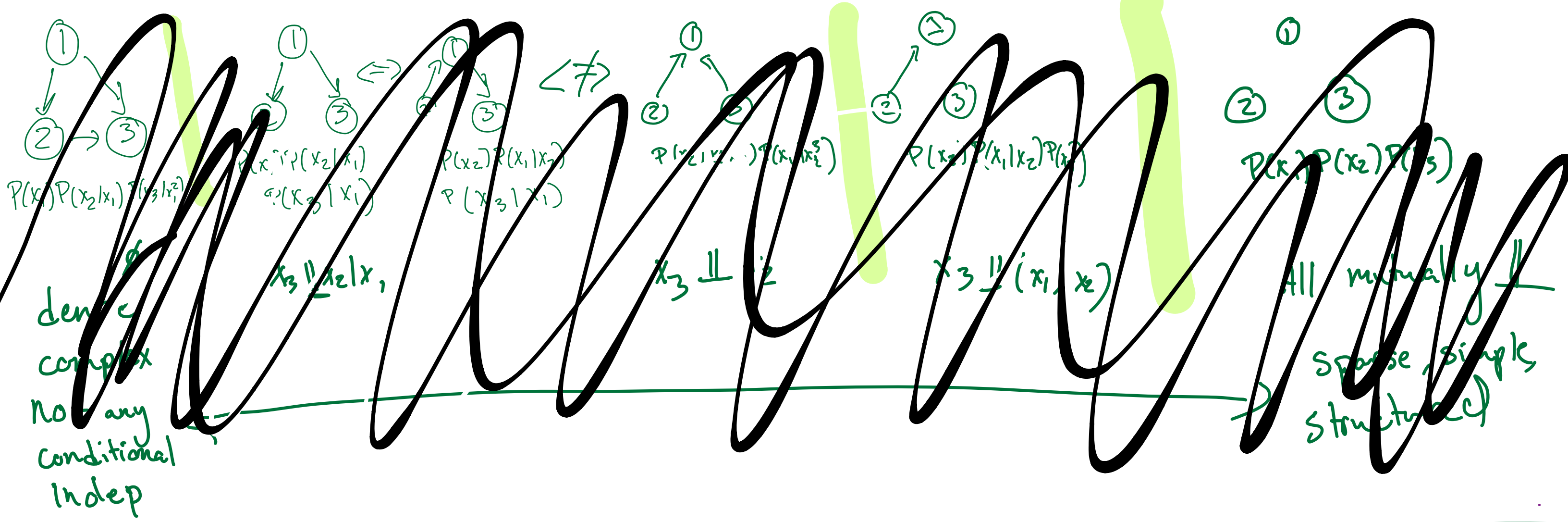


Markov Properties of Directed graphical models

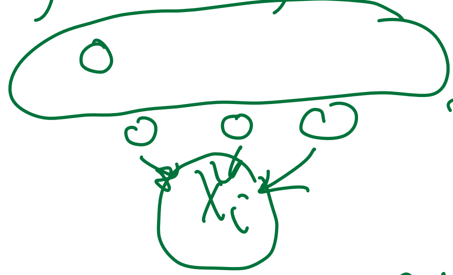
① Directed Ordered Markov Property

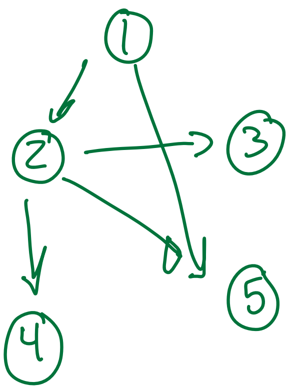
Given a DAG $G=(V,E)$ $\not p$ $(1,\dots,n)$

