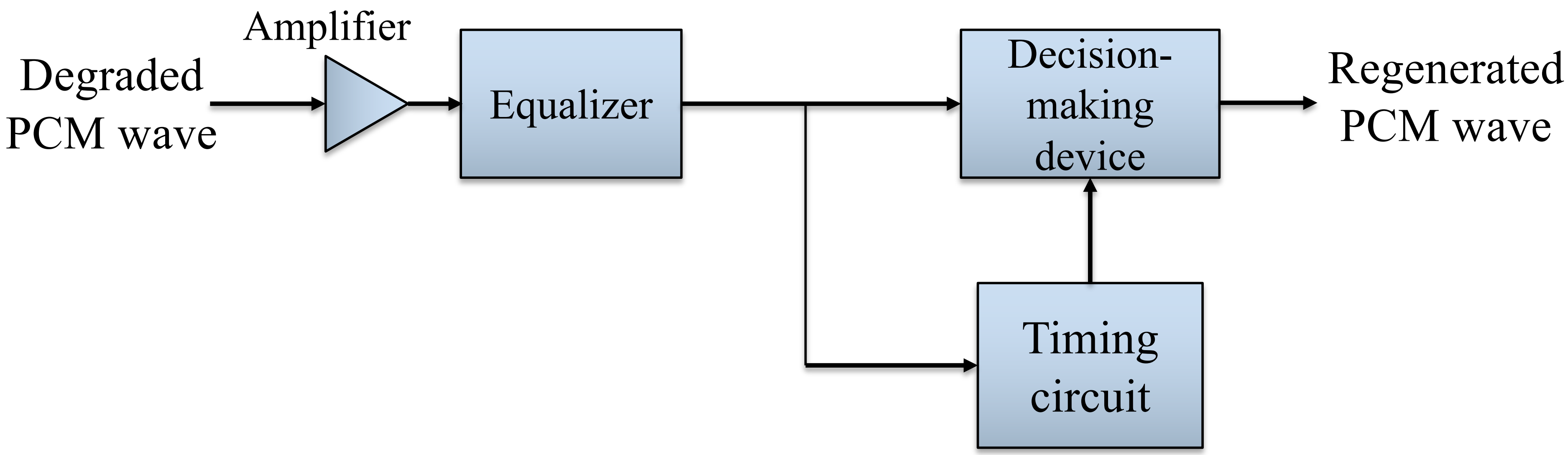
Part 7 Time-Division Multiplexing and Delta-Sigma Modulation

Regeneration

- Regenerative repeater for PCM system
  - It can completely remove the distortion if the decision making device makes the right decision (on 1 or 0).



Decoding & Filtering

- After regenerating the received pulse at the last stage, the receiver decodes and regenerates the original message signal (with acceptable quantization error).
- Finally, a lowpass reconstruction filter whose cutoff frequency is equal to the message bandwidth W is applied at the end (to remove the unnecessary high-frequency components due to “quantization”).

Impact of Noise in PCM Systems

- Two major noise sources in PCM systems
  - (Message-independent) Channel noise
  - (Message-dependent) Quantization noise
    - The quantization noise is often under designer's control, and can be made negligible by taking adequate number of quantization levels.

Impact of Noise in PCM Systems

- The main effect of channel noise is to introduce bit errors.
  - Notably, the *symbol error rate* is quite different from the *bit error rate*.
  - A symbol error may be caused by one-bit error, or two-bit error, or three-bit error, or …; so, in general, one cannot derive the *symbol error rate* from the *bit error rate* (or vice versa) unless some special assumption is made.
  - Considering the reconstruction of original analog signal, a bit error in the most significant bit is more harmful than a bit error in the least significant bit.

Error Threshold

- E_b/N_0
  - E_b: Transmitted signal energy per information bit
    - E.g., information bit is encoded using three-times repetition code, in which each code bit is transmitted using one $\pm\sqrt{E_c}$ symbol with symbol energy E_c.
    - Then $E_b = 3\ E_c$.
  - N_0: One-sided noise spectral density
- The bit error rate (BER) is a function of E_b/N_0 and transmission speed (and implicitly bandwidth, etc).

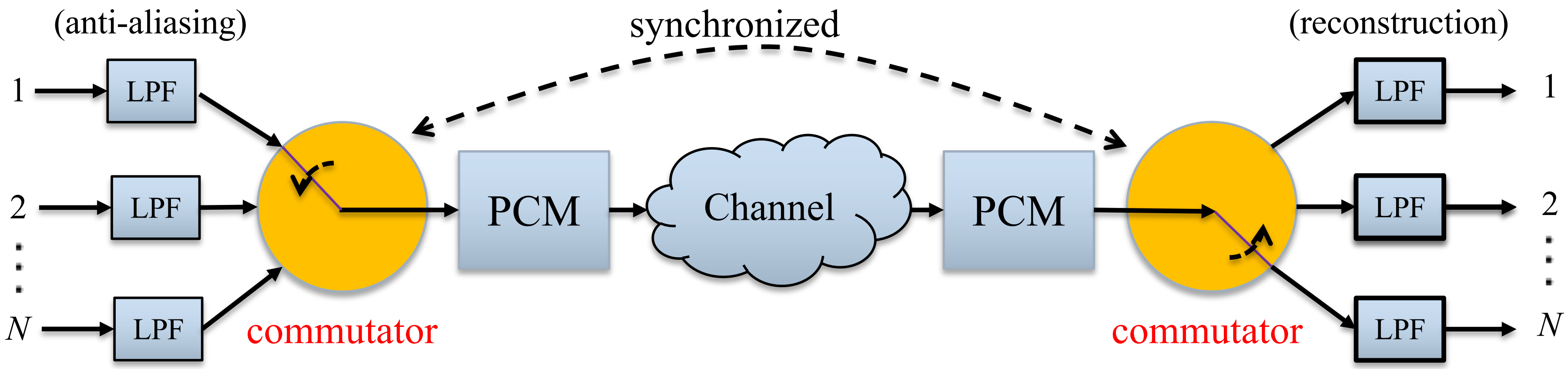
Error Threshold

- Error threshold
  - The minimum E_b/N_0 to achieve the required BER.
- By knowing the error threshold, one can always add a regenerative repeater when E_b/N_0 is about to drop below the threshold; hence, long-distance transmission becomes feasible.
  - Unlike the analog transmission, of which the distortion will accumulate for long-distance transmission.

Time-Division Multiplexing System

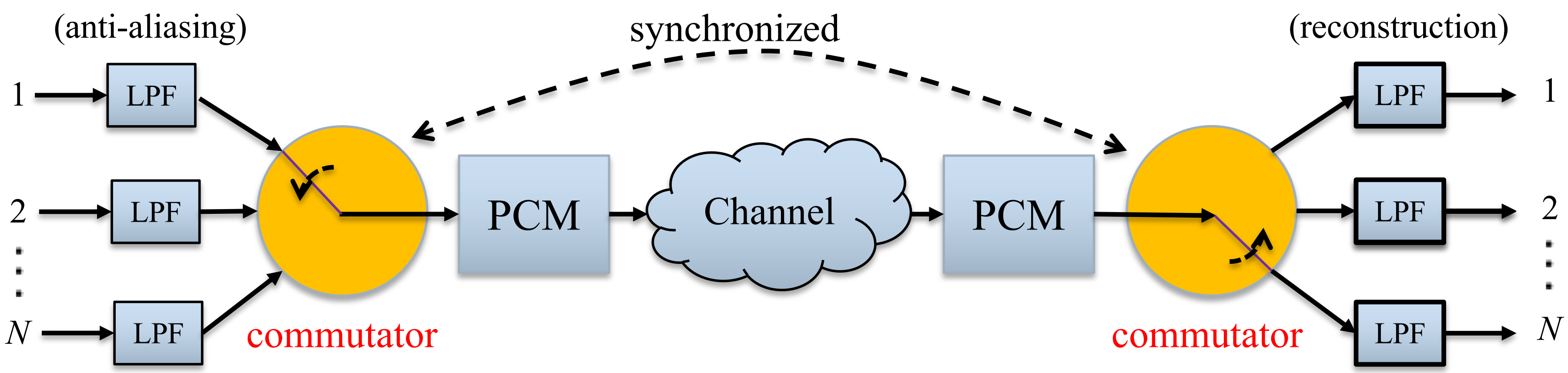
- An important feature of sampling process is a *conservation-of-time*.
  - In principle, the communication link is used only at the sampling time instances.
- Hence, it may be feasible to put other message's samples between adjacent samples of this message on a time-shared basis.
- This forms the time-division multiplex (TDM) system.
  - A joint utilization of a common communication link by a plurality of independent message sources.

Time-Division Multiplexing System



- The commutator (1) takes a narrow sample of each of the N input messages at a rate f_s slightly higher than $2W$, where W is the cutoff frequency of the anti-aliasing filter, and (2) interleaves these N samples inside the sampling interval T_s.

Time-Division Multiplexing System



- The price we pay for TDM is that N samples be squeezed in a time slot of duration T_s.

Time-Division Multiplexing System

- Synchronization is essential for a satisfactory operation of the TDM system.
  - One possible procedure to synchronize the transmitter clock and the receiver clock is to set aside a code element or pulse at the end of a frame, and to transmit this pulse every other frame only.

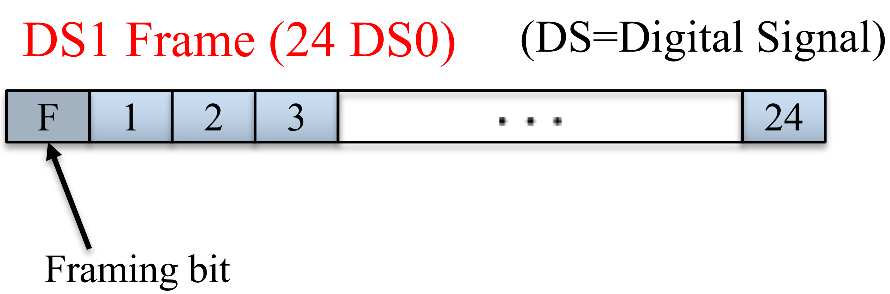
T1 System

- T1 system
  - Carries **24** 64kbps voice channels with regenerative repeaters spaced at approximately 2-km intervals.
  - Each voice signal is essentially limited to a band from 300 to 3100 Hz.
    - Anti-aliasing filter with $W = 3.1$ KHz
    - Sampling rate = 8 KHz ($> 2W = 6.2$ KHz)
  - ITU G.711 μ-law is used with $\mu = 255$.
  - Each frame consists of $24 \times 8 + 1 = 193$ bits, where a single bit is added at the end of the frame for the purpose of synchronization.

