

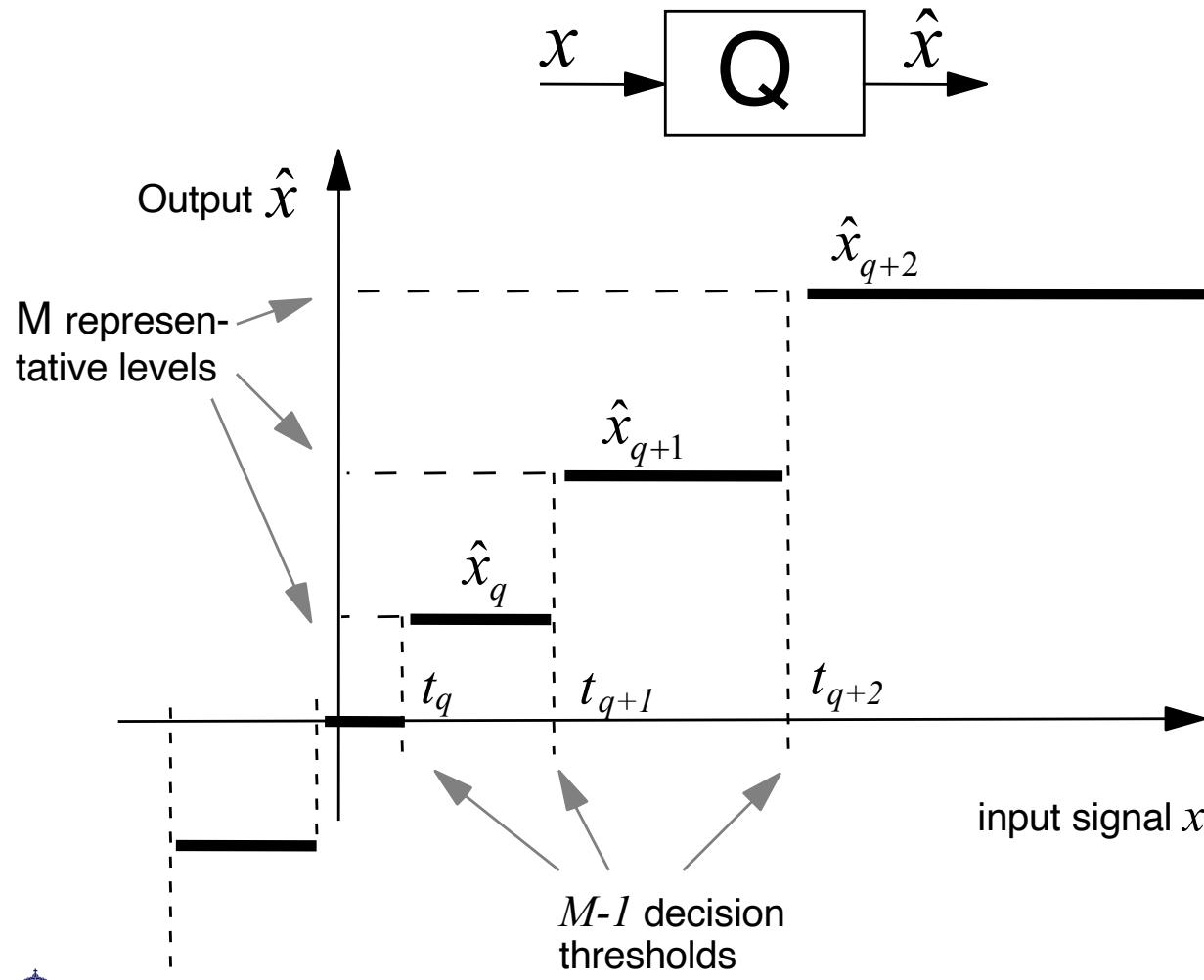
# Quantization

- Scalar quantizer
- Lloyd-Max scalar quantizer
  - Quantizer design algorithm
  - High rate approximation
- Entropy-constrained scalar quantizer
  - Quantizer design algorithm
  - High rate approximation
- Vector quantization
- Lattice vector quantization

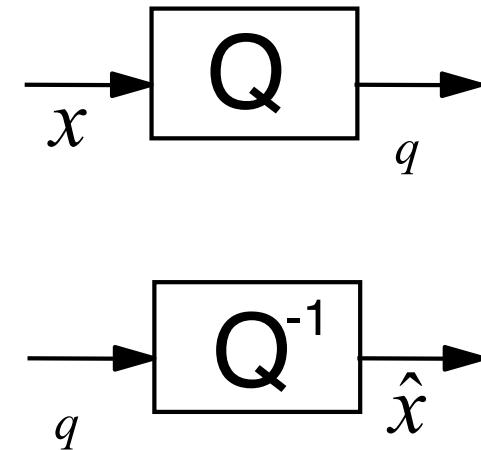


# Quantization

Input-output characteristic of a scalar quantizer



Sometimes, this convention is used:



# Lloyd-Max Scalar Quantizer

- Problem: For a signal  $x$  with given PDF  $f_X(x)$  find a quantizer with  $M$  representative levels such that

$$d = MSE = E \left[ (X - \hat{X})^2 \right] \rightarrow \min.$$

- Solution: Lloyd-Max quantizer

[Lloyd, 1957][Max, 1960]

