Quantum Graphical Models and Belief Propagation

David Poulin

Center for the Physics of Information
California Institute of Technology

Joint work with: Matt Leifer and Ersen Bilgin

Classical and Quantum Information Theory
Santa Fe, March 2008
1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
Graphical models

Belief propagation

Quantum graphical models

Quantum belief propagation

Examples
- Quantum turbo-codes
- Many-body simulations
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:

- A (sparse) graph \( G = (V, E) \).
- Random variables \( u \), each associated with a vertex \( u \in V \).
- An efficiently specifiable distribution \( P(V) = P(u_1, u_2, \ldots) \).
- Edges \( e = (u, v) \) encode some kind of dependency relation in \( P \).
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

- Bayesian networks (artificial intelligence).
- Factor graphs (image recognition).
- Tanner graphs (coding theory).
- Markov networks (statistical physics).
- etc.

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(u_1, u_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A,B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.

- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A, B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C | B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A, B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A,B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Graphical models

Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A, B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C | B) = 0$.
- $P(A, B, C) = P(A)P(B | A)P(C | B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A | B)P(B | C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A | B)P(B)P(C | B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A, B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Let \( A, B, \) and \( C \) be three random variables with distribution \( P(A, B, C) \). We say that \( A \) and \( C \) are independent given \( B \) if

- Conditional mutual information vanishes \( I(A : C | B) = 0 \).
- \( P(A, B, C) = P(A)P(B|A)P(C|B) \) which suggests \( A \to B \to C \).
- \( P(A, B, C) = P(A|B)P(B|C)P(C) \) which suggests \( A \leftarrow B \leftarrow C \).
- \( P(A, B, C) = P(A|B)P(B)P(C|B) \) which suggests \( A \leftarrow B \to C \).

Re-defining the mutual distribution \( P(A : B) = \frac{P(A, B)}{P(A)P(B)} \), we can characterize conditional independence by

- \( P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C) \) which does not suggest a causal relation.
Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A, B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \to B \to C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \to C$.

Defining the mutual distribution $P(A : B) = \frac{P(A,B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.
Conditional independence

Let $A$, $B$, and $C$ be three random variables with distribution $P(A, B, C)$. We say that $A$ and $C$ are independent given $B$ if

- Conditional mutual information vanishes $I(A : C|B) = 0$.
- $P(A, B, C) = P(A)P(B|A)P(C|B)$ which suggests $A \rightarrow B \rightarrow C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ which suggests $A \leftarrow B \leftarrow C$.
- $P(A, B, C) = P(A|B)P(B)P(C|B)$ which suggests $A \leftarrow B \rightarrow C$.

Defining the mutual distribution $P(A : B) = \frac{P(A,B)}{P(A)P(B)}$, we can characterize conditional independence by

- $P(A, B, C) = P(A)P(B)P(C)P(A : B)P(B : C)$ which does not suggest a causal relation.

![Diagram](image_url)
Given a graph $G = (V, E)$ and a distribution $P(V)$, the pair $(G, P(V))$ forms a Markov Random Field iff:

- For all $U \subset V$, $I(U : V - U - n(U)|n(U)) = 0$.
- The correlations are shielded by the neighbors.
Markov random fields

Given a graph $G = (V, E)$ and a distribution $P(V)$, the pair $(G, P(V))$ forms a Markov Random Field iff:

- For all $U \subset V$, $I(U : V - U - n(U)|n(U)) = 0$.
- The correlations are shielded by the neighbors.
Given a graph \( G = (V, E) \) and a distribution \( P(V) \), the pair \((G, P(V))\) forms a Markov Random Field iff:

- For all \( U \subset V \), \( I(U : V - U - n(U) | n(U)) = 0 \).
- The correlations are shielded by the neighbors.

David Poulin (Caltech)
Given a graph $G = (V, E)$ and a distribution $P(V)$, the pair $(G, P(V))$ forms a Markov Random Field iff:

- For all $U \subset V$, $I(U : V - U - n(U)|n(U)) = 0$.
- The correlations are shielded by the neighbors.
Given a graph $G = (V, E)$ and a distribution $P(V)$, the pair $(G, P(V))$ forms a **Markov Random Field** iff:

- For all $U \subset V$, $I(U : V - U - n(U)|n(U)) = 0$.
- The correlations are shielded by the neighbors.

![Diagram of a graph with nodes and edges, illustrating the concept of Markov Random Fields.](attachment:diagram.png)
Given a graph $G = (V, E)$ and a distribution $P(V)$, the pair $(G, P(V))$ forms a Markov Random Field iff:

- For all $U \subset V$, $I(U : V - U - n(U)|n(U)) = 0$.
- The correlations are shielded by the neighbors.

![Graphical model diagram](image-url)
Theorem (Hammersley-Clifford)

The pair \((G, P(V))\) is a positive \((P > 0)\) random Markov field iff

\[
P(V) = \frac{1}{Z} \prod_{C \in c(G)} \psi(C).
\]

Special case: bifactor states (pairwise RMF)

When largest clique size is 2 (2d square lattice) or when \(\psi(C)\) is trivial for \(|C| > 2\), MRF are of the form

\[
P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u, v) \in E} \nu(u : v)
\]

\[
= \frac{1}{Z} \exp \left\{ -\beta \left( \sum_v h_v + \sum_{\langle u, v \rangle} k_{uv} \right) \right\}.
\]
**Hammersley-Clifford Theorem**

**Theorem (Hammersley-Clifford)**

*The pair \((G, P(V))\) is a positive \((P > 0)\) random Markov field iff*

\[
P(V) = \frac{1}{Z} \prod_{C \in \mathcal{C}(G)} \psi(C).
\]

**Special case: bifactor states (pairwise RMF)**

When largest clique size is 2 (2d square lattice) or when \(\psi(C)\) is trivial for \(|C| > 2\), MRF are of the form

\[
P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v)
= \frac{1}{Z} \exp \left\{ -\beta \left( \sum_v h_v + \sum_{\langle u,v \rangle} k_{uv} \right) \right\}.
\]
Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
Description of the algorithm