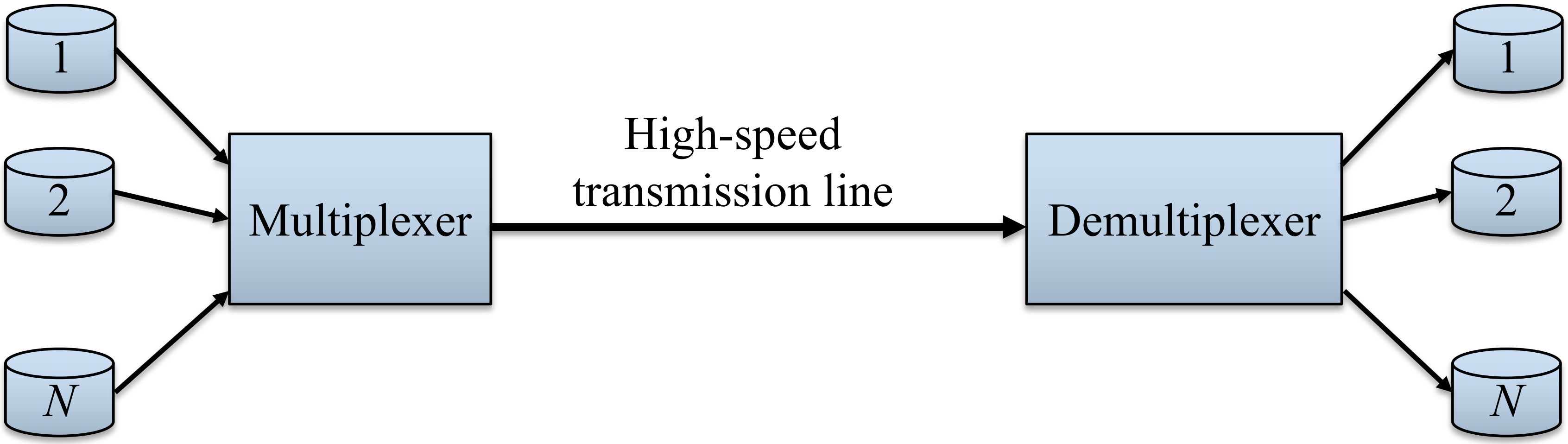
$$193\ \text{bit}/\text{frame} \times \frac{1}{1\ \text{sample (from each of 24 voice channels)}/\text{frame}} \times 8000\ \text{sample}/\text{sec} = 1.544\ \text{Megabits}/\text{sec}$$

T1 System

- In addition to the 193 bits per frame (i.e., 1.544 Megabits per second), a telephone system must also pass signaling information such as “dial pulses” and “on/off-hook.”
  - The least significant bit of each voice channel is deleted in every sixth frame, and a signaling bit is inserted in its place.



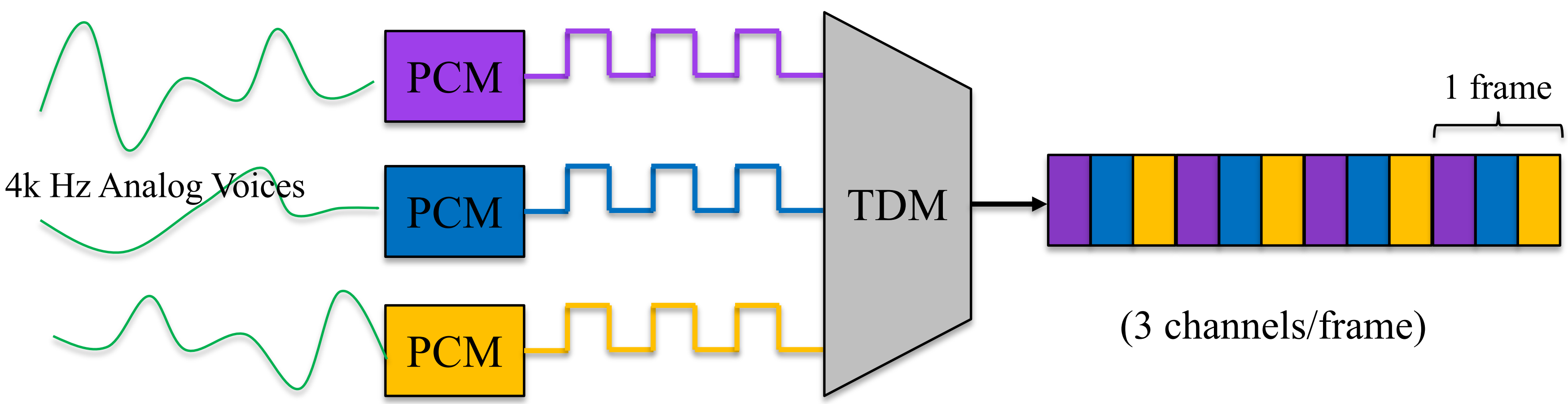
Digital Multiplexers



- The introduction of digital multiplexer enables us to combine digital signals of various natures, such as computer data, digitized voice signals, digitized facsimile and television signals.

Digital Multiplexers

- The multiplexing of digital signals is accomplished by using a *bit-by-bit interleaving* procedure with a selector switch that sequentially takes a (or more) bit from each incoming line and then applies it to the high-speed common line.



Digital Multiplexers

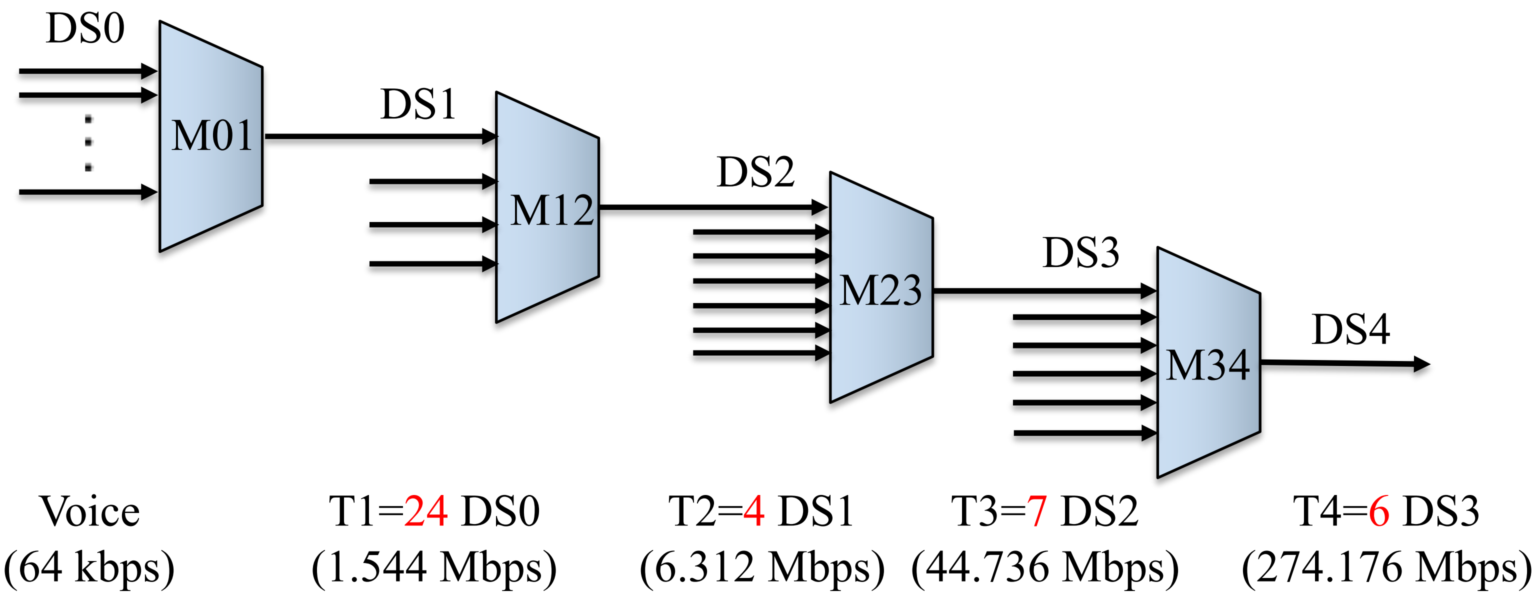
- Digital multiplexers are categorized into two major groups.
  1. 1st Group: Multiplex digital computer data for TDM transmission over public switched telephone network.
     - Require the use of modem technology.
  2. 2nd Group: Multiplex low-bit-rate digital voice data into high-bit-rate voice stream.
     - Accommodate in the hierarchy that is varying from one country to another.
     - Usually, the hierarchy starts at 64 Kbps, named a *digital signal zero* (DS0).

North American Digital TDM Hierarchy

- *The first level hierarchy*
  - Combine 24 DS0 to obtain a primary rate DS1 at 1.544 Mb/s (T1 transmission)
- *The second-level multiplexer*
  - Combine 4 DS1 to obtain a DS2 with rate 6.312 Mb/s
- *The third-level multiplexer*
  - Combine 7 DS2 to obtain a DS3 at 44.736 Mb/s
- *The fourth-level multiplexer*
  - Combine 6 DS3 to obtain a DS4 at 274.176 Mb/s
- *The fifth-level multiplexer*
  - Combine 2 DS4 to obtain a DS5 at 560.160 Mb/s

North American Digital TDM Hierarchy

- The combined bit rate is higher than the multiple of the incoming bit rates because of the addition of *bit stuffing* and *control signals*.



North American Digital TDM Hierarchy

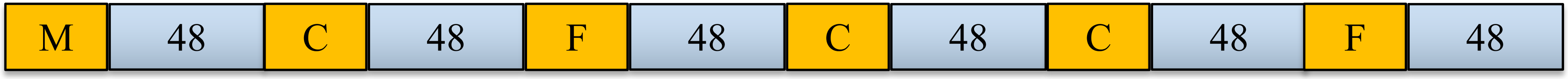
- Basic problems involved in the design of multiplexing system
  - Synchronization should be maintained to properly recover the interleaved digital signals.
  - Framing should be designed so that individual can be identified at the receiver.
  - Variation in the bit rates of incoming signals should be considered in the design.
    - A 0.01% variation in the propagation delay produced by a 1°F decrease in temperature will result in 100 fewer pulses in the cable of length 1000-km with each pulse occupying about 1 meter of the cable.

Digital Multiplexers

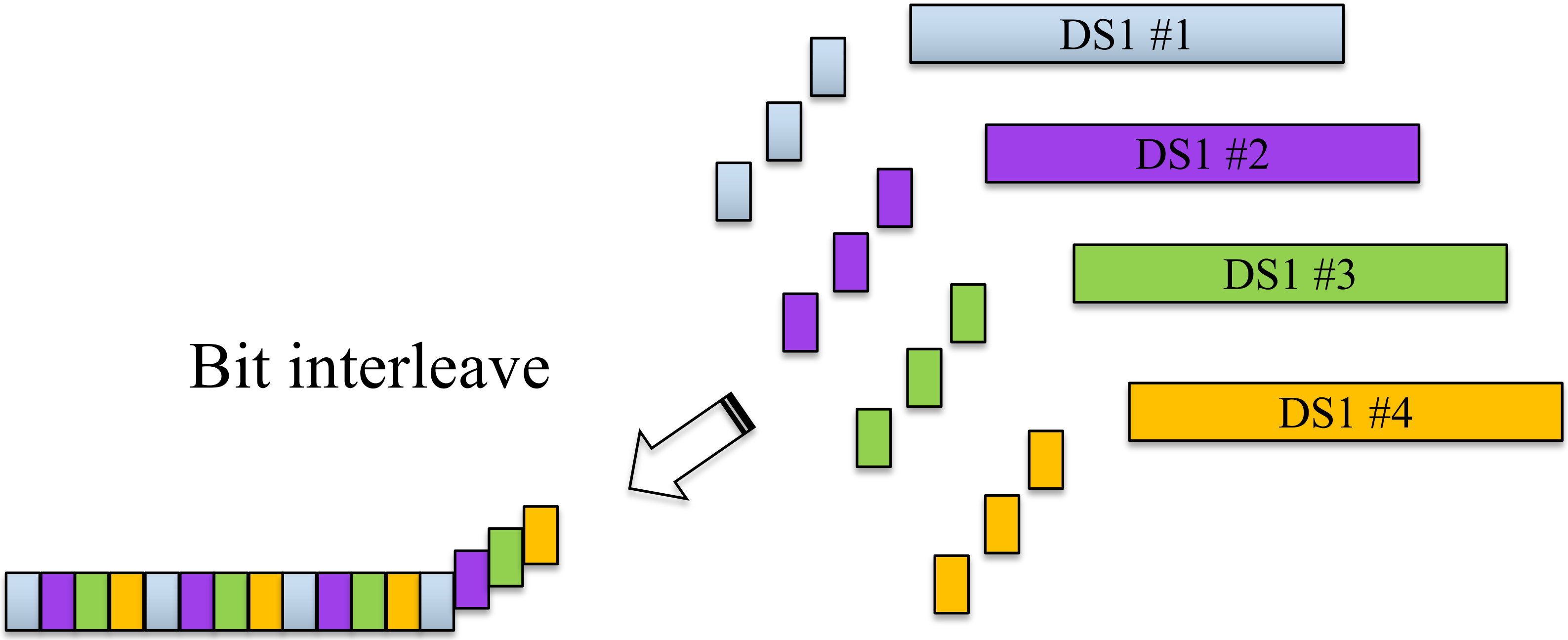
- Synchronization and rate variation problems are resolved by *bit stuffing*.
- AT&T M12 (second-level multiplexer)
  - 24 control bits are stuffed, and separated by sequences of 48 data bits (12 from each DS1 input).

