ML-Enhanced Generalized Langevin Equation for Transient Anomalous Diffusion in Polymer Dynamics

Gian-Michele Cherchi *
ICCF
CNRS, Université Clermont-Auvergne
F-63000, Clermont-Ferrand, France

Alain Dequidt †
ICCF
CNRS, Université Clermont-Auvergne
F-63000, Clermont-Ferrand, France

Arnaud Guillin
Laboratoire de Mathématiques Blaise Pascal
CNRS, Université Clermont-Auvergne
F-63000, Clermont-Ferrand, France

Vincent Barra
Laboratory of Informatics, Modeling and Optimization of the Systems
CNRS, Clermont Auvergne INP
F-63000, Clermont-Ferrand, France

Nicolas Martzel
Michelin
63040, Clermont-Ferrand, France

Patrice Hauret
Michelin
63040, Clermont-Ferrand, France

Abstract

In this work, we introduce an ML framework to generate long-term single-polymer dynamics by exploiting short-term trajectories from molecular dynamics (MD) simulations of homopolymer melts. Even with current advances in machine learning for MD, these polymeric materials are difficult to simulate and characterize due to prohibitive computational costs when long timescales are involved. Our method relies on a 3D neural autoregressive (NAR) model for collective variables (CVs), which enhances the Generalized Langevin Equation capabilities in modeling diffusion phenomena. ML-GLE is capable of reproducing long-term single polymer statistical properties, predicting the diffusion coefficient, and resulting in an enormous acceleration in terms of simulation time. Moreover, it is also scalable with system size.

1 Introduction

Accurate mesoscale MD simulations of physical systems require huge computational resources, given the number of degrees of freedom (d.o.f) and the complexity of interactions involved. Data-driven dimensionality reductions, like *coarse-graining* (CG) [1–6], are not sufficient, and accessing long timescales remains expensive. In polymer melts (Fig.1), single polymers undergo *Transient Anomalous Diffusion* (TAD); their mean square displacement (MSD), $\langle X^2(t)\rangle$, is characterized by a long anomalous *subdiffusive* timescale, t^ν with $\nu \in (0,1)$, before reaching a linear diffusive one, where $\langle X^2(t)\rangle = Dt$.

*The work was conducted at **SimatLab**. Url: https://simatlab.com/english.

†Corresponding Author, alain.dequidt@uca.fr

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

Resorting to the description of single polymers reduces the number of d.o.fs, and the governing effective dynamics can be described by the Generalized Langevin Equation (GLE), for a proper selection of CVs [7–9]. However, their dynamics is *non-Markovian* and memory effects are difficult to model and reproduce faithfully. Parametric GLE models of TAD exist, but they fail when no data is available on the transient and diffusive regimes [10, 11].

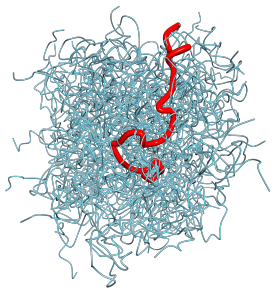


Figure 1: A snapshot of a CG MD simulation of a *Butadiene Rubber* melt. In red, a tagged polymer. ML-GLE exploits short *single-polymer* trajectories to generate long-term stochastic dynamics and predict D.

Several studies have proposed a data-driven parametrization of GLEs [12–20]. Chorin and Lu's seminal work [21] has outlined the connection between nonlinear autoregressive models and the governing equations of the effective stochastic dynamics of a subset of variables of interest. In [22], a GLE description based on a nonlinear autoregressive model for a CV is proposed, with data from a MD simulation of a magnetic system. In contrast, this work proposes, to our knowledge, the first application of data-driven GLEs employing neural networks, to a diffusion problem in soft matter.

Our framework, called ML-GLE, accelerates the polymer Center of Mass (C.o.M) dynamics computation. It connects it to physically motivated faster CVs through a parametrized solution of an *ansatz* GLE, able to model TAD [11]. We independently train a set of NAR generative models, each of them accounting for a different 3D CV. In addition, we improve its learning capabilities by exploiting the symmetries of the physical system. As a result, the 3D CV stochastic dynamics is stable in the long term. Since training and generation are fast and we only use short *downsampled* CG MD trajectories, the computational cost to estimate asymptotic relaxation properties (MSD) reduces dramatically.