**Task (basic case)**

Given a graph $G = (V, E)$ and a bifactor distribution $P(V)$ on $G$, compute marginals

$$P(v) = \sum_{V \setminus v} P(V).$$

**Algorithm architecture**

- One processor per random variable $v$.
- Messages exchanged between processors related by an edge.
- Outgoing messages at $v$ depend on local "fields" $\mu(v)$ and $\nu(u : v)$ and received messages at $v$.
- The marginal $P(v)$ is estimated by a belief $b(v)$ that depends on the received messages at $v$ and the local fields.
- Exact when $G$ is a tree and complexity = $\text{diameter}(G)$.
- Good heuristic on loopy graphs.
Belief propagation

Description of the algorithm

Task (basic case)

Given a graph $G = (V, E)$ and a bifactor distribution $P(V)$ on $G$, compute marginals

$$P(v) = \sum_{V-v} P(V).$$

Algorithm architecture

- One processor per random variable $v$.
- Messages exchanged between processors related by an edge.
- Outgoing messages at $v$ depend on local "fields" $\mu(v)$ and $\nu(u : v)$ and received messages at $v$.
- The marginal $P(v)$ is estimated by a belief $b(v)$ that depends on the received messages at $v$ and the local fields.
- Exact when $G$ is a tree and complexity = diameter($G$).
- Good heuristic on loopy graphs.
Belief propagation

Description of the algorithm

Task (basic case)

Given a graph $G = (V, E)$ and a bifactor distribution $P(V)$ on $G$, compute marginals

$$P(v) = \sum_{V \setminus v} P(V).$$

Algorithm architecture

- One processor per random variable $v$.
- Messages exchanged between processors related by an edge.
  - Outgoing messages at $v$ depend on local "fields" $\mu(v)$ and $\nu(u : v)$ and received messages at $v$.
  - The marginal $P(v)$ is estimated by a belief $b(v)$ that depends on the received messages at $v$ and the local fields.
- Exact when $G$ is a tree and complexity = diameter($G$).
- Good heuristic on loopy graphs.
Description of the algorithm

Task (basic case)
Given a graph $G = (V, E)$ and a bifactor distribution $P(V)$ on $G$, compute marginals

$$P(v) = \sum_{V \setminus v} P(V).$$

Algorithm architecture
- One processor per random variable $v$.
- Messages exchanged between processors related by an edge.
- Outgoing messages at $v$ depend on local "fields" $\mu(v)$ and $\nu(u : v)$ and received messages at $v$.
- The marginal $P(v)$ is estimated by a belief $b(v)$ that depends on the received messages at $v$ and the local fields.
- Exact when $G$ is a tree and complexity = $\text{diameter}(G)$.
- Good heuristic on loopy graphs.
## Description of the algorithm

### Task (basic case)

Given a graph \( G = (V, E) \) and a bifactor distribution \( P(V) \) on \( G \), compute marginals

\[
P(v) = \sum_{V-v} P(V).
\]

### Algorithm architecture

- One processor per random variable \( v \).
- Messages exchanged between processors related by an edge.
- Outgoing messages at \( v \) depend on local "fields" \( \mu(v) \) and \( \nu(u : v) \) and received messages at \( v \).
- The marginal \( P(v) \) is estimated by a belief \( b(v) \) that depends on the received messages at \( v \) and the local fields.
- Exact when \( G \) is a tree and complexity = \text{diameter}(G).
- Good heuristic on loopy graphs.
Belief propagation

Description of the algorithm

Task (basic case)

Given a graph \( G = (V, E) \) and a bifactor distribution \( P(V) \) on \( G \), compute marginals

\[
P(v) = \sum_{V-v} P(V).
\]

Algorithm architecture

- One processor per random variable \( v \).
- Messages exchanged between processors related by an edge.
- Outgoing messages at \( v \) depend on local "fields" \( \mu(v) \) and \( \nu(u:v) \) and received messages at \( v \).
- The marginal \( P(v) \) is estimated by a belief \( b(v) \) that depends on the received messages at \( v \) and the local fields.
- Exact when \( G \) is a tree and complexity = diameter(\( G \)).
- Good heuristic on loopy graphs.
Belief propagation

Description of the algorithm

Task (basic case)

Given a graph $G = (V, E)$ and a bifactor distribution $P(V)$ on $G$, compute marginals

$$P(v) = \sum_{V \setminus v} P(V).$$

Algorithm architecture

- One processor per random variable $v$.
- Messages exchanged between processors related by an edge.
- Outgoing messages at $v$ depend on local "fields" $\mu(v)$ and $\nu(u : v)$ and received messages at $v$.
- The marginal $P(v)$ is estimated by a belief $b(v)$ that depends on the received messages at $v$ and the local fields.
- Exact when $G$ is a tree and complexity $= \text{diameter}(G)$.
- Good heuristic on loopy graphs.
Belief propagation algorithm

**Algorithm**

- **Initialization** \( m_{u \rightarrow v}(v) = cte. \)
- **Iterations** \( m_{u \rightarrow v}(v) \propto \sum_u \mu(u) \nu(u : v) \prod_{v' \in n(u) - v} m_{v' \rightarrow u}(u). \)

- **Beliefs** \( b(u) \propto \mu(u) \prod_{v \in n(u)} m_{v \rightarrow u}(u). \)
- \( b(u, v) \propto \mu(u) \mu(v) \nu(u : v) \prod_{w \in n(u) - v} m_{w \rightarrow u}(u) \prod_{w \in n(v) - u} m_{w \rightarrow v}(v). \)
Algorithm

- **Initialization** \( m_{u \rightarrow v} (v) = cte \).
- **Iterations** \( m_{u \rightarrow v} (v) \propto \sum_u \mu (u) \nu (u : v) \prod_{v' \in n(u) - v} m_{v' \rightarrow u} (u) \).
- **Beliefs** \( b(u) \propto \mu (u) \prod_{v \in n(u)} m_{v \rightarrow u} (u) \).
- \( b(u, v) \propto \mu (u) \mu (v) \nu (u : v) \prod_{w \in n(u) - v} m_{w \rightarrow u} (u) \prod_{w \in n(v) - u} m_{w \rightarrow v} (v) \).
Belief propagation

Algorithm

- Initialization \( m_{u \rightarrow v}(v) = \text{cte.} \)
- Iterations \( m_{u \rightarrow v}(v) \propto \sum_u \mu(u) \nu(u : v) \prod_{v' \in n(u) - v} m_{v' \rightarrow u}(u) \).

