Molecular Latent Space Simulators

Hythem Sidky, a Wei Chen, b and Andrew L. Ferguson a,d

Small integration time steps limit molecular dynamics (MD) simulations to millisecond time scales. Markov state models (MSMs) and equation-free approaches learn low-dimensional kinetic models from MD simulation data by performing configurational or dynamical coarse-graining of the state space. The learned kinetic models enable the efficient generation of dynamical trajectories over vastly longer time scales than are accessible by MD, but the discretization of configurational space and/or absence of a means to reconstruct molecular configurations precludes the generation of continuous all-atom molecular trajectories. We propose latent space simulators (LSS) to learn kinetic models for continuous all-atom simulation trajectories by training three deep learning networks to (i) learn the slow collective variables of the molecular system, (ii) propagate the system dynamics within this slow latent space, and (iii) generatively reconstruct molecular configurations. We demonstrate the approach in an application to Trp-cage miniprotein to produce novel ultra-long synthetic folding trajectories that accurately reproduce all-atom molecular structure, thermodynamics, and kinetics at six orders of magnitude lower cost than MD. The dramatically lower cost of trajectory generation enables greatly improved sampling and greatly reduced statistical uncertainties in estimated thermodynamic averages and kinetic rates.

1 Introduction

Molecular dynamics (MD) simulates the dynamical evolution of molecular systems by numerically integrating the classical equations of motion. Modern computer hardware and efficient and scalable simulation algorithms have enabled the simulation of billion and trillion-atom systems. Advancing the barrier in time scale has proven far more challenging. Stability of the numerical integration requires time steps on the order of femtoseconds commensurate with the fastest atomic motions, which limits simulations to microseconds on commodity processors and milliseconds on special purpose hardware. Enhanced sampling techniques apply accelerating biases and analytical corrections to recover thermodynamic averages, but – except in special cases and the limit of small biases – no analogous approaches exist to recover unbiased dynamical trajectories from biased simulations.

The MD algorithm propagates a molecular configuration \( \mathbf{x}_t \) at time \( t \) to \( \mathbf{x}_{t+\tau} \) via transition densities \( p_\tau(\mathbf{x}_{t+\tau}|\mathbf{x}_t) \). Assuming ergodicity, the probability density over microstates converges to the stationary distribution as \( \lim_{\tau \to \infty} q_\pi(\mathbf{x}) = \pi(\mathbf{x}) \). Breaking the time scale barrier requires a surrogate model for \( p_\tau(\mathbf{x}_{t+\tau}|\mathbf{x}_t) \) that can be more efficiently evaluated and with larger time steps than MD. Accurately approximating this propagator in the high-dimensional \( N \)-atom configurational space \( \mathbf{x} \in \mathbb{R}^{3N} \) is intractable.

In general, for sufficiently large \( \tau \) there is an emergent low-dimensional simplicity that admits accurate modeling of the dynamics by a low-dimensional propagator \( p_\tau(\psi_{t+\tau}|\psi_t) \) within a latent space \( \psi \in \mathbb{R}^{m<3N} \). The relation between MD and latent space dynamics can be represented as

\[
\begin{align*}
\mathbf{x}_t & \xrightarrow{E} \psi_t \\
\text{MD} \downarrow & \quad \downarrow p \\
\mathbf{x}_{t+\tau} & \xleftarrow{D} \psi_{t+\tau}
\end{align*}
\]

This scheme defines three learning problems: (i) encoding \( E \) of molecular configurations \( \mathbf{x} \) to the latent space \( \psi \), (ii) propagation \( P \) of the latent space dynamics according to transition densities \( p_\tau(\psi_{t+\tau}|\psi_t) \), and (iii) decoding (or generating) \( D \) of molecular configurations from the latent space.

Markov state models (MSM) and the equation-free approach of Kevrekidis and co-workers respectively employ configurational and dynamical coarse graining to parameterize low-dimensional propagators, but both methods lack molecular decoders. Recently, numerous deep learning approaches have been proposed to learn \( E, P, \) and \( D \) from MD trajectories, including time-lagged autoencoders, time-lagged variational autoencoders, and time-lagged autoencoders with propagators.