M_0	48	C_I	48	F_0	48	C_I	48	C_I	48	F_1	48
M_1	48	C_{II}	48	F_0	48	C_{II}	48	C_{II}	48	F_1	48
M_1	48	C_{III}	48	F_0	48	C_{III}	48	C_{III}	48	F_1	48
M_1	48	C_{IV}	48	F_0	48	C_{IV}	48	C_{IV}	48	F_1	48
Subframe markers		Stuffing indicators		Frame markers		Stuffing indicators		Stuffing indicators		Frame markers	

DS2 M-Subframe: 294 bits (6 blocks)



For each block:



AT&T M12 Multiplexer

- The control bits are labeled F, M, and C.
  - Frame markers: In sequence of $F_0F_1F_0F_1F_0F_1F_0F_1$, where $F_0 = 0$ and $F_1 = 1$.
  - Subframe markers: In sequence of $M_0M_1M_1M_1$, where $M_0 = 0$ and $M_1 = 1$.
  - Stuffing indicators: In sequences of $C_\text{I}\, C_\text{I}\, C_\text{I}\, C_\text{II}\, C_\text{II}\, C_\text{II}\, C_\text{III}\, C_\text{III}\, C_\text{III}\, C_\text{IV}\, C_\text{IV}\, C_\text{IV}$, where all three bits of C_j equal 1's indicate that a stuffing bit is added in the position of the first information bit associated with the first DS1 bit stream that follows the F_1-control bit in the same subframe, and three 0's in $C_jC_jC_j$ imply no stuffing.
    - The receiver should use majority law to check whether a stuffing bit is added.

AT&T M12 Multiplexer

- These stuffed bits can be used to balance (or maintain) the nominal input bit rates and nominal output bit rates.
  - S = nominal bit stuffing rate
    - The rate at which stuffing bits are inserted when both the input and output bit rates are at their nominal values.
  - f_{in} = nominal input bit rate
  - f_{out} = nominal output bit rate
  - M = number of bits in a frame
  - L = number of information bits (input bits) for one input stream in a frame

$$\frac{L-1}{f_{in}} = 185.88082902\ \mu s \le \frac{M}{f_{out}} = 186.31178707\ \mu s \le \frac{L}{f_{in}} = 186.52849741\ \mu s$$

AT&T M12 multiplexer

- For M12 framing,

$$f_{in} = 1.544 \text{ Mbps}$$

$$f_{out} = 6.312 \text{ Mbps}$$

$$M = 288 \times 4 + 24 = 1176 \text{ bits}$$

$$L = 288 \text{ bits}$$

$$\frac{M}{f_{out}} = S\frac{4(L-1)}{4f_{in}} + (1-S)\frac{4L}{4f_{in}}$$

$$\text{Duration of a frame} = \frac{M}{f_{out}} = \underbrace{S\frac{L-1}{f_{in}}}_{\text{One bit is replaced by a stuffed bit.}} + (1-S)\frac{L}{f_{in}}$$

$$\Rightarrow S = L - \frac{f_{in}}{f_{out}}M = 288 - \frac{1.544}{6.312}1176 = 0.334601$$

AT&T M12 Multiplexer

- Allowable tolerances to maintain nominal output bit rates
  - A sufficient condition for the existence of S such that the nominal output bit rate can be matched.

$$\max_{S\in[0,1]}\left[S\frac{L-1}{f_{in}}+(1-S)\frac{L}{f_{in}}\right]\geq\frac{M}{f_{out}}\geq\min_{S\in[0,1]}\left[S\frac{L-1}{f_{in}}+(1-S)\frac{L}{f_{in}}\right]$$

$$\Leftrightarrow\frac{L}{f_{in}}\geq\frac{M}{f_{out}}\geq\frac{L-1}{f_{in}}\Leftrightarrow\frac{L}{M}f_{out}\geq f_{in}\geq\frac{L-1}{M}f_{out}$$

$$\boxed{1.5458=\frac{288}{1176}6.312\geq f_{in}\geq\frac{287}{1176}6.312=1.54043}$$

AT&T M12 Multiplexer

- This results in an allowable tolerance range:

$$1.5458-1.54043=6.312/1176=5.36735\text{ kbps}$$

- In terms of ppm (pulse per million pulses),

$$\frac{10^6-b_{ppm}}{1.54043}=\frac{10^6}{1.544}=\frac{10^6+a_{ppm}}{1.5458}$$

$$\Rightarrow a_{ppm}=1164.8 \text{ and } b_{ppm}=2312.18$$

  - This tolerance is already much larger than the expected change in the bit rate of the incoming DS1 bit stream.

Summary of PCM

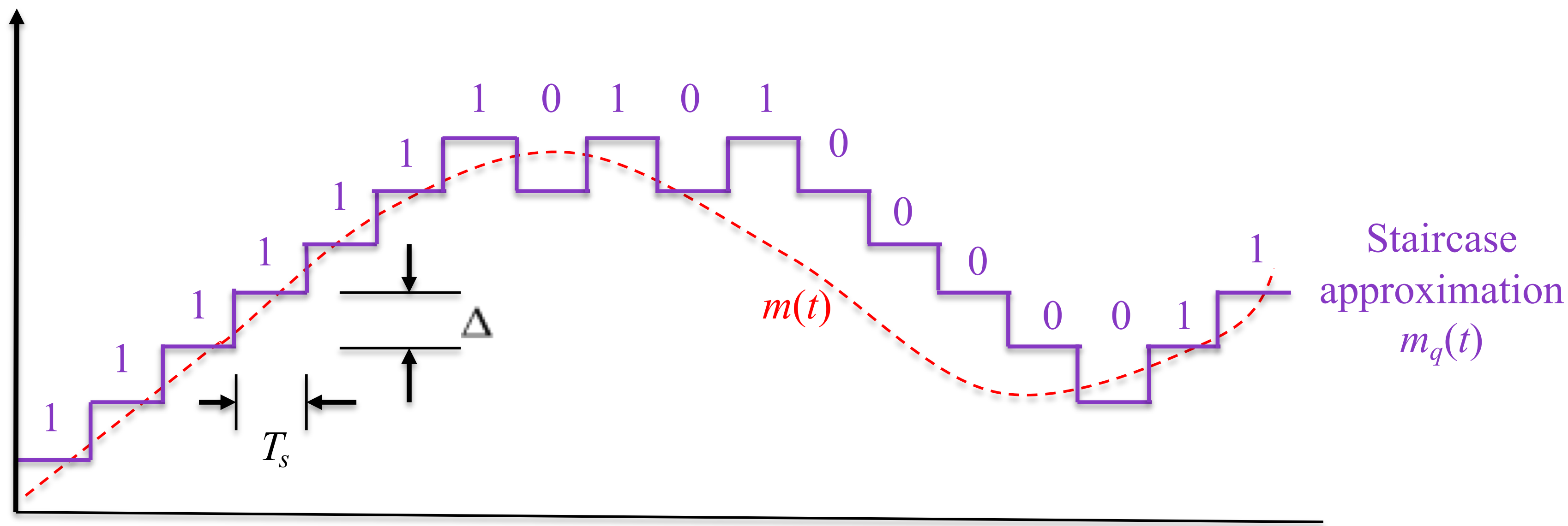
- Virtues of PCM systems
  - Robustness to channel noise and interference
  - Efficient regeneration of coded signal along the transmission path
  - Efficient exchange of increased channel bandwidth for improved signal-to-noise ratio, obeying an exponential law.
  - Uniform format for different kinds of baseband signal transmission; hence, facilitate their integration in a common network.
  - Message sources are easily dropped or reinserted in a TDM system.
  - Secure communication through the use of encryption/decryption.

Summary of PCM

- Two limitations of PCM system (in the past)
  - Complexity
  - Bandwidth
- Nowadays, with the advance of VLSI technology, and with the availability of wideband communication channels (such as fiber) and compression technique (to reduce the bandwidth demand), the above two limitations are greatly released.

Delta Modulation

- Delta Modulation (DM)
  - The message is oversampled (at a rate much higher than the Nyquist rate) to purposely increase the correlation between adjacent samples.
  - Then, the difference between adjacent samples is encoded instead of the sample value itself.


1
1
1
1
1
1
1
1
0
1
0
1
0
0
0
0
0
1
1
Δ
T_s
$m(t)$
Staircase approximation $m_q(t)$

Math Analysis of Delta Modulation

Let $m[n] = m(nT_s)$.

Let $m_q[n]$ be the DM approximation of $m(t)$ at time nT_s.

Then

$$m_q[n] = m_q[n-1] + e_q[n] = \sum_{j=-\infty}^{n} e_q[j],$$

where $e_q[n] = \Delta \cdot \mathrm{sgn}(m[n] - m_q[n-1])$.