**Beliefs** \( b(u) \propto \mu(u) \prod_{v \in n(u)} m_{v \rightarrow u}(u) \).
- \( b(u, v) \propto \mu(u) \mu(v) \nu(u : v) \prod_{w \in n(u) - v} m_{w \rightarrow u}(u) \prod_{w \in n(v) - u} m_{w \rightarrow v}(v) \).
Belief propagation algorithm

Algorithm

- Initialization \( m_{u \rightarrow v}(v) = cte \).
- Iterations \( m_{u \rightarrow v}(v) \propto \sum_u \mu(u) \nu(u : v) \prod_{v' \in n(u) - v} m_{v' \rightarrow u}(u) \).

![Belief propagation diagram]

- Beliefs \( b(u) \propto \mu(u) \prod_{v \in n(u)} m_{v \rightarrow u}(u) \).
- \( b(u, v) \propto \mu(u) \mu(v) \nu(u : v) \prod_{w \in n(u) - v} m_{w \rightarrow u}(u) \prod_{w \in n(v) - u} m_{w \rightarrow v}(v) \).
Quantum graphical models

Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
A (sparse) graph $G = (V, E)$.
- Each vertex $u$ is associated a quantum system (spin) $u$ with Hilbert space $\mathcal{H}_u$.
- An efficiently specifiable quantum state $\rho_V$ on $\mathcal{H}_V = \bigotimes_{u \in V} \mathcal{H}_u$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?
- Many possible generalizations of classical bifactor states.
- They have applications in different contexts:
  - Quantum many-body.
  - Quantum error correction.
A (sparse) graph $G = (V, E)$.

Each vertex $u$ is associated a quantum system (spin) $u$ with Hilbert space $\mathcal{H}_u$.

An efficiently specifiable quantum state $\rho_V$ on $\mathcal{H}_V = \bigotimes_{u \in V} \mathcal{H}_u$.

Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?

- Many possible generalizations of classical bifactor states.
- They have applications in different contexts:
  - Quantum many-body.
  - Quantum error correction.
A (sparse) graph $G = (V, E)$.

Each vertex $u$ is associated a quantum system (spin) $u$ with Hilbert space $\mathcal{H}_u$.

An efficiently specifiable quantum state $\rho_V$ on $\mathcal{H}_V = \bigotimes_{u \in V} \mathcal{H}_u$.

Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?

- Many possible generalizations of classical bifactor states.
- They have applications in different contexts:
  - Quantum many-body.
  - Quantum error correction.
A (sparse) graph $G = (V, E)$.

Each vertex $u$ is associated with a quantum system (spin) $u$ with Hilbert space $\mathcal{H}_u$.

An efficiently specifiable quantum state $\rho_V$ on $\mathcal{H}_V = \bigotimes_{u \in V} \mathcal{H}_u$.

Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?

- Many possible generalizations of classical bifactor states.
- They have applications in different contexts:
  - Quantum many-body.
  - Quantum error correction.
A (sparse) graph $G = (V, E)$.
Each vertex $u$ is associated a quantum system (spin) $u$ with Hilbert space $H_u$.
An efficiently specifiable quantum state $\rho_V$ on $H_V = \bigotimes_{u \in V} H_u$.
Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?

- Many possible generalizations of classical bifactor states.
  - They have applications in different contexts:
    - Quantum many-body.
    - Quantum error correction.
A (sparse) graph $G = (V, E)$.
Each vertex $u$ is associated a quantum system (spin) $u$ with Hilbert space $\mathcal{H}_u$.
An efficiently specifiable quantum state $\rho_V$ on $\mathcal{H}_V = \bigotimes_{u \in V} \mathcal{H}_u$.
Edges $e = (u, v)$ encode some kind of dependency relation in $\rho_V$.

How to specify $\rho_V$?

- Many possible generalizations of classical bifactor states.
- They have applications in different contexts:
  - Quantum many-body.
  - Quantum error correction.
Bifactor state: \( P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v) \).

Quantum generalization: \( \mu_u \) and \( \nu_{u : v} \) operators on \( \mathcal{H}_u \) and \( \mathcal{H}_u \otimes \mathcal{H}_v \) respectively.

Problems:

- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: \( A \star^{(n)} B = (A^{1/2n} B^{1/n} A^{1/2n})^n \)

- \( n = 1 \): \( A \star B = A^{1/2} B A^{1/2} \) (measurement, QEC).
- \( n = \infty \): \( A \otimes B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
Bifactor state: \( P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v) \).
Quantum generalization: \( \mu_u \) and \( \nu_{u:v} \) operators on \( \mathcal{H}_u \) and \( \mathcal{H}_u \otimes \mathcal{H}_v \) respectively.

Problems:

- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: \( A \star^{(n)} B = (A^{1/2n} B^{1/n} A^{1/2n})^n \)

- \( n = 1 \): \( A \star B = A^{1/2} BA^{1/2} \) (measurement, QEC).
- \( n = \infty \): \( A \odot B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
Bifactor state: $P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v)$.
Quantum generalization: $\mu_u$ and $\nu_{u:v}$ operators on $\mathcal{H}_u$ and $\mathcal{H}_u \otimes \mathcal{H}_v$ respectively.

Problems:
- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: $A \star^{(n)} B = (A_{2^n}^1 B_{2^n}^1 A_{2^n}^1)^n$
- $n = 1$: $A \star B = A_2^1 B A_2^1$ (measurement, QEC).
- $n = \infty$: $A \circ B = \exp(\log A + \log B)$ (Hamiltonian, many-body).
- Intermediate $n$: Trotter-Suzuki decomposition.
Bifactor state: $P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v)$. 
Quantum generalization: $\mu_u$ and $\nu_{u:v}$ operators on $\mathcal{H}_u$ and $\mathcal{H}_u \otimes \mathcal{H}_v$ respectively.

