



Fig. 3.9: **The effect of randomness on the forest margin.** (a) Forest posterior for $\rho = 50$ (small randomness). (b) Forest posterior for $\rho = 5$. (c) Forest posterior for $\rho = 2$ (highest randomness). These experiments have used $D = 2, T = 400$ and axis-aligned weak learners. The bottom row shows 1D posteriors computed along the white dashed line. Increasing randomness produces less well defined separating surfaces. The optimal separating surface, *i.e.* the loci of points where the class posteriors are equal (shown in black) moves towards the left of the margin-maximizing line (shown in green in all three experiments). As randomness increases individual training points have less influence on the separating surface.

3.4.3 Max-margin in multiple classes

Since classification forests can naturally apply to more than 2 classes how does this affect their maximum-margin properties? We illustrate this point with a multi-class synthetic example. In fig. 3.11a we have a linearly separable four-class training set. On it we have trained two