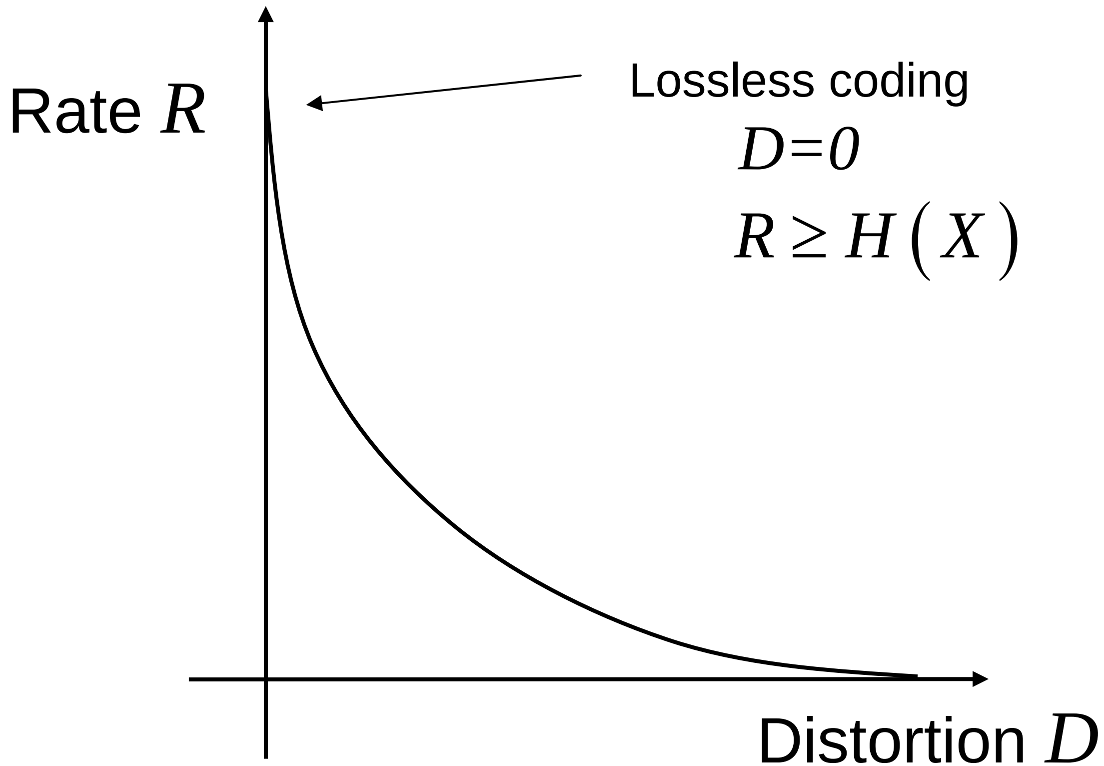
Lossy compression

<u>Goal:</u> Lower the bit-rate R by introducing some (acceptable) distortion D of the signal X.

