LATENT SEQUENCE GENERATION OF STEERED MOLECULAR DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper, we use a LSTM-VAE model framework in order to learn latent representations that are conditioned by potential energy through TorchMD, while being able to autoregressively generate sequences of a 10 deca-alanine system. While previous work have used generative deep learning methods for learning latent representations and predicting motion of molecules, this paper tackles with the latent representations for steered molecular dynamics (SMD).

1 INTRODUCTION

Molecular Dynamics (MD) is the methodology of simulating the forces that are exerted in a particle. MD does this by integrating Newton's equations of motion in order to evolve a system through time. By understanding the dynamics of the system, researchers can be able to model how certain entities are able to interact with each other and create function. In particular, the structure and dynamics of proteins is necessary in order to understand how proteins function.

However, in order to investigate conformational states of a protein through MD, there is a high computational cost. In order to remedy such computational tasks, research has been done in advanced sampling techniques to lessen the number of computations. One technique is to utilize steered molecular dynamics (SMD) Izrailev et al. (1999) in order to accelerate transitions from one state to another by adding a force bias. But inducing bias such that a molecular can go from state A to state B is not enough, as not every path would be biologically relevant due to how much energy is needed for the molecule to cross certain potential energy barriers. Thus, a minimum-action or low energy pathway must be extracted, usually from expert knowledge.

Previously, there has been work in using Reinforcement Learning in order to use steered molecular dynamics as an environment and learn perturbations from state A to state B that minimizes the amount of potential energy (work) Ho et al. (2022b). However, there are issues with having steered molecular dynamics, because it becomes a bottleneck as it utilizes MD, which requires to integrate equations of motion. Thus, a data-driven model would a be step in making this faster and also hopefully more generalizable.

Work has been done though in making data driven models to molecular dynamics. One work involves using LSTMs of a history of raw data of the system and then predicting it in a future time step Gupta et al. (2020). Another utilizes cGANs in order to condition in a specific collected variable, which is a numerical value that describes the molecule Cava et al. (2021). And another paper used GANs in order to learn simultaneously the mappings of 3d coordinates to latent space, and the evolution of one latent time step to another Endo et al. (2018).

In this paper, we diverge from previous work in which GANs were used to generate pathways, and instead use Variational Autoencoders Kingma & Welling (2013) in order to learn a latent representation of all the different protein states of a protein system called 10 deca-alanine Park et al. (2015).

2 EXPERIMENTS

2.1 DATASET

The data used is a 10 deca-alanine molecule Park et al. (2015) was simulated by biased forces in order to stretch. While the 10 deca-alanine molecule is made up of 104 atoms, for the purposes of

