

Boosting as Maximum Entropy

- corresponding **optimization problem**:

$$\min_{D \in \mathcal{P}} \text{RE}(D \parallel \text{uniform}) \leftrightarrow \max_{D \in \mathcal{P}} \text{entropy}(D)$$

- where

$$\begin{aligned} \mathcal{P} &= \text{feasible set} \\ &= \left\{ D : \sum_i D(i) y_i g_j(x_i) = 0 \quad \forall j \right\} \end{aligned}$$

- $\mathcal{P} \neq \emptyset \Leftrightarrow$ weak learning assumption does **not** hold
 - in this case, $D_t \rightarrow$ (unique) solution
- if weak learning assumption **does** hold then
 - $\mathcal{P} = \emptyset$
 - D_t can **never** converge
 - dynamics are fascinating but unclear in this case