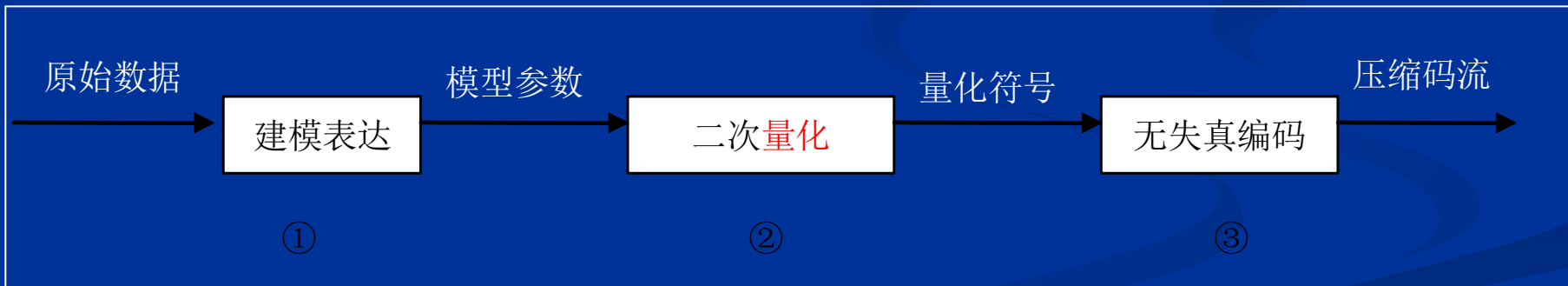


# 第九章 标量量化

- 量化：最简单、最重要的有失真编码
- 量化问题
- 各种量化方法
  - 固定码长的均匀量化
    - 均匀分布
    - 已知pdf的非均匀分布
    - 未知/变化分布
  - pdf最优的非均匀量化
  - 压扩量化
  - 熵约束量化
- 应用

# 量化

- 量化：用一个很小的集合表示一个大集合（可能是无限大）的过程
  - 如：给定信源为 $[-10.0, 10.0]$ 之间的实数
  - 量化： $Q(x) = \lfloor x+0.5 \rfloor$ 
    - $[-10.0, 10.0] \rightarrow \{-10, -9, \dots, -1, 0, 1, 2, \dots, 9, 10\}$
    - 失真：2.95, 3.16, 3.05  $\rightarrow$  3
- 量化是有失真压缩的一个有效工具



- 对模拟信号，量化还包括A/D转换中的一次量化

# 量化过程

## ■ 两个方面

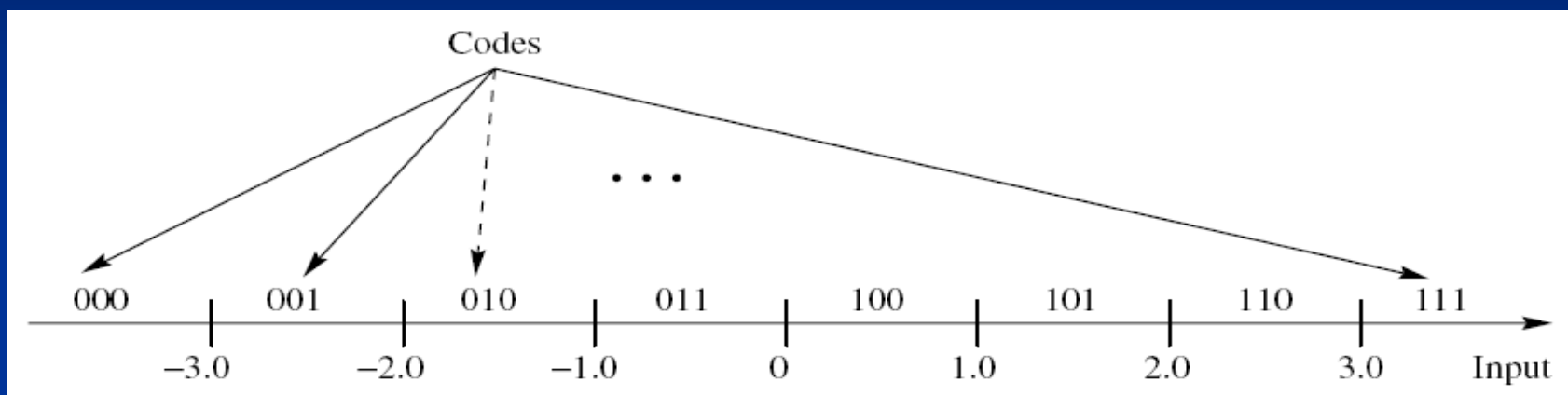
### ■ 编码器映射

- 将一个范围内的值映射成一个区间，每个区间用唯一码字表示
- 若信源是模拟的  $\rightarrow$  A/D 转换器
- **有关信源的知识**有助于选择合适的范围

### ■ 解码器映射

- 将码字映射成某个范围内的一个值
  - 如果输出为模拟的  $\rightarrow$  D/A 转换器
  - **信源分布的知识**有助于选择较好的近似
- 量化器 = 编码器 + 解码器

# 例：量化

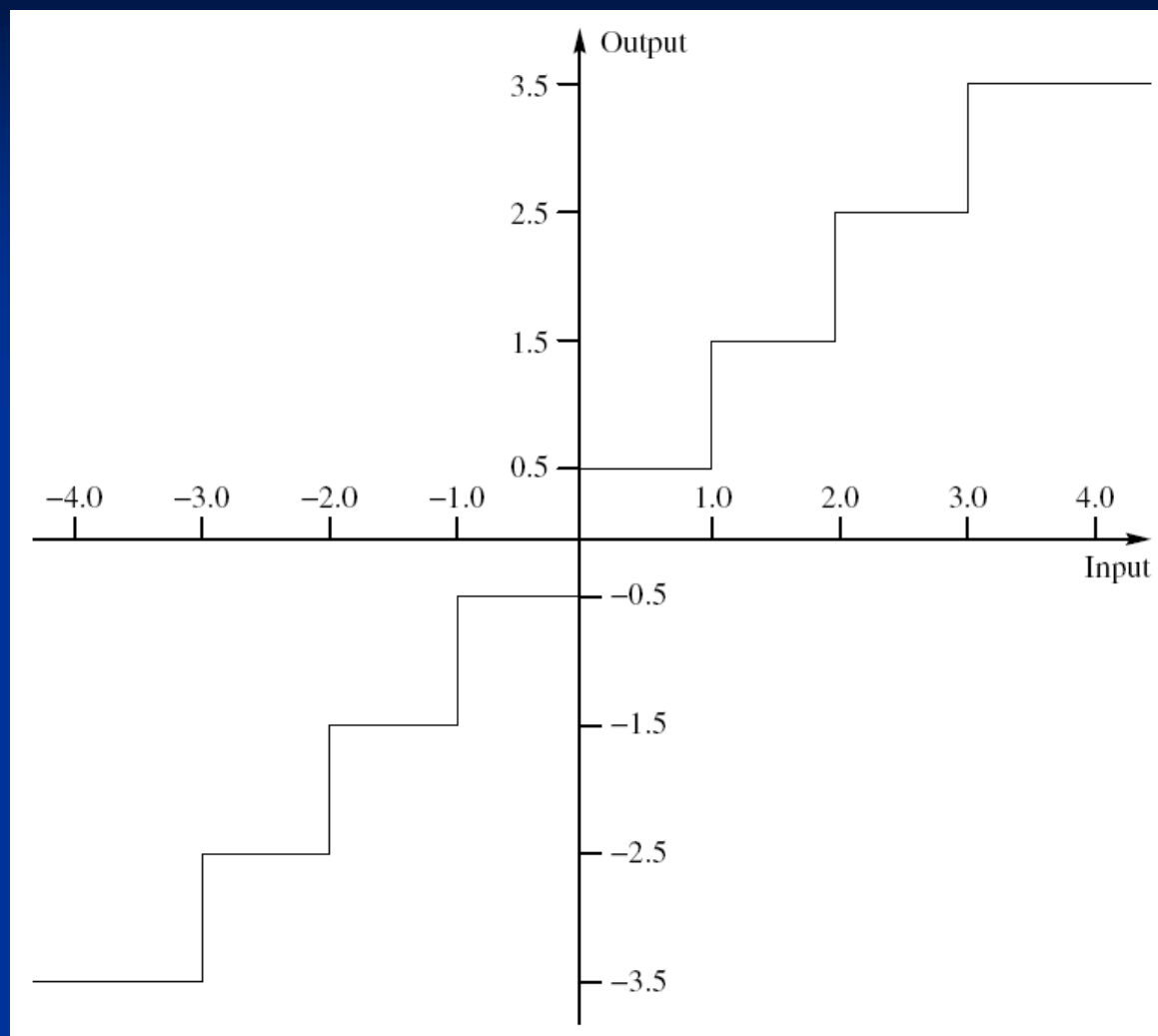


Input Codes	Output
000	-3.5
001	-2.5
010	-1.5
011	-0.5
100	0.5
101	1.5
110	2.5
111	3.5

$t$	$4 \cos(2\pi t)$	A/D Output	D/A Output	Error
0.05	3.804	111	3.5	0.304
0.10	3.236	111	3.5	-0.264
0.15	2.351	110	2.5	-0.149
0.20	1.236	101	1.5	-0.264