Problems:

- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: $A \star^{(n)} B = (A^{1 \over 2n} B^{1 \over n} A^{1 \over 2n})^n$

- $n = 1$: $A \star B = A^{1 \over 2} B A^{1 \over 2}$ (measurement, QEC).
- $n = \infty$: $A \circ B = \exp(\log A + \log B)$ (Hamiltonian, many-body).
- Intermediate $n$: Trotter-Suzuki decomposition.
Bifactor state: \( P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v). \)

Quantum generalization: \( \mu_u \) and \( \nu_{u:v} \) operators on \( \mathcal{H}_u \) and \( \mathcal{H}_u \otimes \mathcal{H}_v \) respectively.

Problems:
- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: \( A \star^{(n)} B = (A^{1/2n} B^{1/n} A^{1/2n})^n \)

- \( n = 1 \): \( A \star B = A^{1/2} B A^{1/2} \) (measurement, QEC).
- \( n = \infty \): \( A \otimes B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
Bifactor state: $P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u:v)$.

Quantum generalization: $\mu_u$ and $\nu_{u:v}$ operators on $\mathcal{H}_u$ and $\mathcal{H}_u \otimes \mathcal{H}_v$ respectively.

Problems:

- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: $A \star^{(n)} B = (A^{\frac{1}{2n}} B^{\frac{1}{n}} A^{\frac{1}{2n}})^n$

- $n = 1$: $A \star B = A^{\frac{1}{2}} B A^{\frac{1}{2}}$ (measurement, QEC).
- $n = \infty$: $A \otimes B = \exp(\log A + \log B)$ (Hamiltonian, many-body).
- Intermediate $n$: Trotter-Suzuki decomposition.
Bifactor state: \[ P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v). \]
Quantum generalization: \( \mu_u \) and \( \nu_{u:v} \) operators on \( \mathcal{H}_u \) and \( \mathcal{H}_u \otimes \mathcal{H}_v \) respectively.

Problems:
- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: \[ A \star^{(n)} B = (A^{1/2n} B^{1/n} A^{1/2n})^n \]
- \( n = 1 \): \( A \star B = A^{1/2} B A^{1/2} \) (measurement, QEC).
- \( n = \infty \): \( A \circ B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
Bifactor state: \( P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u:v). \)

Quantum generalization: \( \mu_u \) and \( \nu_{u:v} \) operators on \( \mathcal{H}_u \) and \( \mathcal{H}_u \otimes \mathcal{H}_v \) respectively.

Problems:

- Ambiguity in order of the terms.
- Not necessarily positive.

Define the family of products: \( A \star^{(n)} B = (A^{1_{2n}} B^{1_n} A^{1_{2n}})^n \)

- \( n = 1: A \star B = A^{1_2} B A^{1_2} \) (measurement, QEC).
- \( n = \infty: A \odot B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
In analogy with the classical case, define

- **Conditional state** $\rho_{A|B}^{(n)} = \rho_B^{-1} \star^{(n)} \rho_{AB}$.

- **Mutual state** $\rho_{A:B}^{(n)} = (\rho_A^{-1} \rho_B^{-1}) \star^{(n)} \rho_{AB}$.
Quantum conditional independence

Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C|B) = 0$ which implies:

- $\rho_{ABC} = \rho_A \star^{(n)} \rho_{B|A} \star^{(n)} \rho^{(n)}_{C|B}$ which suggests $A \rightarrow B \rightarrow C$.
- $\rho_{ABC} = \rho_C \star^{(n)} \rho^{(n)}_{A|B} \star^{(n)} \rho^{(n)}_{B|C}$ which suggests $A \leftarrow B \leftarrow C$.
- $\rho_{ABC} = \rho_B \star^{(n)} \rho^{(n)}_{A|B} \star^{(n)} \rho^{(n)}_{C|B}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

$\rho_{ABC} = (\rho_A \rho_B \rho_C) \star^{(n)} (\rho^{(n)}_{A:B} \rho^{(n)}_{B:C})$ is a quantum bifactor network.
Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C|B) = 0$ which implies:

- $\rho_{ABC} = \rho_A \star^{(n)} \rho_{B|A}^{(n)} \rho_{C|B}^{(n)}$ which suggests $A \rightarrow B \rightarrow C$.
- $\rho_{ABC} = \rho_C \star^{(n)} \rho_{A|B}^{(n)} \rho_{B|C}^{(n)}$ which suggests $A \leftarrow B \leftarrow C$.
- $\rho_{ABC} = \rho_B \star^{(n)} \rho_{A|B}^{(n)} \rho_{C|B}^{(n)}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

$\rho_{ABC} = (\rho_A \rho_B \rho_C) \star^{(n)} (\rho_{A:B}^{(n)} \rho_{B:C}^{(n)})$ is a quantum bifactor network.
Quantum conditional independence

Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C|B) = 0$ which implies:

- $\rho_{ABC} = \rho_A \star^{(n)} \rho_{B|A} \star^{(n)} \rho_{C|B}$ which suggests $A \rightarrow B \rightarrow C$.
- $\rho_{ABC} = \rho_C \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{B|C}$ which suggests $A \leftarrow B \leftarrow C$.
- $\rho_{ABC} = \rho_B \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{C|B}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

- $\rho_{ABC} = (\rho_A \rho_B \rho_C) \star^{(n)} (\rho_{A:B}^{(n)} \rho_{B:C}^{(n)})$ is a quantum bifactor network.
Quantum conditional independence

Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C | B) = 0$ which implies:

- $\rho_{ABC} = \rho_A \star^{(n)} \rho_{B|A} \star^{(n)} \rho_{C|B}$ which suggests $A \rightarrow B \rightarrow C$.
- $\rho_{ABC} = \rho_C \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{B|C}$ which suggests $A \leftarrow B \leftarrow C$.
- $\rho_{ABC} = \rho_B \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{C|B}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

- $\rho_{ABC} = (\rho_A \rho_B \rho_C) \star^{(n)} (\rho_{A:B}^{(n)} \rho_{B:C}^{(n)})$ is a quantum bifactor network.
Quantum conditional independence

Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C|B) = 0$ which implies:

1. $\rho_{ABC} = \rho_A \star^{(n)} \rho_{B|A} \star^{(n)} \rho_{C|B}$ which suggests $A \rightarrow B \rightarrow C$.
2. $\rho_{ABC} = \rho_C \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{B|C}$ which suggests $A \leftarrow B \leftarrow C$.
3. $\rho_{ABC} = \rho_B \star^{(n)} \rho_{A|B} \star^{(n)} \rho_{C|B}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

