properties implied by $\mathcal{G}$, we may decompose the marginal distribution $p(x_s \mid y)$ as

$$p(x_s \mid y) = \frac{p(y \mid x_s)\, p(x_s)}{p(y)} = \alpha p(x_s)\, p(y_s \mid x_s) \prod_{t \in N(s)} p\left(y_{t \setminus s} \mid x_s\right) \tag{2.20}$$

where α denotes a normalization constant[7] which is independent of x. From this decomposition, we see that for the purposes of calculating $p(x_s \mid y)$, the conditional likelihood $p\left(y_{t\setminus s} \mid x_s\right)$ is a *sufficient statistic* of the data in the subtree rooted at node t. Using the conditional independencies implied by $\mathcal{G}$, we can derive a self–consistent set of equations relating the conditional likelihoods at neighboring nodes:

$$\begin{aligned} p\left(y_{t\setminus s} \mid x_s\right) &= \frac{p\left(x_s \mid y_{t\setminus s}\right) p\left(y_{t\setminus s}\right)}{p(x_s)} = \alpha \int_{x_t} \frac{p\left(x_s, x_t \mid y_{t\setminus s}\right)}{p(x_s)}\, dx_t \\ &= \alpha \int_{x_t} \frac{p(x_s, x_t)\, p\left(y_{t\setminus s} \mid x_s, x_t\right)}{p(x_s)}\, dx_t \\ &= \alpha \int_{x_t} \frac{p(x_s, x_t)\, p\left(y_{t\setminus s} \mid x_t\right)}{p(x_s)}\, dx_t \\ &= \alpha \int_{x_t} \left(\frac{p(x_s, x_t)}{p(x_s)\, p(x_t)}\right) p(x_t)\, p(y_t \mid x_t) \prod_{u \in N(t)\setminus s} p\left(y_{u\setminus t} \mid x_t\right)\, dx_t \end{aligned} \tag{2.21}$$

Note that the prior model appears in equation (2.21) solely in terms of the potential functions found in the canonical tree–based factorization (2.19).

It is possible to modify equations (2.20, 2.21) to use the potential terms appearing in an arbitrary pairwise factorization (2.14). In particular, if we "unwrap" the conditional likelihood terms in equation (2.20) by repeated application of the local consistency conditions (2.21), we obtain

$$\begin{aligned} p(x_u \mid y) &= \alpha \int_{x_{\mathcal{V}\setminus u}} \prod_{(s,t)\in\mathcal{E}} \frac{p(x_s, x_t)}{p(x_s)\, p(x_t)} \prod_{s\in\mathcal{V}} p(x_s)\, p(y_s \mid x_s)\, dx_{\mathcal{V}\setminus u} \\ &= \alpha \int_{x_{\mathcal{V}\setminus u}} \prod_{(s,t)\in\mathcal{E}} \psi_{s,t}(x_s, x_t) \prod_{s\in\mathcal{V}} p(y_s \mid x_s)\, dx_{\mathcal{V}\setminus u} \end{aligned} \tag{2.22}$$

where the second equality follows because equations (2.19) and (2.14) provide two

[7]By convention, the normalization constant α is always chosen so that the function it multiplies integrates to unity. Thus, the specific numerical value of α may change from equation to equation.