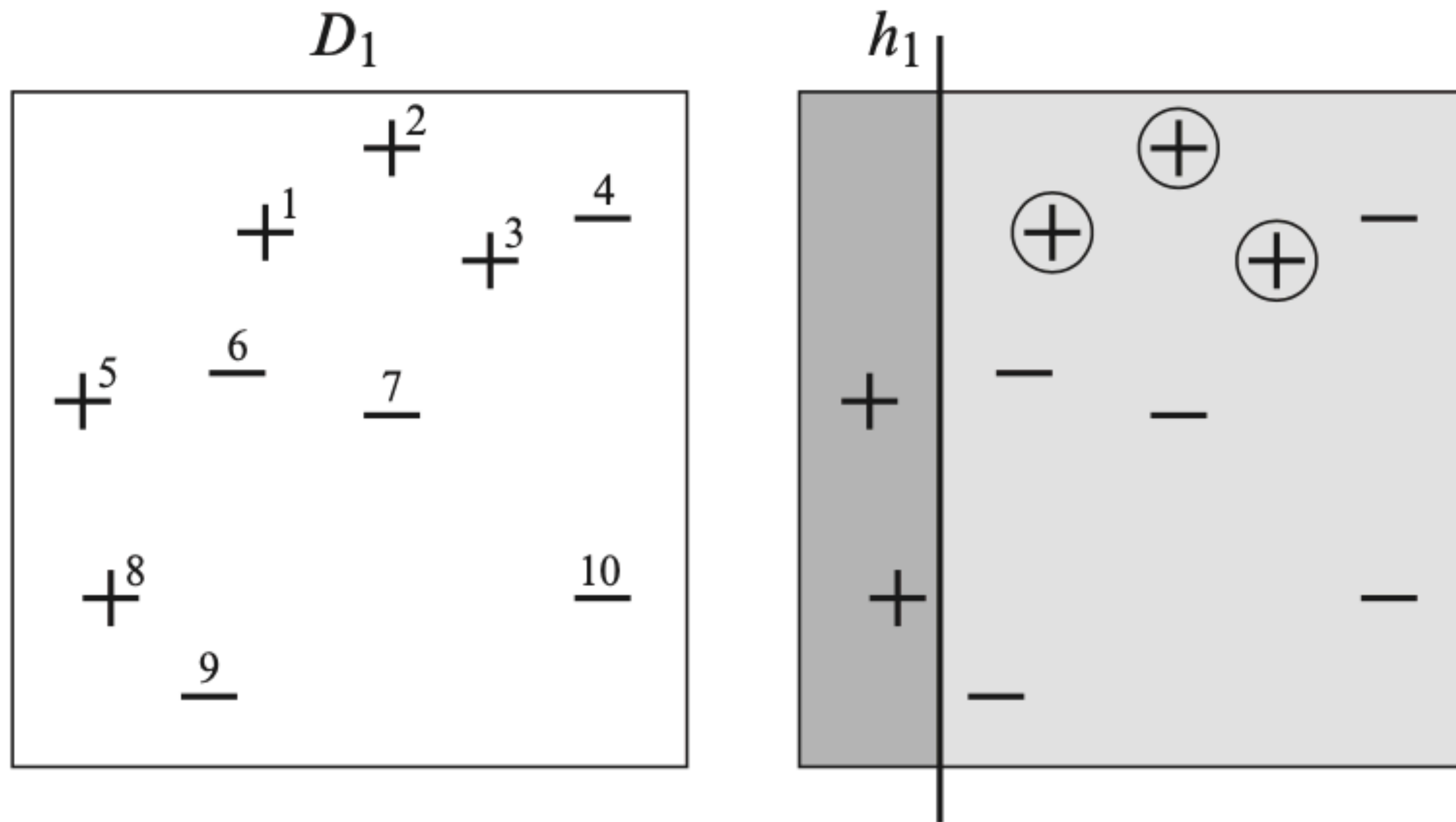# Boosting: a toy example
