Factor graphs, cont'd

BP on FGs: Input $G=(V\cup F, E)$, T
Output $\{\hat{P}(x_i)\}_{i\in V}$

Initialize $\{(m_{i\to a}, m_{a\to i})\}_{(i,a)\in E} = 1$ or random
$\underbrace{[\;]}_{|X|}$

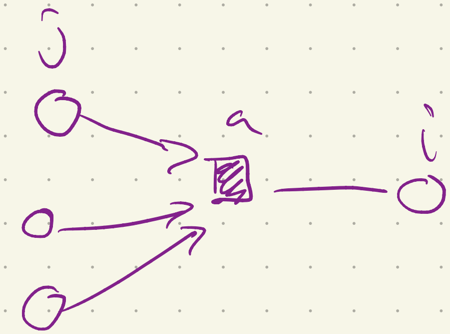
For $t=1 \dots T$

Update $m_{i\to a}$:

$$m_{i\to a}(x_i) = \prod_{b\in\partial i\setminus\{a\}} \tilde{m}_{a\to i}(x_i)$$

Update $\tilde{m}_{a\to i}$:

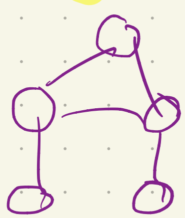
$$\tilde{m}_{a\to i}(x_i) = \sum_{\partial a\setminus\{i\}} f_a(x_{\delta a}) \prod_{j\in\delta a\setminus\{i\}} m_{j\to a}(x_j)$$



Compute Marginal: $m_i(x_i) = \prod_{a\in\partial i} \tilde{m}_{a\to i}(x_i)$

Claim: BP on a FG which is a tree is exact

Notes: Pairwise MRF



BP - is apx $O(|X|^2)$

FG which is a tree

BP is exact

$O(|X|^{\text{max-degree}})$

Example [Decode Low-density parity check]

Def [LDPC] is a family of codes defined as $G=(V\cup F, E)$ with factors acting as parity checks

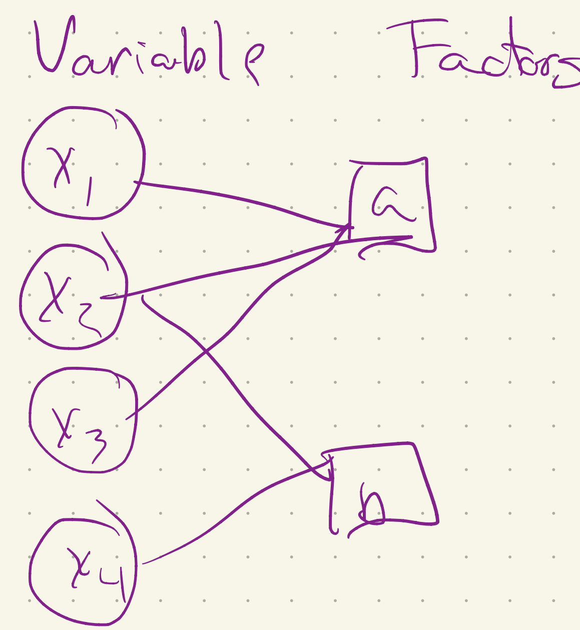
$$f_a(x_{\partial a}) = \mathbb{I}[\oplus x_{\partial a}=0]$$

$\oplus \rightarrow$ XOR, checking parity of the string $x_{\partial a}$

Def A codebook

$\{x \in \mathcal{X}^n \mid x \text{ satisfies all factor parity checks}\}$

$\hookrightarrow \{(0\,0\,0\,0),\ (0\,1\,1\,1),\ (1\,0\,1\,0),\ (1\,1\,0\,1)\}$



Transmitting a codeword $X \in$ Codebook corresponds to sending X over a noisy channel w. behavior $\Pr[Y_i | X_i]$

