Symbolwise GC Decoding: Connecting SPA Decoding and BFE Minimization

Pascal O. Vontobel
Information Theory Research Group
Hewlett-Packard Laboratories Palo Alto

46th Allerton Conference, Monticello, IL, September 26, 2008

Overview of Talk

Bethe free energy and . . .

- symbolwise graph-cover decoding,
- EXIT charts,
- asymptotic growth rate of the average Hamming weight distribution for code ensembles.

Symbolwise graph-cover decoding

Communication Model (Part 1)

Information word: $\mathbf{u} = (u_1, \ldots, u_k) \in \mathcal{U}^k$

Sent codeword: $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{C} \subseteq \mathcal{X}^n$

Received word: $\mathbf{y} = (y_1, \ldots, y_n) \in \mathcal{Y}^n$

Communication Model (Part 1)

Information word: $\mathbf{u} = (u_1, \ldots, u_k) \in \mathcal{U}^k$

Sent codeword: $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{C} \subseteq \mathcal{X}^n$

Received word: $\mathbf{y} = (y_1, \ldots, y_n) \in \mathcal{Y}^n$

Decoding: Based on $\mathbf{y}$ we would like to estimate the transmitted codeword $\hat{\mathbf{x}}$ or the information word $\hat{\mathbf{u}}$.

Communication Model (Part 1)

Information word: $\mathbf{u} = (u_1, \ldots, u_k) \in \mathcal{U}^k$

Sent codeword: $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{C} \subseteq \mathcal{X}^n$

Received word: $\mathbf{y} = (y_1, \ldots, y_n) \in \mathcal{Y}^n$

Decoding: Based on $\mathbf{y}$ we would like to estimate the transmitted codeword $\hat{\mathbf{x}}$ or the information word $\hat{\mathbf{u}}$.

Depending on what criterion we optimize, we obtain different decoding algorithms.

Symbolwise MAP/ML Decoding

Symbolwise MAP/ML Decoding

Minimizing the symbol error probability (for each $i = 1, \ldots, k$) results in symbol-wise MAP decoding.

For each $i = 1, \ldots, k$:

$$\hat{u}_i^{\text{symbol}}(\mathbf{y}) = \underset{u_i \in \mathcal{U}}{\operatorname{argmax}} \; P_{U_i|\mathbf{Y}}(u_i|\mathbf{y}) = \underset{u_i \in \mathcal{U}}{\operatorname{argmax}} \; P_{U_i,\mathbf{Y}}(u_i, \mathbf{y})$$

Symbolwise MAP/ML Decoding

Minimizing the symbol error probability (for each $i = 1, \ldots, k$) results in symbol-wise MAP decoding.

For each $i = 1, \ldots, k$:

$$\hat{u}_i^{\text{symbol}}(\mathbf{y}) = \operatorname*{argmax}_{u_i \in \mathcal{U}} P_{U_i|\mathbf{Y}}(u_i|\mathbf{y}) = \operatorname*{argmax}_{u_i \in \mathcal{U}} P_{U_i,\mathbf{Y}}(u_i, \mathbf{y})$$

$$= \operatorname*{argmax}_{u_i \in \mathcal{U}} \sum_{\substack{\mathbf{u} \in \mathcal{U}^k, \; \mathbf{x} \in \mathcal{X}^n \\ u_i \text{ fixed}}} P_{\mathbf{UXY}}(\mathbf{u}, \mathbf{x}, \mathbf{y}) \; .$$

Symbolwise MAP/ML Decoding

For simplicity,

Symbolwise MAP/ML Decoding

For simplicity,

- the encoding will be systematic, i.e.,

$$x_i = u_i \quad (\text{for } i = 1, \ldots, k)\ ;$$

Symbolwise MAP/ML Decoding

For simplicity,

- the encoding will be systematic, i.e.,

$$x_i = u_i \quad (\text{for } i = 1, \ldots, k) \; ;$$

- the channel input alphabet is $\mathcal{X} = \{0, 1\}$;

Symbolwise MAP/ML Decoding

For simplicity,

- the encoding will be systematic, i.e.,

$$x_i = u_i \quad (\text{for } i = 1, \ldots, k) \; ;$$

- the channel input alphabet is $\mathcal{X} = \{0, 1\}$;
- and all codewords in the code $\mathcal{C}$ are assumed to be equally likely a priori, i.e.,

$$P_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2^k} \cdot [\mathbf{x} \in \mathcal{C}] \; .$$

Symbolwise MAP/ML Decoding

Symbolwise MAP/ML Decoding

Therefore, computing for each $i = 1, \ldots, n$ the marginals

$$\eta_i(0) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 0}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) ,$$

$$\eta_i(1) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 1}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) ,$$

Symbolwise MAP/ML Decoding

Therefore, computing for each $i = 1, \ldots, n$ the marginals

$$\eta_i(0) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 0}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \,,$$

$$\eta_i(1) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 1}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \,,$$

we obtain

$$\hat{x}_i^{\text{symbol}}(\mathbf{y}) = \operatorname*{argmax}_{x_i \in \mathcal{X}} \eta_i(x_i) \,.$$

Symbolwise MAP/ML Decoding

Therefore, computing for each $i = 1, \ldots, n$ the marginals

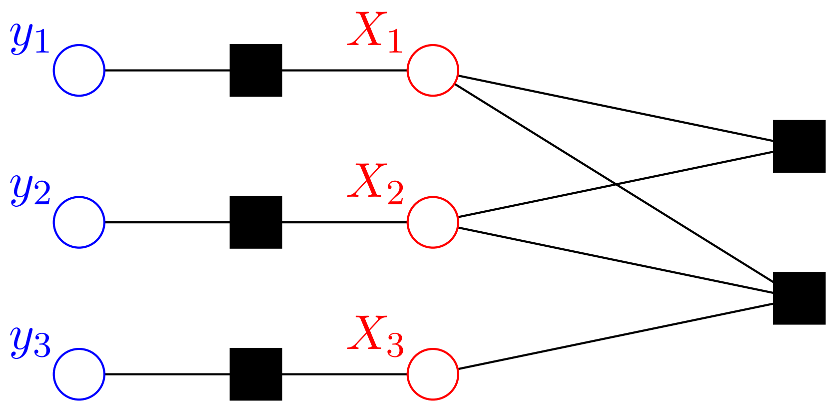
$$\eta_i(0) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 0}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \ ,$$

$$\eta_i(1) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 1}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \ ,$$

we obtain

$$\hat{x}_i^{\text{symbol}}(\mathbf{y}) = \operatorname*{argmax}_{x_i \in \mathcal{X}} \eta_i(x_i) \ .$$

Assume that the joint pmf of $\mathbf{X}$ and $\mathbf{Y} = \mathbf{y}$ is given by the following factor graph:

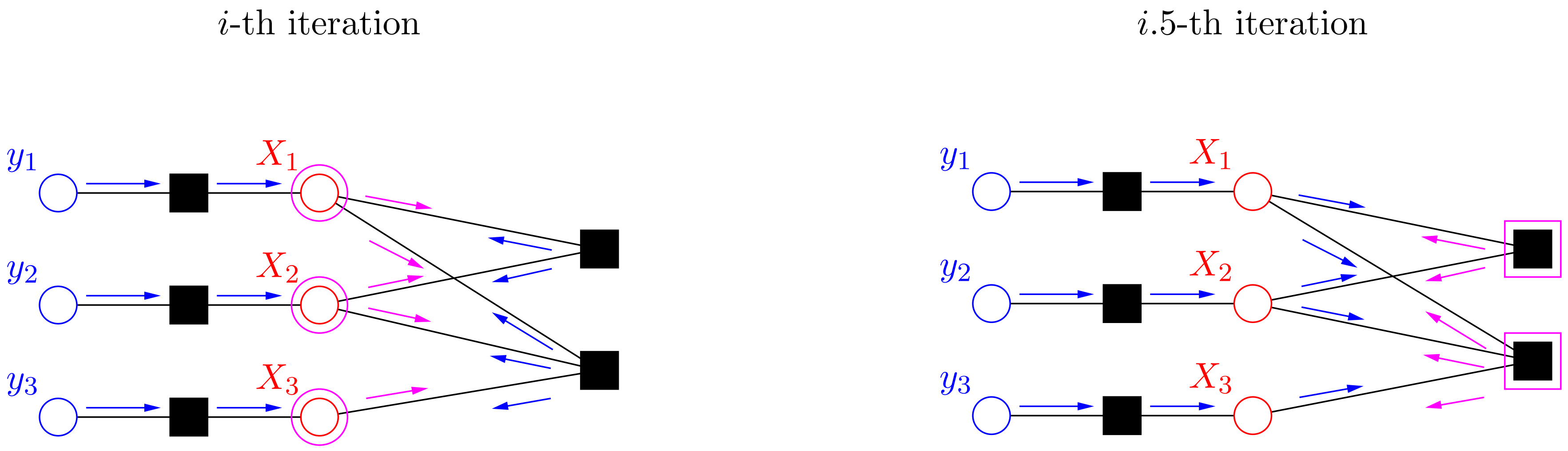


Sum-Product Algorithm Decoding

Sum-Product Algorithm Decoding

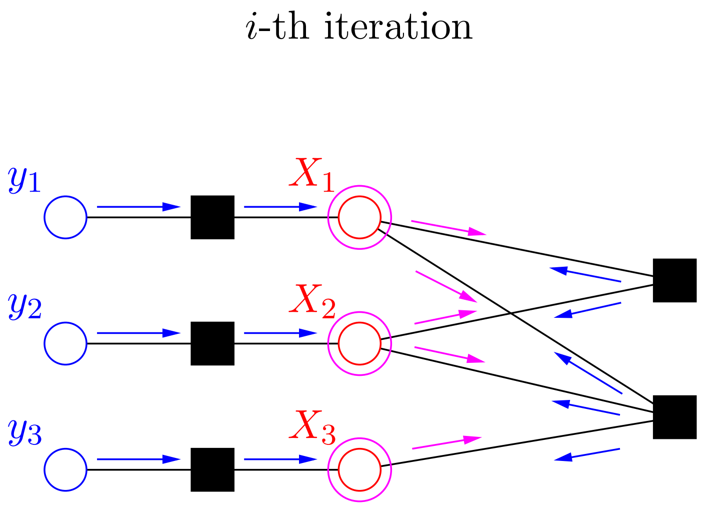
Computing these marginals is computationally intractable in general, therefore we can try to use sub-optimal algorithms like the sum-product algorithm (SPA) to obtain approximations to these marginals.

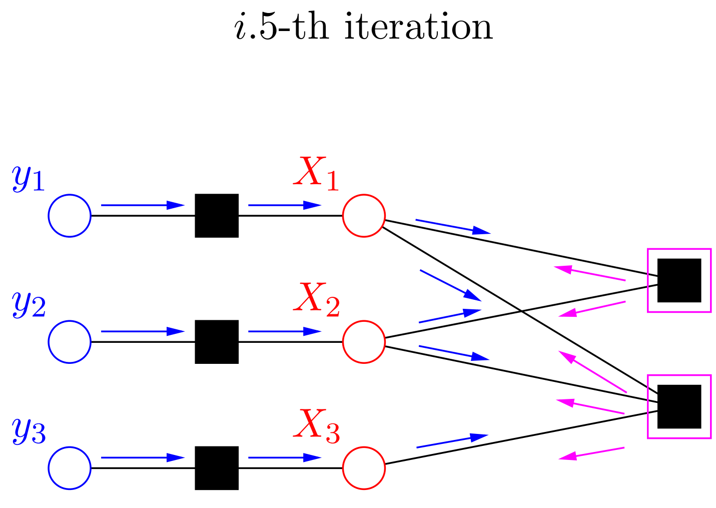
Sum-Product Algorithm Decoding



Computing these marginals is computationally intractable in general, therefore we can try to use sub-optimal algorithms like the sum-product algorithm (SPA) to obtain approximations to these marginals.

Sum-Product Algorithm Decoding





The SPA is an algorithm that operates locally on a factor graph:

- it sends messages along the edges,
- combines local messages to produce new local messages at the vertices.

Symbolwise MAP/ML Decoding

Therefore, computing for each $i = 1, \ldots, n$ the marginals

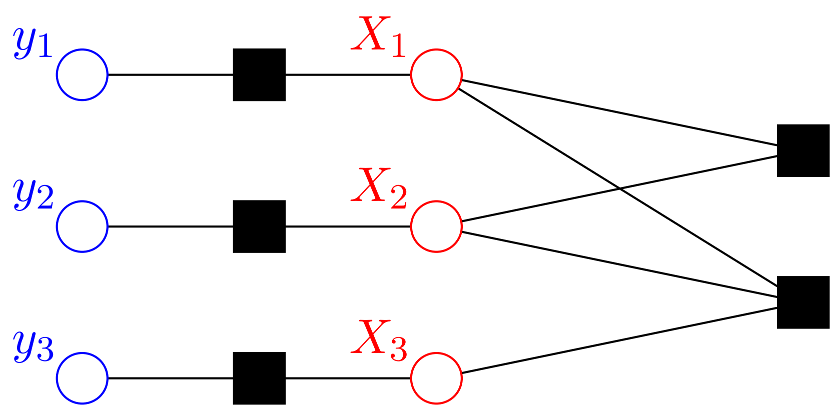
$$\eta_i(0) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 0}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \,,$$

$$\eta_i(1) \triangleq \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ x_i = 1}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \,,$$

we obtain

$$\hat{x}_i^{\text{symbol}}(\mathbf{y}) = \underset{x_i \in \mathcal{X}}{\text{argmax}}\ \eta_i(x_i) \,.$$

Assume that the joint pmf of $\mathbf{X}$ and $\mathbf{Y} = \mathbf{y}$ is given by the following factor graph:

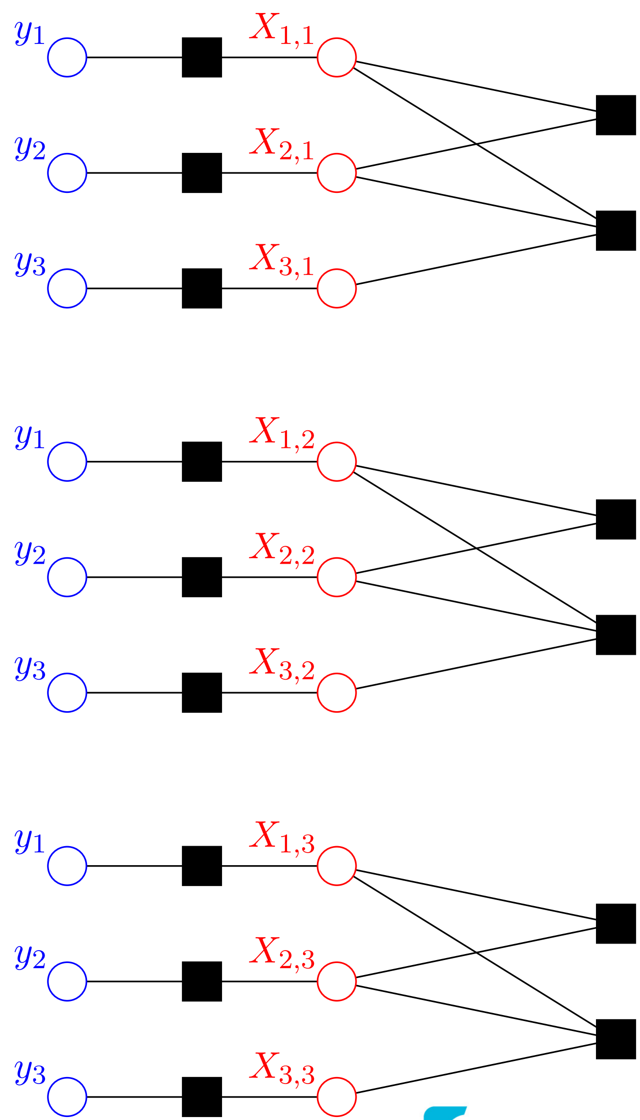


Towards Symbolwise Graph-Cover Decoding

Let us computer these marginal M times in parallel with repeated received vector $\mathbf{y}$. (Here $M = 3$.)

Towards Symbolwise Graph-Cover Decoding

Let us computer these marginal M times in parallel with repeated received vector $\mathbf{y}$. (Here $M = 3$.)

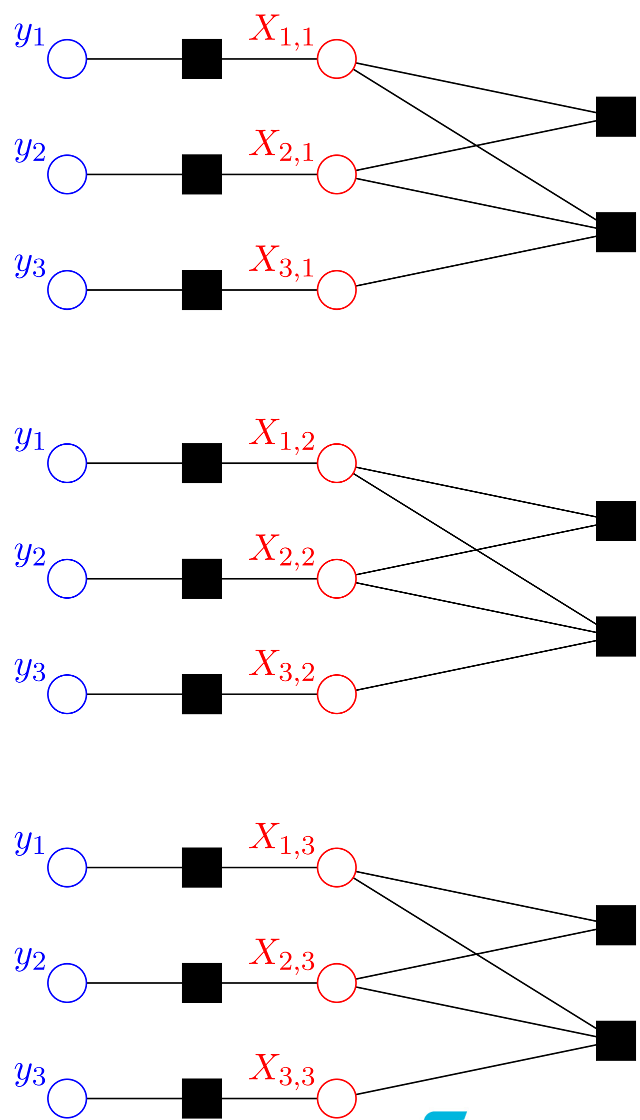


Towards Symbolwise Graph-Cover Decoding

Let us computer these marginal M times in parallel with repeated received vector $\mathbf{y}$. (Here $M = 3$.)

