Scaling autoregressive models for lattice thermodynamics

Xiaochen Du[a], **Sulin Liu**[b], **Rafael Gómez-Bombarelli**[b]

[a] *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA*

[b] *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA*

* Presenting author

1. Introduction

Understanding the equilibrium distribution of states is the first step towards generating realistic materials under experimental conditions. Materials exist at finite temperatures; thus it is not a single structure that defines a static material property. For example, molecules have a distribution of conformations, each with distinct properties [1] while catalyst surfaces evolve as a function of reactants, temperature (T), and external chemical potentials (μ) [2, 3].

2. Results

In this work, we focus on periodic lattice systems that can be modeled using discrete states.

2.1 Related work

Previously, [4–6] adapted autoregressive methods (ARMs) used in image and text generation to learn lattice thermodynamics. However, these ARMs are hampered in flexibility due to: (1) their fixed generation order and (2) requiring expensive evaluation of conditional probabilities over the entire sequence length, L, during training. Due to the former, arbitrary in-/out-painting tasks for conditional generation are out of scope for these fixed-order (FO) ARMs. Due to the latter, the computational graph of automatic differentiation scales as $\mathcal{O}(L^2)$, limiting training samples to a modest lattice size, thereby hampering the accuracy of thermodynamic observables.

2.2 Contribution

We develop any-order (AO) ARMs that improve on FO-ARMs of previous works and marginalization models (MAMs) that scale training to larger lattice systems capable of sampling across T and μ (Fig. 1). We test our method on up to 20×20 Ising lattices and $4 \times 4 \times 8$ CuAu alloys, benchmarking on variational free energies, free energy estimations, and phase diagrams.

2.3 Figures and tables

Results summary for the Ising model is provided in Table 1. Fig. 2 contains detailed results for the 10×10 Ising model. Results summary for the CuAu alloy is provided in Table 2.

Conclusion

We developed any-order autoregressive and marginalization models to improve scalability in lattice thermodynamics sampling. Our results show that any-order models enable flexible in-/outpainting while matching or surpassing fixed-order models. Marginalization models further scale training to larger lattices, overcoming ARM limitations. Applied to the Ising model and CuAu alloy, our approach efficiently captures thermodynamic observables. Future work will extend these methods to larger lattice sizes and supercells, and integrate additional physical constraints for improved accuracy and scalability.

Acknowledgments

The authors thank Hoje Chun, James Damewood, Alexander Hoffman, and Juno Nam for helpful discussions and review of the work. X.D. acknowledges support from the U.S. National Science Foundation Graduate Research Fellowship under Grant No. 2141064. The authors are grateful for computation time allocated on the NERSC Perlmutter cluster under Project No. m4737 and on the MIT SuperCloud cluster.

References

[1] Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, September 2019.

[2] Zhi Wei Seh, Jakob Kibsgaard, Colin F. Dickens, Ib Chorkendorff, Jens K. Nørskov, and Thomas F. Jaramillo. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355(6321):eaad4998, January 2017. Publisher: American Association for the Advancement of Science.

[3] Albert Bruix, Johannes T. Margraf, Mie Andersen, and Karsten Reuter. First-principles-based multiscale modelling of heterogeneous catalysis. *Nature Catalysis*, 2(8):659–670, August 2019. Number: 8 Publisher: Nature Publishing Group.

[4] Dian Wu, Lei Wang, and Pan Zhang. Solving Statistical Mechanics Using Variational Autoregressive Networks. *Physical Review Letters*, 122(8):080602, February 2019. Publisher: American Physical Society.

