Improvement of Heatbath Algorithm in LFT using Generative model

41st Lattice Conference, Liverpool, UK

Ankur Singha

ML group, BIFOLD TU Berlin

Collaborators: Ali Faraz, Prof. Dipankar Chakraborti, Prof. Vipul Arora, Dr. Shinichi Nakajima.

Introduction

Generative Models for Sampling Lattices

- **Albergo et al. (2019)**: Normalizing flows for Phi4 theory.
- **Nicoli et al. (2020, 2021)**: Used NF for thermodynamic observables and mode collapse in Phi4 theory.
- **Kanwar et al. (2020)**, **Boyda et al. (2020)**: Applied forward KL optimization of NF for U(1), SU(N) gauge theories.
- **Singha et al. (2022, 2023)** : Conditional Normalizing flows for LFT.
- **Gerdes et al. (2022, 2023)**: Continuous Normalizing flows for LFT.
- **Caselle et al. (2022)**: Stochastic Normalizing flows for LFT.
- **Wang et al. (2023)**: Diffusion model for LFT.
- **Abbott et al. (2023, 2024)**: Sampled QCD fields with gauge-equivariant flow models.

Exploring the Heat-Bath Algorithm

- **Motivation**: Generative models are promising but complex. Heat-Bath offers a simpler, exact method.
- **Objective**: Investigate Heat-Bath algorithm's performance, especially construct a good proposal.

Introduction: Gibbs Sampling & Heat-Bath

Objective: A Markov Chain Monte Carlo (MCMC) algorithm used to sample from a joint probability distribution $P(x_1, x_2, \ldots, x_n)$.

Steps in Gibbs Sampling:

1. **Initialization:** Start with an initial state $(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$.
2. **Iterative Sampling:**

$$
\begin{aligned}
x_1^{(t+1)} &\sim P(x_1 \mid x_2^{(t)}, x_3^{(t)}, \ldots, x_n^{(t)}) \\
x_2^{(t+1)} &\sim P(x_2 \mid x_1^{(t+1)}, x_3^{(t)}, \ldots, x_n^{(t)}) \\
&\vdots \\
x_n^{(t+1)} &\sim P(x_n \mid x_1^{(t+1)}, x_2^{(t+1)}, \ldots, x_{n-1}^{(t+1)})
\end{aligned}
$$

3. This is a sequential single site update from the conditional distribution.

Gibbs sampling & Heat-Bath

In lattice theories the number of conditions can be few only, like local theories,

$$P(x_1 \mid x_2^{(t)}, x_3^{(t)}, \dots, x_n^{(t)}) = P(x_1 \mid x_2^{(t)}, x_3^{(t)}, x_4^{(t)}, x_5^{(t)})$$

- If the conditional distribution can be factorise further, then parallel sampling is possible: Heatbath.
- This makes Heat-bath as faster sampling method.
- Sampling is done at two steps:

$$X_{odd}^{(t)} \sim P(X_{odd} | X_{even})$$
$$X_{even}^{(t+1)} \sim P(X_{even}^{t} | X_{odd})$$

Heat-Bath: Discrete Case

- **Ising**: Probabilities for state $s_i = +1$ and $s_i = -1$,

$$P(s_i = +1) = \frac{e^{\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}}, \quad P(s_i = -1) = \frac{e^{-\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}}$$

  - Draw a random number u from a uniform distribution $U(0, 1)$.
  - Set $s_i = +1$ if $u < P(s_i = +1)$, otherwise set $s_i = -1$.
- **q state spins**:Probabilities

$$P(s_i = k) = \frac{e^{\beta J n_k}}{\sum_{l=1}^{q} e^{\beta J n_l}}$$

  - Sample s_i from this distribution:
  - Draw a random number u from a uniform distribution $U(0, 1)$.
  - Use cumulative probabilities to determine the state k:

$$k = \text{smallest integer such that} \sum_{m=1}^{k} P(s_i = m) > u$$

Heatbath: Continuous Case

Probability Distribution:

$$P(\varphi_x) \propto \exp(-S(\varphi_x))$$

Cumulative Distribution Function:

$$F(\varphi_x) = \int_{-\infty}^{\varphi_x} P(\varphi')\, d\varphi'$$

Numerical Integration:

Evaluating the CDF $F(\varphi_x)$ non-trivial, often requires numerical methods. Thus need looks alternate methods such as rejection sampling.

Heat-Bath: Rejection Sampling

Action for Scalar LFT:

- The local action is given by:

$$S_{loc}(\varphi_{i,j}, \lambda, m^2, \kappa_{i,j}) = \left(m^2 + 4\right)\varphi_{i,j}^2 + \lambda\varphi_{i,j}^4 - 2\varphi_{i,j}\kappa_{i,j}$$

- Where:

$$\kappa_{i,j} = \varphi_{i+1,j} + \varphi_{i,j+1} + \varphi_{i-1,j} + \varphi_{i,j-1}$$

- λ and m are theory parameters.