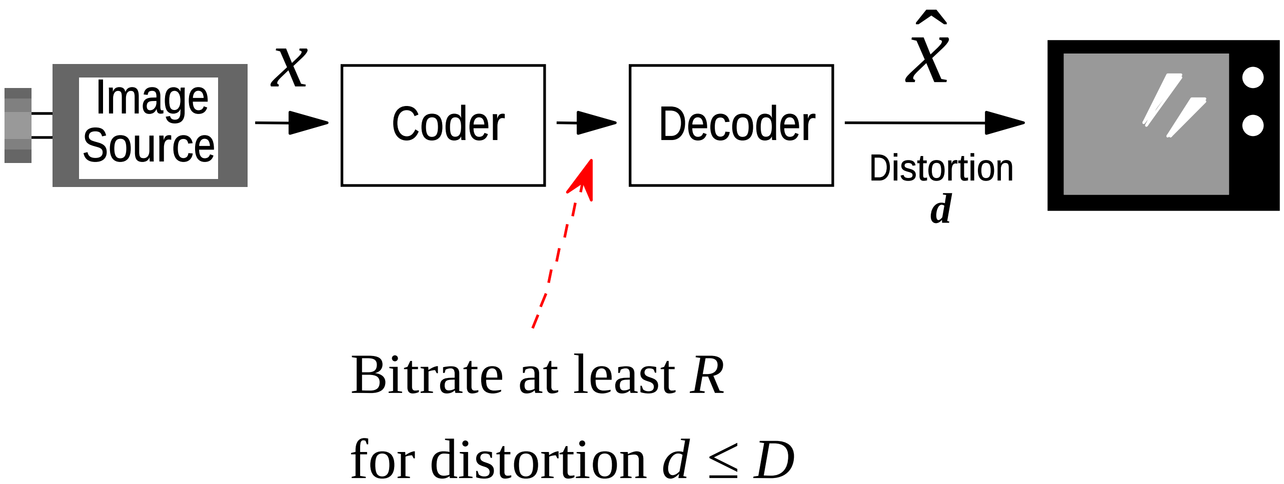


Topics in lossy compression

- Information theoretical bounds for lossy compression: *rate distortion theory*
  - R-D function for memoryless Gaussian sources with mean-squared error distortion criterion
  - R-D function for Gaussian sources with memory
  - R-D function for images
- Practical lossy compression techniques: *quantization*
  - Scalar quantization and vector quantization
  - Quantizer design for fixed length codes
  - Quantizer design for variable length codes
  - Embedded quantizers

Rate distortion theory

- Rate distortion theory calculates the minimum transmission bit-rate R for a required picture quality



- Results of rate distortion theory are obtained without consideration of a specific coding method

Distortion

- Symbol (signal, image . . .) x sent, $\hat{x}$ received
- Single-letter distortion measure:

$$\boxed{\begin{aligned} &\rho(x,\hat{x}) \geq 0 \\ &\rho(x,\hat{x}) = 0 \quad \text{for } x = \hat{x} \end{aligned}}$$

- Average distortion:

$$\boxed{d(X,\hat{X}) = E\left\{\rho(X,\hat{X})\right\} = \sum_x \sum_{\hat{x}} f_{X,\hat{X}}(x,\hat{x})\,\rho(x,\hat{x})}$$

- Distortion criterion:

$$\boxed{d(X,\hat{X}) \leq D}$$

Maximum permissible average distortion

Mutual information

- "Mutual information" is the average information that random variables X and Y convey about each other
  - Reduction in uncertainty about X, if Y is observed
  - Reduction in uncertainty about Y, if X is observed

$$
\boxed{\begin{aligned} I(X;Y) &= H(X) - H(X \mid Y) = H(Y) - H(Y \mid X) \\ &= \sum_x \sum_y f_{X,Y}(x,y) \log_2 \frac{f_{X,Y}(x,y)}{f_X(x) f_Y(y)} \end{aligned}}
$$

- Properties

$$0 \le I(X;Y) = I(Y;X)$$

$$I(X;Y) \le H(X)$$

$$I(X;Y) \le H(Y)$$

Rate distortion function

- Definition: $$\boxed{R(D) = \inf_{f_{\hat{X}|X} : d(X,\hat{X}) \le D} \{I(X;\hat{X})\}}$$
- Shannon's Source Coding Theorem (and converse):
  For a given maximum average distortion D, the rate distortion function $R(D)$ is the (achievable) lower bound for the transmission bit-rate.
- $R(D)$ is continuous, monotonically decreasing for $R>0$ and convex
- Equivalently use distortion-rate function $D(R)$

Extension to continuous random variables

- Differential entropy

$$h(X) = -E\{\log_2 f_X(X)\} = -\int_x f_X(x)\log_2 f_X(x)dx$$

- Differential conditional entropy

$$h(X|Y) = -E\{\log_2 f_{X|Y}(X,Y)\} = -\iint_{x,y} f_{X,Y}(x,y)\log_2 f_{X|Y}(x,y)dxdy$$

