Neural Networks and Neural Operators for the Committor Problem

Alkis Boukas [1] Bill Chen [2] Mentor: Maria Cameron [3]

[1]Cornell University [2]Northwestern University [3]University of Maryland, College Park

Transition Path Theory

Consider a system governed by overdamped Langevin Dynamics

$$dx = -\nabla V(x)dt + \sqrt{2\beta^{-1}}dW$$

- $V(x)$ is a smooth potential, $\beta \propto \frac{1}{T}$, dW is Brownian noise

The committor function q gives the probability that a particle starting at x arrives at attractor B before attractor A.

q satisfies the Boundary Value Problem:

$$\beta^{-1}e^{\beta V}\nabla \cdot (e^{-\beta V}\nabla q) = 0, x \in \Omega$$
$$q = 0, x \in \partial A$$
$$q = 1, x \in \partial B$$

The committor function is integral to studying rare events in chemical reactions and nonlinear oscillators.

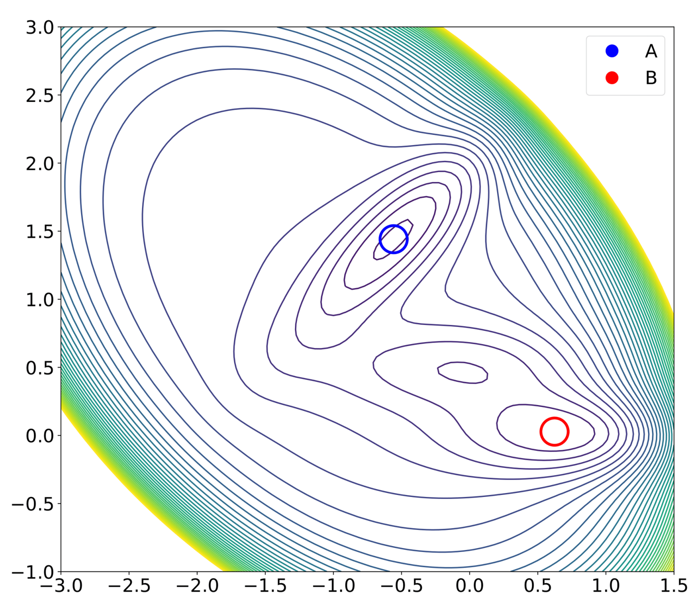

Figure 1. Attractors on Mueller Potential Contour Map

Neural Networks

A neural network of L layers is a sequence of compositions of the form

$$\mathcal{N}(x;\theta) = \sigma \circ \mathcal{L}_L \circ \cdots \circ \sigma \circ \mathcal{L}_1$$

- $\mathcal{L} = A\phi_\theta(x) + b$, where ϕ is an operator with parameters θ, A is a weight matrix, b are bias terms
- σ is an nonlinear activation function (ReLU, tanh) applied pointwise
- The size refers to the number of trainable parameters
- the depth refers to the number of activation layers

Training a neural network refers to optimizing the values of A, ϕ, b for each layer using a training set.

The model is then evaluated on a test set.

Project Objective

Adapt different neural network architectures to solve the committor problem cheaper, faster, or more accurately than traditional finite element methods (FEM).

Approach 1: Neural Operator

- Typical solvers for the committor (FEM, PINNs) solve the committor for one set of parameters (i.e. β)
- In contrast, neural operators learn the solution operator, which allows for quick computation of committors with different parameters
- We use the Fourier Neural Operator (FNO) architecture, a highly successful architecture that solves PDEs with high accuracy (Li et al., 2021).

Architecture of Fourier Neural Operator

Given points of input functions a_j (coefficients) and u_j (solutions)

1. Projective layer P sends data $a_j \to \nu_t$
2. Multiple Fourier Layers $\sigma(W\nu_t x + \mathcal{F}^{-1}(\mathcal{F}(G_\theta \cdot \nu_t)))$ sends $\nu_t \to \nu_{t+1}$
   - Apply Fast Fourier Transform $\mathcal{F}$
   - Performs convolution in Fourier Space with G_θ
   - Weights and biases W, b are applied along with activation function σ
   - Inverse Fast Fourier transform $\mathcal{F}^{-1}$
3. Projective Layer Q sends ν_T to u_j

Computational Result: Rugged Mueller Potential

- A Fourier Neural Operator was trained on the committor for the Rugged Mueller's potential which includes a periodic term
- The 3 parameters are β, γ (amplitude of noise), k (periodicity of noise) which introduces computational difficulty in training
- 5500 epoch model test set error: MAE: 2.6e-5, weighted MAE: 4.9e-3

