Generative models for sampling of lattice field theories

Matija Medvidović[1,2], Juan Carrasquilla[3,4], Lauren E. Hayward[4], and Bohdan Kulchytskyy[4,5]

[1]Center for Computational Quantum Physics, Flatiron Institute, New York, USA, [2]Department of Physics, Columbia University, New York, USA, [3]Vector Institute for Artificial Intelligence, MaRS Centre, Toronto, Ontario, Canada, [4]Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5, Canada, [5]Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada

Abstract

We explore a self-learning Markov chain Monte Carlo method based on the Adversarial Non-linear Independent Components Estimation Monte Carlo, which uses generative models and artificial neural networks. Applying this method to the scalar φ^4 lattice field theory in the weak-coupling regime and, we greatly increase the system sizes explored to date with this self-learning technique. Our approach does not rely on a pre-existing training set of samples, as the agent systematically improves its performance by bootstrapping samples collected by the model itself. We evaluate the performance of the trained model by examining its mixing time and study the ergodicity of generated samples. When compared to methods such as Hamiltonian Monte Carlo (HMC), this approach provides unique advantages like speed of inference and compressed representation of Monte Carlo proposals for potential use in downstream tasks. [1]

Monte Carlo Simulations

A key practical concern in MCMC simulations is the autocorrelation that exists between Monte Carlo samples. Reducing the autocorrelation time [2] enables a Markov chain to become shorter while maintaining the same statistical predictive capacity. Such optimization can be achieved through a tailored design of the proposal distribution in an MCMC update.

A-NICE MC

Here we examine a general approach for a self-training Markov Chain Monte Carlo (MCMC) known as Adversarial Nonlinear Independent Components Estimation Monte Carlo (A-NICE MC) [3], where a neural network is optimized to minimize the autocorrelation in a Markov chain:

- Starting from noise, it generates a good sample in a small number of steps. This is required to reach equilibrium as quickly as possible.
- Starting from a good sample, it generates another, decorrelated good sample in as few steps as possible. This is required to continue generating good samples once equilibrium is reached.

The pairwise discriminator

In contrast to more standard GAN implementations, A-NICE MC uses a discriminator network that jointly scores pairs of samples, allowing for an estimation of autocorrelation time.

Lattice field theory

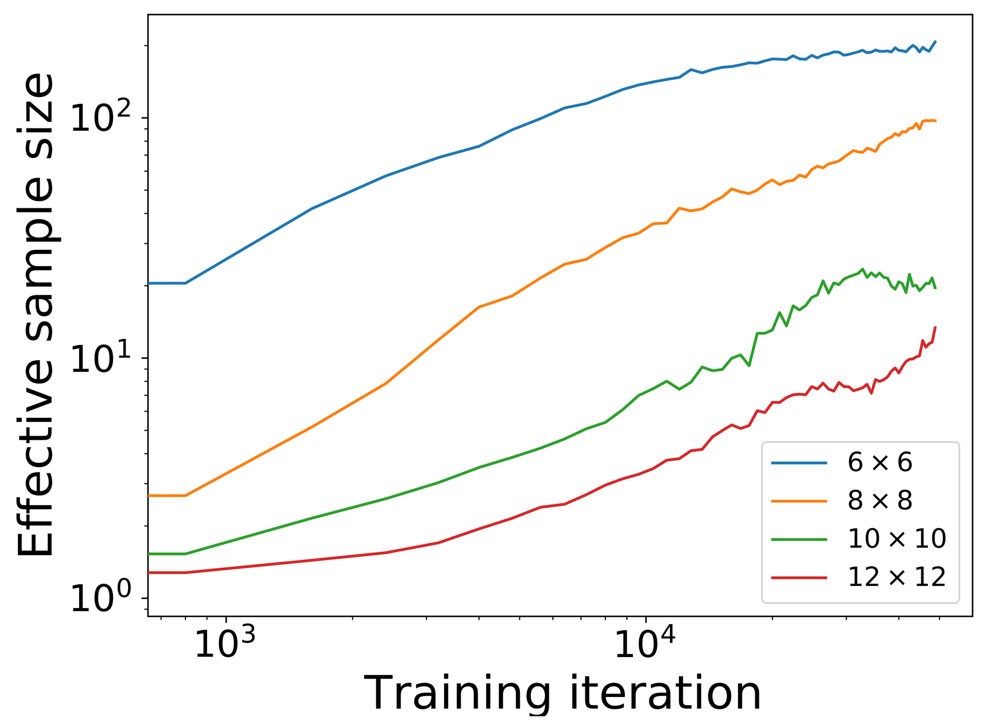
We apply A-NICE-MC a classical φ^4 lattice field theory in two dimensions. We sample a Boltzmann-type distribution of the form $P(\varphi) \propto \exp(-S(\varphi))$, where