[5] Kim A. Nicoli, Shinichi Nakajima, Nils Strodthoff, Wojciech Samek, Klaus-Robert

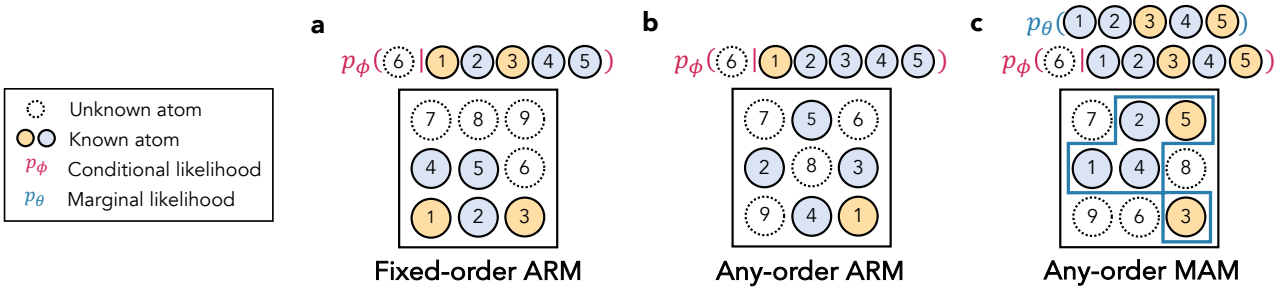

Fig. 1: Autoregressive models investigated in this work.

Table 1: Ising model generation results. The best result for each lattice size is bolded. FO models cannot perform arbitrary in-/out-painting tasks. 20×20 Ising ARM models exceed GPU capacity. $\langle \cdot \rangle$ is equivalent to $\mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})}[\cdot]$.

	5×5	10×10	20×20	
Model	Trained	Trained	10×10 In-/out-painted	Trained
ARM		$\langle \log p(\mathbf{x}) - \log f(\mathbf{x}) \rangle$ (↓)		
FO-MLP	-26.26	-104.69	-	-
AO-MLP	**-26.27**	**-104.77**	**-415.13**	-
MAM		$\langle \log p(\mathbf{x}) - \log f(\mathbf{x}) \rangle$ (↓)		
AO-MLP	-26.22	-98.90	-401.31	Mode collapse
AO-Transformer	-26.25	-104.66	**-415.13**	-414.34

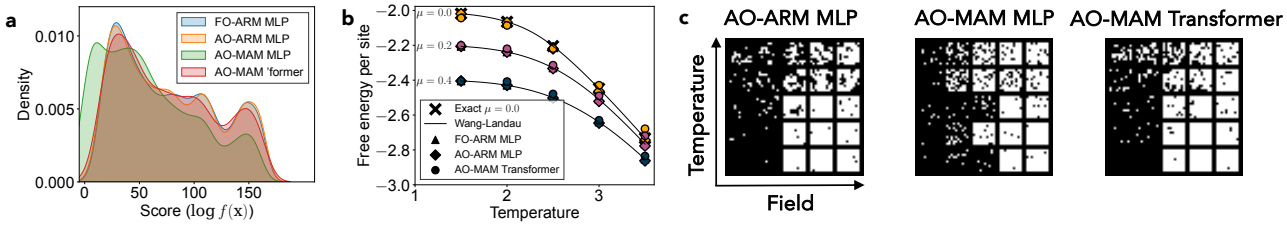

Fig. 2: 10×10 Ising results. (a) Score distribution, (b) free energies, and (c) generated samples.

Table 2: CuAu model generation results with the best for each lattice size bolded.

	$2 \times 2 \times 4$	$4 \times 4 \times 8$
Model	Trained	$2 \times 2 \times 4$ Inpainted
ARM	$\langle \log p(\mathbf{x}) - \log f(\mathbf{x}) \rangle$ (↓)	
FO-MLP	**-32.92**	-
AO-MLP	**-32.92**	-215.29
MAM	$\langle \log p(\mathbf{x}) - \log f(\mathbf{x}) \rangle$ (↓)	
AO-MLP	-31.87	**-223.16**
AO-Transformer	-32.84	-217.06

Müller, and Pan Kessel. Asymptotically unbiased estimation of physical observables with neural samplers. *Physical Review E*, 101(2):023304, February 2020. Publisher: American Physical Society.

[6] James Damewood, Daniel Schwalbe-Koda, and Rafael Gómez-Bombarelli. Sampling lattices in semi-grand canonical ensemble with autoregressive machine learning. *npj Computational Materials*, 8(1):1–10, April 2022. Number: 1 Publisher: Nature Publishing Group.