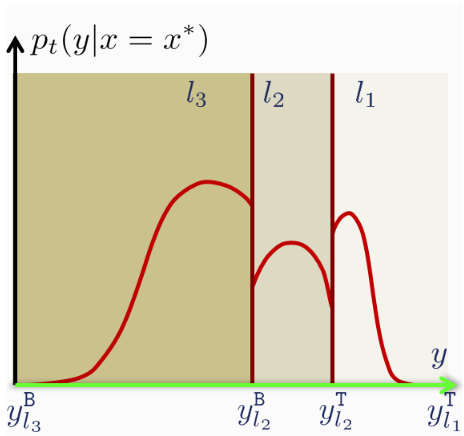

Fig. 5.13: **The tree conditional is a piece-wise Gaussian.** See text for details.

with ϕ denoting the 1D cumulative normal distribution function

$$\phi_{t,l}(y|x = x^*) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{y - \mu_{y|x,l}}{\sqrt{2\sigma^2_{y|x,l}}}\right)\right].$$

Finally, the forest conditional is:

$$p(y|x = x^*) = \frac{1}{T}\sum_{t=1}^{T} p_t(y|x = x^*)$$

Figure 5.14 shows the forest conditional distribution computed for five fixed values of x. When comparing *e.g.* the conditional $p(y|x = x_3)$ in fig. 5.14 with the distribution in 4.10b we see that now the conditional shows three very distinct modes rather than a large, uniformly uninformative mass. Although some ambiguity remains (it is inherent in the training set) now we have a more precise description of such ambiguity.

5.6.2 Sampling from conditional densities

We conclude this chapter by discussing the issue of efficiently drawing random samples from the conditional model $p(y|x)$.