2 Background

The Generalized Langevin Equation (GLE) is a stochastic integro-differential equation, accouting for memory effects and being able to reproduce TAD [23]. In absence of external force fields, it reads as follows,

$$m\ddot{\mathbf{X}}(t) = -\int_0^t \Gamma(t-s)\dot{\mathbf{X}}(s)ds + \mathbf{F}(t) \tag{1}$$

where $\mathbf{X}(t)$ is a spatial coordinate (e.g. C.o.M), $\Gamma(t)$ the kernel function and $\mathbf{F}(t)$ a colored noise. The information on the unresolved environmental variables is contained in the kernel function, which is seldom known and usually intractable. However, if the kernel is *integrable*[3], $\lim_{t\to\infty} \mathbb{E}[X^2(t)] \sim t$, and a solution can be thus obtained in the zero-mass limit [11], yielding,

$$\mathbf{X}(t) = \alpha\mathbf{B}(t) + \beta\sum_{j=1}^{N-1}\mathbf{z}_j(t), \tag{2}$$

where $\mathbf{B}(t)$ is a Brownian process modeling C.o.M diffusion, and $\{\mathbf{z}_j(t)\}_{j=1}^{N-1}$ is a set of faster stochastic processes, which are responsible for the anomalous behaviour and implicitly carry information about the environment. We notice there exists a similarity between the latter and single polymer modes of fluctuation [24], called *normal modes*, in the prototypical case of polymer diffusion. We assume therefore that sufficient information on the long-term behaviour is contained in the short-term dynamics of the k slowest modes. The components of the latter are given by the Discrete Cosine Transform (DCT), $\mathbf{z}_{j,t} = \sum_{n=1}^{N}\mathbf{x}_{n,t}\cos\left[\frac{\pi}{N}\left(j+\frac{1}{2}\right)n\right]$, where $\{\mathbf{x}_{n,t}\}_{n=1}^{N}$ stands for monomer coordinates. The result is a *non-Markovian* 3D discrete time stochastic process $\{\mathbf{z}_{j,t}\}_{t>0},\ \forall j \in \{1,...,k\}$.

[3]More details on kernel families and integrability conditions, as well as TAD can be found in [23]

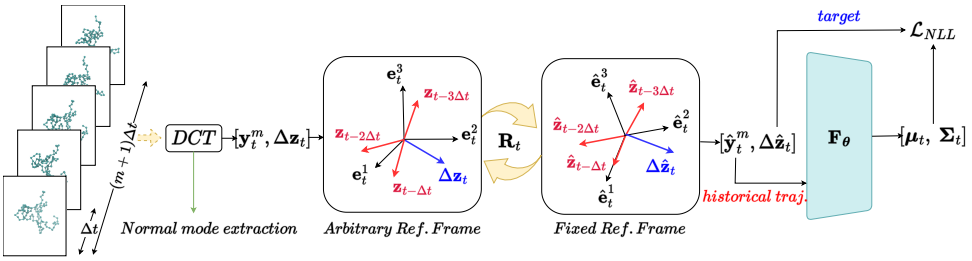


Figure 2: **Training scheme**. Mode j (subscript dropped) is extracted from polymer configurations with a DCT. Subtrajectory of length m and target are rotated to a fixed reference frame, such that $\hat{\mathbf{e}}_t^1$ is always aligned with $\hat{\mathbf{z}}_{t-\Delta t}$, while $\hat{\mathbf{z}}_{t-2\Delta t}$, lies in the plane formed by $\hat{\mathbf{e}}_t^1, \hat{\mathbf{e}}_t^2$.

3 ML-GLE framework

3.1 Neural Autoregressive Generative Model for non-Markovian processes

We denote as $\mathbf{y}_{j,t}^m = \{\mathbf{z}_{j,t-i\Delta t}\}_{i=1}^m \in \mathbb{R}^{3\times m}$ the 3D mode historical trajectory and as $\Delta\mathbf{z}_{j,t} \in \mathbb{R}^{3\times 1}$ its successive time increment vector. Since the whole system is at equilibrium, it is safe to assume that the processes possess a finite size memory of order m, corresponding to a time lag $t_{max} = m\Delta t$, and that $p(\mathbf{z}_{j,t}|\mathbf{y}_{j,t}^m)$ is stationary. We propose therefore to approximate it with a parametrized distribution $p_\theta(\Delta\mathbf{z}_{j,t}|\mathbf{y}_{j,t}^m)$ for *first* differences, followed by an integration step. A 3D Gaussian distribution constitutes a good *ansatz*, since the modes stationary distribution is also Gaussian, hence,

