# Performance Analysis and Optimal Filter Design for Sigma-Delta Modulation via Duality with DPCM

Or Ordentlich
Joint work with Uri Erez

ISIT 2015, Hong Kong
June 15, 2015

# Oversampled Data Conversion

- $X(t)$ is a stationary Gaussian process with $S_X(f) = 0, \forall |f| > f_{\text{max}}$
- Sampling $X(t)$ at Nyquist's rate gives the discrete process $X_n$
- Sampling $X(t)$ at $L$ x Nyquist's rate gives the discrete process $X_n^L$

# Oversampled Data Conversion

- $X(t)$ is a stationary Gaussian process with $S_X(f) = 0, \forall |f| > f_{\max}$
- Sampling $X(t)$ at Nyquist's rate gives the discrete process $X_n$
- Sampling $X(t)$ at $L$ x Nyquist's rate gives the discrete process $X_n^L$

**Rate-Distortion 101**

- The number of bits per second for describing both processes with distortion $D$ is equal
- Normalizing by the number of samples per second gives

$$R_{X^L}(D) = \frac{1}{L} \cdot R_X(D)$$

# Oversampled Data Conversion

- $X(t)$ is a stationary Gaussian process with $S_X(f) = 0, \forall |f| > f_{\max}$
- Sampling $X(t)$ at Nyquist's rate gives the discrete process $X_n$
- Sampling $X(t)$ at $L$ x Nyquist's rate gives the discrete process $X_n^L$

**Rate-Distortion 101**

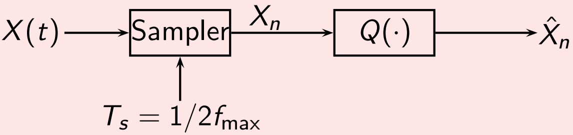
- The number of bits per second for describing both processes with distortion $D$ is equal
- Normalizing by the number of samples per second gives

$$R_{X^L}(D) = \frac{1}{L} \cdot R_X(D)$$

**In data conversion fast low-resolution ADCs are often preferable over slow high-resolution ADCs**

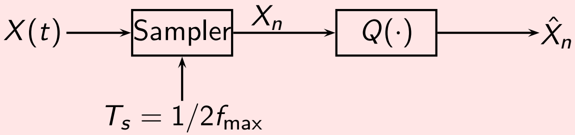
# ΣΔ Modulation

## Standard Data Conversion

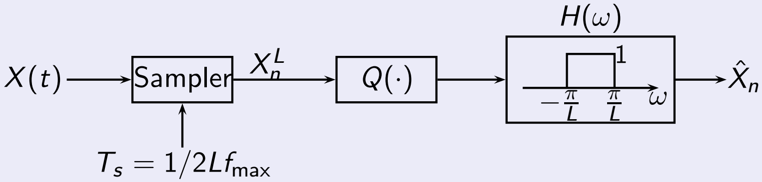


# ΣΔ Modulation

## Standard Data Conversion

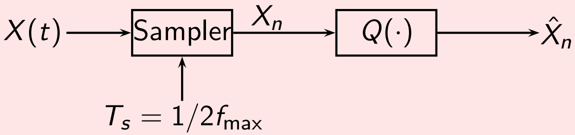



## Oversampled Data Conversion

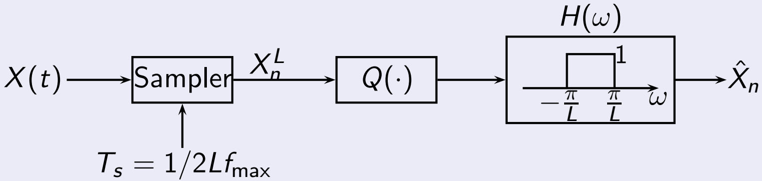


# ΣΔ Modulation

## Standard Data Conversion



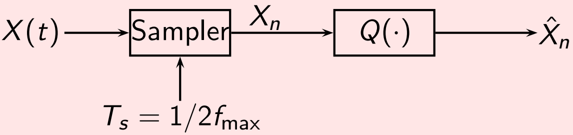

## Oversampled Data Conversion



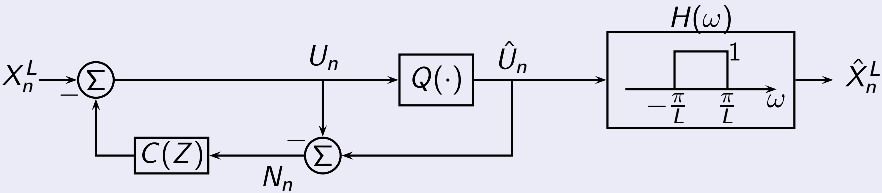

- Oversampling reduces the MSE distortion by $1/L$
  $\Rightarrow$ Not good enough, want exponential decay with $L$

