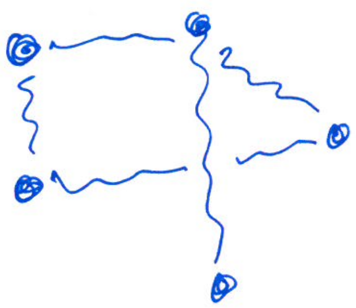
Learning a (very) quantum local Hamiltonian from local measurements.

Joint work w. Eyal Bairy, Netanel Lindner } Technion

arXiv: 1807.04564

Setup:

Quantum Hamiltonians:



$$H = \sum_{\langle i,j \rangle} h_{ij}$$

h_{ij} - Interaction energy between qudits i,j.

Eigenstates $|\phi_0\rangle, |\phi_1\rangle, \ldots$

Eigenvalues: $\epsilon_0 \le \epsilon_1 \le \ldots$

H is relevant mainly in two setups:

(i) Dynamics: $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$

(ii) Theremodynamics $\rho_T = \frac{1}{Z_T} e^{-\frac{1}{T}H}$ $\qquad Z_T = Tr(e^{-\frac{1}{T}H})$

Main Problem:

Learn H from local measurements. $\boxed{H}$

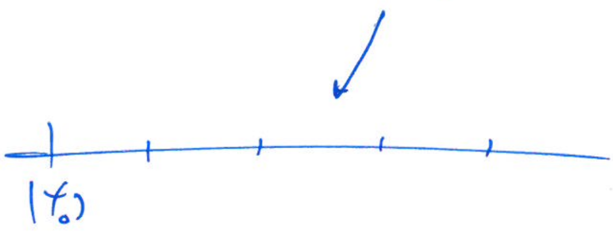
(*) Experiments

(*) Verification of quantum Hardware.

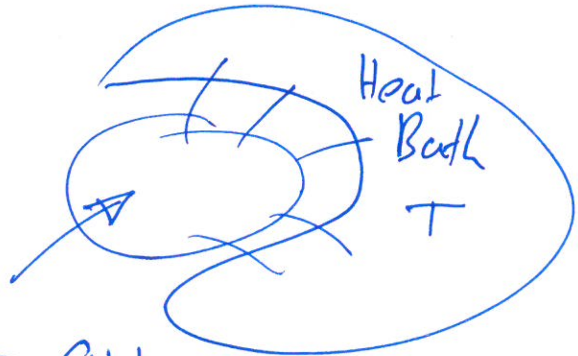
> Sometimes the quantum problem may turn out to be easier than the classical Problem.

Two setups:

(–) $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$



(–)



← Sampling from ρ.

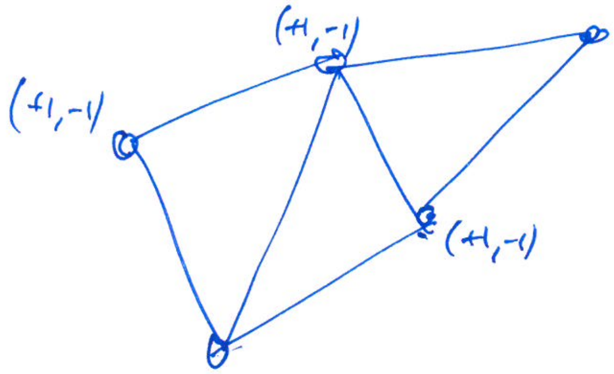
$\frac{1}{Z}e^{-\frac{1}{T}H}$ ← Gibbs state.

Note: when $T \to 0$, $\rho \to |\phi_0\rangle\langle\phi_0|$

because: $\rho = \frac{1}{Z}\left[e^{-\frac{1}{T}\epsilon_0} |\phi_0\rangle\langle\phi_0| + e^{-\frac{1}{T}\epsilon_1} |\phi_1\rangle\langle\phi_1| + \dots \right]$

Classical Problem:

Learning of Boltzmann Machines
(special case of Graphical models,
AKA Markov Random Fields (MRF))



$G = (V, E)$

$$P_\omega(z_1, \dots, z_n) \equiv \frac{1}{Z(\omega)} e^{\overbrace{\sum_{ij} \omega_{ij} z_i z_j + \sum_i \theta_i z_i}^{H_{Ising}(\vec{z})}}$$

$$\equiv \frac{1}{Z(\omega, \theta)} e^{H_{Ising}(\vec{z})} = \Phi_1(z_1, z_2) \cdot \Phi_2(z_2, z_4) \cdot \dots$$

$\eta \equiv \min \{ |\omega_{ij}|, |\theta_i| \}$

$\lambda \equiv \max_i \left\{ \sum_{j \in N(i)} |\omega_{ij}| + |\theta_i| \right\}$

d - maximal local degree of the graph.