3比特量化器

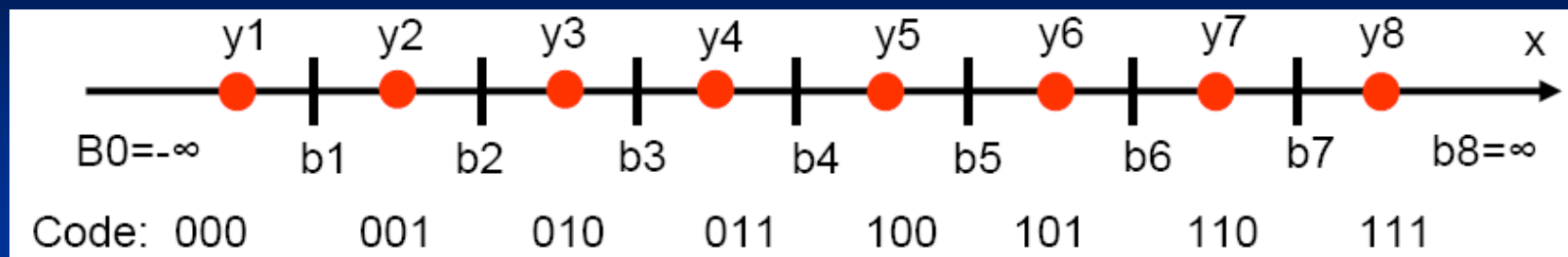
# 量化器的输入输出映射



编码器的区间构造与解码器的重构水平选择都影响保真度，所以将二者当成一个整体看待：  
量化器

量化器 = 编码器 + 解码器

# 量化问题形式化描述

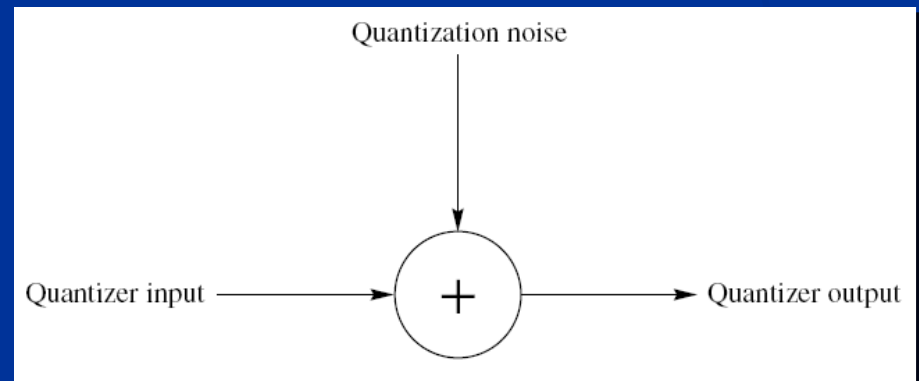


- 输入：
  - $X$  – 随机变量
  - $f_X(x)$  – 概率密度函数 (pdf)
- 输出：
  - $\{b_i\}_{i=0..M}$  决策边界
  - $\{y_i\}_{i=1..M}$  重构水平
- 满足率失真准则
- 离散过程通常用连续分布近似，可以大幅简化设计过程
  - 如：对像素差分用Laplacian建模
  - 若信源无界，则第一个/最后一个边界 =  $\pm\infty$

# 量化误差

$$Q(x) = y_i, b_{i-1} < x \leq b_i$$

- 均方量化误差(Mean squared quantization error, MSQE)
- 量化误差亦称为
  - 量化噪声
  - 量化失真



$$D = MSE = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x) dx = \sum_{i=1}^M \int_{b_{i-1}}^{b_i} (x - y_i)^2 f(x) dx$$

# 量化问题形式化描述 (2)

- 比特率  $w$  / 定长度码字
  - $R = \lceil \log_2 M \rceil$ ,  $M$  为量化器输出的数目
  - 如:  $M = 8 \rightarrow R = 3$
- 量化器设计问题
  - 给定:
    - 输入 pdf  $f_X(x)$  和  $M$
  - 求:
    - 决策边界  $\{b_i\}$
    - 重构水平  $\{y_i\}$
  - 使得:
    - MSQE 最小



# 量化问题形式化描述 (3)

- 比特率  $w$  / 变长码字
  - $R$  取决于决策边界
  - 例:

$$P(y_i) = \int_{b_{i-1}}^{b_i} f_X(x) dx$$

$$R = \sum_{i=1}^M l_i \int_{b_{i-1}}^{b_i} f_X(x) dx$$

$y_1$	1110
$y_2$	1100
$y_3$	100
$y_4$	00
$y_5$	01
$y_6$	101
$y_7$	1101
$y_8$	1111

# 量化问题形式化描述 (4)

## ■ 码率最优形式化描述

- 给定：
  - 失真限制  $\underline{\sigma_q^2 \leq D^*}$
- 求：
  - $\{b_i\}, \{y_i\}$  和 二进制码字
- 使得：
  - $R$  最小

## ■ 失真最优化形式化描述

- 给定：
  - 码率限制  $\underline{R \leq R^*}$
- 求：
  - $\{b_i\}, \{y_i\}$  和 二进制码字
- 使得：
  - $\sigma_q^2$  最小

区间分割和区间的表示决定失真

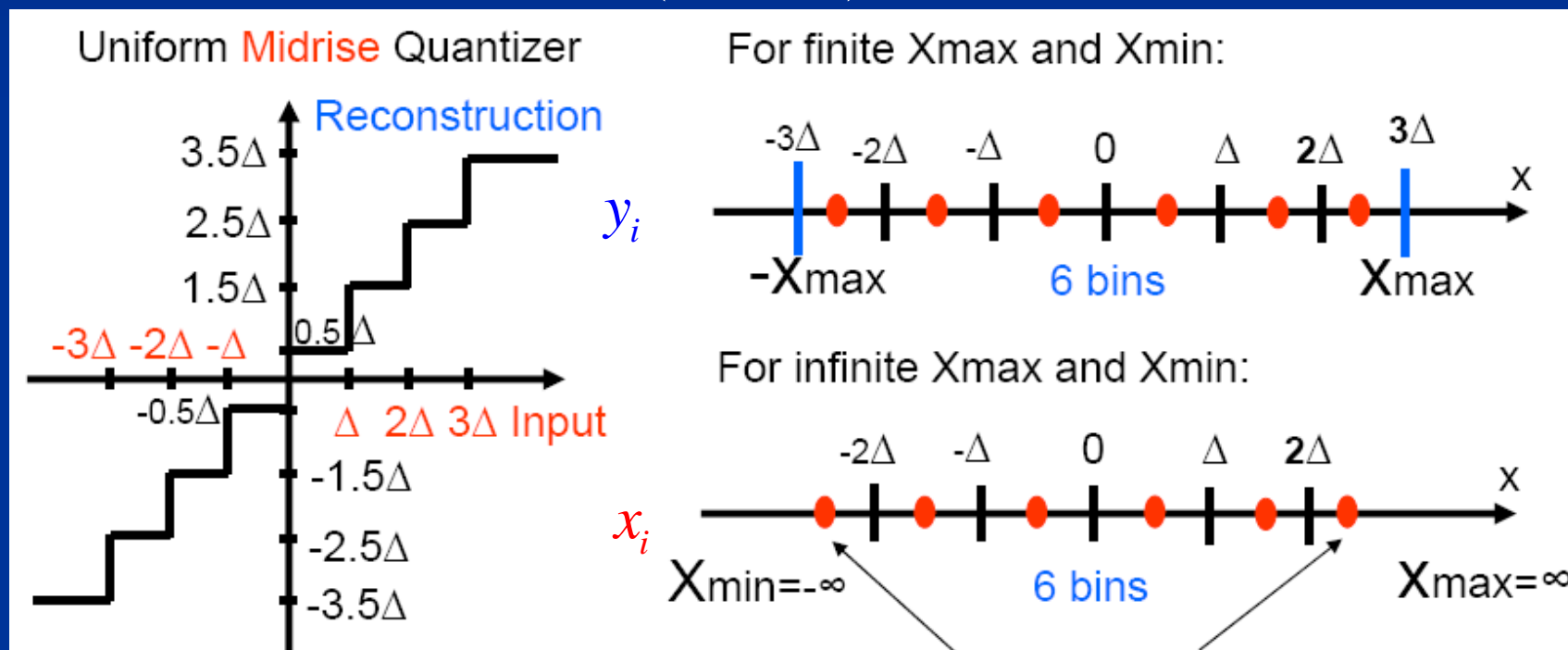
区间分割和区间的二进制表示决定码率

# 均匀量化

- 所有区间大小相同
  - 边界为等间隔分布，空间大小均为 $\Delta$
  - 除了最外面两个区间
- 重构
  - 通常选择区间中点
- 中升(Midrise)量化器
  - 0 不是一个输出水平
- 中平(Midtread)量化器
  - 0 是一个输出水平

# 均匀量化 (Midrise)

- $b_i, y_i$  在空间上均匀分布, 空间均为  $\Delta$ 
  - 对内部区间,  $y_i = 1/2(b_{i-1} + b_i)$

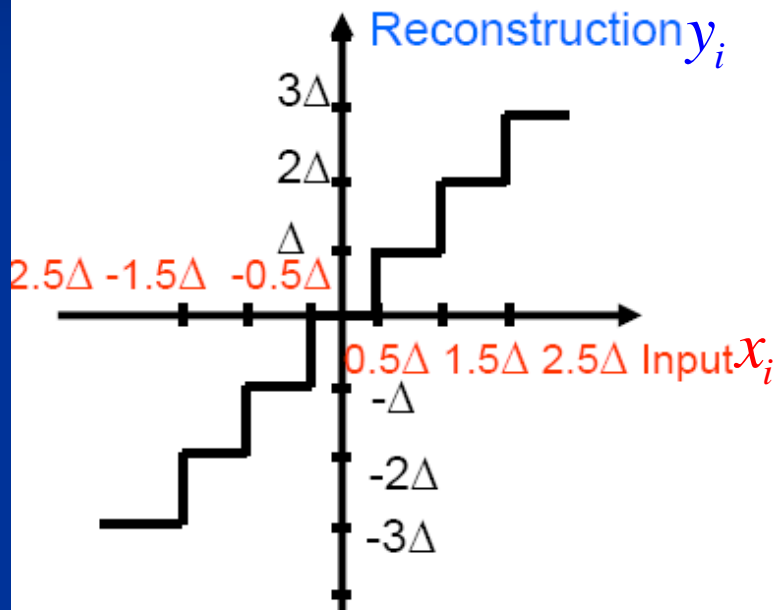


共有偶数个重构水平  
0不是一个重构水平

最外部的两个重构水平内部  
仍是一个步长大小

# 均匀量化 (Midtread)

Uniform Midtread Quantizer

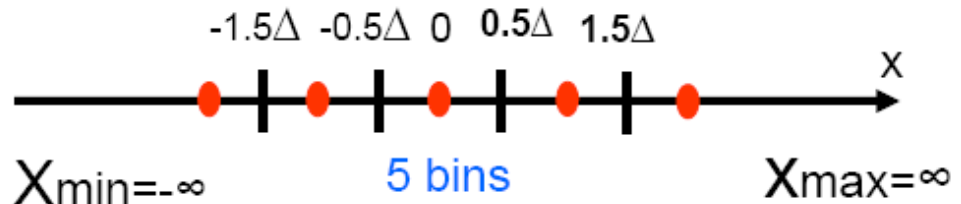


- Odd number of recon levels
- 0 is a recon level
- Desired in image/video coding