$$p_\theta(\Delta\mathbf{z}_{j,t}|\mathbf{y}_{j,t}^m) \sim \mathcal{N}(\Delta\mathbf{z}_{j,t}|\boldsymbol{\mu}_{j,\theta}(\mathbf{y}_{j,t}^m), \boldsymbol{\Sigma}_{j,\theta}(\mathbf{y}_{j,t}^m)) \tag{3}$$

where $\boldsymbol{\mu}_{j,\theta} : \mathbb{R}^{3\times m} \to \mathbb{R}^3, \boldsymbol{\Sigma}_{j,\theta} : \mathbb{R}^{3\times m} \to \mathbb{R}^{3\times 3}$ are two neural networks parametrizing the conditional mean and covariance matrix, depending on instances of historical trajectories. Each $\mathbf{y}_{j,t}^m$ feeds a network $\mathbf{F}_{j,\theta}$ (MLP) which encodes input-output correlations and forecasts the two distribution parameters. $\Delta\mathbf{z}_{j,t}$ is used as target of the supervised training scheme, as shown in Fig. 2.

Symmetries. Polymers diffuse isotropically in the absence of external driving perturbations breaking spherical symmetry. Thus, the conditional distribution should be *invariant* under any arbitrary global rotation, i.e. $p(\Delta\mathbf{z}_{j,t}|\mathbf{y}_{j,t}^m) = p(\mathbf{R}\Delta\mathbf{z}_{j,t}|\mathbf{R}\circ\mathbf{y}_{j,t}^m)$, where $\mathbf{R} \in \mathrm{SO}(3)$, and is applied *element-wise* to $\mathbf{y}_{j,t}^m$. This observation is crucial for asymptotic generation stability because training in a fixed reference frame spares the network from learning rotations in $\mathrm{SO}(3)$, and avoids the need for *data augmentation*, or more complicated architectures. $\mathbf{R}_{j,t} = [\hat{\mathbf{e}}_{j,t}^1, \hat{\mathbf{e}}_{j,t}^2, \hat{\mathbf{e}}_{j,t}^3]$ is defined from the last three elements of $\mathbf{y}_{j,t}^m$, following the *Gram-Schmidt* orthonormalization. The result is a fixed orthonormal set spanning $\mathbb{R}^3$. In this way, optimization is performed in the same reference frame because training instances are transformed accordingly, $\hat{\mathbf{y}}_{j,t}^m = \mathbf{R}_{j,t}\circ\mathbf{y}_{j,t}^m$, $\Delta\hat{\mathbf{z}}_{j,t} = \mathbf{R}_{j,t}\Delta\mathbf{z}_{j,t}$ (see Fig. 2).

Loss function. We optimize on the Negative Log-Likelihood of a 3D Gaussian distribution,

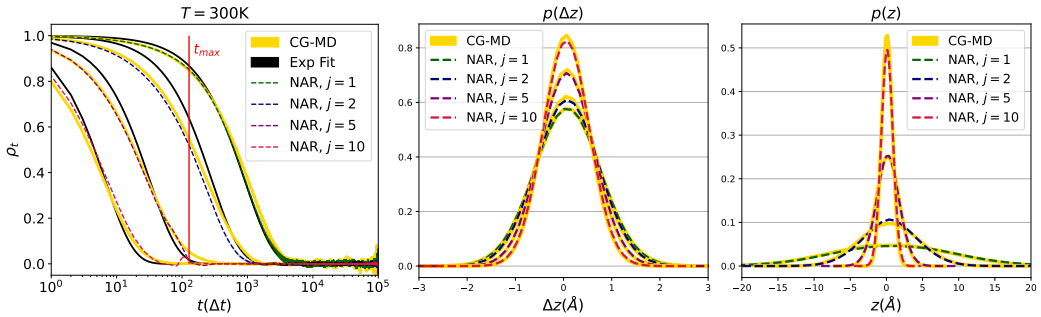
$$\mathcal{L}_{NLL}(\theta) = \mathrm{Tr}(\log\mathbf{D}_{j,t}) + (\Delta\hat{\mathbf{z}}_{j,t} - \boldsymbol{\mu}_{j,t})^T\boldsymbol{\Sigma}_{j,t}^{-1}(\Delta\hat{\mathbf{z}}_{j,t} - \boldsymbol{\mu}_{j,t}) \tag{4}$$

