Coarse-graining auto-encoders for molecular dynamics

Wujie Wang and Rafael Gómez-Bombarelli

Molecular dynamics simulations provide theoretical insight into the microscopic behavior of condensed-phase materials and, as a predictive tool, enable computational design of new compounds. However, because of the large spatial and temporal scales of thermodynamic and kinetic phenomena in materials, atomistic simulations are often computationally infeasible. Coarse-graining methods allow larger systems to be simulated by reducing their dimensionality, propagating longer timesteps, and averaging out fast motions. Coarse-graining involves two coupled learning problems: defining the mapping from an all-atom representation to a reduced representation, and parameterizing a Hamiltonian over coarse-grained coordinates.

We propose a generative modeling framework based on variational auto-encoders to unify the tasks of learning discrete coarse-grained variables, decoding back to atomistic detail, and parameterizing coarse-grained force fields. The framework is tested on a number of model systems including single molecules and bulk-phase periodic simulations.

INTRODUCTION

Coarse-grained (CG) molecular modeling has been used extensively to simulate complex molecular processes with lower computational cost than all-atom simulations. By compressing the full atomistic model into a reduced number of pseudoatoms, CG methods focus on slow collective atomic motions while averaging out fast local motions. Current approaches generally focus on parameterizing coarse-grained potentials from atomistic simulations (bottom-up) or experimental statistics (top-down).

The use of structure-based coarse-grained strategies has enabled important theoretical insights into polymer dynamics and lipid membranes at length scales that are otherwise inaccessible. Beyond efforts to parameterize CG potentials given a pre-defined all-atom to CG mapping, the selection of an appropriate map plays an important role in recovering consistent CG dynamics, structural correlation, and thermodynamics. A poor choice can lead to information loss in the description of slow collective interactions that are important for glass formation and transport. Systematic approaches to creating low-resolution protein models based on essential dynamics have been proposed, but a systematic bottom-up approach is missing for organic molecules of various sizes, resolutions, and functionalities. In general, the criteria for selecting CG mappings are usually based on a priori considerations and chemical intuition. Moreover, although there have been efforts in developing back-mapping algorithms, the statistical connections are missing to reversibly bridge resolutions across scales. We aim to address such multi-scale gaps in molecular dynamics using machine learning.

Recently, machine learning tools have facilitated the development of CG force fields and graph-based CG representations. Here we propose to use machine learning to optimize CG representations and deep neural networks to fit coarse-grained potentials from atomistic simulations. One of the central themes in learning theory is finding optimal hidden representations that capture complex statistical distributions to the highest possible fidelity using the fewest variables. We propose that finding coarse-grained variables can be formulated as a problem of learning latent variables of atomistic distributions. Recent work in unsupervised learning has shown great potential in uncovering the hidden structure of complex data. As a powerful unsupervised learning technique, variational auto-encoders (VAEs) compress data through an information bottleneck that continuously maps an otherwise complex data set into a low-dimensional space and can probabilistically infer the real data distribution via a generating process. VAEs have been applied successfully to a variety of tasks, from image de-noising to learning compressed representations for text, arbitrary grammars, and molecular structures. Recent studies have used VAE-like structures to learn collective molecular motions by reconstructing time-lagged configurations and Markov state models. For the examples mentioned, compression to a continuous latent space is usually parameterized using neural networks. However, coarse-grained coordinates are latent variables in 3D space, and need specially designed computational parameterization to maintain the Hamiltonian structure for discrete particle dynamics.

Motivated by statistical learning theory and advances in discrete optimization, we propose an auto-encoder-based generative modeling framework that (1) learns discrete coarse-grained variables in 3D space and decodes back to atomistic detail via geometric back-mapping; (2) uses a reconstruction loss to help capture salient collective features from all-atom data; (3) regularizes the coarse-grained space with a semi-supervised mean instantaneous force minimization to obtain a smooth coarse-grained free-energy landscape; and (4) variationally finds the highly complex coarse-grained potential that matches the instantaneous mean force acting on the all-atom training data.

