Approximating Free Energy Functions with Machine Learning

Dylan Kupetsky
Mentor: Dr. Huan Lei

Introduction

While free energy functions are often described as the capacity of a system to do work, in application, these functions provide information on biological, thermodynamic, or chemical processes [1]. An important example is the perturbations related to mutations of superoxide dismutase 1, which is suspected to be related to the development of ALS disease [2]. Sampling and reconstruction of these functions therefore can make profound impacts on several areas of research. In this project, we are attempting to simplify the process of approximating such functions by allowing machine learning to construct the equations. We use the Mueller-Brown potential as a benchmark example of this process.

Numerical challanges

In the use of traditional methods, such as radial basis functions, approximating the Mueller-Brown potential works quite well, with Maragliano and Vanden-Eijnden et. Al. citing an error of 4.2e-3. However, in mapping a more complicated, higher dimensional function such as Alanine Depeptide in four torsion angles, the mean relative error is found to be 0.14, nearly 100 times the error of the 3D Mueller-Brown potential. Where machine learning can assist in this error is the fact that upgrading dimensionality with neural networks is quite simple and typically results in little changes to accuracy. In only remains to be seen how this process works and how to sample learning points effectively.

Direct approximation

The first part of this project consisted of experimentation with the possibility that free energy functions may be approximated with neural networks. We begin with some input data, usually random numbers in $-2 \leq x \leq 1$ and $-0.25 \leq y \leq 1.75$ (see figures 1 and 2). We then create a neural network consisting of 3 to 4 hidden layers with depths of 300 neurons and utilize a gradient loss function defined by:

$$\mathcal{L} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\nabla V_{MB}(x_1,x_2) - \nabla\widetilde{V_{MB}}(x_1,x_2)\right)^2} \qquad (1)$$

We use a gradient for the loss function to achieve maximum information in the loss function (as the gradient provides information on magnitude and direction at the sample point).
Within the neural network, we make use of the following customized activation function:

$$g(x) = \frac{2}{1+e^{-Bx}} - 1 \quad , \qquad (2)$$

With B = 1.
Despite some success with this method, there remains the issue of sampling. If in the learning process, the network receives samples from the strongly positive region, the function will not be learned correctly. Therefore, we must implement a better sampling method to receive accurate results dependably.

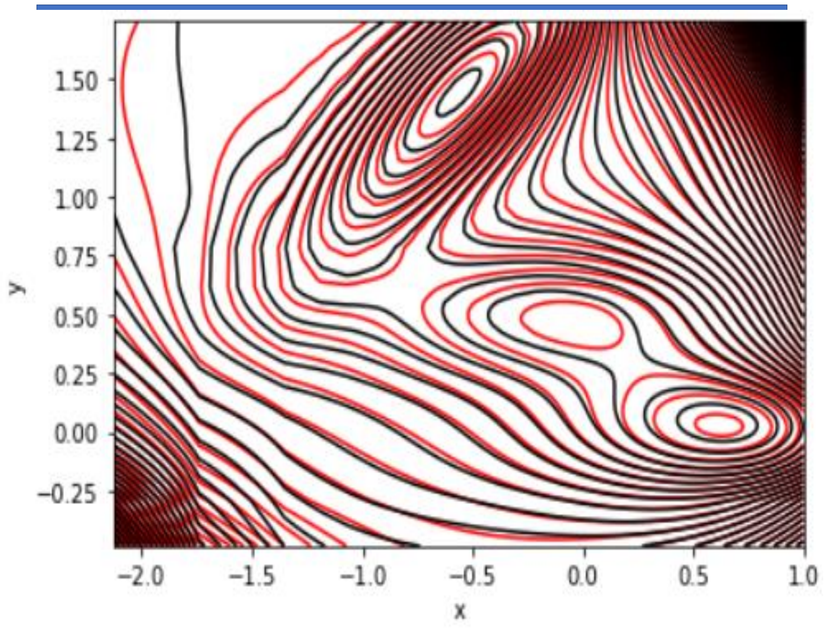

Figure 1: Mueller-Brown potential reconstruction using direct approximation with a TensorFlow neural network. In red is the original Mueller-Brown potential, and the black dashed line is the approximation using TensorFlow. Error = 500, calculated using equation (1). Neural network has a width of 4, depth of 300, 1 epoch of runtime, 40,000 x,y pairs simulated, with an Adam learning rate of .001.