Therefore, for each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$ we compute the marginals

$$\eta_{i,m}(\tilde{x}_{i,m}) \triangleq \sum_{\substack{\tilde{\mathbf{x}} \in \tilde{\mathsf{T}} \\ \tilde{x}_{i,m} \text{ fixed}}} P_{\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}}(\mathbf{y}^{\uparrow M} | \tilde{\mathbf{x}})$$



Towards Symbolwise Graph-Cover Decoding

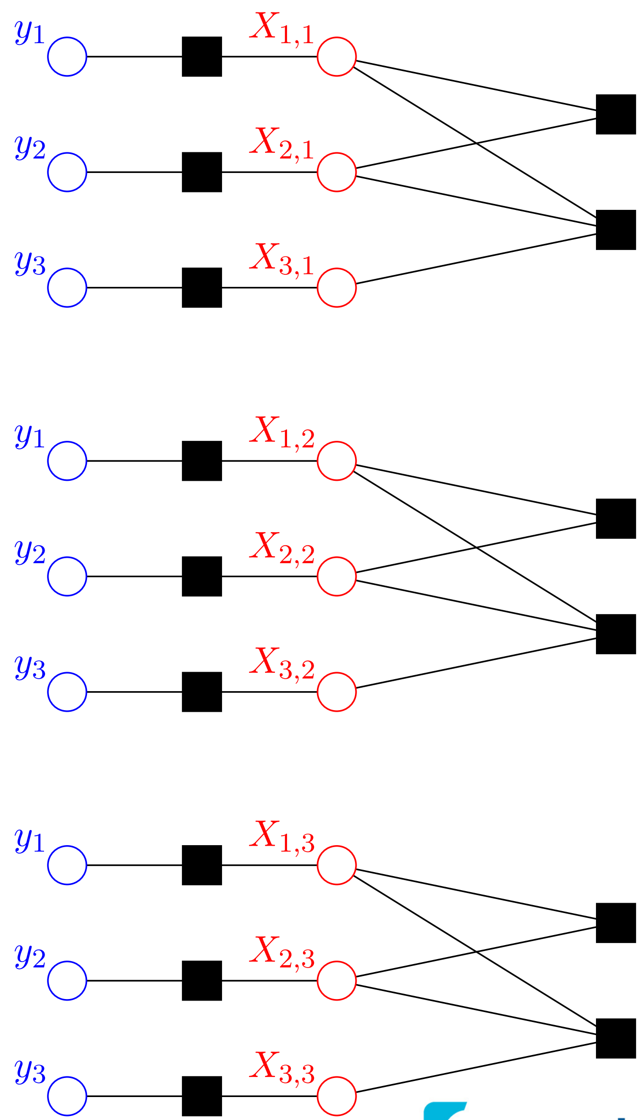
Let us computer these marginal M times in parallel with repeated received vector $\mathbf{y}$. (Here $M = 3$.)

Therefore, for each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$ we compute the marginals

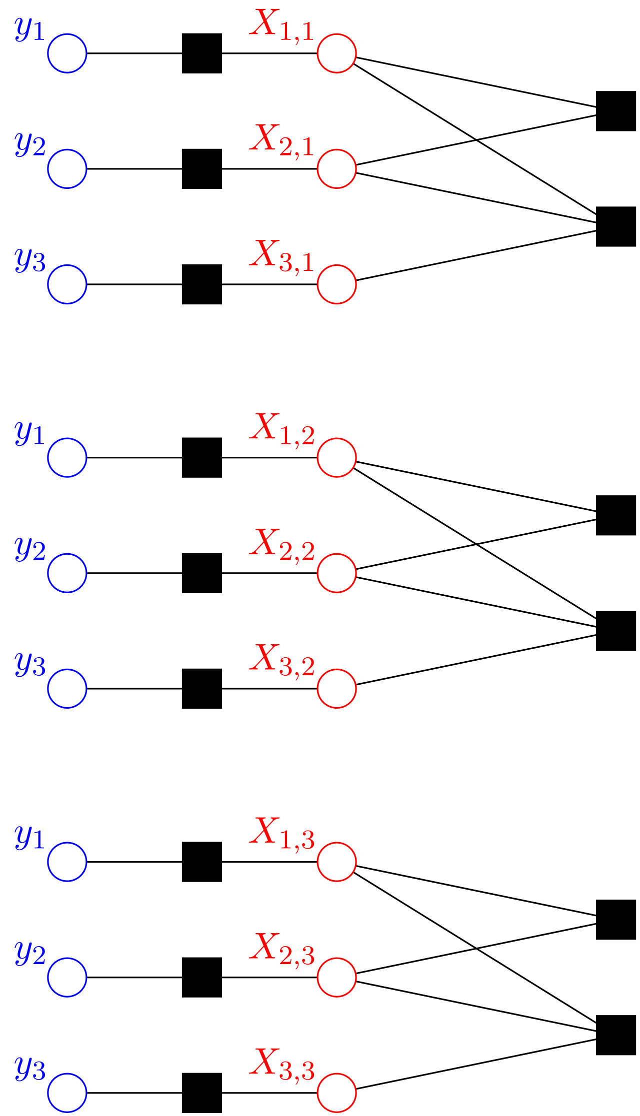
$$\eta_{i,m}(\tilde{x}_{i,m}) \triangleq \sum_{\substack{\tilde{\mathbf{x}} \in \tilde{\mathsf{T}} \\ \tilde{x}_{i,m} \text{ fixed}}} P_{\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}}(\mathbf{y}^{\uparrow M} | \tilde{\mathbf{x}})$$

and we obtain

$$\hat{x}^{\text{symbol}}_{i,m}(\mathbf{y}) = \underset{\tilde{x}_{i,m} \in \mathcal{X}}{\operatorname{argmax}} \ \eta_{i,m}(\tilde{x}_{i,m}) \ .$$

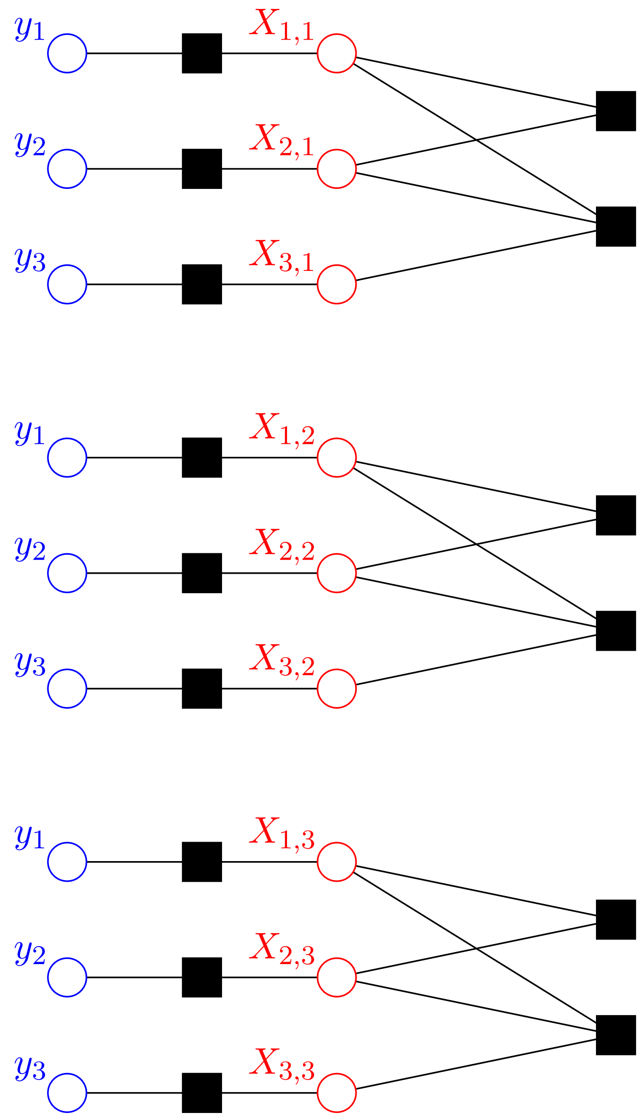


Towards Symbolwise Graph-Cover Decoding



Towards
Symbolwise Graph-Cover Decoding

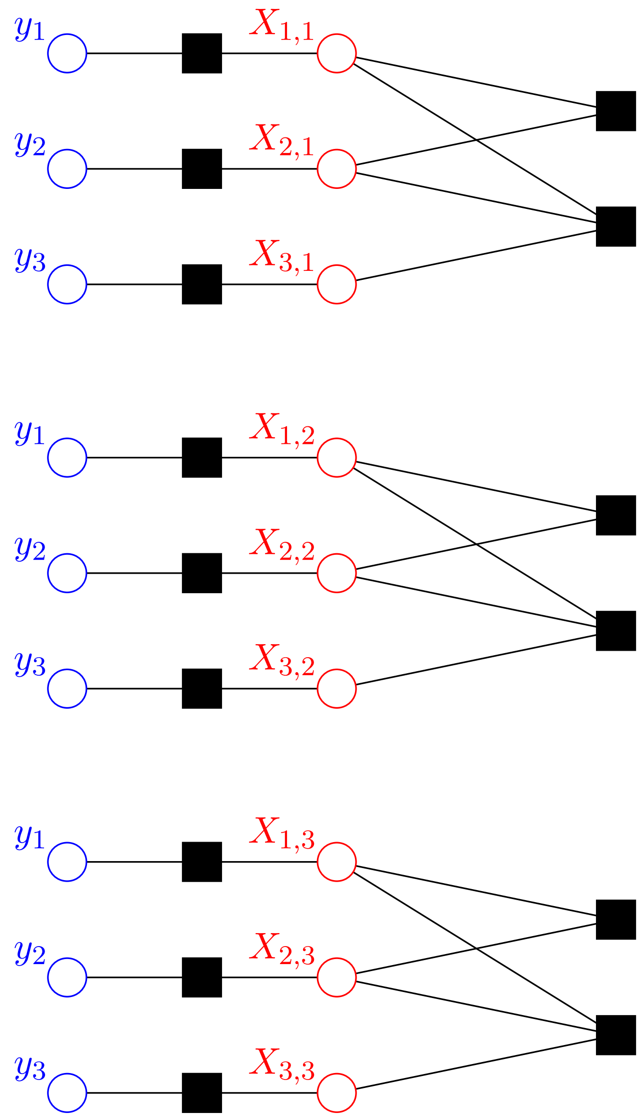
Because the processing is done locally, the SPA cannot distinguish if it is decoding



Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

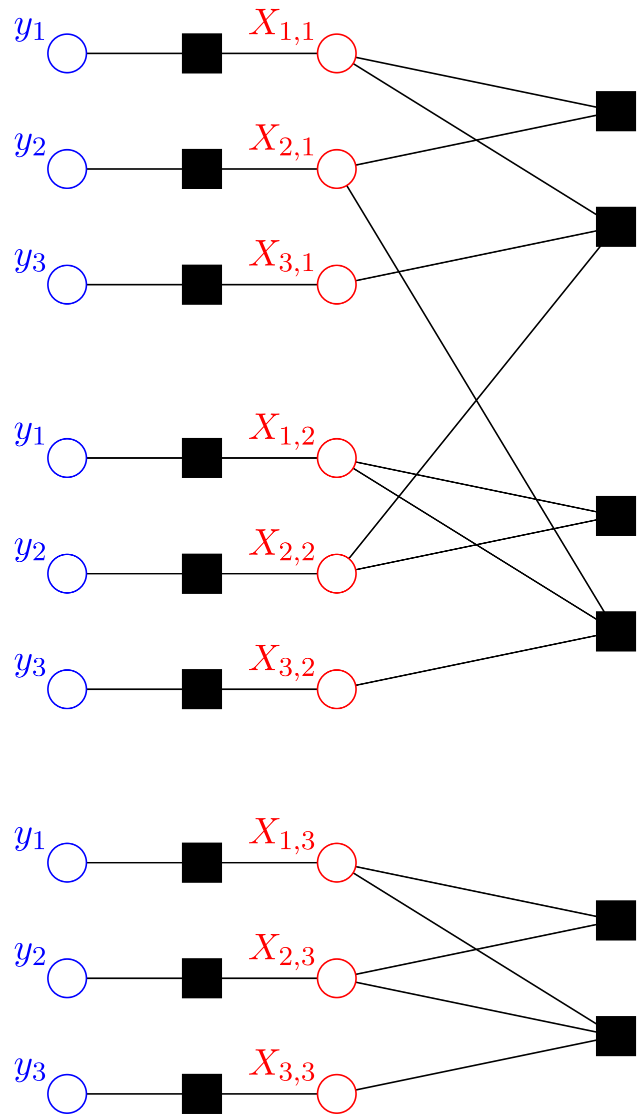
- this overall factor graph,



Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

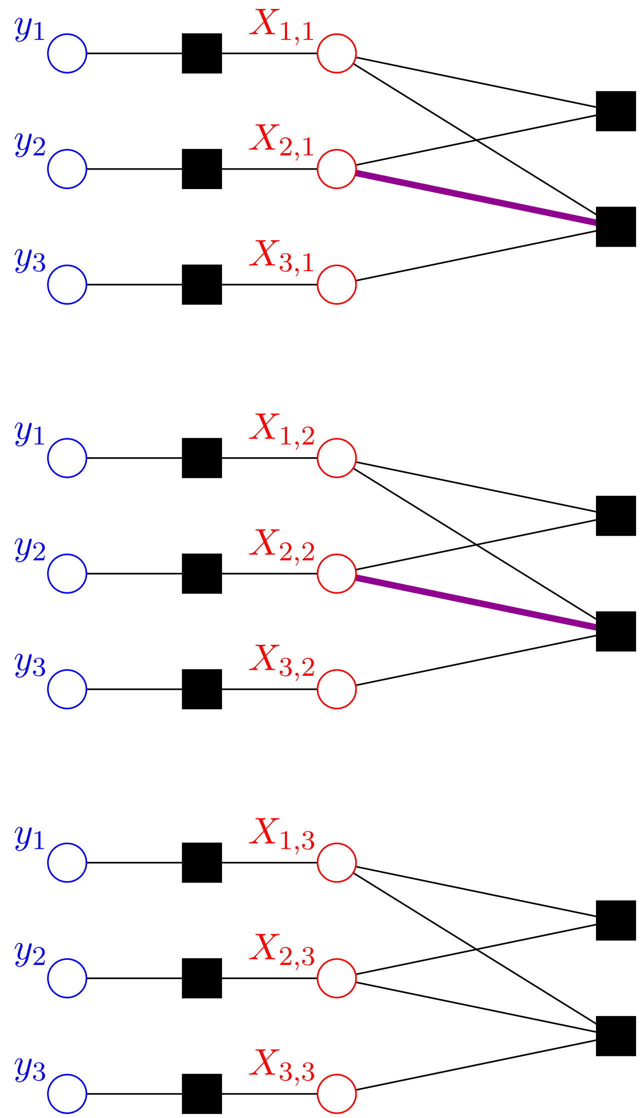
- this overall factor graph,



Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

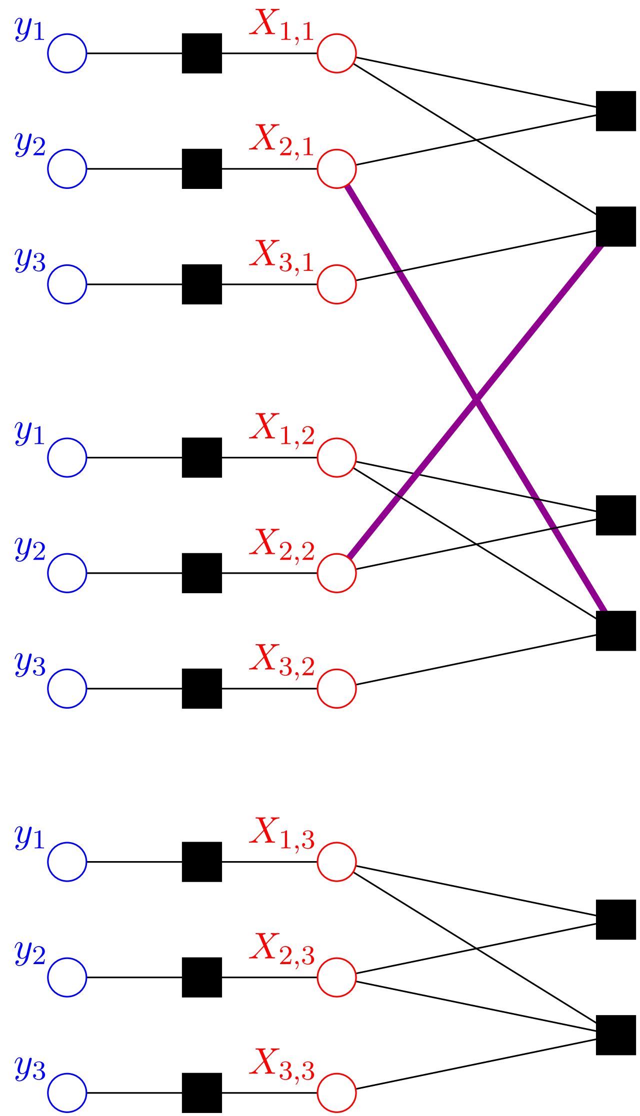
- this overall factor graph,



Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

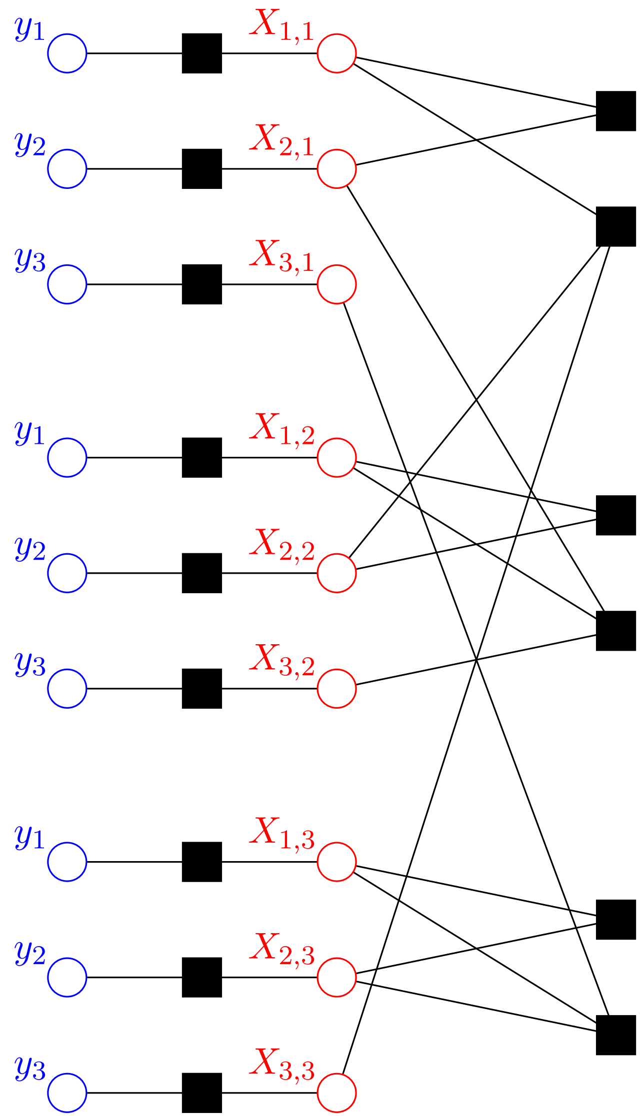
- this overall factor graph,
- or this overall factor graph,



Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

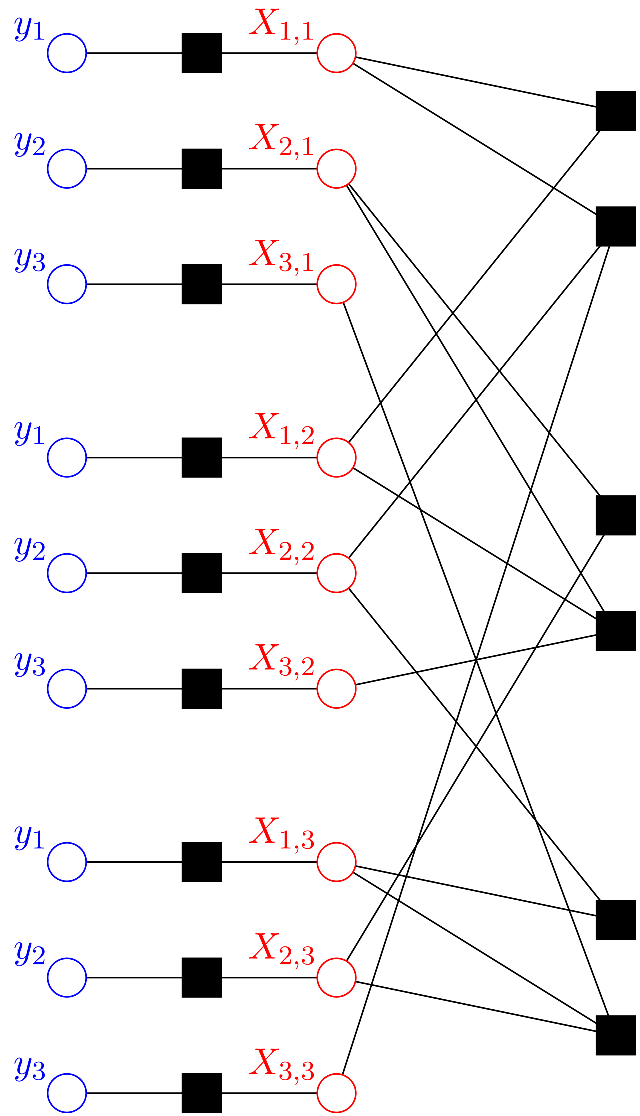
- this overall factor graph,
- or this overall factor graph,
- or this overall factor graph,



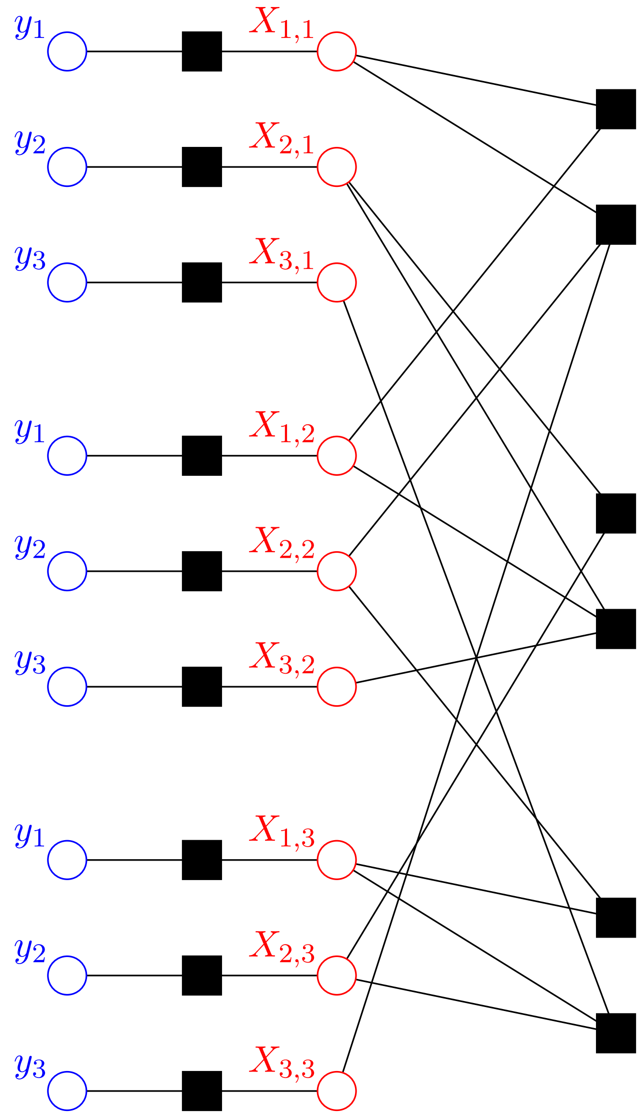
Towards
Symbolwise Graph-Cover Decoding

Because the processing is done locally, the SPA cannot distinguish if it is decoding

- this overall factor graph,
- or this overall factor graph,
- or this overall factor graph,
- or, in fact, any M-fold cover of the original factor graph.



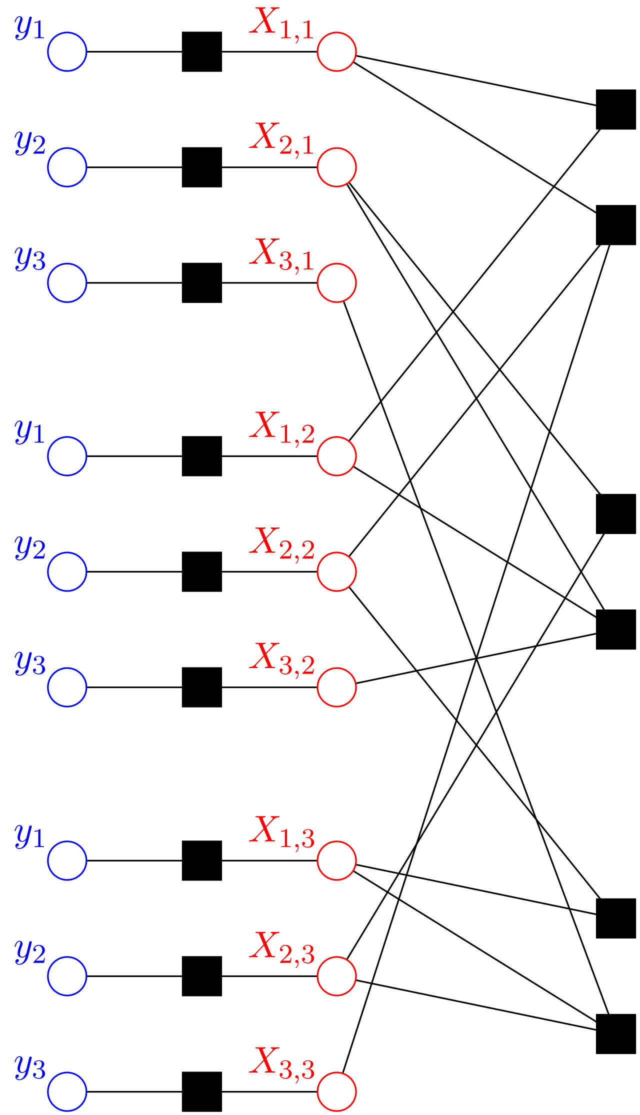
Symbolwise Graph-Cover Decoding



Symbolwise Graph-Cover Decoding

For an M-fold cover $\widetilde{\mathsf{T}}$ we define

$$\eta_{i,m,\widetilde{\mathsf{T}}}(\tilde{x}_{i,m}) \triangleq \sum_{\substack{\widetilde{\mathbf{x}} \in \mathcal{C}(\widetilde{\mathsf{T}}) \\ \tilde{x}_{i,m} \text{ fixed}}} P_{\widetilde{\mathbf{Y}}|\widetilde{\mathbf{X}}}(\mathbf{y}^{\uparrow M} | \widetilde{\mathbf{x}}) \ .$$



Symbolwise Graph-Cover Decoding

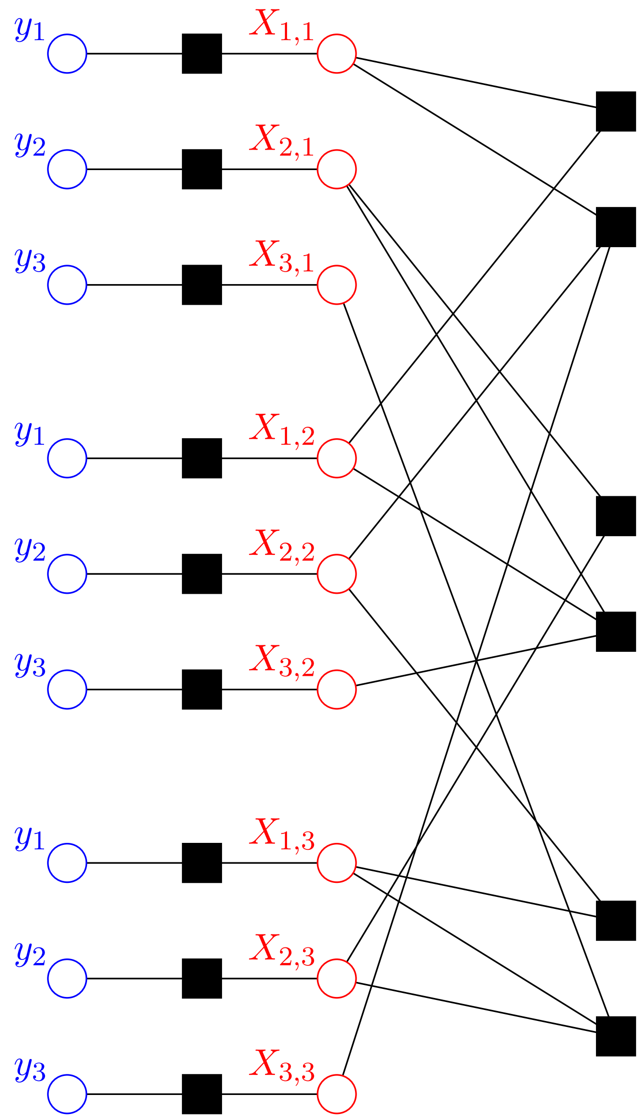
For an M-fold cover $\widetilde{\mathsf{T}}$ we define

$$\eta_{i,m,\widetilde{\mathsf{T}}}(\tilde{x}_{i,m}) \triangleq \sum_{\substack{\widetilde{\mathbf{x}} \in \mathcal{C}(\widetilde{\mathsf{T}}) \\ \tilde{x}_{i,m} \text{ fixed}}} P_{\widetilde{\mathbf{Y}}|\widetilde{\mathbf{X}}}(\mathbf{y}^{\uparrow M} | \widetilde{\mathbf{x}}) \, .$$

Symbolwise graph-cover decoding is then defined to be the decoding algorithm that bases its decision on the averaged marginals

$$\eta_{i,m,M}(\tilde{x}_{i,m}) \triangleq \frac{1}{|\mathcal{T}_M|} \sum_{\widetilde{\mathsf{T}} \in \mathcal{T}_M} \eta_{i,m,\widetilde{\mathsf{T}}}(\tilde{x}_{i,m}) \, ,$$

where the averaging is over the set $\mathcal{T}_M$ of all M-fold covers of the base factor graph.



Symbolwise Graph-Cover Decoding

Symbolwise Graph-Cover Decoding

For each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$, in the limit $M \to \infty$ we have

$$\lim_{M \to \infty} \frac{\big(\eta_{i,m,M}(0),\ \eta_{i,m,M}(1)\big)}{\eta_{i,m,M}(0) + \eta_{i,m,M}(1)} = (1 - \omega_i^*,\ \omega_i^*)\ ,$$

where

$$\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}}\ F_{\text{Bethe}}(\boldsymbol{\omega})\ .$$

Symbolwise Graph-Cover Decoding

For each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$, in the limit $M \to \infty$ we have

$$\lim_{M \to \infty} \frac{\big(\eta_{i,m,M}(0),\ \eta_{i,m,M}(1)\big)}{\eta_{i,m,M}(0) + \eta_{i,m,M}(1)} = \left(1 - \omega_i^*,\ \omega_i^*\right) ,$$

where

$$\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}} \ F_{\mathrm{Bethe}}(\boldsymbol{\omega}) \ .$$

- Let $\mathcal{P}(\mathbf{H})$ be the fundamental polytope associated with $\mathbf{H}$.

Symbolwise Graph-Cover Decoding

For each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$, in the limit $M \to \infty$ we have

$$\lim_{M \to \infty} \frac{\big(\eta_{i,m,M}(0), \ \eta_{i,m,M}(1)\big)}{\eta_{i,m,M}(0) + \eta_{i,m,M}(1)} = (1 - \omega_i^*, \ \omega_i^*) \ ,$$

where

$$\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}} \ F_{\mathrm{Bethe}}(\boldsymbol{\omega}) \ .$$

- Let $\mathcal{P}(\mathbf{H})$ be the fundamental polytope associated with $\mathbf{H}$.
- Then $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = +\infty$ for $\boldsymbol{\omega} \notin \mathcal{P}(\mathbf{H})$.

Symbolwise Graph-Cover Decoding

For each $i = 1, \ldots, n$ and each $m = 1, \ldots, M$, in the limit $M \to \infty$ we have

$$\lim_{M\to\infty} \frac{\big(\eta_{i,m,M}(0),\ \eta_{i,m,M}(1)\big)}{\eta_{i,m,M}(0) + \eta_{i,m,M}(1)} = (1 - \omega_i^*,\ \omega_i^*)\ ,$$

where

$$\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}}\ F_{\text{Bethe}}(\boldsymbol{\omega})\ .$$

- Let $\mathcal{P}(\mathbf{H})$ be the fundamental polytope associated with $\mathbf{H}$.
- Then $F_{\text{Bethe}}(\boldsymbol{\omega}) = +\infty$ for $\boldsymbol{\omega} \notin \mathcal{P}(\mathbf{H})$.
- In the degenerate case where $F_{\text{Bethe}}(\boldsymbol{\omega})$ has multiple global minima, $\boldsymbol{\omega}^* \in \text{conv}\big(\arg\min_{\boldsymbol{\omega}}\ F_{\text{Bethe}}(\boldsymbol{\omega})\big)$.

