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

Sydney Quantum Information Theory Workshop
Coogee, January 2008
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- **Arbitrary channel**: messages are probabilities (NP-hard).
- **Quantum channel**: messages are positive operators.
- **Exact** when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
Example: erasure channel

- Arbitrary channel: messages are probabilities (NP-hard).
- Quantum channel: messages are positive operators.
- Exact when there are no loops (Tanner graph is a tree).
1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Many-body simulations
   - Quantum turbo-codes
Graphical models

2. Belief propagation

3. Quantum graphical models

4. Quantum belief propagation

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

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

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:
- A (sparse) graph $G = (V, E)$.
  - Random variables $u$, each associated with a vertex $u \in V$.
  - An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
  - Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:
- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- Edges $e = (u, v)$ encode some kind of dependency relation in $P$. 
Graphical models

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

Common features:

- A (sparse) graph $G = (V, E)$.
- Random variables $u$, each associated with a vertex $u \in V$.
- An efficiently specifiable distribution $P(V) = P(v_1, v_2, \ldots)$.
- 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 $A \to B \to C$.
- $P(A, B, C) = P(A|B)P(B|C)P(C)$ 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.
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 $A \leftarrow B \to C$.

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

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

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

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

Special case: bifactor states

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
   - Many-body simulations
   - Quantum turbo-codes
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 = \( \text{depth}(G) \).
- 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 = $\text{depth}(G)$.
- 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 $= \text{depth}(G)$.
- 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 = $\text{depth}(G)$.
- 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 = $\text{depth}(G)$.
- 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 = $\text{depth}(G)$.
- 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 = \( \text{depth}(G) \).
- Good heuristic on loopy graphs.
Algorithm

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

Beliefs

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

Algorithm

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

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** $m_{u \rightarrow v}(v) = \text{cte}$.  
- **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 $b(u) \propto \mu(u) \prod_{v \in n(u)} m_{v \rightarrow u}(u)$.  

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

David Poulin (Caltech)

Quantum Belief Propagation

Coogee’08 11 / 29
Algorithm

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

![Graph Diagram]

- **Beliefs** \( b(u) \propto \mu(u) \prod_{v \in n(u)} m_{v \rightarrow u}(u). \)
- \( b(u, v) \propto \mu(u) \mu(v) \nu(u : v) \prod_{w \in n(u) - v} m_{w \rightarrow u}(u) \prod_{w \in n(v) - u} m_{w \rightarrow v}(v). \)
Let $H = \sum_v h_v + \sum_{\langle u,v \rangle} k_{uv}$ be a local Hamiltonian on $G$.

Given a probability $P(V)$, the Gibbs free energy $G = E - TS$ where

- $E = \sum_v P(V)H(V)$ and $S = -\sum_v P(V) \log P(V)$.

$E$ only depends on the two-body probabilities, and so does $S$ if $P$ is a bifactor distribution on a tree (which includes Gibbs distributions)

- $E = \sum_{\langle u,v \rangle} P(u,v)(k_{uv} + h_u + h_v) + \sum_u (1 - d_v)P(u)h_u$
- $S = -\sum_{\langle u,v \rangle} P(u,v) \log P(u,v) - \sum_u (1 - d_v)P(u) \log P(u)$.

$G = \sum_{\langle u,v \rangle} P(u,v)(T \log P(u,v) + E_{uv}) + \sum_u (1 - d_u)P(u)(T \log P(u) + h_u)$.

The Gibbs distribution (bifactor) $P(V) = \frac{1}{Z} e^{-\beta H}$ is the stationary point of $G$.

The Bethe free energy is extending this expression to arbitrary graphs

$G_{Bethe} = \sum_{\langle u,v \rangle} b(u,v)(T \log b(u,v) + E_{uv}) + \sum_u (1 - d_u)b(u)(T \log b(u) + h_u)$. 
Let $H = \sum_{v} h_v + \sum_{\langle u, v \rangle} k_{uv}$ be a local Hamiltonian on $G$. Given a probability $P(V)$, the Gibbs free energy $G = E - TS$ where
\[
E = \sum_{V} P(V)H(V) \quad \text{and} \quad S = - \sum_{V} P(V) \log P(V).
\]
$E$ only depends on the two-body probabilities, and so does $S$ if $P$ is a bifactor distribution on a tree (which includes Gibbs distributions)
\[
E = \sum_{\langle u, v \rangle} P(u, v)(k_{uv} + h_u + h_v) + \sum_{u}(1 - d_v)P(u)h_u
\]
\[
S = - \sum_{\langle u, v \rangle} P(u, v) \log P(u, v) - \sum_{u}(1 - d_v)P(u) \log P(u).
\]
$G = \sum_{\langle u, v \rangle} P(u, v)(T \log P(u, v) + E_{uv}) + \sum_{u}(1 - d_u)P(u)(T \log P(u) + h_u)$.