$$X_i \perp\!\!\!\perp X_{Pr_i \setminus \pi(i)} \mid X_{\pi(i)}$$

$Pr(i)$ = nodes before i

Pf: Chain rule □





Ordering $(1,2,3,4,5)$

$x_3 \perp\!\!\!\perp x_1 \mid x_2$

$x_4 \perp\!\!\!\perp x_1, x_3 \mid x_2$

$x_5 \perp\!\!\!\perp x_1, x_3 \mid x_1, x_2$

Ordering $(1,2,4,5,3)$

$x_4 \perp\!\!\!\perp x_1 \mid x_2$

$x_3 \perp\!\!\!\perp x_1, x_4, x_5 \mid x_2$

⋮

② Directed Local Markov Property

$$X_i \perp\!\!\!\perp X_{nd(i) \setminus \pi(i)} \mid \pi(i)$$

$nd(i)$ = nondescendents of i

set of nodes j where $\nexists$ directed path $i, \dots, j$

x_1 $\emptyset$

x_2 $\emptyset$

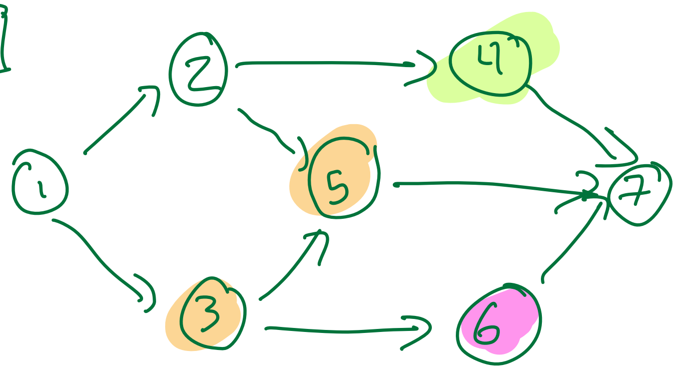
$x_3 \perp\!\!\!\perp x_1, x_4, x_5 \mid x_2$

Q: How do you determine all CI given a G?

$\exists$ CI Test$(A, B, C, G) \to \{0,1\}$ where $CI(A,B,C,G) = 1$ iff $A \perp\!\!\!\perp C \mid B$ in G?

A: Yes, the Bayes Ball Alg.

$\boxed{G}$



$A = \{4\}$

$B = \{3, 5\}$

$C = \{6\}$

is $A \perp\!\!\!\perp C \mid B$?

(3) Directed Global Markov Property

For any $A, B, C \subseteq V$, $X_A \perp\!\!\!\perp X_C \mid X_B$ if A is d-separated by B from C

A is d-separated by B from C if $\forall$ undirected, simple (nonrepeating) paths $a_1 \in A \ldots\ldots c \in C$ the path is blocked by B.

Path $a_1 \ldots\ldots c$ is blocked by B if $\exists$ 3 nodes

$\boxed{\text{Intuition}}$

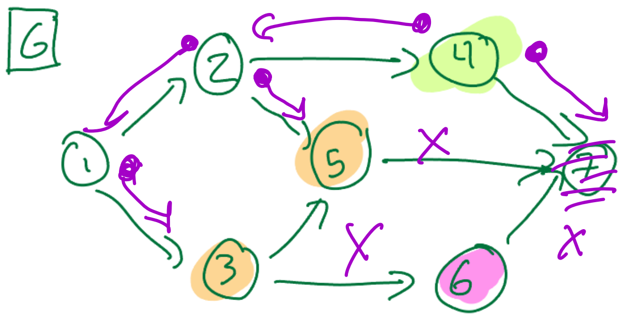
(a) .... $1 \to 2 \to 3$ .... c & $2 \in B$ $\qquad X_1 \perp\!\!\!\perp X_3 \mid X_2$

(a) .... $1 \leftarrow 2 \leftarrow 3$ .... c & $2 \in B$ $\qquad X_1 \perp\!\!\!\perp X_3 \mid X_2$

(a) .... $1 \leftarrow 2 \to 3$ .... c & $2 \in B$ $\qquad X_1 \perp\!\!\!\perp X_3 \mid X_2$

(a) .... $1 \to 2 \leftarrow 3$ .... c & $2 \notin B$ $\qquad X_1 \perp\!\!\!\perp X_3$

Bayes' Ball identifies all CI in a graph



$A = \{4\}$

$B = \{3, 5\}$

$C = \{6\}$

7 acts as a block $\notin B$

Thm: All of the following are equivalent

(a) $P(x)$ factorizes according to $G = (V, E)$ [D Global]

