dataset, the problem is not yet a practical concern in this work.

3 Results

A sensitive test of a potential energy surface for a metal is to calculate its phonon spectrum[11]. Training only to the hybrid DFT dataset produces a qualitatively inaccurate phonon spectrum for lithium, as shown in Fig. 2. By comparison, transfer learning with both pure and hybrid DFT produces reasonable phonon spectra for both prediction heads of the model. The pure-DFT head is more accurate, as expected, since the pure DFT dataset contains only periodic metallic lithium. However, the ωB97X-D3BJ/def2-TZVPD head is still approximately correct, and vastly improved over the equivalent head trained without transfer learning.

Organic reaction barrier tests show a similar trend. As shown in Table 1, the hybrid head of the transfer learning model shows equivalent or slightly improved performance compared to the equivalent model without transfer learning, on a test set composed of Li radical-initiated organic reaction barriers (see Appendix). For the transfer learning model, only the hybrid DFT head gives reasonable results, since the pure DFT dataset is composed of Li only: the transfer learning technique does not produce good results in the complete absence of data for the given output head, which is expected. Even more effective transfer learning might occur if the overlap between the pure and hybrid parts of the dataset were greater: currently no exactly-equivalent geometries exist in the dataset, since the pure DFT is uniformly periodic and the hybrid DFT is uniformly non-periodic. The final entry in Table 1 is the

Table 1: Reaction barrier test set RMSE by model (vs ωB97X-D3BJ/def2-TZVPD reference)

Model	RMSE (kcal/mol)
QRNN, no transfer learning (hybrid DFT head)	4.9
QRNN, transfer learning (hybrid DFT head)	4.8
QRNN, transfer learning (pure DFT head)	35.3
PBE-D3/def2-TZVPD	11.1

equivalent RMSE for the pure DFT functional PBE-D3, demonstrating the relative weakness of this functional for electrolyte reaction barriers.

The model is able to simulate the dynamics of Li metal in contact with ethylene carbonate at 300K, as shown in Fig. 4. However, we do not yet have an appropriate reference comparison to determine the accuracy of the combined simulation, besides extrapolation from our test sets. Comparison to DFT dynamics, such as the work of Kuai et al[2], is costly (especially with hybrid functionals) and not necessarily reliable. Another option is to create non-periodic test sets labeled with hybrid DFT (or higher-level quantum mechanics calculations), allowing comparison to accurate references at lower cost, but with the risk of missing larger-scale effects.

4 Conclusions

We have shown that a simple transfer learning architecture allows effective transfer between pure and hybrid DFT, making simulations of metallic systems and organic reactions more effective. In particular, we show that the phonon spectrum of metallic lithium for such a model is greatly improved over the non-periodic hybrid DFT equivalent, without degrading the accuracy for organic reaction barriers.

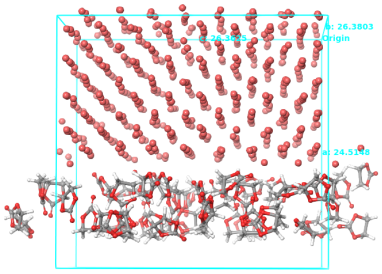

Figure 4: Molecular dynamics snapshot of a lithium metal slab in contact with ethylene carbonate electrolyte at 300K. Pink=Li, Red=O, Gray=C, White=H.

Like other machine learning potential energy surfaces, this model does not have the broad chemical applicability of DFT: functional groups outside the training set (such as phosphates) will likely give larger errors. Likewise, given unusual geometric inputs such as atoms very close together, the energies and forces returned may be very inaccurate, thus ruining the predictive power of a given simulation from that point onward. These unstable regimes are easy to detect in practice because they produce obvious spikes in system energies. Further work is needed in this direction - for example, some models use added atomic repulsion terms to make out-of-bounds geometry samples less likely[13].

Even with potential energy models much faster than DFT, such as this one, all-atom SEI simulations cannot reproduce the full dynamics of SEI formation, since the time scale of hours is too long. Instead, we expect all-atom SEI simulations to aid materials design by predicting structural or material properties which can be used as inputs for conventional or AI-guided design methods.

Being able to simulate millions of timesteps per day on a GPU for systems of 1000 or more atoms, our method promises to be able to study the formulation-property relationships of interest for materials design much faster than pure DFT or hybrid DFT, while avoiding the accuracy pitfalls of training to just a single DFT method. Outside of our demonstrated area of application, lithium batteries, this technique should be applicable to other metal and nonmetal systems which commonly arise in materials science, such as the study of metal corrosion and the design of corrosion-resistant coatings. We anticipate that transfer learning potential energy models will be an important tool for future materials design.