where $\mathbf{D}_{j,t}$ is the diagonal matrix of the LDL^T decomposition of $\boldsymbol{\Sigma}_{j,t}$. This is justified by the fact that $\nabla_\theta\mathcal{L}(\theta) \propto \nabla_\theta \det\boldsymbol{\Sigma}_\theta/\det\boldsymbol{\Sigma}_\theta$, meaning that the *backpropagation* signal would be sensitive to small entries in the covariance matrix, leading to exploding loss values. Hence, $[\boldsymbol{\mu}_{j,t}, \mathbf{D}_{j,t}, \mathbf{L}_{j,t}] = \mathbf{F}_{j,\theta}(\hat{\mathbf{y}}_{j,t}^m)$.

Autoregressive Generation. Once $\hat{\mathbf{y}}_{j,t}^m$ is fed to $\mathbf{F}_\theta$, we use the *reparametrization trick* to sample $\Delta\hat{\mathbf{z}}_{j,t}$, which is then rotated back to the $\mathbf{y}_{j,t}^m$ reference frame. Given $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbb{I}_3)$ one has,

$$\Delta\hat{\mathbf{z}}_{j,t} = \boldsymbol{\mu}_{j,t} + \mathbf{L}_{j,t}\mathbf{D}_{j,t}^{1/2}\boldsymbol{\epsilon}, \qquad \Delta\mathbf{z}_{j,t} = \mathbf{R}_{j,t}^{-1}\Delta\hat{\mathbf{z}}_{j,t}$$

Following an implicit *Euler* integration scheme $\mathbf{z}_{j,t} = \mathbf{z}_{j,t-\Delta t} + \frac{\sigma_{\Delta\mathbf{z}_j}}{\sigma_{\mathbf{z}_j}}\Delta\mathbf{z}_{j,t}\Delta t$, a new value of $\mathbf{z}_{j,t}$ is therefore obtained.
$\sigma_{\Delta\mathbf{z}_j}/\sigma_{\mathbf{z}_j}$ are the empirical standard deviations and are needed for consistency, since $\Delta\mathbf{z}_j, \mathbf{z}_j$ are both standardized. The new input trajectory is obtained with a sliding window on the newly sampled value.



(a) NACF with 99% confidence intervals

(b) Δz_j and z_j empirical distributions

Figure 3: NAR generation results vs. MD for the $1st$, $2nd$, $5th$ and $10th$ mode, at $T = 300\,\mathrm{K}$. Figure **a)** demonstrates the important deviations of MD NACFs from the corresponding exponential fit (in *black*), characteristic of classical *Markovian* OU processes.

3.2 Learning the GLE parameters

Rewriting Eq. (2) as a *lag-t* finite difference equation, squaring and taking expectations, yields $\mathbb{E}[\Delta X_t^2] = \alpha^2 t + \beta^2 \sum_{j=1}^{k} \mathbb{E}[\Delta z_{j,t}^2]$. In the latter, we imposed modes statistical independence, $\mathbb{E}[\Delta z_{i,t} \Delta z_{j,t}] = \sigma_{ij,t}^2 \delta_{ij}$, which allows training k separate *neural networks*. We fit the GLE parameters on short-time ensemble averages datapoints, hence,

$$\alpha_k^*, \beta_k^* = \arg\min_{\alpha,\beta} \left\{ \sum_{t \in \mathcal{P}} \left(\langle \Delta X_t^2 \rangle - \alpha^2 t - \beta^2 \sum_{j=1}^{k} \langle \Delta z_{j,t}^2 \rangle \right)^2 \right\}, \; s.t. \; \alpha, \beta \geq 0 \tag{5}$$

where $\mathcal{P} = \{1, t_{max}/2, t_{max}\}$ and $t_{max} = m\Delta t$.

4 Results

We have performed 11 MD simulation of 100 *cis-Polybutadyene* polymer chains with $N = 100$ monomers, from 300 K to 400 K, every 10 K, in NVT conditions, with dissipative particle dynamics in LAMMPS[4]. CG potentials are obtained from an *All-Atom* MD simulation of the same system (see [25]). We used a small $\delta t = 50\,\mathrm{fs}$ and snapshots were saved each $\Delta t = 2000\delta t$. Simulations run for $T_{sim} = 10^5 \Delta t$ but, for training, $T_{train} = 0.01 T_{sim}$. For each polymer, only the slowest $k = 12$ modes are computed from monomer coordinates. Subtrajectory training instances are finally obtained with a lag-1 sliding window of length $m + 1$ up to T_{train} for each of the 100 polymer trajectories. We therefore obtain a larger set of subtrajectories, used for *mini-batch* training, after random shuffling.