(b) $P(x)$ satisfies the directed ordered ~~global~~ Markov prop

(c) $P(x)$ satisfies the directed DAG local Markov prop

$\boxed{\text{Pf}}$

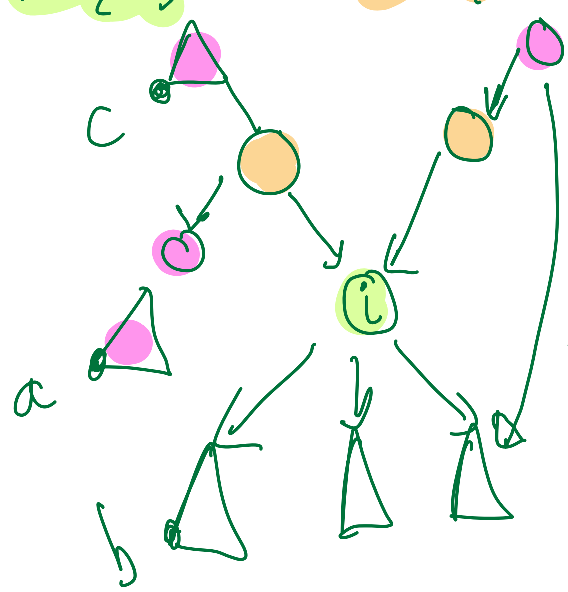
b ⇒ c

$A = \{i\}$ $B = \pi_i$ $C = Nd\, i \setminus \pi_i$

$c \perp\!\!\!\perp i \mid \pi_i$

$a \perp\!\!\!\perp i \mid \pi_i$

$b \not\perp\!\!\!\perp i \mid \pi_i$



descendents of i

---

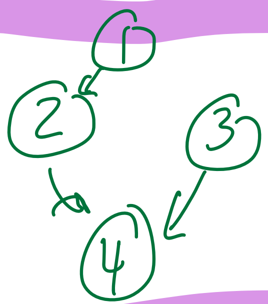
c ⇒ a any topological ordering $(1, \dots, n)$

$$P(x) = \prod_i P(x_i \mid x_1^{i-1}) = \prod_i P(x_i \mid x_{\pi(i)})$$

chain local MP

$a \Rightarrow b$ in a separate PDF online (by ind)

Let $I(G)$ = set of all $\perp\!\!\!\perp$ implied by $G=(V,E)$



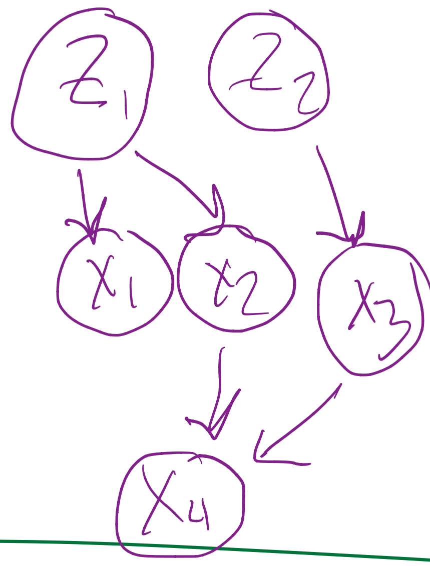
$$I(G) = \{ X_3 \perp\!\!\!\perp X_1, X_2$$
$$X_4 \perp\!\!\!\perp X_1 \mid X_2, X_3 \}$$

Let $I(D)$ be all $\perp\!\!\!\perp$ implied by factorization of P

Consider $Z_1 \perp Z_2 \sim \text{Bern}(1/3)$

$$X_1 = Z_1 \qquad X_2 = Z_1 \qquad X_3 = Z_2, \; X_4 = \underset{\text{XOR}}{X_2 \oplus X_3}$$

$$I(P) = \{ X_3 \perp\!\!\!\perp X_1, X_2$$
$$X_1 \perp\!\!\!\perp X_3, X_4 \mid X_2$$
$$X_2 \perp\!\!\!\perp X_3, X_4 \mid X_1 \}$$



Remarks ① $\exists P(X)$ s.t. no DAG perfectly captures $\perp\!\!\!\perp$ (ie $I(G) \neq I(P)$)

[Equality of RV's is a weird edge case]

$X_1 = X_2$ $\qquad$ $X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_3 = X_1 + Z$ $\qquad$ $X_2 \not\perp\!\!\!\perp X_3 \mid X_1$

