The Journal of Open Source Software

VOTCA: multiscale frameworks for quantum and classical simulations in soft matter

Björn Baumeier [1,2¶], **Jens Wehner** [1,2,3], **Nicolas Renaud** [3], **Felipe Zapata Ruiz** [3], **Rene Halver** [4], **Pranav Madhikar** [1,2], **Ruben Gerritsen** [1,2], **Gianluca Tirimbo** [1,2], **Javier Sijen**[1,2], **David Rosenberger** [5], **Joshua S. Brown** [6], **Vivek Sundaram** [1,2], **Jakub Krajniak** [7], **Marvin Bernhardt** [8], **and Christoph Junghans** [9¶]

1 Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands **2** Institute for Complex Molecular Systems, Eindhoven University of Technology, the Netherlands **3** Netherlands eScience Center, the Netherlands **4** Forschungszentrum Jülich, Jülich, Germany **5** Freie Universität Berlin, Berlin, Germany **6** Oak Ridge National Laboratory, Oak Ridge, TN, USA **7** Katholieke Universiteit Leuven, Leuven, Belgium **8** Technische Universität Darmstadt, Darmstadt, Germany **9** Los Alamos National Laboratory, Los Alamos, New Mexico, USA ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor:

Submitted: 27 May 2024
Published: unpublished

Summary

Many physical phenomena in liquids and soft matter are multiscale by nature and can involve processes with quantum and classical degrees of freedom occurring over a vast range of length- and timescales. Examples range from structure formation processes of complex polymers or even polymer blends on the classical side to charge and energy transport and conversion processes involving explicit electronic and therefore quantum information.

The Versatile Object-Oriented Toolkit for Coarse-graining Applications (VOTCA) provides multiscale frameworks built on a comprehensive set of methods for the development of classical coarse-grained potentials (VOTCA-CSG) as well as state-of-the art excited state electronic structure methods based on density-functional and many-body Green's function theories, coupled in mixed quantum-classical models and used in kinetic network models (VOTCA-XTP).

The Journal of Open Source Software

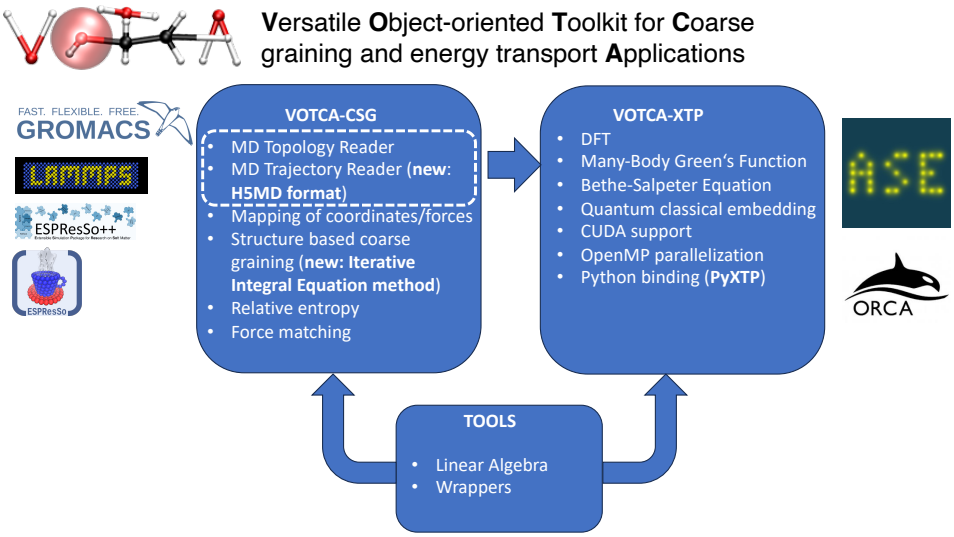

Figure 1: Overview of the differnt VOTCA modules and external interfaces. The trajectory reader of VOTCA-CSG in the dashed line box are reused by VOTCA-XTP.