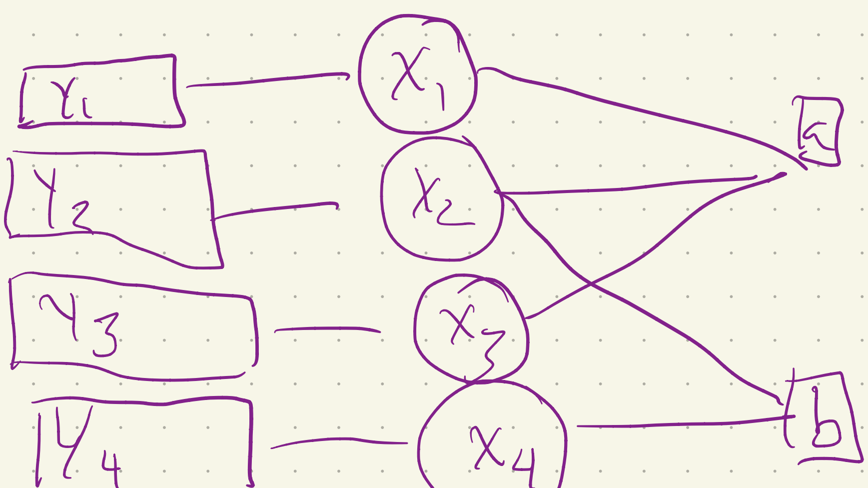
& we need to recover X_i's from Y_i's.

	$X_i = 1$	$X_i = 0$
$Y_i = 1$	.7	.3
$Y_i = 0$	.3	.7

Strategy: Use BP to estimate $\Pr[X_i | Y_1 \dots Y_n]$ $\forall i \in [n]$

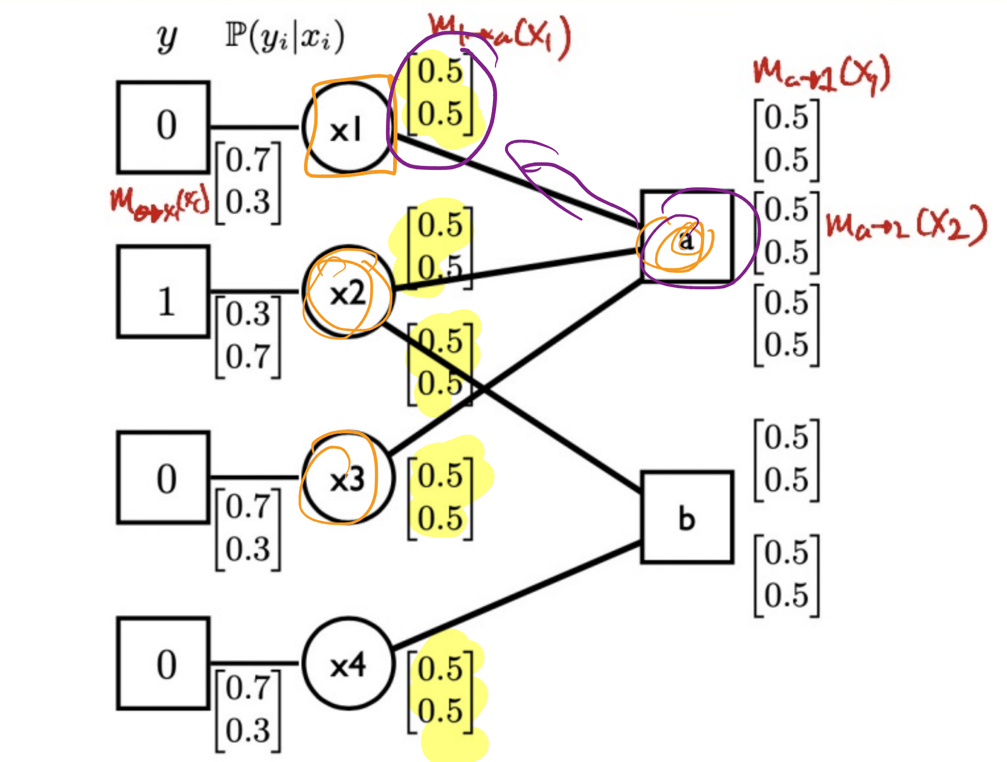
& guess $\tilde{X}_i = \frac{1}{2}\left(\text{sign}\left(\log\left[\frac{\Pr[X_i = 1 | Y]}{\Pr[X_i = 0 | Y]}\right]\right) + 1\right)$

What is our FG for BP example



$$f_a = \begin{array}{c|cccc} & \multicolumn{4}{c}{x_2, x_3} \\ x_1 & 0,0 & 0,1 & 1,0 & 1,1 \\ 0 & 1 & 0 & 0 & 1 \\ & 0 & 1 & 1 & 0 \end{array}$$

$$f_b = \begin{array}{c|cc} & \multicolumn{2}{c}{x_2} \\ x_4 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}$$



$$\nu_{i\to a}^{(t+1)}(x_i) = \mathbb{P}(y_i|x_i)\prod_{b\in\partial i\setminus\{a\}} \tilde{\nu}_{b\to i}^{(t)}(x_i)$$

$$\tilde{\nu}_{a\to i}^{(t+1)}(x_i) = \sum_{x_{\partial a\setminus\{i\}}} \prod_{j\in\partial a\setminus\{i\}} \nu_{j\to a}(x_j)\mathbb{I}(\oplus x_{\partial a} = 0)$$