- $M-1$  decision thresholds exactly half-way between representative levels.
- $M$  representative levels in the centroid of the PDF between two successive decision thresholds.
- Necessary condition

$$t_q = \frac{1}{2} (\hat{x}_{q-1} + \hat{x}_q) \quad q = 1, 2, \dots, M-1$$
$$\hat{x}_q = \frac{\int_{t_q}^{t_{q+1}} x f_X(x) dx}{\int_{t_q}^{t_{q+1}} f_X(x) dx} \quad q = 0, 1, \dots, M-1$$

# Iterative Lloyd-Max Quantizer Design

1. Guess initial set of representative levels  $\hat{x}_q \quad q = 0, 1, 2, \dots, M - 1$
2. Calculate decision thresholds

$$t_q = \frac{1}{2}(\hat{x}_{q-1} + \hat{x}_q) \quad q = 1, 2, \dots, M - 1$$

3. Calculate new representative levels

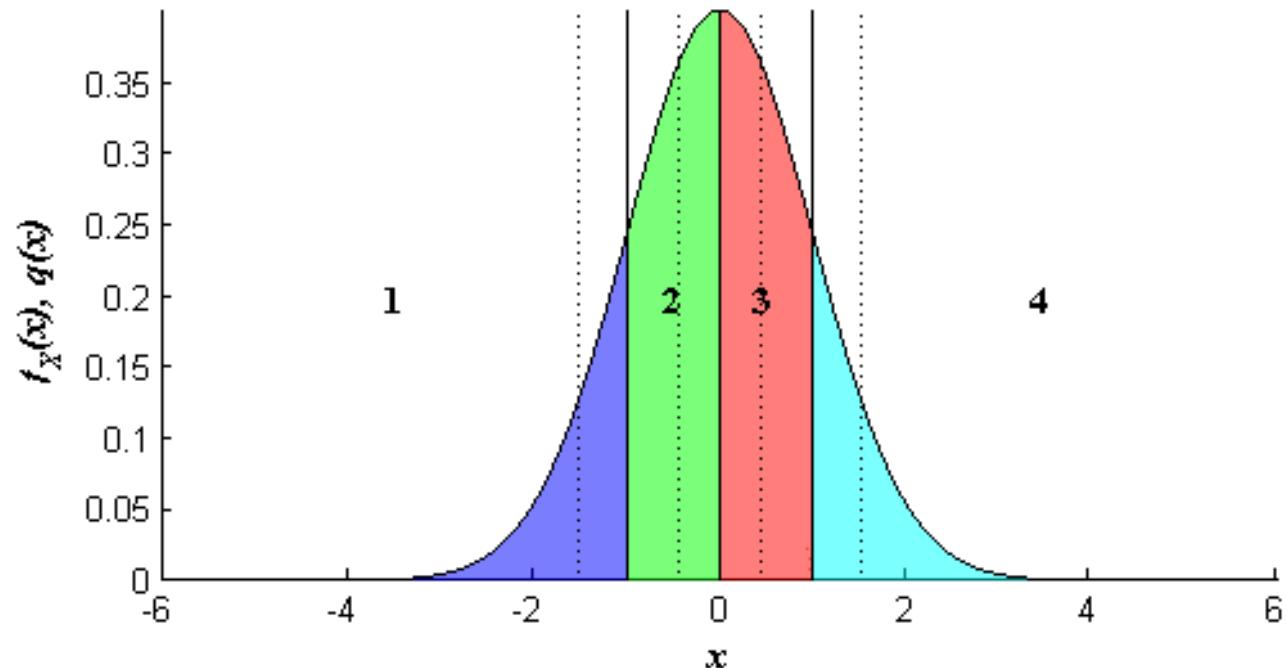
$$\hat{x}_q = \frac{\int_{t_q}^{t_{q+1}} x \cdot f_X(x) dx}{\int_{t_q}^{t_{q+1}} f_X(x) dx} \quad q = 0, 1, \dots, M - 1$$

4. Repeat **2.** and **3.** until no further distortion reduction



# Example: Lloyd-Max Design I

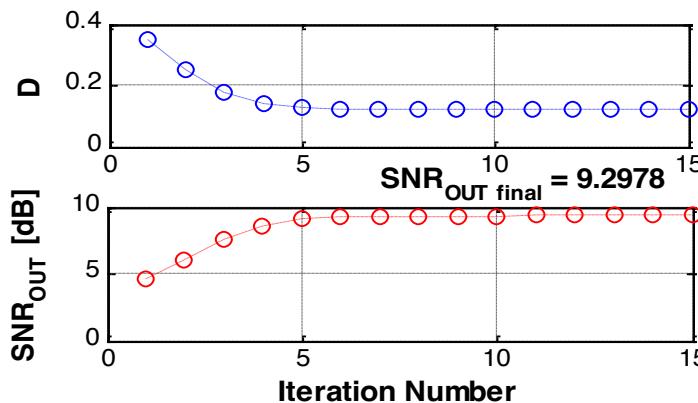
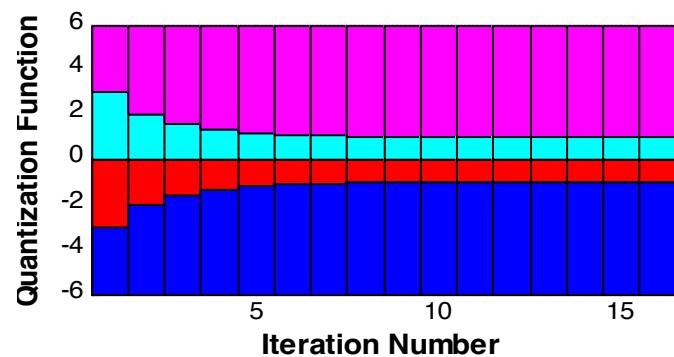
- $X$  zero-mean, unit-variance Gaussian r.v.
- Design scalar quantizer with 4 quantization indices with minimum expected distortion  $D^*$
- Optimum quantizer, obtained with the Lloyd algorithm
  - Decision thresholds  $-0.98, 0, 0.98$
  - Representative levels  $-1.51, -0.45, 0.45, 1.51$
  - $D^* = 0.12$
  - $D^* = 9.30$  dB



