Sub-Exponential ML algorithm for predicting ground state properties

Lauren Preston[1] and Shivashankar[1]

Abstract. Analysing properties of ground state of a quantum systems, is an important problem with applications in various domains. Recently, Huang et al. [2021] demonstrate how machine learning algorithms can be used to efficiently solve this problem with formal guarantees. However this method requires an exponential amount of data to train. In this work we show a method with improved efficiency for a wide class of energy operator. In particular, we show an ML-based method for predicting ground state properties for structured Hamiltonian with sub-exponential scaling in training data. The method relies on efficiently learning low-degree approximation of the energy operator.

1 Introduction

The challenge of predicting the ground state of a quantum system is an important one that has applications in a variety of fields, such as quantum machine learning[Arunachalam and de Wolf, 2017, Biamonte et al., 2017, Schuld and Killoran, 2019], variational quantum algorithms[Cerezo et al., 2021, Gibbs et al., 2022], experimental quantum physics [Carleo and Troyer, 2017, Sharir et al., 2020] and quantum benchmarking[Scott, 2008, Levy et al., 2021]. Huang et al. [2021] show that machine learning can be applied with a classical-shadow [Huang et al., 2020] based representation of quantum states, to effectively handle this problem. Nevertheless, this approach has a sample complexity that is exponential in the amount of training data. Specifically, their proposed algorithm has a sample complexity of $\mathcal{O}(n^{\frac{c}{\epsilon}})$ for a prediction error ϵ. As such, when the prediction error ϵ is small, a significant amount of training data is required to achieve that error. In this work, we propose a method that is more effective than previous approaches for a diverse range of energy operators. The method relies upon low-degree approximation to the energy. Our method uses a representation known as classical shadow, which is a condensed classical description of a many-body quantum state [Flammia and Preskill, 2022]. This description can be created in quantum experiments and can be used to predict many of the attributes of the state. Our method improves over existing ML based algorithms for predicting ground state features in terms of sample complexity. Specifically, for structured Hamiltonians with sub-exponential scaling of training data

Classical ML could be used to generalize from training data that are obtained from either quantum experiments or classical simulations; the same rigorous performance guarantees apply in either case. Even if the training data are generated classically, it could be more efficient and more accurate to use ML to predict properties for new values of the input x, rather than doing new simulations which could be computationally very demanding and of unverified reliability. Promising insights into quantum many-body physics are already being obtained using classical ML based on classical simulation

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

data [Deng et al., 2017, Nomura et al., 2017, Zhang et al., 2017, Vargas-Hernández et al., 2018, Schütt et al., 2019, Zhang et al., 2020, Kawai and Nakagawa, 2020].

2 Preliminaries and Related Work

2.1 Formulation

Consider an m-dimensional vector $x \in [-1,1]^m$ that parameterizes an n-qubit gapped geometrically local Hamiltonian given as

$$H(x) = \sum_j h_j(\boldsymbol{x}_j), \tag{1}$$

$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ are $\mathcal{O}(1)$ sized vectors parameterizing the few-body interaction $h_j(\boldsymbol{x}_j)$. For example, x_j might be the coupling coefficients between a node and its neighbours in an Ising model. Let $\rho(x)$ be the ground state of $H(x)$ and O be a sum of geometrically local observables with $\|O\|_\infty \leq 1$.

The goal is to learn a function $h^*(x)$ that approximates the ground state property $\mathrm{tr}(O\rho(x))$,

$$\big(x_\ell, y_\ell\big), \quad \forall \ell = 1, \ldots, N, \tag{2}$$

where $y_\ell \approx \mathrm{tr}(O\rho(x_\ell))$ records the ground state property for $x_\ell \in [-1,1]^m$ sampled from an arbitrary unknown distribution $\mathcal{D}$.

The setting considered in this work is very similar to that in Huang et al. [2021], but we assume the geometry of the n-qubit system to be known.

Adiabatic quantum computation [Farhi et al., 2000, Aharonov et al., 2008, Wan and Kim, 2020], focuses on finding ground states of special Hamiltonian to perform computation. However, unlike these works we would not use any quantum memory, or explicit description of the operator O or any information about an adiabatic path to hamiltonian H.

