

- The minimization is taken over the set of all transition matrices  $p(\hat{x}|x)$  such that  $Ed(X, \hat{X}) \leq D$ , namely the set

$$\left\{ p(\hat{x}|x) : \sum_{x, \hat{x}} p(x) p(\hat{x}|x) d(x, \hat{x}) \leq D \right\}.$$

- Since this set is compact (closed and bounded) in  $\Re^{|\mathcal{X}||\hat{\mathcal{X}}|}$  and  $I(X; \hat{X})$  is a continuous functional of  $p(\hat{x}|x)$ , the minimum value of  $I(X; \hat{X})$  can be attained.
- Alternatively, the minimization can be taken over the set of all joint distributions  $p(x, \hat{x})$  with marginal distribution  $p(x)$ , the given source distribution, such that  $Ed(X, \hat{X}) \leq D$ .