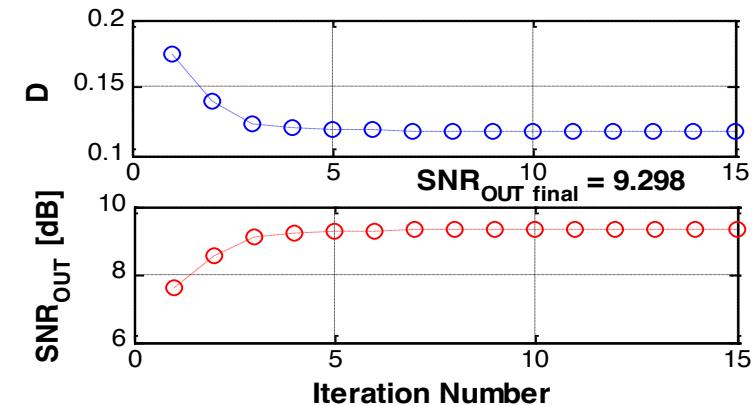
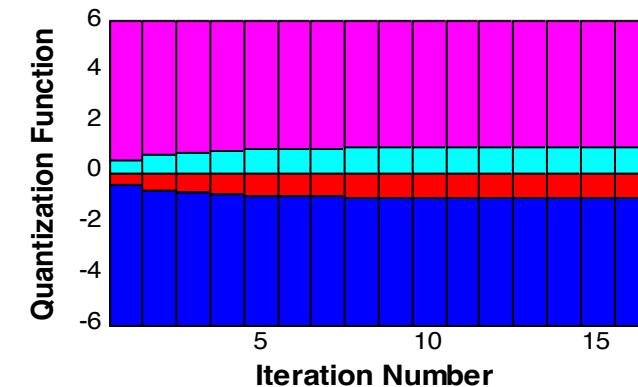
# Example: Lloyd-Max Design II

- Convergence

- Initial quantizer A:  
decision thresholds -3, 0, 3



- Initial quantizer B:  
decision thresholds -0.5, 0, 0.5



- After 6 iterations, in both cases  $(D - D^*)/D^* < 1\%$

# Lloyd Algorithm with Training Data

1. Guess initial set of representative levels  $\hat{x}_q$ ;  $q = 0, 1, 2, \dots, M - 1$
2. Assign each sample  $x_i$  in training set  $\mathcal{T}$  to closest representative  $\hat{x}_q$

$$B_q = \{x \in \mathcal{T} : Q(x) = q\} \quad q = 0, 1, 2, \dots, M - 1$$

3. Calculate new representative levels

$$\hat{x}_q = \frac{1}{\|B_q\|} \sum_{x \in B_q} x \quad q = 0, 1, \dots, M - 1$$

4. Repeat **2.** and **3.** until no further distortion reduction

# Lloyd-Max Quantizer Properties

- Zero-mean quantization error  
$$E[(X - \hat{X})] = 0$$
- Quantization error and reconstruction decorrelated  
$$E[(X - \hat{X})\hat{X}] = 0$$
- Variance subtraction property

$$\sigma_{\hat{X}}^2 = \sigma_X^2 - E[(X - \hat{X})^2]$$



# High Rate Approximation I

- Let  $\Delta_i$  be the step size of the scalar quantizer in cell  $i$
- Local density of centroids  $c_i$ :  $g(c_i) = \frac{1}{\Delta_i}$
- Centroid density function

$$\int_{\mathcal{R}} g(x)dx = M$$

M representative levels

- Expected distortion

$$D = \sum_i p_i \frac{\Delta_i^2}{12} \approx \frac{1}{12} \int_{\mathcal{R}} g^{-2}(x) f_X(x) dx$$

- Optimal constrained resolution scalar quantizer

$$\min_g D(g) \quad \text{s.t.} \quad \int_{\mathcal{R}} g(x)dx = M$$



# High Rate Approximation II

- Approximate solution of the "Max quantization problem," assuming high rate and smooth PDF *[Panter, Dite, 1951]*

$$\frac{1}{g(x)} = \frac{1}{M} \frac{\int_{\mathcal{R}} \sqrt[3]{f_X(x)} dx}{\sqrt[3]{f_X(x)}}$$

Distance between two successive quantizer representative levels

Probability density function of  $x$

- Approximation for the quantization error variance:

$$d = E[(X - \hat{X})^2] \approx \frac{1}{12M^2} \left[ \int_x \sqrt[3]{f_X(x)} dx \right]^3$$



# High Rate Approximation III

- High-rate distortion-rate function for scalar Lloyd-Max quantizer:

$$d(R) \cong \varepsilon^2 \sigma_X^2 2^{-2R}$$

$$\leftarrow M = 2^R$$

$$\text{with } \varepsilon^2 \sigma_X^2 = \frac{1}{12} \left[ \int_x \sqrt[3]{f_X(x)} dx \right]^3$$