$m_{i\to a}(x_i)$

$$m_{a\to 1}(x_1) = \sum_{x_{\partial a\setminus\{1\}}} \prod_{j\in\partial a\setminus\{1\}} m_{j\to a}(x_j) \cdot \mathbb{I}[\oplus x_{\partial a} = 0]$$

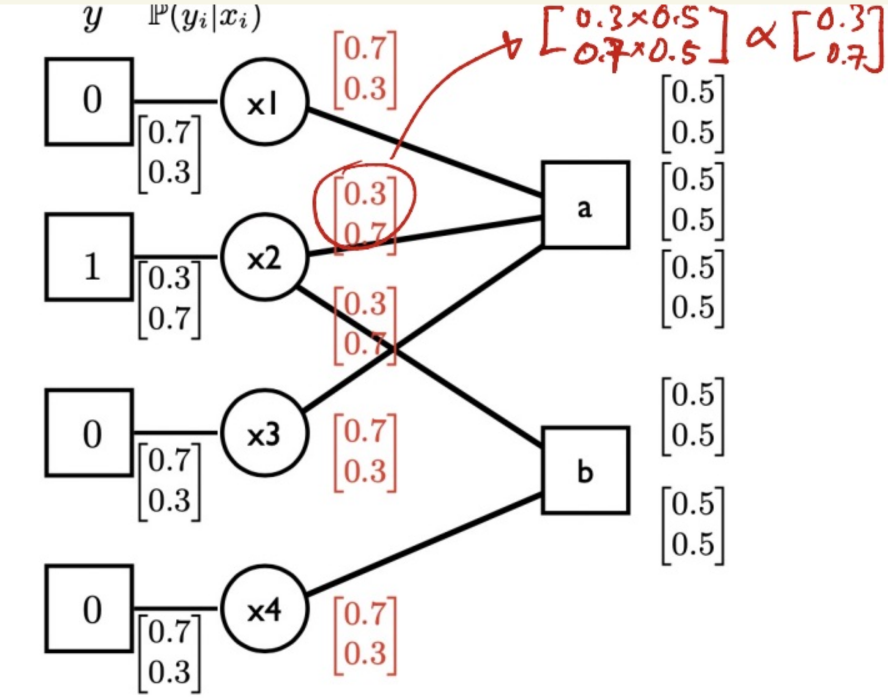
$x_2 = x_3 = 0$ $\quad .3 \cdot .7$

~~$x_2 = 1,\ x_3 = 0$ + $.7 \cdot .7$~~

~~$x_2 = 0,\ x_3 = 1$~~

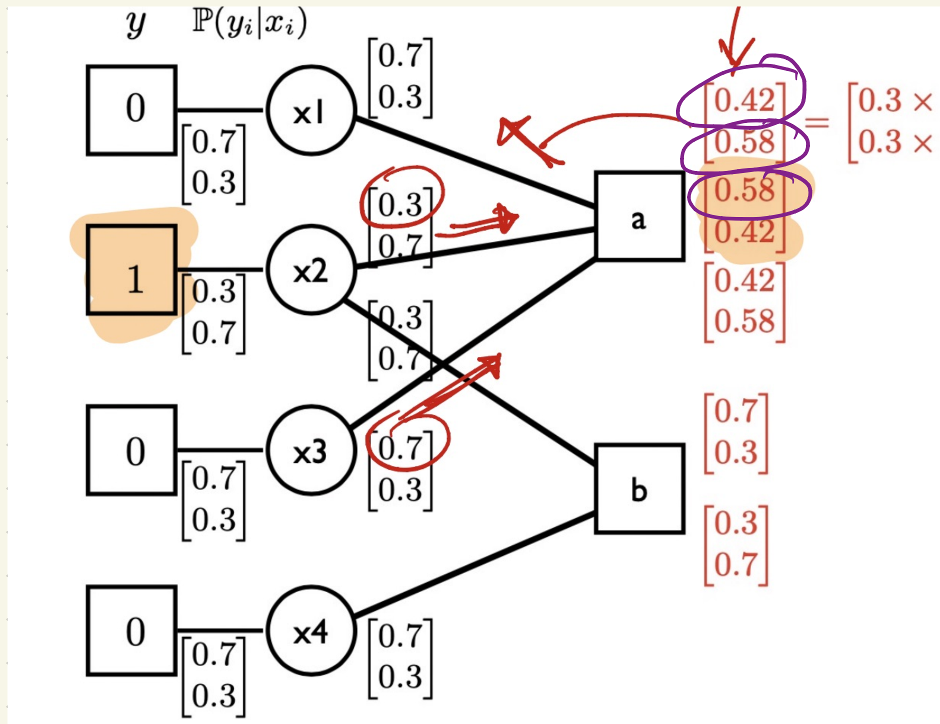
$x_2 = x_3 = 1$ $\quad .7 \cdot .3$

$= .42$



$\nu_{i\to a}^{(t+1)}(x_i) = \mathbb{P}(y_i|x_i)\prod_{b\in\partial i\setminus\{a\}} \tilde{\nu}_{b\to i}^{(t)}(x_i)$

$\tilde{\nu}_{a\to i}^{(t+1)}(x_i) = \sum_{x_{\partial a\setminus\{i\}}} \prod_{j\in\partial a\setminus\{i\}} \nu_{j\to a}(x_j)\mathbb{I}(\oplus x_{\partial a}=0)$