(2) $\forall G, \exists P$ s.t. $I(G) = I(P)$

(3) $\exists G_1 \neq G_2$ s.t. $I(G_1) = I(G_2)$

$\circ \to \circ \to \circ \qquad \circ \to \circ \leftarrow \circ$

<u>Def</u> G is a <u>perfect map</u> for P if $I(P) = I(G)$

<u>Def</u> G is an <u>I-map</u> for $P(X)$ if $I(G) \subseteq I(P)$

<u>Def</u> G is a minimal I-map for $P(x)$ if

1) Remove an edge in $G \Rightarrow$ G no longer an I-map

(4) Min I-MAP isn't unique

- different ordering of vars
- even w = ordering,

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[r] & 3 \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r] \arrow[rr, bend left] & 2 & 3 \end{tikzcd}$$

$x_1 = x_2$

$x_3 = x_1 + z$

Probabilistic GMs

"P factorizes" wrt G

Graph Separation $\Longleftrightarrow$ $I(G) = \bigcap I(P)$

$\Uparrow$

Graphs G $\Longrightarrow$ Factorization $P(x) = \prod$ $\Longrightarrow$ Cond Ind $I(P)$

$\boxed{\text{Undirected GMs}}$

Def Undirected graph $G = (V, E)$ $\nearrow (i,j) = (j,i)$

- A clique $C \subseteq V$ is a set of nodes s.t. $\forall i \neq j$, $i, j \in C$, $(i,j) \in E$

- A clique C is a maximal clique if $\nexists\, i \notin C$ s.t. $\{i\} \cup C$ is a clique

Denote the set of all cliques as $\mathcal{C}$

Def: An *undirected* GM on $G=(V,E)$ is a family of distributions on $X=[X_1,\dots,X_n]$ that factorizes as

[AKA Markov Rand Fields]

$$P(X) = \frac{1}{Z}\prod_{c\in\mathcal{C}} f_c(\pi_c)$$

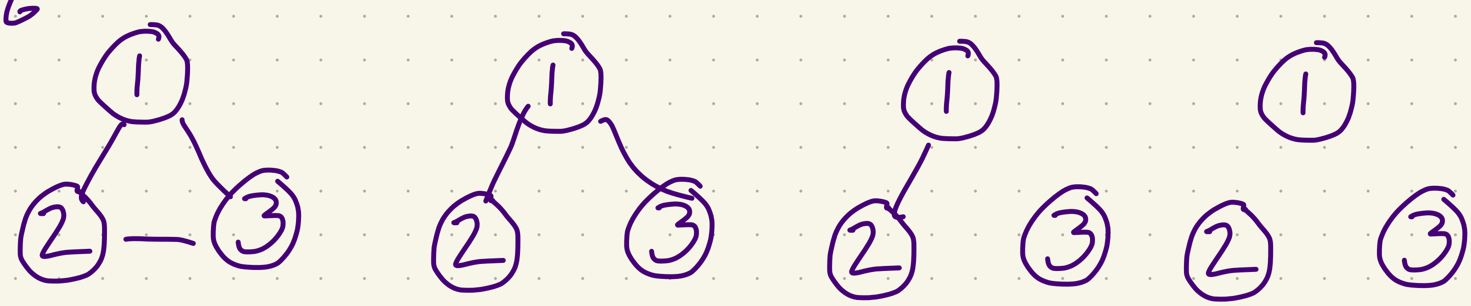
$f_c : \mathcal{X}^{|C|} \to \mathbb{R}_+$ are called *factors* or *computability fns*

$$Z := \sum_{x\in\mathcal{X}^n}\left\{\prod_{c\in\mathcal{C}} f_c(\pi_c)\right\}$$ is called a partition fn.

Unstructured $\longleftrightarrow$ Struct

Dense $P(x)$ dist: $|I(P)|$ small, graph dense — $|I(P)|$ large, graph sparse — Sparse

