Exploring Generative Networks for Manifolds with Non-Trivial Topology

Shiyang Chen

Biagio Lucini, Gert Aarts

Swansea University
Prifysgol Abertawe

Content

- Generative Networks
- Topology Problem
- Our network
- 2D-ϕ^4 theory

Generative Networks

- Estimate Observables via annealed importance sampling,

$$\mathbb{E}_{\phi\sim\mu(\phi)}[\mathcal{O}(\phi)] = \int d\phi \frac{\mu(\phi)}{\mathcal{Z}} \mathcal{O}(\phi) \approx \frac{\sum_{\phi\sim Q(\phi)} \omega(\phi)\mathcal{O}(\phi)}{\sum_{\phi\sim Q(\phi)} \omega(\phi)}, \quad \omega(\phi) = \frac{\mu(\phi)}{Q(\phi)}.$$

$\mu(\phi)$: un-normalized target distribution.

$Q(\phi)$: normalized model distribution.

$\mathcal{Z}$: normalization constant

Precondition

1. Support set $supp\{\mu\} \subset supp\{Q\}$;
2. Distribution Q not far away from μ.

Support set $supp\{\mu\} = \{\forall\phi | \mu \neq 0\}$

Q could be estimated by neural networks!

Generative Networks

- Estimate Observables via annealed importance sampling,

$$\mathbb{E}_{\phi\sim\mu(\phi)}[\mathcal{O}(\phi)] = \int d\phi \frac{\mu(\phi)}{\mathcal{Z}} \mathcal{O}(\phi) \approx \frac{\sum_{\phi\sim Q(\phi)} \omega(\phi)\mathcal{O}(\phi)}{\sum_{\phi\sim Q(\phi)} \omega(\phi)}, \quad \omega(\phi) = \frac{\mu(\phi)}{Q(\phi)}.$$

Available neural network:

Train on Function

Normalizing flow (NF),
[Kim A. Nicoli, et al., arXiv:2007.07115 [hep-lat]]
[Ryan Abbott, et al., arXiv:2207.08945 [hep-lat]]

Continuous flow,
[Mathis Gerdes, et al., arXiv:2207.00283 [hep-lat]]

Hamiltonian neural network,
[Sam Greydanus, et al., arXiv:1906.01563 [cs.NE]]

Stochastic NF
[Michele Caselle. et al., arXiv:2210.03139 [hep-lat]]

Train on Data

Langevin based diffusion model,
[Lingxiao Wang,et al., arXiv:2309.17082 [hep-lat]]

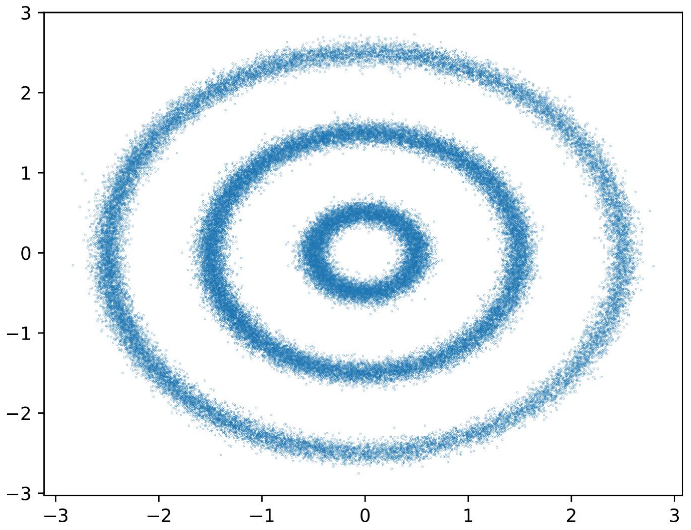
A-Nice-MC
[Jiaming Song, et al., arXiv:1706.07561 [stat.M]]

NF using adversarial learning
[Vikas Kanaujia, et al., arXiv:2401.15948 [cs.LG]]

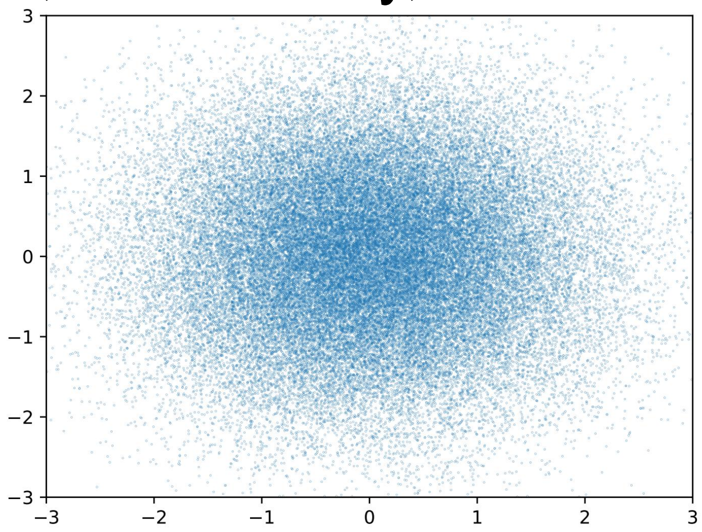
Topology Problem

Topology: Geometry feature preserved under continuous maps.