For finite  $X_{\max}$  and  $X_{\min}$ :



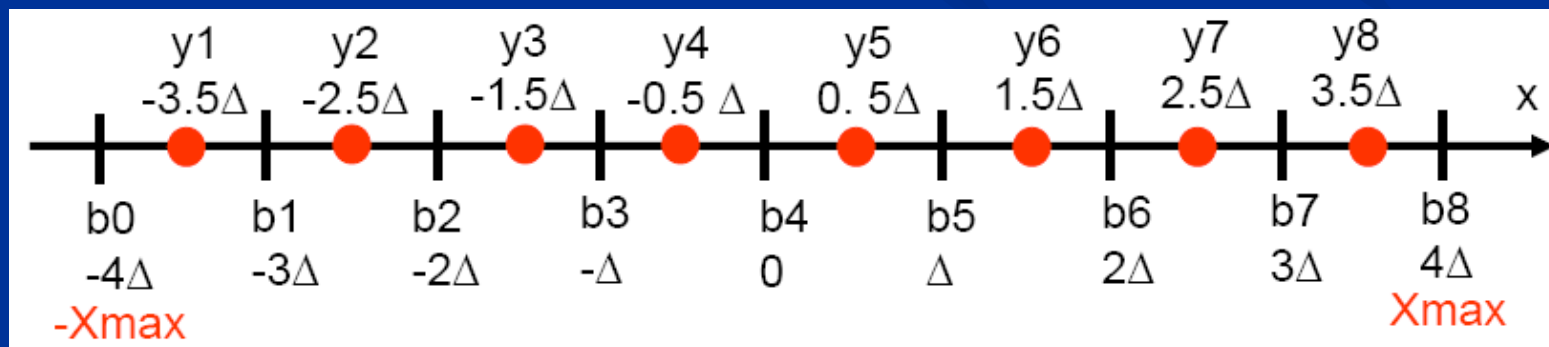
For infinite  $X_{\max}$  and  $X_{\min}$ :



# 均匀分布的均匀量化

## —以Midrise量化器为例

- 输入：均匀分布的信源  $Uniform[-X_{\max}, X_{\max}]$ :  $f_X(x) = 1/2X_{\max}$
- 输出： $M$ 个水平的均匀量化器（对Midrise量化， $M$ 为偶数）
  - 步长： $\Delta = 2X_{\max}/M$
  - 失真： $\sigma_q^2 = 2 \sum_{i=1}^{M/2} \int_{(i-1)\Delta}^{i\Delta} \left( x - \frac{2i-1}{2} \Delta \right)^2 \frac{1}{2X_{\max}} dx = \frac{\Delta^2}{12}$



# 均匀分布的均匀量化

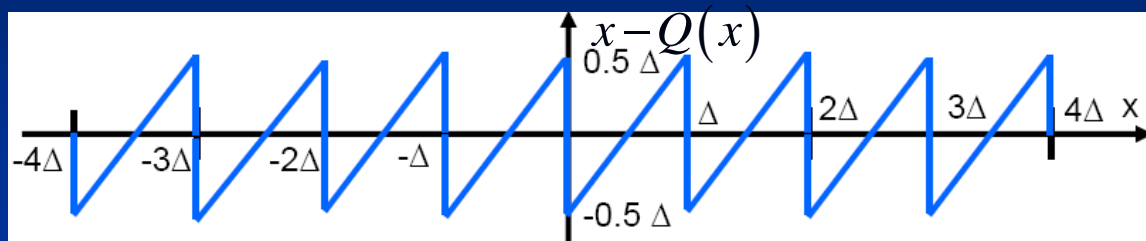
- 若考虑量化误差

- $q = x - Q(x)$

- $q \in [-\Delta/2, \Delta/2]$

- 量化误差也是均匀分布:  $f_Q(q) = 1/\Delta$

- 所以 $q$ 的方差为  $\sigma_q^2 = \Delta^2/12$



$$\sigma_s^2 = (2X_{\max})^2 / 12$$

- 则信噪比为

$$\begin{aligned} \text{SNR(dB)} &= 10 \log_{10} \left( \frac{\sigma_s^2}{\sigma_q^2} \right) = 10 \log_{10} \left( \frac{(2X_{\max})^2}{12} \frac{12}{\Delta^2} \right) = 10 \log_{10} \left( \frac{(2X_{\max})^2}{12} \frac{12}{\left( \frac{2X_{\max}}{M} \right)^2} \right) \\ &= 10 \log_{10} (M^2) = 20 \log_{10} (2^n) = 6.02n \text{ dB} \end{aligned}$$

若 $n \rightarrow n+1$ , 则步长减为一半, 噪声方差减为1/4, SNR增加6 dB

# 例：8 $\rightarrow$ 1, 2, 3 bits/pixel



1 比特量化：  
[0, 127]  $\rightarrow$  64  
[128, 255]  $\rightarrow$  196

2 比特量化：  
边界：{[0, 64, 128,  
196, 255]}  
重构水平{32, 96,  
160, 224}

3 比特量化

变暗、只保留轮廓

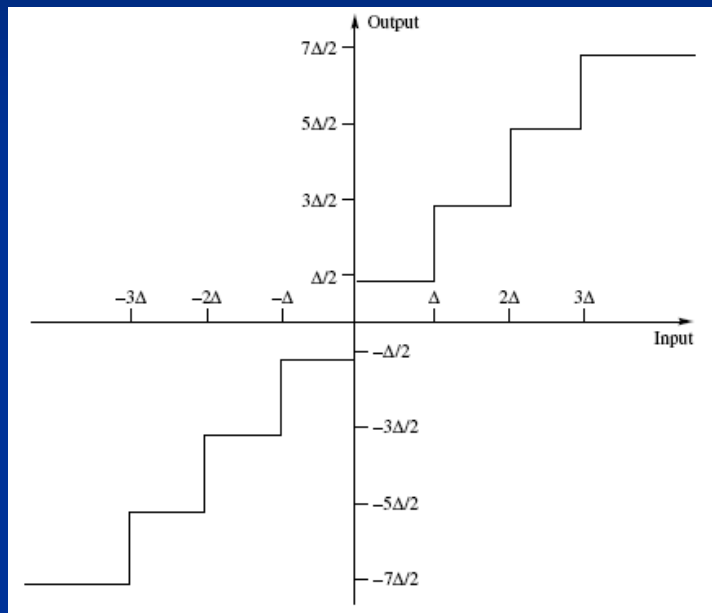


# 非均匀分布的均匀量化器

- 例非均匀分布:  $x \in [-100, 100]$ ,  $P(x \in [-1, 1]) = 0.95$
- 问题
  - 设计一个 8水平的量化器
- 若采用最简单的方法, 步长为25, 则
  - 95%的样本  $([-1, 1])$  被两个数字表示:
    - -12.5 和 12.5
    - 最小QE  $\leq 11.5$
  - 最大 QE = 12.5
- 考虑另一个方案
  - 步长 = 0.3
  - 最大QE = 98.5, 但是95%的时间  $\text{QE} < 0.15$

# 最优化 MSQE

将失真表示为步长的函数。要求出步长的最优值，即对方程求导并置0



$$\sigma_q^2 = 2 \sum_{i=1}^{M/2-1} \int_{(i-1)\Delta}^{i\Delta} \left( x - \frac{2i-1}{2} \Delta \right)^2 f_X(x) dx$$

$$+ 2 \int_{(M/2-1)\Delta}^{\infty} \left( x - \frac{M-1}{2} \Delta \right)^2 f_X(x) dx$$

对特定的pdf, 可数值求解

$$\frac{\partial \sigma_q^2}{\partial \Delta} = - \sum_{i=1}^{M/2-1} (2i-1) \int_{(i-1)\Delta}^{i\Delta} \left( x - \frac{2i-1}{2} \Delta \right) f_X(x) dx$$

$$- (M-1) \int_{(M/2-1)\Delta}^{\infty} \left( x - \frac{M-1}{2} \Delta \right) f_X(x) dx = 0$$

