Efficient learning of commuting hamiltonians on lattices

Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, Mehdi Soleimanifar

March 23, 2021

Background: In a recent work [AAKS20] we constructed an algorithm to learn the hamiltonian from a Gibbs state at any constant temperature. The algorithm is sample-efficient (polynomially tight) when the learning is required for small ℓ_2 error. It is time-efficient above critical temperatures and for stoquastic hamiltonians.

In this note, we consider the Gibbs state of a commuting hamiltonian and provide an algorithm that is both sample-efficient and time-efficient at any constant temperature (and works for small ℓ_∞ error).

TL;DR: Effective hamiltonian of the reduced state of a 'commuting Gibbs state' is also local. Thus, learning can be performed locally.

0.1 Notation and effective hamiltonian

Fix a D-dimensional lattice and let each spin have dimension d. Consider a k-local hamiltonian

$$H = \sum_{\ell=1}^{m} h_\ell \tag{1}$$

with $\|h_\ell\| \leq 1$ ($\forall \ell$, where $\|.\|$ denotes the ℓ_∞ norm) and assume that $[h_\ell, h_{\ell'}] = 0$ ($\forall \ell, \ell'$). Let h_R denote the hamiltonian restricted to a region R. Let $\rho_\beta = \frac{e^{-\beta H}}{\mathrm{Tr}\left(e^{-\beta H}\right)}$ be the Gibbs state. For any region R on the lattice, define the effective hamiltonian $H_R = \frac{-1}{\beta} \log \mathrm{Tr}_{R^c}\left(\rho_\beta\right)$. Let ∂R be the boundary of R, and $\partial_- R$ be the inner boundary of R. The following lemma says that the effective Hamiltonian is local. It is not known to hold in the general case, except above critical temperatures [KKBa20, Theorem 2].

Lemma 1. *It holds that*

$$H_R = \alpha_R I + h_R + \Phi,$$

where Φ is only supported on $\partial_- R$ and $[\Phi, h_R] = 0$. Here, α_R is some real number and $\|\Phi\| \leq 2|\partial R|$.

Proof. We can write $H = h_R + h_{\partial R} + h_{R^c}$. Consider

$$\mathrm{Tr}_{R^c}\left(e^{-\beta H}\right) = e^{-\beta h_R} \mathrm{Tr}_{R^c}\left(e^{-\beta(h_{\partial R} + h_{R^c})}\right).$$

Define $e^{-\beta\Phi} := \mathrm{Tr}_{R^c}\left(e^{-\beta(h_{\partial R}+h_{R^c})}\right)$. It is clear that $[\Phi, h_R] = 0$ and hence H_R has the form as stated in the lemma. In order to bound the norm of Φ, we proceed as follows. Consider,

$$h_R + h_{R^c} - |\partial R| I \preceq H \preceq h_R + h_{R^c} + |\partial R| I.$$

Since every term commutes, we can exponential the Lowner inequality to obtain

$$e^{-\beta|\partial R|} e^{-\beta h_R} \otimes e^{-\beta h_{R^c}} \preceq e^{-\beta H} \preceq e^{\beta|\partial R|} \otimes e^{-\beta h_R} e^{-\beta h_{R^c}}.$$

Tracing out the region R^c, this means that

$$e^{-\beta|\partial R|}\mathrm{Tr}\left(e^{-\beta h_{R^c}}\right)e^{-\beta h_R}\preceq \mathrm{Tr}_{R^c}\left(e^{-\beta H}\right)\preceq e^{\beta|\partial R|}\mathrm{Tr}\left(e^{-\beta h_{R^c}}\right)e^{-\beta h_R}.$$

Thus, the ratio between largest and smallest eigenvalues of $e^{\beta h_R}\mathrm{Tr}_{R^c}\left(e^{-\beta H}\right)=e^{-\beta\Phi}$ is upper bounded by $e^{2\beta|\partial R|}$. This completes the proof. □