- Some example values for  $\varepsilon^2$

uniform	1
Laplacian	$\frac{9}{2} = 4.5$
Gaussian	$\frac{\sqrt{3}\pi}{2} \cong 2.721$



# High Rate Approximation IV

- Partial distortion theorem: Each interval makes an (approximately) equal contribution to overall mean-squared error.

$$\begin{aligned} & \Pr\{t_i \leq X < t_{i+1}\} E\left[\left(X - \hat{X}\right)^2 \middle| t_i \leq X < t_{i+1}\right] \\ & \cong \Pr\{t_j \leq X < t_{j+1}\} E\left[\left(X - \hat{X}\right)^2 \middle| t_j \leq X < t_{j+1}\right] \quad \text{for all } i, j \end{aligned}$$

# Entropy-Constrained Scalar Quantizer

- Lloyd-Max quantizer optimum for fixed-rate encoding. How can we do better for variable-length encoding of the quantizer index?
- Problem: For a signal  $x$  with given pdf  $f_X(x)$  find a quantizer

$$d = MSE = E \left[ (X - \hat{X})^2 \right] \rightarrow \min.$$

$$\text{s.t. } R = H(\hat{X}) = - \sum_{q=0}^{M-1} p_q \log_2 p_q$$

- Solution: Lagrangian cost function

$$J = d + \lambda R = E \left[ (X - \hat{X})^2 \right] + \lambda H(\hat{X}) \rightarrow \min.$$

# Iterative Entropy-Constrained SQ Design

1. Guess initial set of representative levels  $\hat{x}_q$ ;  $q = 0, 1, 2, \dots, M-1$  and corresponding probabilities  $p_q$
2. Calculate  $M-1$  decision thresholds

$$t_q = \frac{\hat{x}_{q-1} + \hat{x}_q}{2} - \lambda \frac{\log_2 p_{q-1} - \log_2 p_q}{2(\hat{x}_{q-1} - \hat{x}_q)} \quad q = 1, 2, \dots, M-1$$

3. Calculate  $M$  new representative levels and probabilities  $p_q$

$$\hat{x}_q = \frac{\int_{t_q}^{t_{q+1}} x f_X(x) dx}{\int_{t_q}^{t_{q+1}} f_X(x) dx} \quad q = 0, 1, \dots, M-1$$

4. Repeat **2.** & **3.** until no further reduction in Lagrangian cost



# ECSQ Design with Training Data

1. Guess initial set of representative levels  $\hat{x}_q$ ;  $q = 0, 1, 2, \dots, M - 1$  and corresponding probabilities  $p_q$
2. Assign each sample  $x_i$  in training set  $\mathcal{T}$  to representative minimizing Lagrangian cost  $J_{x_i}(q) = (x_i - \hat{x}_q)^2 - \lambda \log_2 p_q$

$$B_q = \{x \in \mathcal{T} : Q_\lambda(x) = q\} \quad q = 0, 1, 2, \dots, M - 1$$

3. Calculate new representative levels and probabilities  $p_q$

$$\hat{x}_q = \frac{1}{\|B_q\|} \sum_{x \in B_q} x \quad q = 0, 1, \dots, M - 1$$

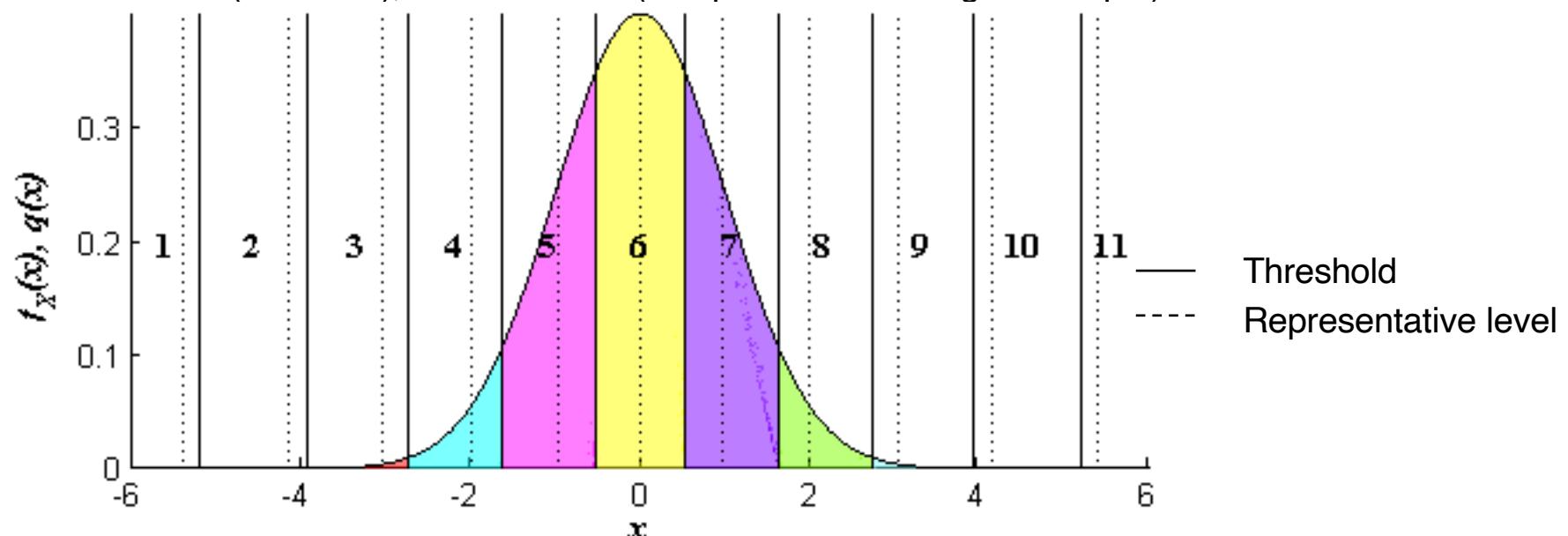
4. Repeat **2.** and **3.** until no further reduction in overall cost

$$J = \sum_{x_i} J_{x_i} = \sum_{x_i} (x_i - Q(x_i))^2 - \lambda \log_2 p_{q(x_i)}$$