We prove that given $\epsilon = \Theta(1)$, the improved ML algorithm can use a dataset size of

$$N = \mathcal{O}\left(\log\left(n\right)\right), \tag{3}$$

to learn a function $h^*(x)$ with an average prediction error of at most ϵ,

$$\mathop{\mathbb{E}}_{x\sim\mathcal{D}} \left|h^*(x) - \mathrm{tr}(O\rho(x))\right|^2 \leq \epsilon, \tag{4}$$

The correctness of the method relies upon results on optimizing k-local Hamiltonians Dinur et al. [2006], Barak et al. [2015], Harrow and Montanaro [2017], Anshu et al. [2021], Flammia and Preskill [2022]. The efficiency of the method is based upon recent results in quantum Bohnenblust-Hille inequalities Rouzé et al. [2022], Bohnenblust and Hille [1931].

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

2.2 Classical Shadows

Classical shadows are an efficient classical representations of quantum systems. The fundamental idea of this representations is similar to random projections. We use the terminology borrowed from earlier works van Enk and Beenakker [2012], Ohliger et al. [2013], Paini and Kalev [2019], Huang et al. [2020]. An n-qubit quantum state ρ can be approximated by performing randomized single-qubit Pauli measurements on T copies of ρ. If we measure every qubit of the state ρ in a random Pauli basis X, Y or Z, and collect the observations, and repeat the procedure T times, we are left with a set of measurement

$$S_T(\rho) = \left\{ |s_i^{(t)}\rangle : \ i \in \{1, \ldots, n\}, \ t \in \{1, \ldots, T\} \right\}$$

where each $|s_i^{(t)}\rangle \in \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |\mathrm{i}+\rangle, |\mathrm{i}-\rangle\}$ corresponds to an eigenstate of the corresponding Pauli operator.

Each element is a highly structured single-qubit pure state, and there are nT of them in total. So, $3nT$ bits suffice to store the entire collection in classical memory. The randomized measurements can be performed in actual physical experiments or through classical simulations. Resulting data can then be used to approximate the underlying n-qubit state ρ:

$$\rho \approx \sigma_T(\rho) = \frac{1}{T} \sum_{t=1}^{T} \sigma_1^{(t)} \otimes \cdots \otimes \sigma_n^{(t)} \quad \text{where} \quad \sigma_i^{(t)} = 3|s_i^{(t)}\rangle\langle s_i^{(t)}| - \mathbb{I}, \tag{5}$$

and $\mathbb{I}$ denotes the 2×2 identity matrix. This *classical shadow* representation Huang et al. [2020] asymptotically reproduces the global density matrix. By the Hoeffding-Chernoff bound, one can also show that with $\bar{T} = \mathcal{O}(\log(n)/\epsilon^2)$ one can get an ϵ-accurate approximation of the density matrix as well. This, implies that with $T > \bar{T}$ experiments, we can use $\sigma_T(\rho)$ to predict local functions (like expectation values).

As detailed next, this classical shadow representation is utilized by Huang et al. [2021] to build an ML algorithm for estimation of local properties of ground states ρ.

2.3 Predicting ground states of quantum many-body systems

Huang et al. [2021] consider the task of predicting ground state properties for finite many-body systems. For this purpose they propose training an ML algorithm on a dataset collected from quantum experiments over a parametric family of Hamiltonians $H(x)$. Before the training of the ML algorithm, many Hamiltonians $H(x)$ are sampled, the classical shadow of the corresponding ground state $\rho(x)$ of $H(x)$ is obtained. The full training data of size N is given by $\left\{x_\ell \to \sigma_T(\rho(x_\ell))\right\}_{\ell=1}^{N}$, where T is the number of measurements in the construction of the classical shadows at each value of x_ℓ.

The ML models is trained on this size-N training data, such that when given the input x_ℓ, the ML predicts a vector representation $\hat{\sigma}(x)$ that approximates $\sigma_T(\rho(x_\ell))$.