Symbolwise Graph-Cover Decoding

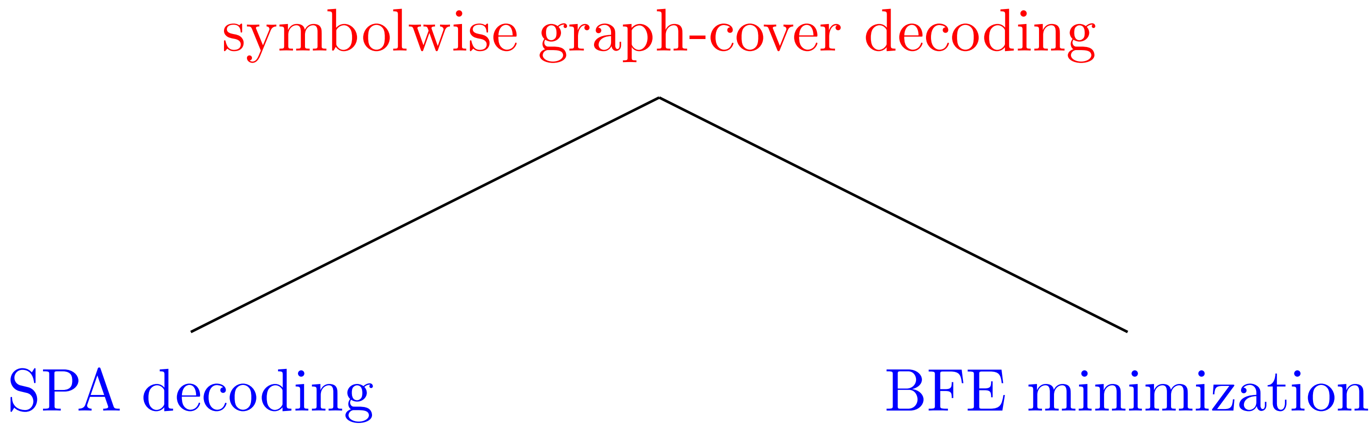
Symbolwise Graph-Cover Decoding

symbolwise graph-cover decoding

SPA decoding

BFE minimization

Symbolwise Graph-Cover Decoding



Theorem (Yedidia/Freeman/Weiss):
Fixed points of the sum-product algorithm correspond to stationary points of the Bethe free energy.

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

$$F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega}) \ .$$

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

$$F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega}) \ .$$

Therefore,

$$\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}} \ +F_{\mathrm{Bethe}}(\boldsymbol{\omega})$$

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

$$F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega}) \; .$$

Therefore,

$$\begin{aligned} \boldsymbol{\omega}^{*} &= \arg\min_{\boldsymbol{\omega}} \; +F_{\mathrm{Bethe}}(\boldsymbol{\omega}) \\ &= \arg\max_{\boldsymbol{\omega}} \; -F_{\mathrm{Bethe}}(\boldsymbol{\omega}) \end{aligned}$$

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

$$F_{\text{Bethe}}(\boldsymbol{\omega}) = U_{\text{Bethe}}(\boldsymbol{\omega}) - H_{\text{Bethe}}(\boldsymbol{\omega}) \; .$$

Therefore,

$$\begin{aligned} \boldsymbol{\omega}^* &= \arg\min_{\boldsymbol{\omega}} \; +F_{\text{Bethe}}(\boldsymbol{\omega}) \\ &= \arg\max_{\boldsymbol{\omega}} \; -F_{\text{Bethe}}(\boldsymbol{\omega}) \\ &= \arg\max_{\boldsymbol{\omega}} \; -U_{\text{Bethe}}(\boldsymbol{\omega}) + H_{\text{Bethe}}(\boldsymbol{\omega}) \; . \end{aligned}$$

Maximum-Entropy Interpretation of Symbolwise Graph-Cover Decoding

$$F_{\text{Bethe}}(\boldsymbol{\omega}) = U_{\text{Bethe}}(\boldsymbol{\omega}) - H_{\text{Bethe}}(\boldsymbol{\omega}) \,.$$

Therefore,

$$\begin{aligned} \boldsymbol{\omega}^* &= \arg\min_{\boldsymbol{\omega}} \; +F_{\text{Bethe}}(\boldsymbol{\omega}) \\ &= \arg\max_{\boldsymbol{\omega}} \; -F_{\text{Bethe}}(\boldsymbol{\omega}) \\ &= \arg\max_{\boldsymbol{\omega}} \; -U_{\text{Bethe}}(\boldsymbol{\omega}) + H_{\text{Bethe}}(\boldsymbol{\omega}) \,. \end{aligned}$$

So, apart from the biasing $-U_{\text{Bethe}}(\boldsymbol{\omega})$ term based on the observed channel output, we take the maximum-entropy solution, i.e., the pseudo-codeword $\boldsymbol{\omega}$ whose associated set of codewords in the graph covers has maximum Bethe entropy.

LABS hp

Symbolwise Graph-Cover Decoding

This maximum-entropy result has to be understood as follows.

Symbolwise Graph-Cover Decoding

This maximum-entropy result has to be understood as follows.

- To any length-Mn codeword $\tilde{\mathbf{c}}$ in an M-cover $\widetilde{\mathsf{T}}$ we can associate the length-n pseudo-codeword $\boldsymbol{\omega}$ where

$$\omega_i \triangleq \frac{1}{M} \sum_{m=1}^{M} \tilde{c}_{i,m} \ .$$

Symbolwise Graph-Cover Decoding

This maximum-entropy result has to be understood as follows.

- To any length-Mn codeword $\tilde{\mathbf{c}}$ in an M-cover $\widetilde{\mathsf{T}}$ we can associate the length-n pseudo-codeword $\boldsymbol{\omega}$ where

$$\omega_i \triangleq \frac{1}{M} \sum_{m=1}^{M} \tilde{c}_{i,m} \ .$$

- With this, $H_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is the asymptotic growth rate of the number of codewords in all M-covers with pseudo-codeword $\boldsymbol{\omega}$, $M \to \infty$.

Symbolwise Graph-Cover Decoding

Some comments:

Symbolwise Graph-Cover Decoding

Some comments:

- Strictly speaking, the Bethe free energy is a function not only $\boldsymbol{\omega}$, but also of other variables, namely of
  - the variables $\{\beta_j\}_j$ associated with the check nodes of T,
  - the variables $\{\gamma_e\}_e$ associated with the edges of T.

Symbolwise Graph-Cover Decoding

Some comments:

- Strictly speaking, the Bethe free energy is a function not only $\boldsymbol{\omega}$, but also of other variables, namely of
  - the variables $\{\beta_j\}_j$ associated with the check nodes of T,
  - the variables $\{\gamma_e\}_e$ associated with the edges of T.
- For computing $H_{\mathrm{Bethe}}(\boldsymbol{\omega})$ we can leverage results on the asymptotic growth rate of the average Hamming weight of protograph-based LDPC codes. Cf.
  - [Fogal/McEliece/Thorpe, 2005],
  - papers by Divsalar, Ryan, et al. (2005–).

Bethe Free Energy and Its Lagrange Dual

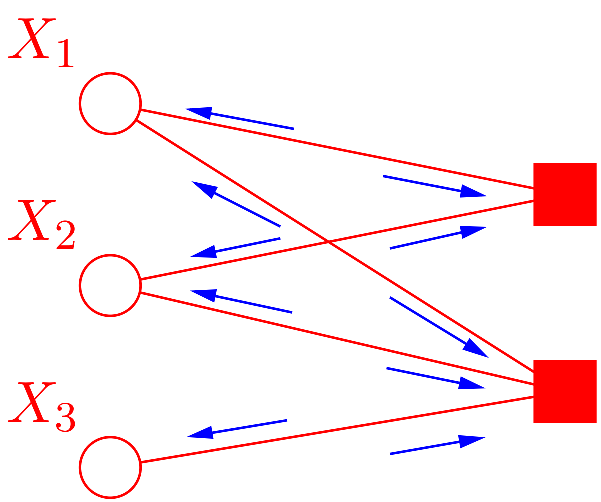
Lagrangian for the Bethe Free Energy

Lagrangian for the Bethe Free Energy

$$
\begin{aligned}
L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq F_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \\
&+ \sum_{e \in \mathcal{E}} \overrightarrow{\mu}_{e,0} \left(\gamma_{e,0} - \sum_{\mathbf{b}_j :\ b_{j,e}=0} \beta_j \right) \\
&+ \sum_{e \in \mathcal{E}} \overrightarrow{\mu}_{e,1} \left(\gamma_{e,1} - \sum_{\mathbf{b}_j :\ b_{j,e}=1} \beta_j \right) \\
&+ \sum_{e \in \mathcal{E}} \overleftarrow{\mu}_{e,0} \big(\gamma_{e,0} - \omega_{e,0} \big) \\
&+ \sum_{e \in \mathcal{E}} \overleftarrow{\mu}_{e,1} \big(\gamma_{e,1} - \omega_{e,1} \big) .
\end{aligned}
$$

Lagrangian for the Bethe Free Energy

$$L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq F_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma})$$



$$+ \sum_{e \in \mathcal{E}} \overrightarrow{\mu}_{e,0} \left(\gamma_{e,0} - \sum_{\mathbf{b}_j :\ b_{j,e}=0} \beta_j \right)$$

$$+ \sum_{e \in \mathcal{E}} \overrightarrow{\mu}_{e,1} \left(\gamma_{e,1} - \sum_{\mathbf{b}_j :\ b_{j,e}=1} \beta_j \right)$$

$$+ \sum_{e \in \mathcal{E}} \overleftarrow{\mu}_{e,0} \big(\gamma_{e,0} - \omega_{e,0} \big)$$

$$+ \sum_{e \in \mathcal{E}} \overleftarrow{\mu}_{e,1} \big(\gamma_{e,1} - \omega_{e,1} \big) \ .$$

Lagrange Dual of Bethe Free Energy

Lagrange Dual of Bethe Free Energy

Lagrange Dual of Bethe Free Energy:

$$F'_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq \min_{\boldsymbol{\omega}} \min_{\boldsymbol{\beta}} \min_{\boldsymbol{\gamma}} \; L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$$

Lagrange Dual of Bethe Free Energy

Lagrange Dual of Bethe Free Energy:

$$F'_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq \min_{\boldsymbol{\omega}} \min_{\boldsymbol{\beta}} \min_{\boldsymbol{\gamma}} \ L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$$

Lagrange Dual of Bethe Free Energy

Lagrange Dual of Bethe Free Energy:

$$F'_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq \min_{\boldsymbol{\omega}} \min_{\boldsymbol{\beta}} \min_{\boldsymbol{\gamma}} \; L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$$

Lagrange Pseudo-Dual of Bethe Free Energy:

$$F^{\#}_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) \triangleq \min_{\boldsymbol{\omega}} \min_{\boldsymbol{\beta}} \max_{\boldsymbol{\gamma}} \; L_{\text{Bethe}}(\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$$

[Regalia and Walsh, 2007]

Bethe Free Energy for the BEC

Lagrange Pseudo-Dual for BEC

Lagrange Pseudo-Dual for BEC

- For the BEC, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0,0)\,, \quad (0,-\infty)\,, \quad (-\infty,0)\,.$$

Lagrange Pseudo-Dual for BEC

- For the BEC, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0, 0)\ , \quad (0, -\infty)\ , \quad (-\infty, 0)\ .$$

- Assuming that the all-zeros codeword was sent, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0, 0) \quad \text{and} \quad (0, -\infty)\ .$$

Lagrange Pseudo-Dual for BEC

- For the BEC, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0,0)\ , \quad (0,-\infty)\ , \quad (-\infty,0)\ .$$

- Assuming that the all-zeros codeword was sent, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0,0) \quad \text{and} \quad (0,-\infty)\ .$$

- Let $\overrightarrow{x}$ count the fraction of $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1}) = (0,0)$ messages.

Lagrange Pseudo-Dual for BEC

- For the BEC, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0,0)\,, \quad (0,-\infty)\,, \quad (-\infty,0)\,.$$

- Assuming that the all-zeros codeword was sent, it is sufficient to consider only pairs $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1})$ and $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,0})$ that take on the values

$$(0,0) \quad \text{and} \quad (0,-\infty)\,.$$

- Let $\overrightarrow{x}$ count the fraction of $(\overrightarrow{\mu}_{e,0}, \overrightarrow{\mu}_{e,1}) = (0,0)$ messages.
- Let $\overleftarrow{x}$ count the fraction of $(\overleftarrow{\mu}_{e,0}, \overleftarrow{\mu}_{e,1}) = (0,0)$ messages.

Lagrange Pseudo-Dual for BEC

Lagrange Pseudo-Dual for BEC

- Consider an LDPC code whose Tanner graph has
  - length-n variable nodes,
  - E edges,
  - vertex-perspective degree distributions $L(x)$ and $R(x)$,
  - edge-perspective degree distributions $\lambda(x)$ and $\rho(x)$.

Lagrange Pseudo-Dual for BEC

- Consider an LDPC code whose Tanner graph has
  - length-n variable nodes,
  - E edges,
  - vertex-perspective degree distributions $L(x)$ and $R(x)$,
  - edge-perspective degree distributions $\lambda(x)$ and $\rho(x)$.
- Let ε be the erasure probability of the BEC.

Lagrange Pseudo-Dual for BEC

- Consider an LDPC code whose Tanner graph has
  - length-n variable nodes,
  - E edges,
  - vertex-perspective degree distributions $L(x)$ and $R(x)$,
  - edge-perspective degree distributions $\lambda(x)$ and $\rho(x)$.
- Let ε be the erasure probability of the BEC.
- Let $\tilde{\varepsilon}$ be the actual fraction of errors that was introduced by the channel.

Lagrange Pseudo-Dual for BEC

- Consider an LDPC code whose Tanner graph has
  - length-n variable nodes,
  - E edges,
  - vertex-perspective degree distributions $L(x)$ and $R(x)$,
  - edge-perspective degree distributions $\lambda(x)$ and $\rho(x)$.
- Let ε be the erasure probability of the BEC.
- Let $\tilde{\varepsilon}$ be the actual fraction of errors that was introduced by the channel.
- Locally tree-like (LTL) assumption: the number of iterations does not exceed $\mathrm{girth}/4$.

Lagrange Pseudo-Dual for BEC

Lagrange Pseudo-Dual for BEC

With these definitions,

$$F^{\#}_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) = \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = \overline{U}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) - \overline{H}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \ ,$$

with

$$\overline{U}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = E \cdot \frac{h(\tilde{\varepsilon}) + D(\tilde{\varepsilon} \| \varepsilon)}{L'(1)}$$

$$\overline{H}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = E \cdot \left(\frac{\tilde{\varepsilon} L(\overleftarrow{x})}{L'(1)} + \frac{R(1 - \overrightarrow{x})}{R'(1)} + \overrightarrow{x}(1 - \overleftarrow{x}) - \frac{1}{R'(1)} \right) .$$

Lagrange Pseudo-Dual for BEC

With these definitions,

$$F^{\#}_{\text{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}}) = \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = \overline{U}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) - \overline{H}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \, ,$$

with

$$\overline{U}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = E \cdot \frac{h(\tilde{\varepsilon}) + D(\tilde{\varepsilon} \| \varepsilon)}{L'(1)}$$

$$\overline{H}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) = E \cdot \left(\frac{\tilde{\varepsilon} L(\overleftarrow{x})}{L'(1)} + \frac{R(1 - \overrightarrow{x})}{R'(1)} + \overrightarrow{x}(1 - \overleftarrow{x}) - \frac{1}{R'(1)} \right) .$$

Remark: $\overline{H}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ is also known as trial entropy,
cf. [Méasson/Montanari/Urbanke, 2005].

Lagrange Pseudo-Dual for BEC

Lagrange Pseudo-Dual for BEC

If LTL assumption is fulfilled, it can be shown that at each iteration of the peeling decoder, $F^{\#}_{\mathrm{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$ stays constant or decreases by 1.

Lagrange Pseudo-Dual for BEC

If LTL assumption is fulfilled, it can be shown that at each iteration of the peeling decoder, $F^{\#}_{\mathrm{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$ stays constant or decreases by 1.

$\Rightarrow$ $F^{\#}_{\mathrm{Bethe}}(\overrightarrow{\boldsymbol{\mu}}, \overleftarrow{\boldsymbol{\mu}})$ can serve as a Lyapunov function for the BEC.

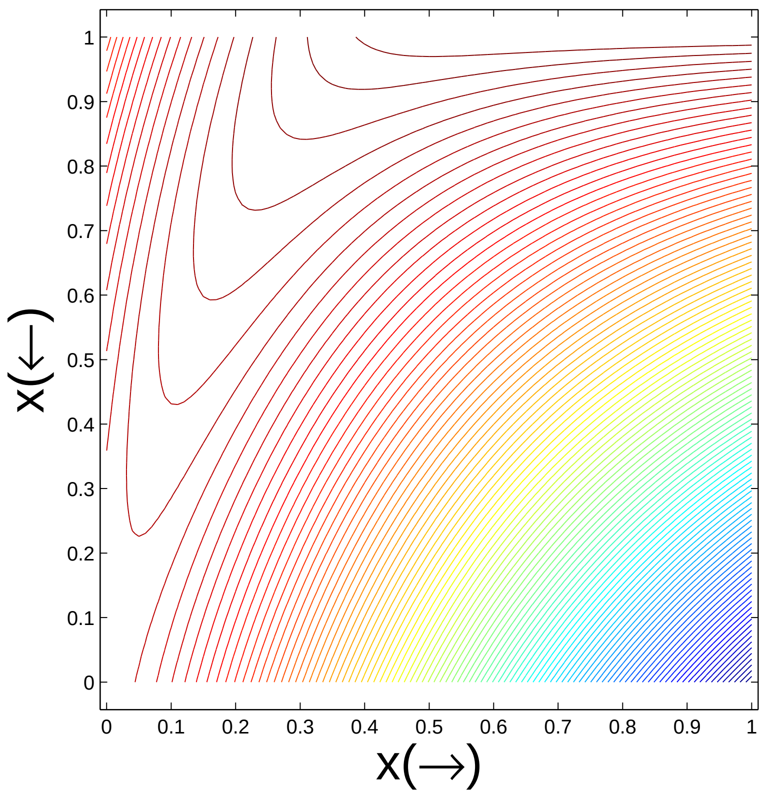
Lagrange Pseudo-Dual for BEC

Lagrange Pseudo-Dual for BEC

- Choose a $(3, 6)$-regular LDPC code of length n.

