where

\[
p_j = \sum_i p_{ij}
\]

\[
p_{ji} = \frac{p_{ij}}{\sum_j p_{ij}}
\]

\[
\mu_j = \sum_i \mu_{ij} p_{ji}
\]

\[
\Sigma_j = \sum_i \Sigma_{ij} p_{ji} + \sum_i (\mu_{ij} - \mu_j)(\mu_{ij} - \mu_j)^T p_{ji}
\]

In the junction tree literature, this is called “weak marginalization”. It can be applied to any conditionally Gaussian model, not just switching SSMs.

The GPB2 algorithm requires running \(M^2\) Kalman filters at each step. A cheaper alternative, known as interacting multiple models (IMM), can be obtained by first collapsing the prior to a single Gaussian (by moment matching), and then updating it using \(M\) different Kalman filters, one per value of \(S_t\); see Figure 40. Unfortunately, it is hard to extend IMM to the smoothing case, unlike GPB2, a smoothing version of which is discussed in Section 6.1.3.

### 5.3.2 Viterbi approximation

If there are a large number of discrete variables, it may be too slow to perform \(M^2\) or even \(M\) KF updates, as required by GPB2 and IMM. Instead, one can enumerate the discrete values in a priori order of probability. (Computing their posterior probability is as expensive as an exact update step.) This makes sense for a DBN for fault diagnosis, where there is a natural ordering on the discrete values: one fault is much more likely than two faults, etc. However, it would not be applicable to the data association DBN in Figure 23, where there is no such ordering.