Statement of need

VOTCA was originally developed as a platform for development and comparison of coarse-graining (CSG) methods. Since the last software publication in 2015 VOTCA-CSG was strengthened by adding more methods, more examples and involving more developers. Many users have used VOTCA to compare different coarse-graining strategies on a neutral ground and, if needed, proceeded with a more specialized package based on the gained insight. Such other coarse-graining packages include, among others, BOCS (Dunn et al., 2018), DeePCG (Zhang et al., 2018), IBIsCO (Karimi-Varzaneh et al., 2011), MagiC (Mirzoev & Lyubartsev, 2013) and OpenMSCG (Peng et al., 2023), some of which are not open-source or specialized in one method. Others have stopped being developed or lack contributions from the greater community. It is difficult to build all-inclusive community package for coarse-graining as it is sometimes hard to consolidate different development styles and constantly changing priorities from sponsors that leave little time for good software engineering practices. In this context we would like to point out that there is a fork of the VOTCA package (**mpip-votca?**) that contains some feature e.g. Kernel-based machine learning methods (Scherer et al., 2020), that has not been merged.

Next to strengthening the coarse-graining functionality of VOTCA, another major development direction taken since 2015 is the addition of explicit quantum-mechanical modules aiming at the simulation of static and dynamic properties of electronically excited states in complex molecular environments using multiscale frameworks. Specifically, the VOTCA-XTP part provides an open-source implementation of many-body Green's functions methods (known as GW-BSE) with the capability of linking this quantum-mechanical description of a molecular system to a classical (microelectrostatic) environment model and to perform rate-based dynamic simulations with kinetic Monte-Carlo. Basic GW-BSE functionality in molecular settings has also more recently been supported in other packages, such as Turbomole (Balasubramani et al., 2020), ADF (Velde et al., 2001), PySCF (Sun et al., 2020), or MOLGW (Bruneval et al., 2016), but these are either commercial or do not provide links to a multiscale framework for complex environments and dynamics.

Coarse-Graining

In the coarse-graining part of VOTCA, VOTCA-CSG, we made a lot of improvements to the inverse Monte Carlo (IMC) method and have added a new iterative approach the so-called iterative integral equation (IIE) method, which are both described in detail below and in reference therein.

Inverse Monte Carlo updates

The inverse Monte-Carlo Method introduced by Lyubartsev and Laaksonen in 1995 (A. P. Lyubartsev & Laaksonen, 1995) is a structure-based coarse graining method, whose goal it is to find an effective pair potential between particles, which reproduces the radial distribution function (RDF) of a reference system (ref) at the coarse grained (CG) resolution. IMC has been part of VOTCA since its first release. In the original implementation the pair potential was determined by iteratively solving a set of linear equations:

$$(A^T A)\Delta U_{ij} = A^T(g_{ij}^{n-1} - g_{ij}^{ref}), \tag{1}$$

where g_{ij} is the RDF between particles i and j, n indicates the iteration counter, ΔU_{ij} is the potential update term, and A and A^T are the Jacobian and its corresponding transpose. The Jacobian A is defined as:

$$A = \frac{\partial g_{ij}}{\partial U_{ij}}, \tag{2}$$

where U_{ij} is the pair potential between particles i and j.

Rosenberger *et al.* (Rosenberger et al., 2016), among others (A. Lyubartsev et al., 2010; Murtola et al., 2007; Tóth, 2003), demonstrated that the standard IMC method can suffer from numerical instabilities and/or slow convergence. Therefore, a Tikhonov regularization for IMC has been implemented in VOTCA. This regularization changes the linear equations according to (Rosenberger et al., 2016):

$$(A^T A + \lambda I)\Delta U_{ij} = A^T(g_{ij}^{n-1} - g_{ij}^{ref}), \tag{3}$$

where λ determines the strength of the regularization and I is the identity matrix. One can perform a singular value decomposition of the Jacobian A to determine an initial value for λ(Rosenberger et al., 2016). As a rule of thumb λ should at least be at the order of the smallest singular values squared.