3D subtrajectories of length $m = 128$ are encoded, after flattening, with an MLP made of 2 hidden layers with 512 neurons, and mapped to a latent vector of size 12. This provides the input for two separate dense layers outputting the conditional mean vector and elements of the LDL^T decomposition. Training is executed with Adam optimizer with learning rate $\eta = 10^{-4}$, using *early-stopping*.
One can appreciate the statistical relevance of our results when comparing the empirical normalized autocorrelations (NACF) and PDFs of the NAR generated modes with respect to the MD ones (see Fig. 3). In addition, the comparison between the normal modes MD NACFs and their corresponding exponential fit makes it clear that a simple OU process, as proposed in the GLE solution, is not flexible enough to reproduce the desired dynamics. Our approach attempts therefore to approximate the conditional *non-Markovian* distribution in a parametric way. After optimizing on Eq. (5), we use Eq. (2) to generate C.o.M synthetic dynamics starting from 300 MD initial trajectories. As shown in Fig. 4, ML-GLE is able to capture TAD and predict the diffusion coefficient extrapolating on short trajectories of length t_{max}. This method is also *scalable*, as increasing system size would imply more subtrajectory data, consequently requiring even shorter simulations. The method's consistency was tested on all available temperatures, yielding satisfying results. Also note that the information about

[4]Training data is available at https://huggingface.co/datasets/gian-michele/meltBR/settings

the stationary PDFs or ACFs is not included in the training architecture or loss function and in spite of that the NAR model is able to reproduce the correct statistical properties, hinting at the fact that we are approximating the correct *coarse-grained* stochastic integrator for the normal modes.

5 Limitations

Training data comes from a CG system, which implies an upper bound on the method accuracy; this is determined by the quality of the CG approximation as compared to the *All-Atom* simulation.
We implicitly made the independence assumption on trajectories, but single polymers are not necessarily independent, as they come from the same system. Small correlations may be present when training on subtrajectory *batches*, effectively reducing the dataset diversity. This issue tends to be less relevant for larger systems, as more statistics on distant less correlated single polymers are available.
In addition, although with low probability, the *normal modes* empirical distribution tails indicate unrealistic out-of-distribution sampling. A probable diagnosis for this problem resides in the spatial symmetry of the Gaussian approximation when sampling on extreme values.

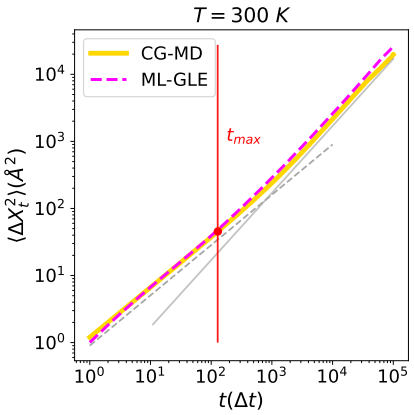


Figure 4: MSD at $T = 300\,\mathrm{K}$. Estimated diffusion coefficients with 95% confidence intervals, $\hat{D}_{MD} = (0.114 \pm 0.011)\mathring{A}^2/\Delta t$, $\hat{D}_{GLE} = (0.125 \pm 0.007)\mathring{A}^2/\Delta t$.

6 Conclusion

In this work, we presented an ML framework for *Transient Anomalous Diffusion* of single polymers in homopolymer melt simulations, predicting its transient time and diffusion coefficient. We develop a 3D neural autoregressive model for CVs and use physical symmetries to generate long-term stable dynamics, offering a first example of how GLE and deep learning can help accelerate slow diffusion phenomena in MD simulations. Future research includes investigating the model generalization capabilities: advanced architectures like Transformers networks are targeted. They could be trained on data at different thermodynamical coordinates, that would act as conditioning static features.