Can we learn H (and the underlying graph $G=(V,E)$) from N samples?

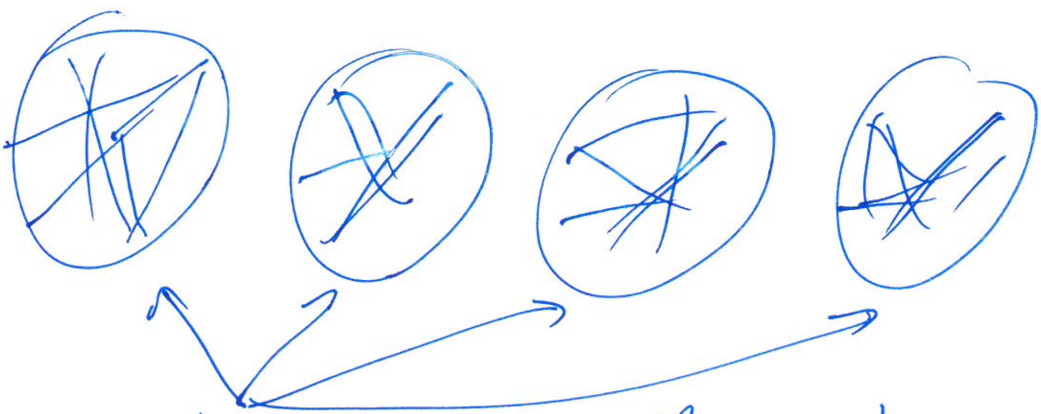
How many samples are needed?

<u>Information – theoretic [Lower-bound]</u>

(Santharam & Wainwright 2009)

To learn the underlying Graph, one needs at least exponential.

$$N = \frac{e^{(\lambda/4)} d\eta \log(\frac{nd}{4}-1)}{128\, e^{\lambda \frac{3\eta}{2}}}$$

$G =$ 

$d+1$-cliques. η = const.

Family of graphs G_{rs} → remove (r,s) from G.

Upper bounds

Algorithm by Klivans & Meka (2017)
arXiv: 1706.06274

A multiplicative-update based alg that is able to recover ω, θ and reconstruct G

with

$$N = O\left(\frac{1}{\eta^4} e^{O(\lambda)} \cdot \log\left(\frac{n}{\rho\eta}\right)\right)$$

with failure prob $\rho > 0$.

$$\text{Prob}(z_1 \ldots z_n) = \frac{1}{Z} e^{\sum_{ij} \omega_{ij} z_i z_j + \sum_i \theta_i z_i}$$

To learn $\{\omega_{ij}, \theta_i\}$ we can focus on specific i

$$Y \equiv \frac{1-z_i}{2} \in \{0,1\} \quad \vec{X} \equiv (z_j \mid j \neq i)$$

$$\Rightarrow \text{Prob}(Y=1 \mid \vec{X}) = \mathbb{E}(Y \mid \vec{X} = \vec{x}) = \sigma(\vec{\omega} \cdot \vec{x} + \theta)$$

where $\sigma(s) \equiv \dfrac{1}{1+e^{-s}}$ - sigmoid function.

$\theta = -\theta_i \qquad \vec{\omega} = \{\omega_{ij} \mid j \neq i\}$

$\Rightarrow$ Learning w_{ij}, θ_i can be done if we learn $\vec{\omega}, \vec{\theta}$ by sampling random variables $(Y, \vec{X})$ that dist' according to $\underline{\mathbb{E}(Y \mid \vec{X} = \vec{x}) = \sigma(\vec{\omega} \cdot \vec{x} + \theta)}$

<u>Going to the quantum world.</u>

$$H_{classic}(z_1 \ldots z_n) = \sum_{ij} \omega_{ij} z_i z_j + \sum_i \theta_i z_i$$

$$= \langle z_1 \ldots z_n | \underbrace{\sum_{ij} \omega_{ij} Z_i Z_j + \sum_i \theta_i Z_i}_{H_{classic}} | z_1 \ldots z_n \rangle$$

$$\mathrm{Prob}(z_1 \ldots z_n) = \langle z_1 \ldots z_n | \underbrace{\tfrac{1}{z} e^{H_{classi}}}_{\rho} | z_1 \ldots z_n \rangle$$

$$H_Q = \sum_{ij} \sum_{\alpha\beta} \omega_{ij}^{\alpha\beta} \sigma_i^\alpha \otimes \sigma_j^\beta + \sum_i \sum_\alpha \theta_i^\alpha \sigma_i^\alpha$$

$\sigma^\alpha = \{X, Y, Z\} \leftarrow$ Pauli Matrices.

$$\rho = \frac{1}{Z} e^{H_Q}$$

We can sample ρ in different bases $\to \{|0\rangle, |1\rangle\}$
$\{|+\rangle, |-\rangle\}$

etc..

