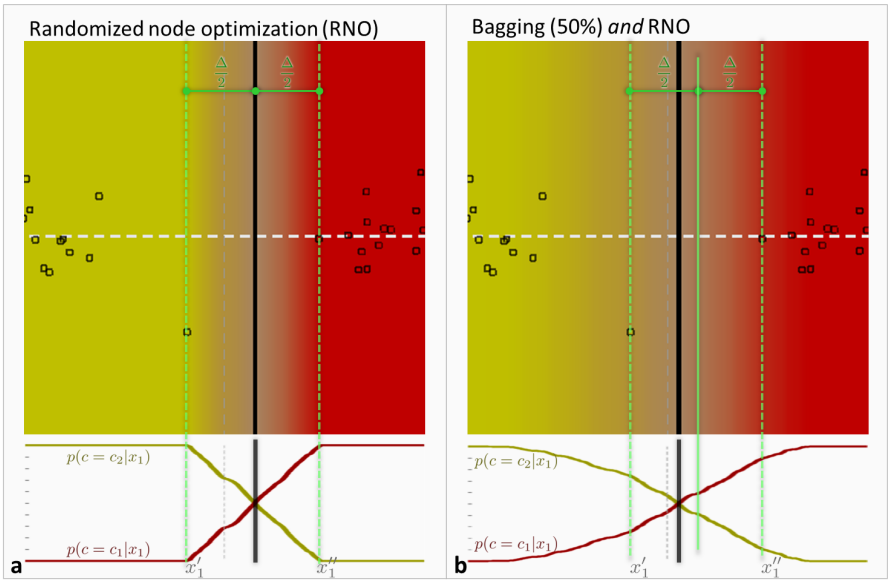

Fig. 3.12: **Max-margin: bagging v randomized node optimization. (a)** Posterior for forest trained with randomized node optimization. **(b)** Posterior for forest trained with bagging. In bagging, for each tree we use 50% random selection of training data with replacement. Loci of optimal separation are shown as black lines. In these experiments we use $\rho = 500, D = 2, T = 400$ and axis-aligned weak learners. Areas of high entropy have been shown strongly grey to highlight the separating surfaces.

We observe that the introduction of training set randomization leads to smoother posteriors whose optimal boundary (shown as a vertical black line) does *not* coincide with the maximum margin (green, solid line). Of course this behaviour is controlled by how much (training set) randomness we inject in the system. If we were to take all training data then we would reproduce a max-margin behaviour (but it would not be bagging). One advantage of bagging is increased training speed (due to reduced training set size). More experiments and comparisons are available in [1]. In the rest of the paper we use the RNO randomness model because it allows us to use all available training data and en-