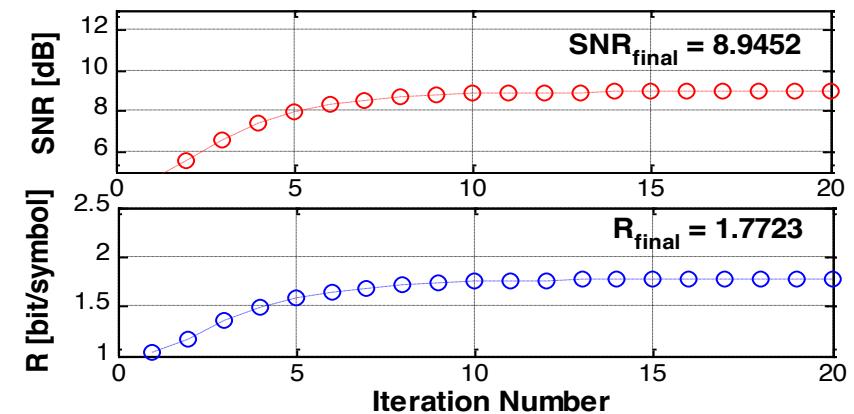
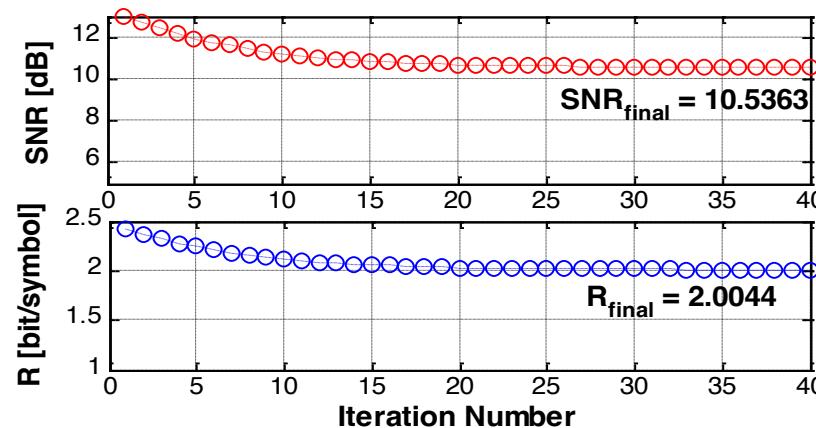
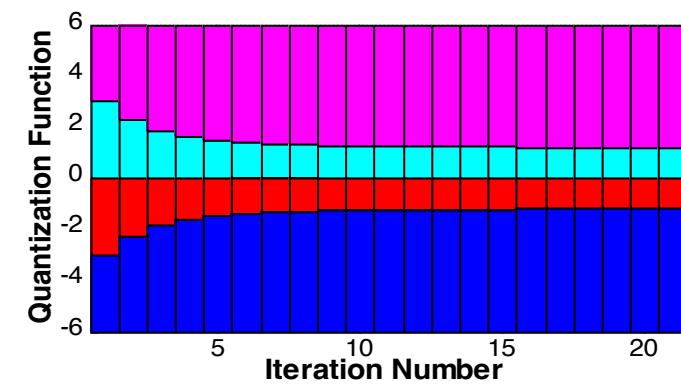
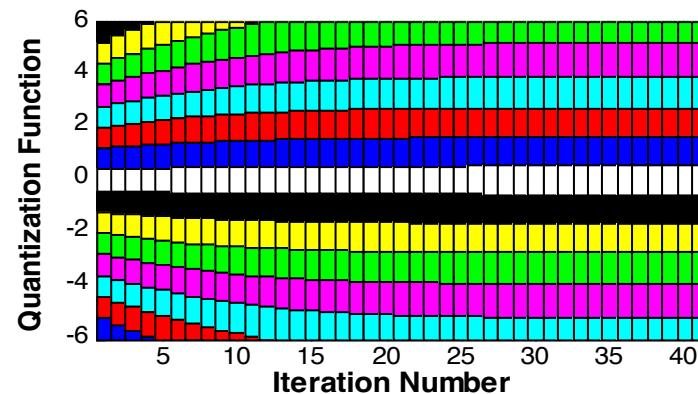
# Example: ECSQ Design I

- $X$  zero-mean, unit-variance Gaussian r.v.
- Design entropy-constrained scalar quantizer with rate  $R \approx 2$  bits, and minimum distortion  $D^*$
- Optimum quantizer, obtained with the entropy-constrained Lloyd algorithm
  - 11 intervals (in  $[-6,6]$ ), almost uniform
  - $D^* = 0.09$  (10.53 dB),  $R = 2.0035$  bits (compare to fixed-length example)



