Fig. 2.8: **Ensemble model. (a)** The posteriors of four different trees (shown with different colours). Some correspond to higher confidence than others. **(b)** An ensemble posterior $p(y|\mathbf{v})$ obtained by averaging all tree posteriors. **(c)** The ensemble posterior $p(y|\mathbf{v})$ obtained as product of all tree posteriors. Both in (b) and (c) the ensemble output is influenced more by the more informative trees.

can impose a minimum information gain. Tree growing may also be stopped when a node contains less that a defined number of training points. Avoiding growing full trees has repeatedly been demonstrated to have positive effects in terms of generalization. In this work we avoid further post-hoc operations such as tree pruning [42] to keep the training process as simple as possible.

2.2.7   Summary of key model parameters

In summary, the parameters that most influence the behaviour of a decision forest are:

- The forest size T;
- The maximum allowed tree depth D;
- The amount of randomness (controlled by ρ) and its type;
- The choice of weak learner model;
- The training objective function;
- The choice of features in practical applications.

Those choices directly affect the forest predictive accuracy, the *accuracy of its confidence*, its generalization and its computational efficiency.