$\rho_{ABC} = (\rho_A \rho_B \rho_C) \star^{(n)} (\rho_{A:B} \rho_{B:C})$ is a quantum bifactor network.
Quantum conditional independence

Given three quantum systems $A$, $B$, and $C$ and a joint state $\rho_{ABC}$, we say that $A$ and $C$ are independent given $B$ if $I(A : C|B) = 0$ which implies:

- $\rho_{ABC} = \rho_A \star (n) \rho_{B|A} \star (n) \rho_{C|B}$ which suggests $A \rightarrow B \rightarrow C$.
- $\rho_{ABC} = \rho_C \star (n) \rho_{A|B} \star (n) \rho_{B|C}$ which suggests $A \leftarrow B \leftarrow C$.
- $\rho_{ABC} = \rho_B \star (n) \rho_{A|B} \star (n) \rho_{C|B}$ which suggests $A \leftarrow B \rightarrow C$.

These conditions differ for different values of $n$ and differ between each other.

- $\rho_{ABC} = (\rho_A \rho_B \rho_C) \star (n) (\rho_{A:B} \rho_{B:C})$ is a quantum bifactor network.
Theorem

For $n = \infty$, all conditions are equivalent and imply conditional independence.

Theorem

For $n = 1$, the first two conditions are equivalent and imply conditional independence.

Theorem (Quantum Hammersley-Clifford)

If $(\rho_V, G)$ is a positive quantum Markov network, then

$$
\rho_V = \bigotimes_{C \in \mathcal{E}(G)} \sigma_C = \exp \left\{ -\beta \sum_{C \in \mathcal{E}(G)} h_C \right\}.
$$
Quantum graphical models

Quantum conditional independence

Theorem

For \( n = \infty \), all conditions are equivalent and imply conditional independence.

Theorem

For \( n = 1 \), the first two conditions are equivalent and imply conditional independence.

Theorem (Quantum Hammersley-Clifford)

If \((\rho_V, G)\) is a positive quantum Markov network, then

\[
\rho_V = \bigotimes_{C \in \mathcal{C}(G)} \sigma_C = \exp \left\{ -\beta \sum_{C \in \mathcal{C}(G)} h_C \right\}.
\]
For $n = \infty$, all conditions are equivalent and imply conditional independence.

For $n = 1$, the first two conditions are equivalent and imply conditional independence.

If $(\rho_V, G)$ is a positive quantum Markov network, then

$$
\rho_V = \bigotimes_{C \in \mathcal{C}(G)} \sigma_C = \exp \left\{ -\beta \sum_{C \in \mathcal{C}(G)} h_C \right\}.
$$
Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
The algorithm

Cut and paste from previous section.
Don't forget to search for $\Pi$ and replace by $\star^{(n)}$. 

M. Hastings '07
The algorithm

Cut and paste from previous section.
Don't forget to search for $\Pi$ and replace by $\star^{(n)}$.

M. Hastings ’07
Let $G = (V, E)$ be a graph and let

$$\rho_V = \frac{1}{Z} \left( \bigotimes_{u \in V} \mu_u \right) \star^{(n)} \left( \prod_{(u, v) \in E} \nu_{u:v} \right)$$

be a bifactor state on $G$.

**Theorem**

If $G$ is a tree and $(G, \rho_V)$ is a quantum Markov random field, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V \setminus u} \{ \rho_V \}$ in a time proportional to $\text{depth}(G)$.

**Theorem**

If $G$ is a tree and $n = 1$, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V \setminus u} \{ \rho_V \}$ in a time proportional to $\text{depth}(G)$. 
Let $G = (V, E)$ be a graph and let

$$\rho_V = \frac{1}{Z} \left( \bigotimes_{u \in V} \mu_u \right) \star^{(n)} \left( \prod_{(u, v) \in E} \nu_{u:v} \right)$$

be a bifactor state on $G$.

**Theorem**

If $G$ is a tree and $(G, \rho_V)$ is a quantum Markov random field, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V-u}\{\rho_V\}$ in a time proportional to $\text{depth}(G)$.

**Theorem**

If $G$ is a tree and $n = 1$, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V-u}\{\rho_V\}$ in a time proportional to $\text{depth}(G)$.
Let $G = (V, E)$ be a graph and let

$$\rho_V = \frac{1}{Z} \left( \bigotimes_{u \in V} \mu_u \right) \ast^{(n)} \left( \prod_{(u,v) \in E} \nu_{u,v} \right)$$

be a bifactor state on $G$.

**Theorem**

If $G$ is a tree and $(G, \rho_V)$ is a quantum Markov random field, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V-u} \{ \rho_V \}$ in a time proportional to $\text{depth}(G)$.

**Theorem**

If $G$ is a tree and $n = 1$, then the beliefs $b_u$ converge to the correct marginals $\rho_u = \text{Tr}_{V-u} \{ \rho_V \}$ in a time proportional to $\text{depth}(G)$.
Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
Turbo code performances on depolarization channel

- Rate is fixed at $\frac{1}{9}$.
- Error probability decreases as number of encoded qubits increases.
- Error-free "phase transition" at 0.1.
- With finite size, $10^{-4}$ threshold around $\epsilon = 0.08$.

Best performance to date at this rate.

Poulin, Tillich, and Ollivier’07.
Turbo code performances on depolarization channel

- Rate is fixed at $\frac{1}{9}$.
- Error probability decreases as number of encoded qubits increases.
- Error-free "phase transition" at 0.1.
- With finite size, $10^{-4}$ threshold around $\epsilon = 0.08$.

Best performance to date at this rate.

Poulin, Tillich, and Ollivier’07.
Rate is fixed at $\frac{1}{9}$.

Error probability decreases as the number of encoded qubits increases.

Error-free "phase transition" at 0.1.

With finite size, $10^{-4}$ threshold around $\epsilon = 0.08$.

Best performance to date at this rate.

Poulin, Tillich, and Ollivier’07.
Quantum turbo-codes

Turbo code performances on depolarization channel

- Rate is fixed at $\frac{1}{9}$.
- Error probability decreases as number of encoded qubits increases.
- Error-free "phase transition" at 0.1.
- With finite size, $10^{-4}$ threshold around $\epsilon = 0.08$.