---

a Pritzker School of Molecular Engineering, University of Chicago, Chicago, USA. E-mail: andrewferguson@uchicago.edu

b Department of Physics, University of Illinois at Urbana-Champaign, Urbana, USA.
Training these networks requires a time-lagged reconstruction term \( |x_{t+\tau} - D \circ P \circ E(x_t)| \) within the loss, which can cause the network to fail to approximate the true slow modes\(^{29}\). Further, time-lagged autoencoders and time-lagged variational autoencoders do not learn valid propagators\(^{30}\), and the inherent stochasticity of MD appears to frustrate learning of the propagator and decoder in time-lagged autoencoders with propagators\(^{19}\).

Deep generative MSMs (DeepGenMSM) simultaneously learn a fuzzy encoding to metastable states and “landing probabilities” to decode molecular configurations\(^{40}\). The method computes a proper propagator and generatively decodes novel molecular structures, but—as with all MSM-based approaches—it configurationally discretizes the latent space and relies on the definition of long-lived metastable states.

In this work, we propose molecular latent space simulators (LSS) as a means to train kinetic models over limited MD simulation data that are capable of producing novel all-atom molecular trajectories at orders of magnitude lower cost. The LSS can be conceived as means to augment conventional MD by distilling a trajectory at orders of magnitude lower cost. The LSS is based on three deep learning networks independently trained to (i) learn an encoding \( E \) of any all-atom structural, thermodynamic, or kinetic observable. Based on this encoding \( E \), one can (ii) simultaneously learn a state-free propagator \( \mathcal{F} \) and (iii) a decoding \( D \) to decode molecular configurations from the latent space to molecular configurations\(^{19}\). The method computes a proper propagator and generatively decodes novel molecular structures, but—as with all MSM-based approaches—it configurationally discretizes the latent space and relies on the definition of long-lived metastable states.

The LSS is based on three deep learning networks independently trained to (i) learn an encoding \( E \) into a latent space of slow variables using state-free reversible VAMPnets (SRV)\(^{31}\), (ii) learn a propagator \( \mathcal{F} \) to evolve the system dynamics within this latent space using mixture density networks (MDN)\(^{42,43}\), and (iii) learn a decoding \( D \) from the latent space to molecular configurations using a conditional Wasserstein generative adversarial network (cWGAN)\(^{44}\). Separation of the learning problems in this manner makes training and deployment of the LSS modular and simple. The stochastic nature of the MDN propagator means that the trained kinetic model generates novel trajectories and does not simply recapitulate copies of the training data. The approach is distinguished from MSM-based approaches in that it requires no discretization into metastable states\(^{19,40}\). The continuous formulation of the propagator in the slow latent space shares commonalities with the equation-free approach\(^{29,31}\) but we eschew parameterizing a stochastic differential equation in favor of a simple and efficient deep learning approach, and also equip our simulator with a generative molecular decoder.

### 2 Methods

A schematic diagram of the LSS and the three deep networks of which it is comprised is presented in Fig. 1. We describe each of the three component networks in turn and then describe LSS training and deployment.

#### 2.1 Encoder: State-free Reversible VAMPnets

The transfer operator \( \mathcal{F} \) at a lag time \( \tau \) is the propagator of probability distributions over microstates with respect to the equilibrium density \( u(x) = q(x)/\pi(x) \) under transition densities \( p_c(x_{t+\tau} | x_t) \). For sufficiently large \( \tau \) the dynamics may be approximated as Markovian so \( p_c(x_{t+\tau} | x_t) \) is time homogenous,

\[
u_{t+\tau}(x) = \mathcal{F} \circ u_t(x) = \frac{1}{\pi(x)} \int dy \, p_c(x | y) u_t(y) \pi(y).
\]

