Accelerating quantum Monte Carlo simulations using neural networks: applications to the Holstein model and beyond

Shaozhi Li[1], Philip M. Dee[1], Ehsan Katami[2], and **Steven Johnston**[1]

[1]Department of Physics and Astronomy, The University of Tennessee, Knoxville, Tennessee 37996, USA

[2]Department of Physics and Astronomy, San José State University, San José, California 95192, USA

Abstract

Monte Carlo (MC) simulations are essential computational method with widespread use throughout all areas of science and engineering. We present a method for accelerating lattice MC simulations using artificial neural networks that are trained to perform local and global moves in configuration space. Our networks take local spacetime MC configurations as input features and can, therefore, be trained using samples generated by conventional MC runs on smaller lattices before being utilized for simulations of larger systems. This new approach is benchmarked for the case of determinant quantum Monte Carlo (DQMC) studies of the two-dimensional Holstein model. We find that the artificial neural networks are capable of learning an unspecified effective model that accurately and efficiently reproduces the MC configuration weights of the Holstein model and achieve an order of magnitude speedup over the conventional DQMC algorithm. Our approach is broadly applicable to many classical and quantum lattice MC algorithms.

Introduction

Monte Carlo (MC) methods are integral in science and engineering fields. For example, quantum Monte Carlo (QMC) simulations have provide critical and unbiased insights into the Hubbard and Holstein Hamiltonians and beyond.

MC simulations face critical challenges:

1. Computationally expensive “moves” in configuration space.
2. Limitations to small clusters & finites size effects.
3. Long autocorrelation times (e.g. near phase transitions).
4. The infamous *Fermion sign* problem.

Can we teach the computer to overcome one or more of these limitations?

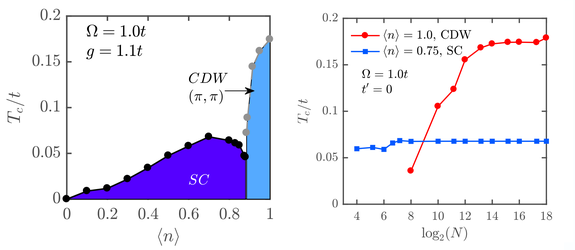
The Holstein Model

The *Holstein* is the simplest model describing itinerant electrons interacting with lattice vibrations:

$$H = -t \sum_{\langle i,j\rangle,\sigma} c^\dagger_{i,\sigma} c_{j,\sigma} - \mu \sum_{i\sigma} c^\dagger_{i,\sigma} c_{i,\sigma}$$
$$+ \Omega \sum_i \left(b^\dagger_i b_i + \frac{1}{2} \right) + g \sum_{i,\sigma} c^\dagger_{i,\sigma} c_{i,\sigma} \left(b^\dagger_i + b_i \right)$$

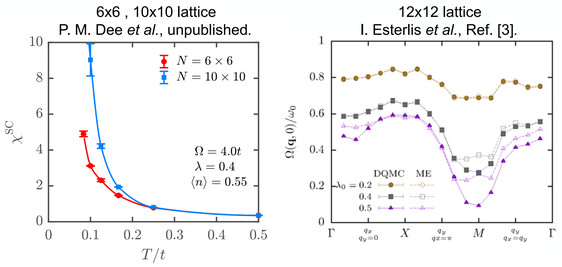
Diagrammatic perturbation theory:

Competing insulating charge-density-wave and superconducting phases (left) and significant finite size effects (right) [1].



Quantum Monte Carlo Simulations:

Limited by long autocorrelation times that restrict simulations to small cluster sizes [2,3]



Determinant QMC & Holstein Model

DQMC simulations of the Holstein model calculate the partition function, expressed as a multi-dimensional integral over phonon fields

$$Z = \mathrm{Tr}\left(e^{-\beta H}\right) = \int W(\{X\})\, d^D X,$$

where the weights are related to matrix determinants

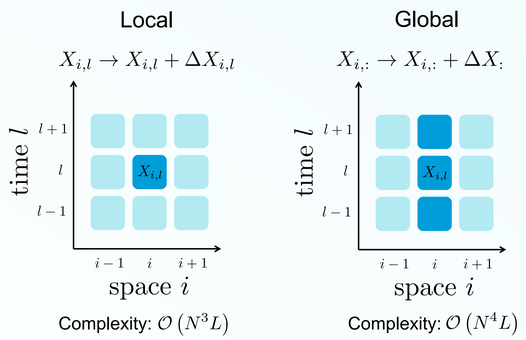
$$W(\{X\}) = e^{-S_{\mathrm{ph}}\Delta\tau} \det(M_\uparrow) \det(M_\downarrow).$$

The displacements are Monte Carlo integrated using the Metropolis-Hastings algorithm, where updates are accepted with probability

$$p = \frac{W(\{X'\})}{W(\{X\})} \propto \frac{\det(M'_\uparrow)\det(M'_\downarrow)}{\det(M_\uparrow)\det(M_\downarrow)}.$$

Performing updates requires expensive determinant calculations.

