Energy based models for tomography of quantum spin-lattice systems

Abhijith J, Marc Vuffray, Andrey Y. Lokhov
Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM 87545
{abhijithj, vuffray, lokhov}@lanl.gov

Abstract

We present a novel method to learn Energy-Based Models (EBM) from quantum tomography data. We represent quantum states via distributions generated by generalized measurements and use state-of-the-art algorithms for energy function learning to obtain a representation of these states as classical Gibbs distributions. Our results show that this method is especially well suited for learning quantum thermal states. For the case of ground states, we find that the learned EBMs often have an effective temperature that makes learning easier, especially in the paramagnetic phase.

1 Introduction

Finding succinct representations for quantum states remains a problem of great interest. Although a general quantum state of n particles resides in a Hilbert space of exponentially large dimension, states of interest often occupy a smaller region of this space. This observation has prompted the use of machine learning algorithms to identify the structure present in physical states, so as to represent them efficiently on a classical computer.

In this work, we will learn generative models for quantum states given measurement data obtained from independent copies of the state. This problem can be seen as a generalization of the standard generative modeling setup, where the data comes from a classical distribution. Numerous models and training methods that were originally developed in the classical machine learning literature have been applied to solve this quantum modeling problem. Translating these models to the quantum regime is often not easy as we have to make sure that the ML model represents a valid quantum state, while still retaining its expressivity. A popular approach, used for pure states, is to use separate neural nets to model the amplitude and phases of a state [14]. In practical applications states are often not pure and modeling mixed states is more challenging. Previous works have attempted to model mixed states by adding ancillary qubits [15], using matrix factorization[16], by exploiting low-rank properties [9], etc. But, mixed states defined as density matrices are direct generalizations of probability distributions, so a natural way to model them is via classical distributions. This approach uses the distribution of measurement outcomes as a model for the quantum density matrix [2].

The above discussion still does not fix the ideal way distributions must be modeled so as to represent states of local quantum systems. In this work, we will deal with states of quantum spins arranged on a lattice. For distributions obtained by measuring these states, we argue that Gibbs distributions of classical spin Hamiltonian provide an effective model. These models which represent a distribution as a classical Gibbs distribution are known as Energy Based Models (EBM) [3, 6, 7].

$$\mu(\underline{\sigma}) \propto \exp(E(\underline{\sigma})). \tag{1}$$

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

The function E is known as the *energy function*. The learning problem in the case of these models amounts to learning a representation for E, given i.i.d samples from μ. Additionally, we must be able to use the learned representation to generate fresh samples.

EBMs are a natural way of modeling classical spin systems at a finite temperature. The power of EBMs comes from the fact that even simple energy functions can give rise to complicated distributions functions with rich structure [10]. Moreover, many features of physical systems like locality and the absence of higher-order interactions are most evident at the level of the energy function. The distribution itself can obscure these properties. Most EBMs arising in classical statistical physics can be learned in a provably efficient fashion using an optimal number of samples drawn from μ [17, 18, 8].

In this work, we will show examples of quantum states that can be effectively modeled using EBMs. The method we outline will not be tied to any specific function family like neural nets. Instead, the practitioner can choose a parametric function family to represent the energy function based on prior information available about the state. We will also show how this choice can be made for any family of states based on an exact learning method. We find that EBMs are effective in modeling even zero-temperature states of quantum systems. This is surprising as low-temperature classical states are usually harder to learn [13]. We observe that this is due to an effective temperature property of EBMs learned from samples generated by ground states.