In equilibrium systems obeying detailed balance \( \pi(x) p_c(y | x) = \pi(y) p_c(x | y) \), the transfer operator \( \mathcal{F} \) is identical to the Koopman operator, self-adjoint with respect to \( \langle a | b \rangle = \int a(x) b(x) \pi(x) dx \), and possesses a complete orthonormal set of eigenfunctions \( \{ \psi_i(x) \} \) with real eigenvalues \( 1 = \lambda_0 > \lambda_1 \geq \lambda_2 \geq \ldots \geq |\lambda| \). The eigenfunctions \( \psi_i(x) \) are defined as

\[
\mathcal{F} \circ \psi_i(x) = \lambda_i \psi_i(x), \quad \langle \psi_i | \psi_j \rangle = \delta_{ij}.
\]

The pair \( \{ \psi_i(x) \} \) corresponds to the equilibrium distribution at \( t \to \infty \) and the remainder to a hierarchy of increasingly quicker relaxing processes with implied time scales \( t_i = -\tau/|\lambda_i| \). The evolution of \( u_t(x) \) after \( k \) applications of \( \mathcal{F} \) is expressed in this basis as,

\[
u_{t+k\tau}(x) = \mathcal{F}^k \circ u_t(x) = \sum_i (\psi_i, u_t)_\pi \exp \left( -\frac{k \tau}{t_i} \right) \psi_i(x).
\]

The variational approach to conformational dynamics (VAC) defines a variational principle to approximate these eigenfunctions as \( \psi_i(x) = \sum_j s_{ij} \chi_j(x) \) within a basis \( \{ \chi_j \} \) by solving for optimal expansion coefficients \( s_{ij} \) \(^{41,47,48}\) SRVs \(^{41}\)-- themselves based on VAMPnets, a deep learning-based method for MSM construction\(^{27}\) and closely related to extended dynamic mode decomposition with dictionary learning\(^{50}\)-- employ deep canonical correlation analysis (DCCA)\(^{51}\) to learn both the optimal expansion coefficients and optimal basis functions as nonlinear transformations of the (featurized) molecular coordinates. This is achieved by training twin-lobed deep neural networks to minimize a VAMP\(r\) loss function \( \mathcal{L}_{SRV} = -\sum_m \lambda_m \). SRVs trained over MD trajectories furnish an encoding \( E \) (Eqn. 1) into a \( m \)-dimensional latent space spanned by \( \{ \psi_i(x) \}_{i=1}^m \) where \( m \) is determined by a gap in the eigenvalue spectrum. This spectral encoding into the leading modes of \( \mathcal{F} \) neglects fast processes with implied timescales \( t_i \ll \tau \) (Eqn. 4) and is an optimal parameterization of the system for a low-dimensional long-time propagator\(^{19}\).

#### 2.2 Propagator: Mixture Density Networks

At sufficiently large \( \tau \) the latent space \( \psi(x) = \{ \psi_i(x) \}_{i=1}^m \) supports an autonomous dynamical system in the leading modes of \( \mathcal{F} \). We train MDNs to learn transition densities \( p_c(\psi_{t+\tau} | \psi_t) \) from MD trajectories projected in the latent space. MDNs combine deep neural networks with mixture density models to overcome poor performance of standard networks in learning multimodal distributions\(^{42,43}\). Transition densities are approximated as a linear combination of \( C \) kernels,

\[
p_c(\psi_{t+\tau} | \psi_t) = \sum_{c=1}^C \alpha_c(\psi_t) \phi_c(\psi_{t+\tau}; \mu_c(\psi_t), \sigma_c(\psi_t)),
\]

where we choose \( \phi_c \) to be \( m \)-dimensional Gaussians. The \( \psi_t \)-dependent Gaussian means \( \mu_c(\psi_t) \), variances \( \sigma_c(\psi_t) \), and linear mixing coefficients \( \alpha_c(\psi_t) \) are learned by a deep feedforward neural network that minimizes the loss function \( \mathcal{L}_{MDN} = \)
$-\sum_{i} \ln p_{t}^{\gamma}(\psi_{i+1}^{\gamma}|\psi_{i}^{\gamma})$, where $\gamma$ indexes pairs of time-lagged training data observations. The normalization $\sum_{\gamma} \alpha_{\gamma}(\psi_{i}) = 1$ is enforced by softmax activations and the $\mu_{t}(\psi_{i})$ bounded using sigmoid activations.