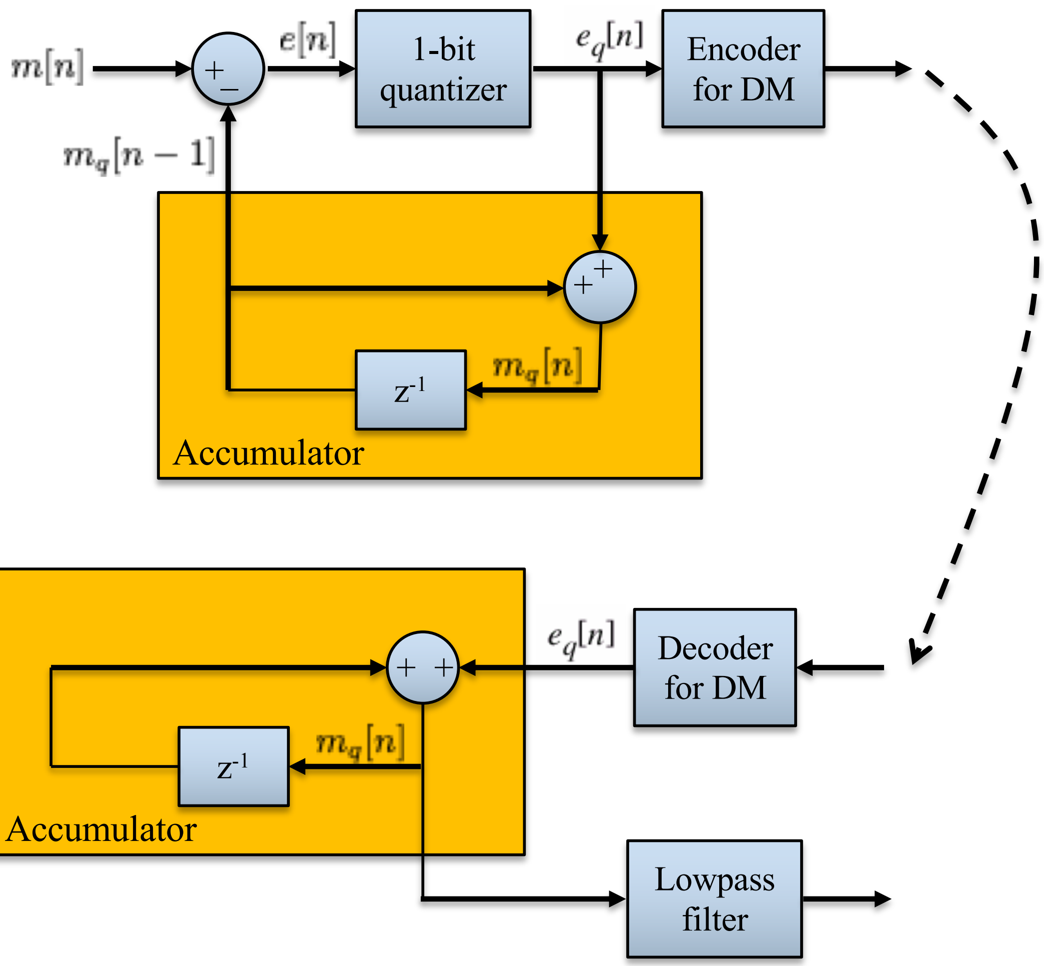
The transmitted code word is $\{[(e_q[n]/\Delta)+1]/2\}_{n=-\infty}^{\infty}$.

$$m_q[n] = m_q[n-1] + e_q[n] = \sum_{j=-\infty}^{n} e_q[n],$$

$$\text{where } e_q[n] = \Delta \cdot \operatorname{sgn}(m[n] - m_q[n-1]).$$

Delta Modulation

- The principle virtue of delta modulation is its simplicity.
  - It only requires the use of *comparator*, *quantizer*, and *accumulator*.

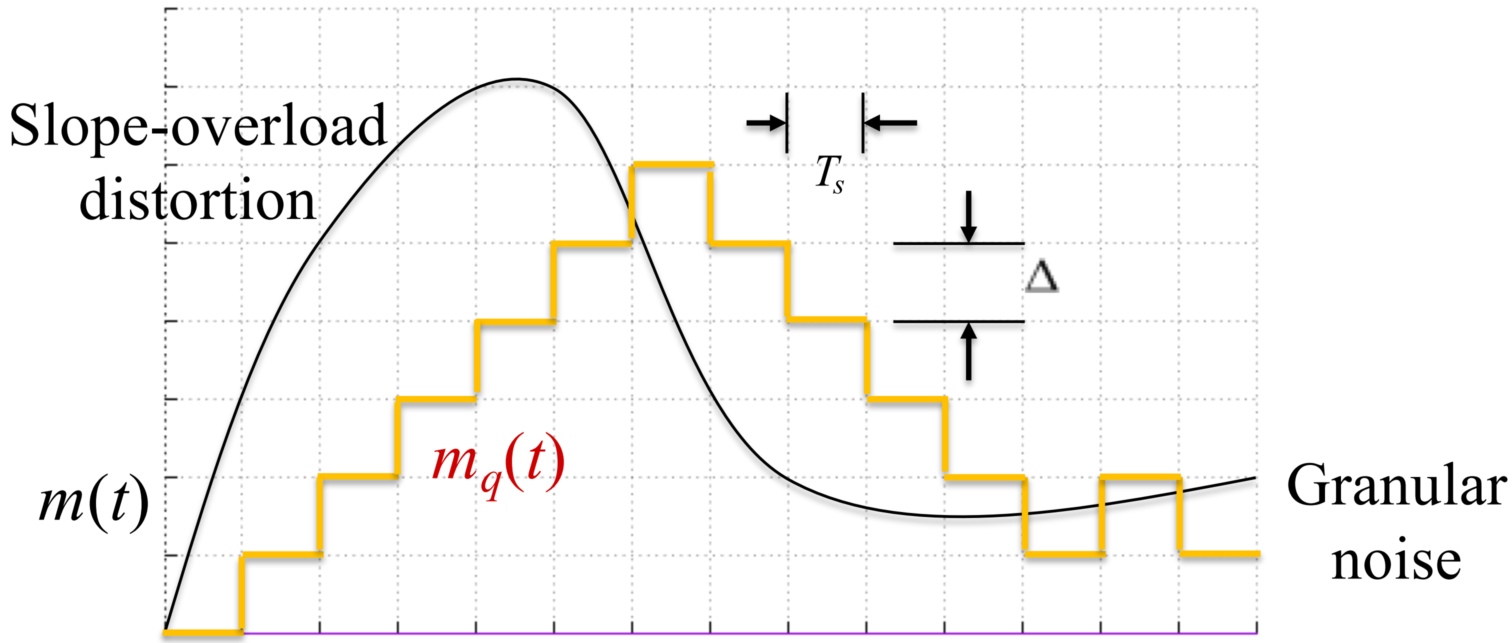


Bandwidth W of $m(t)$

$$\begin{cases} m[n] = m_q[n-1] + e[n] \\ m_q[n] = m_q[n-1] + e_q[n] \end{cases} \Rightarrow m_q[n] - m[n] = e_q[n] - e[n] \quad \text{(See Slide 7-53)}$$

Delta Modulation

- Distortions due to delta modulation
  - Slope overload distortion
  - Granular noise



Delta Modulation

- Slope overload distortion
  - To eliminate the slope overload distortion, it requires

$$\frac{\Delta}{T_s} \geq \max\left|\frac{dm(t)}{dt}\right| \quad \text{(slope overload condition)}$$

  - So, increasing step size Δ can reduce the slope-overload distortion.
  - An alternative solution is to use *dynamic* Δ. (Often, a delta modulation with fixed step size is referred to as a *linear delta modulator* due to its fixed slope, a basic function of linearity.)

Delta Modulation

- Granular noise
  - $m_q[n]$ will hunt around a relatively flat segment of $m(t)$.
  - A remedy is to reduce the step size.

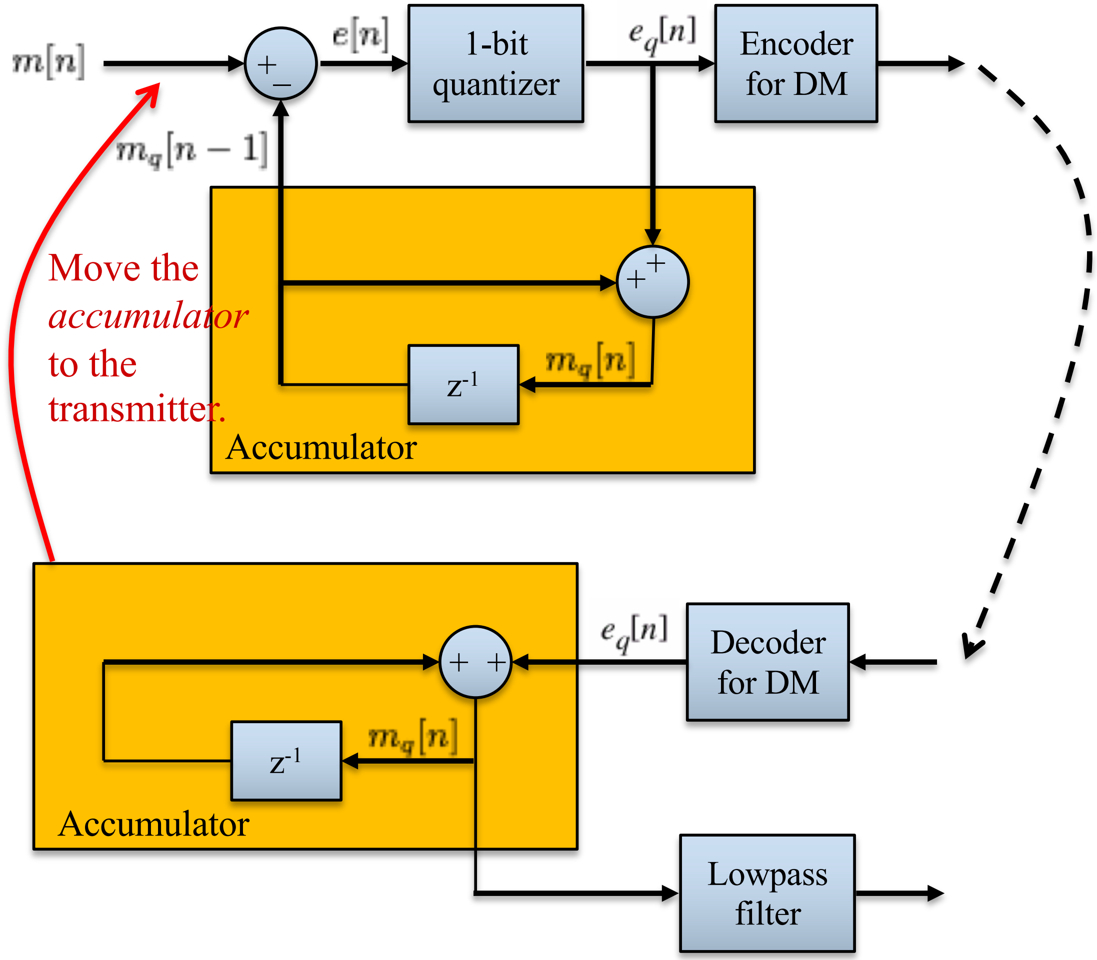
- A tradeoff in step size is therefore resulted for slope overload distortion and granular noise.

Delta-Sigma Modulation

- Delta-sigma modulation
  - In fact, the delta modulation distortion can be reduced by increasing the *correlation* between samples.
  - This can be achieved by integrating the message signal $m(t)$ prior to delta modulation.
  - The "integration" process is equivalent to a pre-emphasis of the low-frequency content of the input signal.