References

[1] Aiping Wang, Sanket Kadam, Hong Li, Siqi Shi, and Yue Qi. Review on modeling of the anode solid electrolyte interphase (sei) for lithium-ion batteries. *npj Computational Materials*, 4(1):1–26, 2018.

[2] Dacheng Kuai and Perla B Balbuena. Solvent degradation and polymerization in the li-metal battery: Organic-phase formation in solid-electrolyte interphases. *ACS Applied Materials & Interfaces*, 14(2): 2817–2824, 2022.

[3] Lars Goerigk, Andreas Hansen, Christoph Bauer, Stephan Ehrlich, Asim Najibi, and Stefan Grimme. A look at the density functional theory zoo with the advanced gmtkn55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics*, 19(48): 32184–32215, 2017.

[4] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[5] Verena A. Neufeld, Hong-Zhou Ye, and Timothy C. Berkelbach. Ground-state properties of metallic solids from ab initio coupled-cluster theory. *The Journal of Physical Chemistry Letters*, 13(32):7497–7503, 2022. doi: 10.1021/acs.jpclett.2c01828.

[6] Narbe Mardirossian and Martin Head-Gordon. ωb97x-v: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Physical Chemistry Chemical Physics*, 16(21):9904–9924, 2014.

[7] Aliaksandr V Krukau, Oleg A Vydrov, Artur F Izmaylov, and Gustavo E Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of chemical physics*, 125(22):224106, 2006.

[8] Thomas M Henderson, Artur F Izmaylov, Gustavo E Scuseria, and Andreas Savin. The importance of middle-range hartree-fock-type exchange for hybrid density functionals. *The Journal of chemical physics*, 127(22):221103, 2007.

[9] Oliver T Unke, Stefan Chmiela, Huziel E Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T Schutt, Alexandre Tkatchenko, and Klaus-Robert Muller. Machine learning force fields. *Chemical Reviews*, 121 (16):10142–10186, 2021.

[10] Steven Dajnowicz, Garvit Agarwal, James M Stevenson, Leif D Jacobson, Farhad Ramezanghorbani, Karl Leswing, Richard A Friesner, Mathew D Halls, and Robert Abel. High-dimensional neural network potential for liquid electrolyte simulations. *The Journal of Physical Chemistry B*, 2022.

[11] Nongnuch Artrith, Alexander Urban, and Gerbrand Ceder. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Physical Review B*, 96(1):014112, 2017.

[12] Akira Takahashi, Atsuto Seko, and Isao Tanaka. Linearized machine-learning interatomic potentials for non-magnetic elemental metals: Limitation of pairwise descriptors and trend of predictive power. *The Journal of chemical physics*, 148(23):234106, 2018.

[13] So Takamoto, Chikashi Shinagawa, Daisuke Motoki, Kosuke Nakago, Wenwen Li, Iori Kurata, Taku Watanabe, Yoshihiro Yayama, Hiroki Iriguchi, Yusuke Asano, et al. Towards universal neural network potential for material discovery applicable to arbitrary combination of 45 elements. *Nature Communications*, 13(1):1–11, 2022.

[14] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15):154104, 2010.

[15] Leif D Jacobson, James M Stevenson, Farhad Ramezanghorbani, Delaram Ghoreishi, Karl Leswing, Edward D Harder, and Robert Abel. Transferable neural network potential energy surfaces for closed-shell organic molecules: Extension to ions. *Journal of Chemical Theory and Computation*, 18(4):2354–2366, 2022.

[16] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.*, 10(44):6615–6620, November 2008.

[17] Stefan Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27(15):1787–1799, November 2006.

[18] Farhad Ramezanghorbani. *Developing Machine Learning Models to Enhance Applicability of Neural Network Potentials in Drug Discovery*. PhD thesis, U. Florida, 2020. URL https://login.lp.hscl.ufl.edu/login?url=https://www-proquest-com.lp.hscl.ufl.edu/dissertations-theses/developing-machine-learning-models-enhance/docview/2533302375/se-2?accountid=10920. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2021-06-07.

[19] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB — an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of chemical theory and computation*, 15(3):1652–1671, 2019.