The trained MDN defines the latent space propagator $P$ (Eqn. 1) and we sample transition densities $p_{t}(\psi_{i+1}^{\gamma}|\psi_{i})$ to advance the system in time (Fig. 1). Propagation is conducted entirely within the latent space and does not require recurrent decoding and encoding to the molecular representations that can lead to accumulation of errors and numerical instability\cite{10,52}. The transition densities are learned from the statistics transitions in the training data and new trajectories are generated by sampling from these transition densities. These new trajectories therefore represent novel dynamical pathways over the latent space and are not simply recapitulations or approximate copies of those in the training data. Successful MDN training is contingent on the low-dimensional and Markovian nature of the latent space dynamics at large $\tau$ discovered by the SRVs.

2.3 Decoder: Conditional Wasserstein GAN

Generative adversarial networks are a leading neural network architecture for generative modeling\cite{53}. We employ a cWGAN\cite{44,54} to decode from the latent space $\psi$ to molecular configurations $x$ by performing adversarial training between a generator $G(z)$ that outputs molecular configurations from inputs $z \sim \mathcal{P}_{\gamma}(z)$ and a critic $\mathcal{C}(x)$ that evaluates the quality of a molecular configuration $x$. The networks are jointly trained to minimize a loss function based on the Wasserstein (Earth Mover’s) distance,

$$\mathcal{L}_{\text{WGAN}} = \max_{w \in \mathcal{W}} \mathbb{E}_{x \sim \mathcal{P}_{\gamma}(x)}[\mathcal{C}(x)] - \mathbb{E}_{z \sim \mathcal{P}_{\gamma}(z)}[\mathcal{C}(G(z))],$$

where $\mathcal{P}_{\gamma}(x)$ is the distribution over molecular configurations observed in the MD training trajectory and $\{\mathcal{C}_{i}\}_{i \in \mathcal{W}}$ is a family of $K$-Lipschitz functions enforced through a gradient penalty\cite{44,54}. To generate molecular configurations consistent with particular states in the latent space we pass $\psi$ as a conditioning variable to $G$ and $\mathcal{C}\mathbb{E}$ and drive the generator with $d$-dimensional Gaussian noise $\mathcal{P}_{\gamma}(z) = \mathcal{N}(0,1) \in \mathbb{R}^{d}$. The noise enables $G$ to generate multiple molecular configurations consistent with each latent space location. We train the cWGAN over $(\mathcal{X}, \mathcal{Y})$ pairs by encoding each frame $\gamma$ of the MD training trajectory into the latent space using the SRV. The trained cWGAN decoder $D$ (Eqn. 1) generates molecular configurations from the latent space trajectory produced by the propagator (Fig. 1).

3 Results and Discussion

3.1 4-well potential

We validate the LSS in an application to a 1D four-well potential\cite{23}$V(x)=2(1.8x+0.8e^{-80x^2}+0.2e^{-80(x-0.5)^2}+0.5e^{-40(x+0.5)^2})$ for which analytical solutions are available. In this simple 1D system we construct the propagator directly in $x=\psi \in \mathbb{R}^{1}$, so encoding and decoding are unnecessary and this test validates that the MDN can learn transition densities $p_{t}(x_{i+1}|x_{i})$ to accurately reproduce the system thermodynamics and kinetics. We generate a $5 \times 10^{6}$ time step Brownian dynamics trajectory in a dimensionless gauge with diffusivity $D=s_{B}T=1000$ and a time step $\Delta t=0.001$\cite{55}. A MDN was trained using Adam\cite{57}, with early stopping over the $[0,1]$ scaled trajectory at a lag time of $\tau=100$, with $\mathcal{C}=8$ Gaussian kernels, and two hidden layers of 100 neurons with ReLU activations\cite{58}. The trained MDN was used to generate a $5 \times 10^{6}$ step trajectory of the same length as the training data. Analytical transition densities were computed by partitioning the domain into 100 equal bins and defining the probability of moving from bin $i$ to bin $j$ as $p(i|j) = C_{i}e^{-|V_{i}-V_{j}|}$ for $|i-j| \leq 1$, where $V_{i}$ is the potential at the center of bin $i$ and $C_{i}$ normalizes the total transition probability of bin $i$\cite{23}.