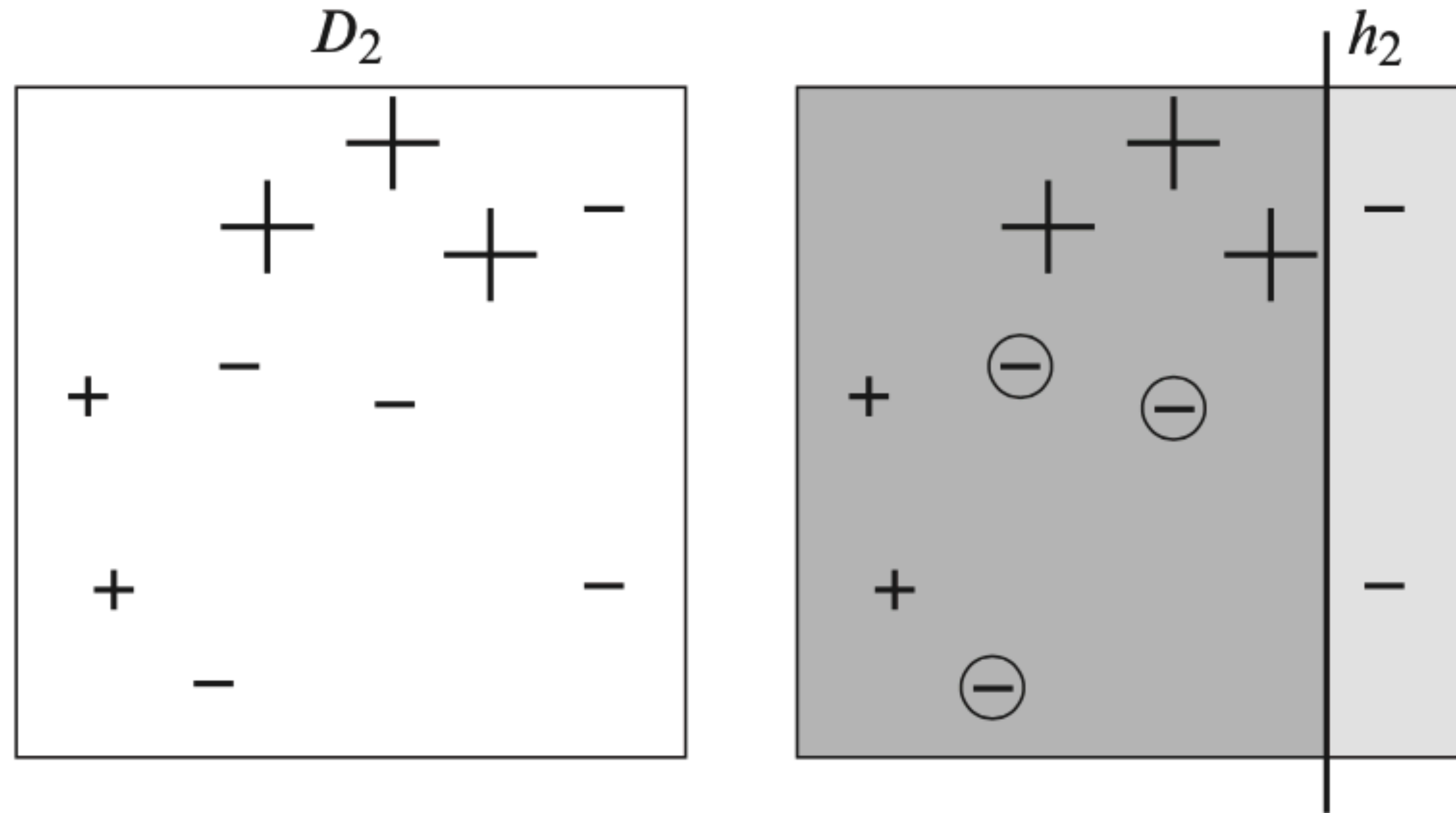
Boosting: a toy example (cont.)