RESULTS

Figure 1 shows the general schematics of the proposed framework, which is based on learning a discrete latent encoding by assigning atoms to coarse-grained particles. In Fig. 1b, we illustrate the computational graph of Gumbel-softmax reparameterization, which continuously relaxes categorical distributions for learning discrete variables. We first apply the coarse-grained auto-encoders to trajectories of individual gas-phase molecules. By variationally optimizing encoder and decoder networks to
minimize the reconstruction loss as in Eq. (1), the auto-encoder picks up salient coarse-grained variables that minimize the fluctuation of encoded atomistic motions conditioned on a linear back-mapping function. We adopt an instantaneous-force regularizer (described in the Methods section), to minimize the force fluctuations of the encoded space. This facilitates the learning of a coarse-grained mapping that corresponds to a smoother coarse-grained free-energy landscape. For the unsupervised learning task, we minimize the following loss function.

$$L_{ae} = \frac{1}{N} \mathbb{E}_{x \sim p(x)} [(D(E(x)) - \mu)^2 + \beta F_{inst}(E(x))^2]$$  \(\text{Eq. (1)}\)

The first term on the right-hand side of Eq. (1) represents the atom-wise reconstruction loss and the second term represents the average instantaneous mean force regularization. The relative weight $\beta$ is a hyperparameter describing the relative importance of the force regularization term. The force regularization loss is discussed in the Methods section and training details in the Supplementary Information.

We show the unsupervised auto-encoding process for gas-phase ortho-terphenyl (OTP) and aniline (C$_6$H$_7$N) in Fig. 2. The results show that the optimized reconstruction loss decreases with increasing coarse-grained resolution and that a small number of coarse-grained atoms have the potential to capture the overall collective motions of the underlying atomistic process. The reconstruction loss represents the information loss of coarse-grained particles to represent collective atomistic motions conditioned on a deterministic back-mapping. In the case of OTP, an intuitive 3-bead mapping is learned that partitions each of the phenyl rings. However, such an encoding loses the configuration information describing the relative rotation of the two side rings, resulting in decoded structures that yield higher error. When the number of coarse-grained degrees of freedom increases to 4, the additional beads are able to encode more configurational information than three-bead models and therefore can decode back into atomistic coordinates with high accuracy. We further apply the auto-encoding framework to a small peptide molecule to examine, as a function of CG resolution, the capacity of the coarse-grained representation to capture the critical collective variables of the underlying atomistic states. Although it is not able to recover the arrangement of hydrogen atoms (Fig. 3), the coarse-grained latent variables of 8 CG atoms can faithfully recover heavy atom positions and represent different collective states in the Ramachandran map as the coarse-grained resolution is increased (Fig. 3).

The regularization term (second term in Eq. (1)) addresses the instantaneous mean forces that arise from transforming the all-atom forces. Inspired by gradient domain regularization in deep learning\cite{41–43} and the role of fluctuations in the generalized Langevin framework,\cite{44} we minimize the average instantaneous force as a regularization term to facilitate the learning of a smooth coarse-grained free-energy surface and to average out fast dynamics. The factor $\beta$ is a hyperparameter that controls the interplay between reconstruction loss and force regularization and is typically set to the highest value for which the CG encoding still uses all allotted dimensions.

In Figs 4, 5, 6, and 7, we demonstrate the applicability of the proposed framework to bulk simulations of liquids for small- (C$_2$H$_6$, C$_3$H$_8$) and long-chain (C$_{24}$H$_{50}$) alkanes. Coarse-grained resolutions of 2 and 3 are used for ethane and propane, respectively, while two coarse-grained resolutions of 8 and 12 are used for the C$_{24}$H$_{50}$ alkane melt. We first train an auto-encoder to obtain the latent coarse-grained variables for ethane, propane, and C$_{24}$H$_{50}$, and subsequently train a neural network-based coarse-grained force field with additional excluded volume interactions using force matching to minimize Eq. (15) (in the case of C$_{24}$H$_{50}$, only the backbone carbon atoms are represented). Coarse-grained simulations are then carried out at the same
density and temperature as the atomistic simulation. We include the training details and model hyperparameters in the Supplementary Information. Coarse-grained forces are evaluated using PyTorch and an MD integrator based on ASE (Atomistic Simulations Environment).