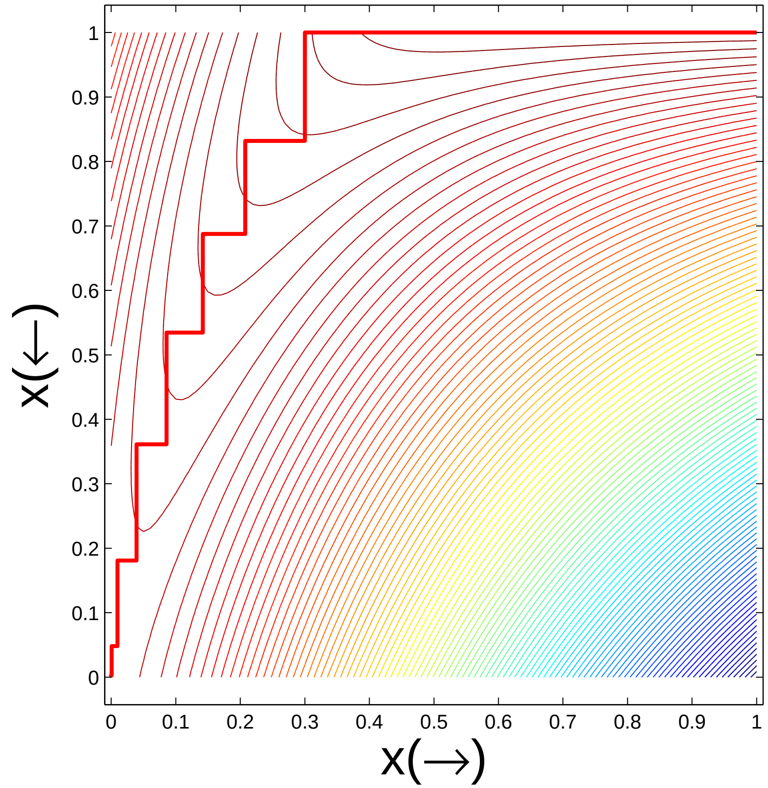
Lagrange Pseudo-Dual for BEC

- Choose a $(3, 6)$-regular LDPC code of length n.
- Plot $\frac{1}{E}\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.3$.



Lagrange Pseudo-Dual for BEC

- Choose a $(3, 6)$-regular LDPC code of length n.
- Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.3$.
- Trajectory under LTL assumption.

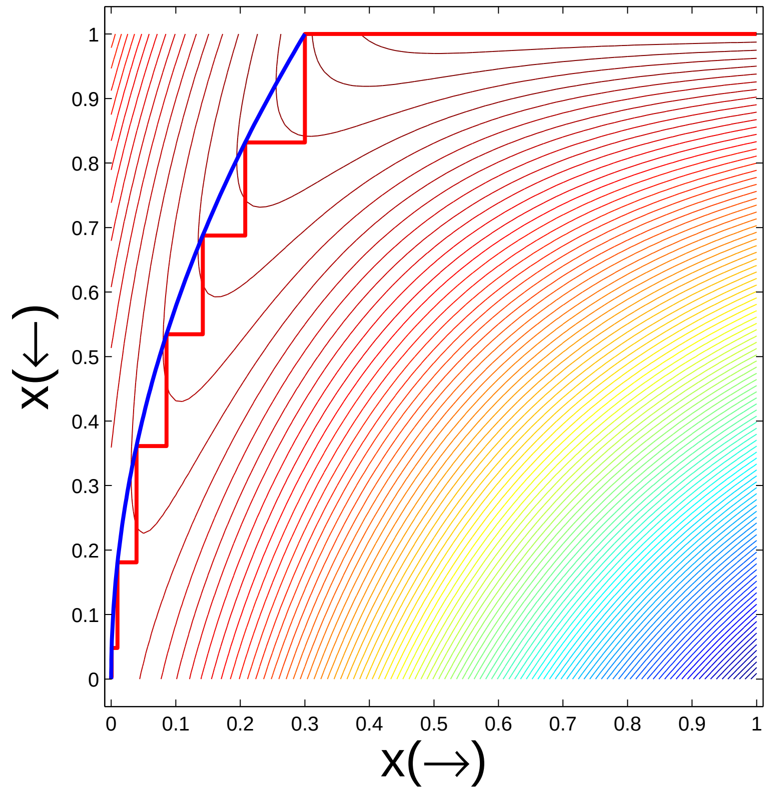


Lagrange Pseudo-Dual for BEC

- Choose a $(3,6)$-regular LDPC code of length n.
- Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.3$.

- Trajectory under LTL assumption.
- Left boundary of trajectory:

$$\overrightarrow{x} = \varepsilon\lambda(\overleftarrow{x}).$$



Lagrange Pseudo-Dual for BEC

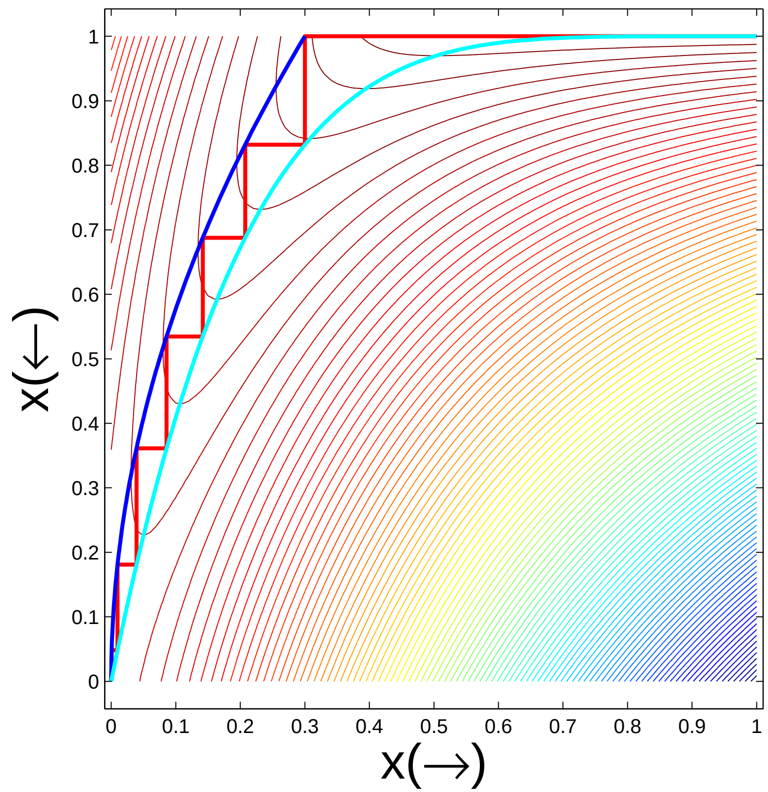
- Choose a $(3,6)$-regular LDPC code of length n.
- Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.3$.

- Trajectory under LTL assumption.
- Left boundary of trajectory:

$$\overrightarrow{x} = \varepsilon\lambda(\overleftarrow{x}).$$

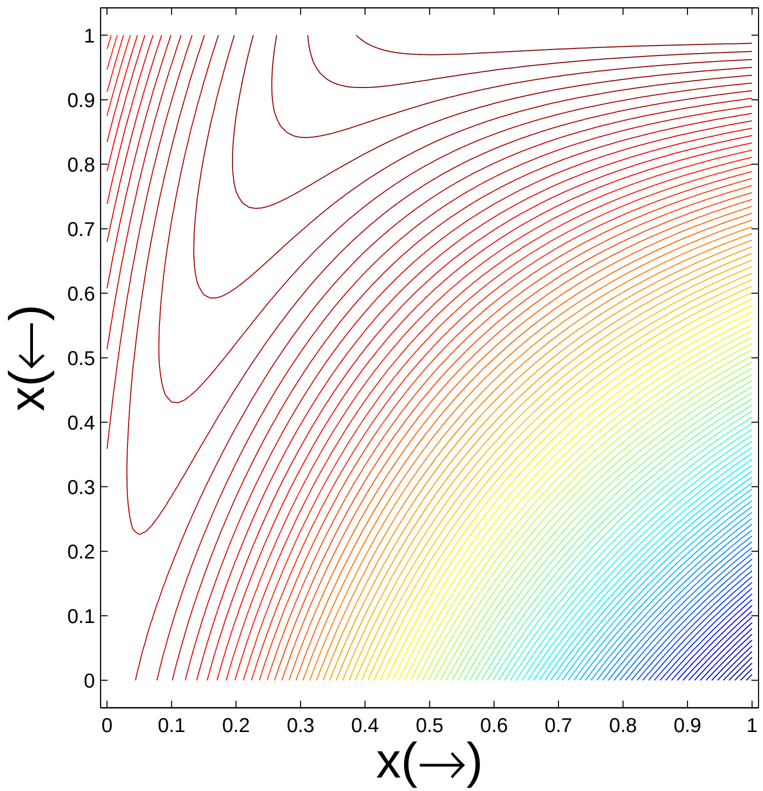
- Right boundary of trajectory:

$$\overleftarrow{x} = 1 - \rho(1 - \overrightarrow{x}).$$



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

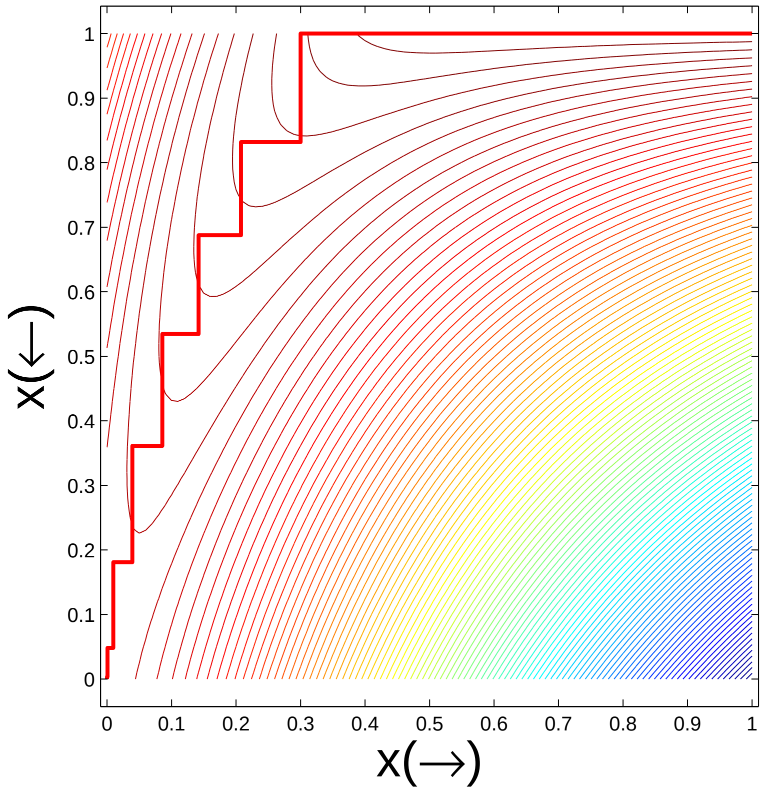


Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta \overline{F}^{\#}_{\text{Bethe}} = n \cdot \big(\text{rate} - \text{cap}_{\text{BEC}(\varepsilon)}\big) \; ,$$



LABS hp

Lagrange Pseudo-Dual for BEC

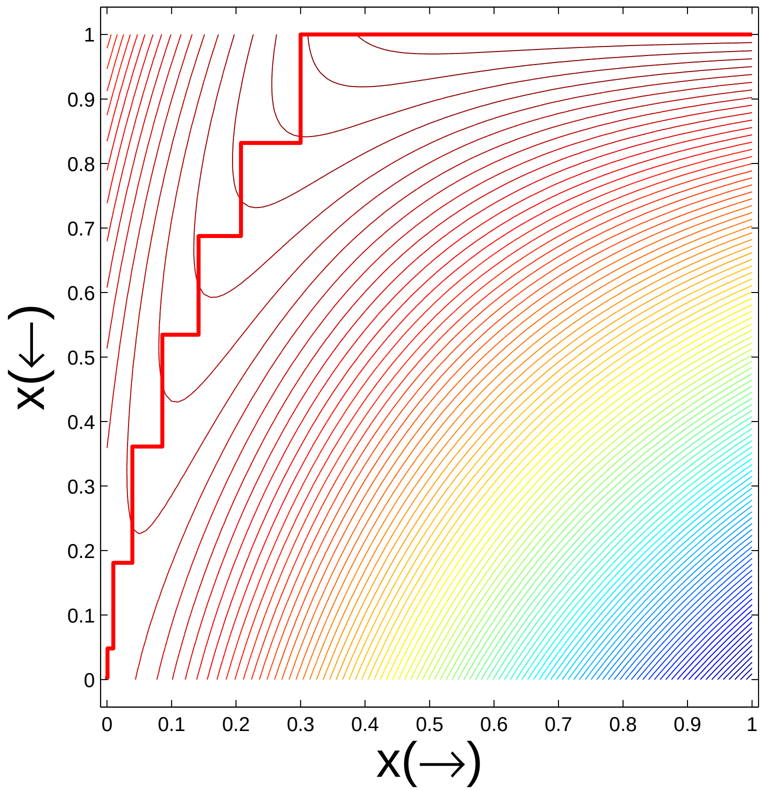
$$\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{end}} - \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{begin}} = \int_{\mathrm{Path}} \nabla \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta \overline{F}^{\#}_{\mathrm{Bethe}} = n \cdot \left(\mathrm{rate} - \mathrm{cap}_{\mathrm{BEC}(\varepsilon)}\right),$$

$\Delta \overline{F}^{\#}_{\mathrm{Bethe}} \leq 0$ implies the necessary condition

$$\mathrm{rate} \leq \mathrm{cap}_{\mathrm{BEC}(\varepsilon)}$$



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

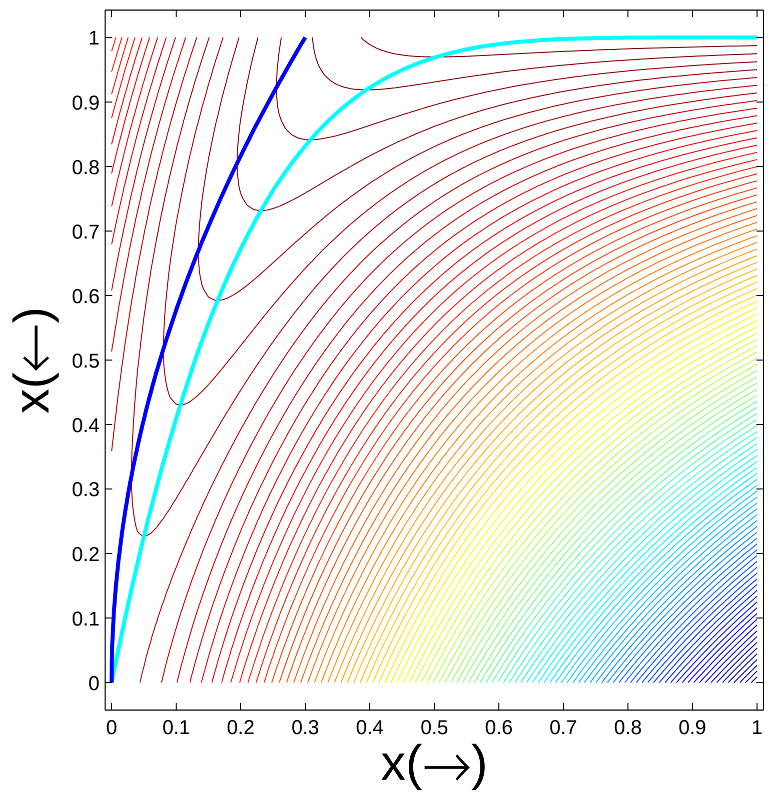
Integral along red path:

$$\Delta \overline{F}^{\#}_{\text{Bethe}} = n \cdot \left(\text{rate} - \text{cap}_{\text{BEC}(\varepsilon)}\right),$$

$\Delta \overline{F}^{\#}_{\text{Bethe}} \leq 0$ implies the necessary condition

$$\text{rate} \leq \text{cap}_{\text{BEC}(\varepsilon)}$$

Gives also the area between the blue and the cyan curve.