2 Details of Learning Algorithm

2.1 Quantum states as distributions

Consider a quantum state represented as a $d \times d$ density matrix. Any q-outcome measurement on this state can be represented by a set of q positive matrices that sum to the identity matrix, $\{M_i | i \in [q]\}, M_i \succcurlyeq 0, \sum_i M_i = I$. This set of matrices form a *Positive Operator Valued Measure*(POVM). The probability of observing outcome i is then given by the Born rule [2, 11],

$$\mu(i) = \text{Tr}(\rho M_i), \quad i \in [q]. \tag{2}$$

For an n-body system, a many-body POVM can be constructed by tensoring together single-body POVMs,

$$\mu(\underline{\sigma}) = \text{Tr}(\rho M_{\sigma_1} \otimes M_{\sigma_2} \otimes \dots M_{\sigma_n}), \quad \underline{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n) \in [q]^n \tag{3}$$

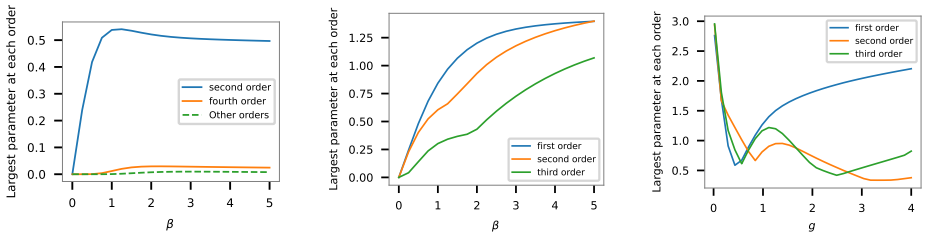
This relation maps any many-body quantum state to a distribution over classical spin states. In this work we will work with n-qubit systems with two standard POVMs, Tetrahedral $(q = 4)$ and $Z2$ $(q = 2)$ POVMs.

The Tetrahedral POVM is more computationally challenging to deal with. But it gives a complete representation of any state i.e given samples from (3) the expectation value of any observable can be estimated. This estimation can be done efficiently if the observable is local. The $Z2$ POVM corresponds to computational basis measurements. Being just a two-outcome POVM, it is computationally easier to deal with. But once we learn a model from this POVM, only diagonal observables can be estimated.

2.2 Learning EBMs by interaction screening

Measurement outcomes from independent copies of a quantum state measured using a POVM are just i.i.d samples drawn from the distribution defined in (3). We represent this distribution by an EBM as in (1) and use the *Interaction Screening*(IS) method to learn the correct energy function. For each spin (σ_u) in the model, the IS method finds a representation for the part of the energy function connected to this spin. This *local energy* for each spin can then be used to find all the single variable conditionals of the model $\mu(\sigma_u | \sigma_{\setminus u})$. These conditionals can in turn be used to generate new samples from the learned model via Gibbs sampling [17, 18, 8, 5].

The IS method can find a representation for the local energies using any parametric function family. If we use polynomials to represent the energy function, then this method can be trained via a convex optimization problem. This variant of the IS method is known to asymptotically sample optimal for the problem of learning Gibbs distributions [18, 13]. On the other hand, we can choose to learn the



(a) (b) (c)

Figure 1: Results of exact learning in the polynomial basis (a) [1D Ferro thermal state, $g = 1$, $n = 5$, $Z2$ POVM] We see that the energy function has predominantly second order terms (b) [1D Ferro thermal state, $g = 1$, $n = 5$, Tetrahedral POVM] The energy function now has third order terms which are non-negligible in size. (c) [1D Anti-ferro ground state, $n = 6$ Tetrahedral POVM] Energy function has third order terms. But the learned energy functions have a non-zero effective temperature at non-zero values of g despite being learned from the ground state. (b) and (c) are learned using a third-order polynomial ansatz, so the energy function could have even higher order terms.

local energies using neural nets. This leads to a non-convex optimization problem in the training phase. But this variant of the IS method can efficiently learn energy functions with higher-order terms. Learning an energy function with $k-$th order terms has $O(n^k)$ memory and time requirements. A neural net parametrization is seen to make these high-order problems significantly more efficient in practice. In either case, if the function family is chosen to exactly represent the local energy function, then the global optima of the IS method can be shown to correspond to the true energy function from which the training samples were drawn, in the limit of infinite samples $(N \to \infty)$ [5]. Further technical details regarding the IS method can be found in Ref. [5].

