Introducing classical molecular dynamics simulation to users of scattering

A. R. McCluskey,[1, 2, *] A. R. Symington,[1] T. Snow,[2, 3] J. Grant,[4] 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]*School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK*
[4]*Computing Services, University of Bath, Claverton Down, Bath, BA2 7AY, UK*
(Dated: January 11, 2019)

Classical molecular dynamics simulations are rapidly becoming a popular component of multi-modal analyses from scattering measurements; such as small angle scattering and diffraction. However, few users of these techniques have formalised training in these methodologies, resulting in frequent use of molecular dynamics simulations as a black box technique. This work discusses an open educational resource designed to introduce classical molecular dynamics to users of scattering, describing possible sources of error in the method. Furthermore, we cover some of the methods that can be used to enable simulation techniques to facilitate in the analysis of scattering data.

I. INTRODUCTION

The use of molecular dynamics simulations to aid in the analysis of experimental data, particularly from small angle scattering (SAS) and diffraction, has grown significantly over the past ten years [1–8]. Figure I shows the growth in the percentage of SAS publications that mention molecular dynamics. It can be seen that there has been a linear growth in such publications, with more than 20 % of all SAS publications mentioning molecular dynamics as of 2018. This alone is a clear indication of the importance that simulation techniques are having on the field of scattering and diffraction.

Normally, the users of scattering and diffraction techniques have a background in experimental science, with little or no formalised training in computational modelling techniques. This may be problematic, as it can lead to the use of molecular simulation as a black-box, with little understanding or consideration for the underlying methodologies. The use of molecular simulations in this fashion can quickly lead to the inclusion of severe, systematic errors. This has lead to the development of tools, such as WAXSiS or SASSIE [9–11], that present easy-to-use interfaces designed to reduce the risk of these errors.

In addition to the development of software packages another method to further limit systematic errors consists of lectures or workshops presenting and introduction to molecular simulation for scattering and diffraction users. For example, the annual ISIS Neutron Training Course now includes a module entitled "An Introduction to Molecular Dynamics for Neutron Scattering". This details the fundamentals of classical molecular dynamics simulation, before presenting applications of these methods in neutron science in turn allowing students an opportunity to gain familiarity with the SASSIE software package [11]. While such lectures and workshops com-
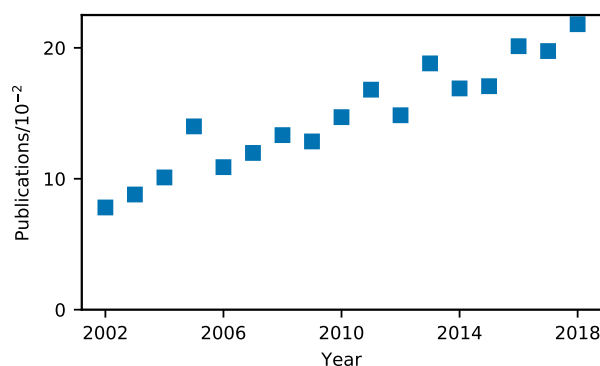

FIG. 1. The annual growth of the percentage of publication that mention "small angle scattering" which also mention "molecular dynamics". Determined from the number of results from a Google Scholar search.

prise an important aspect of training, they are limited in terms of the numbers of students that may attend and the course's location. Ultimately this means that not all scattering and diffraction users may easily access such resources.

There has been a recent movement within the scientific and engineering communities towards technology-enhanced open educational resources (OERs). These are courses, lectures, or learning modules that are made freely available online for use by anyone. This allows the permission for others to engage in the "5R activities": Retain, Reuse, Revise, Remix, and Redistribute [12]. The act of publishing an OER increases the reach of a particular resource as it allows others to pick it up, modify it, and use the modified version in their own teaching. The availability of the Jupyter Notebook framework [13] has facilitated the availability of OERs by improving the ability for students to interact directly with the resource often, but not always, using the Python programming language [14].

Herein, we present an online, open-source, interactive module aimed to introduce members of the scattering

* a.r.mccluskey@bath.ac.uk
† k.edler@bath.ac.uk

and diffraction community to molecular dynamics simulations. This course comprises of six lessons, covering classical methods, introducing molecular dynamics, and showing how this method may interact with scattering in a multi-modal fashion. We leverage the open-source Python library pylj [15] to give a visual and programmatic representation of the interactions between these two techniques. In this paper we will discuss the implementation of the module and detail how the student may get the most from the module (in this work *the student* refers to anyone undertaking the course, regardless of career position).

II.   ASSUMED PRIOR KNOWLEDGE

The OER, entitled *"The interaction between simulation and scattering"* makes use of the Python programming language, this allows for interactive examples of the mathematical and algorithm content to be included. A result of this is that to be able to fully utilise the course some knowledge of, or willingness to learn, Python is required. However, we have attempted to develop the course in such a fashion that an in depth knowledge of Python is not required. It is anticipated that the users of this course would have some familiarity with undergraduate level chemistry or physics, alongside a commensurate understanding of mathematics. This is particularly important when considering the nature, and chemical rationale, of classical interaction potentials.