$$S_\Lambda(\varphi) = \sum_{\mathbf{x}\in\Lambda}\left[-2\kappa\sum_{\mu=1}^{D}\varphi_{\mathbf{x}}\,\varphi_{\mathbf{x}+\hat{\mathbf{e}}_\mu} + (1-2\lambda)\,\varphi_{\mathbf{x}}^2 + \lambda\,\varphi_{\mathbf{x}}^4\right]$$

with coupling constants $\kappa, \lambda \in \mathbb{R}$. We choose $\kappa = 0.21$ and $\lambda = 0.022$ (weak coupling) for easier comparison with standard MCMC approaches.

Training procedure

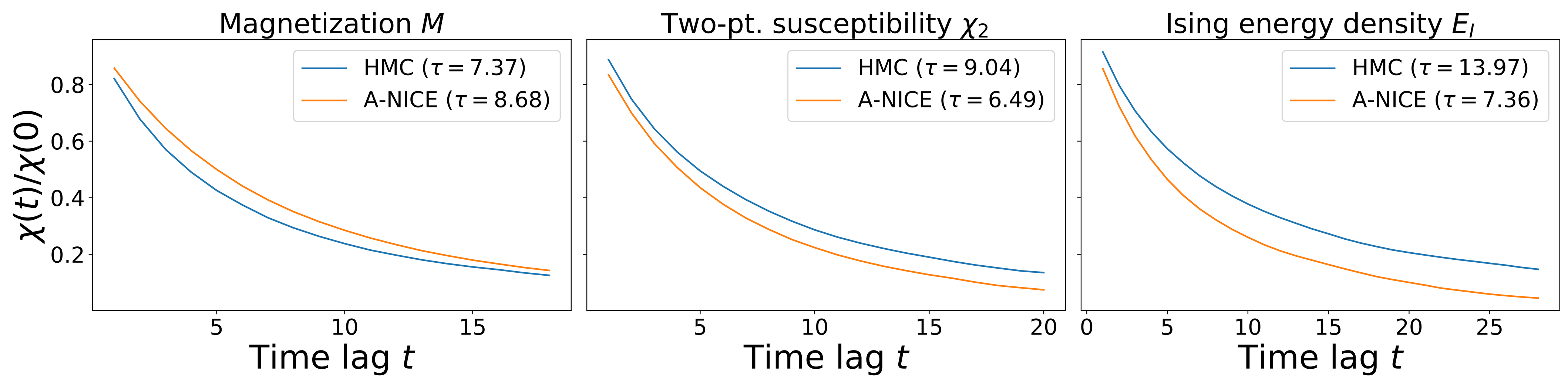
MCMC performance metrics used in this work are the autocorrelation time or, equivalently, the effective sample size. We evaluate effective sample sizes on both A-NICE MC chains during training.