The Brownian dynamics and synthetic LSS stationary distributions are in quantitative agreement with the analytical solution for the stationary density (Fig. 2b) and show very similar ki-
nentic behaviors in their transitions between the four metastable wells (Fig. 2b). This agreement is due to the excellent correspondence between the analytical and learned transition densities (Fig. 2c,d).

Fig. 2 Validation of the LSS in a 1D four-well potential. The MDN propagator predicts (a) a stationary distribution, (b) kinetic transitions, and (c,d) transition densities in excellent accord with analytical and Brownian dynamics results.

### 3.2 Trp-cage miniprotein

We train our LSS over the 208 µs all-atom simulation of the 20-residue TC10b K8A mutant of the Trp-cage mini-protein performed by D.E. Shaw Research. Generation of these MD trajectories would require ~2.5 days (2 million CPU-h) on the special purpose Anton-2 supercomputer or ~6 months on a commodity GPU card.

The SRV encoder was trained over a feature extraction of the trajectory employing backbone and sidechain torsions and Cα pairwise distances as informative and roto-translationally invariant descriptors. We trained a SRV with two hidden layers with 100 neurons, tanh activations, and batch normalization using Adam with a batch size of 200,000, learning rate of 0.01, and early stopping based on the validation VAMP-2 score. A time-averaged latent space was 100 ns was chosen based on convergence of the transfer operator eigenvalues, and a $m=3$-dimensional latent space encoding based on a gap in the eigenvalue spectrum. The MDN propagator was trained over the latent space projection of the MD trajectory at a lag time of $\tau=20$ ns using Adam with early stopping, $C=24$ Gaussian kernels, and two hidden layers of 100 neurons with ReLU activations. The cWGAN decoder comprised a generator and critic with three hidden layers of 200 neurons with Swish activations and a $d=50$-dimensional noise vector. The training loss stabilized after 52 epochs. The cWGAN is trained to generate the Trp-cage Cα backbone by roto-translationally aligning MD training configurations to a reference structure. Training of the full LSS pipeline required ~1 GPU-h on a NVIDIA GeForce GTX 1080 Ti GPU core.

The trained LSS was used to produce $100 \times 208$ µs synthetic trajectories each requiring ~5 s on a single NVIDIA GeForce GTX 1080 Ti GPU core. The LSS trajectories comprise the same total number of frames as the 208 µs all-atom trajectory but contain ~1070 folding/unfolding transitions compared to just 12 in the training data and were generated at six orders of magnitude lower cost. This observation illuminates the crux of the value of the approach: the LSS learns a kinetic model over limited MD training data and is then used to generate vastly longer novel all-atom trajectories that enable the observation of states and events that are only sparsely sampled in the training data. We now validate the thermodynamic, structural, and kinetic predictions of the LSS.

**Thermodynamics.** The free energy profiles projected into the slowest latent space coordinate $F(\psi) = -k_B T \ln(q(\psi))$ show excellent correspondence between the MD and LSS (Fig. 3). The free energy of the folded ($\psi \approx 0$) and unfolded ($\psi \approx 0.9$) basins and transition barriers are in quantitative agreement with a root mean squared error between the aligned profiles of 0.91 $k_B T$. The LSS profiles contain 10-fold lower statistical uncertainties than the MD over the same number of frames due to the 100-fold longer LSS data set enabled by their exceedingly low computational cost.