7 Broader Impact

Polymer melt is an umbrella term encompassing a large variety of highly viscous fluids displaying non-Newtonian and viscoelastic behaviour [24]. Although the only tested system was *Butadiene rubber*, our model is agnostic with respect to chemical details as it exploits physical variables pertaining to single polymers. Consequently, we expect any homopolymer melt simulation at equilibrium could potentially benefit from this work. Related to the broader field of polymer physics and material design, we speculate this approach could be useful to devise better dynamical CG schemes for full-size polymer simulations since the mapping monomers-modes is bijective and allows therefore for polymer configuration reconstruction.
Non-Markovian dynamics is the norm when dealing with complex systems, especially when only some variables are observable. Consequently, we suppose that the neural autoregressive model presented here could be appropriate when one wishes to access the asymptotic statistical properties of *mean-reverting*[5] processes, from either physical simulations or real-world data (e.g. biological or financial time series).
Furthermore, anomalous diffusion has been discovered in numerous systems, mostly in disordered media [26]. Some interesting examples include external tracer particles in biological cells [27, 28], and in artificially crowded systems [29, 30]. Protein simulations studies have reported anomalous subdiffusive behaviour as well [31–33]. All these systems exhibit transient behaviour and therefore this approach could pave the way for data-driven modeling of diffusing macromolecules within the GLE framework, both from *in vitro* and *in vivo* data sources [34].

[5] A stochastic process which asymptotical reverts to its mean

8 Acknowledgements

This work was performed at SimatLab, A joint public-private laboratory dedicated to the modeling of polymeric materials. This laboratory is supported by Michelin, Clermont-Auvergne University (UCA), CHU Clermont and CNRS. SimatLab acknowledges support received from the *Agence Nationale de la Recherche* of the French government through the program "Investissements d'Avenir" (grant no. 16-IDEX-0001 CAP 20-25).
We are grateful to the Mésocentre Clermont-Auvergne for providing help, computing and storage resources. All Computations have been performed on the supercomputer facilities of the Mésocentre Clermont-Auvergne of the Université Clermont Auvergne.

References

[1] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller. Schnet – a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148 (24):241722, Jun 2018. doi: 10.1063/1.5019779. URL https://pubs.aip.org/aip/jcp/article/962591.

[2] Jiang Wang, Simon Olsson, Christoph Wehmeyer, Adrià Pérez, Nicholas E. Charron, Gianni De Fabritiis, Frank Noé, and Cecilia Clementi. Machine learning of coarse-grained molecular dynamics force fields. *ACS Central Science*, 5(5):755–767, May 2019. doi: 10.1021/acscentsci.8b00913.

[3] Wujie Wang and Rafael Gómez-Bombarelli. Coarse-graining auto-encoders for molecular dynamics. *npj Computational Materials*, 5(1):125, Dec 2019. doi: 10.1038/s41524-019-0261-5. URL https://www.nature.com/articles/s41524-019-0261-5.

[4] Zakiya Shireen, Hansani Weeratunge, Adrian Menzel, Andrew W. Phillips, Ronald G. Larson, Kate Smith-Miles, and Elnaz Hajizadeh. A machine learning enabled hybrid optimization framework for efficient coarse-graining of a model polymer. *npj Computational Materials*, 8 (1):224, Nov 2022. doi: 10.1038/s41524-022-00914-4. URL https://www.nature.com/articles/s41524-022-00914-4.

[5] Joseph Bakarji and Daniel M. Tartakovsky. Data-driven discovery of coarse-grained equations. *Journal of Computational Physics*, 434:110219, 2021. doi: 10.1016/j.jcp.2021.110219. URL https://linkinghub.elsevier.com/retrieve/pii/S0021999121001145.

[6] Alain Dequidt and Jose G. Solano Canchaya. Bayesian parametrization of coarse-grain dissipative dynamics models. *The Journal of Chemical Physics*, 143(8), Aug 2015. doi: 10.1063/1.4929557. URL https://pubs.aip.org/jcp/article/143/8/084122/73206/Bayesian-parametrization-of-coarse-grain.

[7] Robert Zwanzig. Nonlinear generalized langevin equations. *Journal of Statistical Physics*, 9(3): 215–220, 1973. doi: 10.1007/BF01008729. URL http://link.springer.com/10.1007/BF01008729.

[8] Hazime Mori. Transport, collective motion, and brownian motion, 1965. URL https://academic.oup.com/ptp/article-lookup/doi/10.1143/PTP.33.423.

