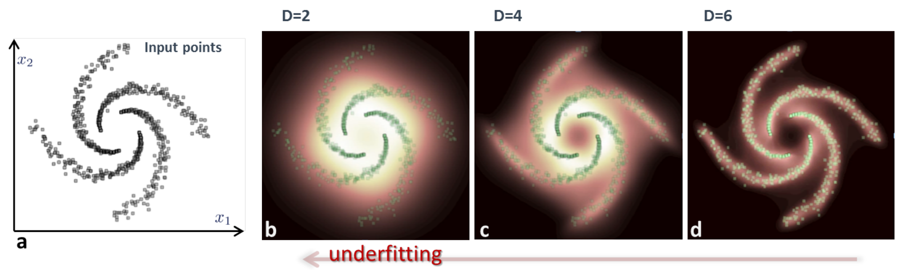

Fig. 5.6: **Density forest applied to a spiral data distribution. (a)** Input unlabelled data points in their 2D feature space. **(b,c,d)** Forest densities for different tree depths D. The original training points are overlaid in green. The complex distribution of input data is captured correctly by a deeper forest, *e.g.* $D = 6$, while shallower trees produce under-fitted, overly smooth densities.

trees yield under-fitting, *i.e.* overly smooth and detail-lacking density estimates. In this example good results are obtained for $D = 6$ as the density nicely captures the individuality of the four spiral arms while avoiding fitting to high frequency noise. Just like in classification and regression here too the parameter D can be used to set a compromise between smoothness of output and the ability to capture structural details.

So far we have described the density forest model and studied some of its properties on synthetic examples. Next we study density forests in comparison to alternative algorithms.

5.4 Comparison with alternative algorithms

This section discusses advantages and disadvantages of density forests as compared to the most common parametric and non-parametric density estimation techniques.

5.4.1 Comparison with non-parametric estimators

Figure 5.7 shows a comparison between forest density, Parzen window estimation and k-nearest neighbour density estimation. The compari-