# 例：最佳步长

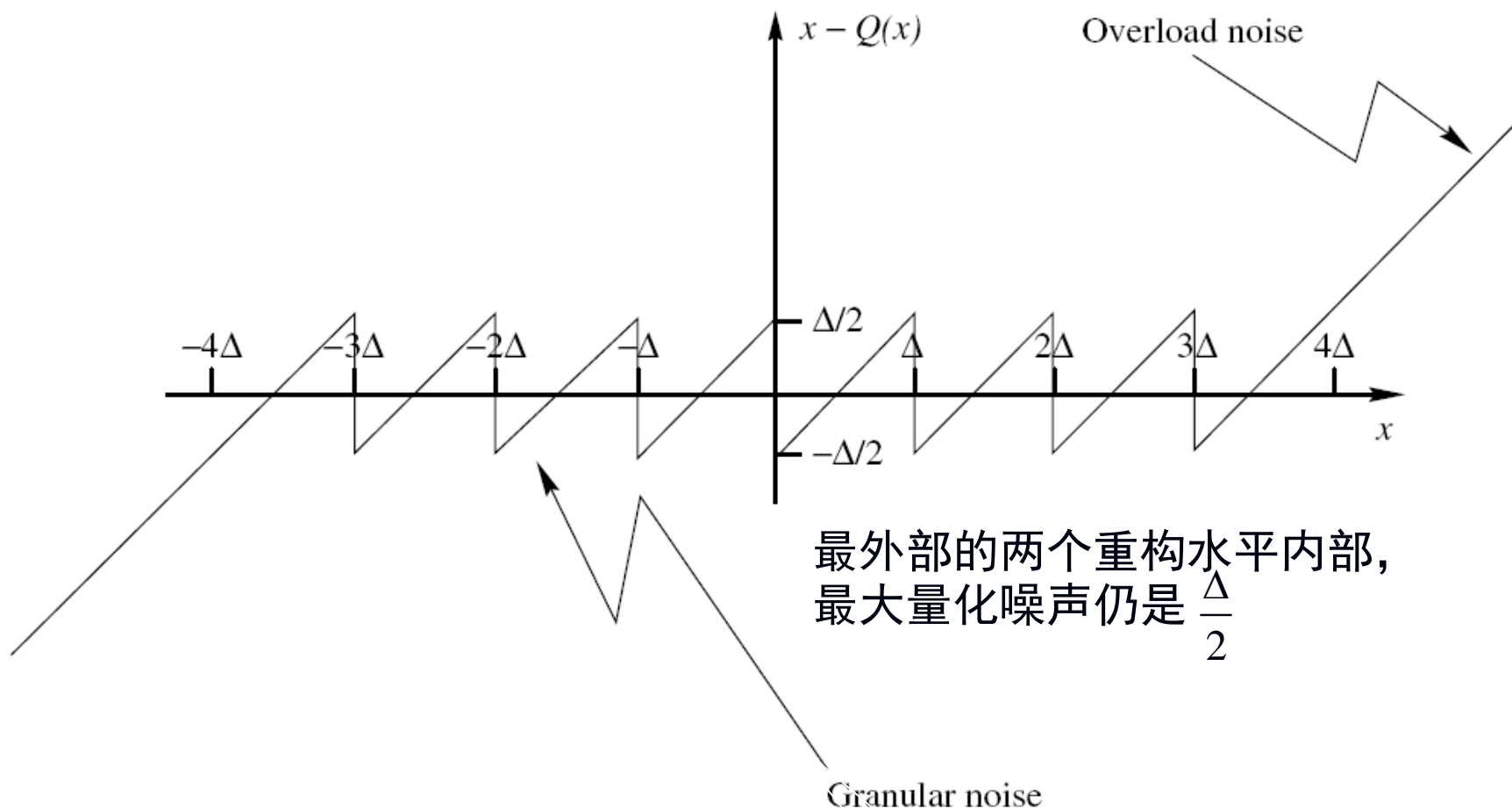
Alphabet Size	Uniform		Gaussian		Laplacian	
	Step Size	SNR	Step Size	SNR	Step Size	SNR
2	1.732	6.02	1.596	4.40	1.414	3.00
4	0.866	12.04	0.9957	9.24	1.0873	7.05
6	0.577	15.58	0.7334	12.18	0.8707	9.56
8	0.433	18.06	0.5860	14.27	0.7309	11.39
10	0.346	20.02	0.4908	15.90	0.6334	12.81
12	0.289	21.60	0.4238	17.25	0.5613	13.98
14	0.247	22.94	0.3739	18.37	0.5055	14.98
16	0.217	24.08	0.3352	19.36	0.4609	15.84
32	0.108	30.10	0.1881	24.56	0.2799	20.46

注：每种分布的方差为1

0附近pdf越大，步长越小  
SNR与比特数的关系与 $6.02n$ 相差越远

# 3比特中升量化器的QE

无界误差



有界误差

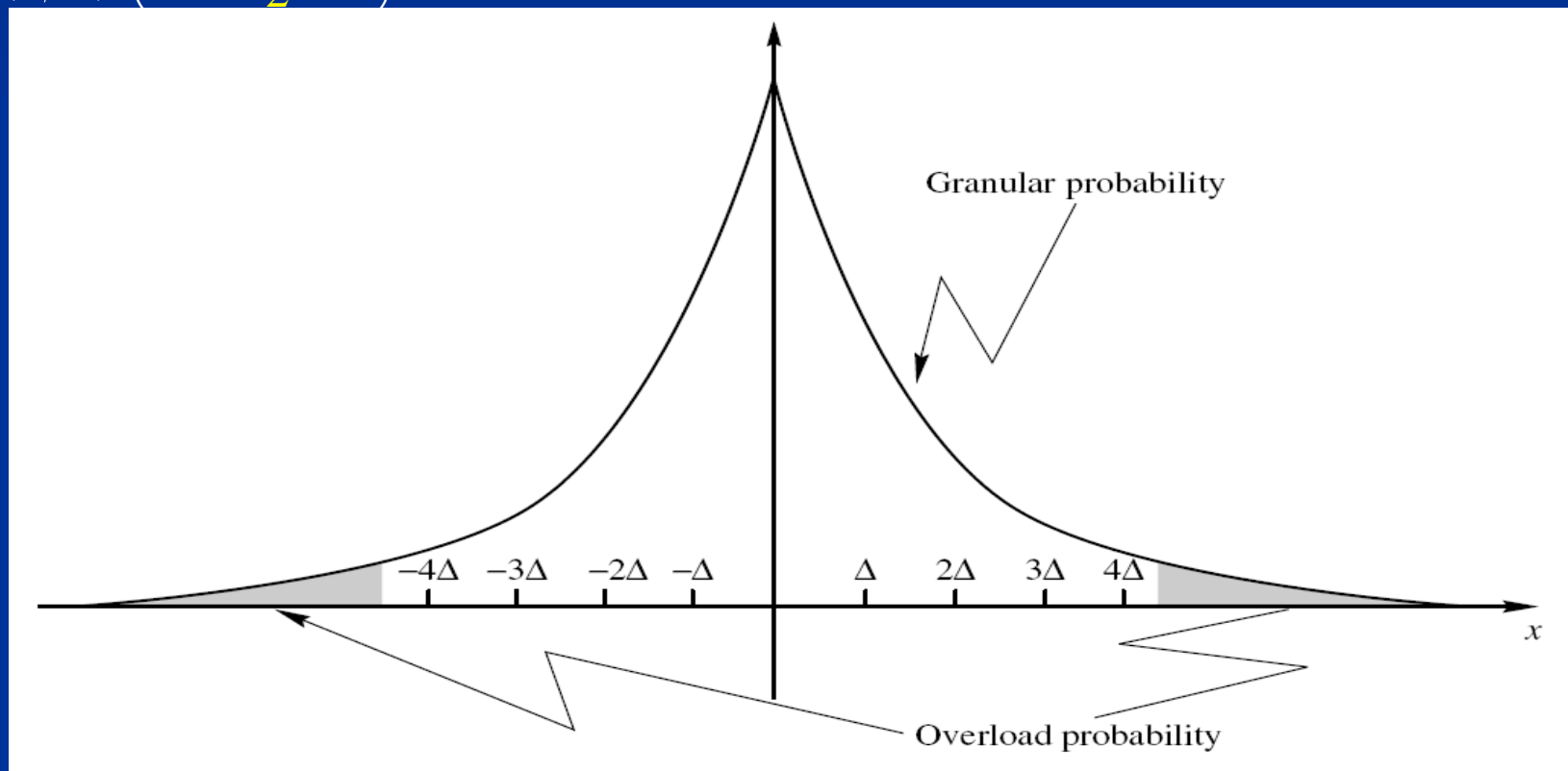
# 过载/粒状区域

$$\sigma_q^2 = 2 \sum_{i=1}^{M/2-1} \int_{(i-1)\Delta}^{i\Delta} \left( x - \frac{2i-1}{2}\Delta \right)^2 f_X(x) dx \quad \text{粒状噪声} \uparrow$$

$$+ 2 \int_{(M/2-1)\Delta}^{\infty} \left( x - \frac{M-1}{2}\Delta \right)^2 f_X(x) dx \quad \text{过载噪声} \downarrow$$

$\Delta \uparrow$

选择步长是一个折中的结果  
尾部较重的分布步长更大



负载因子:  $f_l = \frac{\max \text{ granular value}}{\text{stdev}}$ ,  $f_l = 4 \rightarrow "4\sigma \text{ loading}"$