In particular, they use a Nadaraya-Watson estimator [Nadaraya, 1964, Watson, 1964] (a version of nearest-neighbour regression) with a kernel function [Bierens, 1988].

$$\hat{\sigma}(x) = \frac{1}{N} \sum_{\ell=1}^{N} \kappa(x, x_\ell) \sigma_T(\rho(x_\ell)). \tag{6}$$

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

where $\kappa(x, x_\ell)$ is a kernel function [Bierens, 1988]. The ground state properties are then estimated using these predicted classical representations $\hat{\sigma}(x)$. Specifically, $f_O(x) = \operatorname{tr}(O\rho(x))$ can be predicted efficiently whenever O is a sum of few-body operators. They provide guarantees by using a truncated Fourier (also known as Dirichlet kernel) $\kappa(x, x_\ell) = \sum_{k\in\mathbb{Z}^m, \|k\|_2 \leq \Lambda} \cos(\pi k \cdot (x - x_\ell))$ with cutoff Λ. Their method guarantees that $\mathbb{E}_x \,|\operatorname{tr}(O\hat{\sigma}(x)) - f_O(x)|^2 \leq \epsilon$ for $N = m^{\mathcal{O}(1/\epsilon)}$.

3 Proposed Method

3.1 Idea

Suppose that O is an arbitrary and unknown n-qubit observable, and a distribution $\mathcal{D}$ of n-qubit quantum states. O will correspond to a local ground state property as described earlier. The distribution $\mathcal{D}$, in our case, would correspond to a shadow representation of the ground state of Hamiltonians $H(x)$. More specifically for each $x \in [-1, 1]^m$ we have $\sigma_T(\rho(x))$ in $\mathcal{D}$. The probability over $\mathcal{D}$ is the one naturally induced by this transformation on the distribution over x. Our goal is to find a function $h(\rho)$ which predicts the expectation value $\operatorname{tr}(O\rho)$ of the observable O on the state ρ with a small mean squared error:

$$\mathop{\mathbb{E}}_{\rho\sim\mathcal{D}} |h(\rho) - \operatorname{tr}(O\rho)|^2 \leq \epsilon.$$

We will assume that we can access training data of the form

$$\{\rho_\ell, \operatorname{tr}(O\rho_\ell)\}_{\ell=1}^N, \tag{7}$$

where ρ_ℓ is sampled from the distribution $\mathcal{D}$. In practice, though, we cannot directly access the exact value of the expectation value $\operatorname{tr}(O\rho_\ell)$; instead, we might measure O multiple times in the state ρ_ℓ to obtain an accurate estimate of the expectation value.

A critical aspect of the is that the distribution $\mathcal{D}$, has a specific structure which allows us to learn an efficient approximation to the entire process.

Specifically, the distribution of classical shadow representations $\mathcal{D}$ is *locally flat*. This means that the distribution is unmodified (i.e., the distribution appears flat) when we locally rotate any one of the qubits by a Clifford gate. To see this, recall that the Clifford group normalizes Pauli operators. Hence the composition of a Clifford gate C with a Pauli gate P is equivalent to composition with a different Clifford gate C' and Pauli gate P'. Since the classical shadow $\sigma_T(\rho)$ is obtained by applying Pauli operator P to the state ρ, applying a Clifford gate C to $\sigma_T(\rho)$, is equivalent to applying a different Clifford gate C' with a different Pauli operator P'. However since the classical shadow representation is obtained by applying a randomly chosen Pauli operator to the quantum state, the distribution of classical shadow representations is invariant to applying any Clifford gate, as applying a Clifford gate to the quantum state simply corresponds to applying a different Clifford gate and a different Pauli operator to the classical shadow representation.