The Gibbs distribution (bifactor) $P(V) = \frac{1}{Z} e^{-\beta H}$ is the stationary point of $G$.

The Bethe free energy is extending this expression to arbitrary graphs
\[
G_{\text{Bethe}} = \sum_{\langle u, v \rangle} b(u, v)(T \log b(u, v) + E_{uv}) + \sum_{u}(1 - d_u)b(u)(T \log b(u) + h_u).
\]
Belief propagation

**Bethe free energy**

Let \( H = \sum_v h_v + \sum_{\langle u, v \rangle} k_{uv} \) be a local Hamiltonian on \( G \). Given a probability \( P(V) \), the Gibbs free energy \( G = E - TS \) where

\[
E = \sum_v P(V) H(V) \quad \text{and} \quad S = -\sum_v P(V) \log P(V).
\]

\( E \) only depends on the two-body probabilities, and so does \( S \) if \( P \) is a bifactor distribution on a tree (which includes Gibbs distributions)

\[
E = \sum_{\langle u, v \rangle} P(u, v)(k_{uv} + h_u + h_v) + \sum_u (1 - d_v) P(u) h_u
\]

\[
S = -\sum_{\langle u, v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v) P(u) \log P(u).
\]

\[
G = \sum_{\langle u, v \rangle} P(u, v)(T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u) P(u)(T \log P(u) + h_u).
\]

The Gibbs distribution (bifactor) \( P(V) = \frac{1}{Z} e^{-\beta H} \) is the stationary point of \( G \).

The Bethe free energy is extending this expression to arbitrary graphs

\[
G_{\text{Bethe}} = \sum_{\langle u, v \rangle} b(u, v)(T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u) b(u)(T \log b(u) + h_u).
\]
Let $H = \sum_v h_v + \sum_{\langle u,v \rangle} k_{uv}$ be a local Hamiltonian on $G$. Given a probability $P(V)$, the Gibbs free energy $G = E - TS$ where

- $E = \sum_v P(V) H(V)$ and $S = -\sum_v P(V) \log P(V)$.

$E$ only depends on the two-body probabilities, and so does $S$ if $P$ is a bifactor distribution on a tree (which includes Gibbs distributions)

- $E = \sum_{\langle u,v \rangle} P(u, v)(k_{uv} + h_u + h_v) + \sum_u (1 - d_v) P(u) h_u$
- $S = -\sum_{\langle u,v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v) P(u) \log P(u)$.

$G = \sum_{\langle u,v \rangle} P(u, v)(T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u) P(u)(T \log P(u) + h_u)$.

The Gibbs distribution (bifactor) $P(V) = \frac{1}{Z} e^{-\beta H}$ is the stationary point of $G$.

The Bethe free energy is extending this expression to arbitrary graphs

$G_{\text{Bethe}} = \sum_{\langle u,v \rangle} b(u, v)(T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u) b(u)(T \log b(u) + h_u)$.
Belief propagation

Bethe free energy

Let $H = \sum_v h_v + \sum_{u,v} k_{uv}$ be a local Hamiltonian on $G$. Given a probability $P(V)$, the Gibbs free energy $G = E - TS$ where

- $E = \sum_v P(V)H(V)$ and $S = -\sum_v P(V) \log P(V)$.

$E$ only depends on the two-body probabilities, and so does $S$ if $P$ is a bifactor distribution on a tree (which includes Gibbs distributions)

- $E = \sum_{\langle u,v \rangle} P(u, v)(k_{uv} + h_u + h_v) + \sum_u (1 - d_v)P(u)h_u$
- $S = -\sum_{\langle u,v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v)P(u) \log P(u)$.

$$G = \sum_{\langle u,v \rangle} P(u, v)(T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u)P(u)(T \log P(u) + h_u).$$

The Gibbs distribution (bifactor) $P(V) = \frac{1}{Z} e^{-\beta H}$ is the stationary point of $G$.