G



$P(x)$

$P(x) \propto f_{12}(x_1, x_2)\cdot f_{13}(x_1, x_3)$

$\downarrow$ HW1

$X_2 \perp\!\!\!\perp X_3 \mid X_1$

$P(x) \propto f_{12}(X_1,X_2)\cdot f_3(X_3)$

$\downarrow$

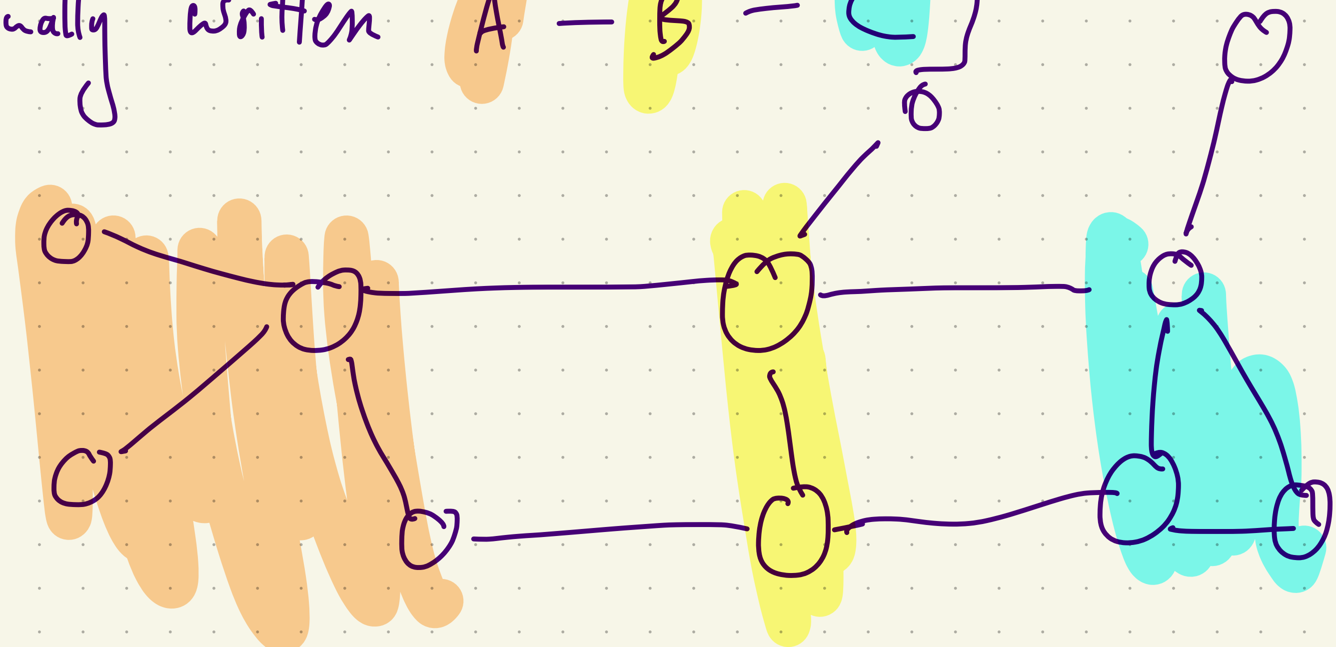
$X_3 \perp\!\!\!\perp X_1, X_2$

X_1, X_2, X_3 mutually ind

Def [Graph Separation]

$B \subseteq V$ separates $A, C \subseteq V$ if every path from any $a \in A$, $c \in C$ passes thru B.

[Casually written $A - B - C$]



Def [Global Markov Propery (G)]

A distribution $P(X)$ satisfies the global markov property if for any $A - B - C$ we have $X_A, X_C \perp\!\!\!\perp X_B$.

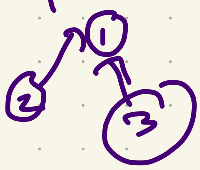
<u>Punchline</u>: Factorization $\Rightarrow$ (G)

$\exists P$ s.t.

$I(G) \subsetneq I(P)$

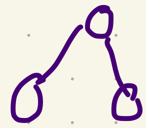
strict $\forall G$

Q: Give a directed Bayesian Net that can't be represented using a MRF



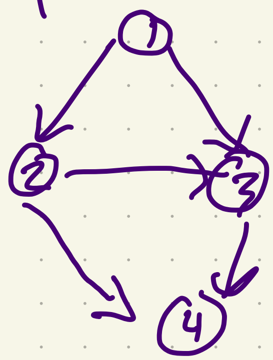
BN

$X_2 \perp\!\!\!\perp X_3$

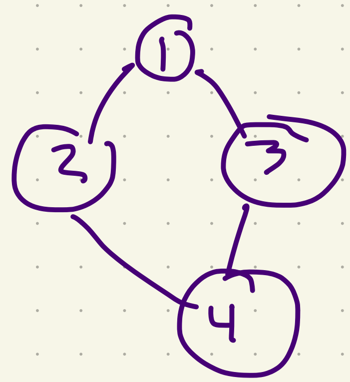


$X_2 \perp\!\!\!\perp X_3 | X_1$

Q: Give a MRF that can't be represented usin a BN



$X_1 \perp\!\!\!\perp X_4 | X_2, X_3$



$X_2 \perp\!\!\!\perp X_3 | X_1, X_4$

Def [Local Markov Property (L)]