Connectivity, Compactness, Countability, etc.



Triple Ring



Normal Distribution

Triple Ring: a disconnected and compact Manifold.

Normal Distribution: a simple connected and non-compact Manifold.

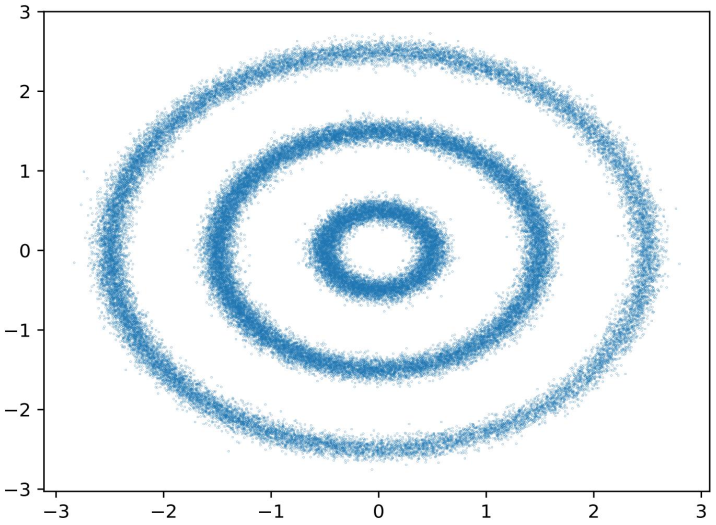
Topology Problem

- Invertible Flows are topology preserved map

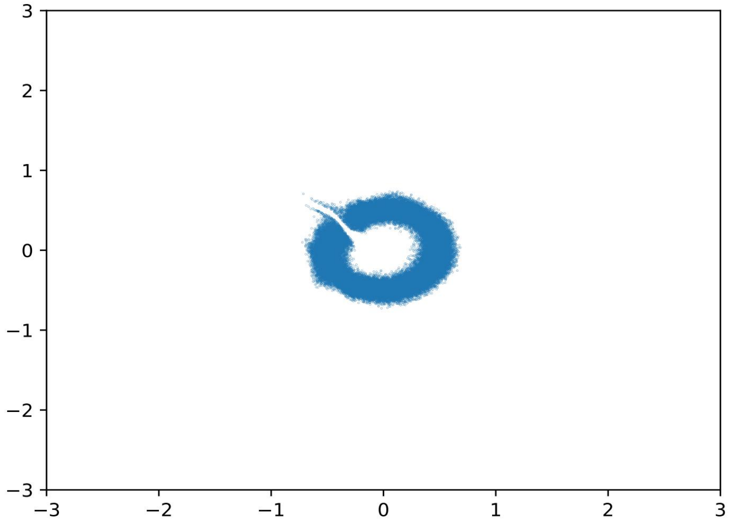
$$Q(\phi) = P(z)\left|\frac{\partial \phi}{\partial z}\right|^{-1}$$

$P(z)$: the prior and Normal distribution

Flow models are differentible, continuous and invertible. → Modeled manifold is diffeomorphism to the prior manifold



Target:Triple Ring



Model: NF

Modeled manifold:

1. open ring
2. simple connected and non-compact
3. outer rings missing

Topology Problem

- Avoid dependence of the choice of hyper-parameters.