The Bethe free energy is extending this expression to arbitrary graphs

$$G_{Bethe} = \sum_{\langle u,v \rangle} b(u, v)(T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u)b(u)(T \log b(u) + h_u).$$
Let $H = \sum_v h_v + \sum_{\langle u,v \rangle} k_{uv}$ be a local Hamiltonian on $G$. Given a probability $P(V)$, the Gibbs free energy $G = E - TS$ where

$$
E = \sum_V P(V) H(V) \quad \text{and} \quad S = - \sum_V P(V) \log P(V).
$$

$E$ only depends on the two-body probabilities, and so does $S$ if $P$ is a bifactor distribution on a tree (which includes Gibbs distributions)

$$
E = \sum_{\langle u,v \rangle} P(u, v) (k_{uv} + h_u + h_v) + \sum_u (1 - d_v) P(u) h_u
$$

$$
S = - \sum_{\langle u,v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v) P(u) \log P(u).
$$

$$
G = \sum_{\langle u,v \rangle} P(u, v) (T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u) P(u) (T \log P(u) + h_u).
$$

The Gibbs distribution (bifactor) $P(V) = \frac{1}{Z} e^{-\beta H}$ is the stationary point of $G$.

The Bethe free energy is extending this expression to arbitrary graphs

$$
G_{\text{Bethe}} = \sum_{\langle u,v \rangle} b(u, v) (T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u) b(u) (T \log b(u) + h_u).
$$
Let \( H = \sum_v h_v + \sum_{\langle u,v \rangle} k_{uv} \) be a local Hamiltonian on \( G \).

Given a probability \( P(V) \), the Gibbs free energy \( G = E - TS \) where

\[
E = \sum_v P(V) H(V) \quad \text{and} \quad S = -\sum_v P(V) \log P(V).
\]

\( E \) only depends on the two-body probabilities, and so does \( S \) if \( P \) is a bifactor distribution on a tree (which includes Gibbs distributions)

\[
E = \sum_{\langle u,v \rangle} P(u, v)(k_{uv} + h_u + h_v) + \sum_u (1 - d_v) P(u) h_u
\]

\[
S = -\sum_{\langle u,v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v) P(u) \log P(u).
\]

\[
G = \sum_{\langle u,v \rangle} P(u, v)(T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u) P(u)(T \log P(u) + h_u).
\]

The Gibbs distribution (bifactor) \( P(V) = \frac{1}{Z} e^{-\beta H} \) is the stationary point of \( G \).

The Bethe free energy is extending this expression to arbitrary graphs

\[
G_{\text{Bethe}} = \sum_{\langle u,v \rangle} b(u, v)(T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u) b(u)(T \log b(u) + h_u).
\]
Bethe free energy

Let \( H = \sum_v h_v + \sum_{\langle u, v \rangle} k_{uv} \) be a local Hamiltonian on \( G \).

Given a probability \( P(V) \), the Gibbs free energy \( G = E - TS \) where

\[
E = \sum_{\langle u, v \rangle} P(u, v) H(V) \quad \text{and} \quad S = -\sum_{\langle u, v \rangle} P(u, v) \log P(u, v).
\]

\( E \) only depends on the two-body probabilities, and so does \( S \) if \( P \) is a bifactor distribution on a tree (which includes Gibbs distributions)

\[
E = \sum_{\langle u, v \rangle} P(u, v) (k_{uv} + h_u + h_v) + \sum_u (1 - d_v) P(u) h_u
\]

\[
S = -\sum_{\langle u, v \rangle} P(u, v) \log P(u, v) - \sum_u (1 - d_v) P(u) \log P(u).
\]

\[
G = \sum_{\langle u, v \rangle} P(u, v) (T \log P(u, v) + E_{uv}) + \sum_u (1 - d_u) P(u) (T \log P(u) + h_u).
\]

The Gibbs distribution (bifactor) \( P(V) = \frac{1}{Z} e^{-\beta H} \) is the stationary point of \( G \).