# Example: ECSQ Design II

- Same Lagrangian multiplier is used in all experiments
  - Initial quantizer A, 15 intervals in  $[-6,6]$ , with the same length
  - Initial quantizer B, 4 intervals in  $[-6,6]$ , with the same length



# High Rate Results for ECSQ

- For MSE distortion and high rates, uniform quantizers (followed by entropy coding) are optimum [Gish, Pierce, 1968]
- Entropy and distortion for smooth PDF and fine quantizer interval  $\Delta$

$$H(\hat{X}) \approx h(X) - \log_2 \Delta$$

$$d \approx \int_{-\Delta/2}^{\Delta/2} \varepsilon^2 d\varepsilon = \frac{\Delta^2}{12}$$

- Distortion rate function

$$d(R) \approx \frac{1}{12} 2^{2h(X)} 2^{-2R}$$

is factor  $\frac{\pi e}{6}$  or 1.53 dB from Shannon Lower Bound

$$D(R) \geq \frac{1}{2\pi e} 2^{2h(X)} 2^{-2R}$$



# Comparison: High Rate Performance of SQ

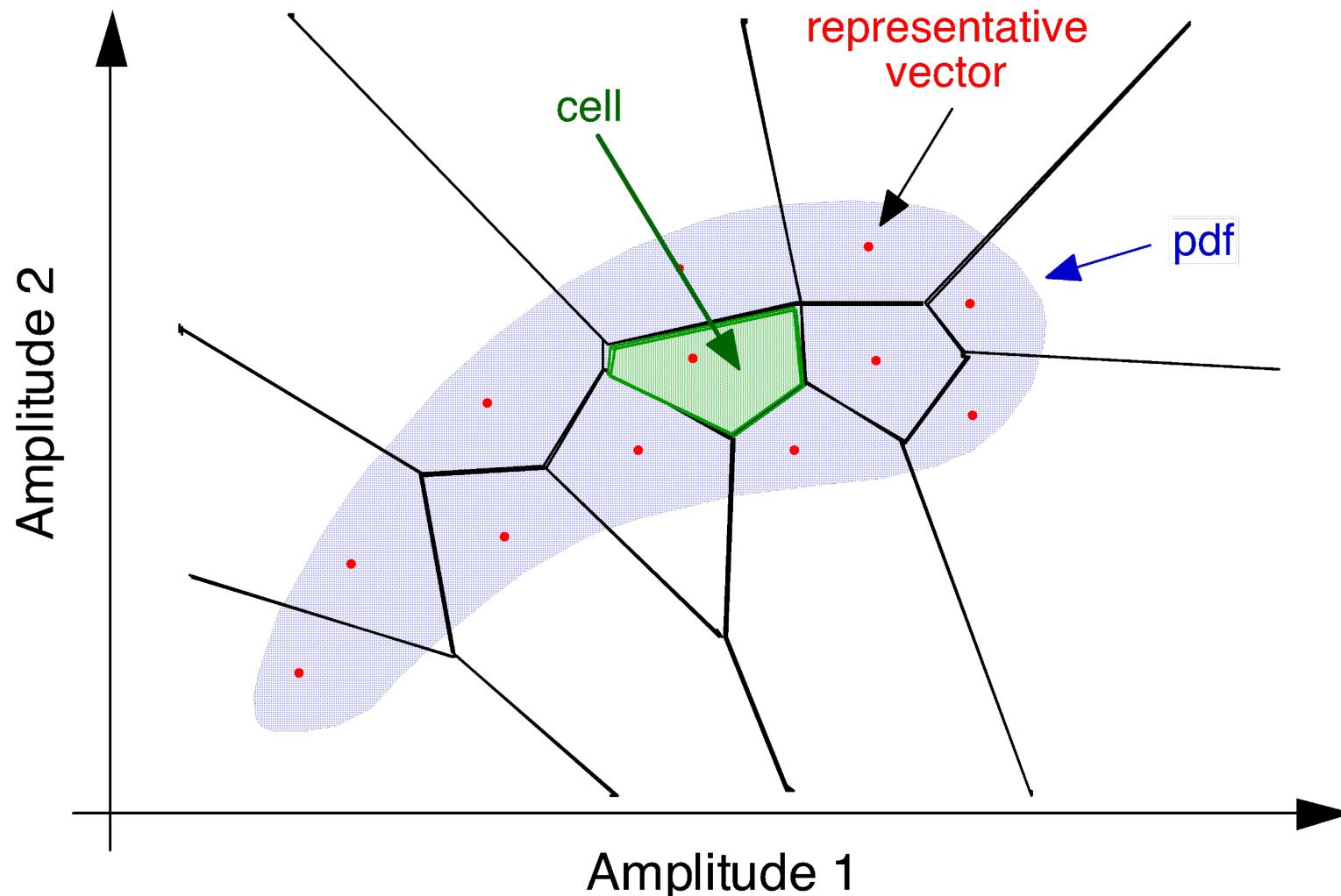
- High-rate distortion-rate function