# ΣΔ Modulation

## Standard Data Conversion

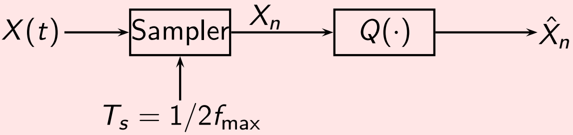



## ΣΔ Modulation

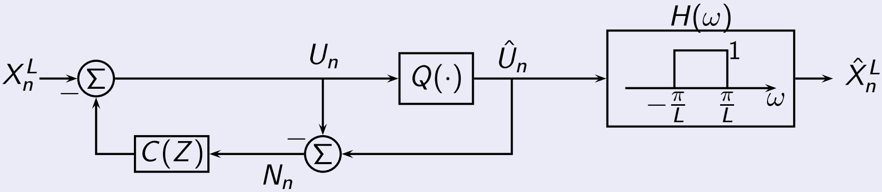


# ΣΔ Modulation

## Standard Data Conversion



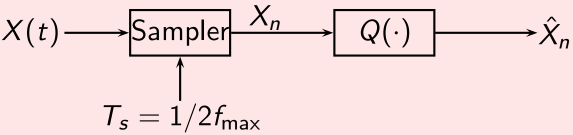

## ΣΔ Modulation



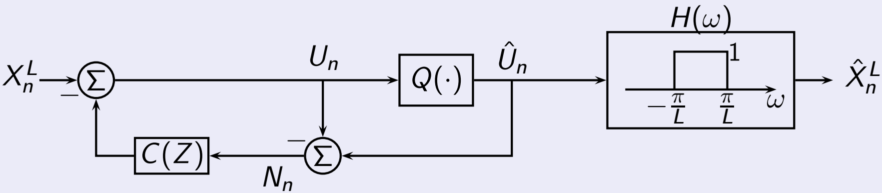

- Our goal is to analyze the **performance** of ΣΔ:
  **Quantization rate vs. MSE distortion**

# ΣΔ Modulation

## Standard Data Conversion



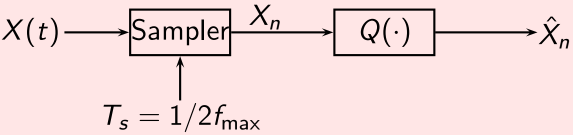

## ΣΔ Modulation



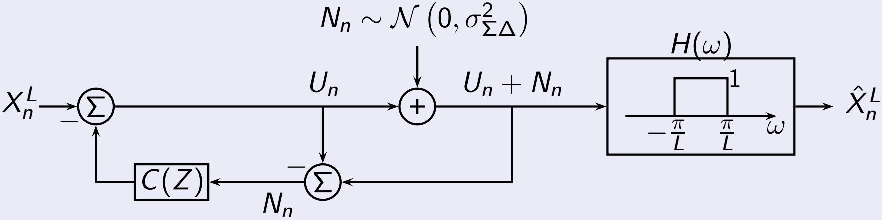

- We will model the ΣΔ modulator by a test-channel

# ΣΔ Modulation

## Standard Data Conversion



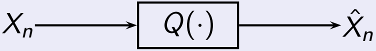

## ΣΔ Modulation



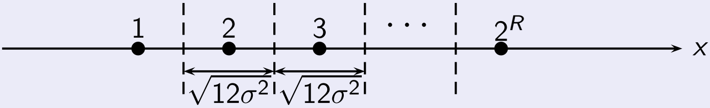

- Will study the tradeoff between $I(U_n; U_n + N_n)$ and the MSE distortion $\mathbb{E}(\hat{X}_n^L - X_n^L)^2$

# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization

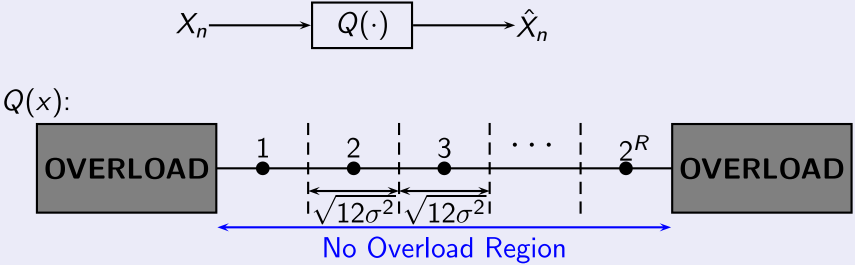



$Q(x)$:



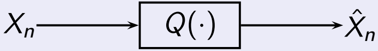
# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization

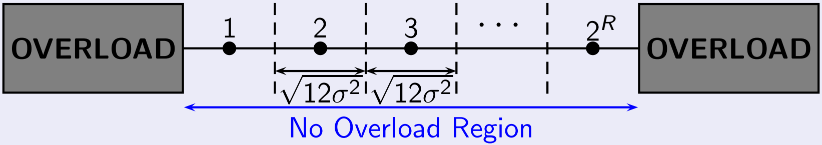


# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization




$Q(x)$:

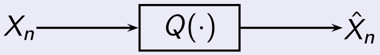


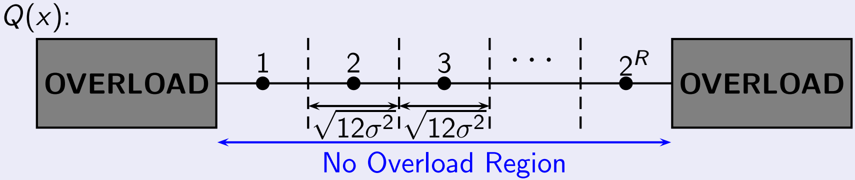

- High-resolution/dithered quantization assumption + no overload

$$\hat{X}_n = X_n + N_n; \quad N_n \sim \text{Uniform}\left(-\frac{\sqrt{12\sigma^2}}{2}, \frac{\sqrt{12\sigma^2}}{2}\right), \quad X_n \perp\!\!\!\perp N_n$$

# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization



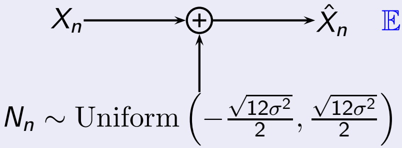

$$|X_n + N_n| < \frac{2^R\sqrt{12\sigma^2}}{2}$$

- High-resolution/dithered quantization assumption + no overload

$$\hat{X}_n = X_n + N_n; \quad N_n \sim \text{Uniform}\left(-\frac{\sqrt{12\sigma^2}}{2}, \frac{\sqrt{12\sigma^2}}{2}\right), \quad X_n \perp\!\!\!\perp N_n$$

# Relevance of Gaussian Test Channel

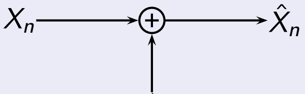
## Uniform Scalar Quantization




$$\mathbb{E}\left(\hat{X}_n - X_n\right)^2 = \sigma^2$$

# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization




$$\mathbb{E}\left(\hat{X}_n - X_n\right)^2 = \sigma^2$$

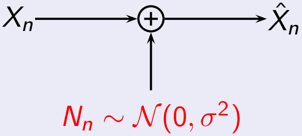
$$N_n \sim \mathrm{Uniform}\left(-\frac{\sqrt{12\sigma^2}}{2}, \frac{\sqrt{12\sigma^2}}{2}\right)$$

- Recalling $X_n \sim \mathcal{N}(0, \sigma_X^2)$, it is easy to show

$$P_{ol} \triangleq \Pr\left(|X_n + N_n| > \frac{2^R\sqrt{12\sigma^2}}{2}\right) \leq 2\exp\left\{-\frac{3}{2}2^{2\left(R - \frac{1}{2}\log\left(1 + \frac{\sigma_X^2}{\sigma^2}\right)\right)}\right\}$$

# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization



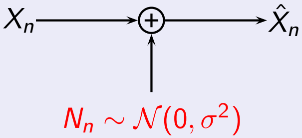

$$\mathbb{E}\left(\hat{X}_n - X_n\right)^2 = \sigma^2$$

- Recalling $X_n \sim \mathcal{N}(0, \sigma_X^2)$, it is easy to show

$$P_{ol} \leq 2\exp\left\{-\frac{3}{2}2^{2(R - I(X_n; X_n + N_n))}\right\}$$

# Relevance of Gaussian Test Channel

## Uniform Scalar Quantization



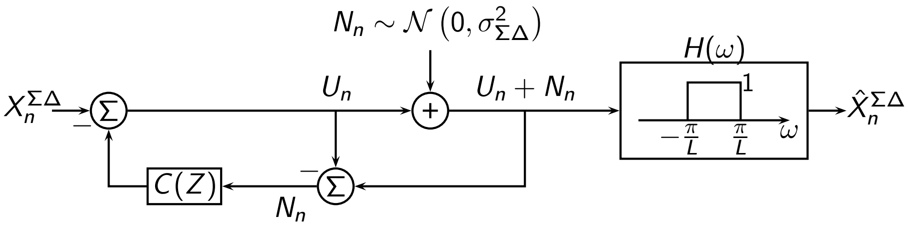

$$\mathbb{E}\left(\hat{X}_n - X_n\right)^2 = \sigma^2$$

- Recalling $X_n \sim \mathcal{N}(0, \sigma_X^2)$, it is easy to show

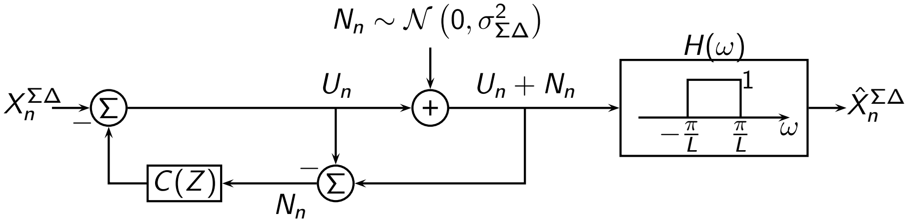
$$P_{ol} \leq 2\exp\left\{-\frac{3}{2}2^{2(R - I(X_n; X_n + N_n))}\right\}$$

Conclusion: the quantizer can be replaced by an AWGN test-channel

# Back to the ΣΔ Test Channel

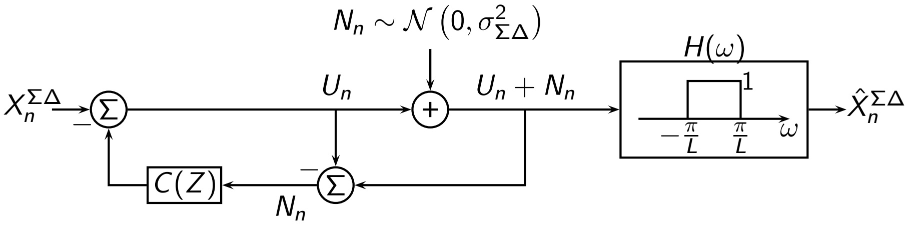


# Back to the ΣΔ Test Channel




- $U_n = X_n^{\Sigma\Delta} - c_n * N_n$
- $U_n + N_n = X_n^{\Sigma\Delta} + (\delta_n - c_n) * N_n$
- $I(U_n; U_n + N_n) = \frac{1}{2} \log \left(1 + \frac{\mathbb{E}(U_n)^2}{\sigma^2_{\Sigma\Delta}}\right)$
- $\hat{X}_n = X_n^{\Sigma\Delta} + h_n * (\delta_n - c_n) * N_n$
- $X_n^{\Sigma\Delta} - \hat{X}_n^{\Sigma\Delta} = h_n * (\delta_n - c_n) * N_n$

