Phase Transitions in Quantum Spin Systems

Ankur Moitra (MIT)

UC Berkeley, February 26th

based on joint work with Ainesh Bakshi, Allen Liu and Ewin Tang

How do local interactions give rise to macroscopic behavior?

How do local interactions give rise to macroscopic behavior?

Running example: Consider **magnetism** – caused by most electron spins being oriented in same direction



How do local interactions give rise to macroscopic behavior?

Running example: Consider **magnetism** – caused by most electron spins being oriented in same direction



How do they know which way to point, if they don't all interact?

In **ferromagnets**, neighboring spin spins prefer to point in the same direction



In **ferromagnets**, neighboring spin spins prefer to point in the same direction



Statistical physics: Define a probabilistic model that captures behavior at thermal equilibrium and study its structural properties

CLASSICAL SPIN SYSTEMS

Definition: an Ising model is a distribution on $\{\pm 1\}^n$ with

inverse temperature Hamiltonian $\mathbb{P}[X = x] = \frac{1}{Z} \exp\left(-\beta H(x)\right)$ partition function

where
$$H(x) = -\sum_{i,j} J_{i,j} x_i x_j - \sum_i h_i x_i$$

CLASSICAL SPIN SYSTEMS

Definition: an **Ising model** is a distribution on $\{\pm 1\}^n$ with

inverse temperature Hamiltonian $\mathbb{P}[X = x] = \frac{1}{Z} \exp\left(-\beta H(x)\right)$ partition function where $H(x) = -\sum_{i,j} J_{i,j} x_i x_j - \sum_i h_i x_i$

Can generalize to higher degree polynomials etc

CONDITIONAL INDEPENDENCE

Often helpful to look at their graph structure:

 $G = (\{X_1, \cdots, X_n\}, E)$ with $E = \{(X_i, X_j) \text{ s.t. } J_{i,j} \neq 0\}$

CONDITIONAL INDEPENDENCE

Often helpful to look at their graph structure:

 $G = (\{X_1, \cdots, X_n\}, E)$ with $E = \{(X_i, X_j) \text{ s.t. } J_{i,j} \neq 0\}$

Markov Property: Two nodes are independent when conditioned on a separator – i.e.

$$X_i \perp X_j | X_U |$$

provided that all paths from X_i to X_j pass through X_U

CONDITIONAL INDEPENDENCE

Often helpful to look at their graph structure:

 $G = (\{X_1, \cdots, X_n\}, E)$ with $E = \{(X_i, X_j) \text{ s.t. } J_{i,j} \neq 0\}$

Markov Property: Two nodes are independent when conditioned on a separator – i.e.

$$X_i \perp X_j | X_U |$$

provided that all paths from X_i to X_j pass through X_U

Caution: The Markov property fails in quantum spin systems

PHASE TRANSITIONS

Dramatic changes in macroscopic properties as temperature varies

e.g. average magnetization







decreasing temperature

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

How do we describe quantum systems at thermal equilibrium?

Definition: The **density matrix** ρ of an n-qubit system is a 2ⁿ x 2ⁿ matrix that is positive semidefinite and trace one

Definition: The **density matrix** ρ of an n-qubit system is a 2ⁿ x 2ⁿ matrix that is positive semidefinite and trace one

Think of it as a generalization of classical distributions for which

$$\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ & \ddots \end{bmatrix}$$

i.e. entries along the diagonal are probabilities of each of the 2^n possible configurations

We will also be interested in operators that act on at most k qubits



We will also be interested in operators that act on at most k qubits



More generally a 2^k x 2^k matrix and a list of qubits, called its support, that it acts on

We will also be interested in operators that act on at most k qubits



More generally a 2^k x 2^k matrix and a list of qubits, called its support, that it acts on

Fhink of it as a generalization of local interactions in a graphical model

How do we describe quantum systems at thermal equilibrium?

How do we describe quantum systems at thermal equilibrium?