By minimizing the instantaneous force-matching loss term according to Eq. (15) in the Methods section, the neural network shows sufficient flexibility to reproduce a reasonably accurate structural correlation function. In the case of C_{24}H_{50} (Figs 6 and 7), the neural network captures the bimodal bond length distribution for the coarse-grained C_{24}H_{50} chains and reproduces the end-to-end distance distribution and mapped monomer pair distribution function accurately. The mean squared displacement plots for all systems demonstrate faster dynamics than the atomistic ground truth due to loss of atomistic friction in the coarse-grained space. For C_{24}H_{50}, we also investigate the decoded C-C structural correlations shown in Fig. 8. The inter-chain structural correlation shows good agreement with the underlying atomistic ground truth, while the C-C bond distances are predicted to be shorter because the coarse-grained super-atoms can only infer average
DISCUSSION

Within the current framework, there are several possibilities for future research directions regarding both the supervised and unsupervised parts.

Here, we have presented a choice of deterministic encoder and decoder. However, such a deterministic CG mapping results, by construction, in an irreversible loss of information. This is reflected in the reconstruction of average all-atom structures instead of the reference instantaneous configurations. To infer the underlying atomistic distributions, past methods have used random structure generation followed by equilibration.\textsuperscript{14-17} By combining this with predictive inference for atomistic back-mapping,\textsuperscript{18} a probabilistic auto-encoder can learn a reconstruction probability distribution that reflects the thermodynamics of the degrees of freedom that were averaged out by the coarse-graining. Using this framework as a bridge between different scales of simulation, generative models can help build better hierarchical understanding of multi-scale simulations.

carbon poses based on the the deterministic inference framework using a linear back-mapping. The prospect of stochastic decoding functions to capture statistical up-scaling is discussed below.

Fig. 4 Comparison between atomistic and CG simulation statistics for liquid propane with a CG resolution of 2 per molecule. a–c are the pair correlation functions and bond length distributions of CG trajectories and mapped atomistic trajectories. d shows the comparison between Mean Squared Displacement for CG and mapped atomistic coordinates, and indicates that CG shows faster dynamics than the atomistic trajectory. e shows the learning of a discrete CG mapping during training of the auto-encoder. The rectangular matrix is a colored representation of matrix $E_{ij}$, the colors showing relative values of the matrix elements.

Fig. 5 Comparison between atomistic and CG simulation statistics for liquid ethane using a CG resolution of 2 per molecule. a–c are the pair correlation functions and bond length distributions of CG trajectories and mapped atomistic trajectories. d shows the comparison between Mean Squared Displacement for CG and mapped atomistic coordinates, and indicates that CG shows faster dynamics than the atomistic trajectory. e shows the learning of a discrete CG mapping during training of the auto-encoder. The rectangular matrix is a colored representation of matrix $E_{ij}$, the colors showing relative values of the matrix elements.
Furthermore, neural network potentials provide a powerful fitting framework to capture many-body correlations. The choice of force-matching approach does not guarantee the recovery of individual pair correlation functions derived from full atomistic trajectories because the cross-correlations among coarse-grained degrees of freedom are not explicitly incorporated. More advanced fitting methods can be incorporated in the current neural network framework to address the learning of structural cross-correlation, including iterative force matching and relative entropy method.

Methods based on force-matching, like other bottom-up approaches such as relative entropy method, attempt to reproduce structural correlation functions at one point in the thermodynamic space. As such, they are not guaranteed to capture non-equilibrium transport properties and are not necessarily transferable among different thermodynamic conditions. The data-driven approach we propose enables learning over different thermodynamic conditions. In addition, this framework opens new routes to understanding how the coarse-grained representation influences transport properties by
training on time-series data. A related example in the literature is to use a time-lagged auto-encoder\textsuperscript{37} to learn a latent representation that best captures molecular kinetics.