# Back to the ΣΔ Test Channel



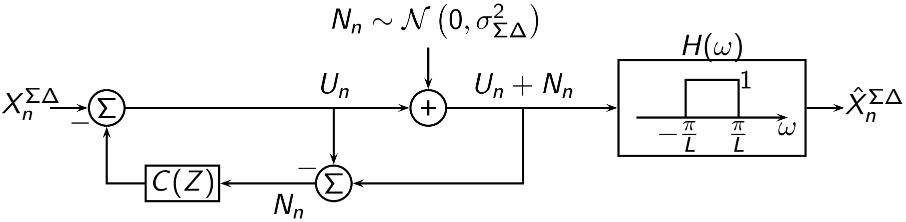

## Proposition - ΣΔ Rate-Distortion Tradeoff

For **any** stationary Gaussian process with variance $\sigma_X^2$ sampled $L$ times above Nyquist's rate

$$I(U_n; U_n + N_n) = \frac{1}{2}\log\left(1 + \frac{1}{2\pi}\int_{-\pi}^{\pi} |C(\omega)|^2 d\omega + \frac{\sigma_X^2}{\sigma^2_{\Sigma\Delta}}\right),$$

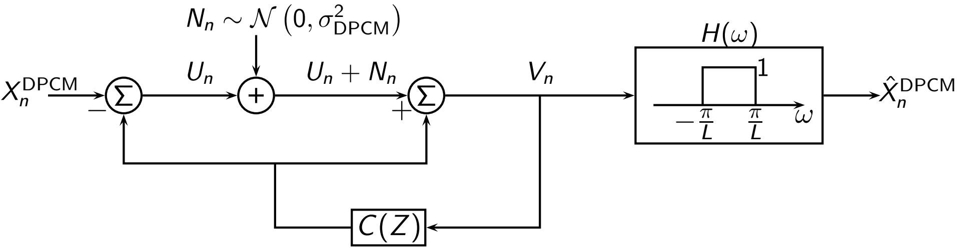
$$D = \sigma^2_{\Sigma\Delta} \cdot \frac{1}{2\pi}\int_{-\pi/L}^{\pi/L} |1 - C(\omega)|^2 d\omega$$

# Back to the ΣΔ Test Channel



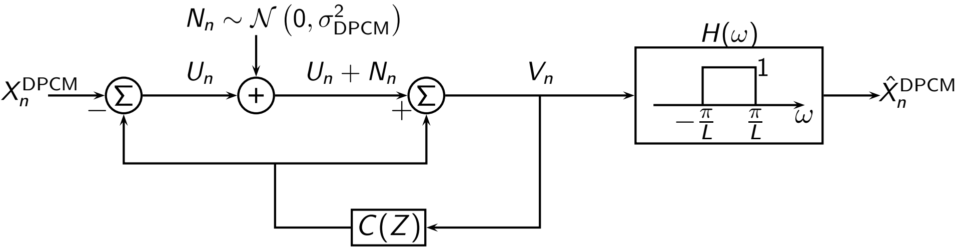

Not clear how to choose $C(Z)$

# Detour: DPCM



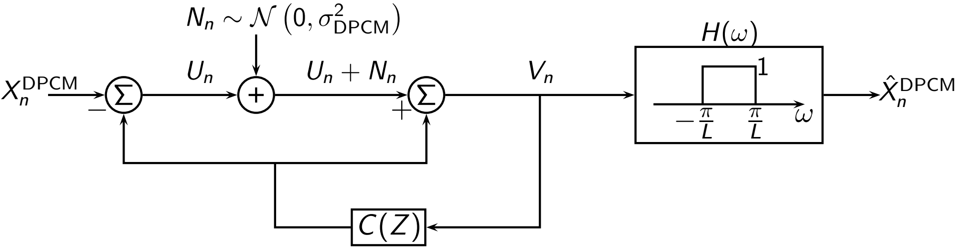
$N_n \sim \mathcal{N}\left(0, \sigma^2_{\text{DPCM}}\right)$
$H(\omega)$
$X_n^{\text{DPCM}}$
$-$
$U_n$
$U_n + N_n$
$+$
$V_n$
1
$-\frac{\pi}{L}$
$\frac{\pi}{L}$
$\omega$
$\hat{X}_n^{\text{DPCM}}$
$C(Z)$

# Detour: DPCM




- Popular for compression of stationary processes (rather than A/D)
- Design depends on $2_{nd}$-order statistics of $\{X_n^{\text{DPCM}}\}$ (in contrast to $\Sigma\Delta$)
- Rate-Distortion tradeoff of DPCM is well understood *(McDonald66, JN84, ZKE08)*