The Bethe free energy is extending this expression to arbitrary graphs

\[
G_{\text{Bethe}} = \sum_{\langle u, v \rangle} b(u, v) (T \log b(u, v) + E_{uv}) + \sum_u (1 - d_u) b(u) (T \log b(u) + h_u).
\]
Belief propagation

Bethe free energy

Theorem (Yedidia, Freeman, and Weiss)

The fixed point of the belief propagation algorithms \( b(u, v) \) and \( b(u) \) are stationary points of the Bethe free energy.

- Exact on tree, complexity = depth(\( G \)).
- May not converge in general, but if it does it corresponds to a Bethe approximation.
- Most successful on trees with only large loops:
  - LDPC and Turbo codes.
  - Spin glasses on Bethe lattices.
  - Random k-SAT assignments.
Bethe free energy

Theorem (Yedidia, Freeman, and Weiss)

The fixed point of the belief propagation algorithms $b(u, v)$ and $b(u)$ are stationary points of the Bethe free energy.

- Exact on tree, complexity $= \text{depth}(G)$.
- May not converge in general, but if it does it corresponds to a Bethe approximation.
- Most successful on trees with only large loops:
  - LDPC and Turbo codes.
  - Spin glasses on Bethe lattices.
  - Random $k$-SAT assignments.
Belief propagation

Bethe free energy

Theorem (Yedidia, Freeman, and Weiss)

The fixed point of the belief propagation algorithms $b(u, v)$ and $b(u)$ are stationary points of the Bethe free energy.

- Exact on tree, complexity = $\text{depth}(G)$.
- May not converge in general, but if it does it corresponds to a Bethe approximation.
- Most successful on trees with only large loops:
  - LDPC and Turbo codes.
  - Spin glasses on Bethe lattices.
  - Random $k$-SAT assignments.
Theorem (Yedidia, Freeman, and Weiss)

The fixed point of the belief propagation algorithms $b(u,v)$ and $b(u)$ are stationary points of the Bethe free energy.

- Exact on tree, complexity $= \text{depth}(G)$.
- May not converge in general, but if it does it corresponds to a Bethe approximation.
- Most successful on trees with only large loops:
  - LDPC and Turbo codes.
  - Spin glasses on Bethe lattices.
  - Random $k$-SAT assignments.
Theorem (Yedidia, Freeman, and Weiss)

The fixed point of the belief propagation algorithms $b(u, v)$ and $b(u)$ are stationary points of the Bethe free energy.

- Exact on tree, complexity $= \text{depth}(G)$.
- May not converge in general, but if it does it corresponds to a Bethe approximation.
- Most successful on trees with only large loops:
  - LDPC and Turbo codes.
  - Spin glasses on Bethe lattices.
  - Random $k$-SAT assignments.
Outline

1. Graphical models
2. Belief propagation
3. Quantum graphical models
4. Quantum belief propagation
5. Examples
   - Many-body simulations
   - Quantum turbo-codes
Bifactor state: $P(V) = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v)$.

Quantum generalization: $\rho_V = \frac{1}{Z} \prod_{v \in V} \mu_v \prod_{(u,v) \in E} \nu_{u:v}$.

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

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).
- $n = \infty$: $A \odot B = \exp(\log A + \log B)$ (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: \( \rho_V = \frac{1}{Z} \prod_{v \in V} \mu(v) \prod_{(u,v) \in E} \nu(u : v) \).

Problems:

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

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

- \( n = 1 \): \( A \star B = A^{\frac{1}{2}} B A^{\frac{1}{2}} \) (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: \( \rho_V = \frac{1}{Z} \prod_{v \in V} \mu_v \prod_{(u,v) \in E} \nu_{u:v}. \)

Problems:

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

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

- \( n = 1: A \star B = A^\frac{1}{2} B A^\frac{1}{2} \) (measurement, QEC).
- \( n = \infty: A \odot B = \exp(\log A + \log B) \) (Hamiltonian, many-body).
- Intermediate \( n \): Trotter-Suzuki decomposition.
Non-commutative generalization

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

Quantum generalization: \( \rho_V = \frac{1}{Z} \prod_{v \in V} \mu_v \prod_{(u,v) \in E} \nu_{u:v} \).