# 方差不匹配的影响

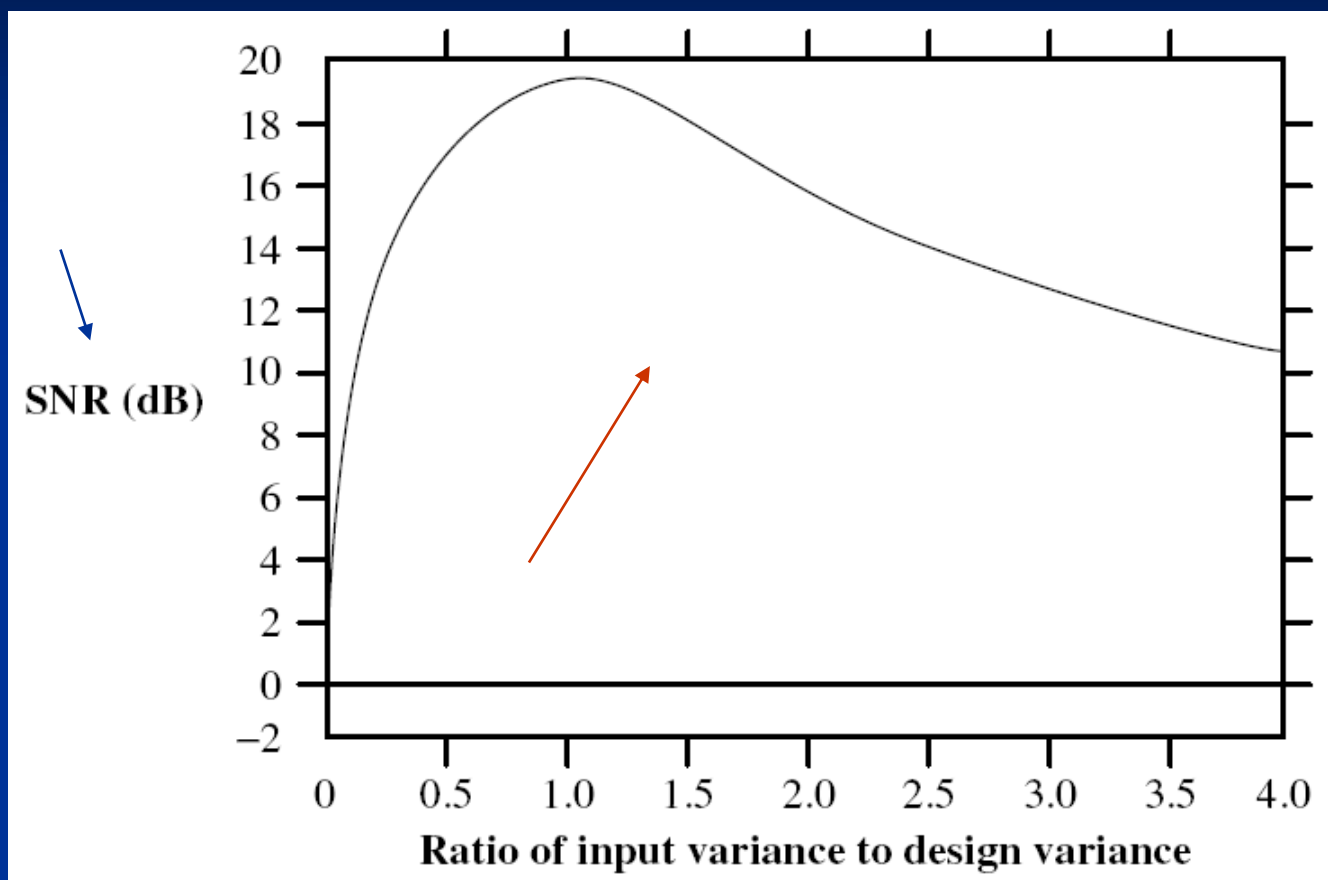
当输入方差小于假定方差，SNR下降很快

MSQE减小，因为过载噪声减小

但是由于输入方差小，SNR中的信号/噪声比率减小很快

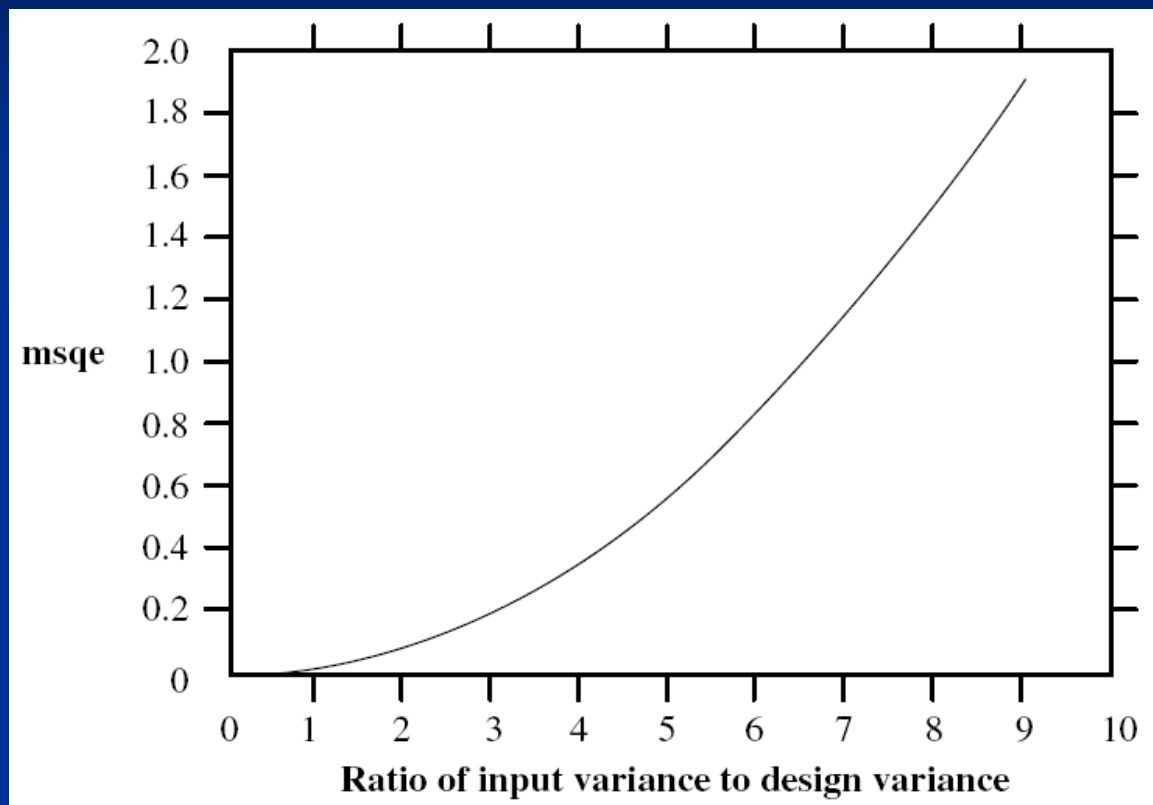
当输入方差大于假定方差

MSQE相应增大，但是由于输入信号的能量增大，信号/噪声比率减小很慢



输入信源：Gaussian，均值为0， $\sigma_x^2 = E(X^2)$   
4比特Gaussian均匀量化器

# 方差不匹配的影响(2)



信噪比的下降并非  
一定与MSQE的增大  
有直接关系

输入信源: Gaussian, 均值为0,  $\sigma_x^2 = E(X^2)$   
4比特Gaussian均匀量化器

# 分布不匹配的影响

8水平量化器, SNR

Input Distribution	Uniform Quantizer	Gaussian Quantizer	Laplacian Quantizer	Gamma Quantizer
Uniform	18.06	15.56	13.29	12.41
Gaussian	12.40	14.27	13.37	12.73
Laplacian	8.80	10.79	11.39	11.28
Gamma	6.98	8.06	8.64	8.76

设计的步长逐渐比“正确的”步长大，性能下降很快  
(类似于方差比预计的方差小)

当失配导致步长小于最优步长时，性能下降的幅度大于量化器步长大于最优值的幅度。



# 自适应量化器

## ■ 思想

- 不是静态方法，而是与真实数据相适应
  - 实际输入的几个统计指标：均值、方差、pdf

## ■ 前向自适应 (离线offline)

- 将信源分块
- 分析块的统计特性
- 设置量化方案
  - 边信道 (Side channel)