LABS hp

Lagrange Pseudo-Dual for BEC

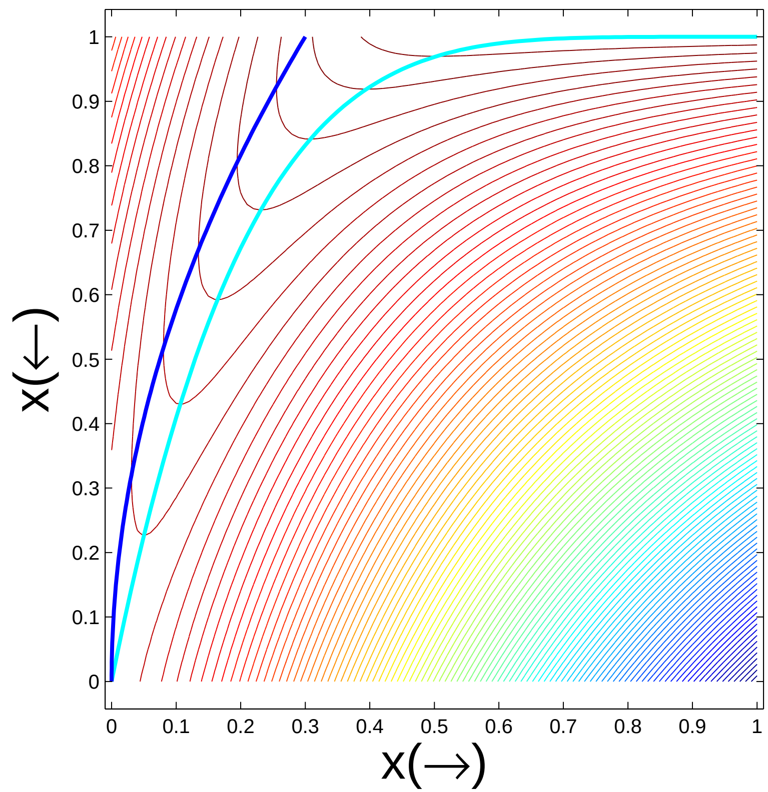
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta \overline{F}^{\#}_{\text{Bethe}} = n \cdot \left(\text{rate} - \text{cap}_{\text{BEC}(\varepsilon)}\right),$$

$\Delta \overline{F}^{\#}_{\text{Bethe}} \leq 0$ implies the necessary condition

$$\text{rate} \leq \text{cap}_{\text{BEC}(\varepsilon)}$$

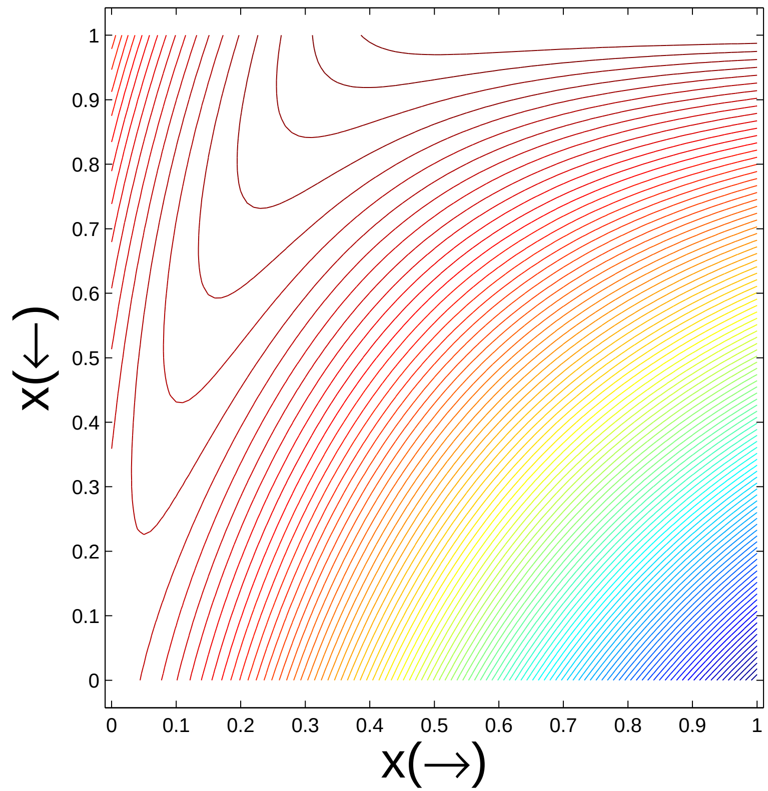


Gives also the area between the blue and the cyan curve.

Area theorem [Ashikhmin et al.]

Lagrange Pseudo-Dual for BEC

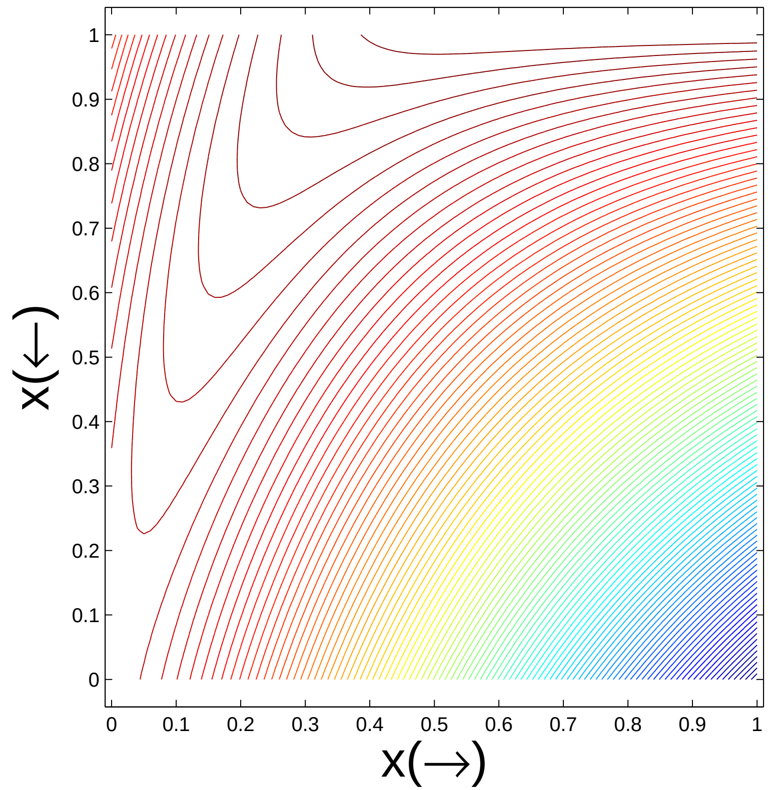
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Necessary condition to reach the point $(\overrightarrow{x}, \overleftarrow{x}) = (0, 0)$: there is some $(\overrightarrow{x}, \overleftarrow{x}) \neq (0, 0)$ in the neighborhood of $(0, 0)$ such that $\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \geq \overline{F}^{\#}_{\text{Bethe}}(0, 0)$.

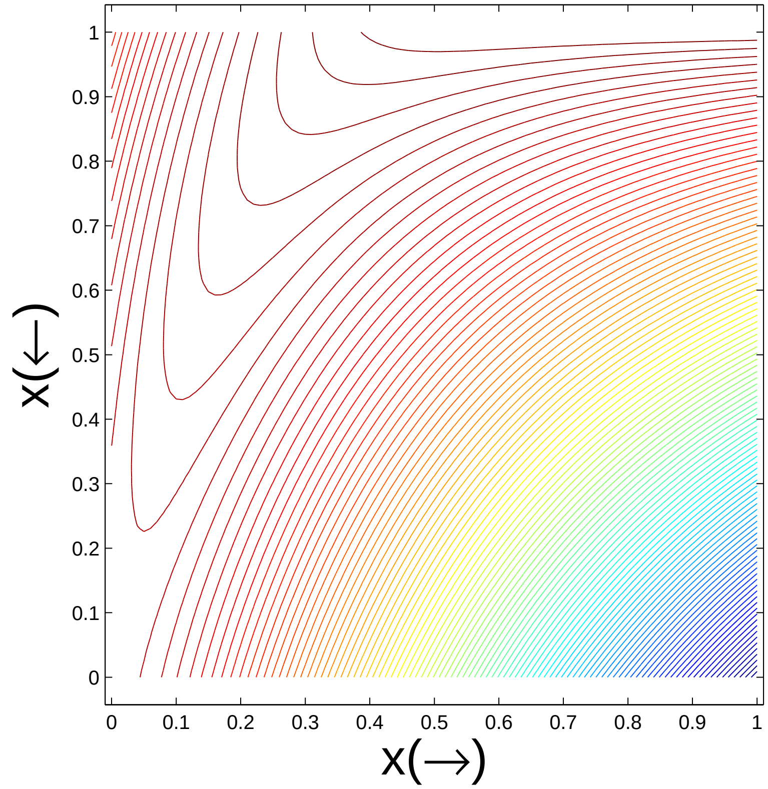


Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Necessary condition to reach the point $(\overrightarrow{x}, \overleftarrow{x}) = (0,0)$: there is some $(\overrightarrow{x}, \overleftarrow{x}) \neq (0,0)$ in the neighborhood of $(0,0)$ such that $\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \geq \overline{F}^{\#}_{\text{Bethe}}(0,0)$.

- Gradient is zero.



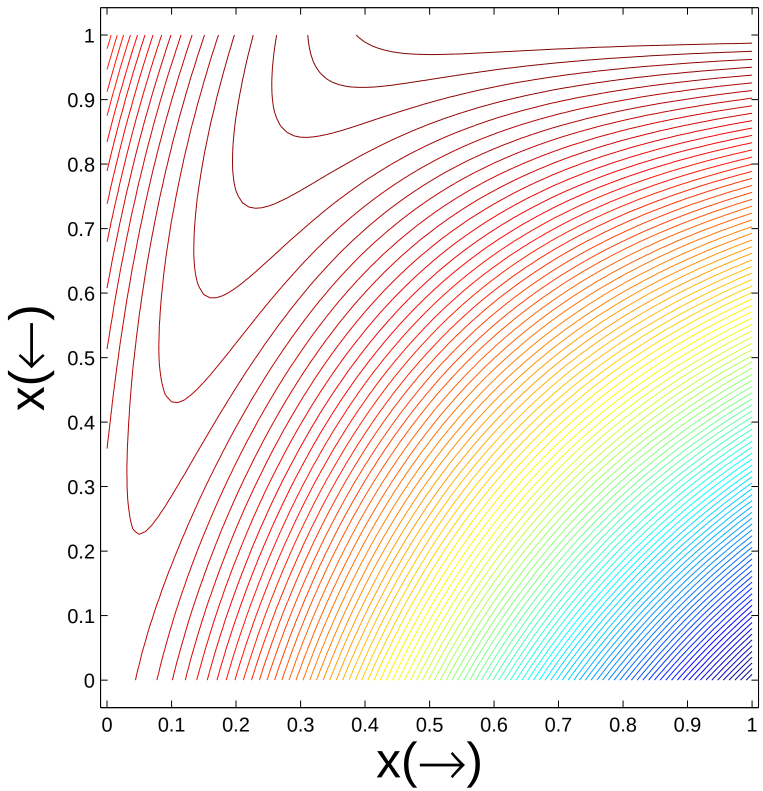
Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Necessary condition to reach the point $(\overrightarrow{x}, \overleftarrow{x}) = (0, 0)$: there is some $(\overrightarrow{x}, \overleftarrow{x}) \neq (0, 0)$ in the neighborhood of $(0, 0)$ such that $\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \geq \overline{F}^{\#}_{\text{Bethe}}(0, 0)$.

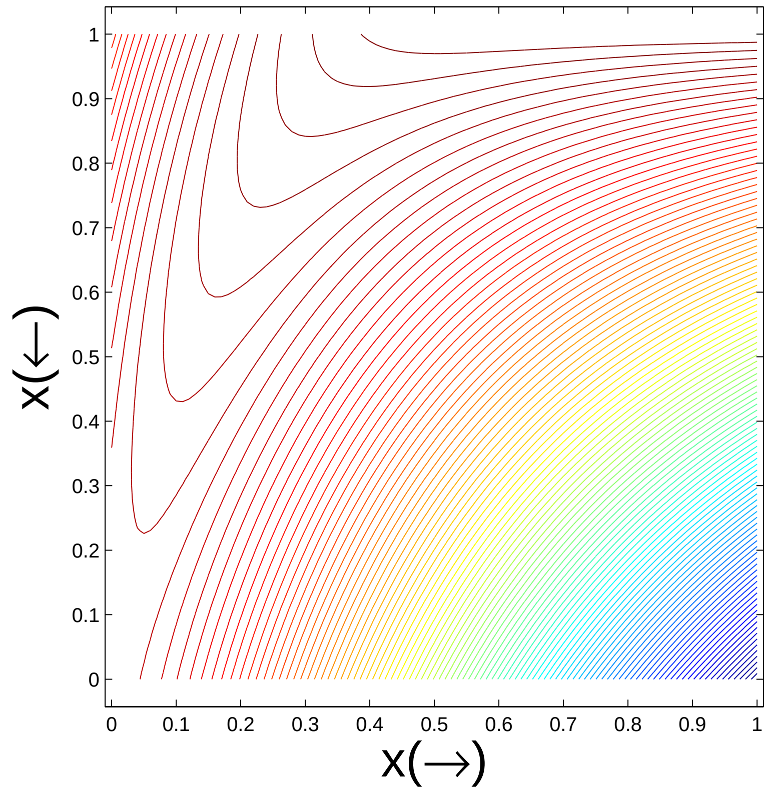
- Gradient is zero.
- Studying the Hessian yields the necessary condition

$$\varepsilon \lambda'(0) \rho(1) \leq 1 \; .$$



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

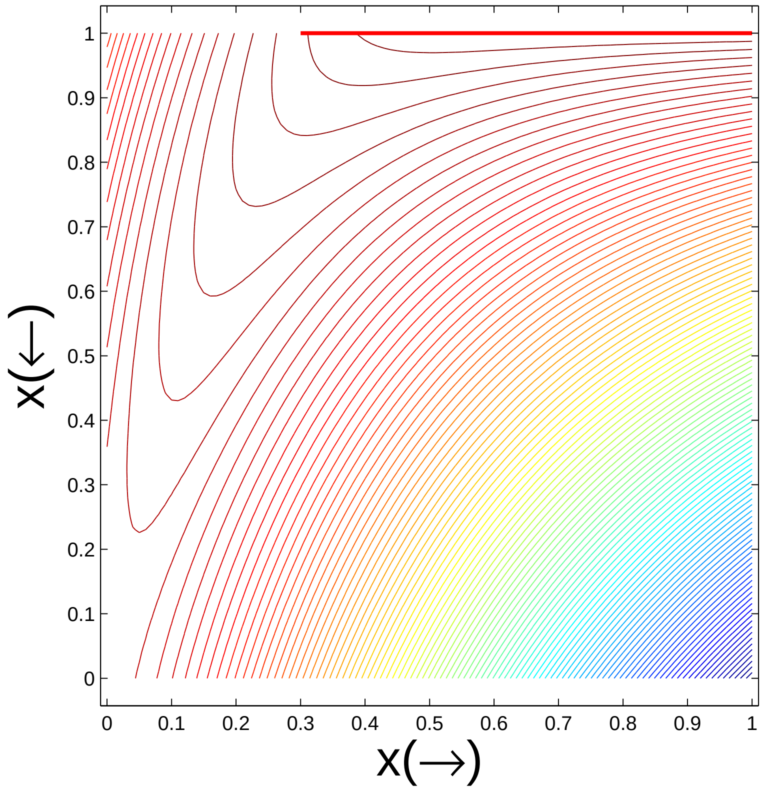


Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta \overline{F}^{\#}_{\text{Bethe}} = -n \cdot (1 - \text{rate}) \cdot R(1 - \varepsilon) \ .$$



Lagrange Pseudo-Dual for BEC

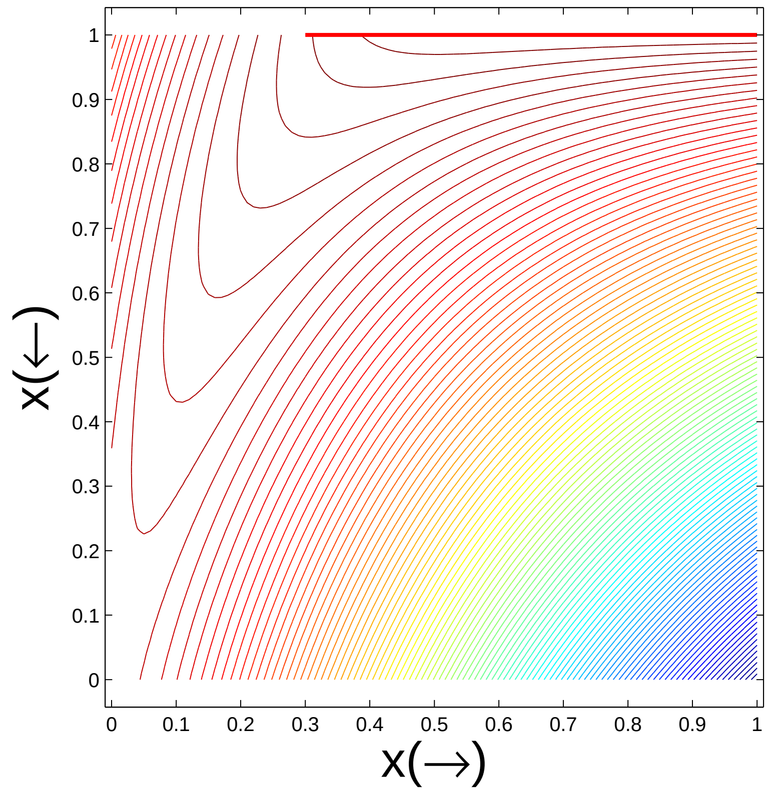
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta\overline{F}^{\#}_{\text{Bethe}} = -n \cdot (1-\text{rate}) \cdot R(1-\varepsilon) \ .$$

