Training and Validation of Moment Tensor Potentials for Potassium Metal

RICHARD MENG[1,2,3], LAURENT BÉLAND[1], AND HAO SUN[1]

[1] Department of Mechanical and Materials Engineering, Queen's University, 45 Union St, Kingston, ON K7L 3N6
[2] contact@richardzjm.com
[3] 17zjm1@queensu.ca

Compiled April 12, 2023

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

Molecular Dynamics has been an invaluable tool in fields such as computational chemistry, materials science, and pharmaceutical drug discovery, for modelling and predicting the properties of compounds.

Atom-scale materials modelling requires a description of the energies and forces associated to interatomic interactions. Quantum chemistry techniques like Density Functional Theory (DFT) or Hartree-Fock methods provide such a description but are limited to systems of modest size. For instance, obtaining the interatomic forces in a system with a few hundred atoms can require tens of hours, even if using a hundred or so computer cores in parallel. Interatomic pair potentials overcome this size limitation and can be based on simple functional forms to reflect the various types of bonding exhibited in solids. Using these potentials, interatomic forces can be calculated in microseconds, enabling simulations that can reach much larger time and length scales, but at the cost of lost accuracy compared to DFT.

In recent years, machine learning potentials have emerged as an alternative that bridges the accuracy/cost compromise between pair potentials and DFT. These include neural network and graph neural network potentials, Gaussian process regression potentials and linear regression potentials. The moment tensor potential (MTP) falls in the latter category. The MTP considers radial interactions by fitting to polynomial basis sets and accounts for angular effects with nested outer products on position vectors. The MTP uses a complete basis set that is analytically differentiable and invariant under rotation, translation, and reflection. MTPs are trained on the output of DFT calculations. The MTP has been of recent interest to the Nuclear Materials group and Queen's MME Professor LK Béland who had been in contact with ECE professor Ryan Grant. Professor Grant specializes in the parallelization of algorithms and high-performance computing and is familiar with the development of LAMMPS, a popular atom-scale simulation software suite. Interfaces are available to apply the MTP in LAMMPS, although some algorithmic and efficiency limitations could be improved upon. In particular, I would like to parallelize some important aspects of the MTP training scheme which are currently only implemented as a serial code.

A. title

here

2. CORRESPONDING AUTHOR

We require manuscripts to identify a single corresponding author. The corresponding author typically is the person who submits the manuscript and handles correspondence throughout the peer review and publication process. If other statements about author contribution and contact are needed, they can be added in addition to the corresponding author designation.

The sections below show examples of different article components.

3. FIGURES AND TABLES

Do not place figures and tables at the back of the manuscript. Figures and tables should be placed and sized as they are likely to appear in the final article.

Figures and Tables should be labelled and referenced in the standard way using the `\label{}` and `\ref{}` commands.

A. Sample Figure

Figure 1 shows an example figure.

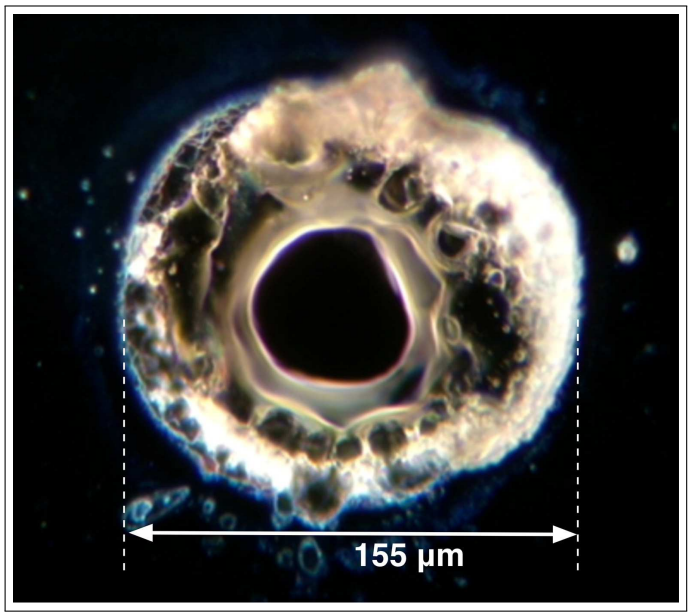

Fig. 1. Dark-field image of a point absorber.

B. Sample Table

Table 1 shows an example table.

Table 1. Shape Functions for Quadratic Line Elements

local node	$\{N\}_m$	$\{\Phi_i\}_m$ $(i = x, y, z)$
$m = 1$	$L_1(2L_1 - 1)$	Φ_{i1}
$m = 2$	$L_2(2L_2 - 1)$	Φ_{i2}
$m = 3$	$L_3 = 4L_1L_2$	Φ_{i3}

4. SAMPLE EQUATION

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $\mathrm{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i}^{n} X_i \qquad (1)$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

5. SAMPLE ALGORITHM

Algorithms can be included using the commands as shown in algorithm 1.

Algorithm 1. Euclids algorithm

```
procedure EUCLID(a, b)                ▷ The g.c.d. of a and b
    r ← a mod b
    while r ≠ 0 do                    ▷ We have the answer if r is 0
        a ← b
        b ← r
        r ← a mod b
    return b                          ▷ The gcd is b
```

A. Supplementary materials in Optica Publishing Group journals

Optica Publishing Group journals allow authors to include supplementary materials as integral parts of a manuscript. Such materials are subject to peer-review procedures along with the rest of the paper and should be uploaded and described using the Prism manuscript system. Please refer to the Author Guidelines for Supplementary Materials in Optica Publishing Group Journals for more detailed instructions on labeling supplementary materials and your manuscript. For preprints submitted to Optica Open a link to supplemental material should be included in the submission, but do not upload the material.