Best performance to date at this rate.

Poulin, Tillich, and Ollivier'07.
Rate is fixed at $\frac{1}{9}$.

Error probability decreases as number of encoded qubits increases.

Error-free "phase transition" at 0.1.

With finite size, $10^{-4}$ threshold around $\epsilon = 0.08$.

Best performance to date at this rate.

Poulin, Tillich, and Ollivier’07.
Fig. 10. Summary of performances of several quantum codes on the 4-ary symmetric channel (depolarizing channel), treated by all decoding algorithms shown in this figure as if the channel were a pair of independent binary-symmetric channels. Each point shows the marginal noise level at which the block error probability is.

In the case of dual-containing codes, this is the noise level at which each of the two identical constituent codes (see (19)) has an error probability of.

As a aid to the eye, lines have been added between the four unicycle codes U; between a sequence of bicycle codes B all of block length with different rates; and between a sequence of of BCH codes with increasing block length. The curve labeled S2 is the Shannon limit if the correlations between errors and errors are neglected, (45).

Points " are codes invented elsewhere. All other point styles denote codes presented for the first time in this paper.

Fig. 11. Summary of performances of several codes on the 4-ary symmetric channel (depolarizing channel). The additional points at the right and bottom are as follows. 3786(B,4SC): a code of construction B (the same code as its neighbor in the figure) decoded with a decoder that exploits the known correlations between errors and errors.

3786(B,D): the same code as the code to its left in the figure, simulated with a channel where the qubits have a diversity of known reliabilities; errors and errors occur independently with probabilities determined from a Gaussian distribution; the channel in this case is not the 4-ary symmetric channel, but we plot the performance at the equivalent value of.

: an algebraically constructed quantum code (not a sparse-graph code) from [10].
1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Quantum turbo-codes
   - Many-body simulations
One dimensional classical system

Consider the 1d classical system with hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$.

Its Gibbs distribution is $(\mu(i) = e^{-\beta h_i}$ and $\nu(i, j) = e^{-\beta J_{ij}}$)

$$\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)}$$

$$= \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots$$

So the partition function can be evaluated step by step:

$$m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)$$

$$m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{i_1\rightarrow i_2}(i_2) \mu(i_2) \nu(i_2, i_3)$$

$$m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{i_2\rightarrow i_3}(i_3) \mu(i_3) \nu(i_3, i_4)$$

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1}\rightarrow i_N} \mu(i_N)$$
Consider the 1d classical system with Hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$. Its Gibbs distribution is $(\mu(i) = e^{-\beta h_i}$ and $\nu(i, j) = e^{-\beta J_{ij}}$)

$$\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)}$$

$$= \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots$$

So the partition function can be evaluated step by step:

$$m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)$$

$$m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{1\rightarrow 2}(i_2) \mu(i_2) \nu(i_2, i_3)$$

$$m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{2\rightarrow 3}(i_3) \mu(i_3) \nu(i_3, i_4)$$

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1}\rightarrow i_N} \mu(i_N)$$
Consider the 1d classical system with hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \).

Its Gibbs distribution is \( (\mu(i) = e^{-\beta h_i} \text{ and } \nu(i, j) = e^{-\beta J_{ij}}) \)

\[
\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)} = \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots
\]

So the partition function can be evaluated step by step:

\[
m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)
\]

\[
m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{1\rightarrow 2}(i_2) \mu(i_2) \nu(i_2, i_3)
\]

\[
m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{2\rightarrow 3}(i_3) \mu(i_3) \nu(i_3, i_4)
\]

\[
\vdots
\]

\[
Z = \sum_{i_N} m_{i_{N-1}\rightarrow i_N} \mu(i_N)
\]
One dimensional classical system

Consider the 1d classical system with Hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \).

Its Gibbs distribution is (\( \mu(i) = e^{-\beta h_i} \) and \( \nu(i, j) = e^{-\beta J_{ij}} \))

\[
\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)} = \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots
\]

So the partition function can be evaluated step by step:

\[
m_{1\to 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)
\]

\[
m_{2\to 3}(i_3) = \sum_{i_2} m_{1\to 2}(i_2) \mu(i_2) \nu(i_2, i_3)
\]

\[
m_{3\to 4}(i_4) = \sum_{i_3} m_{2\to 3}(i_3) \mu(i_3) \nu(i_3, i_4)
\]

\[
\vdots
\]

\[
Z = \sum_{i_N} m_{i_{N-1}\to i_N} \mu(i_N)
\]
One dimensional classical system

Consider the 1d classical system with hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$. Its Gibbs distribution is $(\mu(i) = e^{-\beta h_i}$ and $\nu(i, j) = e^{-\beta J_{ij}}$)

$$\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)}$$

$$= \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots$$

So the partition function can be evaluated step by step:

$$m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)$$

$$m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{i_1\rightarrow 2}(i_2) \mu(i_2) \nu(i_2, i_3)$$

$$m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{i_2\rightarrow 3}(i_3) \mu(i_3) \nu(i_3, i_4)$$

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1}\rightarrow N}(i_N) \mu(i_N)$$
One dimensional classical system

Consider the 1d classical system with Hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$. Its Gibbs distribution is $(\mu(i) = e^{-\beta h_i}$ and $\nu(i, j) = e^{-\beta J_{ij}})$

$$\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)}$$

$$= \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots$$

So the partition function can be evaluated step by step:

$$m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)$$

$$m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{1\rightarrow 2}(i_2) \mu(i_2) \nu(i_2, i_3)$$

$$m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{2\rightarrow 3}(i_3) \mu(i_3) \nu(i_3, i_4)$$

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1}\rightarrow i_N} \mu(i_N)$$
Consider the 1d classical system with Hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \). Its Gibbs distribution is \( (\mu(i) = e^{-\beta h_i} \text{ and } \nu(i, j) = e^{-\beta J_{ij}}) \)

\[
\rho(i_1, i_2, \ldots) = \frac{1}{Z} e^{-\beta H(i_1, i_2, \ldots)} = \frac{1}{Z} \mu(i_1) \nu(i_1, i_2) \mu(i_2) \nu(i_2, i_3) \mu(i_3) \ldots
\]

So the partition function can be evaluated step by step:

\[
m_{1\rightarrow 2}(i_2) = \sum_{i_1} \mu(i_1) \nu(i_1, i_2)
\]

\[
m_{2\rightarrow 3}(i_3) = \sum_{i_2} m_{i_1\rightarrow i_2}(i_2) \mu(i_2) \nu(i_2, i_3)
\]

\[
m_{3\rightarrow 4}(i_4) = \sum_{i_3} m_{i_2\rightarrow i_3}(i_3) \mu(i_3) \nu(i_3, i_4)
\]

\[
\vdots
\]

\[
Z = \sum_{i_N} m_{i_{N-1}\rightarrow i_N} \mu(i_N)
\]
Consider the 1d quantum system with Hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$.

