The Journal of Open Source Software

HylleraasMD: Massively parallel hybrid particle-field molecular dynamics in Python

Morten Ledum*[1], Manuel Carrer[1], Samiran Sen[1], Xinmeng Li[1], Michele Cascella[1], and Sigbjørn Løland Bore[2]

1 Department of Chemistry, and Hylleraas Centre for Quantum Molecular Sciences, University of Oslo, PO Box 1033 Blindern, 0315 Oslo, Norway **2** Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States

DOI: 10.21105/joss.04149

Software

- Review
- Repository
- Archive

Editor: Rachel Kurchin

Reviewers:

- @blakeaw
- @yhtang

Submitted: 07 January 2022
Published: 10 February 2022

Summary

Molecular dynamics (MD) is a computational methodology in which the dynamical behavior of systems of interacting atoms and molecules is investigated by integrating the corresponding classical equations of motion. The analysis of the molecular trajectories yields an incredibly powerful computational microscope with atomic resolution. While prominent examples of molecular dynamics involving all-atom models exist, many systems operate on time- and lengths scales too large, precluding the use of such an approach. The intrinsic complexity of biological soft matter systems has necessitated the development of *coarse-grained* (CG) MD models wherein groups of atoms are treated as individual entities. To probe experimentally relevant length- (nm–μm) and time- (ps–ms) scales, further reduction of computational complexity may be warranted through the removal of explicit particle–particle interactions in favor of particle–density field interactions. Such *hybrid particle–field* (hPF) models recast the interactions between particle pairs into a system of free particles interacting with an external potential dependent on the density, in analogy with self-consistent field theories.

HylleraasMD (named after our affiliate centre, the *Hylleraas Centre for Quantum Molecular Sciences*) (HyMD) is a Python package capable of highly parallel hPF-MD simulations of a wide range of surfactants and other biological systems in a CG representation. At present, it is the only open source implementation of the hPF formalism freely available to computational researchers.

Theoretical background

Hybrid particle–field methods are computationally efficient schemes for simulating mesoscale macromolecular assemblies (Milano & Kawakatsu, 2009). Ordinary MD involves, at every integration step, the calculation of computationally expensive double sum over all particle pairs. Despite numerous clever decompositions of the simulation box, which reduces the formal scaling of this (Frenkel & Smit, 2001), it remains the major computational bottleneck. Hybrid particle–field simulations forego this step completely, instead indirectly coupling particles only through an *interaction energy functional* depending on a slowly varying density field. Exploiting the slow time evolution of the density fields, it is possible to employ multiple time-step algorithms which only seldom impart field impulses on individual particles. Beyond this fundamentally more efficient setup of hPF models, the major advantage over particle–particle approaches is the intrinsically *embarrassingly parallel* nature of a large portion of the

*corresponding author

calculations; inter-MPI communication only being necessary whenever the density field is updated. This is traditionally done every tens–several hundreds of MD steps. Accordingly, the hPF methodology has been successfully applied to polymer melts (Wu et al., 2020, 2021), different phases of lipids and surfactants (Bore et al., 2019, 2020; Carrer et al., 2020; De Nicola et al., 2015; De Nicola et al., 2021; Ledum et al., 2020), and charged surfactants and polypeptides (Bore et al., 2018; Kolli et al., 2018; Schäfer et al., 2020).

Statement of need

Elucidating fundamental aspects of the complexity of biological systems often require atomically resolved mesoscale simulations. One crucial example is the large-scale macromolecular self-assembly of lipids and proteins into eukaryotic cell membranes or intracellular organelle structures. Some such systems are computationally accessible today at the CG-MD level, but this is far from routine and not achievable for the broad scientific community. Hybrid particle–field models allow in principle exploration of such systems at near-atomistic resolution, with good chemical accuracy.

Since the hPF scheme was proposed (Milano et al., 2013), two main codes have been used to perform such simulations. (i) OCCAM (Zhao et al., 2012), a proprietary Fortran software developed by Milano and co-workers; and (ii) GALAMOST (Zhu et al., 2013), a CUDA-GPU accelerated C++ code developed by researchers at Jilin University. Unfortunately, neither are open source and freely available to scientists wishing to run hPF simulations of bio- and soft-matter systems.