Definition: The **quantum Gibbs state** is

$$\rho = \frac{1}{Z} \exp\left(-\beta H\right)$$



How do we describe quantum systems at thermal equilibrium?

Definition: The **quantum Gibbs state** is

$$\rho = \frac{1}{Z} \exp\left(-\beta H\right)$$



By construction, ρ is a density matrix

How do we describe quantum systems at thermal equilibrium?

Definition: The **quantum Gibbs state** is

$$\rho = \frac{1}{Z} \exp\left(-\beta H\right)$$



By construction, ρ is a density matrix

How does temperature affect properties of a quantum system?

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

Challenging to measure entanglement, but easier to say when a state is unentangled

Challenging to measure entanglement, but easier to say when a state is unentangled

Definition: A **product state** satisfies

$$\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$$

where each ρ_i is a 2 x 2 density matrix.

Challenging to measure entanglement, but easier to say when a state is unentangled

Definition: A **product state** satisfies

$$\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$$

where each ρ_i is a 2 x 2 density matrix. A separable state is a mixture of product states

Challenging to measure entanglement, but easier to say when a state is unentangled

Definition: A **product state** satisfies

$$\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$$

where each ρ_i is a 2 x 2 density matrix. A separable state is a mixture of product states

Think of separable states as ones that have only classical correlations

In the limit as $\beta \rightarrow 0$ (i.e. infinite temperature) we have



In the limit as $\beta \rightarrow 0$ (i.e. infinite temperature) we have



Hence the Gibbs state is separable

In the limit as $\beta \rightarrow 0$ (i.e. infinite temperature) we have



Hence the Gibbs state is separable

But this is a non-physical temperature, since it grows with the system size

In the limit as $\beta \rightarrow 0$ (i.e. infinite temperature) we have



Hence the Gibbs state is separable

But this is a non-physical temperature, since it grows with the system size

What happens at reasonable physical temperatures?



We show that heat kills all entanglement

OUR RESULTS

We show that heat kills all entanglement

Theorem [Bakshi, Liu, Moitra, Tang]: There is a constant c > 0 so that for any $\beta \le \frac{c}{dk^2}$ the Gibbs state is separable

Here k is the locality and d is the graph degree
OUR RESULTS

We show that heat kills all entanglement

Theorem [Bakshi, Liu, Moitra, Tang]: There is a constant c > 0 so that for any $\beta \le \frac{c}{dk^2}$ the Gibbs state is separable

Here k is the locality and d is the graph degree

Moreover there is an efficient randomized algorithm that outputs the description of a product state that works under similar parameters

Major goal is to give quantum advantage for useful problems

Major goal is to give quantum advantage for useful problems, e.g.

Show there are efficient quantum Gibbs samplers that succeed where classical algorithms do not

Major goal is to give quantum advantage for useful problems, e.g.

Show there are efficient quantum Gibbs samplers that succeed where classical algorithms do not

Would be applications in **quantum chemistry** for understanding material properties

But Gibbs sampling becomes hard after a point

But Gibbs sampling becomes hard after a point

Theorem [Sly] [Sly, Sun]: Even classical Gibbs sampling is NP-hard for k = 2 and $\beta \ge \frac{c}{d}$ for some constant c'

But Gibbs sampling becomes hard after a point

Theorem [Sly] [Sly, Sun]: Even classical Gibbs sampling is NP-hard for k = 2 and $\beta \ge \frac{c}{d}$ for some constant c'

And so the region you could hope for quantum advantage is now quite narrow for constant locality



INDEPENDENT WORK

Theorem [Rouze, Franca, Alhambra]: Polynomial mixing time bounds for quantum Gibbs sampling at high enough temperature

INDEPENDENT WORK

Theorem [Rouze, Franca, Alhambra]: Polynomial mixing time bounds for quantum Gibbs sampling at high enough temperature

Better dependence on locality (k vs k^2), but worse dependence on degree (d vs $d^{O(1)}$)

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

When are perturbations of the identity still separable?

When are perturbations of the identity still separable?