Its Gibbs distribution is $(\mu_i = e^{-\beta h_i}$ and $\nu_{i:j} = e^{-\beta J_{ij}})$

$$\rho_V = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z} \mu_i \otimes \nu_{i_1:i_2} \otimes \mu_{i_2} \otimes \nu_{i_2:i_3} \otimes \mu_{i_3} \ldots$$

Bottleneck for computing $Z$:

$$\text{Tr}_A\{\mu_A \otimes \nu_{A:B} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C\} \neq \text{Tr}_A\{\mu_A \otimes \nu_{A:B}\} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C$$

But it is equal when $I(A : C|B) = 0$. 
Consider the 1d **quantum** system with hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \).

Its Gibbs distribution is (\( \mu_i = e^{-\beta h_i} \) and \( \nu_{i:j} = e^{-\beta J_{ij}} \))

\[
\rho_V = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z} \mu_{i_1} \otimes \nu_{i_1:i_2} \otimes \mu_{i_2} \otimes \nu_{i_2:i_3} \otimes \mu_{i_3} \ldots
\]

Bottleneck for computing \( Z \):

\[
\text{Tr}_A\{\mu_A \otimes \nu_{A:B} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C\} \neq \text{Tr}_A\{\mu_A \otimes \nu_{A:B}\} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C
\]

But it is equal when \( I(A : C|B) = 0 \).
Consider the 1d \textit{quantum} system with hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \). Its Gibbs distribution is \((\mu_i = e^{-\beta h_i} \text{ and } \nu_{i:j} = e^{-\beta J_{ij}})\)

\[
\rho_V = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z} \mu_{i_1} \otimes \nu_{i_1:i_2} \otimes \mu_{i_2} \otimes \nu_{i_2:i_3} \otimes \mu_{i_3} \ldots
\]

Bottleneck for computing \( Z \):

\[
\text{Tr}_A\{\mu_A \otimes \nu_{A:B} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C\} \neq \text{Tr}_A\{\mu_A \otimes \nu_{A:B}\} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C
\]

But it is equal when \( I(A : C|B) = 0 \).
Consider the 1d \textit{quantum} system with hamiltonian $H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij}$. Its Gibbs distribution is $(\mu_i = e^{-\beta h_i}$ and $\nu_{i:j} = e^{-\beta J_{ij}})$

$$\rho_V = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z} \mu_{i_1} \otimes \nu_{i_1:i_2} \otimes \mu_{i_2} \otimes \nu_{i_2:i_3} \otimes \mu_{i_3} \ldots$$

Bottleneck for computing $Z$:

$$\text{Tr}_A\{\mu_A \otimes \nu_{A:B} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C\} \neq \text{Tr}_A\{\mu_A \otimes \nu_{A:B}\} \otimes \mu_B \otimes \nu_{B:C} \otimes \mu_C$$

But it is equal when $I(A : C|B) = 0$. 
Effective thermal hamiltonian

\[ H = \sum_{i=-\infty}^{\infty} h_i + J_{i,i+1} \quad \text{and} \quad \rho = \frac{1}{Z} \exp\{-\beta H\} \]

Effective Thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)


Effective Thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)

\[ \rho = \frac{1}{Z} \exp\{-\beta H\} \]
Effective thermal Hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1} \{ \rho \} \]

\[ \rho = \frac{1}{Z} \exp \{-\beta H\} \]

Effective Thermal Hamiltonian = \[ \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \]
Effective thermal hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1} \{ \rho \} \]

\[ \rho' = \frac{1}{Z'} \exp\{ -\beta H_{\text{eff}} \} \]

Effective Thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)
Effective thermal hamiltonian

$$\rho' = \text{Tr}_{-\infty \ldots -1} \{\rho\} \quad \rho' = \frac{1}{Z'} \exp\{ -\beta H_{\text{eff}} \}$$

Effective Thermal Hamiltonian = $\sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots$
Effective thermal hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1} \{ \rho \} \quad \rho' = \frac{1}{Z'} \exp\{-\beta H_{\text{eff}}\} \]

Effective Thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)
Effective thermal hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1} \{ \rho \} \]

\[ \rho' = \frac{1}{Z'} \exp \{-\beta H_{\text{eff}}\} \]

Effective Thermal Hamiltonian = \[ \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \]
Effective thermal hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1}\{\rho\} \quad \rho' = \frac{1}{Z'} \exp\{-\beta H_{\text{eff}}\} \]

\[ \ldots (-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots) \]

Effective Thermal Hamiltonian = \[ \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \]
Effective thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)
Effective thermal Hamiltonian

\[ \rho' = \text{Tr}_{-\infty \ldots -1} \{ \rho \} \quad \rho' = \frac{1}{Z'} \exp\{-\beta H_{\text{eff}}\} \]