Authors may also include Supplemental Documents (PDF documents with expanded descriptions or methods) with the primary manuscript. At this time, supplemental PDF files are not accepted for JOCN or PRJ. To reference the supplementary document, the statement "See Supplement 1 for supporting content." should appear at the bottom of the manuscript (above the References heading). Supplemental documents are not accepted for Optica Open preprints.

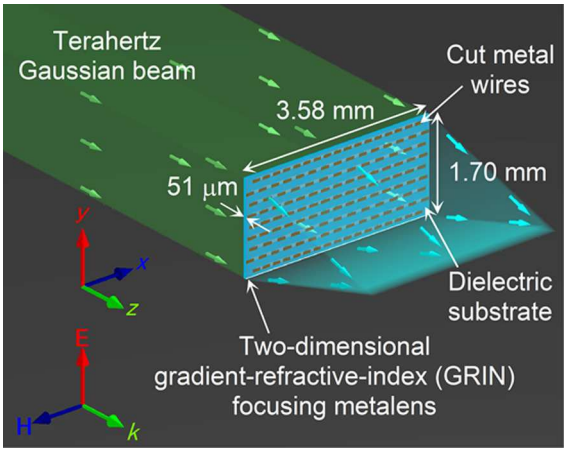

Fig. 2. Terahertz focusing metalens.

B. Sample Dataset Citation

1. M. Partridge, "Spectra evolution during coating," figshare (2014), http://dx.doi.org/10.6084/m9.figshare.1004612.

C. Sample Code Citation

2. C. Rivers, "Epipy: Python tools for epidemiology," Figshare (2014) [retrieved 13 May 2015], http://dx.doi.org/10.6084/m9.figshare.1005064.

6. BACKMATTER

Backmatter sections should be listed in the order Funding/Acknowledgment/Disclosures/Data Availability Statement/Supplemental Document section. An example of backmatter with each of these sections included is shown below.

Funding. Content in the funding section will be generated entirely from details submitted to Prism. Authors may add placeholder text in the manuscript to assess length, but any text added to this section in the manuscript will be replaced during production and will display official funder names along with any grant numbers provided. If additional details about a funder are required, they may be added to the Acknowledgments, even if this duplicates information in the funding section. See the example below in Acknowledgements. For preprint submissions, please include funder names and grant numbers in the manuscript.

Acknowledgments. The section title should not follow the numbering scheme of the body of the paper. Additional information crediting individuals who contributed to the work being reported, clarifying who received funding from a particular source, or other information that does not fit the criteria for the funding block may also be included; for example, "K. Flockhart thanks the National Science Foundation for help identifying collaborators for this work."

Disclosures. Disclosures should be listed in a separate section at the end of the manuscript. List the Disclosures codes identified on the Conflict of Interest policy page. If there are no disclosures, then list "The authors declare no conflicts of interest."

Here are examples of disclosures:

Disclosures. ABC: 123 Corporation (I,E,P), DEF: 456 Corporation (R,S). GHI: 789 Corporation (C).

Disclosures. The authors declare no conflicts of interest.

Data Availability Statement. A Data Availability Statement (DAS) will be required for all submissions beginning 1 March 2021. The DAS should be an unnumbered separate section titled "Data Availability" that immediately follows the Disclosures section. See the Data Availability Statement policy page for more information.

There are four common (sometimes overlapping) situations that authors should use as guidance. These are provided as minimal models, and authors should feel free to include any additional details that may be relevant.

1. When datasets are included as integral supplementary material in the paper, they must be declared (e.g., as "Dataset 1" following our current supplementary materials policy) and cited in the DAS, and should appear in the references.

   Data availability. Data underlying the results presented in this paper are available in Dataset 1, Ref. [3].

2. When datasets are cited but not submitted as integral supplementary material, they must be cited in the DAS and should appear in the references.

   Data availability. Data underlying the results presented in this paper are available in Ref. [3].

3. If the data generated or analyzed as part of the research are not publicly available, that should be stated. Authors are encouraged to explain why (e.g. the data may be restricted for privacy reasons), and how the data might be obtained or accessed in the future.

   Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

4. If no data were generated or analyzed in the presented research, that should be stated.

   Data availability. No data were generated or analyzed in the presented research.

Data availability statements are not required for preprint submissions.

Supplemental document. See Supplement 1 for supporting content.

7. REFERENCES

Note that *Optics Letters* and *Optica* short articles use an abbreviated reference style. Citations to journal articles should omit the article title and final page number; this abbreviated reference style is produced automatically when the *Optics Letters* journal option is selected in the template, if you are using a .bib file for your references.

However, full references (to aid the editor and reviewers) must be included as well on a fifth informational page that will not count against page length; again this will be produced automatically if you are using a .bib file.

Add citations manually or use BibTeX. See [1–5].

REFERENCES

1. Y. Zhang, S. Qiao, L. Sun, Q. W. Shi, W. Huang, L. Li, and Z. Yang, "Photoindehehuced active terahertz metamaterials with nanostructured vanadium dioxide film deposited by sol-gel method," Opt. Express **22**, 11070–11078 (2014).
2. Optica Publishing Group, "Optica," https://opg.optica.org.
3. P. Forster, V. Ramaswamy, P. Artaxo, T. Bernsten, R. Betts, D. Fahey, J. Haywood, J. Lean, D. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz, and R. V. Dorland, "Changes in atmospheric consituents and in radiative forcing," in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth assesment report of Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miler, eds. (Cambridge University Press, 2007).
4. R. McKay, "X-ray crystallography," Ph.D. thesis, Princeton University (1982).
5. V. S. C. Manga Rao and S. Hughes, "Single quantum-dot Purcell factor and β factor in a photonic crystal waveguide," Phys. Rev. B **75** (2007).