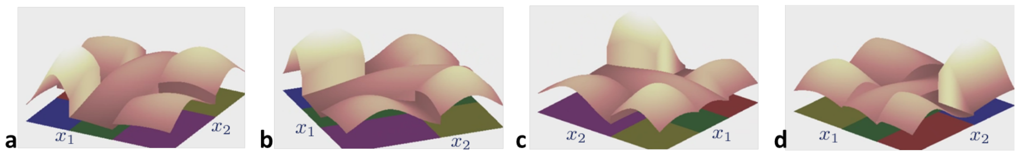

Fig. 5.2: **A tree density is piece-wise Gaussian. (a,b,c,d)** Different views of a tree density $p_t(\mathbf{v})$ defined over an illustrative 2D feature space. Each individual Gaussian component is defined over a bounded domain. See text for details.

leaf (see fig. 5.2). Thus, in order to ensure probabilistic normalization we need to incorporate the partition function Z_t, which is defined as follows:

$$Z_t = \int_{\mathbf{v}} \left(\sum_l \pi_l \, \mathcal{N}(\mathbf{v}; \boldsymbol{\mu}_l, \Lambda_l) \, p(l|\mathbf{v}) \right) \, d\mathbf{v}. \tag{5.4}$$

However, in a density forest each data point reaches exactly *one* terminal node. Thus, the conditional $p(l|\mathbf{v})$ is a delta function $p(l|\mathbf{v}) = [\mathbf{v} \in l(\mathbf{v})]$ and consequently (5.4) becomes

$$Z_t = \int_{\mathbf{v}} \pi_{l(\mathbf{v})} \, \mathcal{N}(\mathbf{v}; \boldsymbol{\mu}_{l(\mathbf{v})}, \Lambda_{l(\mathbf{v})}) \, d\mathbf{v}. \tag{5.5}$$

As it is often the case when dealing with generative models, computing Z_t in high dimensions may be challenging.

In the case of axis-aligned weak learners it is possible to compute the partition function via the cumulative multivariate normal distribution function. In fact, the partition function Z_t is the sum of all the volumes subtended by each Gaussian cropped by its associated partition cell (cuboidal in shape, see fig. 5.2). Unfortunately, the cumulative multivariate normal does not have a close form solution. However, approximating its functional form has is a well researched problem and a number of good numerical approximations exist [39, 71].

For more complex weak-learners it may be easier to approximate Z_t by numerical integration, *i.e.*

$$Z_t \approx \Delta \cdot \sum_i \pi_{l(\mathbf{v}_i)} \, \mathcal{N}(\mathbf{v}_i; \boldsymbol{\mu}_{l(\mathbf{v}_i)}, \Lambda_{l(\mathbf{v}_i)}),$$