Target Distribution:

- The normalized target distribution is:

$$p(\varphi_{i,j}) = \frac{1}{Z}\exp\left(-S_{loc}(\varphi_{i,j}, \lambda, m^2, \kappa_{i,j})\right)$$

- This represents the distribution we want to sample from.

Edwards, R. G., et al. Nuclear Physics B 10.1016/0550-3213(91)90357-4

Heat-Bath Algorithm: Rejection Sampling

Proposal Distribution:

- For Gaussian proposal distribution:

$$g(\varphi) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\varphi-\mu)^2}{2\sigma^2}\right)$$

where,

$$\mu = \frac{\kappa_{i,j}}{m^2+4}, \quad \sigma^2 = \frac{1}{2(m^2+4)}$$

Heat-Bath Algorithm:

1. For each site (i,j):
   - Calculate $\mu = \frac{\kappa_{i,j}}{m^2+4}$ and $\sigma^2 = \frac{1}{2(m^2+4)}$ from the Gaussian approximation.
   - Draw a candidate $\varphi_{i,j}$ from the Gaussian distribution.
   - Accept the candidate with probability:

$$\exp\left(-\lambda\varphi^4\right)$$

   - If rejected, draw a new candidate and repeat until accepted.

Challenges with Continuous Degrees of Freedom

Challenges in Heatbath with rejection sampling

- The primary challenge with this rejection method is finding an effective proposal distribution.
- A poor proposal distribution can result in a high rejection rate, which increases the overall simulation cost.
- The proposal distribution needs to be adjusted for different regions of the action parameters.
- Occasionally, the same type of proposal distribution may not be effective for simulations in parameter regimes where the target distribution has multiple modes.

PBMG Method: Phi4

Target: The conditional distribution of the lattice site (i, j) can be written as

$$p\big(\varphi_{i,j}|\lambda, m^2, \kappa_{i,j}\big) = p\big(\varphi_{i,j}|\mathbf{v}_{i,j}\big) \propto e^{-S_{loc}(\varphi_{i,j}, \mathbf{v}_{i,j})} \tag{1}$$

where, $\mathbf{v}_{i,j} = (\lambda, m^2, \kappa_{i,j})$ is the condition vector for the distribution. Eq. (1) is the target distribution which we model using generative approach.

Proposal: We model the proposal distribution using GMM with K Gaussian components.

$$q(\varphi_{i,j}|\mathbf{v}_{i,j};\theta) = \sum_{k=1}^{K} \pi_k(\mathbf{v}_{i,j};\theta_k)\mathcal{N}(\varphi_{i,j}|\mu_k(\mathbf{v}_{i,j};\theta_k), \sigma_k(\mathbf{v}_{i,j};\theta_k))$$

where, μ_k, σ_k, π_k are the mean, standard deviation and mixing coefficient of the k^{th} Gaussian distribution, which are estimated using NNs with parameters $\theta = \{\theta_k\}_{k=1}^{K}$.

Training: PBMG-Phi4

- **Training data:**
  We generate samples of $\mathbf{v}_{i,j}$ from $p_v(\mathbf{v}_{i,j})$ as the training set. Here, $p_v(\mathbf{v}_{i,j}) = \text{Unif}([2.5, 15] \times [-8, 0] \times [-20, 20])$ as $\lambda \in [2.5, 15]$, $m^2 \in [-8, 0]$ and $\kappa_{i,j} \in [-20, 20]$ are the parameter ranges chosen for training.
- **Loss Function** The training is done by minimizing the KL divergence between the proposal and the target :

$$\mathcal{L} \approx \frac{1}{n} \sum_{r=1}^{n} \left[\frac{1}{N} \sum_{k=1}^{N} [\log q((\varphi_{i,j})_k | (\mathbf{v}_{i,j})_r;) - \log p\big((\varphi_{i,j})_k | (\mathbf{v}_{i,j})_r)] + \|\pi((\mathbf{v}_{i,j})_r; \theta\| \right] \tag{2}$$

PBMG Method: XY Model

The local Hamiltonian for $\varphi_{i,j}$ is

$$H(\varphi_{i,j}) = -\Big[\cos(\varphi_{i,j} - \varphi_{i+1,j}) + \cos(\varphi_{i,j} - \varphi_{i,j+1}) + \cos(\varphi_{i,j} - \varphi_{i-1,j}) + \cos(\varphi_{i,j} - \varphi_{i,j-1})\Big] \quad (3)$$

$$p(\varphi_{i,j}|\mathbf{v}_{i,j}) \propto e^{-\frac{H(\varphi_{i,j})}{T}}. \quad (4)$$

Here, $\mathbf{v}_{i,j} = (\varphi_{i+1,j}, \varphi_{i,j+1}, \varphi_{i-1,j}, \varphi_{i,j-1}, T)$.