    1. Ideal case to the worst output from the network;
    2. Improve the architecture.

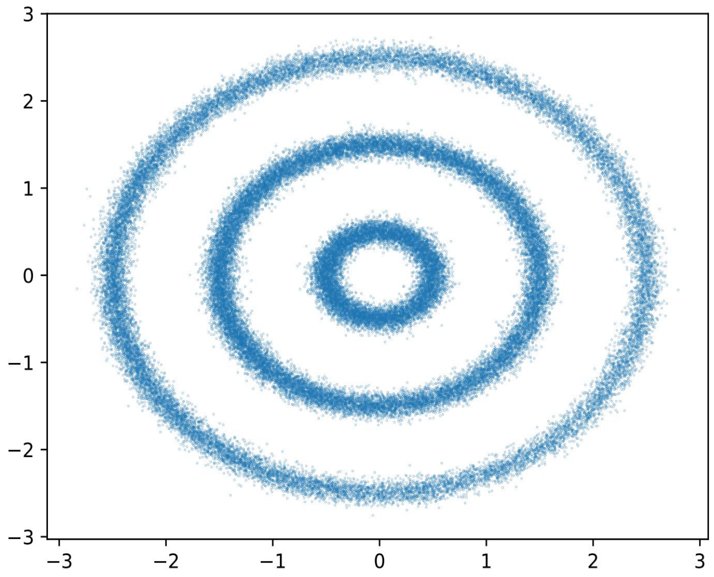
- The diffusion network learns from the action.

    1. Drift and diffusion for forward and backward are estimated by two independent networks.
    2. Results from diffusion will be different from one learning from data.

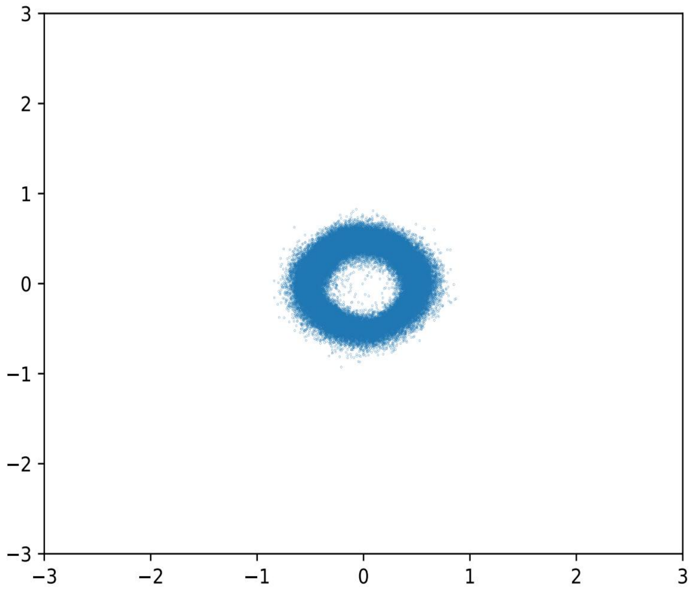
Topology Problem

- Langevin Based Diffusion may suffer from topology problem

carefully chosen hype-parameters and
prior will lead to a better result.



Target:Triple Ring



Model: Diffusion

Modeled manifold:

1. closed ring
2. connected and compact
3. outer rings missing
4. Network trapped at local minimum

Topology Problem

- The expected worst outcome for ideal case (understand from topology side)

$$z_{t+1} = z_t - K_t(z_t)dt + \sqrt{2dt}D_t(z_t)\eta_t, \qquad \eta_t \sim \text{Gauss}$$

1. Ideal case: $K_t(z_t)$ and $D_t(z_t)$ are continuous maps
2. Continuous map preserves topology
3. For two connected vector space, X and Y ,

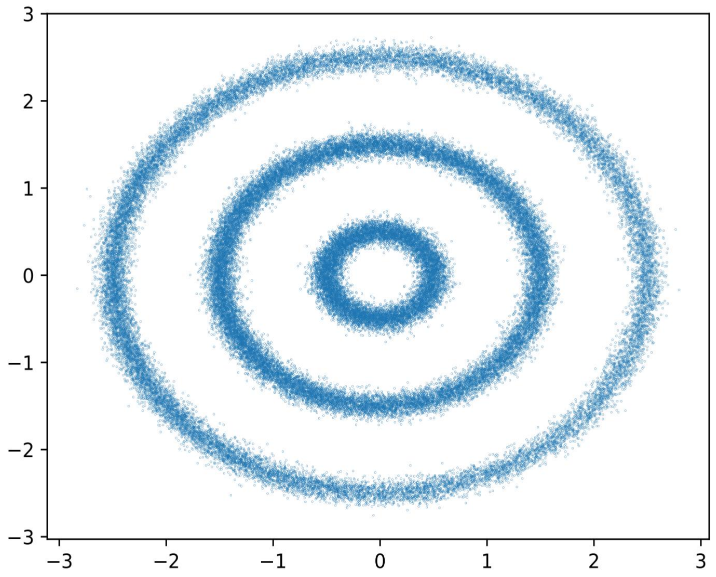
   The vector space $Z_1 = \{x + y | \forall x \in X, \forall y \in Y\}$ and $Z_2 = \{xy | \forall x \in X, \forall y \in Y,\}$ are connected vector space.