Iterative Integral Equation method

The iterative integral equation methods are similar to IMC in that they also aim at reconstructing the RDF of a fine-grained reference system with an effective pair potential. The main difference is in the construction of the Jacobian, which is approximated in IIE methods from integral equation theory.(Delbary et al., 2020) For a molecular fluid, where each molecule is mapped to a single bead, using the Ornstein-Zernicke equation and the hypernetted-chain closure, one arrives at the Jacobian inverse with the form of

$$A^{-1} = \frac{dU}{dg} = \frac{1}{\beta}\left(1 - \frac{1}{g} - \mathcal{F}^{-1}\left(\frac{1}{(1+\rho\hat{h})^2}\right)\mathcal{F}\right). \tag{4}$$

Here, $\hat{h}$ is the Fourier transform of $h = g - 1$ and $\mathcal{F}$ is the Fourier operator. This approximate Jacobian works well for systems with single-bead molecule representations with convergence as fast as IMC, whereas in the general case, convergence is half as fast as IMC.(Bernhardt et al., 2023) The costly sampling of the IMC Matrix is not needed, only an RDF which is calculated on twice the range as the potential.(Bernhardt et al., 2021)

Constraints

When using the IMC or IIE methods described above to find pair potentials, that best reproduce a reference RDF, one can use the Gauss-Newton algorithm and formulate the problem of finding a potential update ΔU as a minimization

$$\arg\min_{\Delta U} ||\Delta g + A\Delta U||_2 \tag{5}$$

where $\Delta g = g - g_{\text{target}}$. In that case, additional constraints can be introduced. For example, it is possible to describe the pressure of a system in terms of the RDF g and the pair potential U. From a target pressure and the current pressure at each iteration, a constraint of the form $B\Delta U = d$ can be described and the constraint is enforced by elimination.(Bernhardt et al., 2021)

Electronic Excitations

The most substantial new feature in the VOTCA package is the addition of explicit quantum-mechanical functionalities in the VOTCA-XTP part. The added methods aim at a first-principles-based multiscale modeling of electronically excited states and their dynamics in complex molecular systems. We very briefly describe the three main modules of XTP in the following.

Density-Functional Theory

Excited state calculations require a reference ground state calculation within density-functional theory. VOTCA-XTP provides both an automated interface to the ORCA package (Neese, 2012) and a lightweight internal DFT engine based on atom-centered Gaussian-type orbitals for method developing and testing. It solves the Kohn-Sham Equations for the molecular orbitals $\phi_n^{\text{KS}}(\mathbf{r})$ with orbital energies $\varepsilon_n^{\text{KS}}$

$$\left\{-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{\text{xc}}(\mathbf{r})\right\}\phi_n^{\text{KS}}(\mathbf{r}) = \varepsilon_n^{\text{KS}}\phi_n^{\text{KS}}(\mathbf{r}), \tag{6}$$

where V_{ext} is the external potential, V_{H} the Hartree potential, and V_{xc} the exchange-correlation potential. VOTCA-XTP also contains functionality for projector-based-embedding DFT-in-DFT ground state calculations (Manby et al., 2012), in which a chosen *active* subregion of a molecular system is embedded into an inactive one, reproducing the total energy of the full system ground state exactly.

Many-Body Green's Functions and the Bethe-Salpeter Equation

Using the ground-state reference, many-body Green's functions theory with the GW approximation first calculayes *single-particle excitations* (electron addition or removal) as solutions to the *quasiparticle equations*

$$\left\{-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r})\right\}\phi_n^{\text{QP}}(\mathbf{r}) + \int \Sigma(\mathbf{r},\mathbf{r}',\varepsilon_n^{\text{QP}})\phi_n^{\text{QP}}(\mathbf{r}')d\mathbf{r}' = \varepsilon_n^{\text{QP}}\phi_n^{\text{QP}}(\mathbf{r}). \tag{7}$$

In place of the exchange-correlation potential in Eq.6, the energy-dependent self-energy operator $\Sigma(\mathbf{r},\mathbf{r}',E)$ occurs in the QP equations. This operator is evaluated using the one-body Green's function in quasi-particle approximation

$$G(\mathbf{r},\mathbf{r}',\omega) = \sum_n \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')}{\omega - \varepsilon_n + i0^+\text{sgn}(\varepsilon_n - \mu)} \tag{8}$$

as

$$\Sigma(\mathbf{r},\mathbf{r}',E) = \frac{i}{2\pi}\int e^{-i\omega 0^+} G(\mathbf{r},\mathbf{r}',E-\omega)W(\mathbf{r},\mathbf{r}',\omega)\,d\omega, \tag{9}$$

where W denotes the dynamically screened Coulomb interaction. Assuming that $\phi_n^{\mathrm{QP}} \approx \phi_n^{\mathrm{KS}}$, the quasiparticle energies can be evaluated perturbatively according to

$$\varepsilon_n^{\mathrm{QP}} = \varepsilon_n^{\mathrm{KS}} + \Delta\varepsilon_n^{GW} = \varepsilon_n^{\mathrm{KS}} + \left\langle \phi_n^{\mathrm{KS}} \middle| \Sigma(\varepsilon_n^{\mathrm{QP}}) - V_{\mathrm{xc}} \middle| \phi_n^{\mathrm{KS}} \right\rangle . \tag{10}$$