Training: PBMG-XY

We use NF to model the proposal distribution $q(\varphi_{i,j}|\mathbf{v}_{i,j};\theta)$. $p_Z(z|\mathbf{v}_{i,j};\theta_B)$ is the base distribution and $f(z;\theta_R)$ is the invertible transformation used in the NF. Here, $\theta = \{\theta_\mathbf{B}, \theta_\mathbf{R}\}$.
Using the change of variables formula,

$$q(\varphi_{i,j}|\mathbf{v}_{i,j};\theta) = p_Z\left(f^{-1}(\varphi_{i,j};\theta_R)|\mathbf{v}_{i,j};\theta_B\right) \left|\det\left(\frac{\partial f^{-1}(\varphi_{i,j};\theta_R)}{\partial \varphi_{i,j}}\right)\right| \quad (5)$$

Here, we use Rational Quadratic Splines (RQS) as the transform f.
The parameters of the RQS flow are learned through neural networks.

Loss Function: PBMG-XY

The base distribution of RQS flow is chosen to be uniform, over a learnable interval w_{int}.

We use four different NNs ($NN_i,\ i = 1, 2, 3, 4$) to learn these parameters.

The Loss function:

$$\mathcal{L} \approx \frac{1}{n} \cdot \frac{1}{N} \sum_{r=1}^{n} \sum_{k=1}^{N} [\log p_Z(z_k|(\mathbf{v}_{i,j})_r; \theta_B) + \\ \log|\det J_f(z_k|(\mathbf{v}_{i,j})_r; \theta_R)|^{-1} - \log p\left(f(z_k; \theta_R)|(\mathbf{v}_{i,j})_r\right)].$$

Here, $p_v(\mathbf{v}_{i,j}) = \text{Unif}([0, 2\pi]^4 \times [0.13, 2.05])$ as the first four components of $\mathbf{v}_{i,j}$ belong to $[0, 2\pi]$ and $T \in [0.13, 2.05]$ is the training range of T.

Results: PBMG vs Heatbath for Phi4 Theory

The trained models are used as proposal for Heatbath sampling.

Acceptance rate: The acceptance rate R on lattice with lattice sites N_{tot} can be calculated as:

$$R = \frac{N_{tot}}{\text{Total attempts}} = \frac{N_{tot}}{\sum_{i=1}^{N} n_i^{(k)}}$$

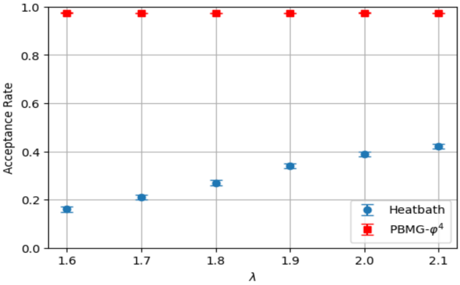

Figure: Acceptance rate of Heatbath vs PBMG for different action parameter values.

Results: PBMG vs Heatbath XY model

We compare PBMG XY model against Heatbath with uniform distribution as proposal:

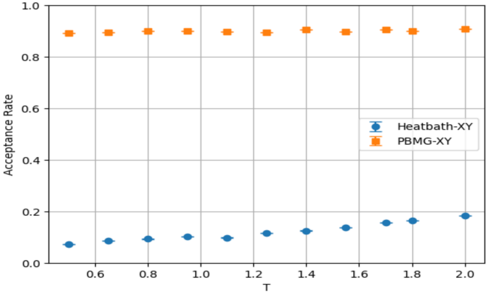

Figure: Acceptance rate of Heatbath vs PBMG for different Temperatures.

For both Phi4 and XY model we consider same training time for the acceptance rate. For XY it is around 90% which further improves with training time.

Results: Observables

The observables such as energy and magnetization are compared using Earth Mover Distance (EMD).

Metric	Lattice size	PBMG-ϕ^4
EMD $\downarrow$	8x8	0.0008
	16x16	0.0030
	32x32	0.0085
	64x64	0.0081

Metric	Lattice size	PBMG-XY
EMD $\downarrow$	8x8	0.0042 ± 0.0036
	16x16	0.0027 ± 0.0026
	32x32	0.0027 ± 0.0034
	64x64	0.0013 ± 0.0014

Summary and Future Outlook:

- We propose a method to train a conditional generative to generate proposal for Heatbath.
- Its a conditional model both on lattice site (not restricted to local sites) and action parameters.
- The model has high acceptance rate and can be used across the wide range of action parameter.
- **Multiscale sampling**: A single model can be utilized for all levels of multi-scale sampling where effective actions need different proposals for Heatbath sampling.
- **Guage Thoery**: We are exploring sampling of Gauge theory using PBMG also theories which cannot be sampled by usual Heatbath.

Thank You!