The above lemma ensures the following identity

$$\mathrm{Tr}_{R^c}\left(\rho_\beta\right)=\frac{e^{-\beta(h_R+\Phi)}}{\mathrm{Tr}\left(e^{-\beta(h_R+\Phi)}\right)}.$$

0.2 Learning algorithm

For every ℓ, let R_ℓ be the smallest region that contains $\mathrm{supp}\left(h_\ell\right)$ in its strict interior (that is, it does not overlap with $\partial_- R_\ell$). We have $|R_\ell|\leq(3k)^D$. Then $\mathrm{Tr}_{R_\ell^c}\left(\rho_\beta\right)=\frac{e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}}{\mathrm{Tr}\left(e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}\right)}$, where Φ_ℓ is only supported $\partial_- R_\ell$. Since $|\Phi_\ell|\leq 2|\partial R_\ell|$, the smallest eigenvalue of $\frac{e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}}{\mathrm{Tr}\left(e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}\right)}$ is at least

$$\frac{e^{-\beta(|R_\ell|+|\partial R_\ell|)}}{d^{|R_\ell|}}\geq e^{-(\beta+\log d)(3k)^D}.$$

The algorithm is as follows. We divide $\{R_\ell\}_{\ell=1}^m$ into different batches, such that within each batch the R_ℓ's don't overlap. Number of batches needed is $(kD)^D$ (a constant). Within each batch, we perform tomography to obtain the classical description of $\frac{e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}}{\mathrm{Tr}\left(e^{-\beta\left(h_{R_\ell}+\Phi_\ell\right)}\right)}$ upto an error of $\epsilon e^{-(\beta+\log d)(3k)^D}$. This gives us a classical description of an operator h' satisfying $\|h'-h_{R_\ell}-\Phi_\ell\|\leq\epsilon$. From this, h_ℓ can be computed by evaluating

$$h'_\ell:=\frac{1}{d^{|R_\ell|}}\sum_{j=1}^{d^{2k}}\sigma_\ell^{(j)}\mathrm{Tr}\left(\sigma_\ell^{(j)}h'\right),$$

where $\{\sigma_\ell^{(j)}\}_{j=1}^{d^{2k}}$ are the Pauli operators in the support of h_ℓ. It can be seen that

$$\|h'_\ell-h_\ell\|\leq d^{2k}\epsilon.$$

In order to perform the tomography in each batch with probability of success $1-\frac{\delta}{\text{number of batches}}$, the number of samples needed is [CW20, BMBO20, HKP20]

$$\frac{e^{2(\beta+\log d)(3k)^D}}{\epsilon^2}\log\left(m\frac{\text{number of batches}}{\delta}\right)\leq\frac{e^{2(\beta+\log d)(3k)^D}}{\epsilon^2}\log\frac{m\,(kD)^D}{\delta}.$$

Thus, setting $d=\mathcal{O}(1)$, total sample complexity is (accounting for all the batches)

$$\frac{e^{\mathcal{O}\left(\beta k^D\right)}}{\epsilon^2}\log\frac{m}{\delta}.$$

Time complexity is roughly

$$m\cdot\frac{e^{\mathcal{O}\left(\beta k^D\right)}}{\epsilon^2}\log\frac{m}{\delta},$$

as the time complexity for processing the data from each sample is roughly $m\cdot e^{\mathcal{O}(\beta k^D)}$.

References

[AAKS20] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar. Sample-efficient learning of quantum many-body systems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 685–691, 2020.

[BMBO20] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E. O'Brien. Nearly optimal measurement scheduling for partial tomography of quantum states. *Phys. Rev. X*, 10:031064, Sep 2020.

[CW20] Jordan Cotler and Frank Wilczek. Quantum overlapping tomography. *Physical Review Letters*, 124(10):100401, 2020.

[HKP20] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, Oct 2020.

[KKBa20] Tomotaka Kuwahara, Kohtaro Kato, and Fernando G. S. L. Brandão. Clustering of conditional mutual information for quantum gibbs states above a threshold temperature. *Physical Review Letters*, 124:220601, 2020.