LATENT SEQUENCE GENERATION OF STEERED MOLECULAR DYNAMICS

John Kevin Cava†,Ankita Shukla†,John Vant†,Shubhra Karmaker‡,Pavan Turaga†,

Ross Maciejewski†,Abhishek Singharoy†
Arizona State University†, Auburn University‡

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 exerted on a particle. MD does this by integrating Newton's equations of motion to evolve a system through time. By understanding the system's dynamics, researchers can model how certain entities can interact with each other and create functions. In particular, the structure and dynamics of proteins are necessary to understand how proteins function.

However, to investigate the 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) to accelerate transitions from one state to another by adding a force bias. However, inducing bias so that a molecule can go from state A to state B is insufficient. However, only some paths would be biologically relevant due to how much energy is needed for the molecule to cross specific potential energy barriers. Thus, a minimum-action or low-energy pathway must be extracted, usually from expert knowledge

Previously, Reinforcement Learning has been used to steer 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 integrating equations of motion. Thus, a data-driven model is a step in making this faster and more generalizable.

Previous works have developed data-driven models for molecular dynamics. One work involves using LSTMs with input being a history of raw data of the system and then predicting it in a future time step Gupta et al. (2020). Another utilizes cGANs to condition in a specific collected variable, a numerical value that describes the molecule Cava et al. (2021). Moreover, another paper used GANs to learn the mappings of 3d coordinates simultaneously 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) to learn a latent representation of all the different protein states of a protein system called 10 deca-alanine Park et al. (2015).

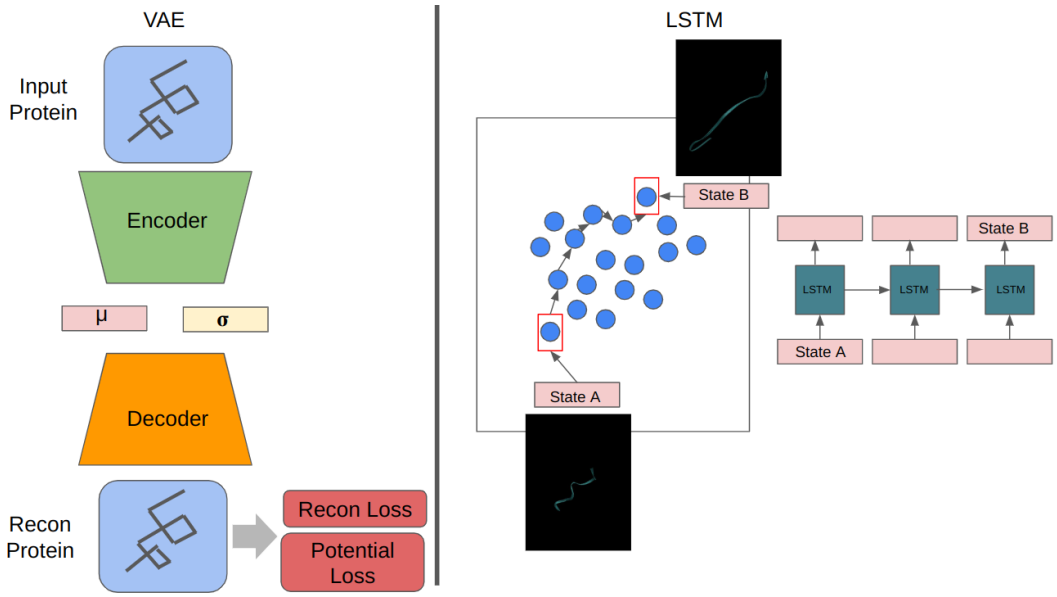

Figure 1: The 2-step training approach. Step 1 (Left) The VAE is trained to learn the latent representation of the molecule. Step 2 (Right) The latent representations of the molecule are used to encode the sequence of the molecular simulation and train a LSTM to autoregressively predict the next latent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

2 Experiments

2.1 Dataset

The data used is a 10 deca-alanine molecule Park et al. (2015) that was simulated by biased forces in order to stretch from an initial unstretched conformation. While the 10 deca-alanine molecule is made up of 104 atoms, for the purposes of the experiments in this paper we use the backbone of the 10 deca-alanine which consists of 40 atoms. We do data augmentation to avoid model overfitting, by rotating the molecule to generate more data. The system of 10 deca-alanine was used due to its illustrative purpose that traversing the End2End distance CV is not normally energetically favorable, and thus exemplifies itself as a toy example of whether a model can learn to keep the molecule biophysically plausible while traversing a path it is unlikely to happen. This is then useful in sampling out of distribution paths.