In summary, we propose to treat the coarse-grained coordinates as latent variables which can be sampled with coarse-grained molecular dynamics. By regularizing the latent space with force regularization, we train the encoding mapping, a deterministic molecular dynamics. By regularizing the latent space with force

ELBO by propagating gradients through the probability distributions imposes a statistical structure over the latent variables. Minimizing the ELBO provides a parameterizable way of inferring complicated distributions of molecular dynamics.

Similar to variational auto-encoders with constraint on the latent space, a coarse-grained latent space should preserve the structure of the molecular mechanics phase space. Noid et al.\textsuperscript{3} have studied the general requirements for a physically rigorous encoding function. In order to address those requirements, the auto-encoder is trained to optimize the reconstruction of atomistic configurations by propagating them through a low-dimensional bottleneck in Cartesian coordinates. Unlike most instances of VAEs, the dimensions of the CG latent space have physical meaning. Since the CG space needs to represent the system in position and momentum space, latent dimensions need to correspond to real-space Cartesian coordinates and maintain the essential structural information of molecules.

We make our encoding function a linear projection in Cartesian space $E(x) : \mathbb{R}^{3n} \rightarrow \mathbb{R}^{3N}$ where $n$ is the number of atoms and $N$ is the desired number of coarse-grained particles.

Let $x$ be the atomistic coordinates and $z$ be the coarse-grained coordinates. The encoding function should satisfy the following requirements:\textsuperscript{104}

1. $z_k = E(x) = \sum_{i=1}^{n} E_i x_{ik} \in \mathbb{R}^3, i = 1 \ldots N, j = 1 \ldots n$,
2. $\sum_{i} E_i = 1$ and $E_i \geq 0$,
3. Each atom contributes to at most one coarse-grained variable $z$ where $E_i$ defines the assignment matrix to coarse-grained variables, $j$ is the atomic index, $i$ is the coarse-grained atom index, and $k$ represents the Cartesian coordinate index. Requirement (2) defines the coarse-grained variables to be a weighted geometric average of the Cartesian coordinates of the contributing atoms. In order to maintain consistency in momentum space after the coarse-grained mapping, the coarse-grained masses are

METHODS

Here we introduce the auto-encoding framework from the generative modeling point of view. The essential idea is to treat coarse-grained coordinates as a set of latent variables that are the most predictive of the atomistic distribution while having a smooth underlying free-energy landscape. We show that this is achieved by minimizing the reconstruction loss and the instantaneous force regularization term. Moreover, under the variational auto-encoding framework, we can understand the force matching as the minimization of the relative entropy between coarse-grained and atomistic distributions in the gradient domain.

Coarse-graining auto-encoding

The essential idea in generative modeling is to maximize the likelihood of the data under the generative process:

$$P(x) = \int P(x|z)P(z)dx$$

(2)

where $z$ are the latent variables that carry the essential information of the distributions and $x$ represents the samples observed in the data. Variational auto-encoders maximize the likelihood of the observed samples by

Fig. 8 Comparison of the decoded backbone carbon atoms with the atomistic ground truth. a the inter-chain C–C radial distributions shows reasonable agreement between the decoded and original distributions. b The decoded carbon backbones show shorter predicted bond length because the decoded structures represent the mean reconstruction of an ensemble of carbon chain poses. c demonstrates the auto-encoding of carbon backbones for C$_{24}$H$_{50}$ molecules. As a result of loss of mapping entropy, the decoded structures show a straight backbone compared to the atomistic ground truth.
Hence, we present an analogous interpretation of the reconstruction loss $g_{ij}$ where entropy (scaled by the variance) de- error and is understood as a Gaussian approximation to the mapping Gaussian, the reconstruction loss yields the term-by-term mean-squared

\begin{equation}
E_{ij} = \frac{C_{ij}}{\sum_{j} C_{ij}}
\end{equation}

(4)

\begin{equation}
\bar{C}_{j} = \text{one hot}(\arg \max \log \phi_{ij})
\end{equation}

