for jointly Gaussian MRFs, all conditional distributions are also Gaussian, and the integrals may be replaced by recursions on the mean and covariance parameters of these Gaussians. Each message update at node s requires $\mathcal{O}(d_s^3)$ operations. The following section explores these updates in detail.

2.3.2 Gaussian Belief Propagation

In this section, we specialize the BP update equations (2.25, 2.26) to the Gaussian graphical models introduced in previous sections. Recall that, for the purpose of calculating $p\left(x_s \mid y\right)$, equation (2.20) shows that the conditional likelihoods $\{p\left(y_{t\backslash s} \mid x_s\right)\}_{t\in N(s)}$ provide a set of sufficient statistics, where

$$p\left(y_{t\backslash s} \mid x_s\right) = \alpha \frac{p\left(x_s \mid y_{t\backslash s}\right)}{p\left(x_s\right)} \tag{2.27}$$

Because x and y are jointly Gaussian, $p\left(x_s\right)$ and $p\left(x_s \mid y_{t\backslash s}\right)$ are also Gaussian, and therefore have a finite–dimensional parameterization in terms of their mean and covariance. When viewed as a function of x_s for a fixed set of observations y, the likelihood $p\left(y_{t\backslash s} \mid x_s\right)$ does *not* correspond to a probability distribution. However, equation (2.27) shows that $p\left(y_{t\backslash s} \mid x_s\right)$ is still proportional to an exponentiated quadratic form in x_s. Since the proportionality constant is independent of x_s, we may parameterize $p\left(y_{t\backslash s} \mid x_s\right)$ with a "mean" vector and "covariance" matrix, just as we would a Gaussian probability distribution. Furthermore, using the structure of Gaussian clique potentials (2.15), it is straightforward to show that the BP messages $m_{ts}\left(x_s\right)$ may also be parameterized by a mean and covariance. This parameterization is important because it allows the integral BP update equation (2.25) to be transformed into a set of linear algebraic equations.

It is possible to develop a set of Gaussian belief propagation equations which act directly on the mean and covariance parameters of the messages $m_{ts}\left(x_s\right)$. However, just as §2.2.3 showed that Gaussian prior models are most naturally parameterized by their sparse *inverse* covariance matrix $J = P^{-1}$, Gaussian BP is simplest when the messages are represented in an *information* form [44]. For a Gaussian distribution