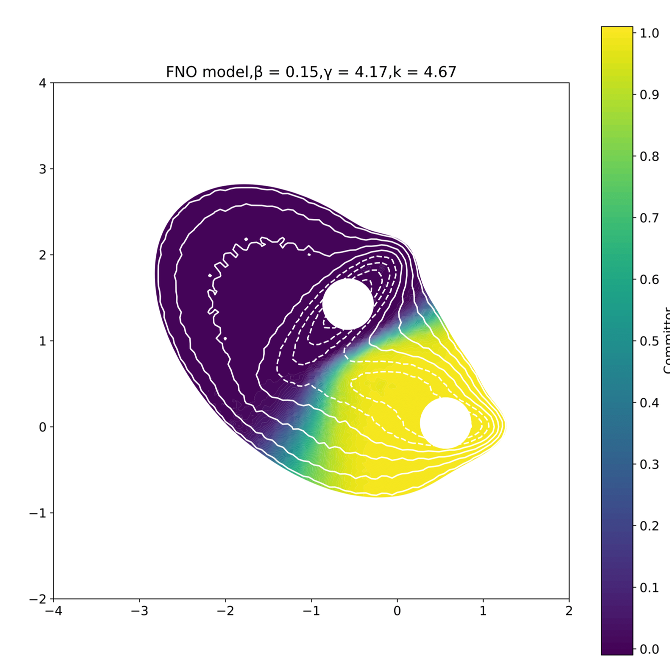

Figure 2. FNO evaluated Committor

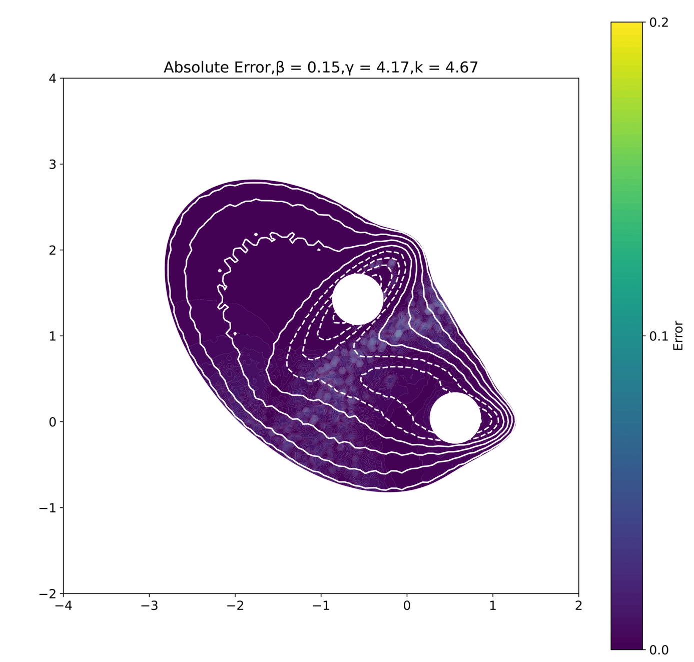

Figure 3. Error from FNO evaluation

Approach 2: Rational Activation Function

Main Idea: Choose $\sigma(x) = \frac{P(x)}{Q(x)} = \frac{\sum_{i=0}^{r_P} a_i x_i}{\sum_{j=0}^{r_Q} b_j x_j}$ $(r_P > r_Q, r_P \approx r_Q + 1)$

Benefits:
- ◦ Gradient does not tend to 0 as $x \to \pm\infty$
- ◦ Gradient is non-zero for negative inputs (Boulle et al., 2020)

Preliminary Results for Mueller Potential

	wMAE			wMRSE		
	Tanh	ReLU	Rational	Tanh	ReLU	Rational
3 L, 200 epochs	8.29e-3	5.65e-3	2.94e-3	2.04e-2	1.54e-2	5.92e-3
3 L, 1000 epochs	4.85e-4	9.51e-3	NaN*	6.86e-4	2.44e-2	NaN*
2 L, 1000 epochs (Yuan et al.)	2.6e-3	N/A	N/A	4.1e-3	N/A	N/A

*Neural network parameters became NaN while training - addressing this issue has proven challenging

Approximation Theory of Rationals

Definition: An ε-approximation of $f : \mathbb{R}^d \to \mathbb{R}$ over $[-1,1]^d$ is a function, $\tilde{f}$, such that $\|f - \tilde{f}\|_\infty \leq \varepsilon$

Novel Results: The size of an ε-approximation by a rational neural network of a tanh neural network is bounded above by $O(\log(\log(\frac{1}{\varepsilon})))$

*Result relies partially on a numeric step

Approach 3: Variational Physics Informed Neural Network (VPINNs)

- Main Idea: q satisfies $a(q, v) = \int_\Omega e^{-\beta V}\nabla q \cdot \nabla v = 0$ for all $v \in V$ such that $V = \{v : v|_{\partial A \cup \partial B} = 0\}$ (Berrone et al., 2021)
- $\mathcal{T}$ is a triangulation of Ω and $\{\phi_i\} \subseteq V$ are the basis functions for $\mathcal{T}$
- Let $\overline{u}$ be a solution to the boundary conditions and define $B : H^1(\Omega) \to \overline{u} + V$ as $Bw = \overline{u} + \Phi w$, where Φ maps w into V
- Train the neural network by minimizing $\mathcal{R}(w) = \sum_{\phi_i} a(Bw, \phi_i)^2$
- Current computational results have been inconclusive

Conclusion

- Neural operators can achieve reasonable accuracy for committors
- Rational neural networks can provide more accurate results than traditional activation functions, but can be harder to train

billchen2025@u.northwestern.edu

Acknowledgement: NSF REU grant DMS-2149913

ab2453@cornell.edu