- The modelled manifold is connected.

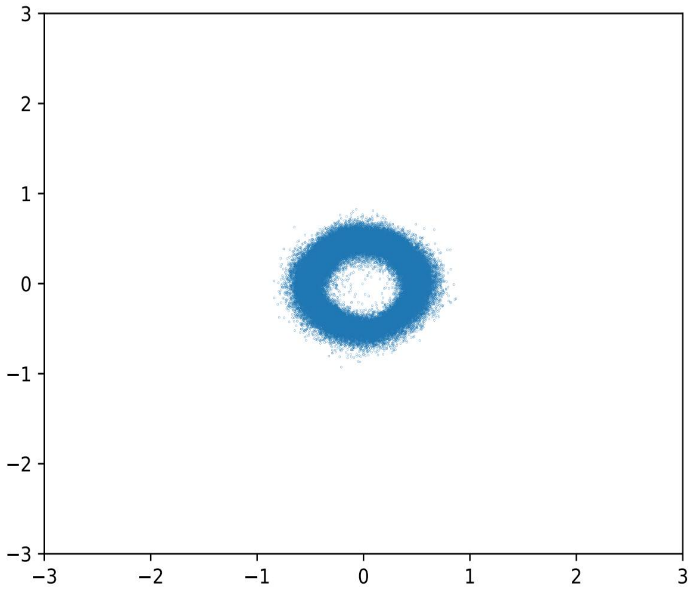
Topology Problem

- Langevin Based Diffusion may suffer from topology problem

carefully chosen hype-parameters and
prior will lead to a better result.



Target:Triple Ring



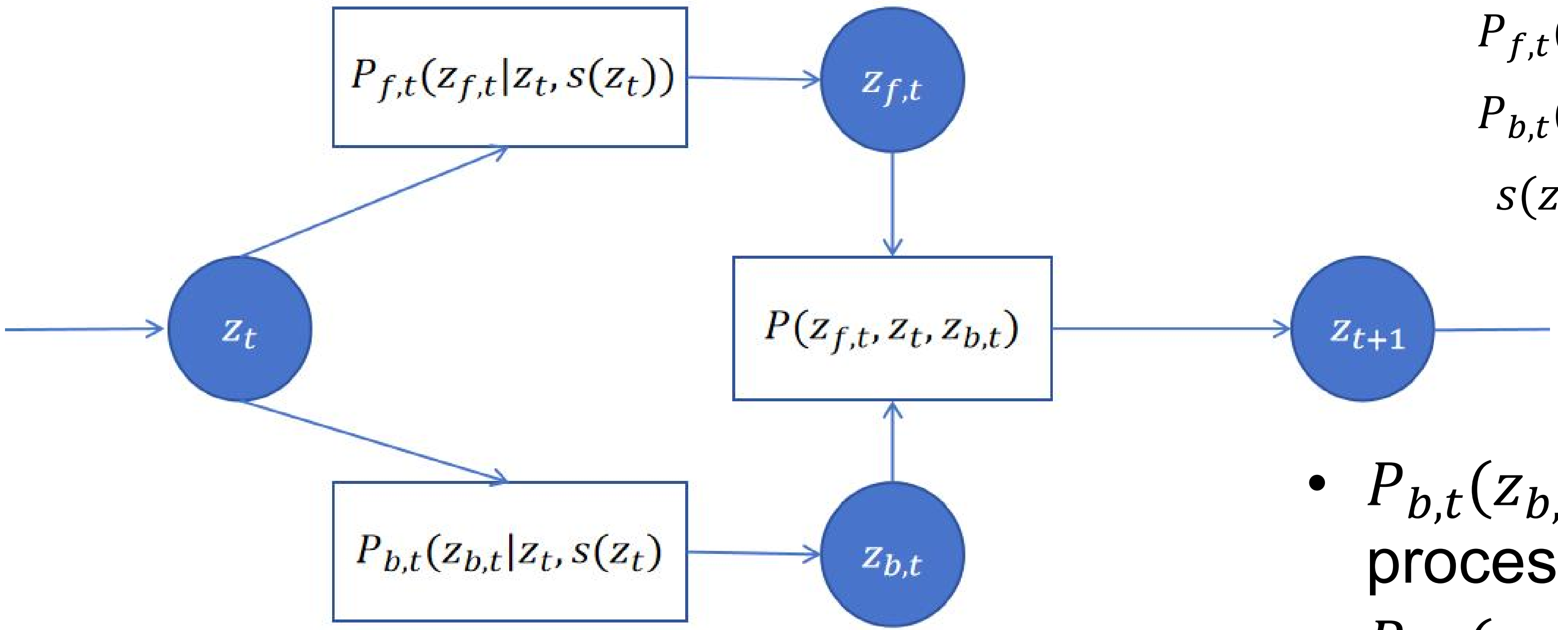
Model: Diffusion

Tips:

1. Action/energy functions suggested to be an augment,

2. ReLu and Leaky ReLu encouraged to use.

Our network



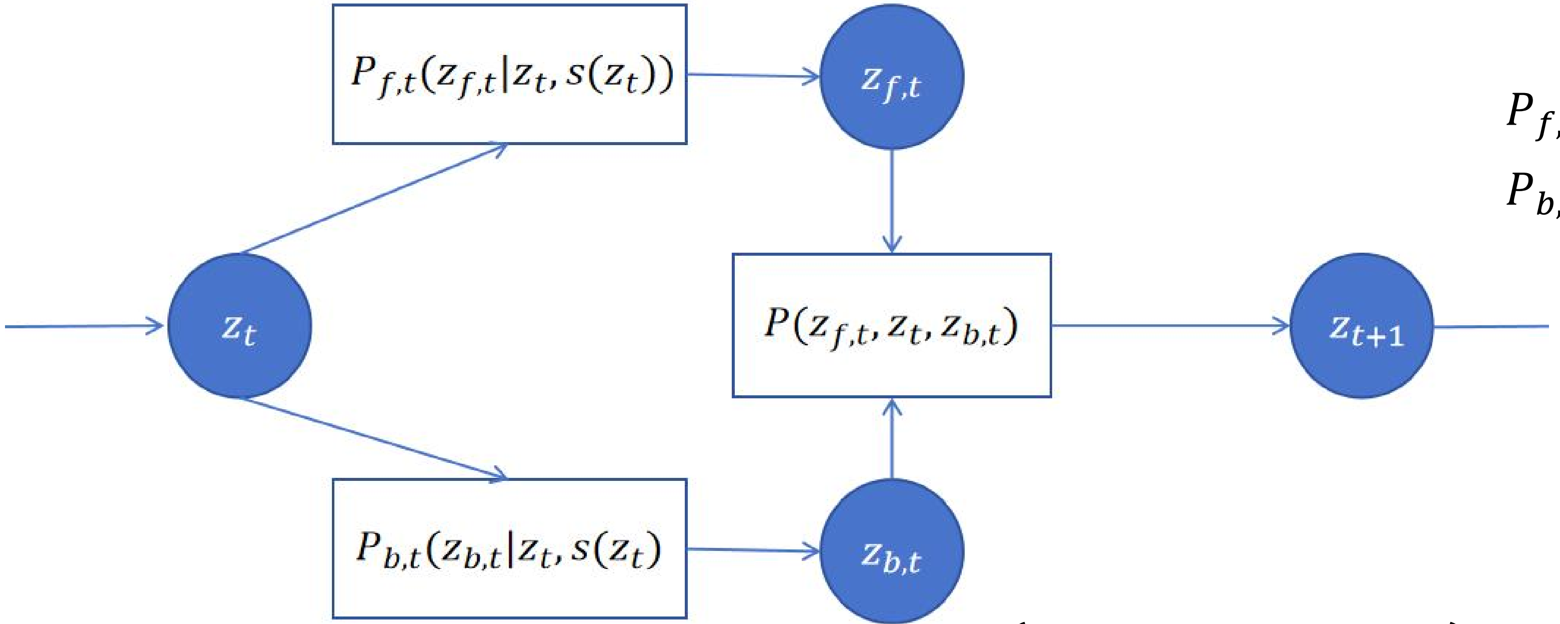
$P_{f,t}(z_{f,t}|z_t, s(z_t))$: forward process

$P_{b,t}(z_{b,t}|z_t, s(z_t))$: backward process

$s(z_t)$: action

- $P_{b,t}(z_{b,t}|z_t, s(z_t))$ is the inverse process of $P_{f,t}(z_{f,t}|z_t, s(z_t))$;
- $P_{f,t}(z_{f,t}|z_t, s(z_t)), P_{b,t}(z_{b,t}|z_t, s(z_t))$ and $P(z_{f,t}, z_t, z_{b,t})$ estimated by networks;
- $P_{f,t}(z_{f,t}|z_t, s(z_t)), P_{b,t}(z_{b,t}|z_t, s(z_t))$ have time dependence;
- Possible way to generate multiple and well seperated models.