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 previously 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 physically-inspired 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 lay-

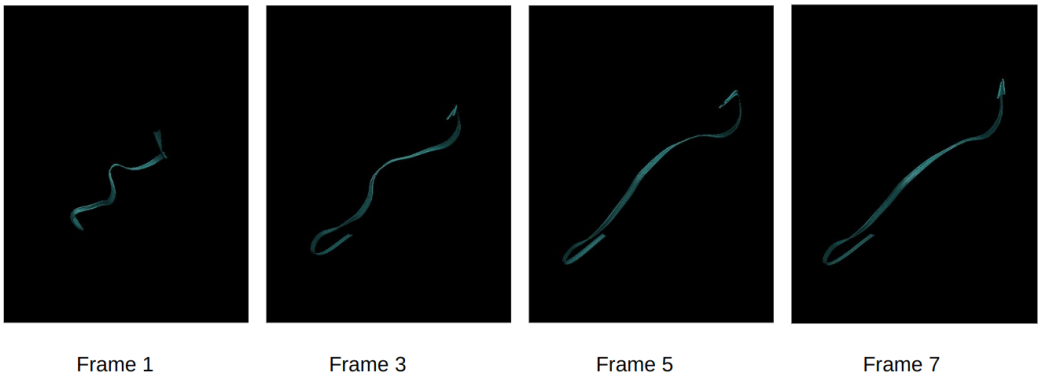

Figure 2: The generated frames from the LSTM-VAE model. The samples illustrate the stretching of the 10 deca-alanine in the End2End distance collected variable from left to right.

ers are accompanied with dropout layer of $p = 0.05$, expect for the last layer. The latent space dimension is fixed to 60 in all our experiments. We train the VAE for 8 epochs.

2.2.2 Step 2: Traversing the latent space using LSTMs

We use a LSTM to traverse the latent space of the molecule from state A to state B, as shown in Figure 1. Once the VAE is trained, and we have a pre-trained encoder, we train the LSTM on a sequence of encoded latent representations of a molecular simulation. For example, for a sequence of 10 frames of 10 deca-alanine, we encode each frame with the encoder which is then used to train the LSTM and predict an "action vector" in the latent space. This action vector defines the direction to go towards the next targeted latent representation. We do this autoregressively for a fixed number of steps and 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 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 model using TorchMD Doerr et al. (2021), 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 is defined on the distances between the atoms at the ends of the 10 deca-alanine, specifically the End2End distance. The main motivation in incorporating collected variables (CVs) as a new loss is that in respect with SMD, CVs are useful in determining how a simulation (trajectory) is progressing. This is important due to the fact that if we were to only minimize potential energy, then it is impossible to traverse a CV since it would be steering the molecule out of a conformation that is energetically favorable. Thus a linear combination of loss between reconstruction, potential energy, and the CV is necessary. 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 a steered molecular dynamics application. We first trained a VAE to 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 encodings of the target frames and train an LSTM to traverse a trajectory of sequences in the latent space.

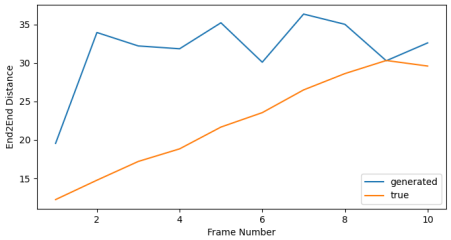

Figure 3: The End2End distance plot of the generated and true sequence frames. The blue line represents the generated sequence, while the orange line is an example ground-truth sequence. The generated sequence shows a change in its End2End distance which indicates stretching of the 10 deca-alanine.

The main motivation in creating this two stage deep learning approach is to mimic work done by Tian et al. (2022) in which a MLP autoencoder was used to sample outlier frames to be used as seeds for further MD simulations. This has been shown to be more efficient than sampling through MD alone. In this paper, we use autoregressive generation from the latent representations to generate "seeds" of sequential frames, instead of just one. This motivation takes inspiration from a SMD technique called "string simulations" Pan et al. (2008) in which from State A to State B you have an initial trajectory with initial frames that act as seeds, and the MD samples new frames for the trajectory. This is done iteratively until a "novel" path has been found.

However, in the case of this paper, we want to explore whether or not it is possible to generate a sequence of frames that reflects non-equilibrium steered molecular dynamics. Based on the results, it is possible for a two stage deep learning model to both learn latent representations, and learn how to traverse the latent representations to generate a sequence of frames which could be considered a trajectory. In Figure 2, we visualize 10 deca-alanine through Visual Molecular Dynamics (VMD) Humphrey et al. (1996). In a qualitative inspection, the model is able to generate a trajectory that is non-equilibrium. In a quantitative analysis, the CV of End2End distance is plotted against the generated frames of the 10 deca-alanine in Figure 3. 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, this inherently makes the competing losses of both minimizing potential energy and minimizing the MSE between 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, due to lack of invariance to rotation and translation. But in terms of other literature such as Tian et al. (2022), the use of a feedforward neural network was sufficient. 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). Though to our knowledge, we haven't seen graph autoencoding generation of steered molecular dynamics. We believe that using a GNN model will 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.

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

Albert C Pan, Deniz Sezer, and Benoît Roux. Finding transition pathways using the string method with swarms of trajectories. *The journal of physical chemistry B*, 112(11):3432–3440, 2008.

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.

Hao Tian, Xi Jiang, Sian Xiao, Hunter La Force, Eric C Larson, and Peng Tao. Last: Latent space-assisted adaptive sampling for protein trajectories. *Journal of Chemical Information and Modeling*, 63(1):67–75, 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.