# Detour: DPCM




## DPCM Rate-Distortion Tradeoff for Flat Low-Pass Process

Let $\{X_n^{\text{DPCM}}\}$ be a stationary Gaussian process with PSD

$$S_X^{\text{DPCM}}(\omega) = \begin{cases} L\sigma_X^2 & \text{for } |\omega| \le \pi/L \\ 0 & \text{for } \pi/L < |\omega| < \pi \end{cases},$$

then $D = \sigma^2_{\text{DPCM}}/L$ and

$$I(U_n; U_n+N_n) = \frac{1}{2}\log\left(1+\frac{1}{2\pi}\int_{-\pi}^{\pi}|C(\omega)|^2 d\omega + \frac{L\sigma_X^2}{\sigma^2_{\text{DPCM}}}\frac{1}{2\pi}\int_{-\pi/L}^{\pi/L}|1-C(\omega)|^2 d\omega\right)$$

# Main Results: ΣΔ-DPCM Duality

Comparing the two rate-distortion characterizations we get

## ΣΔ-DPCM Duality

- Let $\{X_n^{\Sigma\Delta}\}$ be **any** Gaussian stationary process with variance $\sigma_X^2$ whose PSD is zero for all $\omega \notin [-\pi/L, \pi/L]$
- Let $\{X_n^{\text{DPCM}}\}$ be a flat stationary Gaussian process with PSD

$$S_X^{\text{DPCM}}(\omega) = \begin{cases} L\sigma_X^2 & \text{for } |\omega| \leq \pi/L \\ 0 & \text{for } \pi/L < |\omega| < \pi \end{cases}$$

- Let $\sigma_{\Sigma\Delta}^2$ and $\sigma_{\text{DPCM}}^2$ satisfy

$$\frac{\sigma_{\text{DPCM}}^2}{\sigma_{\Sigma\Delta}^2} = L \cdot \frac{1}{2\pi} \int_{-\pi/L}^{\pi/L} |1 - C(\omega)|^2 d\omega$$

**For any choice of $C(Z)$, the ΣΔ and DPCM test-channels achieve the same rate-distortion tradeoff**

# Main Results: Characterization of Optimal $C(Z)$

- For DPCM the optimal $C(Z)$ should minimize the MSE prediction error of $\{X_n^{\text{DPCM}} + N_n\}$ from its past *(ZKE08)*
- For data-converters the filter $C(Z)$ cannot be too complex
- To model this, assume $C(Z)$ must belong to a family $\mathcal{C}$
  e.g., all FIR filters with 5-taps satisfying $|c_i| < 1/2$

# Main Results: Characterization of Optimal $C(Z)$

The ΣΔ-DPCM Duality gives

## Optimal ΣΔ Filter

- Let $\{X_n^{\Sigma\Delta}\}$ be **any** Gaussian stationary process with variance $\sigma_X^2$ whose PSD is zero for all $\omega \notin [-\pi/L, \pi/L]$

**The optimal constrained $C(Z) \in \mathcal{C}$ for ΣΔ modulation with target distortion $D$ is the optimal one-step MSE predictor for $\{S_n + W_n\}$, where $W_n \sim \mathcal{N}(0, LD)$ i.i.d., and $\{S_n\}$ is a flat stationary Gaussian low-pass process with PSD**

$$S_S(\omega) = \begin{cases} L\sigma_X^2 & \textbf{for } |\omega| \leq \pi/L \\ 0 & \textbf{for } \pi/L < |\omega| < \pi \end{cases}$$

# Main Results: Characterization of Optimal $C(Z)$

The $\Sigma\Delta$-DPCM Duality gives

## Optimal $\Sigma\Delta$ Filter

- Let $\{X_n^{\Sigma\Delta}\}$ be **any** Gaussian stationary process with variance $\sigma_X^2$ whose PSD is zero for all $\omega \notin [-\pi/L, \pi/L]$

**The optimal constrained $C(Z) \in \mathcal{C}$ for $\Sigma\Delta$ modulation with target distortion $D$ is the optimal one-step MSE predictor for $\{S_n + W_n\}$, where $W_n \sim \mathcal{N}(0, LD)$ i.i.d., and $\{S_n\}$ is a flat stationary Gaussian low-pass process with PSD**

$$S_S(\omega) = \begin{cases} L\sigma_X^2 & \textbf{for } |\omega| \leq \pi/L \\ 0 & \textbf{for } \pi/L < |\omega| < \pi \end{cases}$$

**The corresponding scalar MI is**

$$I(U_n; U_n + N_n) = \frac{1}{2}\log\left(\frac{\mathbb{E}\left((\delta_n - c_n) * (S_n + W_n)\right)^2}{D}\right)$$

# Main Results: Minimax Optimality of Unconstrained ΣΔ

**Unconstrained DPCM is Rate-Distortion Optimal**

If $\mathcal{C}$ consists of all causal filters, the DPCM architecture attains the optimal rate-distortion function for stationary Gaussian sources *(ZKE08)*

# Main Results: Minimax Optimality of Unconstrained ΣΔ

For flat stationary Gaussian process $\{S_n\}$ with PSD

$$S_S(\omega) = \begin{cases} L\sigma_X^2 & \text{for } |\omega| \leq \pi/L \\ 0 & \text{for } \pi/L < |\omega| < \pi \end{cases}$$

unconstrained DPCM attains $R_S(D) = \frac{1}{2L}\log(\frac{\sigma_X^2}{D})$

# Main Results: Minimax Optimality of Unconstrained $\Sigma\Delta$

For flat stationary Gaussian process $\{S_n\}$ with PSD

$$S_S(\omega) = \begin{cases} L\sigma_X^2 & \text{for } |\omega| \leq \pi/L \\ 0 & \text{for } \pi/L < |\omega| < \pi \end{cases}$$

unconstrained DPCM attains $R_S(D) = \frac{1}{2L}\log(\frac{\sigma_X^2}{D})$

The $\Sigma\Delta$-DPCM Duality gives

## Minimax Optimality of $\Sigma\Delta$ Architecture

- Let $\{X_n^{\Sigma\Delta}\}$ be **any** Gaussian stationary process with variance $\sigma_X^2$ whose PSD is zero for all $\omega \notin [-\pi/L, \pi/L]$