Fact: Any density matrix on n qubits of the form

$$\rho = \frac{I + cE}{2^n}$$
 with ${\rm Tr}(E) = 0$, $\|E\| \le 1$ and $c \le \frac{1}{2^n}$ is separable

When are perturbations of the identity still separable?

Fact: Any density matrix on n qubits of the form

$$\rho = \frac{I + cE}{2^n}$$
 with $\mathrm{Tr}(E) = 0$, $\|E\| \leq 1$ and $c \leq \frac{1}{2^n}$ is separable

This is too weak for our purposes, but there is a sharpening that points the way forward

When are perturbations of the identity still separable?

Fact': Any density matrix on n qubits of the form



When are perturbations of the identity still separable?

Fact': Any density matrix on n qubits of the form

$$\rho = \frac{I+cE}{2^n}$$
 with ${\rm Tr}(E) = 0$, $\|E\| \le 1$ and $c \le \frac{1}{2^n}$ is separable if E has support k $c \le \frac{1}{2^k}$

Can we approximate the Gibbs state by local perturbations of the identity?

This is not so straight forward

this is local, by assumption



but what about this??

This is not so straight forward

this is local, by assumption



but what about this??

Need carefully designed expansions, let's take a detour to explain where they come from

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

CLASSICAL SAMPLING

How do you sample from an Ising model?

Natural Markov chain with local updates

Natural Markov chain with local updates

choose node at random, forget its state



Natural Markov chain with local updates

choose node at random, forget its state



Natural Markov chain with local updates



Natural Markov chain with local updates



Its unique steady state distribution is the Gibbs distribution

A classical thought experiment



A classical thought experiment



A classical thought experiment



A classical thought experiment



A classical thought experiment



First update x_i then update x_i

Claim: At high temperature, there is a good chance that their updates can be made independently, in which case we can ignore their edge

Try to make x_i independent of the rest, but keep only edges that are needed as correction terms



Try to make x_i independent of the rest, but keep only edges that are needed as correction terms



And you end up with small connected components – this is where locality comes from

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

THE PROOF STRATEGY

Definition: We say that a state is **1-vs-all separable** if it can be written as a convex combination of

 $ho_1 \otimes
ho_{[n] \setminus 1}$
THE PROOF STRATEGY

Definition: We say that a state is **1-vs-all separable** if it can be written as a convex combination of

 $\rho_1 \otimes \rho_{[n]\setminus 1}$

i.e. there is no entanglement between first and rest of qubits

THE PROOF STRATEGY

Definition: We say that a state is **1-vs-all separable** if it can be written as a convex combination of

 $\rho_1 \otimes \rho_{[n]\setminus 1}$

i.e. there is no entanglement between first and rest of qubits

Want to show that the Gibbs state is **1-vs-all separable** and proceed by induction

THE PROOF STRATEGY

Definition: We say that a state is **1-vs-all separable** if it can be written as a convex combination of

 $\rho_1 \otimes \rho_{[n] \setminus 1}$

i.e. there is no entanglement between first and rest of qubits

Want to show that the Gibbs state is **1-vs-all separable** and proceed by induction

Key is to define an **extraction operator** and write its expansion with **exponentially decaying coefficients**

THE EXTRACTION OPERATOR

As before, want to remove interactions between one qubit and all the others...

THE EXTRACTION OPERATOR

As before, want to remove interactions between one qubit and all the others. Consider

$$H_1 = \sum_{e \ni 1} H_e$$

THE EXTRACTION OPERATOR

As before, want to remove interactions between one qubit and all the others. Consider

$$H_1 = \sum_{e \ni 1} H_e$$

Definition: The extraction operator is

$$M_1 = e^{\beta (H - H_1)/2} e^{-\beta H/2}$$

THE KEY LEMMA

Lemma: We can write

$$M_{1} = \sum_{k=0}^{\infty} \frac{(\beta/2)^{k}}{k!} f_{k}(H, H_{1}) \text{ where } f_{k}(H, H_{1}) = \sum_{a \in \mathcal{S}} c_{a} E_{a}$$

and...