- Mutual information

$$I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X)$$

- Rate distortion function

$$R(D) = \inf_{f_{\hat{X}|X}:d(X,\hat{X})\le D} \{I(X;\hat{X})\}$$

Shannon lower bound

- It can be shown that $h(X-\hat{X} \mid \hat{X}) = h(X \mid \hat{X})$
- Thus

$$R(D) = \inf_{d \le D} \{h(X) - h(X \mid \hat{X})\}$$

$$= h(X) - \sup_{d \le D} \{h(X \mid \hat{X})\}$$

$$= h(X) - \sup_{d \le D} \{h(X - \hat{X} \mid \hat{X})\}$$

- Ideally, the source coder would introduce i.i.d. errors $X - \hat{X}$ that are statistically independent from the reconstructed signal $\hat{X}$ (not always possible!).
- Shannon lower bound:

$$\boxed{R(D) \ge h(X) - \sup_{d \le D} h(X - \hat{X})}$$

Shannon lower bound (cont.)

- Mean squared error distortion measure:
  Gaussian PDF possesses largest entropy for given variance

$$R(D) \geq h(X) - \sup_{d \leq D} h(X - \hat{X})$$
$$= h\left(X\right) - \tfrac{1}{2}\log_2 2\pi e D$$

- Equivalently

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

- Distortion reduction by 6 dB requires 1 bit/sample

$R(D)$ function for a memoryless Gaussian source and MSE distortion

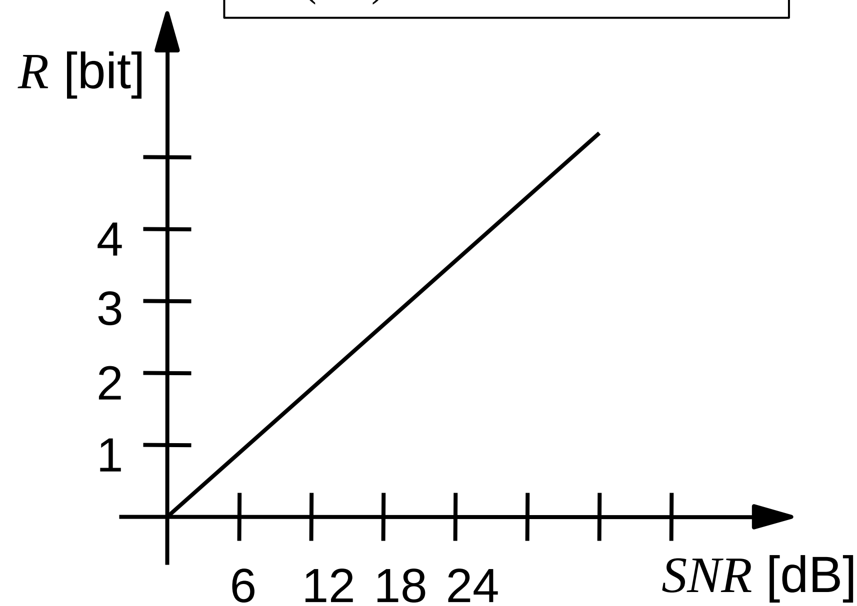
- Gaussian source, variance σ^2
- Mean squared error

$$d = E\{(X-\hat{X})^2\} \le D$$

- Rule of thumb: 6 dB $\cong$ 1 bit
- $R(D)$ for non-Gaussian sources with the same variance σ^2 is always <u>below</u> this Gaussian $R(D)$ curve.