**Unconstrained $\Sigma\Delta$ attains $R_{X^{\Sigma\Delta}}(D) = \frac{1}{2L}\log(\frac{\sigma_X^2}{D})$ universally for all $\{X_n^{\Sigma\Delta}\}$**

# Main Results: Minimax Optimality of Unconstrained $\Sigma\Delta$

For flat stationary Gaussian process $\{S_n\}$ with PSD

$$S_S(\omega) = \begin{cases} L\sigma_X^2 & \text{for } |\omega| \leq \pi/L \\ 0 & \text{for } \pi/L < |\omega| < \pi \end{cases}$$

unconstrained DPCM attains $R_S(D) = \frac{1}{2L}\log(\frac{\sigma_X^2}{D})$

The $\Sigma\Delta$-DPCM Duality gives

## Minimax Optimality of $\Sigma\Delta$ Architecture

- Let $\{X_n^{\Sigma\Delta}\}$ be **any** Gaussian stationary process with variance $\sigma_X^2$ whose PSD is zero for all $\omega \notin [-\pi/L, \pi/L]$

**Unconstrained $\Sigma\Delta$ attains** $R_{X^{\Sigma\Delta}}(D) = \frac{1}{2L}\log(\frac{\sigma_X^2}{D})$ **universally for all** $\{X_n^{\Sigma\Delta}\}$

For $\{X_n^{\Sigma\Delta}\} = \{S_n\}$ this is the optimal RD-function $\Rightarrow$ minimax optimality

# High-Resolution in ΣΔ Modulation?

## Prediction in high-resolution quantization

If the PSD of $\{A_n\}$ is positive for all $\omega$, the optimal predictor of $\{A_n + W_n\}$ from its past approaches the optimal predictor of $\{A_n\}$ from its past
Same is true for the MSE prediction error

# High-Resolution in $\Sigma\Delta$ Modulation?

## Prediction in high-resolution quantization

If the PSD of $\{A_n\}$ is positive for all $\omega$, the optimal predictor of $\{A_n + W_n\}$ from its past approaches the optimal predictor of $\{A_n\}$ from its past
Same is true for the MSE prediction error

## Prediction in $\Sigma\Delta$

$S_S(\omega) = L\sigma_X^2$ for $|\omega| < \frac{\pi}{L}$ and 0 otherwise, $W_n \sim \mathcal{N}(0, LD)$ i.i.d.

- We showed that $C(Z)$ should predict $\{S_n + W_n\}$ from its past
  The quantization rate is $\frac{1}{2}\log\left(\frac{\mathbb{E}((\delta_n - c_n)*(S_n+W_n))^2}{D}\right)$

# High-Resolution in ΣΔ Modulation?

## Prediction in high-resolution quantization

If the PSD of $\{A_n\}$ is positive for all $\omega$, the optimal predictor of $\{A_n + W_n\}$ from its past approaches the optimal predictor of $\{A_n\}$ from its past
Same is true for the MSE prediction error

## Prediction in ΣΔ

$S_S(\omega) = L\sigma_X^2$ for $|\omega| < \frac{\pi}{L}$ and 0 otherwise, $W_n \sim \mathcal{N}(0, LD)$ i.i.d.

- We showed that $C(Z)$ should predict $\{S_n + W_n\}$ from its past
  The quantization rate is $\frac{1}{2}\log\left(\frac{\mathbb{E}((\delta_n - c_n) * (S_n + W_n))^2}{D}\right)$
- For $L > 1$ the prediction error of $\{S_n\}$ from its past can be made arbitrarily small by increasing the filter length

# High-Resolution in ΣΔ Modulation?

## Prediction in high-resolution quantization

If the PSD of $\{A_n\}$ is positive for all $\omega$, the optimal predictor of $\{A_n + W_n\}$ from its past approaches the optimal predictor of $\{A_n\}$ from its past
Same is true for the MSE prediction error

## Prediction in ΣΔ

$S_S(\omega) = L\sigma_X^2$ for $|\omega| < \frac{\pi}{L}$ and 0 otherwise, $W_n \sim \mathcal{N}(0, LD)$ i.i.d.