(5)

$\bar{C}_{j}$ is the coarse-graining one-hot assignment vector for atom $j$ using Gumbel-softmax reparameterization with each vector element $\bar{C}_{ij}$ representing the assignment of atom $j$ to coarse-grained atom $i$, so that requirement (3) is automatically satisfied. Gumbel-softmax reparameterization is a continuous relaxation of Gumbel-max reparameterization for differentiable approximation using the softmax function. Similar parameterization techniques include concrete distributions, REBAR and RELAX. The Gumbel-softmax reparameterization has been applied in various machine learning scenarios involving learning discrete structures, propagating discrete policy gradient in reinforcement learning and generating context-free grammar. The continuously relaxed version of Eq. (5) is:

\begin{equation}
\bar{C}_{j} = \frac{e^{\log \phi_{ij} + u_{ij}/\tau}}{\sum_{i} e^{\log \phi_{ij} + u_{ij}/\tau}}
\end{equation}

(6)

where $u_{ij}$ is sampled from the Gumbel distribution via the inverse transformation $g_{ij} = -\log(-\log(u_{ij}))$, where $u_{ij}$ is sampled from a uniform distribution from 0 to 1. During training, $\tau$ gradually decreases with the training epoch and the one-hot categorical encoding is achieved in the limit of small $\tau$. Therefore, the encoding distribution $Q(\mathbf{z}|x)$ is a linear projection operator parameterized by discrete atom-wise categorical variables.

For the generation of atomistic coordinates conditioned on coarse-grained coordinates, we opt for a simple decoding approach via geometrical projection using a matrix $D$ of dimension $n$ by $N$ that maps coarse-grained variables back to the original space so that $\hat{x} = D(\mathbf{z}) = \sum_{i} D_{i} \bar{z}_{i}$ where $\hat{x}$ are the reconstructed atomistic coordinates. Hence, both the encoding and decoding mappings are deterministic. However, deterministic reconstruction via a low-dimensional space leads to irreversible information loss that is analogous to the mapping entropy introduced in Shell et al. In our experiments, by assuming $P_{0}(x) = \text{Gaussian}$, the reconstruction loss yields the term-by-term mean-squared error and is understood as a Gaussian approximation to the mapping entropy (scaled by the variance) defined by Shell et al.

\begin{equation}
S_{\text{map}} = \mathbb{E}_{x, \mathcal{P}(x)} \log \left( \frac{P(x) \Omega(E(x))}{P(z)} \right)
\end{equation}

(7)

where $\Omega(E(x))$ is the configuration space volume that is mapped to the atomistic coordinates. The latent variable framework provides a clear parameterizable objective whose optimization minimizes the information loss due to coarse-graining by using the following objective as reconstruction loss.

\begin{equation}
\min_{\phi, D_{ij}} L_{AE} = \min_{\phi, D_{ij}} \mathbb{E}_{x} \mathbb{E}_{P(x)} \mathbb{E}_{y} \text{Gumbel}(\beta, 1) \left[ D(E(x, g, \tau)) - x \right]^{2}
\end{equation}

(8)