$$d(R) \cong \varepsilon^2 \sigma_X^2 2^{-2R}$$

- Scaling factor  $\varepsilon^2$

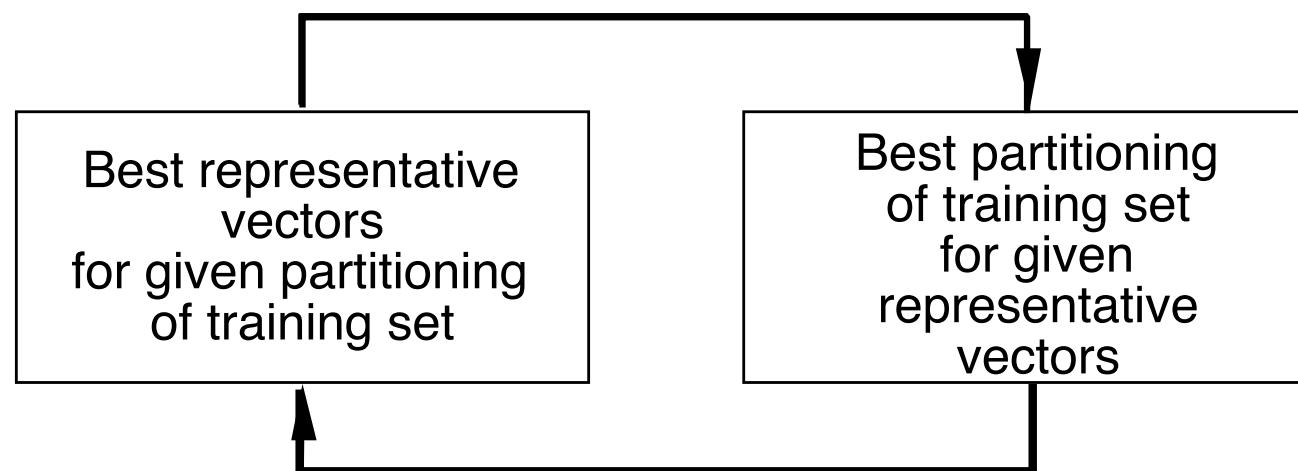
	Shannon LowBd	Lloyd-Max	Entropy-coded
Uniform	$\frac{6}{\pi e} \cong 0.703$	1	1
Laplacian	$\frac{e}{\pi} \cong 0.865$	$\frac{9}{2} = 4.5$	$\frac{e^2}{6} \cong 1.232$
Gaussian	1	$\frac{\sqrt{3}\pi}{2} \cong 2.721$	$\frac{\pi e}{6} \cong 1.423$

# Vector Quantization



# LBG Algorithm

- Lloyd algorithm generalized for VQ *[Linde, Buzo, Gray, 1980]*



- Assumption: fixed code word length
- Code book unstructured: full search

# Design of Entropy-Constrained VQ

- Extended LBG algorithm for entropy-coded VQ  
*[Chou, Lookabaugh, Gray, 1989]*
- Lagrangian cost function: solve unconstrained problem rather than constrained problem

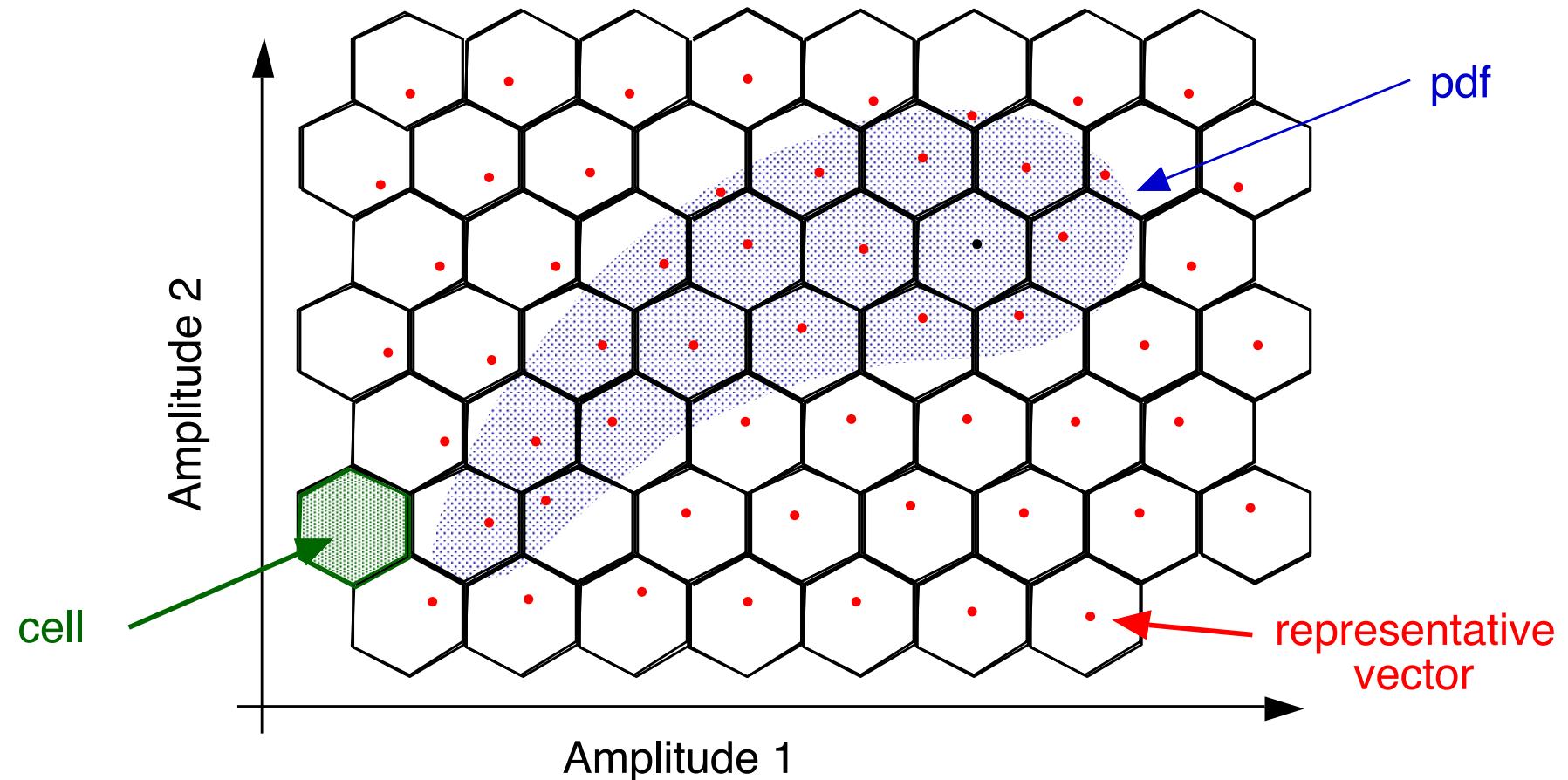
$$J = d + \lambda R = E \left[ \|X - \hat{X}\|^2 \right] + \lambda H(\hat{X}) \rightarrow \min.$$