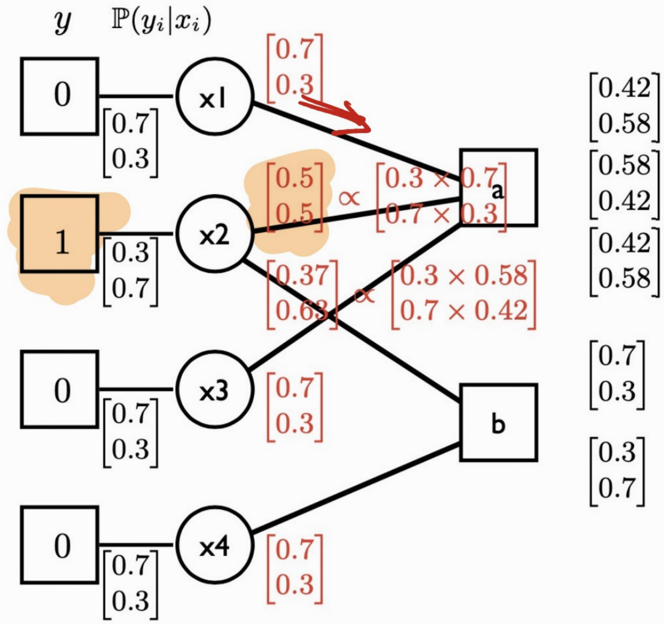


$\nu_{i\to a}^{(t+1)}(x_i) = \mathbb{P}(y_i|x_i)\prod_{b\in\partial i\setminus\{a\}} \tilde{\nu}_{b\to i}^{(t)}(x_i)$

$\tilde{\nu}_{a\to i}^{(t+1)}(x_i) = \sum_{x_{\partial a\setminus\{i\}}} \prod_{j\in\partial a\setminus\{i\}} \nu_{j\to a}(x_j)\mathbb{I}(\oplus x_{\partial a}=0)$

$m_{2\to a}(x_2) = P[y_2|x_2) \prod_{b\in\partial 2\setminus\{a\}} \tilde{m}_{b\to 2}(x_2)$

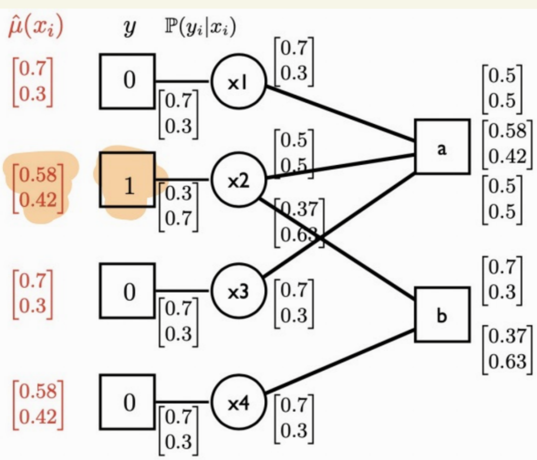
$m_{2\to b}(x_2) = Pr[y_2|x_2] \prod_{a\in\partial 2\setminus\{b\}} \tilde{m}_{a\to 2}(x_2)$



$$\nu^{(t+1)}_{i\to a}(x_i) = \mathbb{P}(y_i|x_i)\prod_{b\in\partial i\setminus\{a\}} \tilde{\nu}^{(t)}_{b\to i}(x_i)$$