$\Delta\frac{1}{n}\overline{F}^{\#}_{\text{Bethe}} \to 0$ implies the necessary condition

$$R(1-\varepsilon) \to 0 \ .$$



Lagrange Pseudo-Dual for BEC

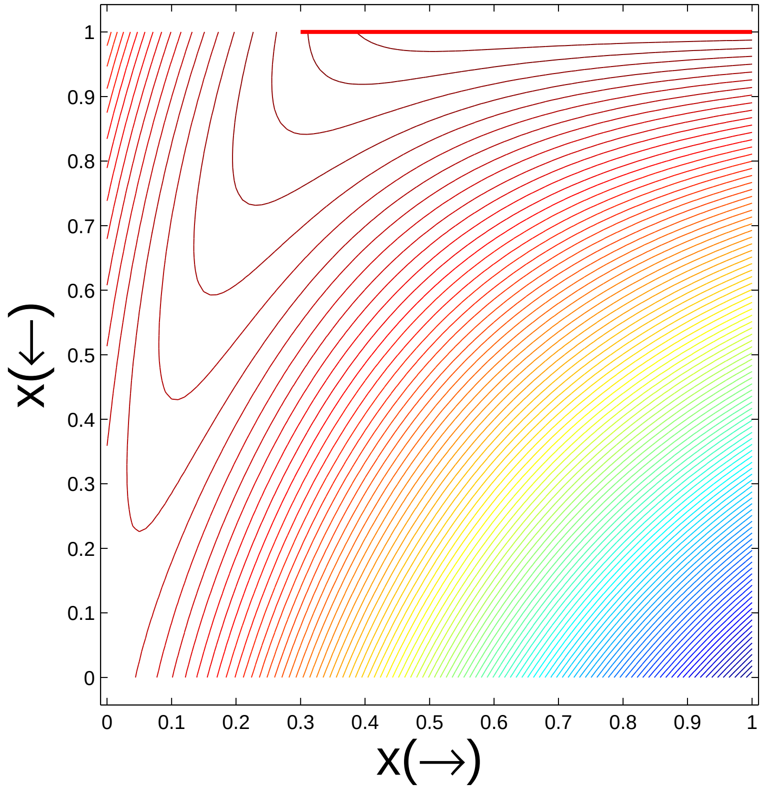
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Integral along red path:

$$\Delta \overline{F}^{\#}_{\text{Bethe}} = -n \cdot (1-\text{rate}) \cdot R(1-\varepsilon) \; .$$

$\Delta \frac{1}{n} \overline{F}^{\#}_{\text{Bethe}} \to 0$ implies the necessary condition

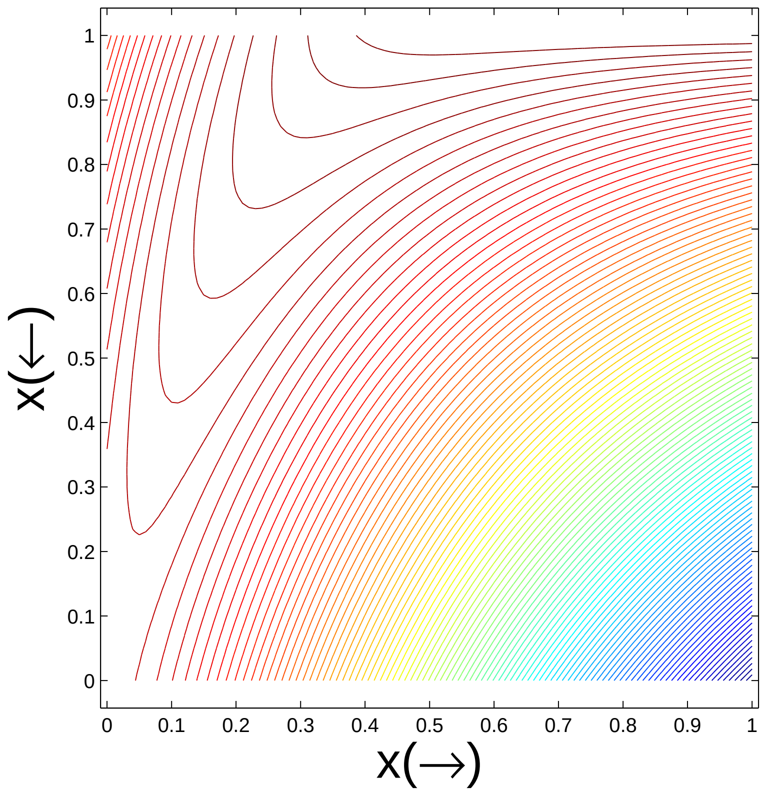
$$R(1-\varepsilon) \to 0 \; .$$



$\Rightarrow$ Average right degree has to go to infinity.

Lagrange Pseudo-Dual for BEC

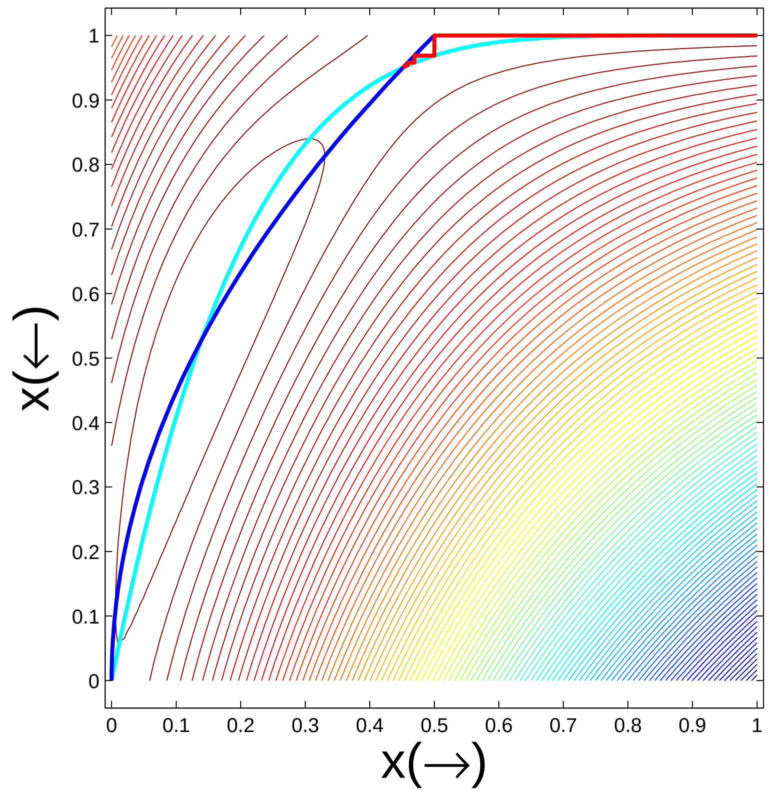
$$\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{end}} - \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{begin}} = \int_{\mathrm{Path}} \nabla \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{end}} - \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{begin}} = \int_{\mathrm{Path}} \nabla \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Plot $\frac{1}{E}\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.5$.

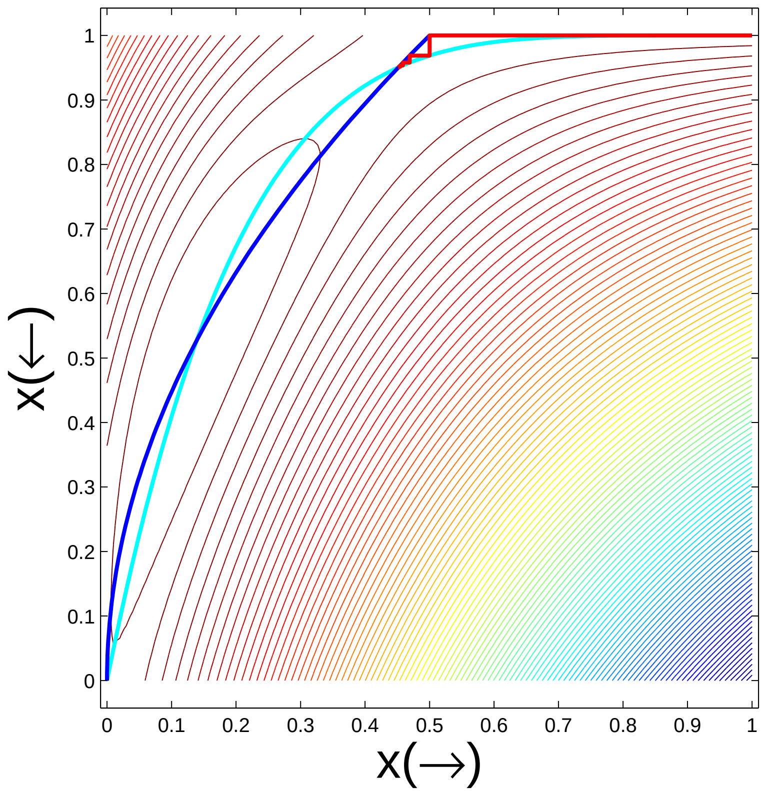


Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x},\overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.5$.

Fixed point corresonds to stationary points of the Bethe free energy.



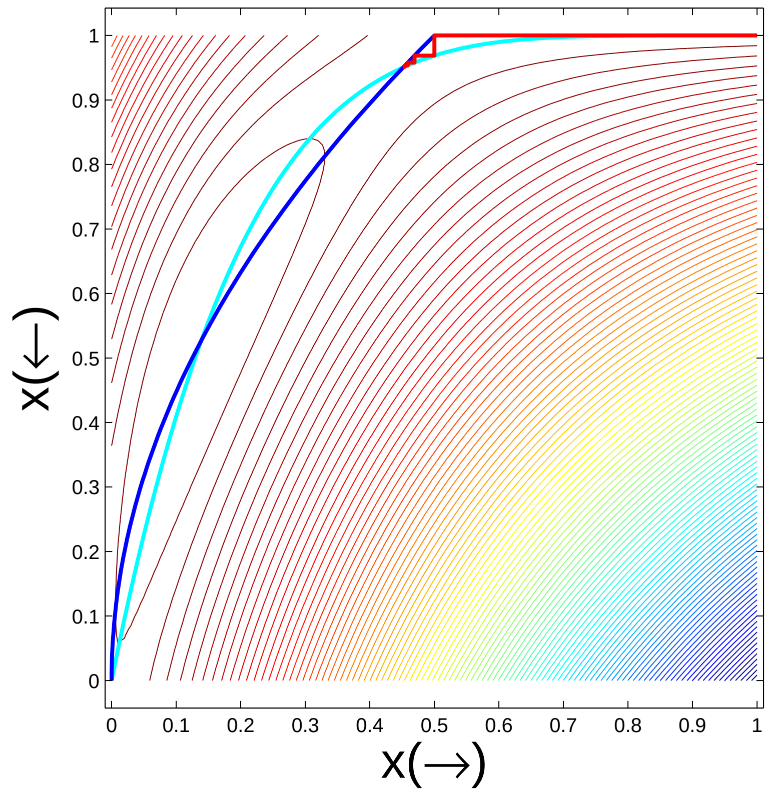
Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.5$.

Fixed point corresonds to stationary points of the Bethe free energy.

$\Rightarrow$ Fixed point corresponds to a pseudo-codword (necessarily in the fundamental polytope).



Lagrange Pseudo-Dual for BEC

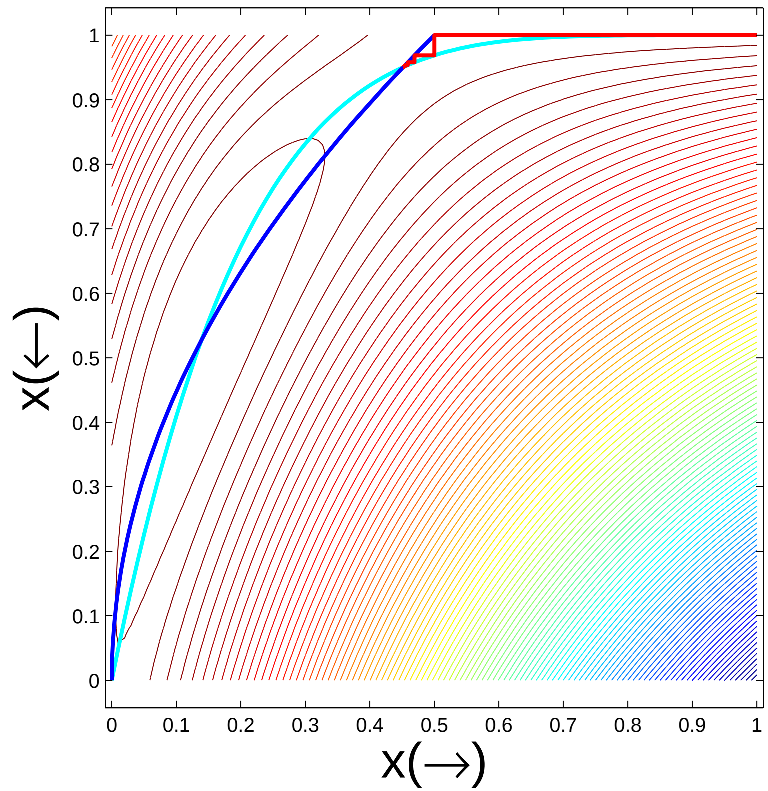
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

Plot $\frac{1}{E}\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$ for $\varepsilon = \tilde{\varepsilon} = 0.5$.

Fixed point corresonds to stationary points of the Bethe free energy.

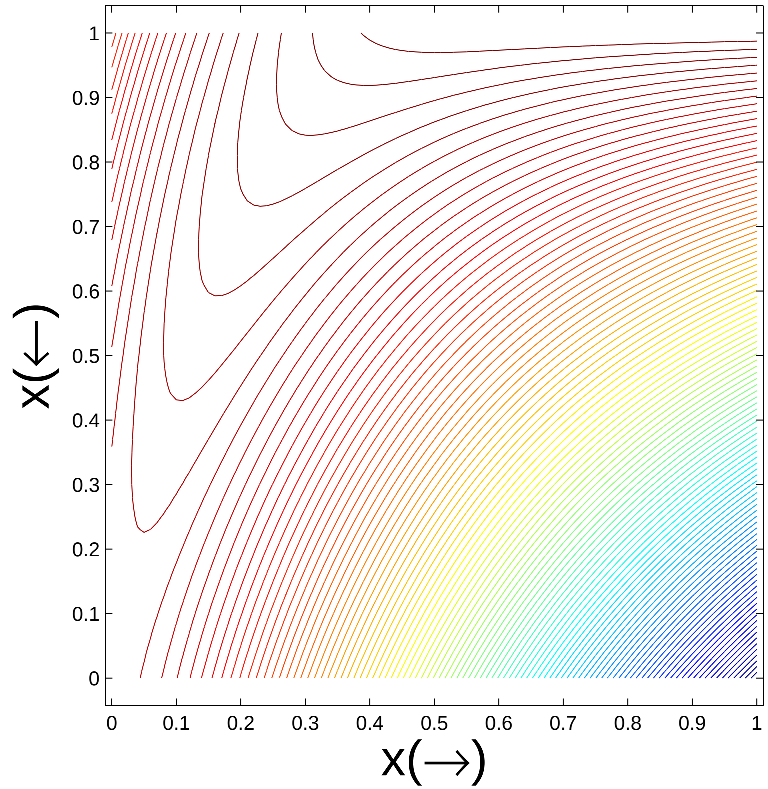
$\Rightarrow$ Fixed point corresponds to a pseudo-codword (necessarily in the fundamental polytope).

$\Rightarrow$ Stopping set equals support of that pseudo-codeword.



Lagrange Pseudo-Dual for BEC

$$\overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{end}} - \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\mathrm{begin}} = \int_{\mathrm{Path}} \nabla \overline{F}^{\#}_{\mathrm{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$



Lagrange Pseudo-Dual for BEC

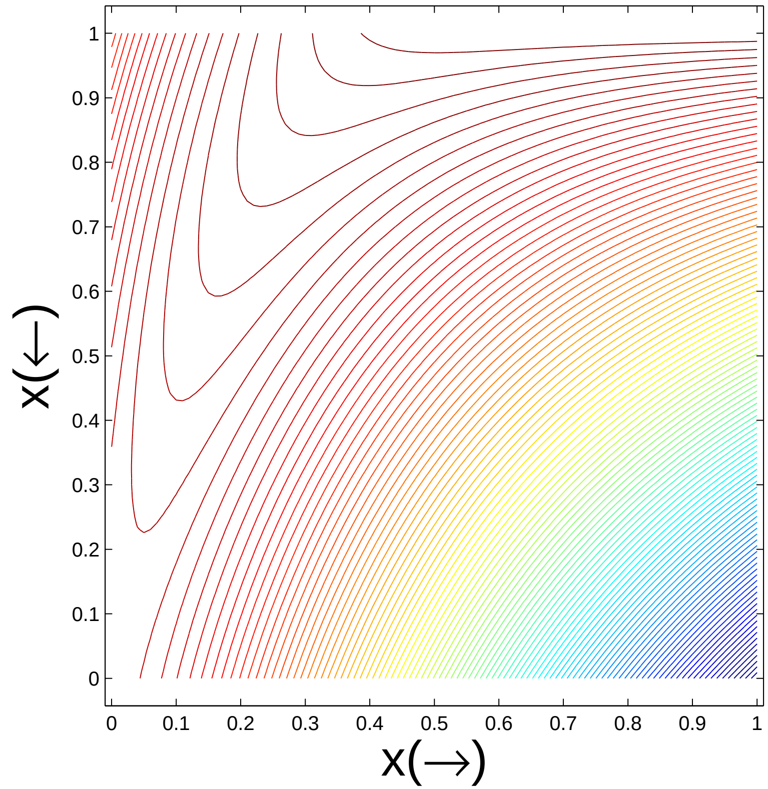
$$\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{end}} - \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})\Big|_{\text{begin}} = \int_{\text{Path}} \nabla \overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x}) \cdot \begin{pmatrix} \mathrm{d}\overrightarrow{x} \\ \mathrm{d}\overleftarrow{x} \end{pmatrix}$$

The connection of SPA and MAP decoding is given by

- Maxwell construction /
- Maxwell decoder.