As the correction $\Delta\varepsilon_n^{GW}$ itself depends on $\varepsilon_n^{\mathrm{QP}}$, Eq.10 needs to be solved self-consistently.

Neutral excitations with a conserved number of electrons can be obtained from the Bethe-Salpeter Equation (BSE) by expressing coupled electron-hole amplitudes of excitation S in a product basis of single-particle orbitals, i.e.,

$$\chi_S(\mathbf{r}_\mathrm{e}, \mathbf{r}_\mathrm{h}) = \sum_v^{\mathrm{occ}} \sum_c^{\mathrm{unocc}} A_{vc}^S \phi_c(\mathbf{r}_\mathrm{e}) \phi_v^*(\mathbf{r}_\mathrm{h}) + B_{vc}^S \phi_v(\mathbf{r}_\mathrm{e}) \phi_c^*(\mathbf{r}_\mathrm{h}), \tag{11}$$

where $\mathbf{r}_\mathrm{e}$ ($\mathbf{r}_\mathrm{h}$) is for the electron (hole) coordinate and A_{vc} (B_{vc}) are the expansion coefficients of the excited state wave function in terms of resonant (anti-resonant) transitions between occupied v and unoccupied c states, respectively. In this basis, the BSE turns into an effective two-particle Hamiltonian problem of the form

$$\begin{pmatrix} \underline{\mathbf{H}}^{\mathrm{res}} & \underline{\mathbf{K}} \\ -\underline{\mathbf{K}} & -\underline{\mathbf{H}}^{\mathrm{res}} \end{pmatrix} \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix} = \Omega_S \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix} . \tag{12}$$

. \end{equation} Specifiaclly, the matrix elements of the blocks $\underline{\mathbf{H}}^{\mathrm{res}}$ and $\underline{\mathbf{K}}$ are calculated as

$$H^{\mathrm{res}}_{vc,v'c'} = D_{vc,v'c'} + \eta K^{\mathrm{x}}_{vc,v'c'} + K^{\mathrm{d}}_{vc,v'c'} \tag{13}$$

$$K_{cv,v'c'} = \eta K^{\mathrm{x}}_{cv,v'c'} + K^{\mathrm{d}}_{cv,v'c'} , \tag{14}$$

with

$$D_{vc,v'c'} = (\varepsilon_c - \varepsilon_v)\delta_{vv'}\delta_{cc'}, \tag{15}$$

$$K^{\mathrm{x}}_{vc,v'c'} = \iint \phi_c^*(\mathbf{r}_\mathrm{e}) \phi_v(\mathbf{r}_\mathrm{e}) v_\mathrm{C}(\mathbf{r}_\mathrm{e}, \mathbf{r}_\mathrm{h}) \phi_{c'}(\mathbf{r}_\mathrm{h}) \phi_{v'}^*(\mathbf{r}_\mathrm{h}) d^3\mathbf{r}_\mathrm{e} d^3\mathbf{r}_\mathrm{h} \tag{16}$$

$$K^{\mathrm{d}}_{vc,v'c'} = -\iint \phi_c^*(\mathbf{r}_\mathrm{e}) \phi_{c'}(\mathbf{r}_\mathrm{e}) W(\mathbf{r}_\mathrm{e}, \mathbf{r}_\mathrm{h}, \omega = 0) \phi_v(\mathbf{r}_\mathrm{h}) \phi_{v'}^*(\mathbf{r}_\mathrm{h}) d^3\mathbf{r}_\mathrm{e} d^3\mathbf{r}_\mathrm{h} . \tag{17}$$

and $\eta = 2$ ($\eta = 0$) for singlet (triplet) excitations. Here, K^{x} is the repulsive exchange interaction originating from the bare Coulomb term v_C, while the direct interaction K^{d} contains the attractive, but screened, interaction W between electron and hole, causing the binding of the electron-hole pair. In Eq.17 it is assumed that the dynamic properties of $W(\omega)$ are negligible, and the computationally less demanding static approximation $\omega = 0$ is employed.