Our network



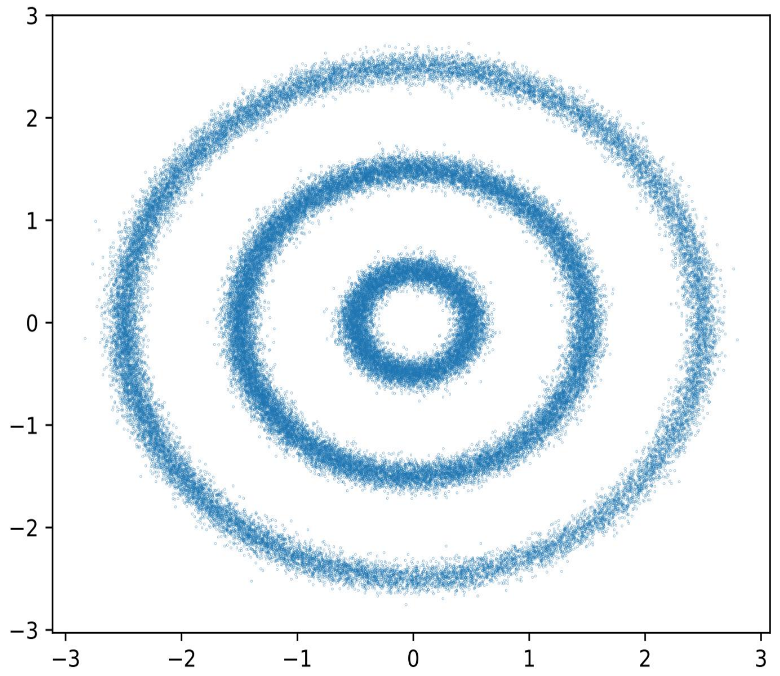
$P_{f,t}(z_{f,t}|z_t, S_t)$: forward process

$P_{b,t}(z_{b,t}|z_t, S_t)$: backward process

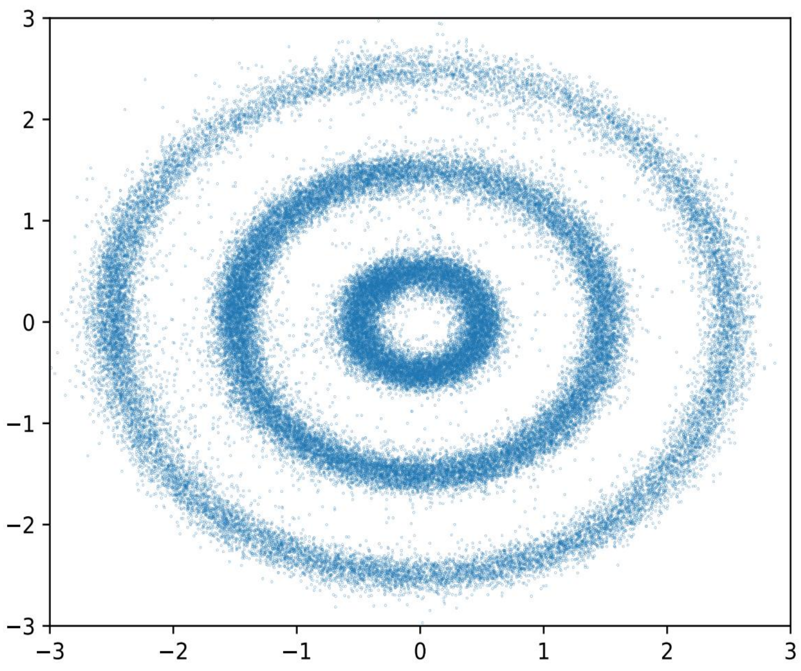
$$P(z_{t+1}) = \delta(z_{t+1} - z_{f,t}) \frac{P\left(z_{t+1}|z_{f,t}, z_t, z_{b,t}\right) P_{f,t}\left(z_{f,t}|z_{t,}, S_t\right)}{P\left(z_{t+1}|z_{b,t+1}, z_{t+1}, z_{f,t+1}\right) P_{b,t+1}\left(z_t|z_{f,t}, S_t\right)} P(z_t)$$
$$+ \delta(z_{t+1} - z_{b,t}) \frac{P\left(z_{t+1}|z_{b,t}, z_t, z_{f,t}\right) P_{b,t}\left(z_{b,t}|z_{t,}, S_t\right)}{P\left(z_{t+1}|z_{f,t+1}, z_{t+1}, z_{b,t+1}\right) P_{f,t+1}\left(z_t|z_{b,t}, S_t\right)} P(z_t)$$