For example: $\text{Prob}(0, 0, +, -) =$

$$= \langle 0|\langle 0|\langle +|\langle -| \rho |0\rangle|0\rangle|+\rangle|-\rangle$$

Can we use these local measurements to learn $\{\omega_{ij}^{\alpha\beta}, \theta_i^\alpha\}$?

No longer have the closed form

$\text{Prob} = \sigma(\ldots)$ ...

Our Approach

$$H_Q = \sum_{ij} \sum_{\alpha\beta} \omega_{ij}^{\alpha\beta} \sigma_i^\alpha \otimes \sigma_j^\beta + \sum_i \sum_\alpha \theta_i^\alpha \sigma_i$$

$$M = O(n^2) \qquad = \sum_{m=1}^{M} \omega_m S_m$$

$$S_m \in \{\sigma_i^\alpha , \sigma_i^\alpha \otimes \sigma_j^\beta\}$$

Let A be an observable and ρ a state s.t. $[H, \rho] = 0$.

Then : $0 = \mathrm{Tr}(A[H,\rho]) = \mathrm{Tr}(\rho[A,H]) =$

$$= \langle [A,H] \rangle = \sum_{m=1}^{M} \omega_m \langle [A, S_m] \rangle = 0$$

$\Rightarrow$ Pick $L > M$ observables (Local)

$A_1, A_2, \ldots, A_L$

$$\forall A_l : \quad \sum_{m=1}^{M} \omega_m \underbrace{\langle [A_l, S_m] \rangle}_{K_{lm}} = 0$$

$$\boxed{K\omega = 0.}$$

$$K = \begin{array}{c} \xleftarrow{\quad M \quad} \\ L \updownarrow \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{array}$$

When $L > M$, we expect the equation $K\omega = 0$ to have a unique solution.

<u>Alg:</u>

① Generate enough samples to estimate $K_{\ell m} = \langle [A_\ell, S_m] \rangle$ to high accuracy.

② Invert $K\omega = 0$.

<u>Analysis:</u>

Define $T = K^{\dagger} \cdot K = M \times M$ PSD.

It has eigenvalues: $\lambda_0 = 0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$
eigenstates: $|c_0\rangle, |c_1\rangle, |c_2\rangle, \ldots$

If we estimate K using N samples then what we actually have is

$$K' = \frac{1}{N} \sum_{s=1}^{N} K^{(s)}$$

entries are single measurement results.

$$K' = K + E$$

$$\Rightarrow T' = K'^{\dagger} K' = (K+E)^{\dagger}(K+E)$$

$$= K^{\dagger}K + \underbrace{K^{\dagger}E + E^{\dagger}K + E^{\dagger}E}$$

$$\equiv T + V \leftarrow \text{Perturbation.}$$

$$\Rightarrow |c_0'\rangle - |c_0\rangle = \sum_{k=1}^{\infty} \left(\sum_{i>0} \frac{|c_i\rangle\langle c_i|}{\lambda_i - \lambda_0} V \right)^k |c_0\rangle$$

Thm: (Matrix chernoff bound)

To get $\frac{\| |c_0'\rangle - |c_0\rangle \|}{\| c_0 \|} \leq \delta$ with Prob $> \frac{1}{2}$

