

① (a), (b), (c)

(only pairwise cliques)
((1-2), (2-3), ..., (N-1, N)).

$$p(\vec{x}) = \frac{1}{Z} \Psi_{1,2}(x_1, x_2) \dots \Psi_{N-1}(x_{N-1}, x_N)$$



(joint distribution obtained as a product of potentials)

$$p(x_n) = \sum_{x_1} \sum_{x_2} \dots \sum_{x_{n-1}} \sum_{x_{n+1}} \dots \sum_{x_N} p(\vec{x})$$

[ask what is the computational complexity]
(replace summations with integrations for Gaussians)

$$= \frac{1}{Z} \sum_{x_1} \sum_{x_2} \dots \sum_{x_{n-1}} \sum_{x_{n+1}} \dots \sum_{x_N} \Psi_{1,2}(x_1, x_2) \dots \Psi_{N-1}(x_{N-1}, x_N)$$

[now what is the computational complexity?]

$$= \left[\frac{1}{Z} \sum_{x_{n+1}} \Psi_{n+1}(x_{n+1}, x_n) \dots \sum_{x_2} \Psi_{2,3}(x_2, x_3) \sum_{x_1} \Psi_{1,2}(x_1, x_2) \right] x$$

$$\left[\sum_{x_{n+1}} \Psi_{n+1}(x_{n+1}, x_n) \sum_{x_{n+2}} \dots \sum_{x_{N-2}} \Psi_{N-2}(x_{N-2}, x_{N-1}) \sum_{x_{N-1}} \Psi_{N-1}(x_{N-1}, x_N) \right]$$

(fn. of x_n)

* Defining $\mu_\alpha(x_n) = \sum_{x_{n+1}} \Psi_{n+1}(x_{n+1}, x_n) \dots \sum_{x_1} \Psi_{1,2}(x_1, x_2)$

$$= \left[\sum_{x_{n+1}} \Psi_{n+1}(x_{n+1}, x_n) \mu_\alpha(x_{n+1}) \right]$$

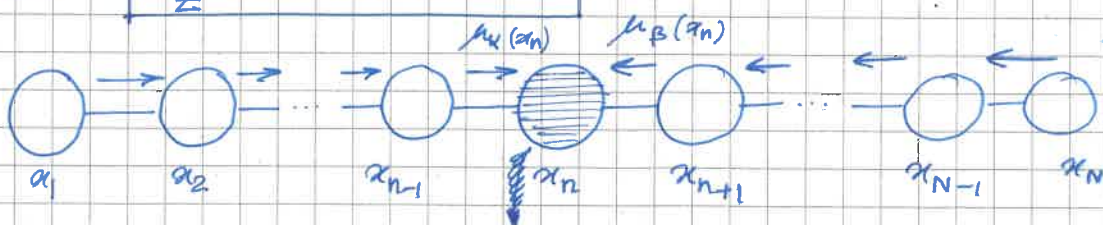
(forward message passing)

and $\mu_\beta(x_n) = \sum_{x_{n+1}} \Psi_n(x_n, x_{n+1}) \mu_\beta(x_{n+1})$

(reverse/backward message passing)

Combining these two definitions,

$$p(x_n) = \left[\frac{1}{Z} \mu_\alpha(x_n) \cdot \mu_\beta(x_n) \right]$$



$$(d) \quad p(x_{n-1}, x_n) = \sum_{x_1} \sum_{x_2} \dots \sum_{x_{n-2}} \sum_{x_{n+1}} \dots \sum_{x_N} p(\bar{x})$$

$$(d) \quad = \sum_{x_1} \sum_{x_2} \dots \sum_{x_{n-2}} \sum_{x_{n+1}} \dots \sum_{x_N} \frac{1}{Z} \Psi_{1,2}(x_1, x_2) \dots \Psi_{n-1,n}(x_{n-1}, x_n) \Psi_n(x_n, x_{n+1}) \dots \Psi_{N-1,N}(x_{N-1}, x_N)$$

$$= \underbrace{\left[\sum_{x_{n-2}} \Psi_{n-2}(x_{n-2}, x_{n-1}) \sum_{x_{n-3}} \Psi_{n-3}(x_{n-2}, x_{n-3}) \dots \sum_{x_1} \Psi_{1,2}(x_1, x_2) \right]}_{= \mu_\alpha(x_{n-1})} \cdot \frac{1}{Z} \times$$

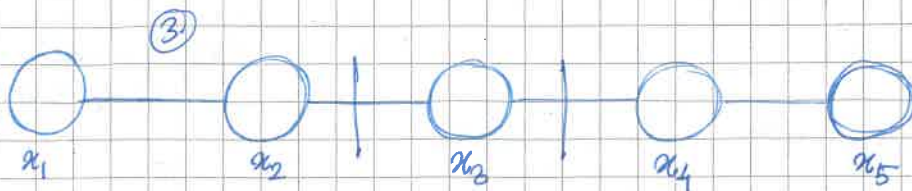
$$\Psi_{n-1,n}(x_{n-1}, x_n) \times$$

$$\underbrace{\left[\sum_{x_{n+1}} \Psi_n(x_n, x_{n+1}) \sum_{x_{n+2}} \Psi_{n+1}(x_{n+1}, x_{n+2}) \dots \sum_{x_N} \Psi_{N-1}(x_{N-1}, x_N) \right]}_{= \mu_\beta(x_n)}$$

$$= \boxed{\frac{1}{Z} \mu_\alpha(x_{n-1}) \Psi_{n-1,n}(x_{n-1}, x_n) \mu_\beta(x_n)}$$

