- PGMs + Markov Properties.
- Inference problems — Efficient algs
  - Variational methods & sampling
- Learning GMs
- Extra

$\boxed{\text{Inference}}$ Given MRF $G=(V,E)$ $P(x) = \frac{1}{Z}\prod_{c\in e} f_c(x_c)$

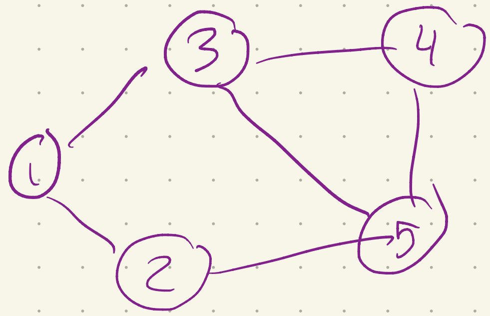
(1) Calculate Marginal $P(x_1)$
(2) Calculate $\arg\max_x P(x)$
(3) Calculate Z
(4) Sample from $P(x)$

(1) $Y \sim$ observation — $X \sim$ Cause, State — $P(X_1) \to$ denoise/estimate

Noisy observation	Pixel values
Spectral measurement	MRI image
Noisy bits recieved	bits sent

★ Elimination Alg (Exact, W.C.RT = $O(|\mathcal{X}|^n)$) → [Computing $P(X=x)$ is #P complete]
For marginalization

$$P(x) = \frac{1}{Z} f_{12}(x_1, x_2)\, f_{13}(x_1, x_3)\, f_{25}(x_2, x_5)\, f_{345}(x_3, x_4, x_5)$$



How to compute $P(X_1)$?

– Brute force $|\mathcal{X}|$ $|\mathcal{X}|^4$

$|\mathcal{X}|$ → enum over X_1

$|\mathcal{X}|^4$ → sum over $X_2 \dots X_5$

Fix elimination order (5, 4, 3, 2)

$$P(x_1) = \frac{1}{Z} \sum_{x_2, x_3, x_4} f_{12}(x_1, x_2)\, f_{13}(x_1, x_3) \underbrace{\sum_{x_5} f_{25}(x_2, x_5)\, f_{345}(x_3, x_4, x_5)}_{m_5(x_2, x_3, x_4)}$$

→ $|\mathcal{X}|^3$ enum, $|\mathcal{X}|$ sum

$$= \frac{1}{Z} \sum_{x_2, x_3} f_{12}(x_1, x_2)\, f_{13}(x_1, x_3) \underbrace{\sum_{x_4} m_5(x_2, x_3, x_4)}_{m_4(x_2, x_3)}$$

→ $|\mathcal{X}|^2$ enum, $|\mathcal{X}|$ sum

$$= \frac{1}{Z} \sum_{x_2} f_{12}(x_1, x_2) \underbrace{\sum_{x_3} f_{13}(x_1, x_3)\, m_4(x_2, x_3)}_{m_3(x_1, x_2)}$$

→ $|\mathcal{X}|^2$ enum, $|\mathcal{X}|$ sum

$$= \frac{1}{Z} \sum_{2} f_{12}(x_1, x_2)\, m_3(x_1, x_2) \rightarrow m_2(x_1)$$

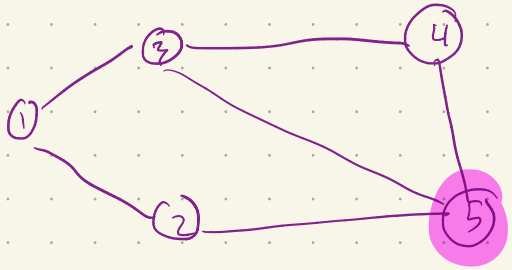
→ $|\mathcal{X}|$ enum, $|\mathcal{X}|$ sum

[Total $|\mathcal{X}|^4$ vs $|\mathcal{X}|^5$]

Def Reconstituted graph of $G=(V,E)$ wrt Elim ordering $P=(p_1, \ldots p_n)$

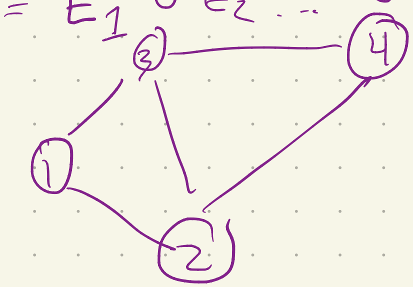
$\tilde{G}=(V,\tilde{E})$ $\quad \tilde{E} = E_1 \cup E_2 \ldots \cup E_n$

