Enhancing the identification of collective variables and their interpretability using machine learning in molecular dynamics

Gohar Ali Siddiqui[1], Prof. Dr. Alessio Gagliardi[1]
[1]Simulations of Nanosystems for Enery Conversion, Technical University of Munich, Hans-Piloty-Strasse 1, 85748 Garching

gohar.siddiqui@tum.de

Molecular dynamics simulations are widely used in several fields of science to study phenomena at the atomistic level e.g. protein folding and ligand binding. However, as soon as an event has an associated energy barrier higher than the thermal energy k_BT, it becomes difficult to sample it in the limited simulation time available with the current computational capabilities [1]. To alleviate this problem, a family of methods have been developed which enhance the sampling of configurational space by adding an external bias to the slowest modes of the system [2]. Regardless of the method, finding the optimal modes (called collective variables or CVs) to bias becomes the crucial to the success of enhanced sampling.

We present here an iterative approach to train a β-Variational Autoencoder (VAE) to find the relevant collective variables of any molecular system. Furthermore, we expand on how to select the dimensionality of the latent space and perform correlation analysis to learn more about the found collective variables. The method is applied on Alanine Dipeptide and a 10-residue miniprotein, Chignolin.

Effect of Solvent

The effect of solvent is crucial in many chemical and biological systems. It can lead to significant changes in free energy landscape of the system. Thus, to input a more complete description of the system into the VAE, we need to incorporate the effect of solvent in the input vector. Taking inspiration from [3], we use a metric of solvent concentration at certain points around the molecule and concatenate these to the input vector. The value of each of these descriptors is defined in the following way;

$$d_i^0 = \sum_j^{r_{ij}<r_{NL}} \frac{1-\left(\frac{r_{ij}}{r_0}\right)^n}{1-\left(\frac{r_{ij}}{r_0}\right)^m} \qquad (1)$$

with r_{NL} being the radius of the descriptor and r_0 being an inner radius, able to fit one water molecule in the center. r_{ij} describes the distance between the descriptor d_i and the oxygen atom of water molecule j.

Interpretation of the collective variables

Interpretation of the multidimensional latent space mapped by the trained VAE is not straightforward. In a completely unknown system one could create and look at trajectories translating in one dimension at a time and see their collective effect. The correlation of the latent space dimensions with the input variables can also provide insights. Since we use general coordinates, this would be useful only if the latent space dimension correlate highly to a distance change. Nevertheless, collective motions will be difficult to interpret just with the correlations. In the systems we studied, we already have some idea of the collective motions, so we can use this knowledge to assess if the latent space dimensions actually converge to some interpretable variable

References

[1] Peters, B. *Reaction Rate Theory and Rare Events* (Elsevier, Oxford, NETHERLANDS, THE, 2017).
[2] Bussi, G. *et al.* Using metadynamics to explore complex free-energy landscapes. en. *Nature Reviews Physics* **2.** Number: 4 Publisher: Nature Publishing Group, 200–212 (Apr. 2020).
[3] Rizzi, V. *et al.* The role of water in hostguest interaction. *Nature Communications* **12,** 93 (2021)

Figures

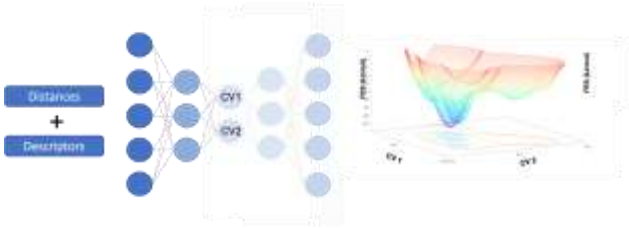
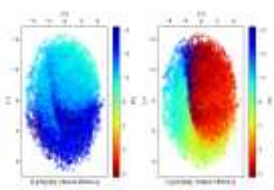
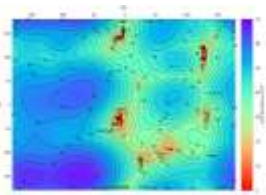


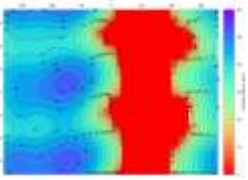
Figure 1. Enhanced sampling using Autoencoder network. Inputs are distances and solvation descriptors and output is used to accelerate sampling of configurational space using on-the-fly probability enhanced sampling



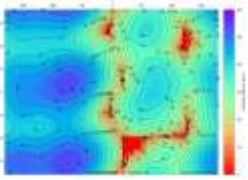
(a) Latent space of trained VAE on Alanine



(b) After 30 nanoseconds with bias.



(c) After 30 nanoseconds without bias.



(c) After 100 nanoseconds without bias.

Figure 2. Enhanced sampling of Alanine Dipeptide

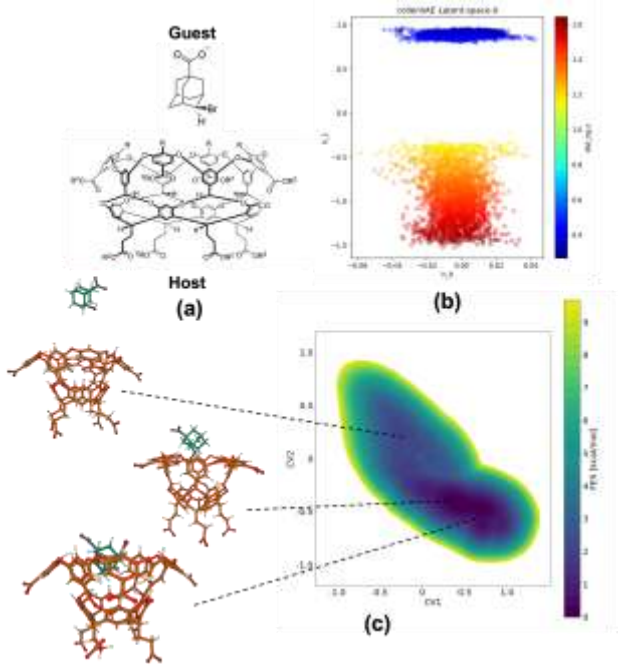


Figure 3. (a) The host-guest system (b) Latent space of VAE trained for host-guest system (c) The free energy surface of host-guest system in the latent space of