HyMD is, to date, the only available open-source hPF simulation software. Furthermore, through a recent reformulation of the hPF formalism (Bore & Cascella, 2020), which decouples the computational mesh grid and the length scale of the particle–grid interaction, a new *Hamiltonian* hPF (HhPF) method has emerged. Currently, HyMD constitutes the only software for performing HhPF simulations, open-source or otherwise. This new scheme has a number of advantages over canonical hPF, such as rigorous energy and momentum conservation, rotationally and translationally-invariant forces, and a tunable coarse-graining length scale representing the size extent of particles. Additionally, the new formulation naturally lends itself to calculation in reciprocal space, enabling us to take advantage of highly optimized FFT algorithms.

Features

Apart from a minimal set of high-performance Fortran kernels, the entirety of HyMD is written in Python. This makes extending the software with new functionality easy, enabling fast prototyping of new features. The key components of HyMD include:

- Standard hPF interaction functionals, with the option to specify *any* (local or otherwise) functional, which is automatically handled through symbolic differentiation and numpy vectorization.
- Density filtering (with any user-provided filter function), enabling canonical hPF or HhPF simulations with tunable coarse-graining scale which can be changed *on-the-fly*.
- Optional explicit electrostatic interactions through our custom long-range Particle-Mesh Ewald.
- All standard intramolecular bonded interactions, including stretching, bending, torsional potentials, and combined bending–torsional potentials describing peptide backbone conformations (Bore et al., 2018).

- Topological reconstruction of permanent peptide chain backbone dipoles, enabling realistic protein conformational simulations (Alemani et al., 2010; Bore et al., 2018; Cascella et al., 2008).

To probe experimentally relevant structures, parallelization through mpi4py is used. A 2D *pencil grid* domain decomposition is employed, separating spatial areas of the simulation box across MPI ranks. The highly scalable PFFT (Pippig, 2013) library is used for reciprocal space calculations, as a backend for the PMESH (Feng et al., 2017) package through which we handle the particle–mesh part of the code. A specialized HDF5 file format for MD trajectories (Buyl et al., 2014) is used to enable massively parallel file IO while maintaining an easy structural organization of quantities calculated for storage.

Availability

HyMD is free and open-source, published under a permissive GNU Lesser General Public License v3.0 (LGPLv3). The source code is available at /github.com/Cascella-Group-UiO/HyMD. Documentation, usage guides, and tutorials can be accessed via cascella-group-uio.github.io/HyMD.

Acknowledgements

This work was supported by the Research Council of Norway through the Centre of Excellence *Hylleraas Centre for Quantum Molecular Sciences* (grant number 262695), by the Norwegian Supercomputing Program (NOTUR) (grant number NN4654K), and by the Deutsche Forschungsgemeinschaft (DFG) within the project B5 of the TRR-146 (project number 233630050).

References

Alemani, D., Collu, F., Cascella, M., & Dal Peraro, M. (2010). A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. *J. Chem. Theory Comput.*, *6*(1), 315–324. https://doi.org/10.1021/ct900457z

Bore, S. L., & Cascella, M. (2020). Hamiltonian and alias-free hybrid particle–field molecular dynamics. *J. Chem. Phys.*, *153*(9), 094106. https://doi.org/10.1063/5.0020733

Bore, S. L., Kolli, H. B., De Nicola, A., Byshkin, M., Kawakatsu, T., Milano, G., & Cascella, M. (2020). Hybrid particle-field molecular dynamics under constant pressure. *J. Chem. Phys*, *152*(18), 184908. https://doi.org/10.1063/5.0007445

Bore, S. L., Kolli, H. B., Kawakatsu, T., Milano, G., & Cascella, M. (2019). Mesoscale electrostatics driving particle dynamics in nonhomogeneous dielectrics. *J. Chem. Theory Comput.*, *15*(3), 2033–2041. https://doi.org/10.26434/chemrxiv.7398719

Bore, S. L., Milano, G., & Cascella, M. (2018). Hybrid particle-field model for conformational dynamics of peptide chains. *J. Chem. Theory Comput.*, *14*(2), 1120–1130. https://doi.org/10.1021/acs.jctc.7b01160

Buyl, P. de, Colberg, P. H., & Höfling, F. (2014). H5MD: A structured, efficient, and portable file format for molecular data. *Comput. Phys. Commun.*, *185*(6), 1546–1553. https://doi.org/10.1016/j.cpc.2014.01.018

Carrer, M., Skrbic, T., Bore, S. L., Milano, G., Cascella, M., & Giacometti, A. (2020). Can polarity-inverted surfactants self-assemble in nonpolar solvents? *J. Phys. Chem. B.* https://doi.org/10.26434/chemrxiv.12388772.v1