Hence, we present an analogous interpretation of the reconstruction loss in Eq. (3) but in the Cartesian space of coarse-grained pseudo-atoms in molecular dynamics. This loss can be optimized by algorithm 1. A regularized version is introduced in section C.
where \( F(z) \) is the mean force and \( b \) represents a family of possible vectors such that \( \mathbb{V} E(x) \neq 0 \). We further define \( F_{\text{inst}}(z) = -\mathbb{V} V(x) \) to be the instantaneous force and its conditional expectation is equal to the mean force \( F(z) \). It is important to note that \( F_{\text{inst}}(z) \) is not unique and depends on the specific choice of \( \mathbb{w} \)\(^{45-46} \) but their conditional averages return the same mean force. For possible \( b \), we further choose \( \mathbb{w} = \mathbb{V} E(x) \) which is a well-studied choice\(^ {45} \), so that:

\[
\mathbb{b} = \mathbb{V} E(x) / \mathbb{V} E(x) = \mathbb{C}
\]

where \( \mathbb{b} \) is a function of \( \mathbb{V} E(x) \). In the case of coarse-graining encodings, \( \mathbb{b} = \mathbb{C} \) where \( \mathbb{C} \) is the encoding matrix formed by concatenating atomwise one-hot vectors as defined in Eq. (6). We adopt the force-matching scheme introduced by Izvekov et al.\(^ {45,46} \), in which the mean-squared error is used to match the mean force and the “coarse-grained force” is the negative gradient of the coarse-grained potential. The optimizing functional, developed based on Izvekov et al., is:

\[
\min_{\mathbb{b}} E_{m} = \min_{\mathbb{b}} \mathbb{E}_{m}[(F(z) + \mathbb{V} V_{CG}(z))^{2}]
\]

where \( \theta \) are the parameters in \( V_{CG} \) and \( V_{CG} \) represents the “coarse-grained forces” which can be obtained from automatic differentiation as implemented in open-source packages like PyTorch.\(^ {45} \) However, to compute the mean force \( F \) would require constrained dynamics\(^ {45} \) to obtain the average of the fluctuating microscopic forces. According to Zhang et al.\(^ {9} \), the force-matching functional can be alternatively formulated by treating the instantaneous mean force as an instantaneous observable with a well-defined average being the mean force \( F(z) \):

\[
F_{\text{inst}}(z) = F(z) + \epsilon(z)
\]

on the condition that \( \mathbb{E}[F_{\text{inst}}] = F(z) \). The original variational functional becomes instantaneous in nature and can be reformulated as the following minimization target:

\[
\min_{\mathbb{b}} E_{\text{inst}} = \min_{\mathbb{b}} \mathbb{E}[F_{\text{inst}}(z) + \mathbb{V} V_{CG}(z)]^{2}
\]

Instead of matching mean forces that need to be obtained from constrained dynamics, our model minimizes \( L_{\text{inst}} \) with respect to \( V_{CG}(z) \) and \( E(z) \). \( L_{\text{inst}} \) can be shown to be related to \( L \) with some algebra: \( L_{\text{inst}} = L + \mathbb{E}[(E(z))]^{2} \).\(^ {18} \) This functional provides a variational way to find a CG mapping and its associated force field functions.

Instantaneous mean force regularization

Here we introduce the gradient regularization term that is designed to minimize the fluctuation in the mean forces. Similar methods involving gradient regularization have been applied in supervised learning computer vision tasks to smooth the loss landscape for improved model generalization.\(^ {11-14} \) In coarse-grained modeling, minimizing the forces is important for learning the slow degrees of freedom and a smoother free-energy surface.

Based on the generalized Langevin equation, the difference between the general force and instantaneous mean force \( \epsilon(E(x)) \) can be approximated as\(^ {4,6,5} \):

\[
\epsilon(E(x), t) = \gamma \frac{dE(x)(t)}{dt} - \int_{0}^{t} \beta(t) \frac{dE(x)(t - \tau)}{dt} d\tau + \tilde{\eta}(t) + \sum_{C_i} c_i \eta_i
\]

where \( \gamma \) is the friction coefficient, \( \beta(t) \) is the memory kernel, \( \tilde{\eta}(t) \) is the colored Gaussian noise, and \( \sum_{C_i} c_i \eta_i \) is the mapped atomistic white noise. To avoid the need for special dynamics when running ensemble calculations, it is desirable to minimize the memory and fluctuation term to yield calculations with fewer fluctuation terms. A related example in the work by Guttenberg et al.\(^ {47} \) who compare the memory heuristics among coarse-grained mapping function. The objective we propose can be optimized by gradient descent to continuously explore the coarse-grained mapping space without iterating over the combinatorial spaces. We perform this regularization by minimizing the mean-squared instantaneous forces over mini-batches of atomic trajectories to optimize the CG mappings.

