A Deep Learning Method for Solving Fokker-Planck Equations

Yao Li, Matthew Dobson and **Jiayu Zhai**
University of Massachusetts Amherst

Mathematical and Scientific Machine Learning
August 19, 2021

Invariant Distribution for SDE systems

The invariant distribution of SDE

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$$

satisfies the stationary Fokker-Planck equation

$$\begin{cases} \mathcal{L}u := -\sum_{i=1}^{n}(b_i u)_{x_i} + \frac{1}{2}\sum_{i,j=1}^{n}(D_{i,j}u)_{x_i x_j} = 0 \\ \int_{\mathbb{R}^n} u\,d\boldsymbol{x} = 1 \end{cases} \tag{1}$$

Its solution is learnt with the loss functional

$$J(u) = \|\mathcal{L}u\|^2_{L^2(D)} + \|u - v\|^2_{L^2(D)}, \tag{2}$$

A Deep Learning Method (Convergence)

Let

$$\boldsymbol{e} = \boldsymbol{v} - \boldsymbol{u}^* \text{ (error from data)},$$

$$\boldsymbol{z} = \bar{\boldsymbol{u}} - \boldsymbol{u}^* \text{ (error of the method)},$$

where $\bar{\boldsymbol{u}}$ is the minimizer of the discretization of $J(u)$.

Theorem

Under suitable assumptions about $\boldsymbol{e}$ and $\boldsymbol{A}$, we have

$$\lim_{h \to 0} \frac{\mathbb{E}[\|\boldsymbol{z}\|^2]}{\mathbb{E}[\|\boldsymbol{e}\|^2]} = 0.$$

Effectiveness of $\mathcal{L}u$ in the Loss Function (Ring Density)

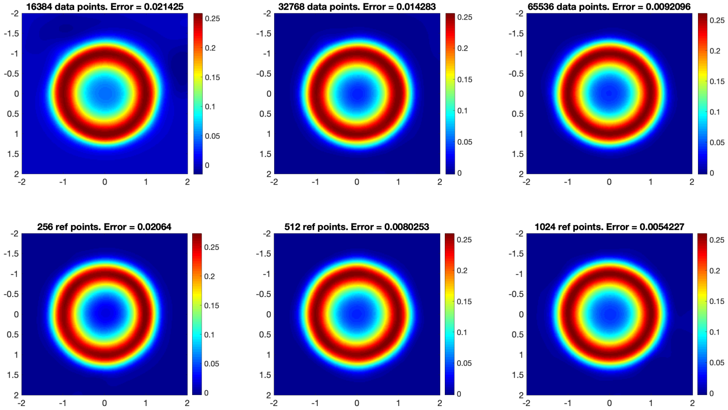
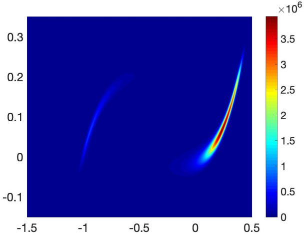


Figure: Comparison between the solutions learnt from: (1) 16384, 32768, 65536 reference data without $\mathcal{L}u$ (first row); (2) 256, 512, 1024 reference data with $\mathcal{L}u$ (second row)

Sampling Collocation Points



Due to the reduced demand on data, we need to choose their collocations carefully. We propose the following sampling strategy.

1. For the concentrated parts:
   - run a trajectory again and select $\alpha\%$ collocation points on the trajectory.
2. For the complement part:
   - $1-\alpha\%$ from uniform distribution on D.

α is set up to be $50-90$

Effectiveness in Correcting Spatial Noise (Gibbs Measure)

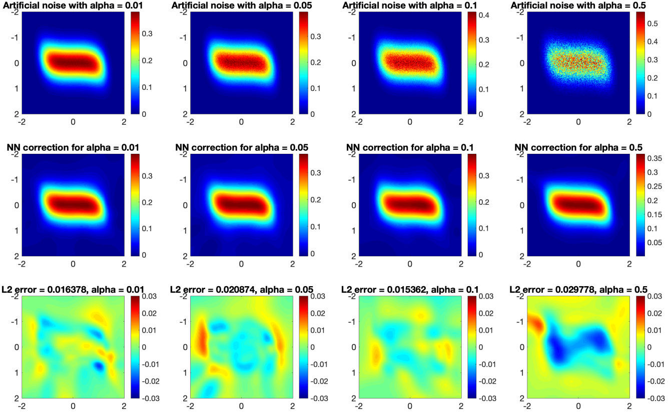


Figure: The ability of neural network representations in rectifying different levels of noise strengths. $v(\mathbf{y}) = r(\mathbf{y})u(\mathbf{y})$, where $r \sim U([1-\alpha, 1+\alpha])$ with $\alpha = 0.01, 0.05, 0.1, 0.5$