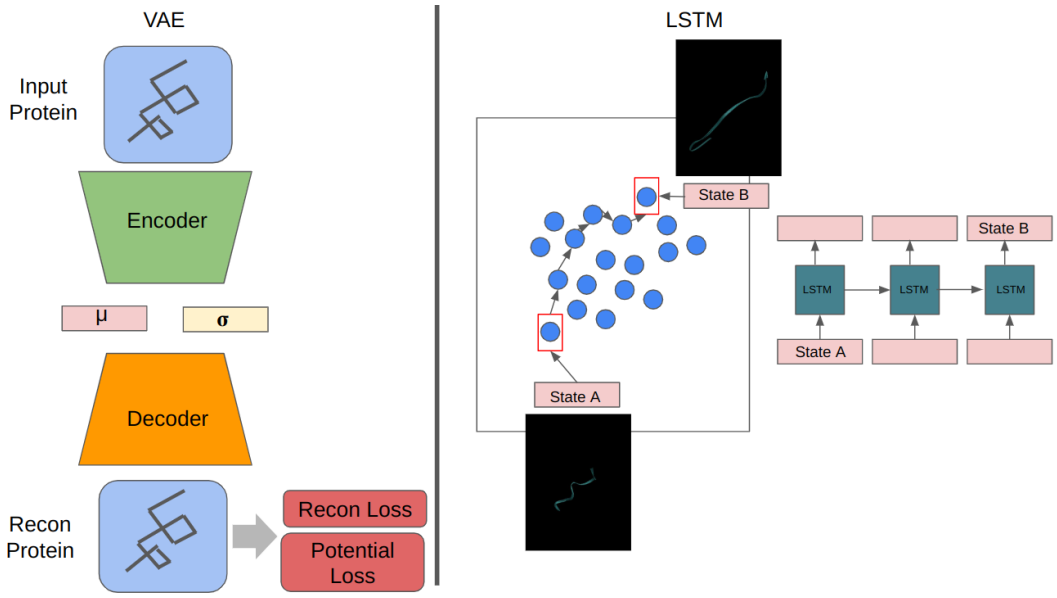

Figure 1: The 2-step training approach. Step 1 (Left) The VAE is trained to learn the latent representation of the molecule. Step 2 (Right) The latent representations of the molecule are used to encode the sequence of the molecular simulation and train a LSTM to autoregressively predict the next latent representation.

the experiments in this paper we use the backbone of the 10 deca-alanine which is 40 atoms. We also do data augmentation to avoid model overfitting, by rotating the molecule to generate more data.

2.2 Approach and Training Details

2.2.1 Step 1: Learning low dimensional representations for molecules using VAE

Instead of utilizing a GAN as works used in MolGAN De Cao & Kipf (2018), Multi-tme step GANs Endo et al. (2018), and cGANs for minimum potential pathways Cava et al. (2021), we use a VAE to encode the representations of the molecule. This was done in order to avoid issues with Cava et al. (2021) in which the generator of the cGAN was conditioned to a numerical value that represents a certain structure through time. This doesn't allow for autoregressive generation of a sequence. While Endo et al. (2018) uses a GAN to traverse through latent space in time and map to the real 3d space, the GAN isn't constrained by a physics loss function. Thus, we propose the VAE to learn the representations alongside the potential energy.

For the VAE, both the encoder and decoder are feedforward neural networks with three fully connected layers. For the encoder the network is: [Linear(120, 100), Linear(100, 100), Linear(100,60)]. For the decoder the network is: [Linear(60, 100), Linear(100,100), Linear(100,120)]. The FC layers are accompanied with dropout layer of p=0.05, expect for the last layers. The latent space dimension is fixed to 60 in our experiments. We train the VAE for 8 epochs.

2.2.2 Step 2: Traversing latent space using LSTMs

We use a LSTM to traverse the latent representations of the molecule from state A to state B, as described in Figure 1. Once the VAE is trained, and we have a pretrained encoder, we train the LSTM on a sequence of encoded latent representations of a molecular simulation. For example, if we have a sequence of 10 frames of 10 deca-alanine, we can encode each frame with the encoder which are then used to train the LSTM to then predict an "action vector" in the latent space. This action vector defines the direction to go to towards the next targeted latent representation. We can do this autoregressively until it hits a predetermined amount of steps we want the LSTM to run, and

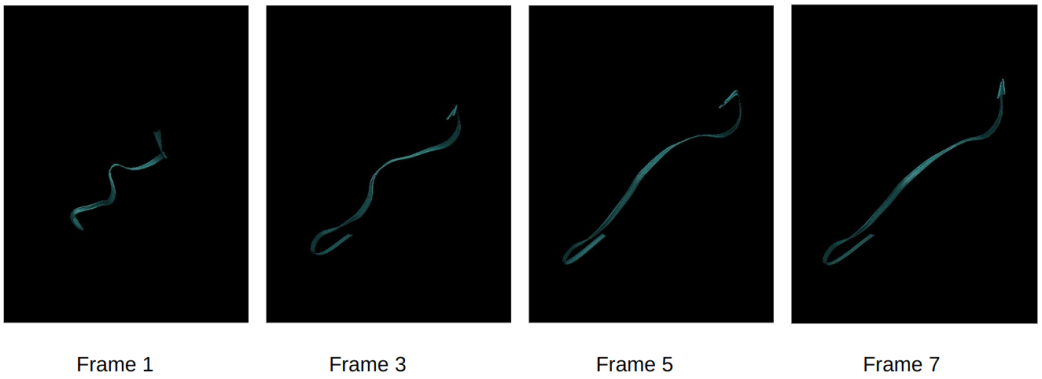

Figure 2: The generated frames from the LSTM-VAE model. The samples illustrate the stretching of the 10 deca-alanine in the End2End distance collected variable.

then compare the last predicted latent vector with the latent representation of state B. We train the LSTM for 22 epochs.