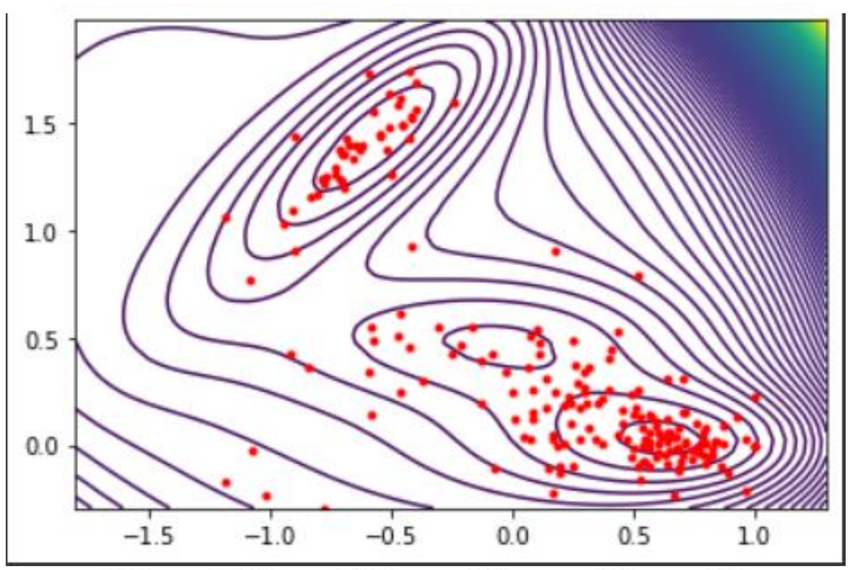

Figure 2: Sampling of the true Mueller-Brown potential using the artificial temperature trajectory. Origin point is (x0,y0) = (1,0), timestep 10^-5, 4e4 steps, friction coefficient 1, and temperature 17e23 K. Instead of tens of thousands of steps, every 200th step is displayed for simplicity.

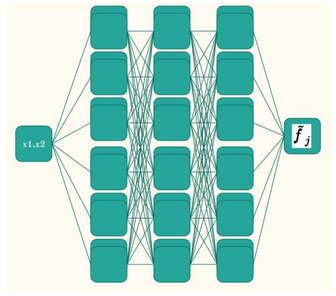

Figure 3: An example of a neural network for approximating a function. As mentioned previously, increasing dimensionality with neural networks simply implies adding samples to the input vector, so that an upgrade from 3D to 4D might imply changing x1,x2 to x1,x2,x3.

Sampling methods

The main issue associated with effective sampling of a free energy function is how the dimensionality may be reduced. We are interested in the free energy minima, and with typical Langevin dynamics simulations, the timescale of exiting such a well is:

$$t^{(1)} = \mathrm{O}\left(\gamma e^{\beta \Delta F}\right) \quad , \qquad (3)$$

With gamma a friction coefficient and beta a measure of temperature. Maragliano and Vanden-Eijnden et. Al. propose a new set of equations utilizing a trajectory with an artificial temperature and friction coefficient, so that the new timescale may be adjusted to be of order 1 [3].
The Mueller-Brown potential is a special case of this change of variables where the artificial temperature is the true temperature. This landscape thus provides an easily testable benchmark of this sampling method.

References

[1] – Maragliano, L., & Vanden-Eijnden, E. (2008). Single-sweep methods for free energy calculations. The Journal of Chemical Physics, 128(18), 184110. https://doi.org/10.1063/1.2907241

[2] – Wells, N. G. M., Tillinghast, G. A., O'Neil, A. L., & Smith, C. A. (2021). Free energy calculations of ALS-causing SOD1 mutants reveal common perturbations to stability and dynamics along the maturation pathway. Protein Science: A Publication of the Protein Society, 30(9), 1804–1817. https://doi.org/10.1002/pro.4132

[3] – Luca Maragliano, Eric Vanden-Eijnden, A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations, Chemical Physics Letters, Volume 426, Issues 1–3, 2006, Pages 168-175, ISSN 0009-2614, https://doi.org/10.1016/j.cplett.2006.05.062. (https://www.sciencedirect.com/science/article/pii/S0009261406007408)

Acknowledgements

- NSF ACRES REU - OAC-1949912
- NSF DMS-2110981
- Thank you to Dr. Lei for the support
- Thank you to Zhiyuan She for helpful discussions