3 Numerical Results

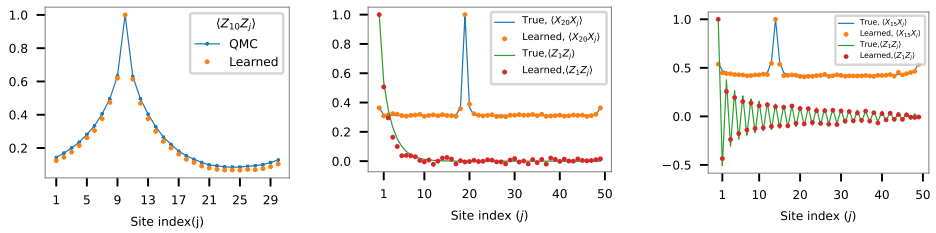
Here we will present results from learning thermal states and ground states of quantum spin models. We will choose our Hamiltonians from the transverse Ising model (TIM) family, $H = J \sum_{\langle i,j \rangle} Z_i Z_j + g \sum_i X_i$. Here the $\langle i, j \rangle$ summation goes over the nearest neighbors on a lattice. We choose the TIM family as samples can be produced from this model easily in 1D using tensor network techniques [4] and in 2D using quantum Monte-Carlo methods [1, 12].

3.1 Exact learning and nature of the energy function

Before we attempt to learn an EBM for any state we have to choose a parametric function family to represent the energy function. To make this choice for a family of states it is instructive to first look at states on a few qubits. For such a small system, the polynomial ansatz can include higher order terms (possibly all orders) and learning can be done in the $N \to \infty$ limit. The results of these experiments for three classes of states are given in Figure 1. In these experiments, the absolute value of the largest interaction term at each order in the energy function is plotted against inverse temperature(β) for thermal states and transverse field strength (g) for ground states. We see that thermal states of the TIM measured in the $Z2$ basis have predominantly second-order terms in their energy function and can be represented using quadratic polynomials. For the other two classes of states, we observe the presence of higher order terms and this requires us to use neural nets to model the energy function in the EBM.

3.2 Results on learning larger states

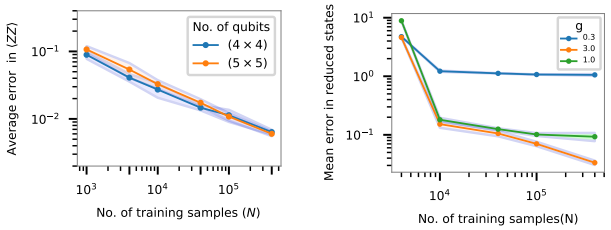
The results of learning systems of larger size from three classes of states are given in Figure 2. We see that both EBM with both polynomial and neural net energy functions are capable of learning long-range behavior of two-body expectation values, even at the critical point. As expected from previous works on the IS method, we see that increasing the number of samples consistently reduces



(a) (b) (c)

Figure 2: Two site expectation values compared with the true model (a) 1D thermal state with $g = 1, \beta = 2$ measured in $Z2$ basis. EBM learned using polynomial ansatz. (b) 1D thermal state $g = 1, \beta = 1$ measured using Tetrahedral basis. (c) 1D ground state $g = 1$. (b) and (c) are learned using neural net ansatz

.



(a) (b)

Figure 3: (a) Error in ZZ expectation values for 2D Ferromagentic TIM $(g = 3.044, \beta = 1)$. The error consistently decreases with sample size and does not show an appreciable increase with system size (b) Average error in one and two body reduced density matrices for the ground state of anti-Ferro TIM at different values of g. Higher values of g lead to higher effective temperatures and more efficient learning $(n = 50)$.

the errors in the learned model for a state at finite temperature (Figure 3a). For the case of ground states, we see that the IS method is more effective in the paramagnetic phase of the model than in the ordered phase (Figure 3b). This can be understood from Figure 1c, which shows the terms in the energy function blowing up as $g \to 0$. This implies that the EBM has a low effective temperature (strength of the terms is proportional to inverse temperature), which makes learning difficult. On the other hand, as $g \to \infty$, we get closer to the uniform superposition state with no correlation between the different parts of the system. This makes learning the distribution easier.