Once the training has been completed, we can then do autoregressive inference by taking an encoded latent representation of a beginning state, such as State A, and then predict for n number of steps. An example of generated sequence is shown in Figure 2.

2.3 Loss Objective

We train the VAE using TorchMD Doerr et al. (2021), to train the model to reconstruct a biologically relevant molecule. In order to do that we define a loss function on the potential energy and a collected variable loss. The potential energy loss is defined on each frame/state of the molecule, whereas our custom collected variable (CV) loss defined on the distance between the atoms at the ends of the 10 deca-alanines, specifically the End2End distance. Since, we are working with a backbone, we only consider potential energies of bonds (V_{bonds}), angles (V_{angles}), dihedrals ($V_{torsion}$). The combined loss function is given below:

$$Loss = Loss_{PE} + Loss_{CV} + Loss_{Recon} \tag{1}$$

$$Loss_{PE} = MSE_{angles} + MSE_{bonds} + MSE_{torsion} \tag{2}$$

Here the potential energy (PE) loss minimize the difference between the potential energy of the input molecule frame and the corresponding reconstructed molecule frame. In addition, to avoid exploding gradients, we take the log of the MSE loss.

The loss objective for the LSTM is defined as the MSE between the encoded true molecular dynamics simulation and the autoregressively generated latent sequence.

3 Discussion

In this paper, we developed a two stage deep learning approach for steered molecular dynamics application. We trained a VAE to to first learn the latent representations of the molecule conditioned with reconstructed molecule potential energies provided by TorchMD.

In order to generate the sequence of frames, we use latent encoding of the target frames and train an LSTM to traverse a trajectory of sequences in the latent space.

In this experiment, for training LSTM we used 100 frames, but for purposes of illustration, we sample a sequence of 10 frames. In Figure 2, we visualize a sequence of a 10 deca-alanine through Visual Molecular Dynamics (VMD) Humphrey et al. (1996). It can be seen that there is a big difference in stretching between Frame 1 and Frame 5. Looking at the generated frames, we can see that the ends of the molecule are closer together and are violating physical constraints. This

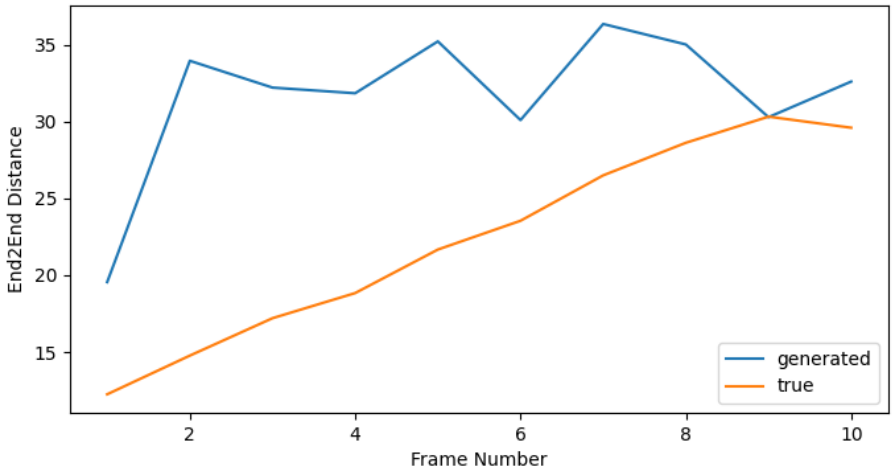

Figure 3: The End2End distance plot of the generated and true sequence frames.

indicates how there are competing objectives with training the model. In Figure 3, we illustrate the difference in collected variable between the generated sequence and the true molecule sequence. We observe that the generated sequence has a difference in End2End distance between Frame 1 to 2, and then stays relatively constant, which corroborates with Figure 1 in the sequence of frames. Since we are trying to learn physics that is biased, that inherently makes the competing losses of minimizing potential energy and minimizing the MSE of the predicted and target frames' collected variable difficult. Overall, we show that we are able to generate a sequence from a starting frame A to an ending frame B autoregressively and achieve change in a particular collected variable - namely, End2End distance.