[20] Lukas Schreder and Sandra Luber. Local approaches for electric dipole moments in periodic systems and their application to real-time time-dependent density functional theory. *The Journal of Chemical Physics*, 155(13):134116, 2021.

[21] Leif Jacobson and Sebastian Ehlert. Energy of supercell is inconsistent with converged kpoint mesh using dftb+. https://github.com/tblite/tblite/issues/5, 2021.

[22] HG Smith, G Dolling, RM Nicklow, PR Vijayaraghavan, and MK Wilkinson. Phonon dispersion curves in lithium. In *Neutron Inelastic Scattering Vol. I. Proceedings of a Symposium on Neutron Inelastic Scattering*, pages "149–155", 1968.

[23] Satoshi Maeda, Yu Harabuchi, Makito Takagi, Tetsuya Taketsugu, and Keiji Morokuma. Artificial force induced reaction (afir) method for exploring quantum chemical potential energy surfaces. *The Chemical Record*, 16(5):2232–2248, 2016.

[24] Daniel S Levine. Automatic determination of chemical reaction mechanisms with neural networks. *Second Workshop on Machine Learning and the Physical Sciences, NeurIPS 2019*, 2019.

[25] Justin M Turney, Andrew C Simmonett, Robert M Parrish, Edward G Hohenstein, Francesco A Evangelista, Justin T Fermann, Benjamin J Mintz, Lori A Burns, Jeremiah J Wilke, Micah L Abrams, et al. Psi4: an open-source ab initio electronic structure program. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(4):556–565, 2012.

[26] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.

[27] Halls, Mathew D. Schrodinger materials science suite. https://www.schrodinger.com/platform/materials-science, 2022. Accessed: 2022-9-23.

[28] Ehlert, Sebastian and Balduf, Ty and Bannwarth, Christoph and Grimme, Stefan. xtb-python. https://github.com/grimme-lab/xtb-python, 2022. Accessed: 2022-9-23.

[29] Art D Bochevarov, Edward Harder, Thomas F Hughes, Jeremy R Greenwood, Dale A Braden, Dean M Philipp, David Rinaldo, Mathew D Halls, Jing Zhang, and Richard A Friesner. Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *International Journal of Quantum Chemistry*, 113(18):2110–2142, 2013.

[30] Xiang Gao, Farhad Ramezanghorbani, Olexandr Isayev, Justin S Smith, and Adrian E Roitberg. TorchANI: A free and open source PyTorch-Based deep learning implementation of the ANI neural network potentials. *J. Chem. Inf. Model.*, 60(7):3408–3415, July 2020.

[31] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, et al. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27): 273002, 2017.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]

(c) Did you discuss any potential negative societal impacts of your work? [No]

(d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]

2. If you are including theoretical results...

(a) Did you state the full set of assumptions of all theoretical results? [N/A]

(b) Did you include complete proofs of all theoretical results? [N/A]

3. If you ran experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No] The dataset is provided; the code is proprietary.

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No]

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [Yes]

(b) Did you mention the license of the assets? [Yes]

(c) Did you include any new assets either in the supplemental material or as a URL? [Yes]

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Appendix

Training and test datasets of labeled data (atomic coordinates, system net charges, system energies, atomic forces, and atomic partial charges) are provided under the MIT license (`https://www.mit.edu/~amini/LICENSE.md`), and are available here: `https://figshare.com/s/c2631e40009be57474df` (note, this is an anonymized link). The dataset is a superset of that from Dajnowicz et al[10] containing 600,390 total datapoints, of which 362,382 are from Dajnowicz et al and 238,008 are new calculations in a format compatible with the old. Of the newly added points, 6902 are periodic systems. GFN2-xTB charges have also been added to all non-periodic systems for which GFN2-xTB labeling converged (581,954 points), including the points from the Dajnowicz et al dataset. The keys are arbitrary integers (stored as strings). Each value is a dictionary with the following properties:

```
{
  'atomicNumbers': [integers representing atomic number for each atom],
  'elements': [strings representing the element for each atom],
  'positions': [[x, y, z] coords of each atom, in Angstroms],
  'charge': integer representing system total charge,
  'multiplicity': integer representing spin multiplicity,
  'lattice': [[lattice vectors, in Angstroms, only exists if periodic]],
  'labels': dict of QM reference data for this point (energy, gradient, etc)
}
```

