An introduction to classical molecular dynamics simulation for experimental scattering users

A. R. McCluskey,[1, 2, *] J. Grant,[3] A. R. Symington,[1] T. Snow,[2, 4]
J. Doutch,[5] B. J. Morgan,[1, †] S. C. Parker,[1] and K. J. Edler[1]

[1]*Department of Chemistry, University of Bath, Claverton Down, Bath, BA2 7AY, UK*
[2]*Diamond Light Source, Harwell Campus, Didcot, OX11 0DE, UK*
[3]*Computing Services, University of Bath, Claverton Down, Bath, BA2 7AY, UK*
[4]*School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK*
[5]*ISIS Facility, Rutherford Appleton Laboratory, STFC, Chilton, Didcot, OX11 0QX, UK*
(Dated: February 4, 2019)

Classical molecular dynamics simulations are a common component of multi-modal analyses from scattering measurements, such as small-angle scattering and diffraction. Users of these experimental techniques often have no formal training in the theory and practice of molecular dynamics simulation, leading to the possibility of these simulations being treated as a "black box" analysis technique. In this article we describe an open educational resource (OER) designed to introduce classical molecular dynamics to users of scattering methods. This resource is available as a series of interactive web pages, which can be easily accessed by students, and as a open source software repository, which can be freely copied, modified, and redistributed by educators. The topic covered in this OER include classical atomistic modelling, parameterising interatomic potentials, molecular dynamics simulations, typical sources of error, and some of the approaches to using simulations in the analysis of scattering data.

I. INTRODUCTION

The use of molecular dynamics simulations to help analyse and interpret experimental data from small-angle scattering and diffraction studies has grown significantly over the past ten years [1–8], with the percentage of small-angle scattering publications that mention molecular dynamics reaching more than 20 % in 2018 (Fig. I). Users of scattering and diffraction techniques often have backgrounds in experimental science, and may have received little formal training in the theory or practice of computational modelling. This can lead to molecular dynamics simulations being used as a "black box", without understanding the underlying methodologies, or considering possible sources of error. The use of molecular simulations without any technical understanding can lead to unintended but severe errors. To help support researchers use molecular dynamics simulation in their analysis of scattering data, while reducing this risk of modelling errors, a number of software tools, such as WAXSiS and SASSIE [9–11], have been developed that present easy-to-use, graphical, web-based user interfaces.

A complementary approach is to organise educational activities, such as lectures or workshops, tailored to introduce molecular simulation techniques to audiences of scattering and diffraction users. One example is the annual ISIS Neutron Training Course, which includes a module titled "An Introduction to Molecular Dynamics for Neutron Scattering". This module covers the fundamentals of classical molecular dynamics simulation,
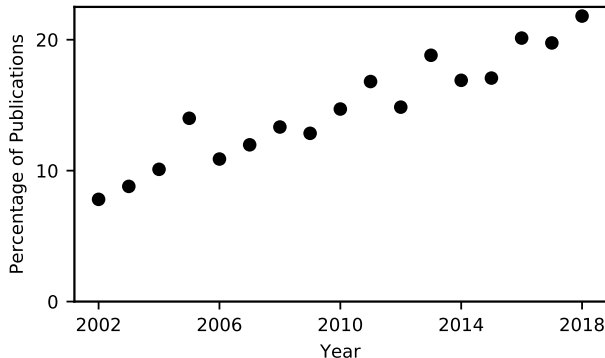

FIG. 1. The annual percentage of publications that mention "small angle scattering" that also mention "molecular dynamics", determined from the numbers of matching Google Scholar results.

presents applications of these methods in neutron science, and gives students practical hands-on experience with the SASSIE software package [11].

While lectures and workshops are an effective tool for education and training, participation can be limited due to difficulties attending in person (due to location or cost) or physical limits on student numbers. An alternative educational strategy gaining popularity within scientific and engineering communities is the publication of technology-enhanced open educational resources (OERs). These are courses, lectures, or learning resources published online that are freely available for use by anyone. In addition to their broad accessibility, these resources have permissive "open" copyright licences that allow their use in the "5R activities": retain, reuse, re-

---
* a.r.mccluskey@bath.ac.uk; andrew.mccluskey@diamond.ac.uk
† b.j.morgan@bath.ac.uk

vise, remix, and redistribute [12]. Publishing a resouce as an OER increases its reach and impact, because others may use it in their own teaching not only in its original form, but are free to modify, and redistribute, the material to better suit their aims. The Jupyter Notebook framework [13] has become a popular platform for OERs that teach computational skills, because it allows authors to include instructional text, images, and other media, alongside executable editable code, in an example of "literate programming" [14]. This format encourages students to directly interact with code examples by running, editing, and rerunning these within the source document [15], supporting exploratory experiential learning [16].