Effective Thermal Hamiltonian = \( \sum_{i=1}^{\infty} h_i + J_{i,i+1} + V_1 + V_2 + V_3 + V_4 \ldots \)
One dimensional quantum system

\[
\begin{align*}
\sigma_{1-4} &= e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})} \\
\sigma'_{2-4} &= \text{Tr}_1\{\sigma_{1-4}\}, \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \\
\sigma_{2-5} &= e^{-\beta(h'_{2-4} + h_5 + J_{45})} \\
\sigma'_{3-5} &= \text{Tr}_2\{\sigma_{2-5}\}, \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \\
\sigma_{3-6} &= e^{-\beta(h'_{3-5} + h_6 + J_{56})} \\
&\vdots \\
Z &= \text{Tr}\{\sigma_{N-3,N-2,N-1,N}\}
\end{align*}
\]
One dimensional quantum system

\[ \sigma_{1-4} = e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})} \]

\[ \sigma'_{2-4} = Tr_1\{\sigma_{1-4}\} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \]

\[ \sigma_{2-5} = e^{-\beta(h'_{2-4} + h_5 + J_{45})} \]

\[ \sigma'_{3-5} = Tr_2\{\sigma_{2-5}\} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \]

\[ \sigma_{3-6} = e^{-\beta(h'_{3-5} + h_6 + J_{56})} \]

\[ \vdots \]

\[ Z = Tr\{\sigma_{N-3,N-2,N-1,N}\} \]
One dimensional quantum system

\[ h'_{2-4} \]

\[ \sigma_{1-4} = e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})} \]

\[ \sigma'_{2-4} = Tr_1 \{ \sigma_{1-4} \} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \]

\[ \sigma_{2-5} = e^{-\beta(h'_{2-4} + h_5 + J_{45})} \]

\[ \sigma'_{3-5} = Tr_2 \{ \sigma_{2-5} \} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \]

\[ \sigma_{3-6} = e^{-\beta(h'_{3-5} + h_6 + J_{56})} \]

\[ \vdots \]

\[ Z = Tr \{ \sigma_{N-3,N-2,N-1,N} \} \]
One dimensional quantum system

\[
\sigma_{1-4} = e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})}
\]

\[
\sigma'_{2-4} = \text{Tr}_1\{\sigma_{1-4}\} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4}
\]

\[
\sigma_{2-5} = e^{-\beta(h'_{2-4} + h_5 + J_{45})}
\]

\[
\sigma'_{3-5} = \text{Tr}_2\{\sigma_{2-5}\} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5}
\]

\[
\sigma_{3-6} = e^{-\beta(h'_{3-5} + h_6 + J_{56})}
\]

\[
\vdots
\]

\[
Z = \text{Tr}\{\sigma_{N-3,N-2,N-1,N}\}
\]
One dimensional quantum system

\[ \sigma_{1-4} = e^{-\beta(h_{1} + h_{2} + h_{3} + h_{4} + J_{12} + J_{23} + J_{34})} \]

\[ \sigma'_{2-4} = Tr_{1}\{\sigma_{1-4}\} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \]

\[ \sigma_{2-5} = e^{-\beta(h'_{2-4} + h_{5} + J_{45})} \]

\[ \sigma'_{3-5} = Tr_{2}\{\sigma_{2-5}\} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \]

\[ \sigma_{3-6} = e^{-\beta(h'_{3-5} + h_{6} + J_{56})} \]

\[ \vdots \]

\[ Z = Tr\{\sigma_{N-3,N-2,N-1,N}\} \]
One dimensional quantum system

\[ \sigma_{3-6} \]

\[ \sigma_{1-4} = e^{-\beta(h_1+h_2+h_3+h_4+J_{12}+J_{23}+J_{34})} \]

\[ \sigma'_{2-4} = \text{Tr}_1\{\sigma_{1-4}\} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \]

\[ \sigma_{2-5} = e^{-\beta(h'_2+h_5+J_{45})} \]

\[ \sigma'_{3-5} = \text{Tr}_2\{\sigma_{2-5}\} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \]

\[ \sigma_{3-6} = e^{-\beta(h'_3+h_6+J_{56})} \]

\[ \vdots \]

\[ Z = \text{Tr}\{\sigma_{N-3,N-2,N-1,N}\} \]
One dimensional quantum system

\[ \sigma_{N-3,N-2,N-1,N} \]

\[
\sigma_{1-4} = e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})}
\]

\[
\sigma'_{2-4} = \text{Tr}_1 \{ \sigma_{1-4} \} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4}
\]

\[
\sigma_{2-5} = e^{-\beta(h'_{2-4} + h_5 + J_{45})}
\]

\[
\sigma'_{3-5} = \text{Tr}_2 \{ \sigma_{2-5} \} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5}
\]

\[
\sigma_{3-6} = e^{-\beta(h'_{3-5} + h_6 + J_{56})}
\]

\[
Z = \text{Tr}\{ \sigma_{N-3,N-2,N-1,N} \}
\]
Examples
Many-body simulations

Critical 1d Ising model

- Replica: Trotter decomposition $N_\tau = 10$ (bifactor $\star^{(10)}$).
- TEBD: Time-evolving block decimation (DMRG $\chi = 150$).
- Sliding window $\ell = 6$ (bifactor $\odot$).

Bilgin and Poulin ’07.
1D anti-ferromagnetic Heisenberg model

- **Bethe Ansatz**: exact (A. Klümper and D. C. Johnston, PRL'00).
- **Sliding window** $\ell = 9$ (bifactor ⋄).
Laumann, Scardicchio, and Sondhi ’07, Bilgin and Poulin.
Quantum Monte Carlo: M.S. Makivić and H.-Q. Ding PRB’91.
10th-order $J/T$ expansion.
Quantum Belief propagation, window size 7.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems. Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).

- Good heuristic for iterative decoding of sparse and quantum turbo codes.
- Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems. Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).

- Good heuristic for iterative decoding of sparse and quantum turbo codes.
- Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems. Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).

- Good heuristic for iterative decoding of sparse and quantum turbo codes.

- Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems. Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).

Good heuristic for iterative decoding of sparse and quantum turbo codes.

Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems. Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).
- Good heuristic for iterative decoding of sparse and quantum turbo codes.
- Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems.

Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov $\Rightarrow$ Gibbs).

Good heuristic for iterative decoding of sparse and quantum turbo codes.

Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details.
Belief propagation operating on graphical models is a powerful, highly parallelizable, heuristic for all sorts of inference problems.

Many of these properties carry over to the quantum realm:

- Half Hammersley-Clifford Theorem (Markov ⇒ Gibbs).

Good heuristic for iterative decoding of sparse and quantum turbo codes.

Good heuristic for many-body systems on graphs with no small loops.

See poster by Ersen Bilgin for more details