Problems:

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

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).
- \( n = \infty \): \( A \bigcirc B = \exp(\log A + \log B) \) (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: \( \rho_V = \frac{1}{Z} \prod_{v \in V} \mu v \prod_{(u,v) \in E} \nu u : v \).

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

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: \( \rho_V = \frac{1}{Z} \prod_{v \in V} \mu_V \prod_{(u,v) \in E} \nu_{u:v} \).

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

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

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

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

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).
- $n = \infty$: $A \odot B = \exp(\log A + \log B)$ (Hamiltonian, many-body).
- Intermediate $n$: Trotter-Suzuki decomposition.
Quantum generalisations

In analogy with the classical case, define

- **Conditional state** \( \rho_A^{(n)}_{|B} = \rho_B^{-1} \ast^{(n)} \rho_{AB} \).
- **Mutual state** \( \rho_{A:B}^{(n)} = (\rho_A^{-1} \rho_B^{-1}) \ast^{(n)} \rho_{AB} \).

A quantum Markov Network is defined as in the classical case, with the von Neuman entropy substituting the Shannon entropy:

\[
I(U : V - n(U) - U|n(U)) = 0
\]
In analogy with the classical case, define

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

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

A quantum Markov Network is defined as in the classical case, with the von Neuman entropy substituting the Shannon entropy:

\[
I(U : V - n(U) - U|n(U)) = 0
\]
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:

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

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

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

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

$$\rho_{ABC} = (\rho_A \rho_B \rho_C) \ast^{(n)} (\rho_{A:B}^{(n)} \rho_{B:C}^{(n)})$$ 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.
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:

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

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

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

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.
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

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

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

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

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

Cut and paste from previous section.
Don’t forget to search for $\prod$ and replace by $\star^{(n)}$. 
Let $G = (V, E)$ be a graph and let

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

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

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

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 $\text{depth}(G)$. 

David Poulin (Caltech)
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 $\text{depth}(G)$. 
Outline

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

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

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

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

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

\[
\vdots
\]

\[
Z = \sum_{i_N} m_{i_{N-1} \rightarrow i_N} \mu(i_N)
\]
Consider the 1d \textbf{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 $(\mu(i) = e^{-\beta h_i}$ 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:

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

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

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

$$\vdots$$

$$Z = \sum_{i_N} m_{N-1 \rightarrow N} \mu(i_N)$$
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:

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

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

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

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1} \rightarrow i_N} \mu(i_N)$$
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:

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

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

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

$$\vdots$$

$$Z = \sum_{i_N} m_{i_{N-1} \rightarrow i_N} \mu(i_N)$$
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 $\nu_{i:j} = e^{-\beta J_{ij}})$.

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

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 **quantum** system with hamiltonian \( H = \sum_i h_i + \sum_{\langle ij \rangle} J_{ij} \). Its Gibbs distribution is \( (\mu_i = e^{-\beta h_i} \text{ and } \nu_{i:j} = e^{-\beta J_{ij}}) \)

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

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 \( I(A : C|B) = 0 \).
One dimensional quantum system

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 $I(A : C|B) = 0$. 
One dimensional quantum system

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 $I(A : C|B) = 0$. 

---

David Poulin (Caltech)  Quantum Belief Propagation  Coogee'08
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

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

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

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

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

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

\[ \vdots \]

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

\[
\begin{align*}
\sigma_{1-4} &= e^{-\beta(h_1 + h_2 + h_3 + h_4 + J_{12} + J_{23} + J_{34})} \\
\sigma'_{2-4} &= Tr_1\{\sigma_{1-4}\} \quad h'_{2-4} = -\frac{1}{\beta} \log \sigma'_{2-4} \\
\sigma_{2-5} &= e^{-\beta(h'_{2-4} + h_5 + J_{45})} \\
\sigma'_{3-5} &= Tr_2\{\sigma_{2-5}\} \quad h'_{3-5} = -\frac{1}{\beta} \log \sigma'_{3-5} \\
\sigma_{3-6} &= e^{-\beta(h'_{3-5} + h_6 + J_{56})} \\
& \vdots \\
Z &= Tr\{\sigma_{N-3,N-2,N-1,N}\}
\end{align*}
\]
One dimensional quantum system

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

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

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

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

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

\[ \vdots \]