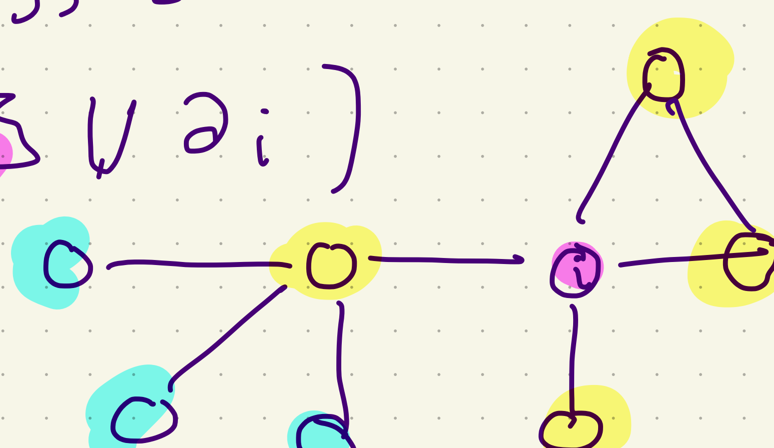
P(X) satisfies (L) w.r.t. G if

$$X_i \perp\!\!\!\perp X_{rest} | X_{\partial_i}$$

∂_i = boundary or neighborhood of i

$j \in \partial_i$ if $(i,j) \in E$

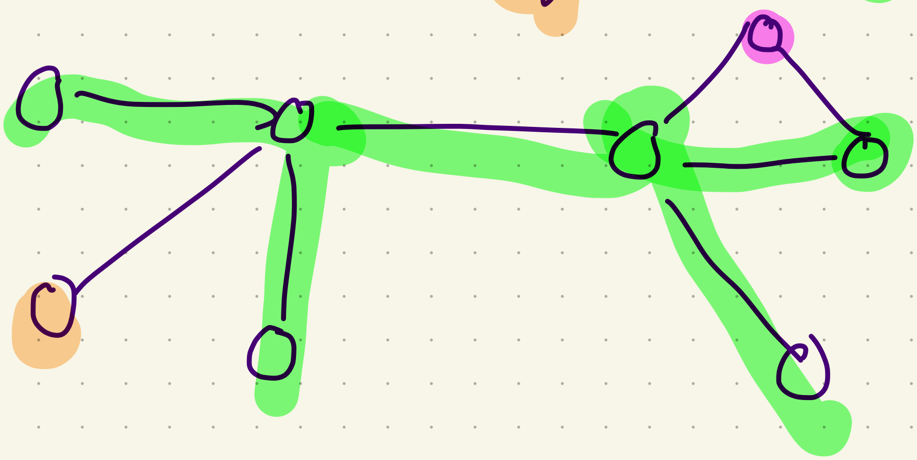
$$X_{rest} = V \setminus (\{i\} \cup \partial_i)$$



Def [Pairwise Markov Property (P)]

$P(X)$ satisfies (P) wrt G if

$$X_i \perp\!\!\!\perp X_j \mid V \setminus \{i,j\} \quad \forall\, i, j \text{ not connected}$$



Claim

$$(G) \Rightarrow (L) \Rightarrow (P)$$

(G) $X_A \perp\!\!\!\perp X_C \mid X_B$

(L) $X_i \perp\!\!\!\perp X_{rest} \mid X_{\partial_i}$

(P) $X_i \perp\!\!\!\perp X_j \mid V \setminus \{i,j\}$ when $(i,j) \notin E$

$P \Rightarrow L$

Proof

Follows from $i - \partial_i - \text{Rest}$

aka i is separated from rest by ∂i.

