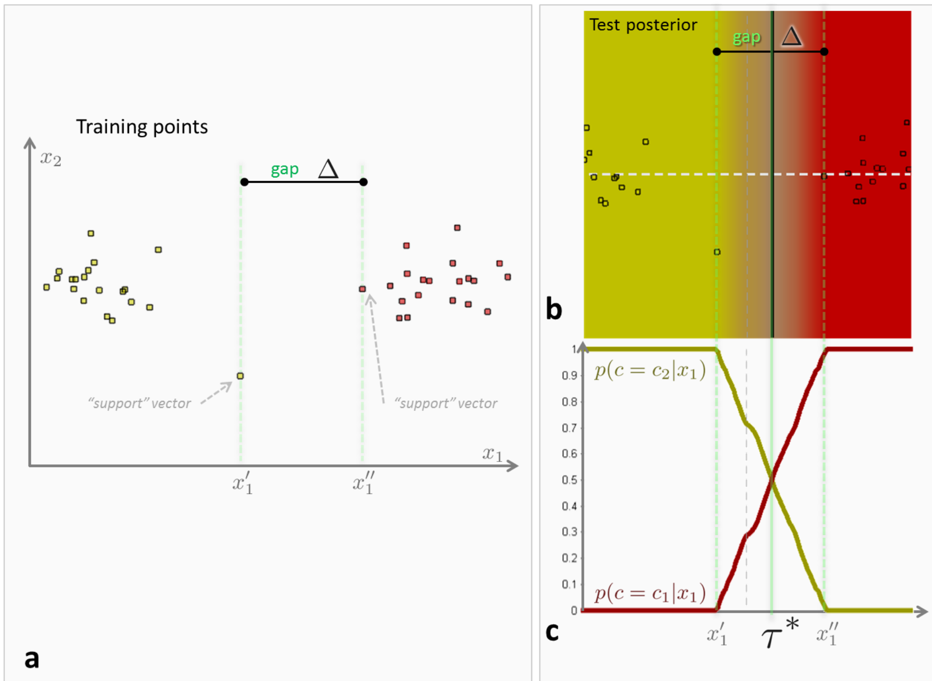

Fig. 3.8: **Forest's maximum-margin properties. (a)** Input 2-class training points. They are separated by a gap of dimension Δ. **(b)** Forest posterior. Note that all of the uncertainty band resides within the gap. **(c)** Cross-sections of class posteriors along the horizontal, white dashed line in (b). Within the gap the class posteriors are linear functions of x_1. Since they have to sum to 1 they meet right in the middle of the gap. In these experiments we use $\rho = 500, D = 2, T = 500$ and axis aligned weak learners.

dle of the gap, *i.e.* the maximum-margin solution. Next, we describe the same concepts more formally.

We are given the two-class training points in fig. 3.8a. In this simple example the training data is not only linearly separable, but it is perfectly separable via vertical stumps on x_1. So we constrain our weak learners to be vertical lines only, *i.e.*

$$h(\mathbf{v}, \boldsymbol{\theta}_j) = [\phi(\mathbf{v}) > \tau] \quad \text{with} \quad \phi(\mathbf{v}) = x_1.$$