- We showed that $C(Z)$ should predict $\{S_n + W_n\}$ from its past
  The quantization rate is $\frac{1}{2}\log\left(\frac{\mathbb{E}((\delta_n - c_n) * (S_n + W_n))^2}{D}\right)$
- For $L > 1$ the prediction error of $\{S_n\}$ from its past can be made arbitrarily small by increasing the filter length
- **High resolution assumption never holds**

# High-Resolution in ΣΔ Modulation?

## Prediction in high-resolution quantization

If the PSD of $\{A_n\}$ is positive for all $\omega$, the optimal predictor of $\{A_n + W_n\}$ from its past approaches the optimal predictor of $\{A_n\}$ from its past

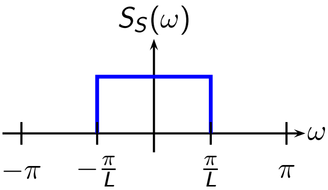
Same is true for the MSE prediction error

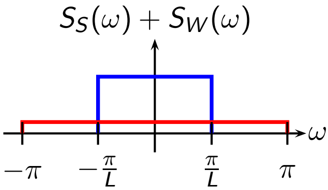
## Prediction in ΣΔ

$S_S(\omega) = L\sigma_X^2$ for $|\omega| < \frac{\pi}{L}$ and 0 otherwise, $W_n \sim \mathcal{N}(0, LD)$ i.i.d.

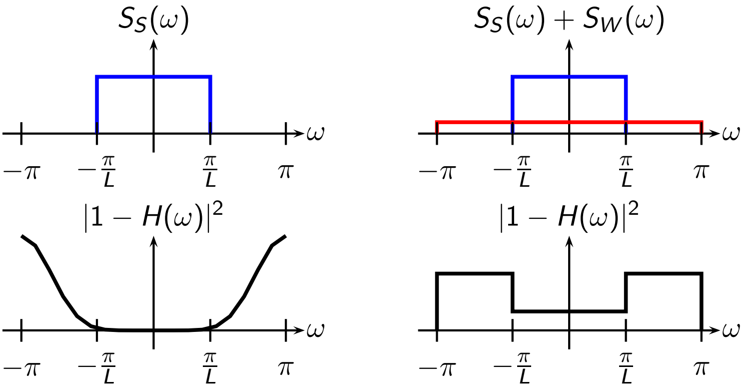
- We showed that $C(Z)$ should predict $\{S_n + W_n\}$ from its past
  The quantization rate is $\frac{1}{2}\log\left(\frac{\mathbb{E}((\delta_n - c_n)*(S_n+W_n))^2}{D}\right)$
- For $L > 1$ the prediction error of $\{S_n\}$ from its past can be made arbitrarily small by increasing the filter length
- **High resolution assumption never holds**
- Nevertheless... this assumption is sometimes erroneously made, leading to inaccurate results

# High-Resolution in ΣΔ Modulation?

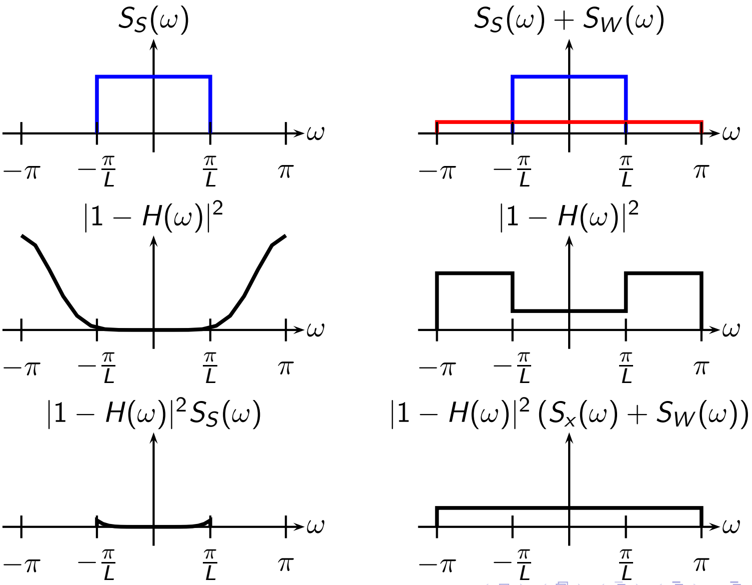


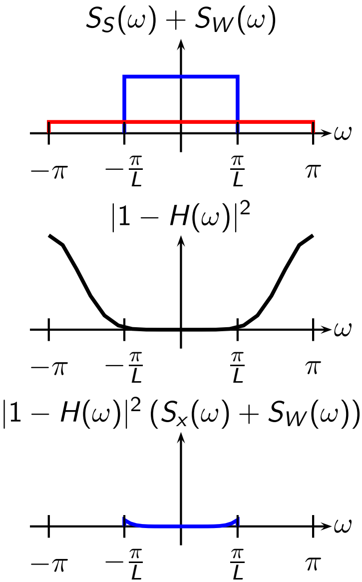
# High-Resolution in ΣΔ Modulation?