Quantum-Classical Embedding

Polarization effects of an environment can have significant impact on electronic excitations. As polarization effects are long-ranged accounting for them requires the treatment of large systems which is infeasible with explicit quantum methods such as DFT-GW-BSE. Instead, the system is split into a small part with to electronically active subsystem to be treated at quantum (QM) level and a large environment part in which electrostatic and polarization effects are accounted for in classical models (MM). In VOTCA-XTP the QM/MM scheme employs distributed atomic multipole representations for molecules in the MM region, which allows treatment of both the effects of static electric fields and the polarization response as a self-consistent reaction field. Specifically, this classical MM energy for the system is evaluated as

$$E_{\mathrm{MM}} = \frac{1}{2} \sum_{\substack{A,B \\ A \neq B}} \sum_{a \in A} \sum_{b \in B} \sum_{tu} (Q_t^a + \Delta Q_t^a) T_{tu}^{ab} Q_u^b, \tag{18}$$

where A and B indicate individual molecules in the system, a and b atoms in the respective molecules, Q_t^a are the static atomic multipole moments of rank t associated to atom a, and T_{tu}^{ab} is the tensor describing the interactions between the multipoles moments Q_t^a and Q_u^b (Stone, 2005). The induced moments ΔQ_t^a are generated by the electric field created by moments t' of atom $a' \neq a$ in molecule A and the one generated by the moment u of atom b in molecule B:

$$\Delta Q_t^a = -\sum_{\substack{A,B\in\mathcal{S}\\A\neq B}}\sum_{b\in B}\sum_{\substack{a'\in A\\a'\neq a}}\sum_{tt'u}\alpha_{tt'}^{aa'}T_{t'u}^{a'b}(Q_u^b+\Delta Q_u^b), \tag{19}$$

with $\alpha_{tt'}^{aa'}$ the isotropic atomic polarizability on each site. To avoid effects of spurious overpolarization, a damped version of the interaction tensor (Thole damping (Stone, 2005)) is used. Then, the static and induced multipoles in the MM region also interact with the electron density in QM region via an additional external potential to Eq.6. At the same time, the explicit electrostatic field from the QM density is included in polarizing the MM region.

Code Structure

For the last couple of years, we have also focused on code hardening and the introduction of better software engineering practices. Original VOTCA was designed as modules in separate repositories, but as many other projects, this turned out to be quite cumbersome hence we switched to a mono-repo. With recent performance improvements in the git tools, the benefits of a single repository by far out-weigh the downside of the very complex workflow of multiple repositories. The module structure still exists in the source code.

Additionally, we have added continuous integration testing through GitHub action for 50+ different compiler and operating system combinations. The also perform continous deployment to the GitHub Docker registry. And releases get rolled into all major linux distributions, HomeBrew, Spack and FreeBSD.

Code Modernization

We did a lot of code refactoring and bumped the C++ standard to 17. We also modernized our usage of CMake and switched to a mostly target-base scheme. An attempt to port our particle structure on top of Cabana (Slattery et al., 2022) was made, due to incompatibilties between Kokkos (Trott et al., 2022) and Eigen, we will delay this effort.

Updates in VOTCA-CSG

The particle and molecule data structure were refactored, and we add support of the H5MD format, which is described below in details.

H5MD support

The recent version of VOTCA supports the `H5MD` (de Buyl et al., 2014) file format, which internally uses `HDF5` (The HDF Group, n.d.) storage. This is a very fast and scalable method for storing molecular trajectories, already implemented in simulation packages such as `LAMMPS`, `ESPResSo++`, and `ESPResSo`. VOTCA recognizes the trajectory file format by the extension. In the case of H5MD, it expects a `.h5` extension. Following the H5MD concepts, the particle trajectories are organized in the `particles` container. This container can handle multiple subsets of the studied system. Therefore, we must define `h5md_particle_group` in the XML topology file to declare which subset of particles to use. The reader handles both coordinates (if present), forces, and velocities.

