



Fig. 5.15: **Sampling from conditional model.** Since x is known and y unknown y -nodes cannot apply the associated split function. When sampling from such a tree a child of a y -node is chosen randomly. Instead, the child of an x -node is selected deterministically. See text for details.

the forest size T , the amount of training randomness ρ etc.

5.7 Quantitative analysis

This section assesses the accuracy of the density estimation algorithm with respect to ground-truth. Figure 5.17a shows a ground-truth probability density function. The density is represented non-parametrically as a normalized histogram defined over the 2D (x_1, x_2) domain.

Given the ground-truth density we randomly sample 5,000 points numerically (fig. 5.17b), via the multivariate inverse probability integral transform algorithm [26]. The goal now is as follows: Given the sampled points only, reconstruct a probability density function which is as close as possible to the ground-truth density.

Thus, a density forest is trained using the sampled points alone. No use is made of the ground-truth density in this stage. Given the trained forest we test it on all points in a predefined domain (not just on the training points, fig. 5.17c). Finally, a quantitative comparison between