An arbitrary observable O can be expanded in terms of the Pauli operator basis:

$$O = \sum_{P\in\{I,X,Y,Z\}^{\otimes n}} \alpha_P P. \tag{8}$$

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

Though there are 4^n Pauli operators, if the distribution $\mathcal{D}$ is locally flat and O has a constant spectral norm, we can approximate the sum over P by a truncated sum

$$O^{(k)} = \sum_{P \in \{I,X,Y,Z\}^{\otimes n} : |P| \leq k} \alpha_P P. \tag{9}$$

including only the Pauli operators P with weight $|P|$ up to k, those acting nontrivially on no more than k qubits. The mean squared error incurred by this truncation decays exponentially with k. Therefore, to learn O with mean squared error ϵ it suffices to learn this truncated approximation to O, where $k = \mathcal{O}(\log(1/\epsilon))$.

Furthermore, using recent work on Bohnenblust-Hille Inequalities [Slote et al., 2023, Volberg and Zhang, 2022] one can show that for O with low norms, only a few large coefficients α_P are relevant. Hence, instead of the complexity of $n^{\mathcal{O}(k)}$ in order to learn O one can get an efficient approximation with $\mathcal{O}(\log n)$ samples.

To make predictions about the expectation value of an operator O for a new quantum state ρ drawn from a distribution $\mathcal{D}$, we need to have some information about ρ. Both ρ and O can be represented using Pauli operators. When we only consider the truncated part of O, the prediction is based only on the corresponding part of ρ. If the reduced density matrices of the states in $\mathcal{D}$ are known, the prediction can be calculated classically. If the states in $\mathcal{D}$ are unknown, the reduced density matrices can be learned efficiently (for small k) using classical shadow tomography, after which the classical calculation can be performed to make a prediction about the expectation value of $\mathrm{tr}(O\rho)$.

3.2 Algorithm Details

Suppose we have obtained a classical dataset by performing N randomized experiments. Recall that a randomized Pauli measurement measures each qubit of a state in a random Pauli basis (X, Y or Z) and produces a measurement outcome of $|\psi^{(\mathrm{out})}\rangle = \bigotimes_{i=1}^n |s_i^{(\mathrm{out})}\rangle$, where $|s_i^{(\mathrm{out})}\rangle \in \mathrm{stab}_1 \triangleq \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |y+\rangle, |y-\rangle\}$. We denote the classical dataset of size N to be

$$S_N \triangleq \left\{ |\psi_\ell^{(\mathrm{in})}\rangle = \bigotimes_{i=1}^n |s_{\ell,i}^{(\mathrm{in})}\rangle,\ |\psi_\ell^{(\mathrm{out})}\rangle = \bigotimes_{i=1}^n |s_{\ell,i}^{(\mathrm{out})}\rangle \right\}_{\ell=1}^N, \tag{10}$$

where $|s_{\ell,i}^{(\mathrm{in})}\rangle, |s_{\ell,i}^{(\mathrm{out})}\rangle \in \mathrm{stab}_1$. Each product state is represented classically with $\mathcal{O}(n)$ bits. Hence, the classical dataset S_N is of size $\mathcal{O}(nN)$ bits.

Let O be an observable with $\|O\| \leq 1$ that is written as a sum of few-body observables, where each qubit is acted by $\mathcal{O}(1)$ of the few-body observables. We denote the Pauli representation of O as $\sum_{Q \in \{I,X,Y,Z\}^{\otimes n}} a_Q Q$. By definition of O, there are $\mathcal{O}(n)$ nonzero Pauli coefficients a_Q. We consider a hyperparameter $\tilde{\epsilon} > 0$; roughly speaking $\tilde{\epsilon}$ will scale inverse polynomially in the dataset size N. For every Pauli observable $P \in \{I,X,Y,Z\}^{\otimes n}$ with $|P| \leq k = \Theta(\log(1/\epsilon))$, the algorithm computes an empirical estimate for the corresponding Pauli coefficient α_P via

$$\hat{x}_P(O) = \frac{1}{N} \sum_{\ell=1}^N \mathrm{tr}\left(P \bigotimes_{i=1}^n |s_{\ell,i}^{(\mathrm{in})}\rangle\langle s_{\ell,i}^{(\mathrm{in})}| \right) \mathrm{tr}\left(O \bigotimes_{i=1}^n \left(3|s_{\ell,i}^{(\mathrm{out})}\rangle\langle s_{\ell,i}^{(\mathrm{out})}| - I \right) \right), \tag{11}$$