Every system needs at least one entry under 'labels': 'PBE-D3' or 'wB97X-D3BJ__def2-TZVPD'. Some label types will be missing from some points, for example xtbCharges for periodic systems. Labels not present for a datapoint are simply excluded from the loss function for that datapoint. Example 'labels' entry:

```
{
  'PBE-D3': {
    'totalEnergy': system energy in Hartree,
    'atomizationEnergy': system energy, in Hartree, minus per-atom offsets,
    'gradient': [[dE/dx, dE/dy, dE/dz] for each atom, in Hartree/Angstrom]
  }
  'wB97X-D3BJ__def2-TZVPD': {
     'totalEnergy': system energy in Hartree,
     'gradient': [[dE/dx, dE/dy, dE/dz] for each atom, in Hartree/Angstrom],
     'dipoleMoment': [system dipole moment vector in electron-Angstroms],
     'atomizationEnergy': system energy, in Hartree, minus per-atom offsets,
     'xtbCharges': [atomic partial charges from GFN2-xTB, in e-]
  }
}
```

The metallic lithium cluster geometries are sampled from common crystal packings (BCC, FCC, SC) supplemented with scaling and Gaussian noise. The organic reaction samples are generated by applying Articial Force-Induced Reactions (AFIR)[23, 24] sampled across the set of electrolytes by selecting atom pairs uniformly at random to induce reactions. The lithium-organic complexes are prepared using AFIR between existing clusters to cause them to react. In addition, we add periodic bulk lithium metal data for the common crystal packings, also supplemented with scaling and Gaussian noise.

All non-periodic DFT calculations in the training dataset are performed using Psi4[25] v1.6, with the same settings as in Dajnowicz et al (density fitting, a 1e-10 DFT basis tolerance, a 1e-10 Schwarz screening threshold)[10] with the addition of two newly available Psi4 acceleration settings: dft_pruning_scheme "robust" and s_orthogonalization "partialcholesky" with s_cholesky_tolerance set to 1e-6 (matching the 1e-6 linear dependency cutoff of Dajnowicz et al). Using Google Compute n2d-highmem CPU nodes with hyperthreading disabled, we observed a CPU time per system of approximately

$$t_{DFT,seconds} = 7\text{e}{-7} * \left(n_{basis}^2 * n_{basis-aux}\right) + 1 \tag{2}$$

Where n_{basis} is the number of basis functions used for the system and $n_{basis-aux}$ is the number of auxiliary basis functions. This agrees with the expected cubic scaling of density-fitting DFT.

All periodic DFT calculations in the training dataset are performed using Quantum Espresso[26] as implemented in Schrodinger Materials Science Suite[27] version 2022-3. The exchange-correlation energy was determined using generalized gradient approximation (GGA) Perdew-Burke-Ernzerhof (PBE) functional[4] coupled with D3 dispersion correction[14]. We used plane wave energy cutoff of 40 Ry to expand the electronic wavefunction. The Brillouin zone integration was performed using 3x3x3 k-point mesh for bulk supercells and 3x3x1 k-point mesh for the surface supercells.

GFN2-xTB charge labeling was performed using the Python API xtb-python[28] with default settings.

The test dataset of electrolyte decomposition reaction barriers was labeled using Jaguar[29] release 2019-4 using the ωB97X-D3(BJ)/def2-TZVPD and PBE-D3/def2-TZVPD levels of theory, with default settings for single-point energies. It consists of Li radical-initiated decomposition reactions for the following common electrolytes: ethylene carbonate, propylene carbonate, vinylene carbonate, dimethyl carbonate, diethyl carbonate, ethyl methyl carbonate, and fluoroethylene carbonate.

Our code is a proprietary fork of TorchANI[30], the original of which is available at https://github.com/aiqm/torchani.

Our phonon and molecular dynamics calculations were performed using our PyTorch code interfaced with the Atomic Simulation Environment (ASE)[31]. The molecular dynamics speed for this setup on a 1000-atom system was approximately 0.4ns/day on a 1080-Ti GPU in our internal cluster. Significantly greater speed is expected for a dedicated GPU implementation.

Our training was performed using the same hyperparameters as Dajnowicz et al[10], except for a training batch size of 128 (which makes periodic training more efficient), fixed epoch size of 200,000

(so that the length of training would not increase with dataset size), and a training/validation split uniformly at random at a ratio of 99:1 (we have not seen an increase in generalization from early stopping on larger validation sets, so we used a small validation set). Each training run was performed on a 1080-Ti GPU in our internal cluster for approximately 48 hours.