Design of VOTCA-XTP

Data structures related to atomistic properties (topology, molecules, segments, fragments, atoms) in XTP are reused or build upon those of CSG. Linear algebra related structures and functionalities are handled by Eigen (Guennebaud et al., 2010) which can be accelerated by internally calling the Intel Math Kernel Library (*Intel Math Kernel Library. Reference Manual*, 2009). Exchange-correlation functionals are provided by the Library of eXchange-Correlation (`LIBXC`) functionals (Lehtola et al., 2018), while `libint` (Valeev, 2024) and `libecpint` (Shaw & Hill, 2017) are used for the evaluation of molecular integrals of many-body operators over Gaussian functions.

VOTCA–XTP is designed as a library, which is linked to very thin executables. These executables provide a variety of calculators by adding keywords on the command line. Virtual interfaces and factory patterns make the addition of new calculators simple. The same architecture is used for external DFT and MD codes, making VOTCA–XTP easily extensible. Lower-level data structures make use of template metaprogramming to support a variety of data types. VOTCA-XTP provides different functionalities in three types of *calculator* classes:

- a collection of tools that do not require information of a mapped MD trajectory, including a specific DFT-GW-BSE calculator in `tools` callable by `xtp_tools`
- analysis and not-high-throughput applications that require a mapped MD trajectory in `calculators` callable by `xtp_run`
- high-throughput, high-performance applications that require a mapped MD trajectory in `jobcalculators` callable by `xtp_parallel`

In general, VOTCA-XTP uses shared-memory parallelization in the heavy calculations involving the quantum methods, with the possibility to seamlessly offload matrix-matrix and matrix-vector operations to GPU via `CUDA`.

PyXTP

The PyXTP python package distributed with VOTCA, contains python bindings to the main functionalities of VOTCA-XTP. These python bindings were created using pybind11 (Jakob et al., 2017) and provide users with a dedicated Atomistic Simulation Environment (ASE) (Larsen et al., 2017) calculator. The use of this calculator not only facilitates the adoption of VOTCA-XTP by non-experts users, but they also allow integrating VOTCA-XTP in the broader ASE ecosystem.

The following snippet of code illustrate the use of PyXTP. This small code optimize the geometry of a CO molecule in the first excited singlet states. As seen in the code, the XTP calculator is used to compute the forces on the nuclei while the geometry optimization itself is driven by ASE functionalities.

```
from pyxtp import xtp
from ase.io import write
from ase.build import molecule
from ase.optimize import QuasiNewton

# create a distorted CO molecule
atoms = molecule('CO')
atoms.rattle()

# instantiate the calculator
calc = xtp(nthreads=2)

# select the state for which to compute the forces
calc.select_force(energy='singlets', level=0, dynamic=False)
```

```
# this allows to change all options
calc.options.dftpackage.functional = 'PBE'
calc.options.dftpackage.basisset = 'def2-svp'
calc.options.dftpackage.auxbasisset = 'aux-def2-svp'

# set up the logger
calc.options.logging_file = 'CO_forces.log'

# set the calculator
atoms.calc = calc

# optimize the geometry
dyn = QuasiNewton(atoms, trajectory='test.traj')
dyn.run(fmax=0.01)
write('final.xyz', atoms)
```

Acknowledgements

We acknowledge contributions from Brigitta Sipocz, Syrtis Major, and Semyeong Oh, and support from Kathryn Johnston during the genesis of this project. We acknowledge support by the Innovational Research Incentives Scheme Vidi of the Netherlands Organisation for Scientific Research (NWO) with project number 723.016.002. Funding is also provided by NWO and the Netherlands eScience Center for funding through project number 027.017.G15, within the Joint CSER and eScience program for Energy Research (JCER 2017).