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

$$\hat{\alpha}_P(O) = \begin{cases} 3^{|P|}\hat{x}_P(O), & \left(\frac{1}{3}\right)^{|P|} > 2\tilde{\epsilon} \text{ and } |\hat{x}_P(O)| > 2 \cdot 3^{|P|/2}\sqrt{\tilde{\epsilon}}\sum_{Q:a_Q\neq 0}|a_Q|, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

The computation of $\hat{x}_P(O)$ and $\hat{\alpha}_P(O)$ can both be done classically. The basic idea of $\hat{\alpha}_P(O)$ is to set the coefficient $3^{|P|}\hat{x}_P(O)$ to zero when the influence of Pauli observable P is negligible. Given an n-qubit state ρ, the algorithm outputs

$$h(\rho, O) = \frac{1}{N}\sum_{\ell=1}^{N}\kappa(x, x_\ell)\sum_{P:|P|\leq k}\hat{\alpha}_P(O)\,\mathrm{tr}(P\rho(x_\ell)). \tag{13}$$

Note that, to make predictions, the ML algorithm only needs the k-body reduced density matrices (k-RDMs) of ρ. The k-RDMs of ρ can be efficiently obtained by performing randomized Pauli measurement on ρ and using the classical shadow formalism Huang et al. [2020].

Theorem 1 (Learning an unknown observable). *Given $\epsilon, \epsilon', \delta > 0$, $\|O\| < 1$ from a training data $\{\rho_\ell, \mathrm{tr}\,(O\rho_\ell)\}_{\ell=1}^{N}$ of size*

$$N = \log(n/\delta)2^{\mathcal{O}(\log(\frac{1}{\epsilon})\log(n))}, \tag{14}$$

where ρ_ℓ is sampled from $\mathcal{D}$, we can learn a function $h(\rho)$ such that

$$\mathop{\mathbb{E}}_{\rho\sim\mathcal{D}}|h(\rho) - \mathrm{tr}(O\rho)|^2 \leq (\epsilon + 2\epsilon') \tag{15}$$

with probability at least $1-\delta$.

The proof the the theorem and the detailed description of the ML algorithm are given in the longer version of the paper. To measure the prediction error of the ML model, we consider the average-case prediction performance under an arbitrary n-qubit state distribution $\mathcal{D}$ invariant under single-qubit Clifford gates, which means that the probability distribution $f_{\mathcal{D}}(\rho)$ of sampling a state ρ is equal to $f_{\mathcal{D}}(U\rho U^\dagger)$ of sampling $U\rho U^\dagger$ for any single-qubit Clifford gate U.

4 Conclusion

Traditional machine learning (ML) offers a strategy that has the potential to be extremely effective in resolving difficult quantum many-body issues in the fields of physics and chemistry. On the other hand, as these algorithms use an exponential amount of training data[Huang et al., 2021], it is not clear whether there is any guaranteed advantage of using ML based methods. In this study, we show that for certain restricted families ML algorithms do not need exponential amount of data. Our proposed method cuts down the required sample complexity, at the cost of considering only local Hamiltonians. We have focused purely on a theoretical analysis, and experiments need to be conducted to assess the efficiency of our approach against existing methods.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

Bibliography

D. Aharonov, W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM review*, 50(4):755–787, 2008.

A. Anshu, D. Gosset, K. J. M. Korol, and M. Soleimanifar. Improved approximation algorithms for bounded-degree local hamiltonians. *Physical Review Letters*, 127(25): 250502, 2021.

S. Arunachalam and R. de Wolf. Guest column: A survey of quantum learning theory. *ACM Sigact News*, 48(2):41–67, 2017.

B. Barak, A. Moitra, R. O'Donnell, P. Raghavendra, O. Regev, D. Steurer, L. Trevisan, A. Vijayaraghavan, D. Witmer, and J. Wright. Beating the random assignment on constraint satisfaction problems of bounded degree. *arXiv preprint arXiv:1505.03424*, 2015.