Two types of Monte Carlo moves are required:



A bosonic effective model

The matrix determinants are due to the electron degrees of freedom. To circumvent them, we define an *effective* bosonic model E_{eff} that depends only on the lattice degrees of freedom

$$W(\{X\}) = e^{-S_{\mathrm{ph}}\Delta\tau} \det(M_\uparrow) \det(M_\downarrow) \equiv e^{-\beta E_{\mathrm{eff}}(\{X\})}.$$

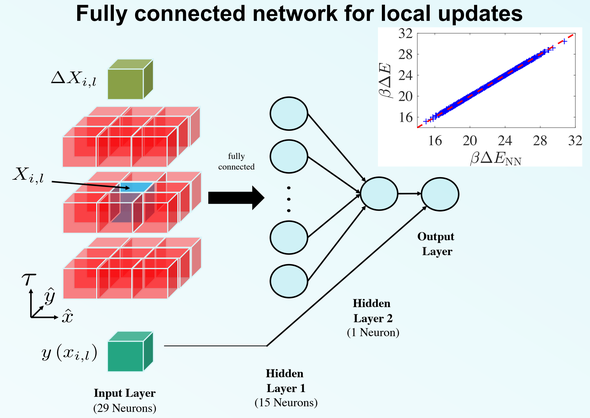
The *self-learning Monte Carlo* [5, 6] methods specify E_{eff} at the onset. Instead, we designed artificial neural networks to learn the effective model from training examples.

The effective model is used propose many updates that are accepted with cumulative probability given by

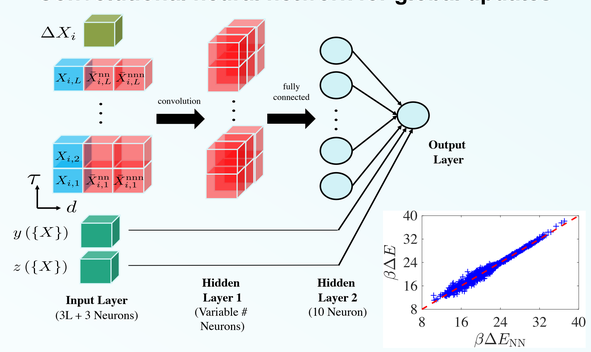
$$p_c = \frac{W(\{X'\})}{W(\{X\})} \frac{e^{-\beta(E_{\mathrm{eff}}(\{X\})}}{e^{-\beta(E_{\mathrm{eff}}(\{X'\})}}.$$

Artificial Neural Networks

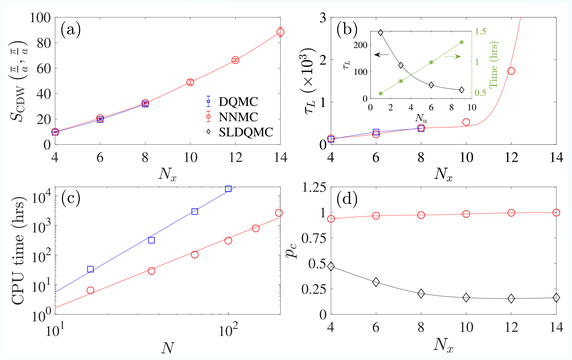
Fully connected network for local updates



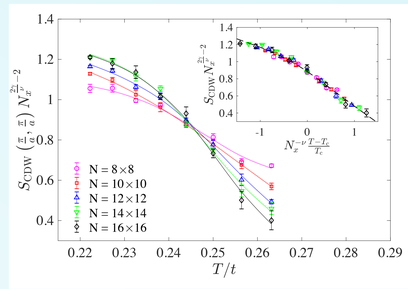
Convolutional neural network for global updates



Performance



Charge-Density-Wave Transition



$\Omega = 0.5t$

$\lambda = 0.5$

$T_c = 0.244/t$

Networks trained on a 6x6 cluster generalize to larger systems and can be used to perform a scaling analysis.

Conclusions

- We designed artificial neural networks capable of **accurately** and **effectively** predicting DQMC moves for the Holstein model.
- The networks use local field configurations and can be easily trained using small systems.
- Accelerated the algorithm by an order of magnitude, allowing for scaling analysis of the CDW transition.
- The method is generalizable across the phase diagram, and can be used to accelerate other types of lattice Monte Carlo simulations.

References & Acknowledgements

1) P. M. Dee et al., Phys. Rev. B **99**, 024514 (2019).
2) M. Hohendler and T. C. Lang, *Computational Many-Particle Physics* (2008).
3) I. Esterlis *et al.*, Phys. Rev. B **97**, 140501(R) (2018).
4) S. Li *et al.*, arXiv:1905.07440 (2019); In press at Phys. Rev. B.
5) J. Liu et al., Phys. Rev. B **95**, 041101 (2017).
6) C. Chen *et al.*, Phys. Rev. B **98**, 041102 (2018).

This work was supported by the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Basic Energy Sciences, Division of Materials Sciences and Engineering. E. K. is supported by the National Science Foundation Grant No. DMR-1609560. An award of computer time was provided by the INCITE program. This research also used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.