Trian network by inverse Kullback–Leibler (KL) divergence

Our network



Target:Triple Ring



Model:Our network

- Good samples from correct support set
- Noisy samples

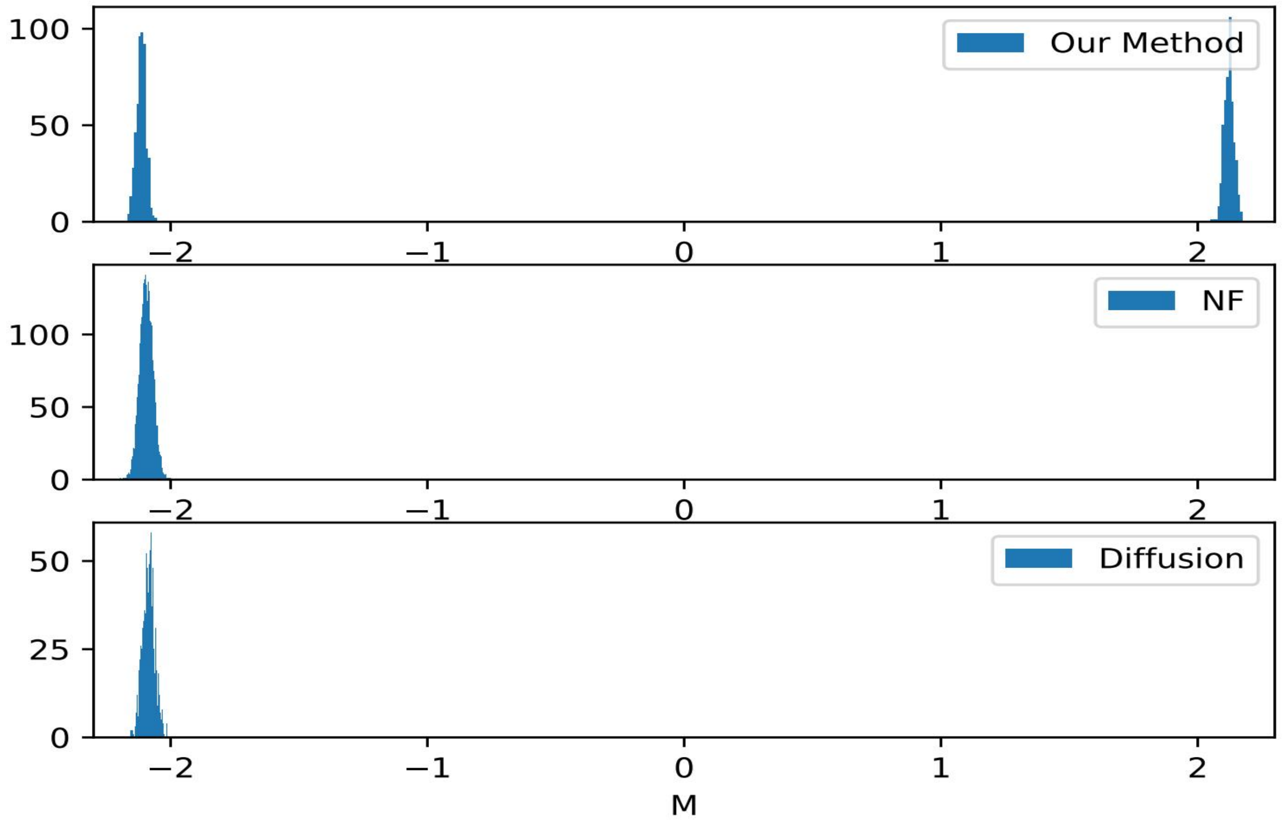
2D-ϕ^4 theory

- 2D-ϕ^4 model has $\mathbb{Z}_2$ symmetry, $S(\phi) = S(-\phi)$,

$$S(\phi) = \sum_{x \in \Lambda} -2\kappa \sum_{\hat{\mu}=1}^{2} \phi_x \phi_{x+\hat{\mu}} + (1-2\lambda)\,\phi_x^2 + \lambda \phi_x^4$$

$$|\Lambda| = N_L \times N_t = 64 \times 32$$

$$\lambda = 0.022, \;\; \kappa = 0.3$$



$$M = \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \phi$$

Flow models for multiple peak

[Daniel C. Hackett, arXiv:2107.00734[hep-lat]]

$$P_{mix}(\phi) = P(\phi) \sum_{b} \frac{t_b^{-1} \circ \phi}{P(t_a \circ t_b^{-1} \circ \phi)} \left| \frac{\partial t_a \circ z}{\partial \phi} \right|$$