Cascella, M., Neri, M. A., Carloni, P., & Dal Peraro, M. (2008). Topologically based multipolar reconstruction of electrostatic interactions in multiscale simulations of proteins. *J. Chem. Theory Comput.*, *4*(8), 1378–1385. https://doi.org/10.1021/ct800122x

De Nicola, A., Kawakatsu, T., Rosano, C., Celino, M., Rocco, M., & Milano, G. (2015). Self-assembly of triton x-100 in water solutions: A multiscale simulation study linking mesoscale to atomistic models. *J. Chem. Theory Comput.*, *11*(10), 4959–4971. https://doi.org/10.1021/acs.jctc.5b00485

De Nicola, A., Soares, T. A., Santos, D. E. S., Bore, S. L., Sevink, G. J. A., Cascella, M., & Milano, G. (2021). Aggregation of lipid a variants: A hybrid particle-field model. *Biochim. Biophys. Acta*, *1865*, 129570. https://doi.org/10.1016/j.bbagen.2020.129570

Feng, Y., Hand, N., & biweidai. (2017). *Rainwoodman/pmesh 0.1.33* (Version 0.1.33) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.1051254

Frenkel, D., & Smit, B. (2001). *Understanding molecular simulation: From algorithms to applications* (Vol. 1). Elsevier. https://doi.org/10.1063/1.881812

Kolli, H. B., De Nicola, A., Bore, S. L., Schäfer, K., Diezemann, G., Gauss, J., Kawakatsu, T., Lu, Z.-Y., Zhu, Y.-L., Milano, G., & Cascella, M. (2018). Hybrid particle-field molecular dynamics simulations of charged amphiphiles in an aqueous environment. *J. Chem. Theory Comput.*, *14*(9), 4928–4937. https://doi.org/10.26434/chemrxiv.6264644.v1

Ledum, M., Bore, S. L., & Cascella, M. (2020). Automated determination of hybrid particle-field parameters by machine learning. *Mol. Phys.*, *118*(19-20), e1785571. https://doi.org/10.1080/00268976.2020.1785571

Milano, G., & Kawakatsu, T. (2009). Hybrid particle-field molecular dynamics simulations for dense polymer systems. *J. Chem. Phys.*, *130*(21), 214106. https://doi.org/10.1063/1.3142103

Milano, G., Kawakatsu, T., & De Nicola, A. (2013). A hybrid particle–field molecular dynamics approach: A route toward efficient coarse-grained models for biomembranes. *Phys. Biol.*, *10*(4), 045007. https://doi.org/10.1088/1478-3975/10/4/045007

Pippig, M. (2013). PFFT: An extension of FFTW to massively parallel architectures. *SIAM J. Sci. Comput.*, *35*(3), C213–C236. https://doi.org/10.1137/120885887

Schäfer, K., Kolli, H. B., Christensen, M. K., Bore, S. L., Diezemann, G., Gauss, J., Milano, G., Lund, R., & Cascella, M. (2020). Beyond the molecular packing model: Understanding morphological transitions of charged surfactant micelles. *Angew. Chem. Int. Ed.*, *59*, 18591–18598. https://doi.org/10.1002/anie.202004522

Wu, Z., Alberti, S. A., Schneider, J., & Müller-Plathe, F. (2021). Knotting behaviour of polymer chains in the melt state for soft-core models with and without slip-springs. *J. Phys-Condens. Mat.*, *33*(24), 244001. https://doi.org/10.1088/1361-648X/abef25

Wu, Z., Milano, G., & Müller-Plathe, F. (2020). Combination of hybrid particle-field molecular dynamics and slip-springs for the efficient simulation of coarse-grained polymer models: Static and dynamic properties of polystyrene melts. *J. Chem. Theory Comput.*, *17*(1), 474–487. https://doi.org/10.1021/acs.jctc.0c00954

Zhao, Y., De Nicola, A., Kawakatsu, T., & Milano, G. (2012). Hybrid particle-field molecular dynamics simulations: Parallelization and benchmarks. *J. Comput. Chem.*, *33*(8), 868–880. https://doi.org/10.1002/jcc.22883

Zhu, Y.-L., Liu, H., Li, Z.-W., Qian, H.-J., Milano, G., & Lu, Z.-Y. (2013). GALAMOST: GPU-accelerated large-scale molecular simulation toolkit. *J. Comput. Chem.*, *34*(25), 2197–2211. https://doi.org/10.1002/jcc.23365