## ■ 后向自适应 (在线online)

- 基于量化器输出自适应
- 无需边信道

# 前向自适应量化器（FAQ）

## ■ 选择块大小 - tradeoff!

### ■ 太大

- 分辨率不够，不能抓住输入的变化
- 延迟增大

### ■ 太小

- 需要更多边信息

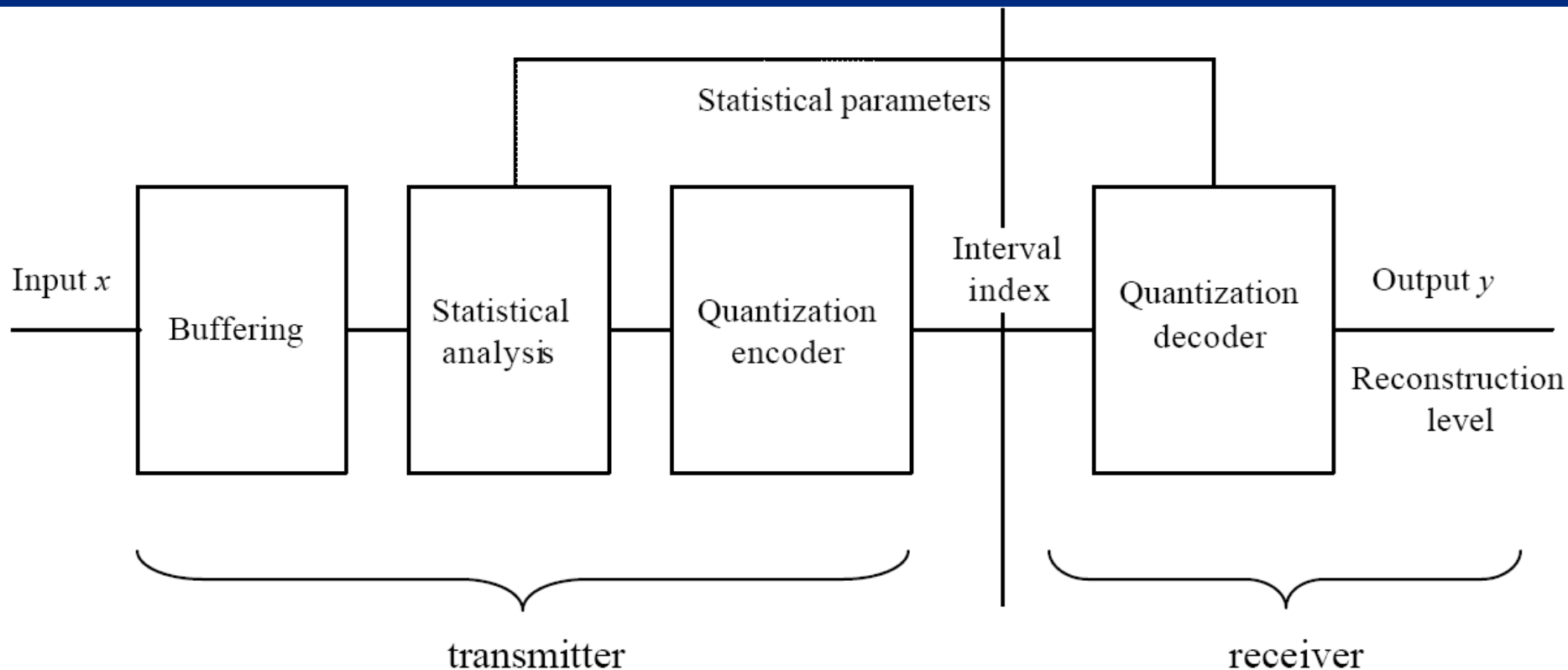
## ■ 假设均值为0（前提）

### ■ 方差估计

$$\hat{\sigma}_q^2 = \frac{1}{N} \sum_{i=0}^{N-1} x_{n+i}^2$$

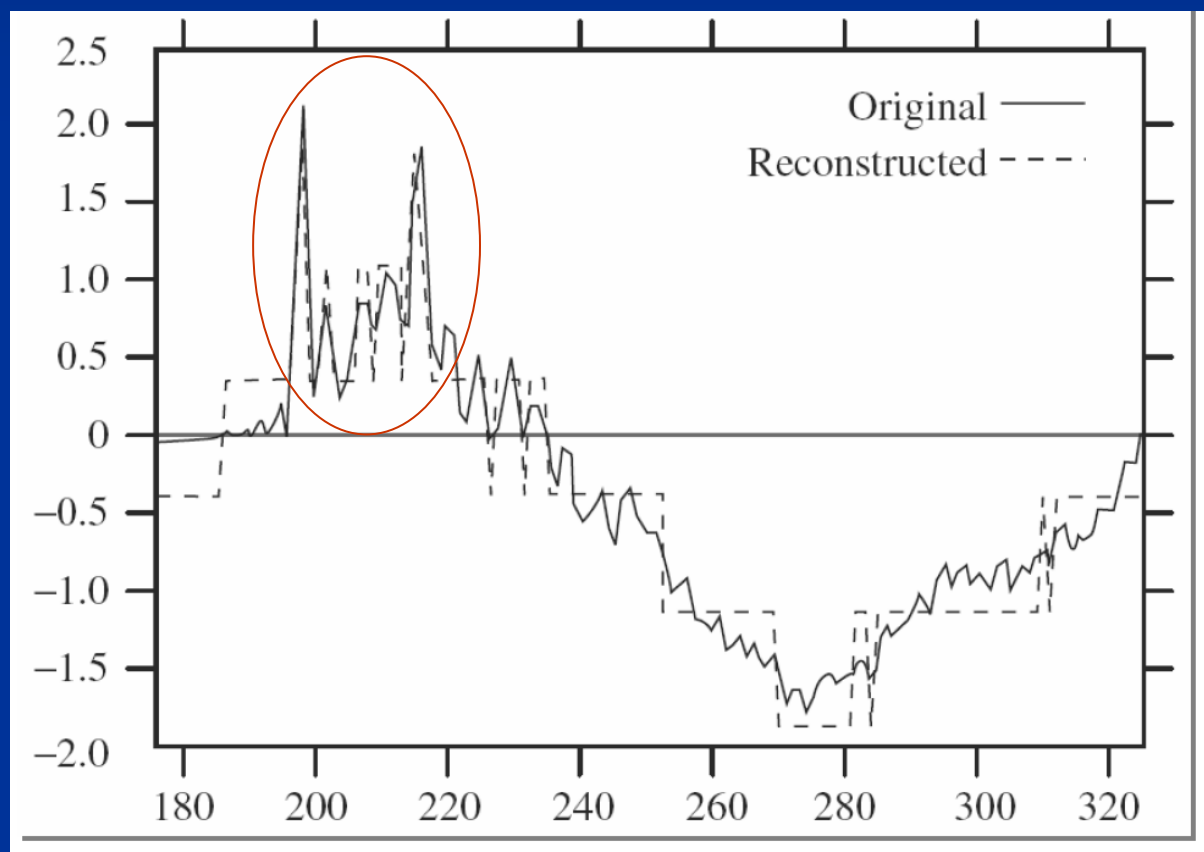
- 方差也应被量化表示,其比特数远远大于样本的量化比特

# 前向自适应量化器 (FAQ)



# 例：语音量化

- 8kHz, 16比特  $\rightarrow$  3比特定长码 男性说test, 4000个样本

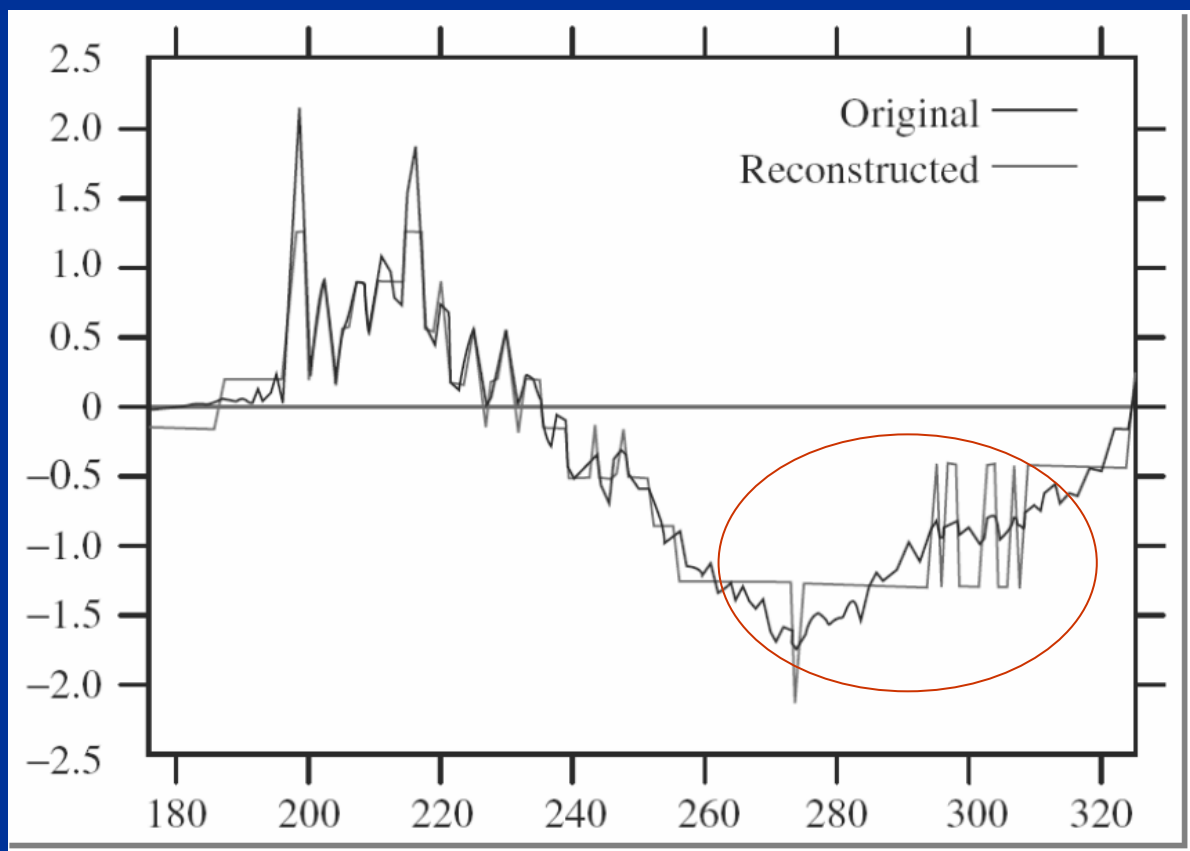


失真较大

# 例：语音量化 (2)

## ■ 16比特 → 3比特FAQ

- 块大小：128个样本
- 方差：用8比特量化



失真较小

红色区域还有提升空间

# 改进的FAQ

## ■ 改进

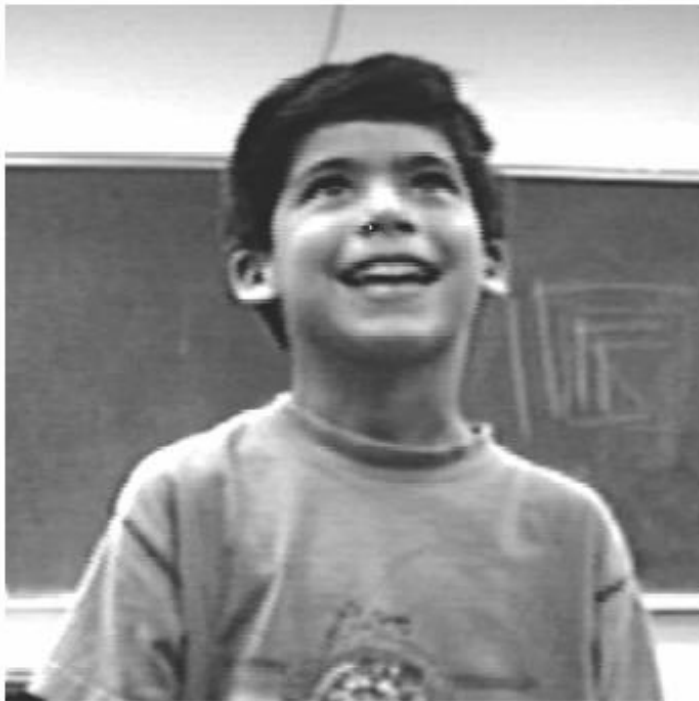
- 假设信源在不同区域内满足均匀pdf，但
- 记录每块的最大/最小值（可作为边信息传送）

## ■ 例：Sena图像

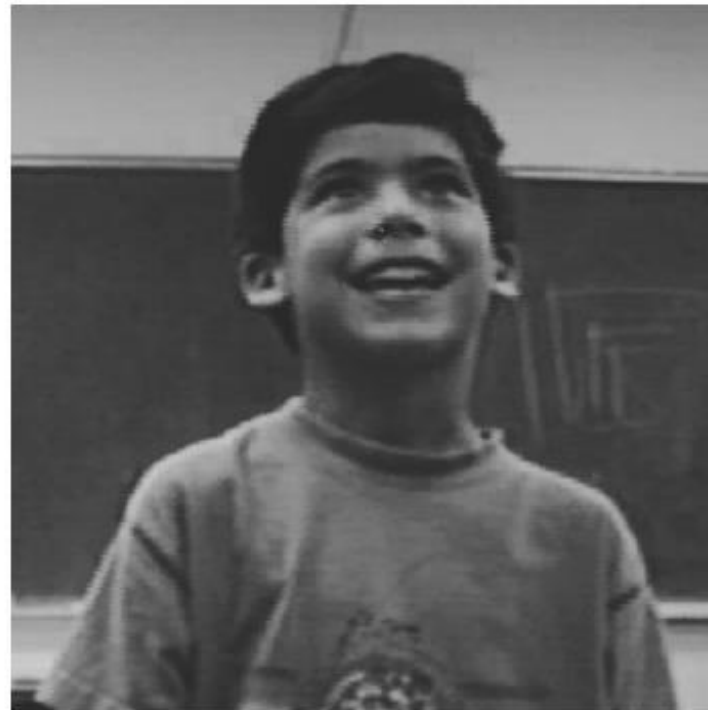
- $8 \times 8$  块（每块的最大值和最小值都用8比特表示）
- 3比特量化
- 共：（边信息开销）

