The constant $\eta$ plays no role since the resampling takes place with probabilities proportional to the importance weights. By resampling particles with probability proportional to $w^\text{[m]}_t$, the resulting particles are indeed distributed according to the product of the proposal and the importance weights $w^\text{[m]}_t$:

$$\eta w^\text{[m]}_t p(x_t | x_{t-1}, u_t) p(x_{0:t-1} | z_{1:t-1}, u_{1:t-1}) = bel(x_{0:t})$$

(Notice that the constant factor $\eta$ here differs from the one in (4.34).) The algorithm in Table 4.4 follows now from the simple observation that if $x^\text{[m]}_0$ is distributed according to $bel(x_{0:t})$, then the state sample $x^\text{[m]}_t$ is (trivially) distributed according to $bel(x_t)$.

As we will argue below, this derivation is only correct for $M \uparrow \infty$, due to a laxness in our consideration of the normalization constants. However, even for finite $M$ it explains the intuition behind the particle filter.

4.3.4 Practical Considerations and Properties of Particle Filters

Density Extraction

The sample sets maintained by particle filters represent discrete approximations of continuous beliefs. Many applications, however, require the availability of continuous estimates, that is, estimates not only at the states represented by particles, but at any point in the state space. The problem of extracting a continuous density from such samples is called density estimation.

Figure 4.5 illustrates different ways of extracting a density from particles. The leftmost graph shows the particles and density of the transformed Gaussian from our standard example (c.f. Figure 4.3). A simple and highly efficient approach to extracting a density from such particles is to compute a Gaussian approximation, as illustrated by the dashed Gaussian in Figure 4.5(b). In this case, the Gaussian extracted from the particles is virtually identical to the Gaussian approximation of the true density (solid line).

Obviously, a Gaussian approximation captures only basic properties of a density, and it is only appropriate if the density is unimodal. Multimodal sample distributions require more complex techniques such as $k$-means clus-