[Méasson/Montanari/Urbanke, 2005].

This connection can also be expressed in terms of $\overline{F}^{\#}_{\text{Bethe}}(\overrightarrow{x}, \overleftarrow{x})$.



Bethe Free Energy and Weight Spectra

Evaluating the Bethe Free Entropy Along the Cube Diagonal

Evaluating the Bethe Free Entropy Along the Cube Diagonal

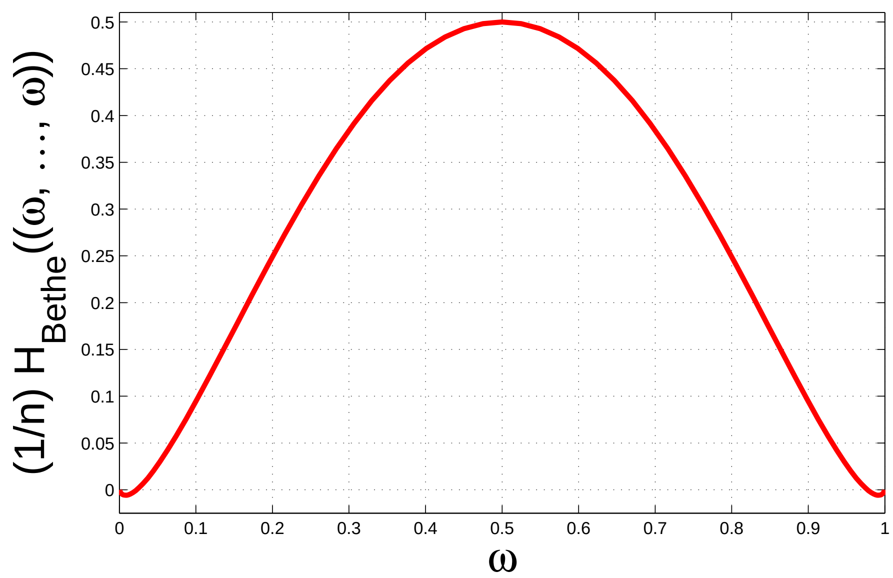
- Take some finite-length (j, k)-regular LDPC code of length n.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j,k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0,1]$ we obtain:

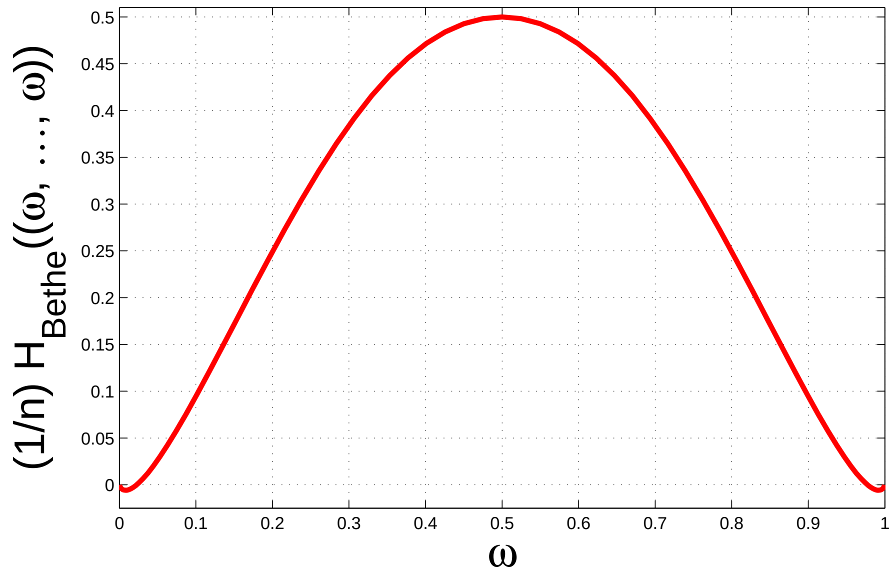
Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain:



Evaluating the Bethe Free Entropy Along the Cube Diagonal

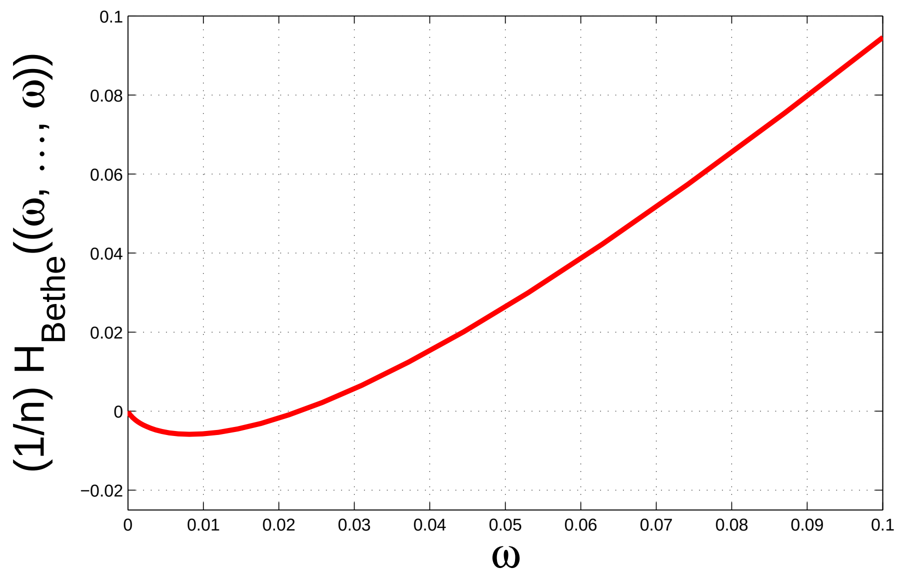
- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain:



- This function equals the exponent of the asymptotic average Hamming weight distribution for Gallager's ensemble of (j, k)-regular LDPC codes!

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain:



- This function equals the exponent of the asymptotic average Hamming weight distribution for Gallager's ensemble of (j, k)-regular LDPC codes!

LABS hp

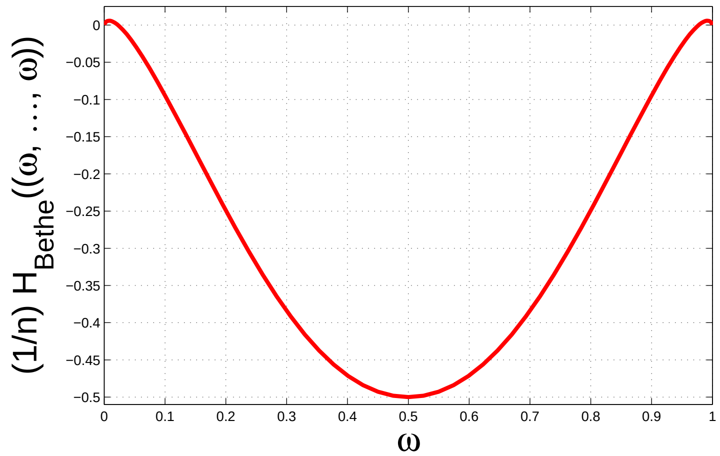
Evaluating the Bethe Free Entropy Along the Cube Diagonal

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.

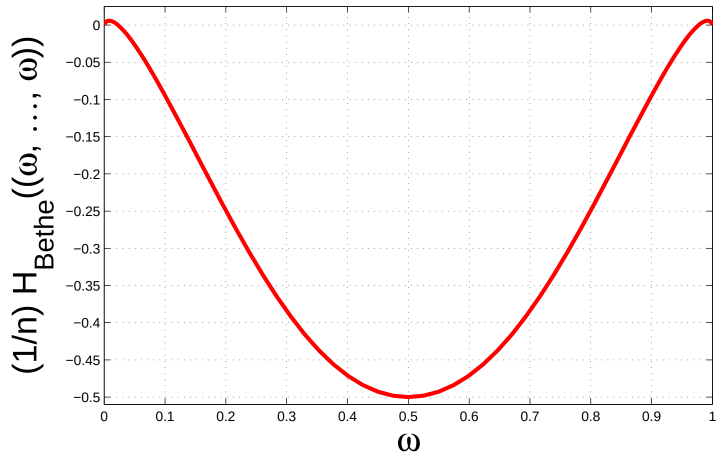
Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



Evaluating the Bethe Free Entropy Along the Cube Diagonal

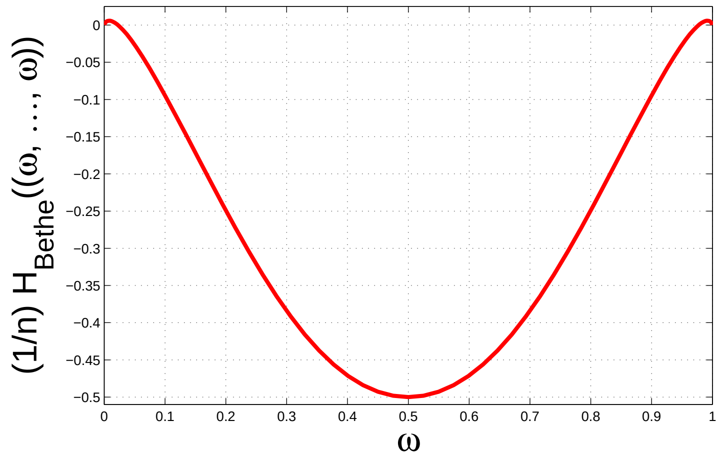
- Let's look at $-\frac{1}{n}H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



- Remember that $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega})$.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

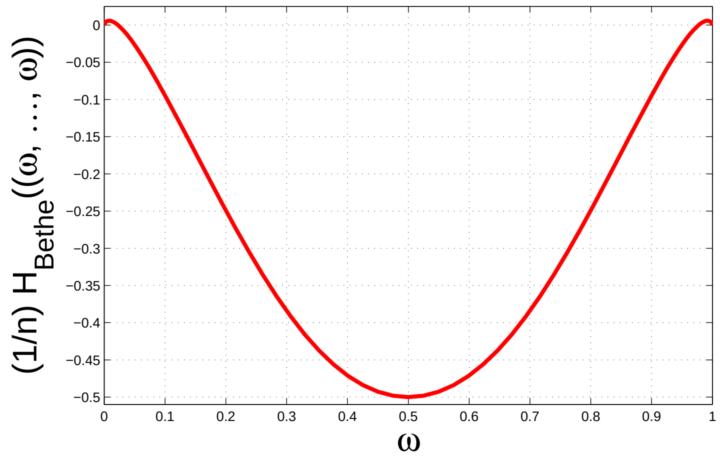
- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



- Remember that $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega})$.
- Remember that $U_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is linear in $\boldsymbol{\omega}$.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

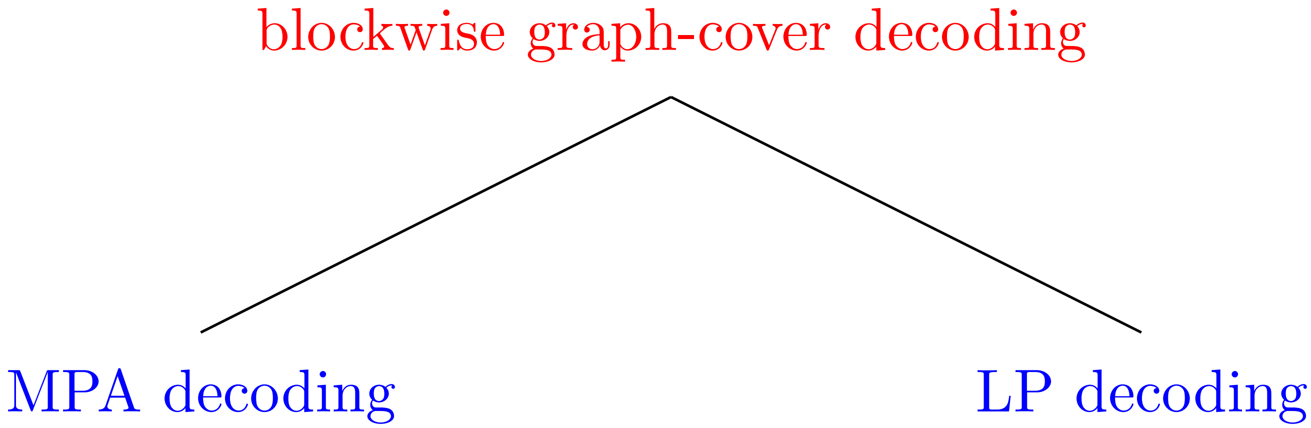
- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



- Remember that $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega})$.
- Remember that $U_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is linear in $\boldsymbol{\omega}$.
- Therefore, we see that for a finite-length code from an ensemble with asymptotically linearly growing minimum Hamming distance, $F_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is not a convex function of $\boldsymbol{\omega}$.

Conclusions

Conclusions



Conclusions

blockwise graph-cover decoding

- MPA decoding
- LP decoding

symbolwise graph-cover decoding

- SPA decoding
- BFE minimization

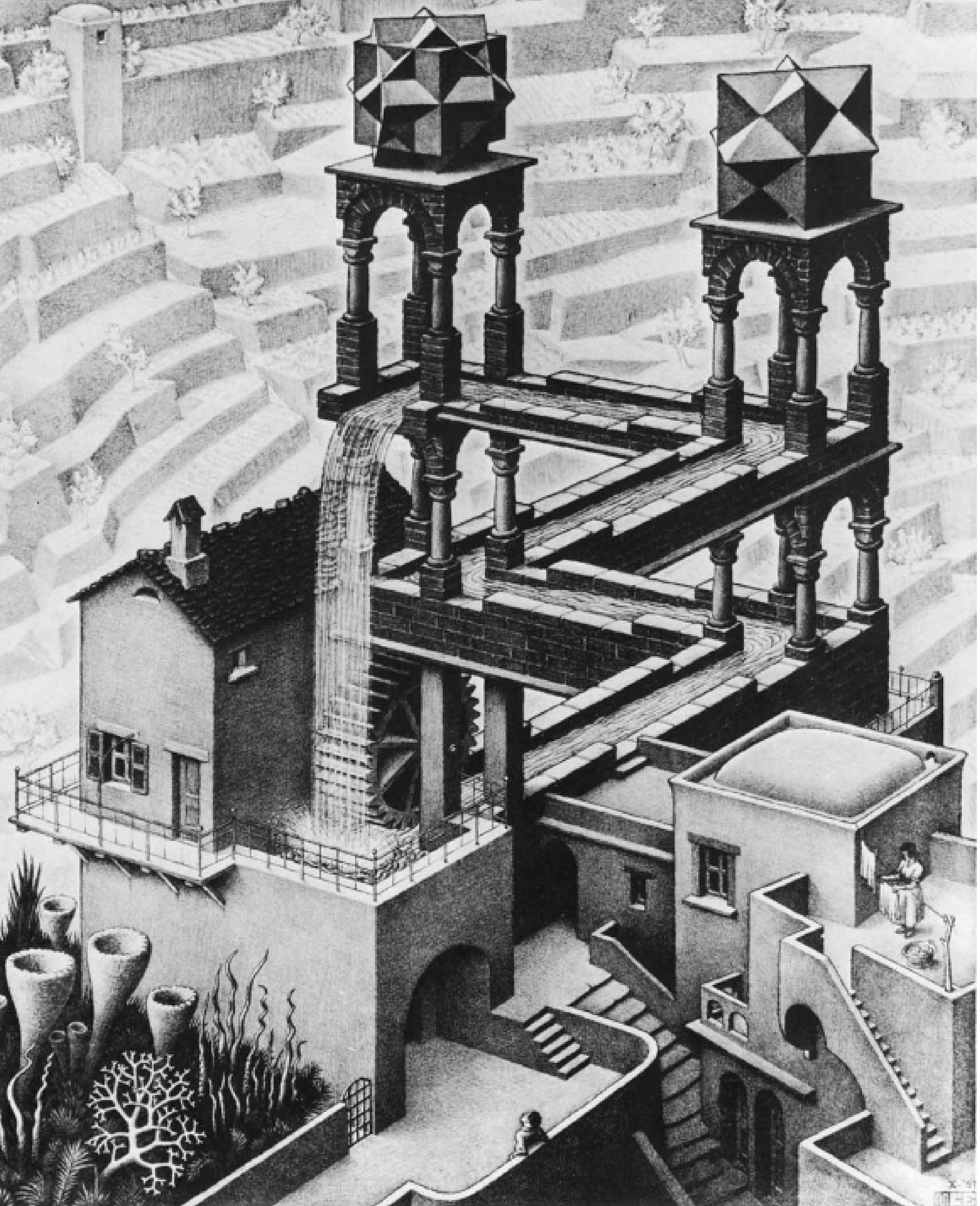
Conclusions

Conclusions

- We have discussed the relevance of Bethe free energy to EXIT charts for the BEC.

Conclusions

- We have discussed the relevance of Bethe free energy to EXIT charts for the BEC.
- We have discussed a connection between Bethe free energy for a finite-length (j, k)-regular LDPC code and the asymptotic growth rate of the average Hamming weight distribution for Gallager's ensemble of (j, k)-regular LDPC codes.



Thank you!

LABS hp