**Structures.** The MD and LSS molecular structures within the folded basin ($\psi \approx 0$) and metastable transition state ($\psi \approx 0.45$) possess a relative $C_{\alpha}$-RMSD of 0.29 nm and 0.37 nm, respectively. Relative to the Trp-cage native state (PDB ID: 2JOF), the MD and LSS folded configurations possess a $C_{\alpha}$-RMSD of 0.20 nm and 0.28 nm, respectively. The mean and standard deviation of the time-averaged radii of gyration ($R_g$) for the MD ($0.87 \pm 0.16$) nm and LSS ($0.87 \pm 0.13$) nm trajectories are indistinguishable with standard errors computed by five-fold block averaging. These results demonstrate that the LSS molecular structures are in excellent accord with MD.
Kinetics. We compare the MD and LSS kinetics through the autocorrelation times corresponding to the relaxation time scales associated with the \( m=3 \) leading kinetic processes. All three time scales are in excellent agreement and again the LSS uncertainties are approximately 10-fold lower than the MD (Table 1).

| Timescale | MD (μs)       | LSS (μs)  |
|----------|---------------|-----------|
| \( t_1 \) | 3.00 ± 0.61   | 2.89 ± 0.12 |
| \( t_2 \) | 0.54 ± 0.37   | 0.43 ± 0.04  |
| \( t_3 \) | 0.45 ± 0.12   | 0.42 ± 0.01  |

We then employ time-lagged independent component analysis (TICA) to determine whether the LSS trajectory possesses the same slow (linear) subspace as the MD. We featureize the trajectories with pairwise \( \alpha \) distances and perform TICA at a lag time of \( \tau = 20 \) ns. Projection of the free energy surfaces into the leading three MD TICA coordinates show that the leading kinetic variance in the MD data is quite accurately reproduced by the LSS (Fig. 4). The only substantive disagreement is absence in the LSS projection of a small high-free energy metastable state at (TIC1 \( \approx 0 \), TIC3 \( \approx -2.5 \)) corresponding to configurations with Pro18 dihedral angles \( \psi \approx (-50)^\circ \). These configurations are only transiently occupied due to rare Pro18 dihedral flips that occur only twice during the 208 μs MD trajectory and are not contained in the \( m=3 \)-dimensional latent space.

As with all data-driven approaches, the primary deficiency of the LSS approach is that the resulting kinetic models are not necessarily transferable to other conditions or systems and are subject to systematic errors due to approximations in the molecular force fields and incomplete sampling of the relevant configurational space in the training data. The latter issue means that although the generated LSS trajectories are – similar to MSM-based and equation-free approaches – largely interpolative. The stochastic nature of the MDN propagator and generative nature of the cWGAN generator means that we may anticipate local extrapolations beyond the exact training configurations. There is no expectation, however, that the trained model will discover new metastable states or kinetic transitions, and certainly not do so with the correct thermodynamic weights or dynamical time scales.

The present work has demonstrated LSS in a data-rich training regime where the MD training data comprehensively samples configurational space. The next step is to establish an adaptive sampling paradigm – similar to that in MSM construction and some enhanced sampling techniques – to enable its application in a data-poor regime. The adaptive sampling approach interrogates the kinetic model to identify under-sampled states and transitions that contribute most to uncertainties in the model predictions (i.e., “known unknowns”) and initializes new MD simulations to collect additional training data in these regions. This interleaving of MD training data collection and model retraining can dramatically reduce the required quantity of training data. Moreover, new simulations initialized in under-sampled regions may also occasionally be expected to transition into new configurational states not present in the initial training data (i.e., “un-
known unknowns". Iterating this process until convergence can expand the range of the trained kinetic model to encompass the relevant configurational space and minimize the cost of training data collection.

Finally, we also envisage applications of the LSS approach to other fields of dynamical modeling where stiff or multi-scale systems of ordinary or partial differential equations, or the presence of activated processes or rare events introduces a separation of time scales between the integration time step and events of interest. For example, there may be profitable adaptations of the approach for dynamical modeling in such fields as cosmology, ecology, immunology, epidemiology, and climatology.

Conflicts of interest

A.L.F. is a consultant of Evozyne and a co-author of US Provisional Patents 62/853,919 and 62/900,420 and International Patent Application PCT/US2020/035206.

Acknowledgements

This work was supported by MICCoM (Midwest Center for Computational Materials), as part of the Computational Materials Science Program funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division and the National Science Foundation under Grant No. CHE-1841805. H.S. acknowledges support from the National Science Foundation Molecular Software Sciences Institute (MoSSSI) Software Fellows program (Grant No. ACI-1547580). We are grateful to D.E. Shaw Research for sharing the Trp-cage simulation trajectories.