Here, we present an online, open-source, interactive learning resource written to introduce members of the scattering and diffraction community to molecular dynamics simulations. This OER comprises six lessons that introduce classical molecular dynamics methods, and show how these can be used in the context of scattering [can this be more specific?] in a multi-modal analysis. We use the open-source Python library pylj [17] to provide simple, but computationally authentic, examples of simulations, that demonstrate visually and programmatically the conceptual relationships between simulation and scattering techniques. In this article, we discuss the structure of the resource and describe how a student may get the most from the resource [18].

II.   ASSUMED PRIOR KNOWLEDGE

The OER, entitled *"The Interaction Between Simulation and Scattering"* makes use of the Python programming language to include interactive examples of mathematical and algorithmic content. To be able to fully use this resource, therefore, requires some knowledge of, or willingness to learn, Python. We have, however, attempted to design the resource in such a fashion that an in-depth knowledge of Python is not a prerequisite. It is anticipated that the users of this resource have some experience of at least undergraduate level chemistry or physics, and a commensurate level of mathematical understanding. This background knowledge will be particularly important for the sections dealing with the functional form and chemical basis of classical interaction potentials.

III.   RESOURCE CONSTRUCTION

The resource is available in two main formats. First, as a series of webpages, hosted at `http://pythoninchemistry.org/sim_and_scat`. Secondly as the source-code repository used to build these webpages, hosted at `https://github.com/pythoninchemistry/sim_and_scat` [19]. The source content consists of a set of Jupyter Notebooks and Markdown files, which are au-

tomatically compiled using the `jupyter-book` tool [20] to generate the web version. This system allows the resulting webpages to include text, equations, and figures, which we use to describe key concepts and to explain details of algorithms, as well as Python code blocks, which we use to provide specific examples. The webpages have Thebelab and BinderHub integrations [21–23], which allow students to launch interactive versions of these webpages that allow execution and editing of the included Python code. We believe that the ability to read the resource as an "interactive document" will help students to develop a deeper understanding of the materials than they would from a traditional static document. The resource is provided under a CC-BY license [24] [Note, expand to explain the use of two resources (both permissive)]. This open-source license allows readers to reuse or remix the material to enhance their own educational platform, and for secondary authors to contribute to improving the source material.the material to be anyone to use the material to enhance their own educational platform and experts in the field may contribute to improving the resource.

IV.   RESOURCE OUTLINE

The resource follows a simple outline that introduces key aspects of molecular dynamics simulations.

A.   Home

The welcome page introduces the resource and gives the student information about how the resource may be used, including details of the Thebelab and BinderHub integrations. This page also contains details about the permitted use and sharing of the content of the resource, and about the resource license. This page also includes a list of authors and contributors.

B.   Classical methods

After a brief introduction, this section introduces concepts related to classical simulation methods. This includes the use of interatomic potential functions, alongside some examples, such as the Lennard-Jones and Buckingham potential models [25, 26]. We then introduce the problem of parameterising a potential model, including the use of higher accuracy quantum mechanical calculations to do so. The presence of off-the-shelf, general potential models are discussed; with the caveat that they may still require system specific optimisations. Finally, we mention mixing rules; again discussing the possible problems that a user may encounter if applying these blindly to specific systems.

C. Molecular dynamics

Having introduced the concept of classical interatomic potentials, we then discuss how these are used in molecular dynamics simulations. We go through different aspects of the method [specifics?] to build up a one-dimensional NVE (constant number, volume, and energy) molecular dynamics simulation, using the Velocity-Verlet algorithm and the Lennard-Jones potential model [25, 27]. This [what does This refer to?] is introduced in terms of Newton's laws of motion and the generalised equations of motion. Finally, we discuss a range of key factors that can affect molecular dynamics simulations, including including choice of simulation ensembles, setting the range cut-off for an interatomic potential, and the use of periodic boundary conditions.

D. `pylj` and the interaction with scattering

The final aspect of the resource covers using molecular dynamics simulations to understand scattering profiles. This is presented as a practical example, using the open-source `pylj` package [17]. We demonstrate a two-dimensional molecular dynamics simulation of argon particles interacting through a Lennard-Jones potential. The student is first shown a working `pylj` simulation and invited to interact with the simulation and the custom plotting functionality of `pylj`. In the final lessons, we present the Debye equation [28] [And it's use in combination with MD data?], and the student is invited to observe the effect of simulation temperature on the resulting scattering profile. We finish by discussing alternative, faster, algorithms for calculating scattering pro-files, such as the Fibonacci sequence or Golden Vectors methods [29, 30] [Why is Fibonacci sequence not title case, but Golden Vectors is?].

V. FUTURE OUTLOOK

We hope that publishing this material as an OER will promote interest from other parties towards reuse, and remixing the material for their own educational purposes. We also hope to implement this material within training at the ISIS Neutron and Muon Source as well as Diamond Light Source. Finally, it is hoped, *via* student and community feedback, that the implementation, content, and pedagogical approach of this resource can be increased over time. We hope that this resource will benefit students and researchers interested in combining molecular dynamics analysis with scattering and diffraction experiments.