References

Balasubramani, S. G., Chen, G. P., Coriani, S., Diedenhofen, M., Frank, M. S., Franzke, Y. J., Furche, F., Grotjahn, R., Harding, M. E., Hättig, C., Hellweg, A., Helmich-Paris, B., Holzer, C., Huniar, U., Kaupp, M., Marefat Khah, A., Karbalaei Khani, S., Müller, T., Mack, F., … Yu, J. M. (2020). TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *The Journal of Chemical Physics*, *152*(18), 184107. https://doi.org/10.1063/5.0004635

Bernhardt, M. P., Hanke, M., & Van Der Vegt, N. F. A. (2023). Stability, speed, and constraints for structural coarse-graining in VOTCA. *Journal of Chemical Theory and Computation*, *19*(2), 580–595. https://doi.org/10.1021/acs.jctc.2c00665

Bernhardt, M. P., Hanke, M., & Vegt, N. F. A. van der. (2021). Iterative integral equation methods for structural coarse-graining. *The Journal of Chemical Physics*, *154*(8), 084118. https://doi.org/10.1063/5.0038633

Bruneval, F., Rangel, T., Hamed, S. M., Shao, M., Yang, C., & Neaton, J. B. (2016). Molgw 1: Many-body perturbation theory software for atoms, molecules, and clusters. *Computer Physics Communications*, *208*, 149–161. https://doi.org/https://doi.org/10.1016/j.cpc.2016.06.019

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

Delbary, F., Hanke, M., & Ivanizki, D. (2020). A generalized newton iteration for computing the solution of the inverse henderson problem. *Inverse Problems in Science and Engineering*, *28*(8), 1166–1190. https://doi.org/10.1080/17415977.2019.1710504

Dunn, N. J. H., Lebold, K. M., DeLyser, M. R., Rudzinski, J. F., & Noid, W. G. (2018). BOCS: Bottom-up open-source coarse-graining software. *The Journal of Physical Chemistry B*, *122*(13), 3363–3377. https://doi.org/10.1021/acs.jpcb.7b09993

Guennebaud, G., Jacob, B., & others. (2010). *Eigen v3*. http://eigen.tuxfamily.org.

Intel Math Kernel Library. Reference Manual. (2009). [Computer software]. Intel Corporation.

Jakob, W., Rhinelander, J., & Moldovan, D. (2017). *pybind11 – seamless operability between c++11 and python*.

Karimi-Varzaneh, H. A., Qian, H.-J., Chen, X., Carbone, P., & Müller-Plathe, F. (2011). IBIsCO: A molecular dynamics simulation package for coarse-grained simulation. *Journal of Computational Chemistry*, *32*(7), 1475–1487. https://doi.org/https://doi.org/10.1002/jcc.21717

Larsen, A. H., Mortensen, J. J., Blomqvist, J., Castelli, I. E., Christensen, R., Dułak, M., Friis, J., Groves, M. N., Hammer, B., Hargus, C., Hermes, E. D., Jennings, P. C., Jensen, P. B., Kermode, J., Kitchin, J. R., Kolsbjerg, E. L., Kubal, J., Kaasbjerg, K., Lysgaard, S., … Jacobsen, K. W. (2017). The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, *29*(27), 273002. http://stacks.iop.org/0953-8984/29/i=27/a=273002

Lehtola, S., Steigemann, C., Oliveira, M. J. T., & Marques, M. A. L. (2018). Recent developments in LIBXC — A comprehensive library of functionals for density functional theory. *SoftwareX*, *7*, 1–5. https://doi.org/10.1016/j.softx.2017.11.002

Lyubartsev, A. P., & Laaksonen, A. (1995). Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Physical Review E*, *52*(4), 3730–3737. https://doi.org/10.1103/PhysRevE.52.3730

Lyubartsev, A., Mirzoev, A., Chen, L., & Laaksonen, A. (2010). Systematic coarse-graining of molecular models by the Newton inversion method. *Faraday Discuss.*, *144*, 43–56. https://doi.org/10.1039/B901511F

Manby, F. R., Stella, M., Goodpaster, J. D., & Miller, T. F. (2012). A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.*, *8*, 2564–2568.