$$\tilde{\nu}^{(t+1)}_{a\to i}(x_i) = \sum_{x_{\partial a\setminus\{i\}}} \prod_{j\in\partial a\setminus\{i\}} \nu_{j\to a}(x_j)\mathbb{I}(\oplus x_{\partial a} = 0)$$

$$m_{2\to b}(x_2) = \Pr(y_2|x_2)\cdot \prod_{a\in\partial 2\setminus\{b\}} \tilde{m}_{a\to 2}(x_2)$$



$$\nu^{(t+1)}_{i\to a}(x_i) = \mathbb{P}(y_i|x_i)\prod_{b\in\partial i\setminus\{a\}} \tilde{\nu}^{(t)}_{b\to i}(x_i)$$

$$\tilde{\nu}^{(t+1)}_{a\to i}(x_i) = \sum_{x_{\partial a\setminus\{i\}}} \prod_{j\in\partial a\setminus\{i\}} \nu_{j\to a}(x_j)\mathbb{I}(\oplus x_{\partial a} = 0)$$

Codebook

{0000, 0111, 1010, 1101}

0100 → y

0000

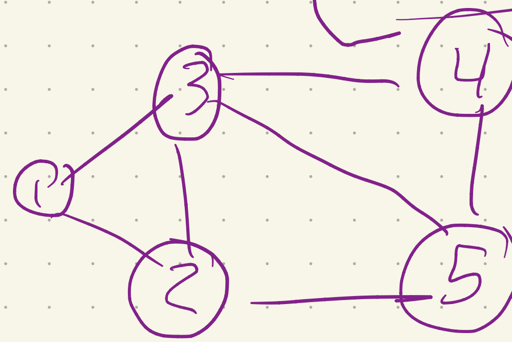
↳ min hamming dist code from y

Junction trees (Elimination algs on junction trees will be exact)

- Fixing tree + order, inference will be easy
- Data structure designed for efficient elimination

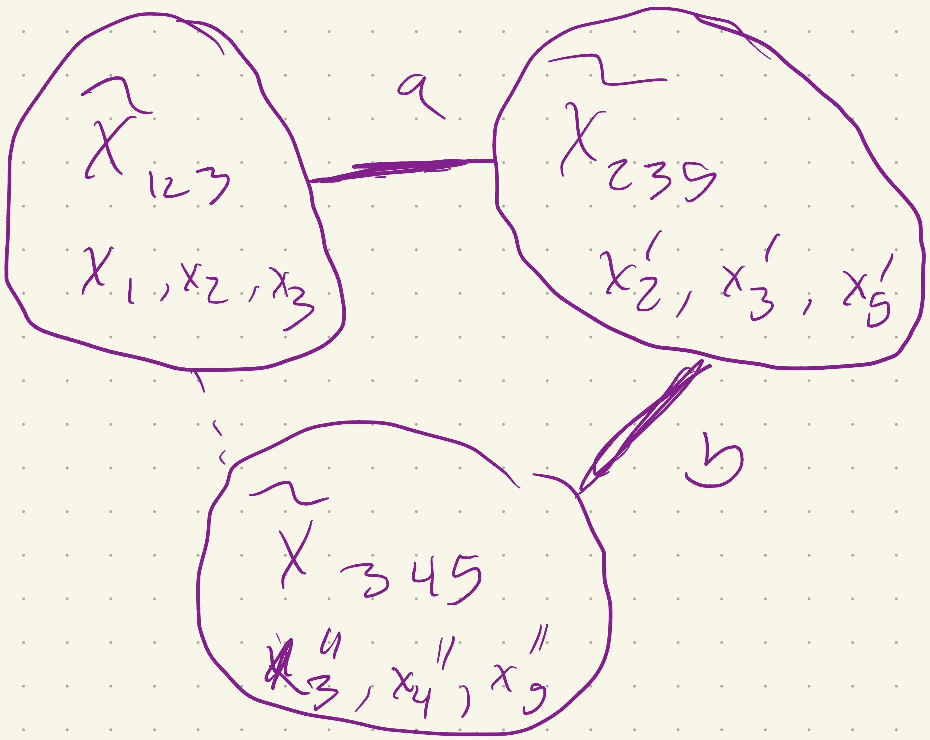
MRF ⇒ Clique tree

(non unique)



$$P(X) \propto f_{123} f_{235} f_{345}$$

- Create a node $\forall c \in \mathcal{C}$
- Create Local copies of variables
- Connect to form a tree



If my clique tree has a set of edges which allow enforcing local variable consistency, then we have global consistency.

Q: when will a clique tree for MRF allow for global consistency?

D