VI. AUTHOR CONTRIBUTIONS

The open education resource was developed, and the manuscript written, by A. R. M. with input from all authors.

ACKNOWLEDGMENTS

[1] J. Pan, F. A. Heberle, S. Tristram-Nagle, M. Szymanski, M. Koepfinger, J. Katsaras, and N. Kučerka, BBA - Biomembranes **1818**, 2135 (2012).

[2] L. Boldon, F. Laliberte, and L. Liu, Nano Rev. **6**, 25661 (2015).

[3] J. S. Hub, Curr. Opin. Struct. Biol. **49**, 18 (2018).

[4] M. T. Ivanović, L. K. Bruetzel, J. Lipfert, and J. S. Hub, Angew. Chemie Int. Ed. **57**, 5635 (2018).

[5] A. East, A. E. Mechaly, G. H. M. Huysmans, C. Bernarde, D. Tello-Manigne, N. Nadeau, A. P. Pugsley, A. Buschiazzo, P. M. Alzari, P. J. Bond, and O. Francetic, Structure **24**, 92 (2016).

[6] M. E. Wall, A. H. van Benschoten, N. K. Sauter, P. D. Adams, J. S. Fraser, and T. C. Terwilliger, Proc. Natl. Acad. Sci. U.S.A **111**, 17887 (2014).

[7] M. E. Wall, IUCrJ **5**, 172 (2018).

[8] M. Satoh, H. Saburi, T. Tanaka, Y. Matsuura, H. Naitow, R. Shimozono, N. Yamamoto, H. Inoue, N. Nakamura, Y. Yoshizawa, T. Aoki, R. Tanimura, and N. Kunishima, FEBS Open Bio. **5**, 557 (2015).

[9] P.-C. Chen and J. S. Hub, Biophys. J. **107**, 435 (2014).

[10] C. J. Knight and J. S. Hub, Nucleic Acids Res. **43**, W225 (2015).

[11] S. J. Perkins, D. W. Wright, H. Zhang, E. H. Brookes, J. Chen, T. C. Irving, S. Krueger, D. J. Barlow, K. J. Edler, D. J. Scott, N. J. Terrill, S. M. King, P. D. Butler, and J. E. Curtis, J. Appl. Crystallogr. **49**, 1861 (2016).

[12] D. Wiley, "Open content - definition," http://opencontent.org/definition (2018).

[13] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, and C. Willing, Positioning and Power in Academic Publishing: Players, Agents and Agendas , 87 (2016).

[14] D. E. Knuth, Comput. J. **27**, 97 (1984).

[15] L. A. Barba, A. Wickenheiser, and R. Watkins, "Cyber-Training: DSE–The Code Maker: Computational Thinking for Engineers with Interactive Contextual Learning," https://doi.org/10.6084/m9.figshare.5662051.v1 (2017).

[16] S. Papert, *Mindstroms* (Basic Books, New York City, US, 1993).

[17] A. R. McCluskey, B. J. Morgan, K. J. Edler, and S. C. Parker, J. Open Source Educ. **1**, 19 (2018).

[18] In this work *the student* refers to anyone working through the OER, regardless of career position.

[19] A. R. McCluskey, J. Grant, A. R. Symington, T. Snow, J. Doutch, B. J. Morgan, S. C. Parker, and K. J. Edler, "Pythoninchemistry/sim_and_scat: Sim_and_scat-0.1," http://doi.org/10.5281/zenodo.2543277 (2019).

[20] S. Lau and C. Holdgraf, "Jupyter/jupyter-book Beta-2," https://github.com/jupyter/jupyter-book (2019).

[21] B. Ragan-Kelley, N. M. Thiéry, S. Corlay, S. Gutsche, C. Holdgraf, and T. Head, "Minrk/thebelab v0.3.3," https://github.com/minrk/thebelab (2019).

[22] B. Ragan-Kelley, Y. Panda, C. Willing, C. Holdgraf, K. Erdogan, T. Head, M. Bussonnier, H. Zarea, M. van Niekerk, C. Yick, and E. Sundell, "Jupyterhub/binderhub FirstLight," https://github.com/jupyterhub/binderhub (2019).

[23] Jupyter and others, in *Proceedings of the 17th Python in Science Conference.* (2018).

[24] "Creative Commons – Attribution 4.0 International – CC-BY-4.0," https://creativecommons.org/licenses/by/4.0/ (2019).

[25] J. E. Lennard-Jones, Proc. Royal Soc. Lond. A. **106**, 463 (1924).

[26] R. A. Buckingham, Proc. Royal Soc. Lond. A. **168**, 264 (1938).

[27] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[28] P. Debye, Ann. Phys. **351**, 809 (1915).

[29] D. I. Svergun, Acta Crystallogr. A **50**, 391 (1994).

[30] M. C. Watson and J. E. Curtis, J. Appl. Crystallogr. **46**, 1171 (2013).