The Journal of Open Source Software

MDCraft: A Python assistant for performing and analyzing molecular dynamics simulations of soft matter systems

Benjamin B. Ye [1¶], **Pierre J. Walker** [1,2], **and Zhen-Gang Wang** [1]

1 Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States **2** Department of Chemical Engineering, Imperial College, London SW7 2AZ, United Kingdom ¶ Corresponding author

DOI: 10.21105/joss.07013

Software

- Review
- Repository
- Archive

Editor: Sarath Menon

Reviewers:

- @raynol-dsouza
- @aazocar

Submitted: 23 June 2024
Published: 15 August 2024

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

MDCraft is a comprehensive Python package designed to enhance research workflows involving molecular dynamics (MD) simulations. It streamlines the entire process—from setting up and executing simulations to analyzing trajectories using sophisticated algorithms and visualizing results—making computational chemistry more accessible to a broader audience. At its core, MDCraft comprises three principal components.

First, the `openmm` module provides user-friendly tools to initialize, optimize, and run simulations, enabling the exploration of various large soft matter systems across different timescales. This module extends the functionality of the OpenMM (Eastman et al., 2017) simulation package by introducing custom force fields, such as the efficient and intuitive Gaussian core model with smeared electrostatics (GCMe) (Ye et al., 2024); incorporating advanced techniques like the slab correction (Ballenegger et al., 2009; Yeh & Berkowitz, 1999) and the method of image charges (Hautman et al., 1989) for charged systems with slab geometries; facilitating coarse-grained MD simulations by scaling physical values by the fundamental quantities (mass m, length d, energy ϵ, and Boltzmann constant $k_{\mathrm{B}}T$); and offering feature-rich readers and writers for topologies and trajectories stored in memory-efficient formats (such as NetCDF).

Second, the `algorithm` and `analysis` modules offer optimized serial and multithreaded algorithms and analysis classes for evaluating structural, thermodynamic, and dynamic properties using thermodynamic state and trajectory data. The analysis classes provide properties including, but not limited to, static and dynamic structure factors (Ashcroft & Langreth, 1967; Faber & Ziman, 1965; Róg et al., 2003), density and potential profiles, end-to-end vector autocorrelation functions for polymers, and Onsager transport coefficients (Fong et al., 2020; Rubinstein & Colby, 2003). The algorithms provide the underlying tools used to perform analysis and are intended to be easily extensible by more-advanced users. These modules are not limited to OpenMM and can also be used with simulation run in other packages, such as LAMMPS (Thompson et al., 2022) and GROMACS (Abraham et al., 2015).

Finally, the `fit` and `plot` modules simplify the post-processing and visualization of data, aiding in the creation of figures suitable for scientific publications. These modules consist of models for curve fitting and helper functions that interface seamlessly with the commonly used SciPy (Virtanen et al., 2020) and Matplotlib (Hunter, 2007) libraries.

Together, these modules provide both novice and experienced MD simulation users with a comprehensive set of tools necessary to conduct computer experiments ranging from simple to complex, all within a single, succinct package.

Statement of need

Although established MD analysis packages such as MDAnalysis (Michaud-Agrawal et al., 2011), MDTraj (McGibbon et al., 2015), freud (Ramasubramani et al., 2020) and pytraj (Roe & Cheatham, 2013) have been around for a considerable time, they primarily focus on the post-simulation analysis. In contrast, MDCraft not only improves upon some of the analysis tools provided, but is designed to also provide comprehensive support throughout the entire simulation process, from initialization to post-processing.

MDCraft is tightly integrated with OpenMM, something that is unique amongst MD analysis packages. OpenMM is a relatively new simulation toolkit that has seen a surge in popularity in recent years due to its class-leading performance and flexibility through support for custom intermolecular forces and integrators for equations of motion (Eastman et al., 2017). Due to its age and design philosophy, OpenMM offers comparatively fewer choices of pair potentials and external forces, and no built-in analysis support. MDCraft fills this gap in two ways. First, the `openmm` module leverages the modularity of OpenMM to provide a suite of custom force fields, problem-solving tools, trajectory readers and writers, and utility functions for unit reduction, topology transformations, and performance optimizations that are not typically available in other simulation packages. Of special significance is the support for GCMe which, as demonstrated in a recent article (Ye et al., 2024), provides significant acceleration compared to other force fields while also remaining physically meaningful. Then, the classes in the `analysis` module enable computing common structural, thermodynamic, and dynamic properties using the topology, trajectory, and state data generated by OpenMM (or other simulation packages).

The `analysis` module also stands out due to the flexibility it affords its end users, in contrast to contemporary MD analysis packages. General users have substantial control over what aspects of the properties to calculate and which method to employ through a plethora of well-documented built-in options in each analysis class, without having to be concerned about the underlying implementations. More advanced users, on the other hand, have the option to work directly with the algorithms in the `algorithms` module for further customization. These analysis functions and classes have proven indispensable in several recent publications (Glisman et al., 2024; Lee et al., 2024; Mantha et al., 2024).