- Unstructured code book: full search for

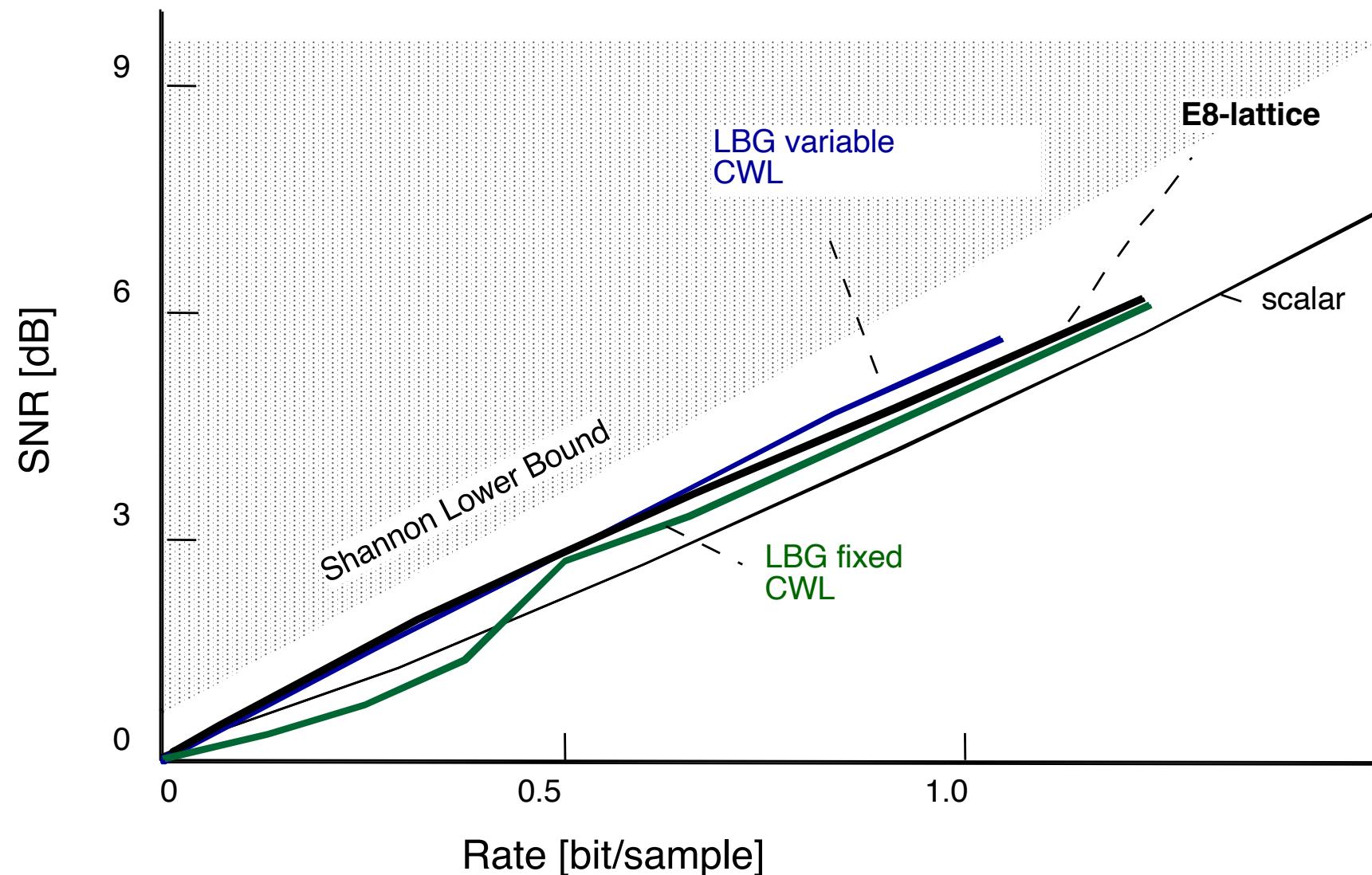
$$J_{x_i}(q) = \|x_i - \hat{x}_q\|^2 - \lambda \log_2 p_q$$

The most general coder structure:  
Any source coder can be interpreted as VQ with VLC!

# Lattice Vector Quantization



# 8D VQ of Memoryless Laplacian Source



# 8D VQ of a Gauss-Markov Source