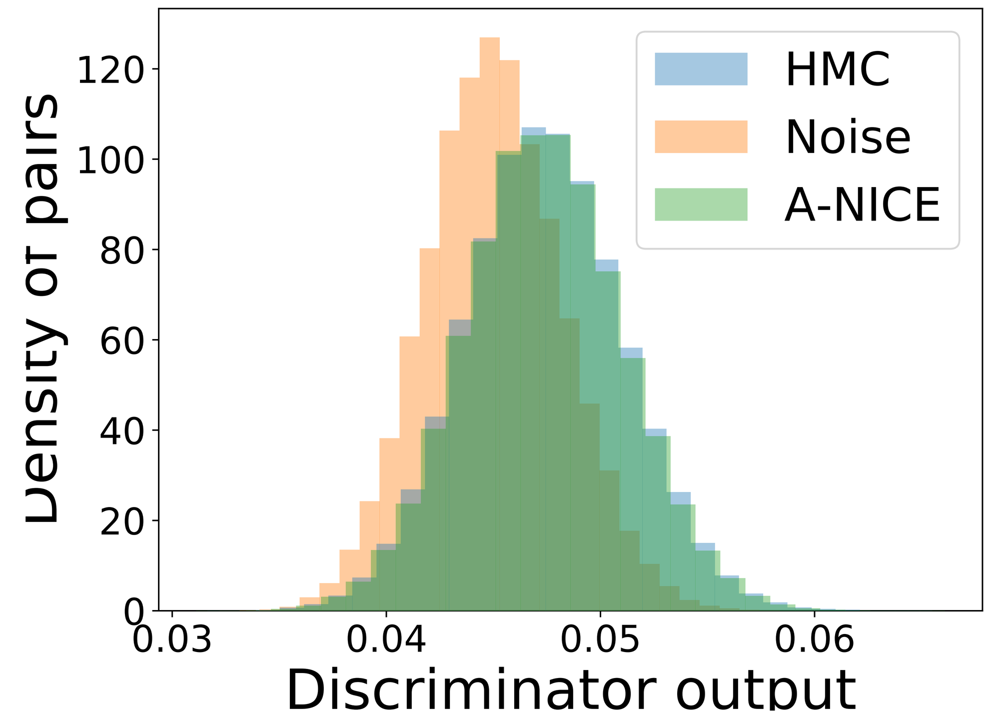
Bootstrapping data

The A-NICE MC model trains on its own data which we call "bootstrapped" datasets:

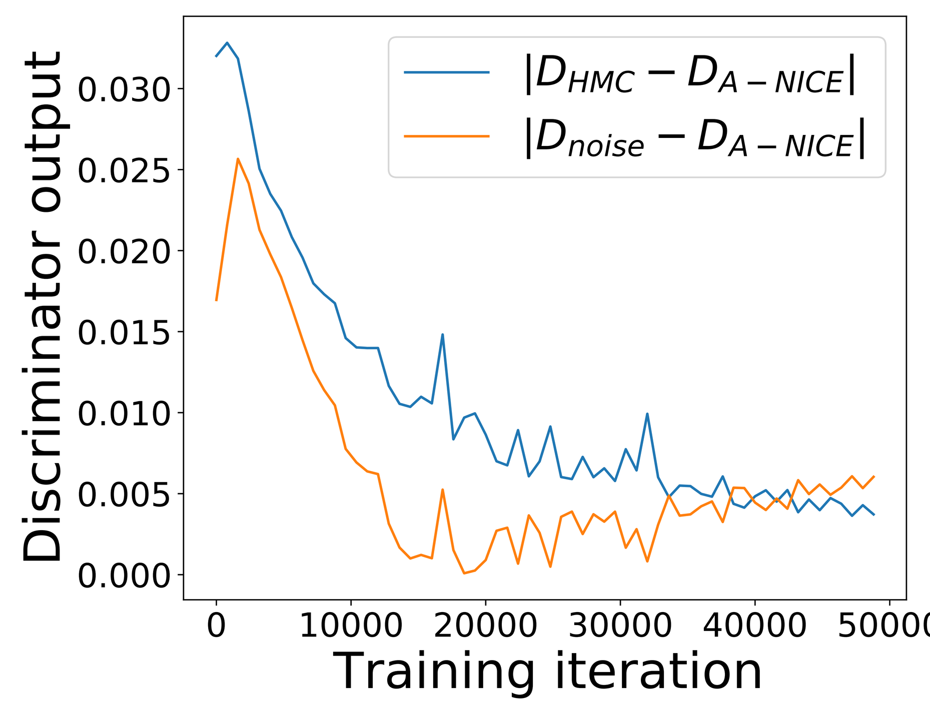
1. The proposal network is sampled to create a dataset $\mathcal{D}_0$. Samples are biased towards the correct distribution by the Metropolis-Hastings accept/reject step.
2. Train the network for a preset number of steps.
3. Repeat previous two steps to generate better and better samples.