↳ add edge (i,j) if $(i, p_1) \in E$, (j, p_1)



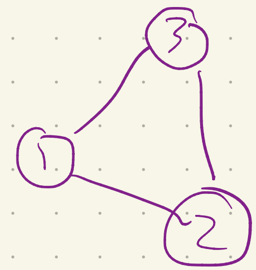
E

$\Rightarrow$



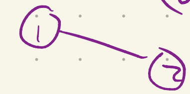
E_1

→



E_2

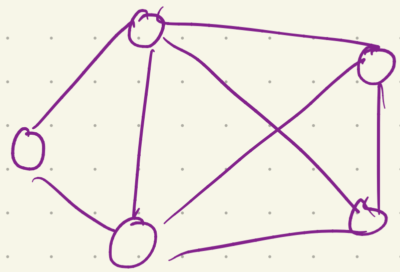
→



E_3

E_4

$\tilde{E} =$



→ tree width is 4

Def Tree Width is $\min_{\text{all orderings}} \{$ max size of a clique in reconstituted graph $\}$

size of largest vertex set in tree decomposition

* Elimination Alg for MAP Inference $\arg\max_x P(x) = x^*$

Q: Can we use $\arg\max_{x_1} P_1(x_1) \stackrel{?}{=} x_1^*$

$$\| $$

$$\arg\max_{x_1} \sum_{x_2,\dots,x_n} P(x_1, x_2^n) \neq \arg\max_{x_1} \max_{x_2 \dots x_n} P(x_1, x_2^n)$$

A: No, not like this

So rather than "pushing in" sums, we "push in" maxes!

Start w. an order, $(5, 4, 3, 2, 1)$

$$\arg\max_{x_1 \dots x_5} P(x) = \cancel{\frac{1}{Z}} \arg\max_{x_1 \dots x_4} f_{12}(x_1, x_2) f_{13}(x_1 x_3) \underbrace{\overbrace{\max_{x_5} f_{25}(x_2 x_5) f_{345}(x_3 x_4 x_5)}^{m^*_5(x_2, x_3, x_4)}}_{\arg\max_{x_5} := x^*_5(x_2, x_3, x_4)}$$

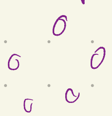
$$= \arg\max_{x_1 \dots x_3} f_{12}(x_1, x_2) f_{13}(x_1 x_3) \underbrace{\max_{x_4} m^*_5(x_2, x_3, x_4)}_{m^*_4(x_2, x_3),\ x^*_4(x_2, x_3)}$$

$$= \arg\max_{x_1, x_2} f_{12}(x_1, x_2) \max_{x_3} \overbrace{f_{13}(x_1, x_3) m^*_4(x_2, x_3)}^{m^*_3(x_1, x_2)}$$

$$= \arg\max_{x_1} \boxed{\max_{x_2} f_{12}(x_1, x_2)\, m^*_3(x_1, x_2)} \rightarrow m^*_2(x_1) \quad x^*_2(x_1) = x_2$$

$$= \arg\max_{x_1} \boxed{m_2(x_1)} = m^*(x_1), \quad x_1^*$$

Tree width is our first notion of simplicity of a graph that we can use to capture computational complexity

small simple	chain	tree	2-d grid	3-d grid	Erdös-Reyni	K_n Complete
 $E = \emptyset$						
$Tw = 1$	$Tw = 2$	$Tw = 2$	$Tw = \sqrt{n}$	$Tw = n^{2/3}$	$p = \frac{c}{n}$	$Tw = n$
$O(n\lvert X\rvert)$	$\lvert X\rvert^2$	$\lvert X\rvert^2$	$\lvert X\rvert^{\sqrt{n}}$	$\lvert X\rvert^{n^{2/3}}$	$\lvert X\rvert^n$	$O(\lvert X\rvert^n)$
max-d $= 0$	2	2	4	6	c	n

Q: Is inference really as hard on E-R as on K_n?

A: Exact inference, yes. Approx?