Delta-Sigma Modulation

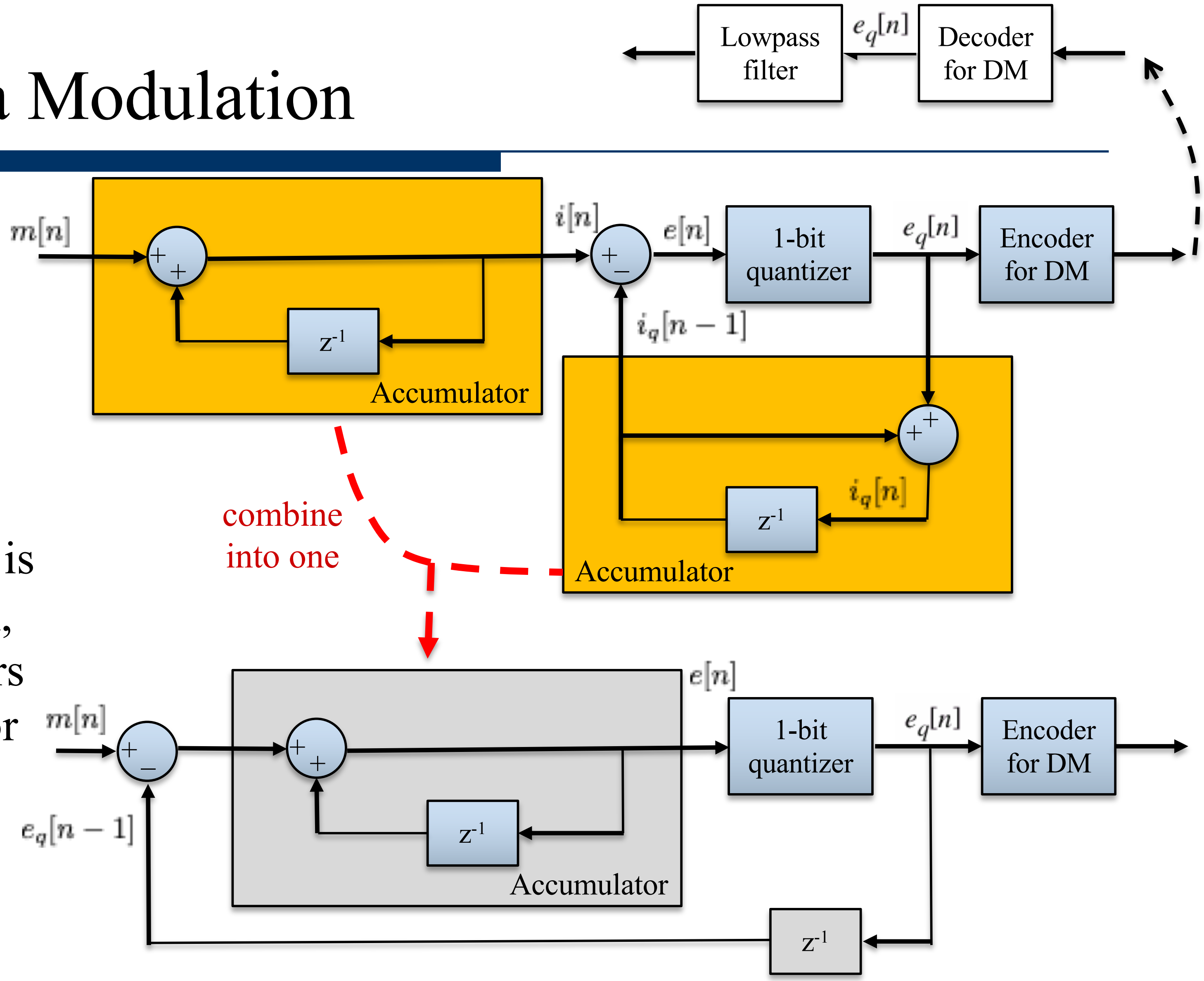
- A side benefit of "integration-before-delta-modulation," which is named *delta-sigma modulation*, is that the receiver design is further simplified (at the expense of a more complex transmitter).



Delta-Sigma Modulation

A straightforward structure

Since integration is a linear operation, the two integrators before comparator can be combined into one after comparator.



Math Analysis of Delta-Sigma Modulation

Let $i[n] = \sum_{i=-\infty}^{n} m[n] \approx \int_{-\infty}^{nT_s} m(t)dt.$

Let $i_q[n]$ be the DM approximation of $i[n]$.

Then, $i_q[n] = i_q[n-1] + \sigma_q[n]$, where $\sigma_q[n] = \Delta \cdot \text{sgn}(i[n] - i_q[n-1])$.

The transmitted code word is $\{[(\sigma_q[n]/\Delta)+1]/2\}_{n=-\infty}^{\infty}$.

Since

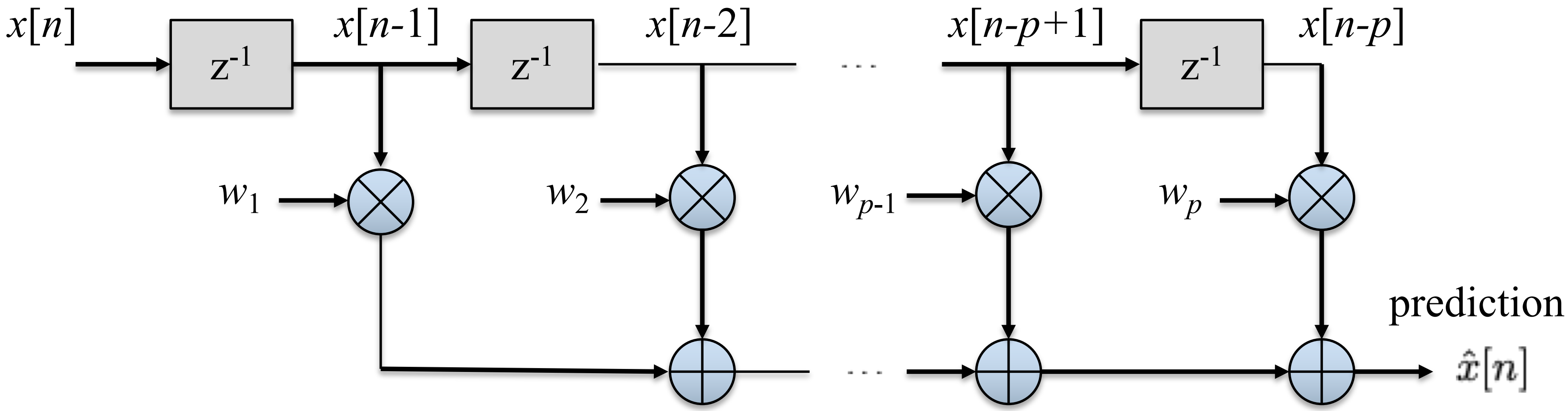
$$\sigma_q[n] = i_q[n] - i_q[n-1] \approx i[n] - i[n-1] = m[n],$$

we only need a lowpass filter to smooth out the received signal at the receiver end. (See the previous slide.)

Delta-Sigma Modulation

- Final notes
  - Delta(-sigma) modulation trades channel bandwidth (e.g., much higher sampling rate) for reduced system complexity (e.g., the receiver only demands a lowpass filter).
  - Can we trade increased system complexity for a reduced channel bandwidth? Yes, by means of *prediction technique*.
  - In Section 3.13, we will introduce the basics of prediction technique. Its application will be addressed in subsequent sections.

Linear Prediction



- Consider a finite-duration impulse response (FIR) discrete-time filter, where p is the prediction order, with linear prediction

$$\hat{x}[n] = \sum_{k=1}^{p} w_k x[n-k]$$

Linear Prediction

- Design objective
  - To find the filter coefficient w_1, w_2, …, w_p so as to minimize *index of performance J*:

$$J = E[e^2[n]], \text{ where } e[n] = x[n] - \hat{x}[n].$$

Let $\{x[n]\}$ be statinoary with autocorrelation function $R_X[k]$.

$$J = E\left[\left(x[n] - \sum_{k=1}^{p} w_k x[n-k]\right)^2\right]$$

$$= E[x^2[n]] - 2\sum_{k=1}^{p} w_k E[x[n]x[n-k]] + \sum_{k=1}^{p}\sum_{j=1}^{p} w_k w_j E[x[n-k]x[n-j]]$$

$$= R_X[0] - 2\sum_{k=1}^{p} w_k R_X[k] + \left(2\sum_{k=1}^{p}\sum_{j>k}^{p} w_k w_j R_X[k-j] + \sum_{k=1}^{p} w_k^2 R_X[0]\right)$$

$$\frac{\partial}{\partial w_i} J = -2R_X[i] + \left(2\sum_{j=i+1}^{p} w_j R_X[i-j] + 2\sum_{k=1}^{i-1} w_k R_X[k-i] + 2w_i R_X[0]\right)$$

$R_X[k-i] = R_X[i-k]$

$$= -2R_X[i] + 2\sum_{j=1}^{p} w_j R_X[i-j] = 0$$

$$\sum_{j=1}^{p} w_j R_X[i-j] = R_X[i] \text{ for } 1 \le i \le p.$$

The above optimality equations are called the *Wiener-Hopf equations* for linear prediction.
It can be rewritten in matrix form as:

$$\begin{bmatrix} R_X[0] & R_X[1] & \cdots & R_X[p-1] \\ R_X[1] & R_X[0] & \cdots & R_X[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_X[p-1] & R_X[p-2] & \cdots & R_X[0] \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} R_X[1] \\ R_X[2] \\ \vdots \\ R_X[p] \end{bmatrix}$$

or $\mathbf{R}_X \mathbf{w} = \mathbf{r}_X \Rightarrow$ Optimal solution $\mathbf{w}_o = \mathbf{R}_X^{-1} \mathbf{r}_X$

Toeplitz (Square) Matrix

- Any square matrix of the form

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{p-1} \\ a_1 & a_0 & \cdots & a_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p-1} & a_{p-2} & \cdots & a_0 \end{bmatrix}_{p\times p}$$

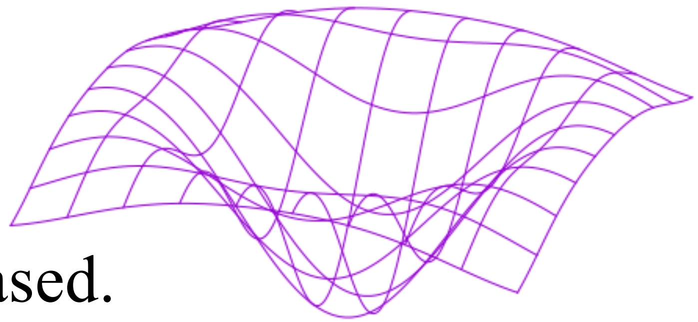
  is said to be *Toeplitz*.
- A Toeplitz matrix (such as $\mathbf{R}_X$) can be uniquely determined by p elements, $[a_0, a_1, \ldots, a_{p-1}]$.

Linear Adaptive Predictor

- The optimal $\mathbf{w}_0$ can only be obtained with the knowledge of autocorrelation function.
- *Question*: What if the autocorrelation function is unknown?
- *Answer*: Use *linear adaptive predictor*.