$$\frac{16}{8 \times 8} = 0.25 \text{ bits/pixel}$$

# 例：改进的FAQ



原始图像：8bits/pixel



量化：3.25 bits/pixel

# 后向自适应量化器（BAQ）

## ■ 观察

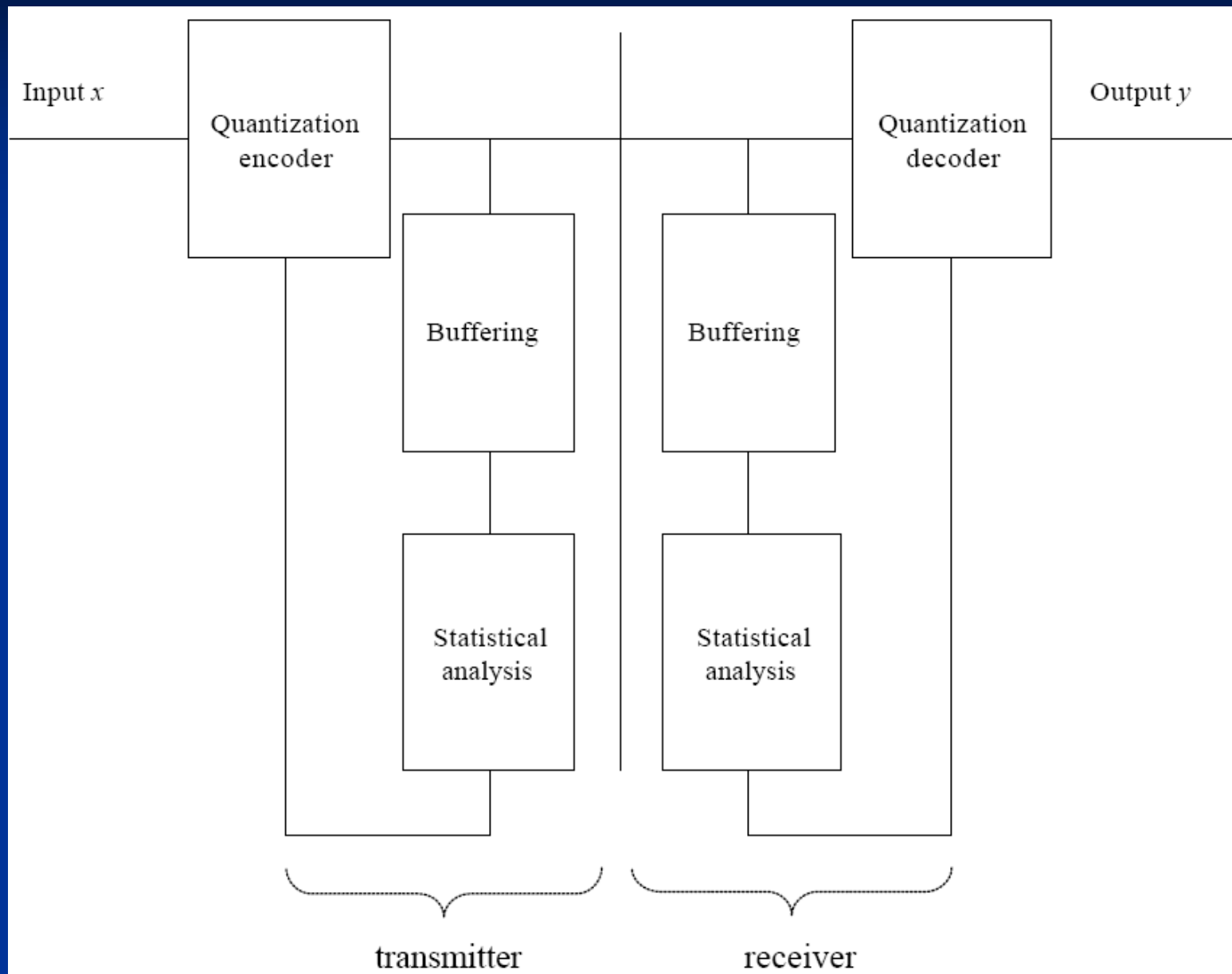
- 解码器只能看到量化器的输出  
→ 只能根据量化器输出进行自适应

## ■ 问题

- 只根据输出，如何减少不匹配信息
- 如果知道pdf，这是可能的...
- ...耐心：观察**长时间的**量化器输出，推测是否发生了不匹配现象
  - 如果匹配，落入某区间的概率与预定的pdf一致，否则，发生了不匹配现象
  - 如果步长比应有的步长小，输入落入外侧区间的数目偏大；相反，输入落入内侧区间的数目偏大



# 后向自适应量化器 (BAQ)



# Jayant 量化器

贝尔实验室的Jayant证明：完全不需要长时间观察量化器输出。在观察单次输出之后就可以调整量化器步长。即：单字记忆量化

## ■ 思想 (Quantization with one word memory)

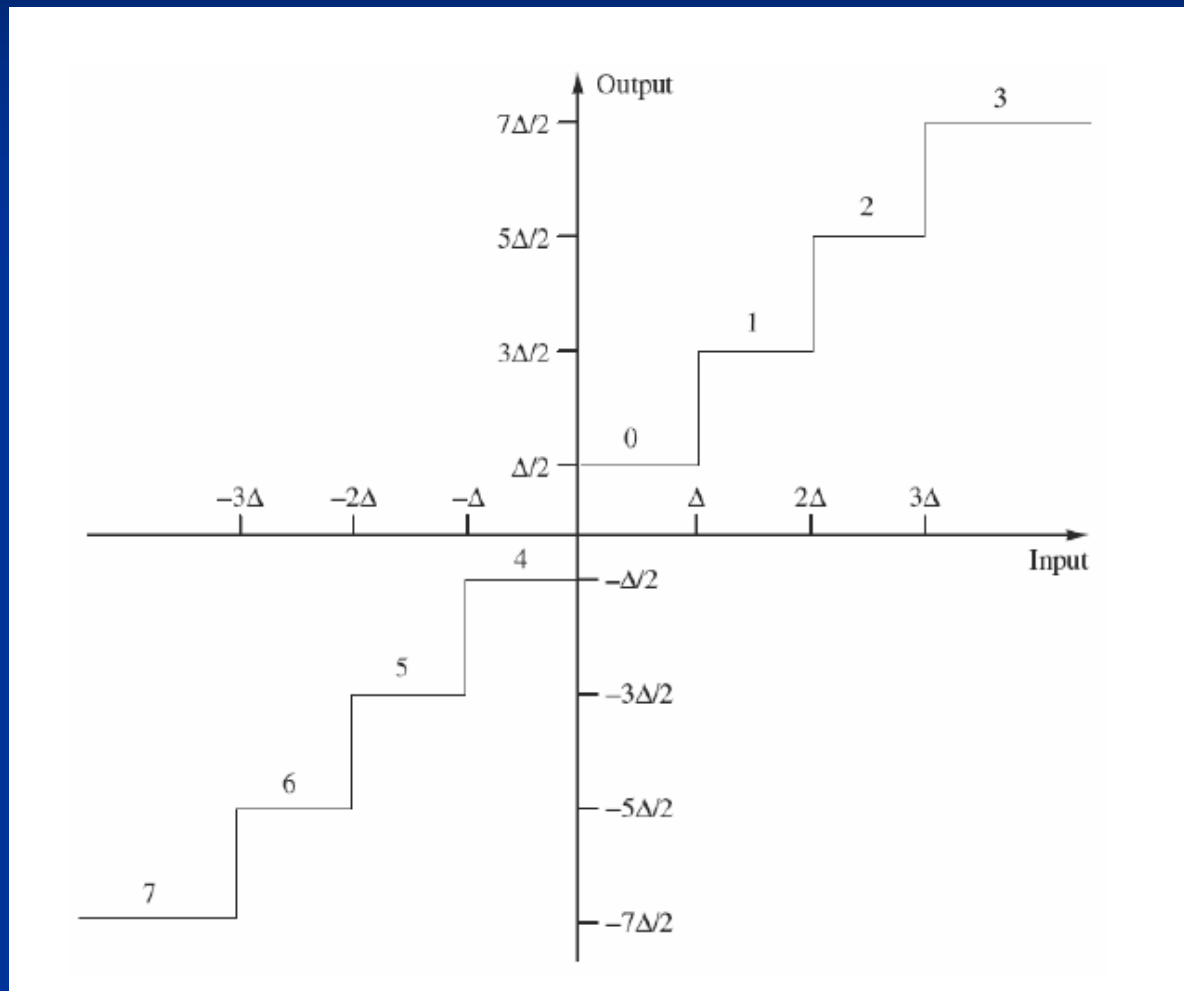
- 如果输入落入外边的输出水平
  - 增大量化步长
- 如果输入落入里边的输出水平
  - 缩短量化步长
- 如果匹配，则增大值和缩短值的乘积应为1

## ■ 乘子： $M_k$

- 如果  $S_{n-1}$  输入落入第 $k$ 个区间，则步长乘以  $M_k$ 
$$\Delta_n = M_k \Delta_{n-1}$$
- 内部： $M_k < 1$ ， 外部： $M_k > 1$

前一个输出在编码器和解码器端的都可见，无需发送边信息

# 3比特Jayant 量化器输出水平



$$M_0 = M_4, M_1 = M_5, M_2 = M_6, M_3 = M_7$$

# 例：Jayant量化器

- $M_0 = M_4 = 0.8, M_1 = M_5 = 0.9$
- $M_2 = M_6 = 1.0, M_3 = M_7 = 1, 2, \Delta_0 = 0.5$
- 输入：0.1 -0.2 0.2 0.1 -0.3 0.1 0.2 0.5 0.9 1.5