1. $\{t_a\}$: a finite group of invertible transformation
2. break invertibility

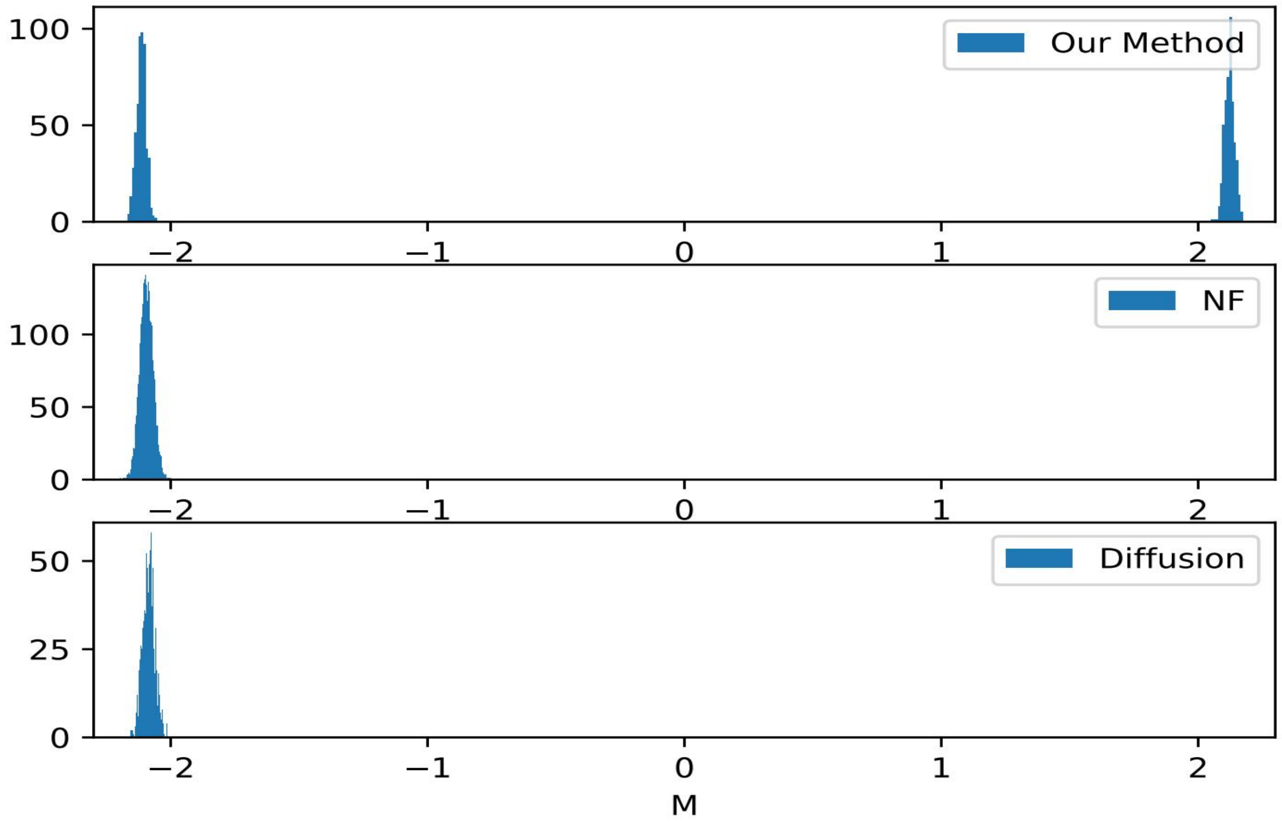
2D-ϕ^4 theory

- 2D-ϕ^4 model has $\mathbb{Z}_2$ symmetry, $S(\phi) = S(-\phi)$,

$$S(\phi) = \sum_{x \in \Lambda} -2\kappa \sum_{\hat{\mu}=1}^{2} \phi_x \phi_{x+\hat{\mu}} + (1-2\lambda)\,\phi_x^2 + \lambda \phi_x^4$$

$|\Lambda| = N_L \times N_t = 64 \times 32$

$\lambda = 0.022, \;\; \kappa = 0.3$



$$M = \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \phi$$

- Invertible NF and diffusion models recover one peak,
- Stochastic mean pathway help generate $\mathbb{Z}_2$configurations

Summary

- Invertible flow based and diffusion (possible) methods suffer from topology problem if the network learns from the action;
- Some part of configurations may not be generated with invertible flows and diffusion models;
- Stochastic mean pathway helps sample for seperated models.