Mirzoev, A., & Lyubartsev, A. P. (2013). MagiC: Software package for multiscale modeling. *Journal of Chemical Theory and Computation*, *9*(3), 1512–1520. https://doi.org/10.1021/ct301019v

Murtola, T., Falck, E., Karttunen, M., & Vattulainen, I. (2007). Coarse-grained model for phospholipid/cholesterol bilayer employing inverse Monte Carlo with thermodynamic constraints. *The Journal of Chemical Physics*, *126*(7), 075101. https://doi.org/10.1063/1.2646614

Neese, F. (2012). The ORCA program system. *WIREs Computational Molecular Science*, *2*(1), 73–78. https://doi.org/10.1002/wcms.81

Peng, Y., Pak, A. J., Durumeric, A. E. P., Sahrmann, P. G., Mani, S., Jin, J., Loose, T. D., Beiter, J., & Voth, G. A. (2023). OpenMSCG: A software tool for bottom-up coarse-graining. *The Journal of Physical Chemistry B*, *127*(40), 8537–8550. https://doi.org/10.1021/acs.jpcb.3c04473

Rosenberger, D., Hanke, M., & Vegt, N. F. A. van der. (2016). Comparison of iterative inverse coarse-graining methods. *The European Physical Journal Special Topics*, *225*(8-9), 1323–1345. https://doi.org/10.1140/epjst/e2016-60120-1

Scherer, C., Scheid, R., Andrienko, D., & Bereau, T. (2020). Kernel-based machine learning for efficient simulations of molecular liquids. *Journal of Chemical Theory and Computation*, *16*(5), 3194–3204. https://doi.org/10.1021/acs.jctc.9b01256

Shaw, R. A., & Hill, J. G. (2017). Prescreening and efficiency in the evaluation of integrals over ab initio effective core potentials. *The Journal of Chemical Physics*, *147*(7), 074108. https://doi.org/10.1063/1.4986887

Slattery, S., Reeve, S. T., Junghans, C., Lebrun-Grandié, D., Bird, R., Chen, G., Fogerty, S., Qiu, Y., Schulz, S., Scheinberg, A., Isner, A., Chong, K., Moore, S., Germann, T., Belak, J., & Mniszewski, S. (2022). Cabana: A performance portable library for particle-based simulations. *Journal of Open Source Software*, *7*(72), 4115. https://doi.org/10.21105/joss.04115

Stone, A. J. (2005). Distributed Multipole Analysis: Stability for Large Basis Sets. *Journal of Chemical Theory and Computation*, *1*(6), 1128–1132. https://doi.org/10.1021/ct050190+

Sun, Q., Zhang, X., Banerjee, S., Bao, P., Barbry, M., Blunt, N. S., Bogdanov, N. A., Booth, G. H., Chen, J., Cui, Z.-H., Eriksen, J. J., Gao, Y., Guo, S., Hermann, J., Hermes, M. R., Koh, K., Koval, P., Lehtola, S., Li, Z., … Chan, G. K.-L. (2020). Recent developments in the PySCF program package. *The Journal of Chemical Physics*, *153*(2), 024109. https://doi.org/10.1063/5.0006074

The HDF Group. (n.d.). *Hierarchical Data Format, version 5*. https://github.com/HDFGroup/hdf5

Tóth, G. (2003). An iterative scheme to derive pair potentials from structure factors and its application to liquid mercury. *The Journal of Chemical Physics*, *118*(9). https://doi.org/10.1063/1.1543142

Trott, C. R., Lebrun-Grandié, D., Arndt, D., Ciesko, J., Dang, V., Ellingwood, N., Gayatri, R., Harvey, E., Hollman, D. S., Ibanez, D., Liber, N., Madsen, J., Miles, J., Poliakoff, D., Powell, A., Rajamanickam, S., Simberg, M., Sunderland, D., Turcksin, B., & Wilke, J. (2022). Kokkos 3: Programming model extensions for the exascale era. *IEEE Transactions on Parallel and Distributed Systems*, *33*(4), 805–817. https://doi.org/10.1109/TPDS.2021.3097283

Valeev, E. F. (2024). *Libint: A library for the evaluation of molecular integrals of many-body operators over gaussian functions*. http://libint.valeyev.net/.

Velde, G. te, Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., Gisbergen, S. J. A. van, Snijders, J. G., & Ziegler, T. (2001). Chemistry with ADF. *J. Comput. Chem.*, *22*(9), 931–967. https://doi.org/10.1002/jcc.1056

Zhang, L., Han, J., Wang, H., Car, R., & E, W. (2018). DeePCG: Constructing coarse-grained models via deep neural networks. *The Journal of Chemical Physics*, *149*(3), 034101. https://doi.org/10.1063/1.5027645