$n$	$\Delta_n$	Input	Output Level	Output	Error	Update Equation
0	0.5	0.1	0	0.25	0.15	$\Delta_1 = M_0 \times \Delta_0$
1	0.4	-0.2	4	-0.2	0.0	$\Delta_2 = M_4 \times \Delta_1$
2	0.32	0.2	0	0.16	0.04	$\Delta_3 = M_0 \times \Delta_2$
3	0.256	0.1	0	0.128	0.028	$\Delta_4 = M_0 \times \Delta_3$
4	0.2048	-0.3	5	-0.3072	-0.0072	$\Delta_5 = M_5 \times \Delta_4$
5	0.1843	0.1	0	0.0922	-0.0078	$\Delta_6 = M_0 \times \Delta_5$
6	0.1475	0.2	1	0.2212	0.0212	$\Delta_7 = M_1 \times \Delta_6$
7	0.1328	0.5	3	0.4646	-0.0354	$\Delta_8 = M_3 \times \Delta_7$
8	0.1594	0.9	3	0.5578	-0.3422	$\Delta_9 = M_3 \times \Delta_8$
9	0.1913	1.5	3	0.6696	-0.8304	$\Delta_{10} = M_3 \times \Delta_9$
10	0.2296	1.0	3	0.8036	0.1964	$\Delta_{11} = M_3 \times \Delta_{10}$
11	0.2755	0.9	3	0.9643	0.0643	$\Delta_{12} = M_3 \times \Delta_{11}$

当输入变化较慢时，量化器可与输入很好适应：语音、图像

# 选择Jayant乘子

- 防止上溢/下溢:  $\Delta_{\min} / \Delta_{\max}$
- 乘子的选择对适应性影响大: 乘子离1越远, 越适应, 但太适应会导致不稳定

平稳性准则:

当量化器与输入匹配, 增大和缩短的乘积为1

$$\prod_{k=0}^M M_k^{n_k} = 1$$

$$\prod_{k=0}^M M_k^{n_k/N} = 1$$

$n_k$  是落入第k个区间的次数  
 $N$  是输入总数

$$\prod_{k=0}^M M_k^{p_k} = 1, \text{ 其中 } p_k = n_k / N$$

令  $M_k$  具有某种结构

$l_k$  取整数值

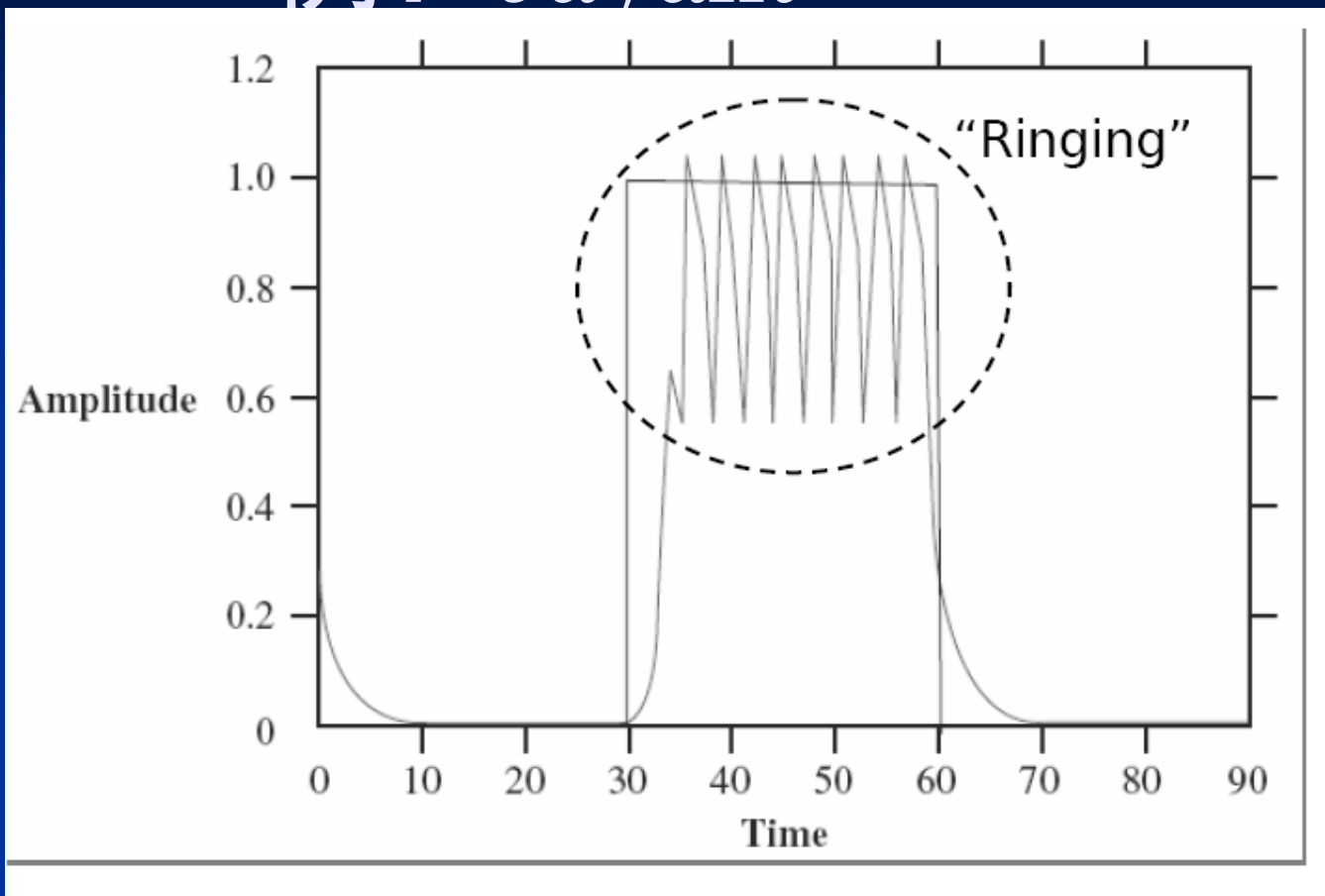
$$M_k = \gamma^{l_k}, \text{ 其中 } \gamma^{l_k} > 1, l_k = \lfloor l_k \rfloor$$

$\gamma$  越大, 适应快;  $\gamma$  较小, 稳定性好

$$\prod_{k=0}^M \gamma^{l_k p_k} = 1 \Rightarrow \sum_{k=0}^M l_k p_k = 0$$

# 例：Jayant

2比特量化器  
输入：0、1  
 $P_0=0.8$ ,  $P_1=0.2$   
选择  $I_0=-1$ ,  $I_1=4$

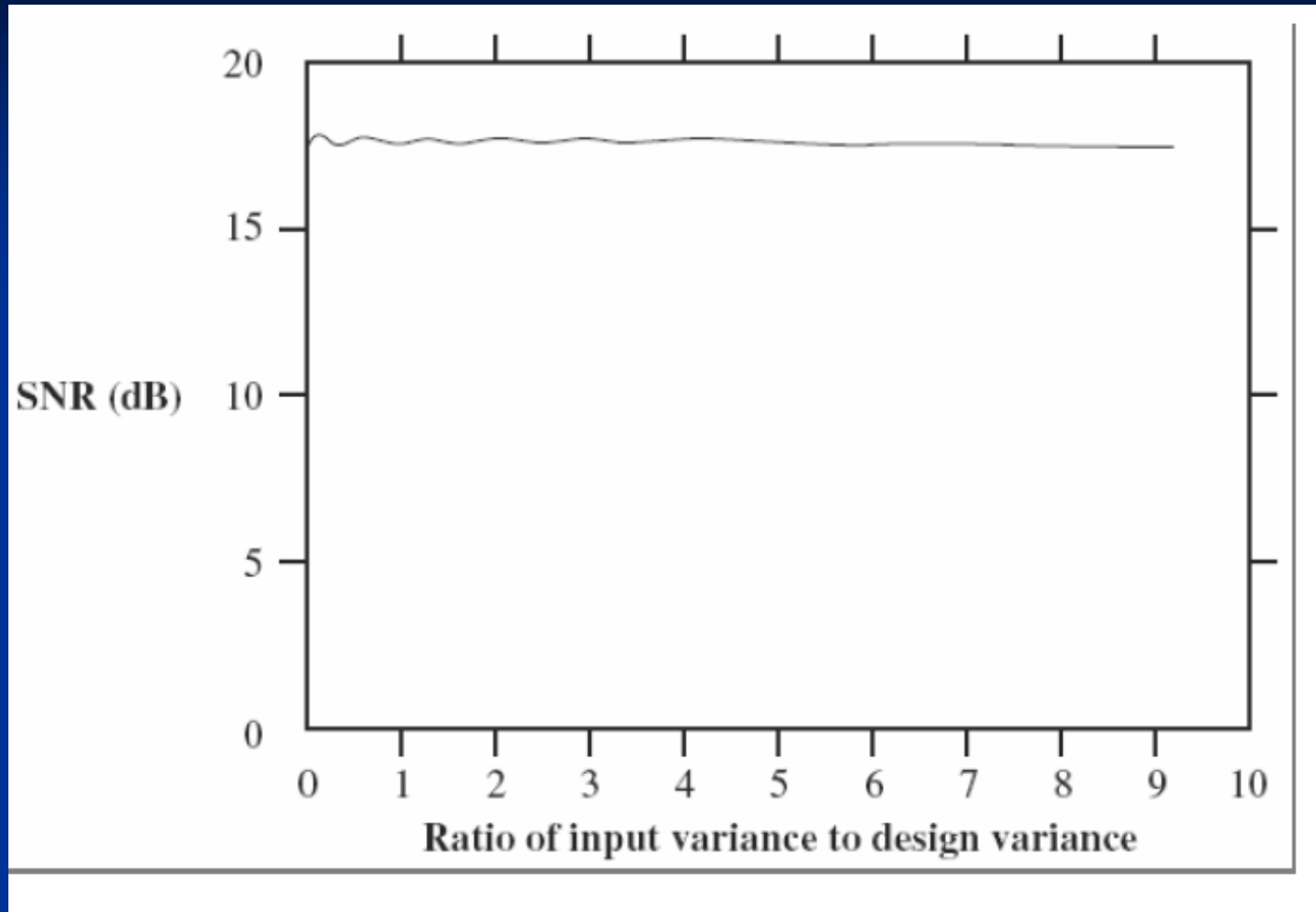


图中结果为 $\gamma$ 接近1

$\gamma$ 越大，适应快，振铃效应减弱

采用两种自适应调整策略：用于快速变化的情况；扩大速度要远快于缩小速度。因为过载误差是无界的，而颗粒噪声是有界的。  
用于输入不变或接近不变的情况 – G.726

# Jayant的性能



当输入方差与假定方差不匹配时，Jayant量化器的性能要远优于非自适应均匀量化器。但当满足匹配时，非自适应量化器的性能要明显优于Jayant量化器的性能。