THE KEY LEMMA

Lemma: We can write

$$M_{1} = \sum_{k=0}^{\infty} \frac{(\beta/2)^{k}}{k!} f_{k}(H, H_{1}) \text{ where } f_{k}(H, H_{1}) = \sum_{a \in \mathcal{S}} c_{a} E_{a}$$

and each $||E_a|| = 1$ and is supported on a connected component of size at most k+1 and contains site 1. Moreover

$$\sum_{a \in \mathcal{S}} |c_a| \le (10d)^k k!$$

THE KEY LEMMA

Lemma: We can write

$$M_{1} = \sum_{k=0}^{\infty} \frac{(\beta/2)^{k}}{k!} f_{k}(H, H_{1}) \text{ where } f_{k}(H, H_{1}) = \sum_{a \in \mathcal{S}} c_{a} E_{a}$$

and each $||E_a|| = 1$ and is supported on a connected component of size at most k+1 and contains site 1. Moreover

$$\sum_{a \in \mathcal{S}} |c_a| \le (10d)^k k!$$

The point is, when $\beta << 1/d$ the terms are exponentially decaying, and can get a handle on entanglement via Fact'

Corollary: Hence we know that $M_1M_1^*$ is separable

Corollary: Hence we know that $M_1M_1^*$ is separable

The key identity is

$$e^{-\beta H} = e^{-\beta (H - H_1)/2} M_1 M_1^* e^{-\beta (H - H_1)/2}$$

Corollary: Hence we know that $M_1M_1^*$ is separable

The key identity is

$$e^{-\beta H} = e^{-\beta (H - H_1)/2} M_1 M_1^* e^{-\beta (H - H_1)/2}$$

Acts on 1 and [n]\1 separately, so preserves separability

Corollary: Hence we know that $M_1M_1^*$ is separable

The key identity is

$$e^{-\beta H} = e^{-\beta (H - H_1)/2} M_1 M_1^* e^{-\beta (H - H_1)/2}$$

Acts on 1 and [n]\1 separately, so preserves separability

Induction is involved because $M_1M_1^*$ is not literally the identity, need a careful potential function argument

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

OUTLINE

Part I: Introduction

- Classical Spin Systems and Phase Transitions
- Density Matrices, Operators and Quantum Spin Systems
- Entanglement and Our Results

Part II: Showing Separability

• Perturbations of the Identity

Intermission: Glauber Dynamics

- Proof Strategy and the Extraction Operator
- The Expansion and its Interpretation via Commutators

$$M_1 = e^{\beta (H - H_1)/2} e^{-\beta H/2}$$

$$M_1 = \left(\sum_{i=0}^{\infty} \frac{(\beta/2)^i}{i!} (H - H_1)^i\right) \left(\sum_{j=0}^{\infty} \frac{(\beta/2)^j}{j!} (-H)^j\right)$$

$$\begin{split} M_1 &= \Big(\sum_{i=0}^{\infty} \frac{(\beta/2)^i}{i!} (H - H_1)^i \Big) \Big(\sum_{j=0}^{\infty} \frac{(\beta/2)^j}{j!} (-H)^j \Big) \\ &= \sum_{k=0}^{\infty} \frac{(\beta/2)^k}{k!} \sum_{j=0}^k \binom{k}{j} (H - H_1^{k-j}) (-H)^j \end{split}$$

$$M_{1} = \left(\sum_{i=0}^{\infty} \frac{(\beta/2)^{i}}{i!} (H-H_{1})^{i}\right) \left(\sum_{j=0}^{\infty} \frac{(\beta/2)^{j}}{j!} (-H)^{j}\right)$$
$$= \sum_{k=0}^{\infty} \frac{(\beta/2)^{k}}{k!} \sum_{j=0}^{k} \binom{k}{j} (H-H_{1}^{k-j}) (-H)^{j}$$
$$f_{k}(H,H_{1})$$