[9] Robert Zwanzig and Mordechai Bixon. Hydrodynamic theory of the velocity correlation function. *Physical Review A*, 2(5):2005–2012, Nov 1970. doi: 10.1103/PhysRevA.2.2005. URL https://link.aps.org/doi/10.1103/PhysRevA.2.2005.

[10] Daniel Molina-Garcia, Trifce Sandev, Hadiseh Safdari, Gianni Pagnini, Aleksei Chechkin, and Ralf Metzler. Crossover from anomalous to normal diffusion: truncated power-law noise correlations and applications to dynamics in lipid bilayers. *New Journal of Physics*, 20(10): 103027, Oct 2018. doi: 10.1088/1367-2630/aae4b2. URL https://iopscience.iop.org/article/10.1088/1367-2630/aae4b2.

[11] Scott A. McKinley, Lingxing Yao, and M. Gregory Forest. Transient anomalous diffusion of tracer particles in soft matter. *Journal of Rheology*, 53(6):1487–1506, 2009. doi: 10.1122/1.3238546. URL https://pubs.aip.org/sor/jor/article/53/6/1487-1506/240578.

[12] Francesca Grogan, Huan Lei, Xiantao Li, and Nathan A. Baker. Data-driven molecular modeling with the generalized langevin equation. *Journal of Computational Physics*, 418:109633, 2020. doi: 10.1016/j.jcp.2020.109633. URL https://linkinghub.elsevier.com/retrieve/pii/S0021999120304071.

[13] Kookjin Lee and Kevin T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *Journal of Computational Physics*, 404: 108973, 2020. doi: 10.1016/j.jcp.2019.108973. URL https://linkinghub.elsevier.com/retrieve/pii/S0021999119306783.

[14] Alexandre Chorin and Panagiotis Stinis. Problem reduction, renormalization, and memory. *Communications in Applied Mathematics and Computational Science*, 1(1):1–27, Dec 2006. doi: 10.2140/camcos.2006.1.1. URL http://msp.org/camcos/2006/1-1/p01.xhtml.

[15] M. Berkowitz, J. D. Morgan, and J. Andrew McCammon. Generalized langevin dynamics simulations with arbitrary time-dependent memory kernels. *The Journal of Chemical Physics*, 78(6):3256–3261, Mar 1983. doi: 10.1063/1.445244. URL https://pubs.aip.org/jcp/article/78/6/3256/820801/Generalized-Langevin-dynamics-simulations-with.

[16] John Fricks, Lingxing Yao, Timothy C. Elston, and M. Gregory Forest. Time-domain methods for diffusive transport in soft matter. *SIAM Journal on Applied Mathematics*, 69(5):1277–1308, 2009. doi: 10.1137/070695186. URL http://epubs.siam.org/doi/10.1137/070695186.

[17] Aram Davtyan, James F. Dama, Gregory A. Voth, and Hans C. Andersen. Dynamic force matching: A method for constructing dynamical coarse-grained models with realistic time dependence. *The Journal of Chemical Physics*, 142(15):154104, Apr 2015. doi: 10.1063/1.4917454. URL https://pubs.aip.org/jcp/article/142/15/154104/212038/Dynamic-force-matching-A-method-for-constructing.

[18] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Langevin equation with colored noise for constant-temperature molecular dynamics simulations. *Physical Review Letters*, 102(2): 020601, Jan 2009. doi: 10.1103/PhysRevLett.102.020601. URL https://link.aps.org/doi/10.1103/PhysRevLett.102.020601.

[19] Michele Ceriotti, Michele Parrinello, Thomas E. Markland, and David E. Manolopoulos. Efficient stochastic thermostatting of path integral molecular dynamics. *The Journal of Chemical Physics*, 133(12):124104, Sep 2010. doi: 10.1063/1.3489925. URL https://pubs.aip.org/jcp/article/133/12/124104/955703/Efficient-stochastic-thermostatting-of-path.

[20] Norbert Schaudinnus, Björn Bastian, Rainer Hegger, and Gerhard Stock. Multidimensional langevin modeling of nonoverdamped dynamics. *Physical Review Letters*, 115(5):050602, Jul 2015. doi: 10.1103/PhysRevLett.115.050602. URL https://link.aps.org/doi/10.1103/PhysRevLett.115.050602.

[21] Alexandre J. Chorin and Fei Lu. Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics. *Proceedings of the National Academy of Sciences*, 112(32): 9804–9809, Aug 2015. doi: 10.1073/pnas.1512080112. URL https://pnas.org/doi/full/10.1073/pnas.1512080112.