\[
\min_{\mathbb{b}} E_{\text{inst}}(z) = \min_{\mathbb{b}} E_{\text{inst}}(z)^2 + \epsilon^{2}(E(x))
\]

In practice, this regularization loss is combined with \( L_{\text{cg}} \) to obtain a coarse-grained mapping with a certain weight \( \rho \) that is added onto the reconstruction loss. We discuss the practical effect of including the regularization term in the Supplementary Information.

DATA AVAILABILITY

Data for training the model is available upon request. An implementation of the algorithm described in the paper is available at https://github.com/learningmatter/Coarse-Graining-Auto-encoders.

Received: 27 March 2019; Accepted: 19 November 2019; Published online: 18 December 2019

REFERENCES

1. Agostino, M. D., Risselada, H. J., Lürick, A., Ungermann, C. & Mayer, A. A tethering complex drives the terminal stage of SNARE-dependent membrane fusion. Nature 551, 634–638 (2017).
2. Huang, D. M. et al. Coarse-grained computer simulations of polymer / fullerene bulk heterojunctions for organic photovoltaic applications. J. Chem. Theory Comput. 6, 1–11 (2010).
3. Noid, W. G. et al. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. J. Chem. Phys. 128, 243116 (2008).
4. Martínk, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & De Vries, A. H. The MARTINI force field: coarse-grained model for biomolecular simulations. J. Phys. Chem. B 111, 7812–7824 (2007).
5. Periole, X., Cavalli, M., Martínk, S.-J. & Ceruso, M. A. Combining an elastic network coarse-grained model with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. J. Chem. Theory Comput. 5, 2531–2543 (2009).
6. Wijesinghe, S., Perahia, D. & Grest, G. S. Polymer topology effects on dynamics of comb polymer melts. Macromolecules 51, 7621–7628 (2018).
7. Salerno, K. M., Agrawal, A., Peters, B. L., Perahia, D. & Grest, G. S. Dynamics in entangled polyethylene melts. Eur. Phys. J. Spec. Topics 225, 1707–1722 (2016).
8. Salerno, K. M., Agrawal, A., Perahia, D. & Grest, G. S. Resolving dynamic properties of polymers through coarse-grained computational studies. Phys. Rev. Lett. 116, 058302 (2016).
9. Xia, W. et al. Energy renormalization for coarse-graining polymers having different segmental structures. Sci. Adv. 5, eaav6883 (2019).
10. Vogele, M., Köfﬁnger, J. & Hummer, G. Hydrodynamics of diffusion in lipid membrane simulations. Phys. Rev. Lett. 120, 268104 (2018).
11. Rudzinski, J. F. & Noid, W. G. Investigation of coarse-grained mappings via an iterative generalized Yvon-Born-Green method. J. Phys. Chem. B 118, 8295–8312 (2014).
12. Noid, W. G. Perspective: coarse-grained models for biomolecular systems. J. Chem. Theory Comput. 13, 9090 (2017).
13. Zhang, Z. et al. A systematic methodology for deﬁning coarse-grained sites in large biomolecules. Biophys. J. 95, 5073–5083 (2008).
14. Peng, J., Yuan, C., Ma, R. & Zhang, Z. Backmapping from multiresolution coarse-grained models to atomic structures of large biomolecules by restrained molecular dynamics simulations using bayesian inference. J. Chem. Theory Comput. 15, 3344–3353 (2019).
15. Chen, L. J., Qian, H. J., Lu, Z. Y., Li, Z. S. & Sun, C. C. An automatic coarse-graining and fine-graining simulation method: application on polyethylene. J. Phys. Chem. B 110, 24093–24100 (2006).
16. Lombardi, L. E., Marti, M. A. & Capece, L. CG2AA: backmapping protein coarse-grained structures. Bioinformatics 32, 1235–1237 (2016).