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

\[ h'_{3-5} \]

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

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

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

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

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

\[ \vdots \]

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

\[ \sigma_{3-6} \]

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

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

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

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

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

\[
\vdots
\]

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

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

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

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

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

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

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

\[ \vdots \]

\[ Z = \text{Tr}\{\sigma_{N-3,N-2,N-1,N}\} \]
Critical 1d Ising model

\[ (\tilde{E} - E) \cdot 10^4 \]

\( \beta \)

- Replica
- TEBD
- Sliding window

Bilgin and Poulin '07.
Phase diagram

Ising spin glass on Cayley tree

Laumann, Scardicchio, and Sondhi ’07, Bilgin and Poulin.
Outline

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

Quantum turbo-codes

Rudiments

Basic block codes

Shor, Steane, 5-qubit, ...

After $\ell$ concatenations
- Remaining error $p_e \propto \epsilon^{2\ell}$.
- Rate $R = (k/n)^\ell$.

As block size $n$ increases
- Remaining error $p_e \propto \exp(-cn)$.
- Fixed rate $R = k/n$. 
### Basic block codes

**Shor, Steane, 5-qubit, ...**

**After $\ell$ concatenations**
- Remaining error $p_e \propto \epsilon^{2\ell}$.
- Rate $R = (k/n)^{\ell}$.

**As block size $n$ increases**
- Remaining error $p_e \propto \exp(-cn)$.
- Fixed rate $R = k/n$. 
Basic block codes

Shor, Steane, 5-qubit, ...

After \( \ell \) concatenations

- Remaining error \( p_e \propto \epsilon^{2\ell} \).
- Rate \( R = (k/n)^\ell \).

As block size \( n \) increases

- Remaining error \( p_e \propto \exp(-cn) \).
- Fixed rate \( R = k/n \).
Rudiments

Basic block codes

Shor, Steane, 5-qubit, ...

After \( \ell \) concatenations

- Remaining error \( p_e \propto \epsilon^{2\ell} \).
- Rate \( R = (k/n)^\ell \).

As block size \( n \) increases

- Remaining error \( p_e \propto \exp(-cn) \).
- Fixed rate \( R = k/n \).
A conflict between giants: Shannon vs Turing

**Information theory (Shannon)**

Random codes are optimal.
- Optimal transmission rate on symmetric classical channels.
- Not quite true of quantum codes, but pretty good.

**Computer science (Turing)**

Decoding a random code is NP-hard.
- Finding ground state of spin-glass.

**Compromise**

- Some randomness in the code design.
- Enough structure for *approximate* iterative decoding.
A conflict between giants: Shannon vs Turing

Information theory (Shannon)

Random codes are optimal.
- Optimal transmission rate on symmetric classical channels.
- Not quite true of quantum codes, but pretty good.

Computer science (Turing)

Decoding a random code is NP-hard.
- Finding ground state of spin-glass.

Compromise

- Some randomness in the code design.
- Enough structure for approximate iterative decoding.
# A conflict between giants: Shannon vs Turing

## Information theory (Shannon)

Random codes are optimal.
- Optimal transmission rate on symmetric classical channels.
- Not quite true of quantum codes, but pretty good.

## Computer science (Turing)

Decoding a random code is NP-hard.
- Finding ground state of spin-glass.

## Compromise

- Some randomness in the code design.
- Enough structure for *approximate* iterative decoding.
Interleaver provides randomness: good code.
Regular pattern of $U$ and $W$: good iterative decoding performances.
Complete control of code parameters: rate, size, memory.

Interleaver provides randomness: good code.

Regular pattern of $U$ and $W$: good iterative decoding performances.

Complete control of code parameters: rate, size, memory.

Interleaver provides randomness: good code.
Regular pattern of $U$ and $W$: good iterative decoding performances.
Complete control of code parameters: rate, size, memory.

Interleaver provides randomness: good code.
Regular pattern of $U$ and $W$: good iterative decoding performances.
Complete control of code parameters: rate, size, memory.

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.
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.
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.
Examples
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.
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.
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 $S_2$ is the Shannon limit if the correlations between errors and errors are neglected, (45). Points "*" are codes in vented 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.


MacKay, Mitchison, McFadden, IEEE’04.