

Neural Control Variates

Srijit Paul, Hyunwoo Oh and Paulo Bedaque

College Park
2024-03-25

2024

March

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.

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

$$S_{\text{link}}^{2D} = \beta(L^2 - \sum_{j=1}^L \sum_{i=1}^L \cos(\theta_{i,j, \text{ link}}^{\mu=0} + \theta_{i+1,j, \text{ link}}^{\mu=1} - \theta_{i,j+1, \text{ link}}^{\mu=0} - \theta_{i,j, \text{ link}}^{\mu=1})) \quad (2)$$

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 which means [**Temporal Gauge**]

$$\theta_{i,j, \text{ link}}^{\mu=0} = 0 \quad \forall i, j \in [1, L]$$

Then the link action becomes

$$S_{\text{link, temporal}}^{2D} = \beta(L^2 - \sum_{j=1}^L \sum_{i=1}^L \cos(\theta_{i,j, \text{ link}}^{\mu=1} - \theta_{i+1,j, \text{ link}}^{\mu=1})) \quad (3)$$

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

$$Z_{\text{link}}^{2D} = \int \prod_{i=1}^L \prod_{j=1}^L d\theta_{i,j, \text{ link}}^{\mu=1} e^{-S_{\text{link, temporal}}^{2D}} \quad (4)$$

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

April

April 1, 2024

Control Variate

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

$$\nabla_{\theta_{i,j}^\mu} g(\theta_{i,j}^\mu) = g(\theta_{i,j}^\mu) \nabla_{\theta_{i,j}^\mu} S_{\text{link}}^{2D} \quad (5)$$

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

$$f(\theta_{i,j}^\mu) = \nabla_{\theta_{i,j}^\mu} g(\theta_{i,j}^\mu) - g(\theta_{i,j}^\mu) \nabla_{\theta_{i,j}^\mu} S_{\text{link}}^{2D} \quad (6)$$

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

$$\langle f(\theta_{i,j}^\mu) \rangle_\infty = 0 \quad (7)$$

Now if we define a gauge invariant observable O as

$$\langle O \rangle_F = \frac{1}{Z} \sum_{a=1}^N O([\theta_{i,j}^\mu]_a) e^{-S_{\text{link}}^{2D}[\theta_{i,j}^\mu]_a} \quad (8)$$

As $N \rightarrow \infty$

$$\langle O \rangle_F \rightarrow \langle O \rangle_\infty \quad (9)$$

such that,

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

Neural Network

The Neural Network is constructed to approximate the function $g(\theta_{i,j}^\mu)$ in Eq. 5. The Neural Network is trained to minimize the loss function \mathcal{L} , which can be defined as,

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