17. Machado, M. R. & Pantano, S. SIRAH tools: mapping, backmapping and visualization of coarse-grained models. Bioinformatics 32, 1568–1570 (2016).
18. Schöberl, M., Zabaras, N. & Koutsourelakis, P.-S. Predictive coarse-graining. J. Comput. Phys. 333, 49–77 (2017).
19. Zhang, L., Han, J., Wang, H., Car, R. & W. E. DeePCG: constructing coarse-grained models via deep neural networks. J. Chem. Phys. 149, 034101 (2018).
20. Bevajia, K. K., Singh, S., An, Y. & Deshmukh, S. A. Machine-learned coarse-grained models. J. Phys. Chem. Lett. 9, 4667–4672 (2018).
21. Lemke, T. & Peter, C. Neural network based prediction of conformational free energies - a new route toward coarse-grained simulation models. J. Chem. Theory Comput. 13, 6213–6221 (2017).
Advances in Neural Information Processing Systems. Vol. 2017, p. 2628–2637 (2017).
56. Grathwohl, W., Choi, D., Wu, Y., Roeder, G. & Duvenaud, D. Backpropagation through the Void: optimizing control variates for black-box gradient estimation. In Proc. International Conference on Learning Representations (2017).
57. Van Den Oord, A., Vinyals, O. & Kavukcuoglu, K. Neural discrete representation learning. In Advances in Neural Information Processing Systems. Vol. 2017, p. 6307–6316 (2017).
58. Wu, Y., Wu, Y., Gkioxari, G. & Tian, Y. Building generalizable agents with a realistic and rich 3D environment. https://arxiv.org/abs/1801.02209 (2018).
59. Chmiela, S. et al. Machine learning of accurate energy-conserving molecular force fields. Sci. Adv., 3, e1603015 (2017).
60. Izvekov, S. & Voth, G. A. A multiscale coarse-graining method for biomolecular systems. J. Chem. Phys. 109, 2469–2473 (2005).
61. Ciccotti, G., Kapral, R. & Vanden-Eijnden, E. Blue moon sampling. Vectorial reaction coordinates, and unbiased constrained dynamics. ChemPhysChem 6, 1809–1814 (2005).
62. Kalligiannakis, E., Harmandaris, V., Katsoulakis, M. A. & Plecháč, P. The geometry of generalized force matching and related information metrics in coarse-graining of molecular systems. J. Chem. Phys. 143, 84105 (2015).
63. DenOtter, W. K. Thermodynamic integration of the free energy along a reaction coordinate in Cartesian coordinates. J. Chem. Phys. 112, 7283–7292 (2000).
64. Izvekov, S. & Voth, G. A. Multiscale coarse-graining of mixed phospholipid/cholesterol bilayers. J. Chem. Theory Comput. 2, 637–648 (2006).
65. Lange, O. F. & Grubmüller, H. Collective Langevin dynamics of conformational motions in proteins. J. Chem. Phys. 124, 214903 (2006).

ACKNOWLEDGEMENTS
W.W. thanks Toyota Research Institute for financial support. R.G.B. thanks MIT DMSE and Toyota Faculty Chair for support. W.W. and R.G.B. thank Prof. Adam P. Willard (Massachusetts Institute of Technology) and Prof. Salvador Leon Cabanillas (Universidad Politécnica de Madrid) for helpful discussions. W.W. thanks Mr. William H. Harris for proofreading the manuscript and helpful discussions.

AUTHOR CONTRIBUTIONS
R.G.B. conceived the project, W.W. wrote the computer software and carried out simulations with contributions from R.G.B.; both authors wrote the manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information is available for this paper at https://doi.org/10.1038/s41524-019-0261-5.
Correspondence and requests for materials should be addressed to R.G.B.
Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2019

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences npj Computational Materials (2019) 125