some of the graph's structure may be lost in the process.

For Gaussian Markov random fields, the prior distribution $p(x)$ is uniquely specified by either the full covariance matrix $P$ or the inverse covariance matrix $J = P^{-1}$. However, because of the sparse structure implied by Theorem 2.2, $J$ provides the more natural and efficient parameterization. Often, it is convenient to decompose $J$ into (pairwise) clique potentials as in equation (2.14). To each edge $(s,t) \in \mathcal{E}$, we associate a clique potential

$$\psi_{s,t}(x_s, x_t) = \exp\left\{ -\frac{1}{2} \begin{bmatrix} x_s^T & x_t^T \end{bmatrix} \begin{bmatrix} J_{s(t)} & J_{s,t} \\ J_{t,s} & J_{t(s)} \end{bmatrix} \begin{bmatrix} x_s \\ x_t \end{bmatrix} \right\} \tag{2.15}$$

where the $J_{s(t)}$ terms are chosen so that for all $s \in \mathcal{V}$,

$$\sum_{t \in N(s)} J_{s(t)} = J_{s,s} \tag{2.16}$$

Straightforward algebraic manipulation shows that any set of clique potentials (2.15) which satisfy the consistency condition (2.16) will define a probability distribution $p(x)$ such that

$$p(x) = \frac{1}{Z} \exp\left\{ -\frac{1}{2} x^T P^{-1} x \right\} = \frac{1}{Z} \prod_{s=1}^{N} \prod_{t=1}^{N} \exp\left\{ -\frac{1}{2} x_s^T J_{s,t} x_t \right\} =$$
$$\frac{1}{Z} \prod_{(s,t) \in \mathcal{E}} \exp\left\{ -\frac{1}{2} \begin{bmatrix} x_s^T & x_t^T \end{bmatrix} \begin{bmatrix} J_{s(t)} & J_{s,t} \\ J_{t,s} & J_{t(s)} \end{bmatrix} \begin{bmatrix} x_s \\ x_t \end{bmatrix} \right\} = \frac{1}{Z} \prod_{(s,t) \in \mathcal{E}} \psi_{s,t}(x_s, x_t) \tag{2.17}$$

where $Z = \left( (2\pi)^N \det P \right)^{1/2}$ is a normalization constant.

The preceding construction shows that it is always possible to represent a Gaussian Markov random field using pairwise clique potentials *without* having to cluster nodes. Note that this construction does not guarantee that the local potential matrices $\left[\begin{smallmatrix} J_{s(t)} & J_{s,t} \\ J_{t,s} & J_{t(s)} \end{smallmatrix}\right]$ will be positive definite. In such cases, $\psi_{s,t}(x_s, x_t)$ is not a bounded function of $x$, and therefore cannot be normalized so that it defines a locally consis-