Idea Behind Linear Adaptive Predictor

- To minimize J, we should update w_i toward the bottom of the J-bowel.



$$g_i \equiv \frac{\partial J}{\partial w_i}$$

  - So, when $g_i > 0$, w_i should be decreased.
  - On the contrary, w_i should be increased if $g_i < 0$.
  - Hence, we may define the update rule as:

$$\hat{w}_i[n+1] = \hat{w}_i[n] - \frac{1}{2}\mu \cdot g_i[n]$$

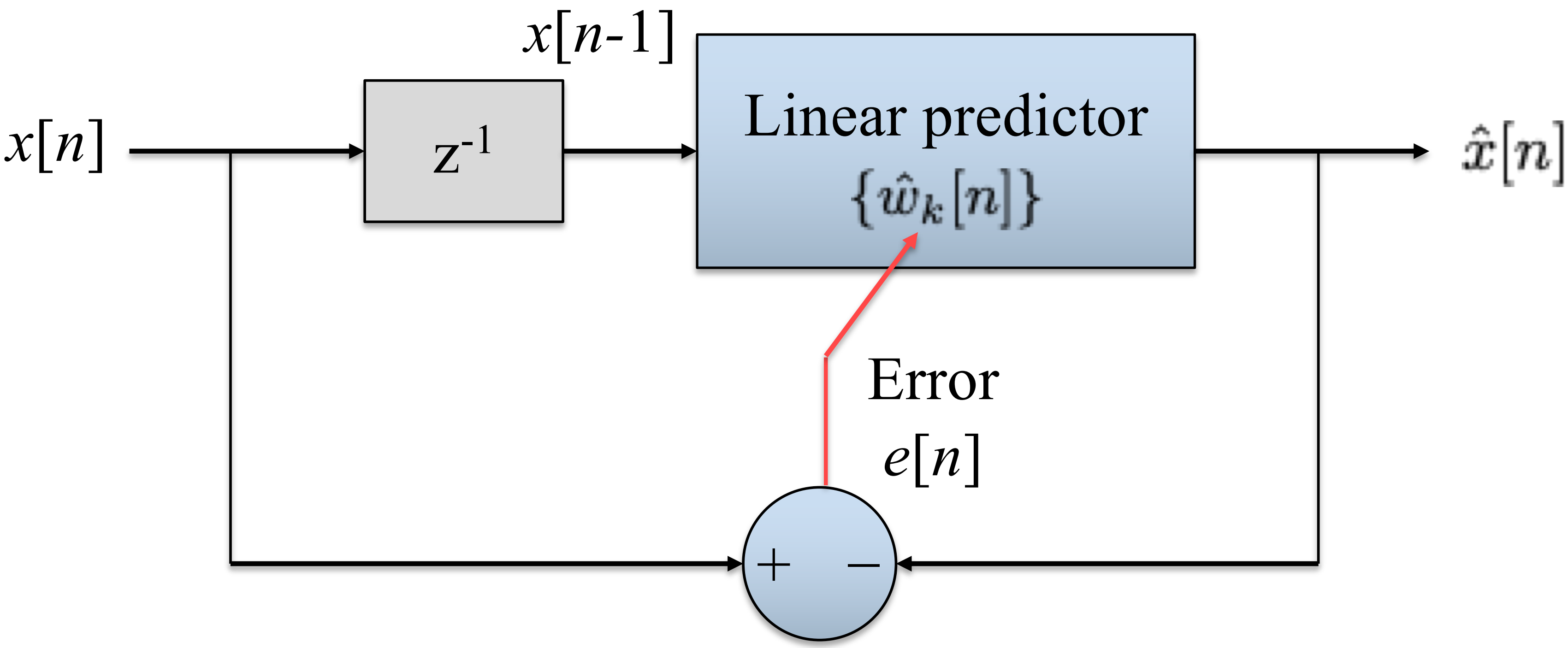
where μ is a chosen constant step size, and ½ is included only for convenience of analysis.

- $g_i[n]$ can be approximated by:

$$\begin{aligned} g_i[n] &\equiv \partial J/\partial w_i = -2R_X[i] + 2\sum_{j=1}^{p} w_j\, R_X[i-j] \\ &\approx -2\,x[n]\,x[n-i] + 2\sum_{j=1}^{p} \hat{w}_j[n]\,x[n-j]\,x[n-i] \\ &= 2\,x[n-i]\left(-x[n] + \sum_{j=1}^{p} \hat{w}_j[n]\,x[n-j]\right) \end{aligned}$$

$$\begin{aligned} \Rightarrow \hat{w}_i[n+1] &= \hat{w}_i[n] + \mu \cdot x[n-i]\left(x[n] - \sum_{j=1}^{p} \hat{w}_j[n]x[n-j]\right) \\ &= \hat{w}_i[n] + \mu \cdot x[n-i]e[n] \end{aligned}$$

Structure of Linear Adaptive Predictor



Least Mean Square Algorithm

- The pairs in the form of the popular least-mean-square (LMS) algorithm for linear adaptive prediction

$$\begin{cases} \hat{w}_j[n+1] = \hat{w}_j[n] + \mu \cdot x[n-j]e[n] \\ e[n] = x[n] - \sum_{j=1}^{p} \hat{w}_j[n]x[n-j] \end{cases}$$

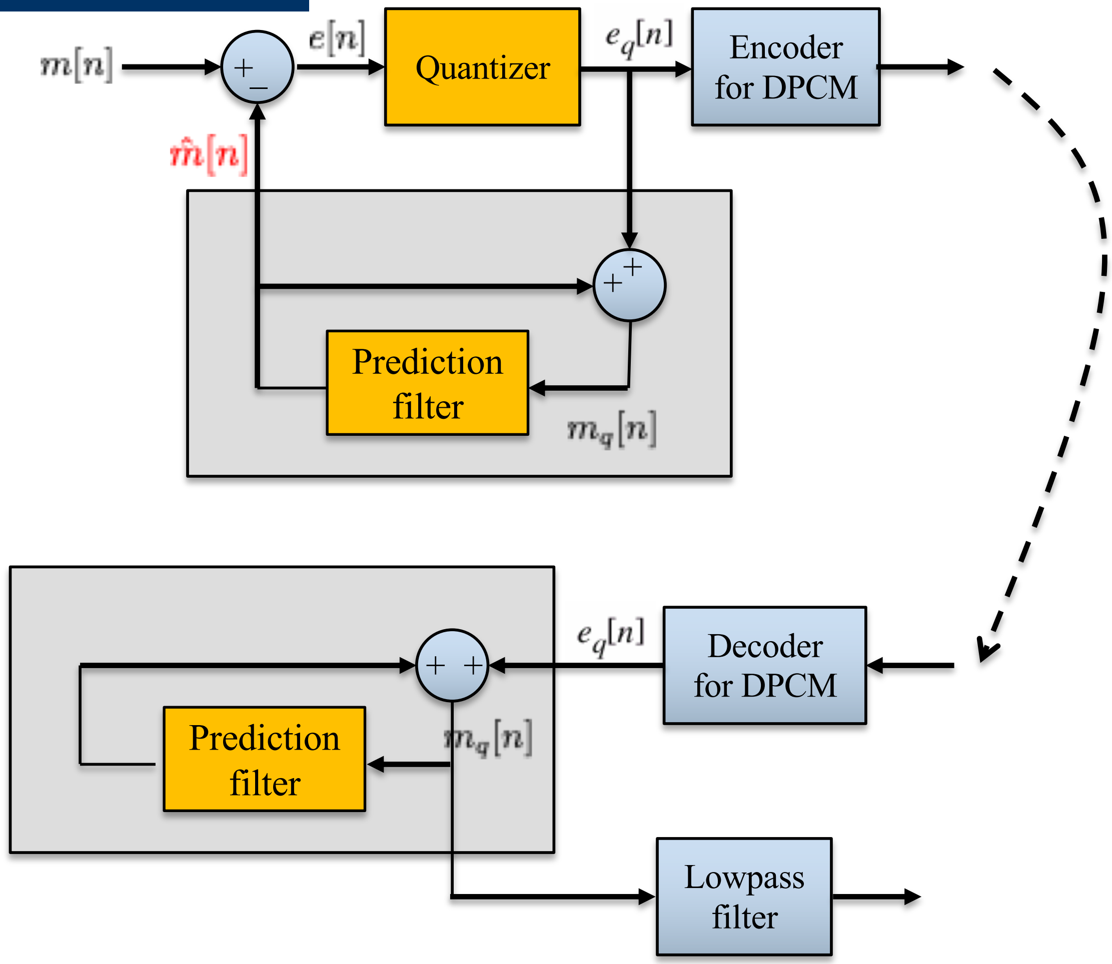
Differential Pulse-Code Modulation (DPCM)

- Basic idea behind *differential pulse-code modulati*on
  - Adjacent samples are often found to exhibit a high degree of correlation.
  - If we can remove this adjacent redundancy before encoding, a more efficient coded signal can be resulted.
  - A way to remove the redundancy is to use *linear prediction*.

DPCM

$$e[n] = m[n] - \hat{m}[n]$$
$$q[n] = e_q[n] - e[n]$$

- For DPCM, the quantization error is on $e[n]$, rather on $m[n]$ as for PCM.
- So, the quantization error $q[n]$ is supposed to be smaller.



Bandwidth W of $m(t)$

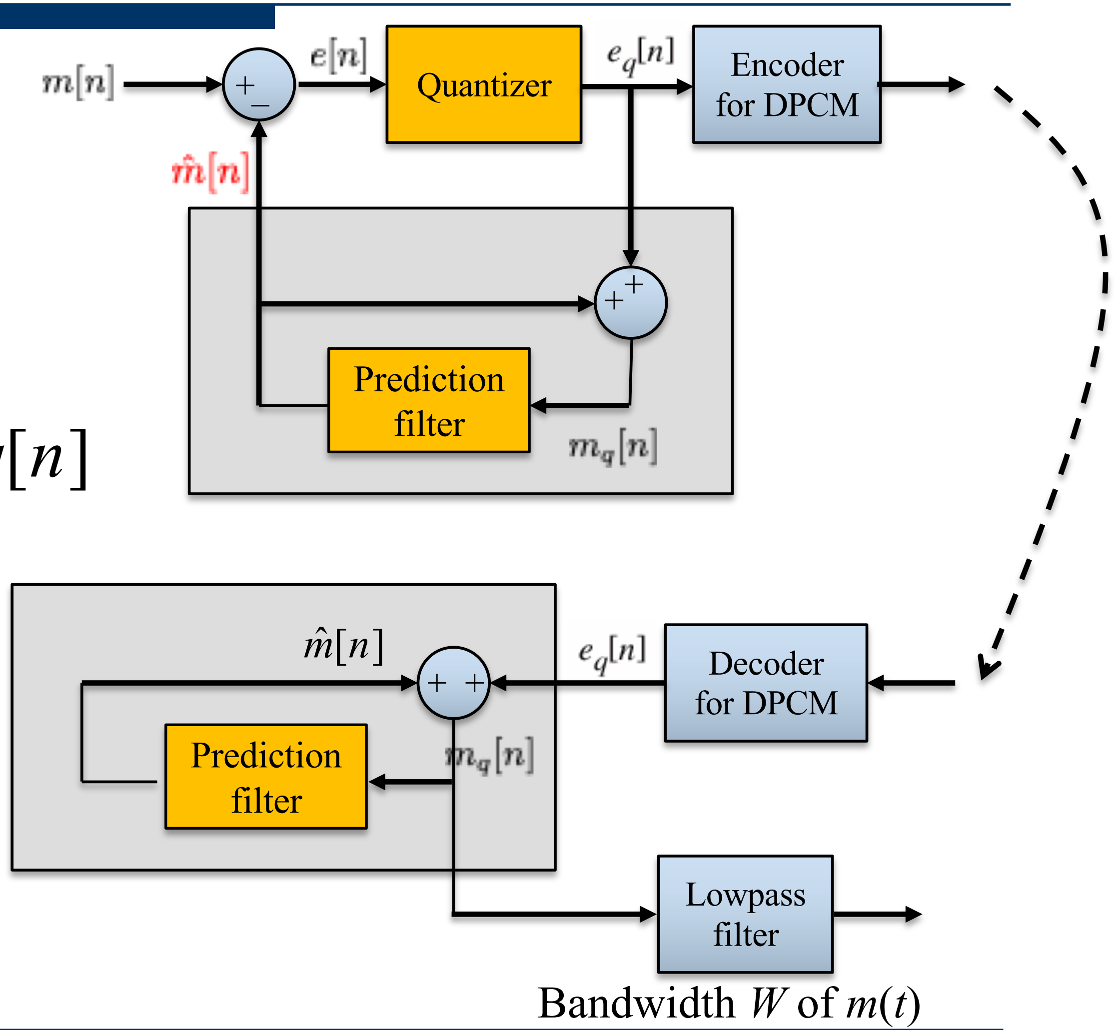
DPCM

- Derive:

$$e_q[n] = e[n] + q[n]$$

$$\Rightarrow \quad m_q[n] = \hat{m}[n] + e_q[n]$$
$$= \hat{m}[n] + e[n] + q[n]$$
$$= m[n] + q[n]$$