Ergodicity



After training we compare discriminator outputs of HMC (assumed ergodic) samples, trained A-NICE samples and random Gaussian noise. trained discriminator penalizes HMC samples and A-NICE samples equally, indicating that A-NICE has been trained to be ergodic. Gausssian noise is penalized slightly differently (weak coupling) but are distinctly different from φ^4 samples.



As the discriminator optimizes, the difference $|D_{\mathrm{HMC}} - D_{\text{A-NICE}}|$ decreases and thus the A-NICE score approaches the decorrelated HMC score.

Autocorrelation

The autocorrelation function χ for an observable $\mathcal{O}$ [2] is commonly defined as

$$\chi(t) = \int_0^\infty \mathrm{d}t' \left[\mathcal{O}(t')\mathcal{O}(t+t') - \langle\mathcal{O}\rangle^2\right]$$

Observables

In the above figure, we examine autocorrelation times following of the following lattice field theory observables:

- **The "magnetization"**: $M \equiv \left\langle \frac{1}{L^D}\sum_{\mathbf{x}\in\Lambda}\varphi_{\mathbf{x}}\right\rangle$
- **The connected two-point function**:

$$G_2(\mathbf{x}) \equiv \frac{1}{L^D}\sum_{\mathbf{y}\in\Lambda}\left[\langle\varphi_{\mathbf{y}}\,\varphi_{\mathbf{x}+\mathbf{y}}\rangle - \langle\varphi_{\mathbf{y}}\rangle\,\langle\varphi_{\mathbf{x}+\mathbf{y}}\rangle\right]$$

and its momentum-space representation $G_2(\mathbf{k}) \equiv \sum_{\mathbf{x}\in\Lambda} G_2(\mathbf{x})\, e^{i\mathbf{k}\cdot\mathbf{x}} = \left\langle|\tilde{\varphi}_{\mathbf{k}}|^2\right\rangle/L^D$ where $\tilde{\varphi}_{\mathbf{k}}$ is the Fourier transform of $\tilde{\varphi}_{\mathbf{x}} \equiv \varphi_{\mathbf{x}} - \langle\varphi_{\mathbf{x}}\rangle$. Related scalar quantities are the two-point susceptibility and the Ising energy density

$$\chi_2 \equiv \sum_{\mathbf{x}\in\Lambda} G_2(\mathbf{x})\,; \quad E_I \equiv \lim_{\lambda\to\infty}\frac{1}{D}\sum_{\mu} G_2(\hat{\mathbf{e}}_\mu)\,.$$

References

[1] M. Medvidovic, J. Carrasquilla, L. E. Hayward, and B. Kulchytskyy, "Generative models for sampling of lattice field theories," `arXiv:2012.01442`.

[2] M. E. J. Newman and G. T. Barkema, "Monte Carlo Methods in Statistical Physics," *Oxford University Press* (1999) .

[3] J. Song, S. Zhao, and S. Ermon, "A-NICE-MC: Adversarial Training for MCMC,". `http://arxiv.org/abs/1706.07561`.