III.   MODULE CONSTRUCTION

The module is currently available online, at `https://arm61.github.io/sim_and_scat`. These webpages were written as a series of Jupyter Notebooks and compiled using the build-system developed within the University of Bath [REFERENCE?]. The use of this system allows for Python code blocks to be easily built into lessons providing algorihtmic details whilst also facilitating interactivity. It is expected that the student would use the online material alongside a locally running Jupyter Notebook (the installation of which is detailed in the module introduction). Interacting with the module in this fashion will allow the student to copy the Python code from the web interface, run it locally, and alter the code as they see fit.

The module is provided under a CC-BY-SA-4.0 license [16], and builds on the growing library of open educational resources. The open-source nature of this license means that anyone may use the material to enhance their own educational platform and experts in the field may contribute to improve the module. The source code for the module is available at `https://github.com/arm61/sim_and_scat`.

IV.   MODULE OUTLINE

The module follows a simple outline to introduce important aspects of molecular dynamics simulations. Code blocks are used to gradually build up the student's understanding of the various concepts.

A.   Getting started

The welcome page, nominally Lesson 0, introduces the course, outlines the content and gives the student information about preparing their computer in order to run the necessary software, including Anaconda Python and the pylj package [15] It is important for the student to have their computer ready with the necessary software before beginning the module. This ensures that they are able to interact fully with the content of each lesson.

B.   Classical methods

Once the necessary software is installed, concepts relating broadly to classical simulation methods are introduced. This includes the functional nature of potential modelling and gives some examples, such as the Lennard-Jones and Buckingham potential models [17, 18]. Potential model parameterisation is briefly covered including the use of higher accuracy quantum mechanical calculations to do so. The presence of off-the-shelf, general potential models are discussed; with particular attention paid to the caveat that they may still require system specific optimisations. Finally, we mention mixing, or combination, rules; again discussing the possible problems that a user may encounter related to system specificity.

C.   Molecular dynamics

Building on the classical potential model methods the student is then presented with molecular dynamics. This is shown by using different aspects for the method to gradually build up a one-dimensional molecular dynamics simulation using NVE (Number, Volume and Energy) ensembles, the Velocity-Verlet algorithm and the Lennard-Jones potential model [17, 19]. This is introduced in terms of Newton's laws of motion and the generalised equations of motion, which should be familiar to most students. Finally, different importance considerations for molecular dynamics simulations are described including ensembles, the potential cut-off, and periodic boundary conditions.

D.   pylj and the interaction with scattering

The final aspect of the module is to utilise a molecular dynamics simulation to understand some scattering pro-

files. This is achieved using the open-source pylj package [15], which allows for a two-dimensional simulation or argon particles, interacting through a Lennard-Jones potential, to be performed. The student is firstly shown a working pylj simulation and invited to interact with the simulation and the custom plotting functionality of pylj. In the final lessons, a formulation of the Debye equation [20] is given and the student is invited to observe the effect of simulation temperature on the resulting scattering profile. There is mention of other, faster, algorithms for the determination of scattering profiles, such as the Fibonacci sequence or Golden Vectors methods [21, 22].

V. FUTURE OUTLOOK

In future, we hope that the nature of the material, as an OER, will promote interest from other parties towards reuse, and remixing the material. Furthermore, we 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 improvement of implementation, materials, and general pedagogy of the module can be achieved, driving up the quality and depth of molecular dynamics analyses performed on results obtained from scattering and diffraction measurements.

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. Hub, Curr. Opin. Struct. Biol. **49**, 18 (2018).

[4] M. T. Ivanović, L. K. Bruetzel, J. Lipfert, and J. S. Hub, Ange. Chem. 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, PNAS **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. Kunishma, 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. 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), this material was created by David Wiley and published freely under a Creative Commons Attribution 4.0 license at http://opencontent.org/definition/. Accessed: 2018-11-19.

[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, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, edited by F. Loizides and B. Schmidt (IOS Press, 2016) pp. 87–90.

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

[15] A. R. McCluskey, B. J. Morgan, K. J. Edler, and S. C. Parker, JOSE **1**, 19 (2018).

[16] C. Commons, "Creative commons – attribution-sharealike 4.0 international – cc-by-sa 4.0," `https://creativecommons.org/licenses/by-sa/4.0/` (2019).

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

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

[19] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 648 (Appendix) (1982).

[20] P. J. W. Debye, Ann. Phys. **351**, 809 (1915).

[21] D. I. Svergun, Acta. Cryst. **A50**, 391 (1994).

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