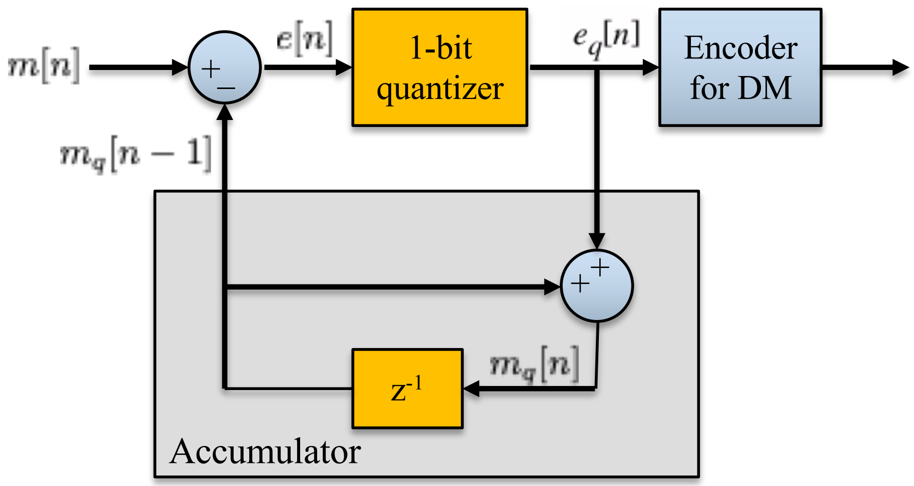
So, we have the same relation between $m_q[n]$ and $m[n]$ (as in Slide 7-32) but with smaller $q[n]$.



DPCM

- Notes
  - DM system can be treated as a special case of DPCM.

Prediction filter => single delay
Quantizer => single-bit



DPCM

- Distortions due to DPCM
  - Slope overload distortion
    - The input signal changes too rapidly for the prediction filter to track it.
  - Granular noise

Processing Gain

- The DPCM system can be described by:

$$m_q[n] = m[n] + q[n]$$

- So, the output signal-to-noise ratio is:

$$SNR_O = \frac{E[m^2[n]]}{E[q^2[n]]}$$

- We can re-write SNR_O as:

$$SNR_O = \frac{E[m^2[n]]}{E[e^2[n]]} \frac{E[e^2[n]]}{E[q^2[n]]} = G_p \cdot SNR_Q$$

where $e[n] = m[n] - \hat{m}[n]$ is the prediction error.

Processing Gain

- In terminologies,

$$\begin{cases} G_p = \dfrac{E[m^2[n]]}{E[e^2[n]]} & \text{processing gain} \\ SNR_Q = \dfrac{E[e^2[n]]}{E[q^2[n]]} & \text{signal to quantization noise ratio} \end{cases}$$

Notably, SNR_Q can be treated as the SNR for system of $e_q[n] = e[n] + q[n]$.

Processing Gain

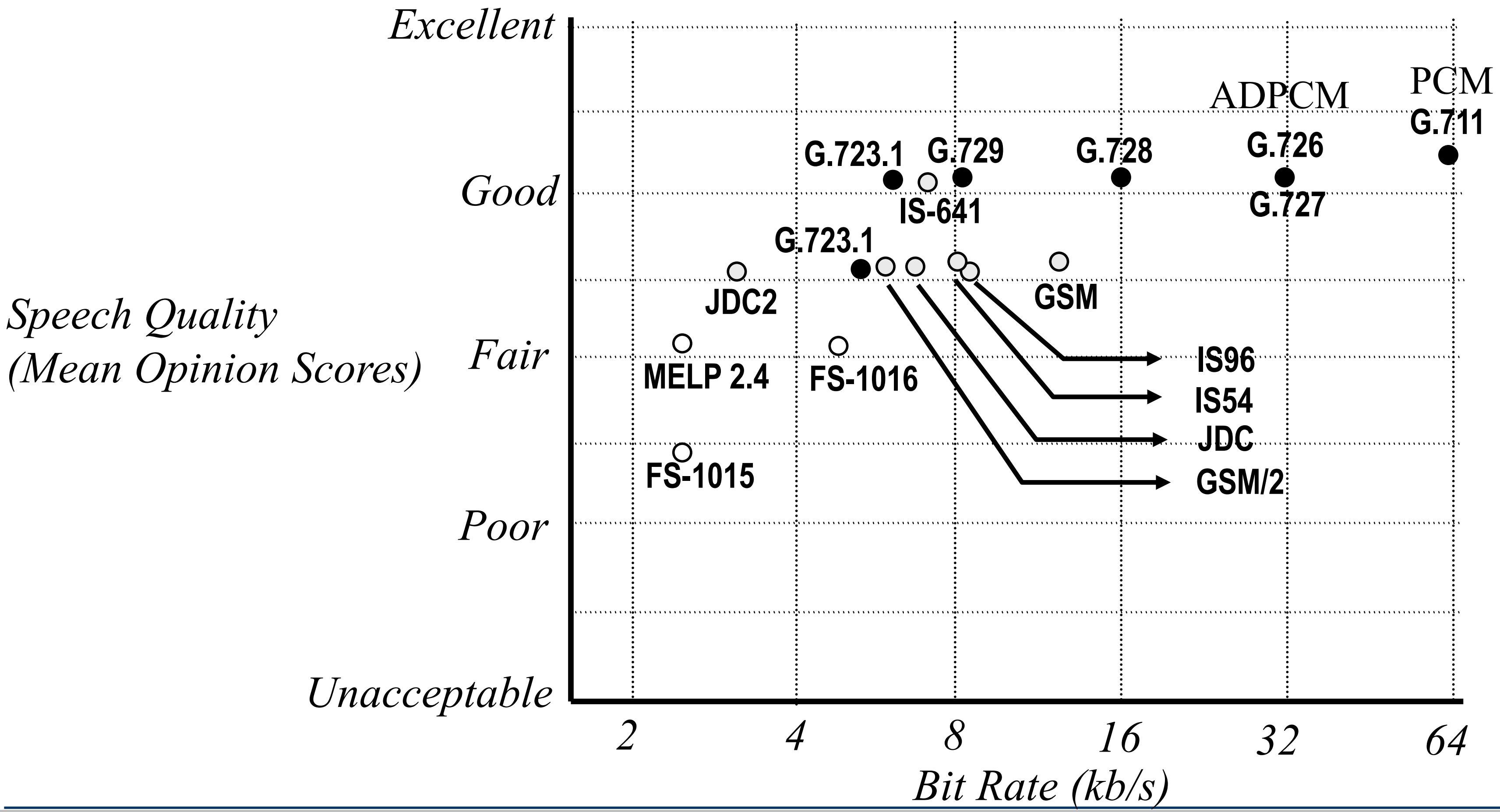
- Usually, the contribution of SNR_Q to SNR_O is fixed.
  - One additional bit in quantization results in 6 dB improvement.
- G_p is the processing gain due to a nice "prediction."
  - The better the prediction is, the larger G_p is.

DPCM

- Final notes on DPCM
  - Comparing DPCM with PCM in the case of voice signals, the improvement is around 4-11 dB, depending on the prediction order.
  - The greatest improvement occurs in going from no prediction to first-order prediction, with some additional gain, resulting from increasing the prediction order up to 4 or 5, after which little additional gain is obtained.
  - For the same sampling rate (8KHz) and signal quality, DPCM may provide a saving of about 8~16 Kbps compared to standard PCM (64 Kpbs).

DPCM

Source: IEEE Communications Magazine, September 1997.



IS = Interim Standard GSM = Global System for Mobile Communications JDC = Japanese Digital Cellular
FS = Federal Standard MELP = Mixed-Excitation Linear Prediction

Adaptive Differential Pulse-Code Modulation

- *Adaptive prediction* is used in DPCM.
- Can we also combine *adaptive quantization* into DPCM to yield a comparably voice quality to PCM with 32 Kbps bit rate? The answer is YES from the previous figure.
  - 32 Kbps: 4 bits for one sample, and 8 KHz sampling rate
  - 64 Kbps: 8 bits for one sample, and 8 KHz sampling rate
- So, "adaptive" in ADPCM means being responsive to changing level and spectrum of the input speech signal.

Adaptive Quantization

- Adaptive quantization refers to a quantizer that operates with a *time-varying* step size $\Delta[n]$.
- $\Delta[n]$ is adjusted according to the power of input sample $m[n]$.
  - Power = variance, if $m[n]$ is zero-mean.

$$\Delta[n] = \phi \cdot \sqrt{E[m^2[n]]}$$

  - In practice, we can only obtain an estimate of $E[m^2[n]]$.

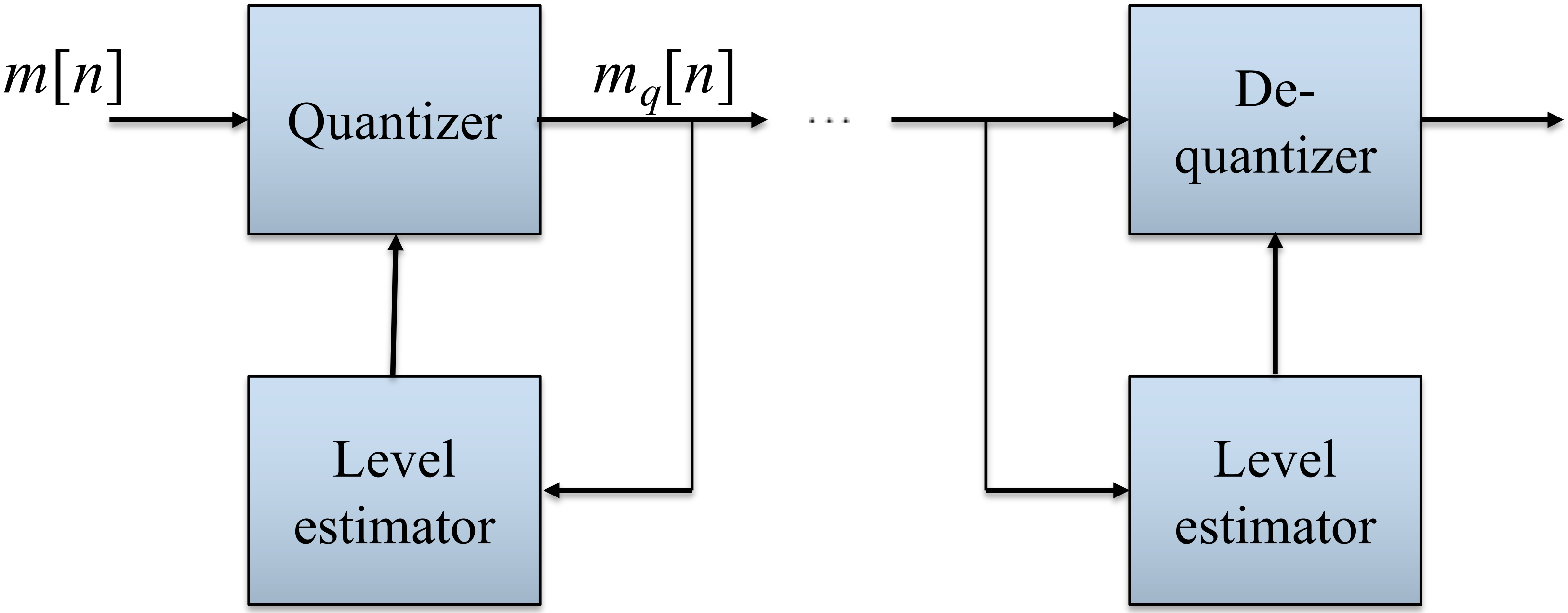
Adaptive Quantization

- The estimate of $E[m^2[n]]$ can be done in two ways:
  - Adaptive quantization with forward estimation (AQF)
    - Estimate based on *unquantized* samples of the input signals.
  - Adaptive quantization with backward estimation (AQB)
    - Estimate based on *quantized* samples of the input signals.