Notes and references

1. D. Frenkel and B. Smit, Understanding Molecular Simulation: From algorithms to applications, Academic Press, San Diego, 2002.
2. J. E. Stone, D. J. Hardy, I. S. Ufimtsev and K. Schulten, Journal of Molecular Graphics and Modelling, 2010, 29, 116–125.
3. D. E. Shaw, J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Isrovich, J. S. Kuskin, R. H. Larson, T. Layman, L. S. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y. H. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpuzza, U. Ben Sacher, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang and C. Young, SC ’14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2014, pp. 41–53.
4. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, Journal of Computational Chemistry, 2005, 26, 1781–1802.
5. E. Chow, C. A. Rendleman, K. J. Bowers, R. O. Dror, D. H. Hughes, J. Gullingsrud, F. D. Sacerdoti and D. E. Shaw, Desmond performance on a cluster of multicore processors (DESRES/TR–2008-01), DE Shaw Research technical report, 2008.
6. J. Glaser, T. D. Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millan, D. C. Morse and S. C. Glotzer, Computer Physics Communications, 2015, 192, 97–107.
7. S. Plimpton, Fast parallel algorithms for short-range molecular dynamics (SAND-91-1144), Sandia National Laboratory technical report, 1993.
8. F. F. Abraham, R. Walkup, H. Gao, M. Duchaineau, T. D. De La Rubia and M. Seager, Proceedings of the National Academy of Sciences of the United States of America, 2002, 99, 5793–5797.
9. F. F. Abraham, R. Walkup, H. Gao, M. Duchaineau, T. D. De La Rubia and M. Seager, Proceedings of the National Academy of Sciences of the United States of America, 2002, 99, 5777–5782.
10. N. Tchapte, S. Seckler, M. Heinen, J. Vrabec, F. Gratl, M. Horsch, M. Bernreuther, C. W. Glass, C. Niethammer, N. Hammer, B. Krischok, M. Resch, D. Kranzlmüller, H. Hasse, H.-J. Bungartz and F. Neumann, The International Journal of High Performance Computing Applications, 2019, 33, 838–854.
11. R. Elber, The Journal of Chemical Physics, 2016, 144, 060901.
12. G. M. Torrie and J. P. Valleau, Journal of Computational Physics, 1977, 23, 187–199.
13. I. R. Torrie and J. P. Valleau, Journal of Chemical Physics, 1967, 47, 4766–4772.
14. C. Abrams and G. Bussi, Entropy, 2014, 16, 163–199.
15. Y. Miao and J. A. McCammon, Molecular Simulation, 2016, 42, 1046–1055.
16. H. Sidky, W. Chen and A. L. Ferguson, Molecular Physics, 2020, 118, 1–21.
17. J. D. Chodera, W. C. Swope, F. Noé, J.-H. Prinz, M. R. Shirts and V. S. Pande, The Journal of Chemical Physics, 2011, 134, 06B612.
18. L. Donati and B. G. Keller, The Journal of Chemical Physics, 2018, 149, 072335.
19. F. Noé, arXiv preprint arXiv:1812.07669, 2018, https://arxiv.org/abs/1812.07669v1.
20. A. Fernández, Annalen der Physik, 2020, 532, 1–5.
21. B. E. Husic and V. S. Pande, Journal of the American Chemical Society, 2018, 140, 2386–2396.
22. V. S. Pande, K. Beauchamp and G. R. Bowman, Methods, 2010, 52, 99–105.
23. J. H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schtte and F. Noé, Journal of Chemical Physics, 2011, 134, 174105.
24. G. R. Bowman, V. S. Pande and F. Noé, An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation, Springer Science & Business Media, 2013, vol. 797.
25. H. Sidky, W. Chen and A. L. Ferguson, Journal of Physical Chemistry B, 2019, 123, 7999–8009.
26. C. Wehmeyer, M. K. Scherer, T. Hempel, B. E. Husic, S. Olsson and F. Noé, Living Journal of Computational Molecular Science, 2019, 1, 1–12.
Engineering, 2018, 20, 26–38.