We can compute

$$M_{1} = \left(\sum_{i=0}^{\infty} \frac{(\beta/2)^{i}}{i!} (H-H_{1})^{i}\right) \left(\sum_{j=0}^{\infty} \frac{(\beta/2)^{j}}{j!} (-H)^{j}\right)$$
$$= \sum_{k=0}^{\infty} \frac{(\beta/2)^{k}}{k!} \sum_{j=0}^{k} \binom{k}{j} (H-H_{1}^{k-j}) (-H)^{j}$$
$$f_{k}(H,H_{1})$$

Now let's try to interpret these expressions

A RECURRENCE

Easy to check that this expression

$$f_k(H, H_1) = \sum_{j=0}^k \binom{k}{j} (H - H_1)^{k-j} (-H)^j$$

A RECURRENCE

Easy to check that this expression

$$f_k(H, H_1) = \sum_{j=0}^k \binom{k}{j} (H - H_1)^{k-j} (-H)^j$$

satisfies a natural recurrence

$$f_k(H, H_1) = (H - H_1)f_{k-1}(H, H_1) - f_{k-1}(H, H_1)H$$

A RECURRENCE

Easy to check that this expression

$$f_k(H, H_1) = \sum_{j=0}^k \binom{k}{j} (H - H_1)^{k-j} (-H)^j$$

satisfies a natural recurrence

$$f_k(H, H_1) = (H - H_1)f_{k-1}(H, H_1) - f_{k-1}(H, H_1)H$$
$$= [H, f_{k-1}(H, H_1)] - H_1f_{k-1}(H, H_1)$$

where [A, B] = AB - BA is the commutator

UNDERSTANDING COMMUTATORS

Key Fact: For any A and B, we have

$$[A,B] = \begin{cases} 0 & \text{if their supports are disjoint} \\ & \text{supported on the union of their} \\ & \text{supports otherwise} \end{cases}$$

UNDERSTANDING COMMUTATORS

Key Fact: For any A and B, we have

$$[A,B] = \begin{cases} 0 & \text{if their supports are disjoint} \\ & \text{supported on the union of their} \\ & \text{supports otherwise} \end{cases}$$

The commutator is linear, so we can write

$$[H, f_{k-1}(H, H_1)] = \sum_{a} \lambda_a [E_a, f_{k-1}(H, H_1)]$$

UNDERSTANDING COMMUTATORS

Key Fact: For any A and B, we have

$$[A,B] = \begin{cases} 0 & \text{if their supports are disjoint} \\ & \text{supported on the union of their} \\ & \text{supports otherwise} \end{cases}$$

The commutator is linear, so we can write

$$[H, f_{k-1}(H, H_1)] = \sum_{a} \lambda_a [E_a, f_{k-1}(H, H_1)]$$

Hence, for each new term, its support grows by one incident edge

If our base graph is



If our base graph is



we can visualize the support of the terms in f_1



If our base graph is



If k = 2 and our base graph is



And the support of the terms in f_2



If k = 2 and our base graph is



And the support of the terms in f_2









Can also track how the coefficients grow

Does entanglement exhibit a sharp phase transition?

Does entanglement exhibit a sharp phase transition?

And does that transition happen earlier than NP-hardness

Does entanglement exhibit a sharp phase transition?

And does that transition happen earlier than NP-hardness

Prove strong bounds on conditional mutual information?

Does entanglement exhibit a sharp phase transition?

And does that transition happen earlier than NP-hardness

Prove strong bounds on conditional mutual information?

Related to classical notions like spatial mixing

Summary:

- New physical law: At high temperature, quantum spin systems have zero entanglement
- Proof via carefully designed expansions and the extraction operator
- For what kind of quantum spin systems should we expect quantum advantage in preparing the Gibbs state?
Summary:

- New physical law: At high temperature, quantum spin systems have zero entanglement
- Proof via carefully designed expansions and the extraction operator
- For what kind of quantum spin systems should we expect quantum advantage in preparing the Gibbs state?

Thanks! Any Questions?