The application of MDCraft extends across various domains within computational chemistry and materials science. Researchers can utilize it to study the low-level mechanisms involved in supercapacitors, polymer gels, drug delivery systems, and nanomaterial synthesis, thus highlighting its versatility and broad applicability in cutting-edge scientific research.

Acknowledgements

We acknowledge contributions from Alec Glisman and Dorian Bruch in the development of this package and financial support from Hong Kong Quantum AI Lab, AIR@InnoHK of the Hong Kong Government.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Ashcroft, N. W., & Langreth, D. C. (1967). Structure of Binary Liquid Mixtures. I. *Physical Review*, *156*(3), 685–692. https://doi.org/10.1103/PhysRev.156.685

Ballenegger, V., Arnold, A., & Cerdà, J. J. (2009). Simulations of non-neutral slab systems with long-range electrostatic interactions in two-dimensional periodic boundary conditions.

J. Chem. Phys., *131*(9), 094107. https://doi.org/10.1063/1.3216473

Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., & Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.*, *13*(7), e1005659. https://doi.org/10.1371/journal.pcbi.1005659

Faber, T. E., & Ziman, J. M. (1965). A theory of the electrical properties of liquid metals: III. The resistivity of binary alloys. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, *11*(109), 153–173. https://doi.org/10.1080/14786436508211931

Fong, K. D., Self, J., McCloskey, B. D., & Persson, K. A. (2020). Onsager Transport Coefficients and Transference Numbers in Polyelectrolyte Solutions and Polymerized Ionic Liquids. *Macromolecules*, *53*(21), 9503–9512. https://doi.org/10.1021/acs.macromol.0c02001

Glisman, A., Mantha, S., Yu, D., Wasserman, E. P., Backer, S., & Wang, Z.-G. (2024). Multivalent Ion-Mediated Polyelectrolyte Association and Structure. *Macromolecules*, *57*(5), 1941–1949. https://doi.org/10.1021/acs.macromol.3c02437

Hautman, J., Halley, J. W., & Rhee, Y. -J. (1989). Molecular dynamics simulation of water beween two ideal classical metal walls. *J. Chem. Phys.*, *91*(1), 467–472. https://doi.org/10.1063/1.457481

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Lee, S., Walker, P., Velling, S., Chen, A., Taylor, Z., Fiori, C., Gandhi, V., Wang, Z.-G., & Greer, J. (2024). *Molecular Control via Dynamic Bonding Enables Material Responsiveness in Additively Manufactured Metallo-Polyelectrolytes*. https://doi.org/10.21203/rs.3.rs-3643582/v1

Mantha, S., Glisman, A., Yu, D., Wasserman, E. P., Backer, S., & Wang, Z.-G. (2024). Adsorption Isotherm and Mechanism of Ca^{2+} Binding to Polyelectrolyte. *Langmuir*, *40*(12), 6212–6219. https://doi.org/10.1021/acs.langmuir.3c03640

McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., Schwantes, C. R., Wang, L.-P., Lane, T. J., & Pande, V. S. (2015). MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.*, *109*(8), 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015

Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. (2011). MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.*, *32*(10), 2319–2327. https://doi.org/10.1002/jcc.21787

Ramasubramani, V., Dice, B. D., Harper, E. S., Spellings, M. P., Anderson, J. A., & Glotzer, S. C. (2020). Freud: A software suite for high throughput analysis of particle simulation data. *Computer Physics Communications*, *254*, 107275. https://doi.org/10.1016/j.cpc.2020.107275

Roe, D. R., & Cheatham, T. E. I. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, *9*(7), 3084–3095. https://doi.org/10.1021/ct400341p

Róg, T., Murzyn, K., Hinsen, K., & Kneller, G. R. (2003). *N* Moldyn: A program package for a neutron scattering oriented analysis of molecular dynamics simulations: *n* Moldyn. *Journal of Computational Chemistry*, *24*(5), 657–667. https://doi.org/10.1002/jcc.10243

Rubinstein, M., & Colby, R. H. (2003). *Polymer Physics*. Oxford University Press. https://doi.org/10.1093/oso/9780198520597.001.0001

Thompson, A. P., Aktulga, H. M., Berger, R., Bolintineanu, D. S., Brown, W. M., Crozier, P. S., In 'T Veld, P. J., Kohlmeyer, A., Moore, S. G., Nguyen, T. D., Shan, R., Stevens, M. J., Tranchida, J., Trott, C., & Plimpton, S. J. (2022). LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.*, *271*, 108171. https://doi.org/10.1016/j.cpc.2021.108171

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Ye, B. B., Chen, S., & Wang, Z.-G. (2024). GCMe: Efficient Implementation of the Gaussian Core Model with Smeared Electrostatic Interactions for Molecular Dynamics Simulations of Soft Matter Systems. *J. Chem. Theory Comput.*, acs.jctc.4c00603. https://doi.org/10.1021/acs.jctc.4c00603

Yeh, I.-C., & Berkowitz, M. L. (1999). Ewald summation for systems with slab geometry. *J. Chem. Phys.*, *111*(7), 3155–3162. https://doi.org/10.1063/1.479595