J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

H. J. Bierens. The nadaraya-watson kernel regression function estimator. 1988.

H. F. Bohnenblust and E. Hille. On the absolute convergence of dirichlet series. *Annals of Mathematics*, pages 600–622, 1931.

G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

D.-L. Deng, X. Li, and S. D. Sarma. Machine learning topological states. *Physical Review B*, 96(19):195145, 2017.

I. Dinur, E. Friedgut, G. Kindler, and R. O'Donnell. On the fourier tails of bounded functions over the discrete cube. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 437–446, 2006.

E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. *arXiv preprint quant-ph/0001106*, 2000.

S. T. Flammia and J. Preskill. Learning noisy quantum experiments. *arXiv preprint*, 2022.

J. Gibbs, M. Caro, A. Sornborger, and P. Coles. Quantum simulation with machine learning. *arXiv preprint*, 2022.

A. W. Harrow and A. Montanaro. Extremal eigenvalues of local hamiltonians. *Quantum*, 1:6, 2017.

H.-Y. Huang, R. Kueng, and J. Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill. Provably efficient machine learning for quantum many-body problems. 2021.

H. Kawai and Y. O. Nakagawa. Predicting excited states from ground state wavefunction by supervised quantum machine learning. *Machine Learning: Science and Technology*, 1(4):045027, 2020.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5

R. Levy, D. Luo, and B. K. Clark. Classical shadows for quantum process tomography on near-term quantum computers. *arXiv preprint arXiv:2110.02965*, 2021.
E. A. Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9 (1):141–142, 1964.
Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada. Restricted boltzmann machine learning for solving strongly correlated quantum systems. *Physical Review B*, 96 (20):205152, 2017.
M. Ohliger, V. Nesme, and J. Eisert. Efficient and feasible state tomography of quantum many-body systems. *New Journal of Physics*, 15(1):015024, 2013.
M. Paini and A. Kalev. An approximate description of quantum states. *arXiv preprint arXiv:1910.10543*, 2019.
C. Rouzé, M. Wirth, and H. Zhang. Quantum talagrand, kkl and friedgut's theorems and the learnability of quantum boolean functions. *arXiv preprint arXiv:2209.07279*, 2022.
M. Schuld and N. Killoran. Quantum machine learning in feature hilbert spaces. *Physical review letters*, 122(4):040504, 2019.
K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nature communications*, 10(1):5024, 2019.
A. J. Scott. Optimizing quantum process tomography with unitary 2-designs. *Journal of Physics A: Mathematical and Theoretical*, 41(5):055308, 2008.
O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua. Deep autoregressive models for the efficient variational simulation of many-body quantum systems. *Physical review letters*, 124(2):020503, 2020.
J. Slote, A. Volberg, and H. Zhang. Noncommutative bohnenblust-hille inequality in the heisenberg-weyl and gell-mann bases with applications to fast learning. *arXiv preprint arXiv:2301.01438*, 2023.
S. J. van Enk and C. W. Beenakker. Measuring tr ρ n on single copies of ρ using random measurements. *Physical review letters*, 108(11):110503, 2012.
R. A. Vargas-Hernández, J. Sous, M. Berciu, and R. V. Krems. Extrapolating quantum observables with machine learning: Inferring multiple phase transitions from properties of a single phase. *Physical review letters*, 121(25):255702, 2018.
A. Volberg and H. Zhang. Noncommutative bohnenblust–hille inequalities. *arXiv preprint arXiv:2210.14468*, 2022.
K. Wan and I. H. Kim. Fast digital methods for adiabatic state preparation. *arXiv preprint arXiv:2004.04164*, 2020.
G. S. Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.
Y. Zhang, R. G. Melko, and E.-A. Kim. Machine learning z 2 quantum spin liquids with quasiparticle statistics. *Physical Review B*, 96(24):245119, 2017.
Y. Zhang, P. Ginsparg, and E.-A. Kim. Interpreting machine learning of topological quantum phase transitions. *Physical Review Research*, 2(2):023283, 2020.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36030-5_5