4 Conclusion

In this work, we have shown how EBMs can be used to model quantum states given data from quantum tomography experiments. These models are seen to be especially effective when dealing with quantum thermal states and ground states in a paramagnetic phase. We have also observed that the learned models have an effective temperature property even in the case of zero temperature states that can often aid learning. The IS method used here for learning these models is flexible enough to work with different parametric function families and can make used of prior information available about the nature of the true energy function representing the model. Further extensions (not shown in this short paper) of the IS method are possible to make use of known symmetries of states like translational invariance and permutation invariance.

5 Impact statement.

We don't expect the work presented here to have any direct ethical implications on human society.

Acknowledgements The research presented here was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project numbers 20210674ECR and 20220545CR-CNL.

References

[1] Sergey Bravyi, David P Divincenzo, Roberto Oliveira, and Barbara M Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Information & Computation*, 8(5):361–385, 2008.

[2] Juan Carrasquilla, Giacomo Torlai, Roger G Melko, and Leandro Aolita. Reconstructing quantum states with generative models. *Nature Machine Intelligence*, 1(3):155–161, 2019.

[3] Yilun Du and Igor Mordatch. Implicit generation and modeling with energy based models. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[4] Matthew Fishman, Steven R. White, and E. Miles Stoudenmire. The ITensor software library for tensor network calculations, 2020.

[5] Abhijith J, Andrey Lokhov, Sidhant Misra, and Marc Vuffray. Learning of discrete graphical models with neural networks. *Advances in Neural Information Processing Systems*, 33:5610–5620, 2020.

[6] Taesup Kim and Yoshua Bengio. Deep directed generative models with energy-based probability estimation. *arXiv preprint arXiv:1606.03439*, 2016.

[7] Rithesh Kumar, Sherjil Ozair, Anirudh Goyal, Aaron Courville, and Yoshua Bengio. Maximum entropy generators for energy-based models. *arXiv preprint arXiv:1901.08508*, 2019.

[8] Andrey Y Lokhov, Marc Vuffray, Sidhant Misra, and Michael Chertkov. Optimal structure and parameter learning of Ising models. *Science advances*, 4(3):e1700791, 2018.

[9] Abhijeet Melkani, Clemens Gneiting, and Franco Nori. Eigenstate extraction with neural-network tomography. *Physical Review A*, 102(2):022412, 2020.

[10] Marc Mézard, Giorgio Parisi, Nicolas Sourlas, Gérard Toulouse, and Miguel Virasoro. Replica symmetry breaking and the nature of the spin glass phase. *Journal de Physique*, 45(5):843–854, 1984.

[11] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[12] Anders W Sandvik. An introduction to quantum monte carlo methods. In *Strongly Correlated Magnetic and Superconducting Systems*, pages 109–135. Springer, 1997.

[13] Narayana P Santhanam and Martin J Wainwright. Information-theoretic limits of selecting binary graphical models in high dimensions. *IEEE Transactions on Information Theory*, 58(7):4117–4134, 2012.

[14] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Neural-network quantum state tomography. *Nature Physics*, 14(5):447–450, 2018.

[15] Giacomo Torlai and Roger G Melko. Latent space purification via neural density operators. *Physical review letters*, 120(24):240503, 2018.

[16] Filippo Vicentini, Riccardo Rossi, and Giuseppe Carleo. Positive-definite parametrization of mixed quantum states with deep neural networks. *arXiv preprint arXiv:2206.13488*, 2022.

[17] Marc Vuffray, Sidhant Misra, Andrey Lokhov, and Michael Chertkov. Interaction screening: Efficient and sample-optimal learning of Ising models. In *Advances in Neural Information Processing Systems*, pages 2595–2603, 2016.

[18] Marc Vuffray, Sidhant Misra, and Andrey Y Lokhov. Efficient learning of discrete graphical models. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124017, 2022.