[22] Pinchen Xie, Roberto Car, and Weinan E. Ab initio generalized langevin equations. *ArXiv*, 2022. doi: 10.48550/ARXIV.2211.06558. URL https://arxiv.org/abs/2211.06558.

[23] Scott A. McKinley and Hung D. Nguyen. Anomalous diffusion and the generalized langevin equation. *SIAM Journal on Mathematical Analysis*, 50(5):5119–5160, 2018. doi: 10.1137/17M115517X. URL https://epubs.siam.org/doi/10.1137/17M115517X.

[24] M. Doi and S. F. Edwards. *The theory of polymer dynamics*, volume 27. Oxford Science Publications, 1989. doi: 10.1002/pol.1989.140270706.

[25] R. L. Nkepsu Mbitou, F. Goujon, A. Dequidt, B. Latour, J. Devémy, R. Blaak, N. Martzel, C. Emeriau-Viard, J. Tchoufag, S. Garruchet, E. Munch, P. Hauret, and P. Malfreyt. Consistent and transferable force fields for statistical copolymer systems at the mesoscale. *Journal of Chemical Theory and Computation*, Oct 2022. doi: 10.1021/acs.jctc.2c00945. URL https://doi.org/10.1021/acs.jctc.2c00945.

[26] Jean-Philippe Bouchaud and Antoine Georges. Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics Reports*, 195(4–5):127–293, 1990. doi: 10.1016/0370-1573(90)90099-N. URL https://linkinghub.elsevier.com/retrieve/pii/037015739090099N.

[27] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical Review Letters*, 103(1):018102, Jul 2009. doi: 10.1103/PhysRevLett.103.018102. URL https://link.aps.org/doi/10.1103/PhysRevLett.103.018102.

[28] Georg Seisenberger, Martin U. Ried, Thomas Endreß, Hildegard Buning, Michael Hallek, and Christoph Brauchle. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science*, 294(5548):1929–1932, Nov 2001. doi: 10.1126/science.1064103. URL https://www.science.org/doi/10.1126/science.1064103.

[29] Kamilla Norregaard, Ralf Metzler, Christine M. Ritter, Kirstine Berg-Sørensen, and Lene B. Oddershede. Manipulation and motion of organelles and single molecules in living cells. *Chemical Reviews*, 117(5):4342–4375, Mar 2017. doi: 10.1021/acs.chemrev.6b00638. URL https://pubs.acs.org/doi/10.1021/acs.chemrev.6b00638.

[30] Jedrzej Szymanski and Matthias Weiss. Elucidating the origin of anomalous diffusion in crowded fluids. *Physical Review Letters*, 103(3):038102, Jul 2009. doi: 10.1103/PhysRevLett.103.038102. URL https://link.aps.org/doi/10.1103/PhysRevLett.103.038102.

[31] Matti Javanainen, Henrik Hammaren, Luca Monticelli, Jae-Hyung Jeon, Markus S. Miettinen, Hector Martinez-Seara, Ralf Metzler, and Ilpo Vattulainen. Anomalous and normal diffusion of proteins and lipids in crowded lipid membranes. *Faraday Discuss.*, 161:397–417, 2013. doi: 10.1039/C2FD20085F. URL http://xlink.rsc.org/?DOI=C2FD20085F.

[32] Kejia Chen, Bo Wang, and Steve Granick. Memoryless self-reinforcing directionality in endosomal active transport within living cells. *Nature Materials*, 14(6):589–593, 2015. doi: 10.1038/nmat4239. URL https://www.nature.com/articles/nmat4239.

[33] Jae-Hyung Jeon, Matti Javanainen, Hector Martinez-Seara, Ralf Metzler, and Ilpo Vattulainen. Protein crowding in lipid bilayers gives rise to non-gaussian anomalous lateral diffusion of phospholipids and proteins. *Physical Review X*, 6(2):021006, Apr 2016. doi: 10.1103/PhysRevX.6.021006. URL https://link.aps.org/doi/10.1103/PhysRevX.6.021006.

[34] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sørensen, Lene Oddershede, and Ralf Metzler. In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Physical Review Letters*, 106(4):048103, Jan 2011. doi: 10.1103/PhysRevLett.106.048103. URL https://link.aps.org/doi/10.1103/PhysRevLett.106.048103.