Neural Control Variates

Srijit Paul, Hyunwoo Oh and Paulo Bedaque

College Park
2024-03-25

2024

March

March 25, 2024

Following the idea proposed in 2312.08228.

U(1) Model Application

- Conflict between the U(1) model plaquette picture and the link picture.

Problem: The Action in the U(1) model is a sum of cosine of the angle of the plaquettes, but in the link picture it is the cosine of the sum of the angles of the links.

$$S^{2D}_{\text{plaquette}} = \beta(L^2 - \sum_{\text{j}=1}^{L}\sum_{\text{i}=1}^{L} \cos(\theta_{\text{i,j, plaquette}})) \tag{1}$$

$$S^{2D}_{\text{link}} = \beta(L^2 - \sum_{\text{j}=1}^{L}\sum_{\text{i}=1}^{L} \cos(\theta^{\mu=0}_{\text{i,j, link}} + \theta^{\mu=1}_{\text{i+1,j, link}} - \theta^{\mu=0}_{\text{i,j+1, link}} - \theta^{\mu=1}_{\text{i,j, link}})) \tag{2}$$

Can we show that the plaquette picture is equivalent to the link picture in the Haar measure when we compute the partition function?

Solution: The Haar measure is invariant under the transformation of the link variables. If we choose a gauge where all link variables pointing along the time direction are rotated to the unit element ($e^{i\theta} = 1 \implies \theta = 0$) which means in [**Temporal Gauge**]

$$\theta^{\mu=0}_{\text{i,j, link}} = 0 \;\; \forall \text{ i, j} \in [1, L]$$

Then the link action becomes

$$S^{2D}_{\text{link, temporal}} = \beta(L^2 - \sum_{\text{j}=1}^{L}\sum_{\text{i}=1}^{L} \cos(\theta^{\mu=1}_{\text{i,j, link}} - \theta^{\mu=1}_{\text{i+1,j, link}})) \tag{3}$$

Now we can see that the link action is equivalent to the plaquette action in the Haar measure. If we compute the partition function using the link action in the temporal gauge as

$$Z^{2D}_{\text{link}} = \int \prod_{\text{i}=1}^{L}\prod_{\text{j}=1}^{L} d\theta^{\mu=1}_{\text{i,j, link}} e^{-S^{2D}_{\text{link, temporal}}} \tag{4}$$

and we perform a change of variables such that $\theta^{\mu=1}_{\text{i,j, link}} - \theta^{\mu=1}_{\text{i+1,j, link}} = \theta'_{\text{i,j}}$ then we can see that the partition function is equivalent to the plaquette action in the Haar measure.

Control Variate

Using Stein Operator to construct a Neural Network to approximate a function $g(\theta^{\mu}_{\text{i,j}})$ such that

$$\nabla_{\theta^{\mu}_{\text{i,j}}} g(\theta^{\mu}_{\text{i,j}}) = g(\theta^{\mu}_{\text{i,j}}) \nabla_{\theta^{\mu}_{\text{i,j}}} S^{2D}_{\text{link}} \tag{5}$$

The Control Variate $f(\theta^{\mu}_{\text{i,j}})$ is then defined as

$$f(\theta^{\mu}_{\text{i,j}}) = \nabla_{\theta^{\mu}_{\text{i,j}}} g(\theta^{\mu}_{\text{i,j}}) - g(\theta^{\mu}_{\text{i,j}}) \nabla_{\theta^{\mu}_{\text{i,j}}} S^{2D}_{\text{link}} \tag{6}$$

By Schwinger-Dyson equation, we can show that the expectation value of the Control Variate in infinite statistics limit is zero.

$$\langle f(\theta^{\mu}_{\mathrm{i,j}})\rangle_{\infty} = 0 \tag{7}$$

Now if we define a gauge invariant observable O as

$$\langle O\rangle_F = \frac{1}{Z}\sum_{\mathrm{a}=1}^{N} O([\theta^{\mu}_{\mathrm{i,j}}]_a)e^{-S^{2D}_{\mathrm{link}}[\theta^{\mu}_{\mathrm{i,j}}]_a} \tag{8}$$

As $N \to \infty$

$$\langle O\rangle_F \to \langle O\rangle_{\infty} \tag{9}$$

such that,

$$\langle O\rangle_F - \sqrt{\langle O^2\rangle_F - \langle O\rangle_F^2} \leq \langle O\rangle_{\infty} \leq \langle O\rangle_F + \sqrt{\langle O^2\rangle_F - \langle O\rangle_F^2} \tag{10}$$

April

April 20, 2024

Neural Network

The Neural Network is constructed to approximate the function $g(\theta^{\mu}_{\mathrm{i,j}})$ in Eq. 6. The Neural Network is trained to minimize the naive loss function $\mathcal{L}$, which can be defined as,

$$\mathcal{L} = \frac{1}{N}\sum_{\mathrm{a}=1}^{N}\{O([\theta^{\mu}_{\mathrm{i,j}}]_a) - f([\theta^{\mu}_{\mathrm{i,j}}]_a)\}^2 \tag{11}$$

where N is the number of samples, O is the observable, and f is the Control Variate.

Single Plaquette Model in 2D

The Neural Network is trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. The Neural Network is trained for 100 epochs.