~~***~~

8/7



$$p(x_2 | x_3, x_5)$$

$$= \frac{p(x_2, x_3, x_5)}{p(x_3, x_5)}$$

$$= \frac{p(x_2 | x_3) p(x_5 | x_3) p(x_3)}{p(x_3) p(x_5 | x_3)} = p(x_2 | x_3)$$

$$x_2 \perp\!\!\!\perp x_5 | x_3$$

(d-separation)

So to find $p(x_2 | x_3) = \frac{p(x_2, x_3)}{p(x_3)}$

$$p(x_3) = \frac{1}{Z} \mu_\alpha(x_3) \mu_\beta(x_3), \quad (\text{message passing})$$

$$p(x_3, x_2) = \frac{1}{Z} \mu_\alpha(x_2) \Psi_{2,3}(x_2, x_3) \mu_\beta(x_3) \quad (\text{message passing for two variables})$$

$$\therefore p(x_2 | x_3) = \boxed{\frac{\mu_\alpha(x_2) \Psi_{2,3}(x_2, x_3)}{\mu_\alpha(x_3)}}$$

$$p(x) = \prod_{s \in \text{ne}(x)} f_s(x_s)$$

$$= \prod_{s \in \text{ne}(x)} F_s(x, X_s)$$

Calculating the marginal $p(x)$ means:

$$* p(x) = \sum_{\bar{x} \setminus \{x\}} \prod_{s \in \text{ne}(x)} F_s(x, X_s)$$

$$= \prod_{s \in \text{ne}(x)} \sum_{\bar{x} \setminus \{x\}} F_s(x, X_s)$$

$$= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x) \quad (\text{product of all messages from fac. nodes connected to } x)$$

$$* \mu_{f_s \rightarrow x}(x) = \sum_{X_s} F_s(x, X_s)$$

[black + green blanket]

$$= \sum_{(x_1, x_2, \dots, x_n)} f_s(x, x_1, \dots, x_n) \cdot G_1(x_1, X_{s_1}) \dots G_n(x_n, X_{s_n})$$

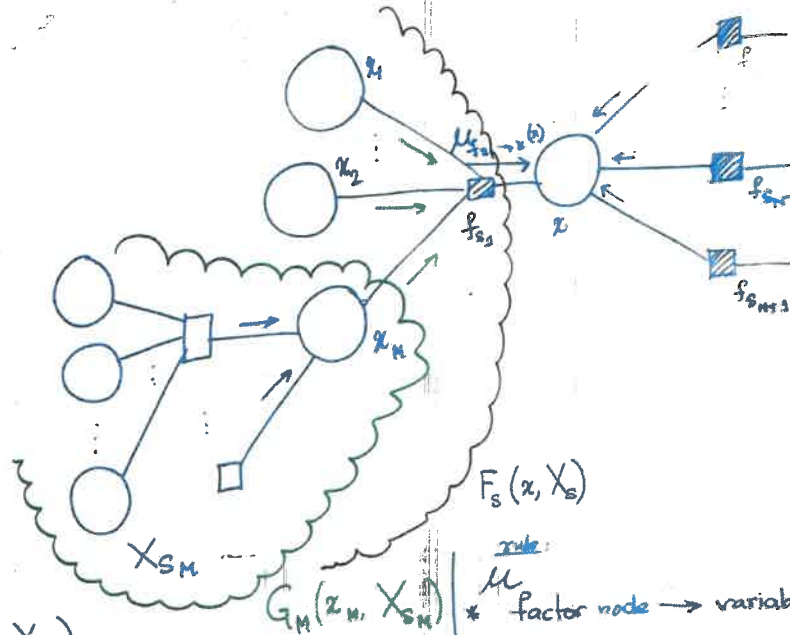
$$= \sum_{(x_1, \dots, x_n)} f_s(x, x_1, \dots, x_n) \prod_{m \in \text{ne}(f_s) \setminus \{x\}} \sum_{X_{s_m}} G_m(x_m, X_{s_m})$$

$$= \sum_{(x_1, \dots, x_n)} f_s(x, x_1, \dots, x_n) \prod_{m \in \text{ne}(f_s) \setminus \{x\}} \mu_{x_m \rightarrow f_s}(x_m)$$

$$* \mu_{x_m \rightarrow f_s}(x_m) = \sum_{X_{s_m}} G_m(x_m, X_{s_m})$$

$$= \sum_{X_{s_m}} \prod_{l \in \text{ne}(x_m) \setminus \{f_s\}} f_l(x_m, x_{s_l})$$

$$= \prod_{l \in \text{ne}(x_m) \setminus \{f_s\}} \mu_{x_l \rightarrow x_m}(x_m)$$



rule:

μ factor node \rightarrow variable node
 = weighted sum of incoming product of messages at factor node

μ variable node \rightarrow factor node
 = product of incoming messages at variable node

Initialize:

$$\mu_{\text{var} \rightarrow \text{fac}} = 1$$

$$\mu_{f_s \rightarrow \text{var}} = f_s(\text{var})$$