Boosting: a toy example (cont.)

Source: Schapire and Freund, Boosting: Foundations and Algorithms, Section 1.2
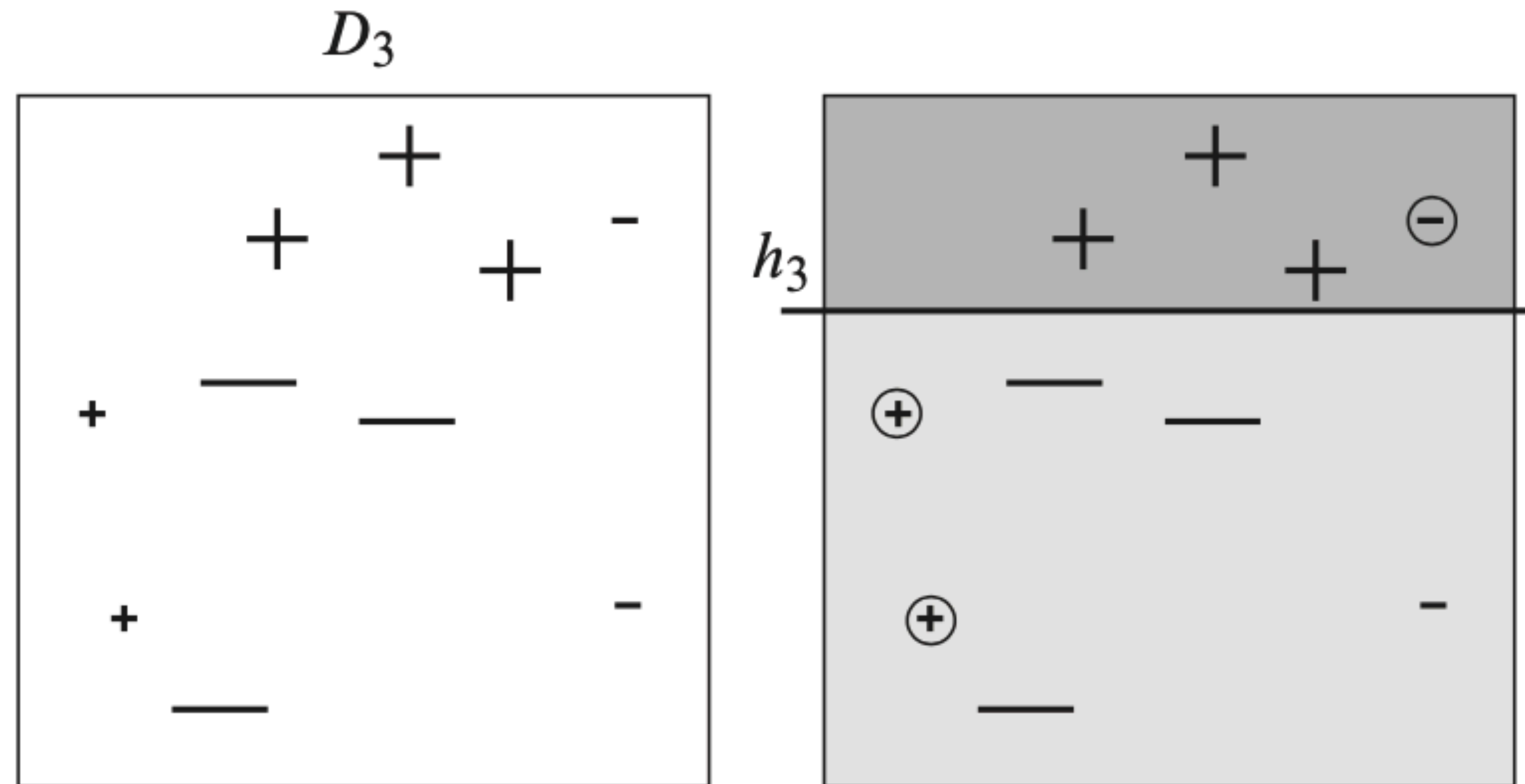
Boosting: a toy example (cont.)
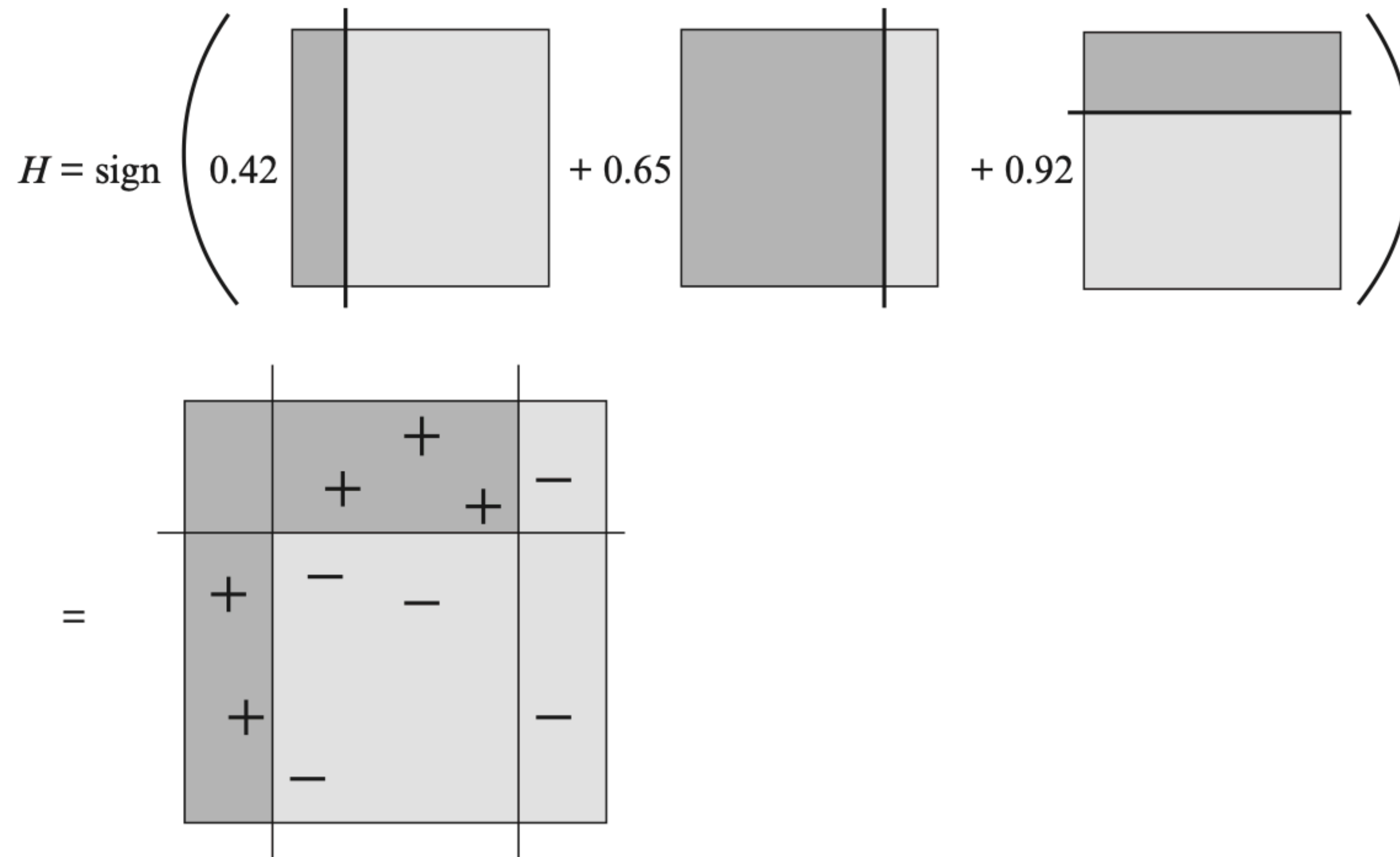
AdaBoost

Given: $(x_1, y_1), \ldots, (x_m, y_m)$ where $x_i \in \mathcal{X}$, $y_i \in \{-1, +1\}$.

Initialize: $D_1(i) = 1/m$ for $i = 1, \ldots, m$.

For $t = 1, \ldots, T$:

- Train weak learner using distribution D_t.

- Get weak hypothesis $h_t : \mathcal{X} \to \{-1, +1\}$.

- Aim: select h_t to minimalize the weighted error:

$$\epsilon_t \doteq \mathbf{Pr}_{i \sim D_t}[h_t(x_i) \neq y_i].$$

- Choose $\alpha_t = \dfrac{1}{2} \ln \left(\dfrac{1 - \epsilon_t}{\epsilon_t} \right).$

- Update, for $i = 1, \ldots, m$:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases}$$

$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t},$$

where Z_t is a normalization factor (chosen so that D_{t+1} will be a distribution).

Output the final hypothesis:

$$H(x) = \text{sign} \left(\sum_{t=1}^{T} \alpha_t h_t(x) \right).$$