$L \Rightarrow P$

Pf $X_i \perp\!\!\!\perp X_{V \setminus \{i\} \cup \partial_i} \mid X_{\partial_i}$ Fix $j \neq i$, $(i,j) \notin E$

(a) $\Rightarrow X_i \perp\!\!\!\perp X_{V \setminus \{i\} \cup \partial_i} \mid X_{V \setminus \{i,j\}}$

(b) $\Rightarrow X_i \perp\!\!\!\perp X_j \mid X_{V \setminus \{i,j\}}$

(a) Follows b/c

$$A \perp\!\!\!\perp B \mid C \Rightarrow A \perp\!\!\!\perp B \mid C \cup h(B) \quad \forall h$$

& $X_{\partial_i} \subseteq V \setminus \{i, j\}$

(b) Follows b/c $j \notin \partial i$ $j \neq i$

So $j \in V \setminus \{i\} \cup \partial i$

And $X \perp\!\!\!\perp Y \mid Z \Rightarrow X \perp\!\!\!\perp h(Y) \mid Z \quad \forall h$

[Data Processing inequality] $X - Z - Y - h(Y)$

$\boxed{P \Rightarrow G}$ Claim: If $P(X) > 0 \ \forall X$, then $P \Rightarrow G$

[P is a PMF, $P(x) > 0 \quad \forall x = [x_1, \ldots, x_n] \in \mathcal{X}^n$]

$\boxed{\text{Lemma}}$ [Intersection lemma]

For $P(X) > 0$ if

$$X_A \perp\!\!\!\perp X_B \mid (X_C, X_D)$$

& $$X_A \perp\!\!\!\perp X_C \mid (X_B, X_D)$$

then

$$X_A \perp\!\!\!\perp (X_B, X_C) \mid X_D$$

Pf of Claim $\boxed{(G)\ X_i \perp\!\!\!\perp X_j \mid X_{V\setminus\{i,j\}} \text{ if } (i,j)\notin E \Rightarrow X_A \perp\!\!\!\perp X_C \mid X_B}$

Let's show by induction based on $|B|$

$= n-2, n-3 \ldots$

1

① $|B| = n-2$, $A = \{i\}$ $B = \{j\}$

then G, P are identical.

② Suppose (G) holds $\forall\ |B| \geq S$

will show G for B w. $|B| = s-1$

$< n-2$

Suppose $|A| \geq 2$ [WLOG], $|B| = S$

$A - B - C$ (G holds)

$\Rightarrow A\setminus\{i\} - B \cup \{i\} - C$

$\nearrow$

$\tilde{B}$, size $s+1$ so $X_{A\setminus\{i\}} \perp\!\!\!\perp X_C \mid X_{B\cup\{i\}}$ (a)

$\underset{\hat{A}}{\{i\}} - \underbrace{B \cup A\setminus\{i\}}_{|\tilde{B}| > S} - C$ also separated

(b) $X_C \perp\!\!\!\perp X_i \mid X_{B\cup A\setminus\{i\}}$

Intersection lemma + (a), (b)

$\Rightarrow X_C \perp\!\!\!\perp X_A \mid X_B$ $\forall A, B, C$ s.t. $A - B - C$

& $|B| = S$ ...

If $P(X) \not> 0$, then there's a counterexample (HW1).

Def [Factorization (F)]

We say $P(x)$ factorizes w.r.t. G if

$$P(x) = \frac{1}{Z} \prod_{c \in \mathcal{C}} f_c(x_c)$$

maximal clique set.

Claim $(F) \Rightarrow (G)$

Idea behind pf: $p(x) \propto \prod_{c \in \mathcal{C}} f_c(x_c)$ all cliques

$+A - B - C$

(No factor has $a \in A$ & $c \in C$)

$$\Rightarrow \quad p(x) \propto F_{AB}(x_A, x_B) \cdot F_{BC}(x_B, x_C)$$

In HW1 $\longrightarrow X_A \perp\!\!\!\perp X_C \mid X_B$

$(G) \Rightarrow F$ if $P(x) > 0$ [Hammersley Clifford Thm]

So, if $P(x) > 0$ $(F) \Leftrightarrow (G) \Leftrightarrow (L) \Leftrightarrow (P)$

if $p(x) \not> 0$ $G \not\Rightarrow F$