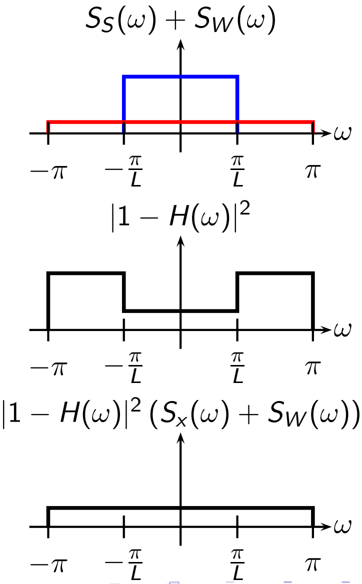


# High-Resolution in ΣΔ Modulation?

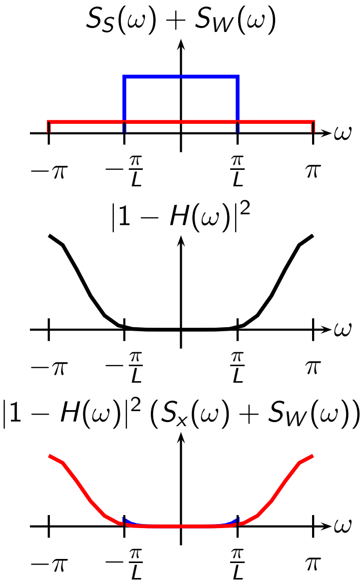


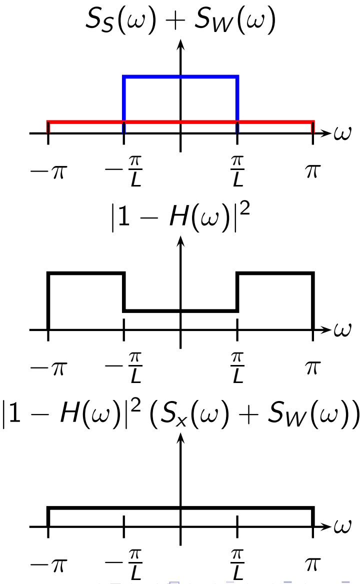
# High-Resolution in ΣΔ Modulation?



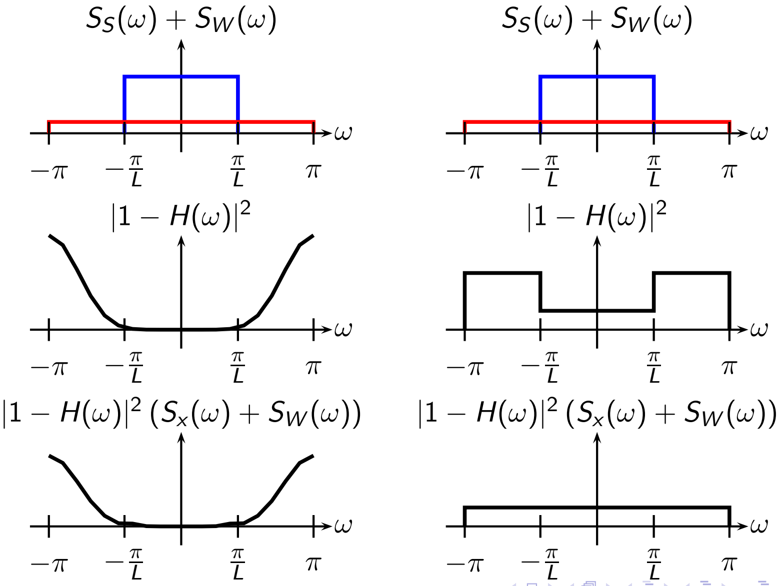



# High-Resolution in ΣΔ Modulation?

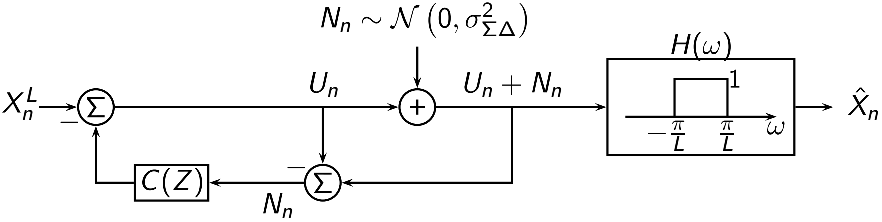


# High-Resolution in ΣΔ Modulation?



# From Test-Channel Back to a Data Converter

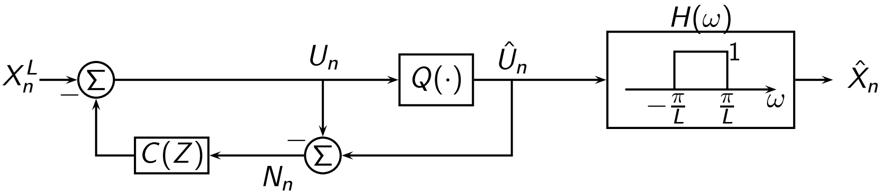



## Performance of ΣΔ Modulator

- Let $0 < P_e < 1$, and $R = I(U_n; U_n + N_n) + \delta(P_e)$ where

$$\delta(P_e) \triangleq \frac{1}{2} \log \left( -\frac{2}{3} \ln \frac{P_e}{2N} \right)$$

# From Test-Channel Back to a Data Converter




## Performance of $\Sigma\Delta$ Modulator

- Let $0 < P_e < 1$, and $R = I(U_n; U_n + N_n) + \delta(P_e)$ where

$$\delta(P_e) \triangleq \frac{1}{2} \log \left( -\frac{2}{3} \ln \frac{P_e}{2N} \right)$$

- With probability $\geq 1 - P_e$ no overload occurs within the block
- If no overload occurs within the block the MSE distortion is smaller than $D\frac{1+o_N(1)}{1-P_e}$

# Summary

- We established a duality between DPCM for flat low-pass processes and $\Sigma\Delta$ modulation for the compound class of oversampled processes
- Using this duality we found the optimal feedback filter for $\Sigma\Delta$
- We showed that the $\Sigma\Delta$ architecture is robust and minimax optimal for this compound class
- DPCM with unconstrained filter is robust. For constrained filters it isn't
- Our analysis was information-theoretic, but remains relevant for $\Sigma\Delta$ modulators with scalar quantizers