AQF

- AQF is in principle a more accurate estimator. However, it requires
  - an additional *buffer* to store unquantized samples for the learning period.
  - explicit transmission of level information to the receiver (the receiver, even without noise, only has the quantized samples).
  - a processing delay (around 16 ms for speech) due to buffering and other operations for AQF.
- The above requirements can be relaxed by using AQB.

AQB

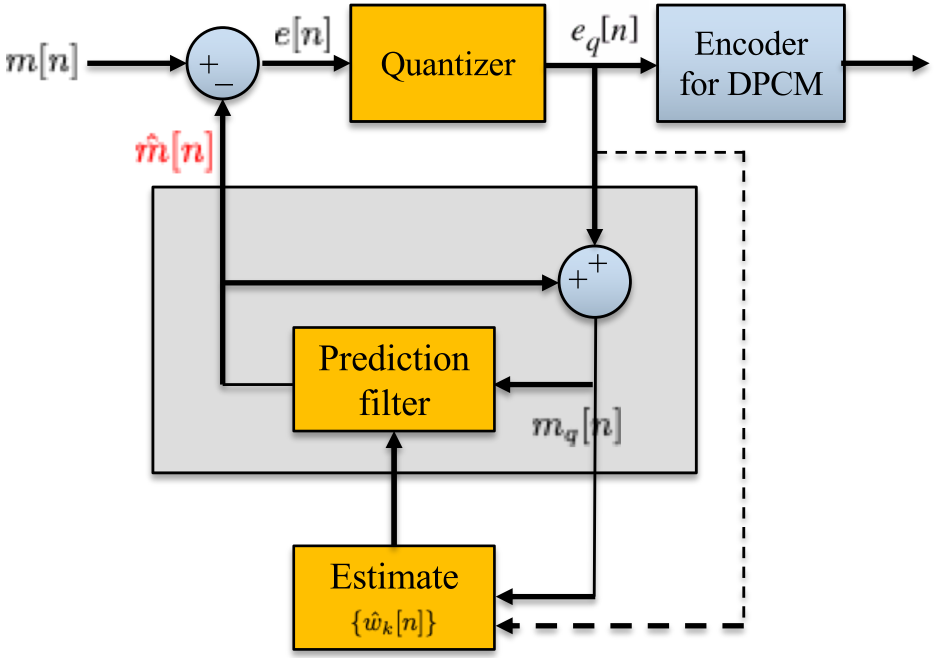


A possible drawback for a feedback system is its potential unstability. However, stability in this system can be guaranteed if $m_q[n]$ is bounded.

APF and APB

- Likewise, the prediction approach used in ADPCM can be classified into:
  - Adaptive prediction with forward estimation (APF)
    - Prediction based on *unquantized* samples of the input signals.
  - Adaptive prediction with backward estimation (APB)
    - Prediction based on *quantized* samples of the input signals.
- The pro and con of APF/APB is the same as AQF/AQB.
- APB/AQB are a preferred combination in practical applications.

ADPCM



Adaptive prediction with backward estimation (APB).

MPEG Audio Coding Standard

- The ADPCM and various voice coding techniques introduced above did not consider the *human auditory perception*.
- In practice, a consideration on human auditory perception can further improve the system performance (from the human standpoint).
- The MPEG-1 standard is capable of achieving *transparent*, "*perceptually lossless" compression* of stereophonic audio signals at high sampling rate.
  - A human subjective test shows that a 6-to-1 compression ratio are "perceptually indistinguishable" to human.

Characteristics of Human Auditory System

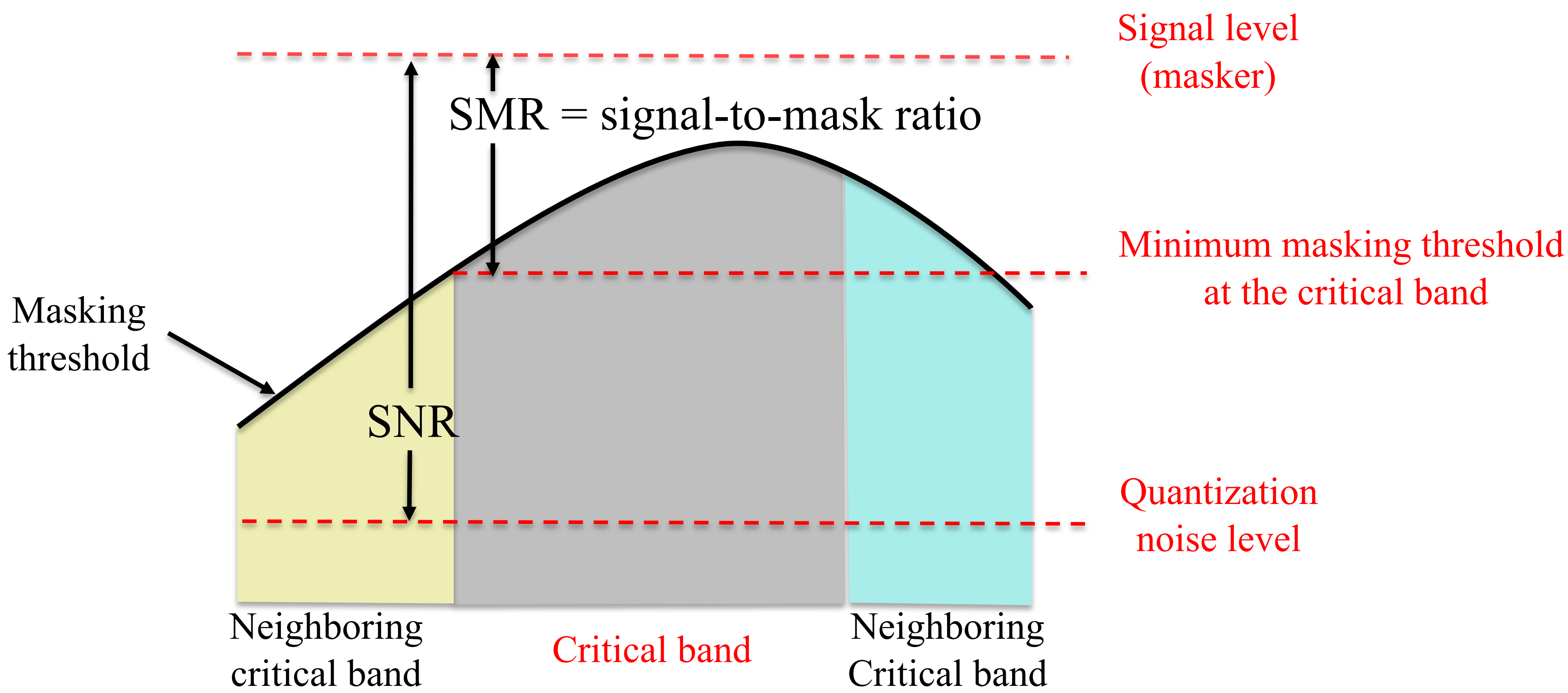
- Psychoacoustic characteristic of human auditory system
  - Critical band
    - The inner ear will scale the power spectra of incoming signals nonlinearly in the form of limited frequency bands called the *critical bands*.
    - Roughly, the inner ear can be modeled as 25 selective overlapping band-pass filters with bandwidth < 100Hz for the lowest audible frequencies, and up to 5kHz for the highest audible frequencies.

Characteristics of Human Auditory System

- Auditory masking
  - When a low-(power-)level signal (i.e., the maskee) and a high-(power-)level signal (i.e., the masker) occur simultaneously in the same critical band, and are close to each other in frequency, the low-(power-)level signal will be made *inaudible* (i.e., masked) by the high-(power-)level signal, if the low-(power-)level one lies below a *masking threshold*.

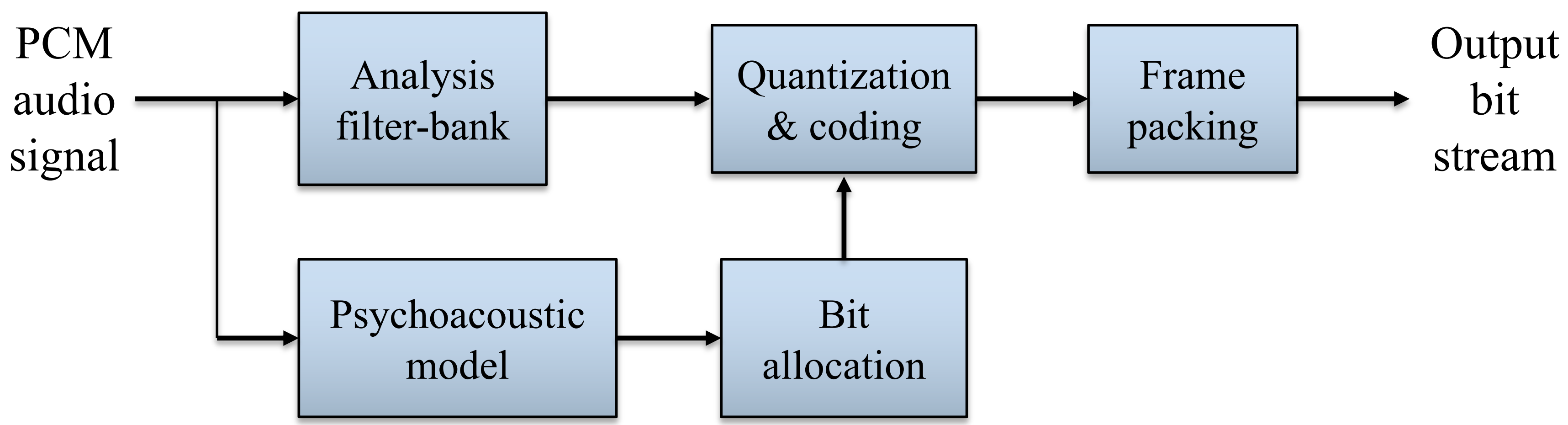
Characteristic of Human Auditory System

- *Masking threshold* is frequency-dependent.



Within a critical band, the quantization noise is inaudible as long as the NMR = SMR - SNR for the pertinent quantizer is negative.

MPEG Audio Coding Standard



MPEG Audio Coding Standard

- Analysis filter-bank
  - Divide the audio signal into a proper number of subbands, which is a compromise design for computational efficiency and perceptual performance.
- Psychoacoustic model
  - Analyze the spectral content of the input audio signal and thereby compute the signal-to-mask ratio (SMR).
- Quantization & coding
  - Decide how to apportion the available number of bits for the quantization of the subband signals.
- Frame packing
  - Assemble the quantized audio samples into a decodable bit stream.

Summary

- TDM (Time-Division Multiplexing)
- PCM, DM, DPCM, ADPCM
- MPEG audio coding