However, we believe that there are a few limitations in the current experimental setup. Specifically, the architecture used is a feedforward neural network that limits the model to a specific molecule or system. However, in literature, Graph Neural Networks have been used for molecule representations Winter et al. (2022)as well as to simulate physics by predicting the velocity Sanchez-Gonzalez et al. (2020) and potential energy Thölke & De Fabritiis (2022) but not in a protein steered molecular simulation. We believe that using a GNN model would allow us to improve the learned latent space representations.

In the future, we plan to explore Latent Diffusion Models Zeng et al. (2022); Ho et al. (2022a) to improve the diversity in the generated protein frames.

REFERENCES

John Kevin Cava, John Vant, Nicholas Ho, Ankita Shukla, Pavan Turaga, Ross Maciejewski, and Abhishek Singharoy. Towards conditional generation of minimal action potential pathways for molecular dynamics, 2021. URL https://arxiv.org/abs/2111.14053.

Nicola De Cao and Thomas Kipf. Molgan: An implicit generative model for small molecular graphs. 2018. doi: 10.48550/ARXIV.1805.11973. URL https://arxiv.org/abs/1805.11973.

Stefan Doerr, Maciej Majewski, Adrià Pérez, Andreas Krämer, Cecilia Clementi, Frank Noe, Toni Giorgino, and Gianni De Fabritiis. TorchMD: A deep learning framework for molecular simulations. *Journal of Chemical Theory and Computation*, 17(4):2355–2363, mar 2021. doi: 10.1021/acs.jctc.0c01343. URL https://doi.org/10.10212Facs.jctc.0c01343.

Katsuhiro Endo, Katsufumi Tomobe, and Kenji Yasuoka. Multi-step time series generator for molecular dynamics. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr. 2018. doi: 10.1609/aaai.v32i1.11863. URL https://ojs.aaai.org/index.php/AAAI/article/view/11863.

Chitrak Gupta, John Kevin Cava, Daipayan Sarkar, Eric Wilson, John Vant, Steven Murray, Abhishek Singharoy, and Shubhra Kanti Karmaker. Mind reading of the proteins: Deep-learning to forecast molecular dynamics. *bioRxiv*, pp. 2020–07, 2020.

Jonathan Ho, Tim Salimans, Alexey Gritsenko, William Chan, Mohammad Norouzi, and David J Fleet. Video diffusion models. *arXiv preprint arXiv:2204.03458*, 2022a.

Nicholas Ho, John Kevin Cava, John Vant, Ankita Shukla, Jake Miratsky, Pavan Turaga, Ross Maciejewski, and Abhishek Singharoy. Learning free energy pathways through reinforcement learning of adaptive steered molecular dynamics. *bioRxiv*, 2022b. doi: 10.1101/2022.10.04.510845. URL https://www.biorxiv.org/content/early/2022/10/05/2022.10.04.510845.

William Humphrey, Andrew Dalke, and Klaus Schulten. Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

Sergei Izrailev, Sergey Stepaniants, Barry Isralewitz, Dorina Kosztin, Hui Lu, Ferenc Molnar, Willy Wriggers, and Klaus Schulten. Steered molecular dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas: Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21–24, 1997*, pp. 39–65. Springer, 1999.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Sanghyun Park, Fatemeh Khalili, and Johan Strümpfer. Stretching deca-alanine, 2015.

Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *International conference on machine learning*, pp. 8459–8468. PMLR, 2020.

Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.

Robin Winter, Marco Bertolini, Tuan Le, Frank Noé, and Djork-Arné Clevert. Unsupervised learning of group invariant and equivariant representations. *arXiv preprint arXiv:2202.07559*, 2022.

Xiaohui Zeng, Arash Vahdat, Francis Williams, Zan Gojcic, Or Litany, Sanja Fidler, and Karsten Kreis. Lion: Latent point diffusion models for 3d shape generation, 2022. URL https://arxiv.org/abs/2210.06978.