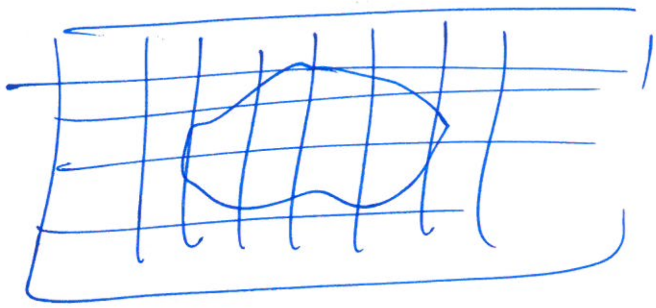
it is enough to use

$$N = O\left(\frac{M^2 L \log(L)}{\lambda_1 \delta^2} \right) \text{ samples.}$$

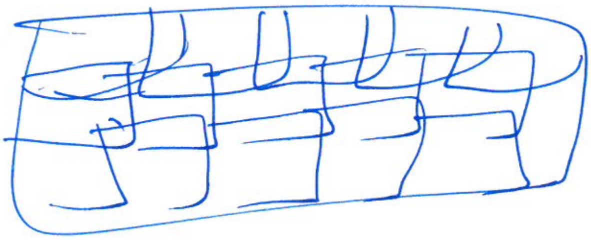
$\Rightarrow$ If the spectral gap $\lambda_1 > 0$ is $\Omega(\mathrm{Poly}(n))$, we have a polynomial complexity.

Further comments.

(*) If H is on a grid w. n.n. interactions

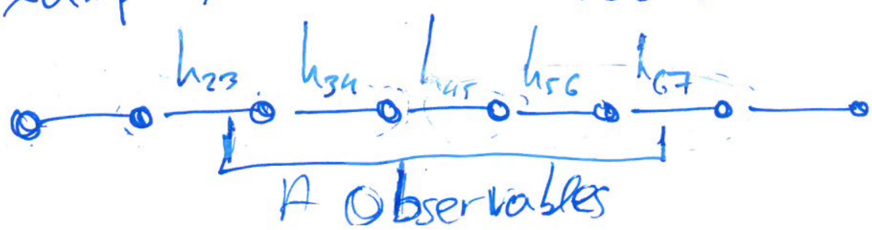


Then if we have a patch L we can recover $H|_L$ only using measurements of observables in L. → Linear time reconstruction



(*) We can always take the observables A_ℓ to be local, so $[A_\ell, S_m]$ is also local.

(*) For example, on a chain:



If we consider all possible observables A

then: $$K_{\ell m} = \mathrm{Tr}(\rho [A_\ell, S_m])$$

Take $\{A_\ell\}$ to be an orthonormal basis.
$$= \mathrm{Tr}(A_\ell [S_m, \rho])$$

$$\Rightarrow [S_m, \rho] = \sum_\ell K_{\ell m} A_\ell$$

$$\Rightarrow T_{mm'} = \sum_\ell (K^\dagger)_{m\ell} K_{\ell m'} =$$
$$= \mathrm{Tr}\left([S_m, \rho]^\dagger [S_{m'}, \rho]\right)$$

If $\rho = |\epsilon\rangle\langle\epsilon|$

$$[S_m, \rho] = S_m|\epsilon\rangle\langle\epsilon| - |\epsilon\rangle\langle\epsilon|S_m$$

$$\Rightarrow \mathrm{Tr}\left([S_m, \rho]^\dagger [S_{m'}, \rho]\right) =$$

$$\mathrm{Tr}\left((|\epsilon\rangle\langle\epsilon|S_m - S_m|\epsilon\rangle\langle\epsilon|)(S_{m'}|\epsilon\rangle\langle\epsilon| - |\epsilon\rangle\langle\epsilon|S_{m'})\right)$$

$$= 2\left[\langle\epsilon|S_m S_{m'}|\epsilon\rangle - \langle\epsilon|S_m|\epsilon\rangle\langle\epsilon|S_{m'}|\epsilon\rangle\right]$$

$$\Rightarrow H = \sum \omega_m S_m$$
$$\Rightarrow \langle\epsilon|H^2|\epsilon\rangle - (\langle\epsilon|H|\epsilon\rangle)^2 = 0.$$

This was the starting point of

- Qi and Ranard (2017)
- Chertkov and Clark (2018)
- Greiter, Schnells and Thomale (2018)

We can also do dynamics:

$$\rho(t) = e^{-iHt}\rho(0)\,e^{iHt}$$

$$\frac{d}{dt}\rho(t) = -i[H,\rho(t)]$$

$$\Rightarrow \rho_{avg}(T) \equiv \frac{1}{T}\int_0^T \rho(t)\,dt$$

$$\Rightarrow [\rho_{avg}, H] = \frac{1}{T}\int_0^T [\rho(t), H]\,dt =$$

$$= i\,\frac{1}{T}\int_0^T \frac{d}{dt}\rho(t)\,dt = \frac{i}{T}\left[\rho(T)-\rho(0)\right] \to 0$$