$$R(D) = \frac{1}{2}\log_2\left(\frac{\sigma^2}{D}\right)$$

$$D(R) = \sigma^2 2^{-2R}$$



$$SNR = 10\log_{10}\left(\frac{\sigma^2}{D}\right) \text{ [dB]}$$

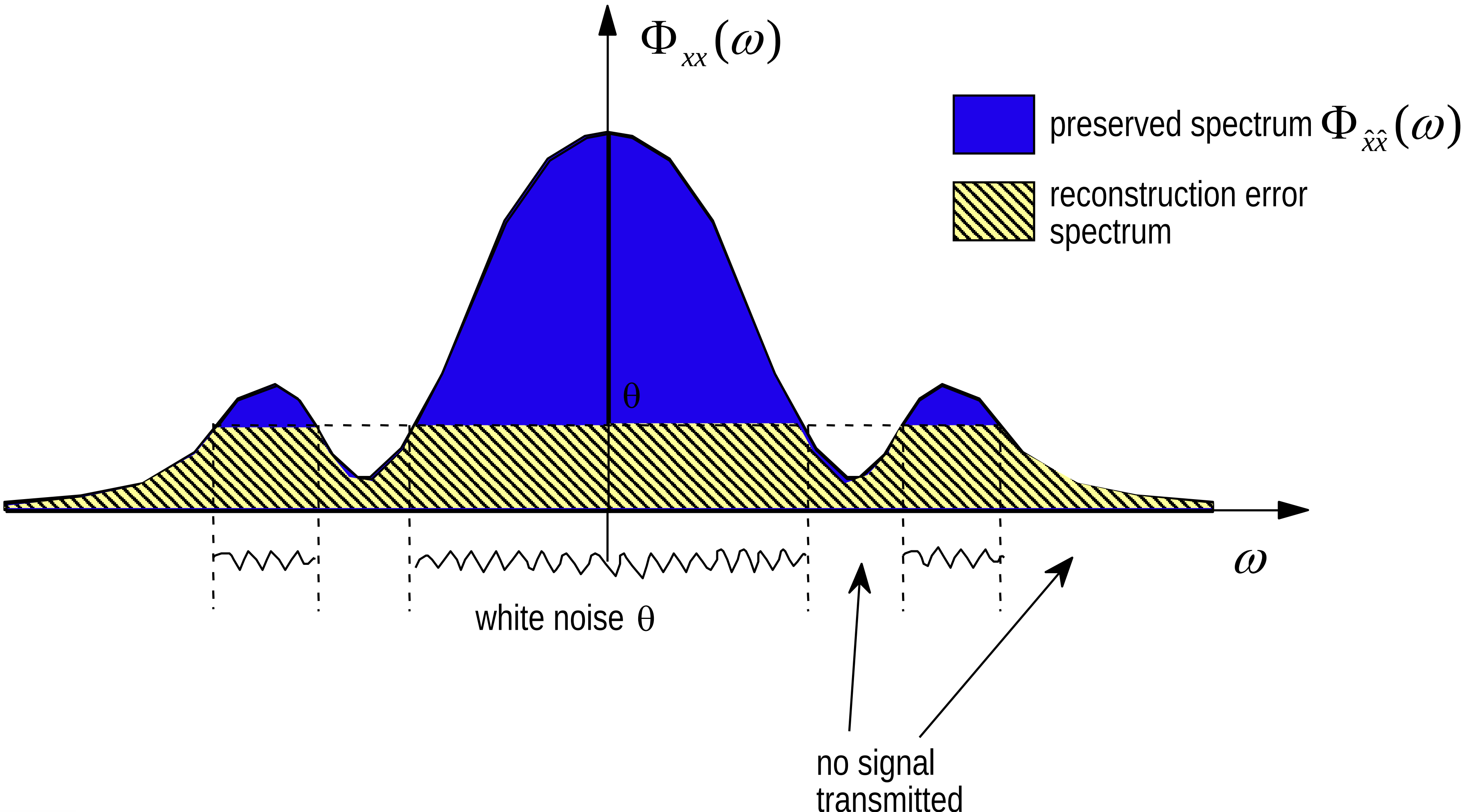
$R(D)$ for Gaussian source with memory

- Stationary, band-limited, jointly Gaussian source with power spectral density $\Phi_{xx}(\omega)$
- Mean squared error distortion $d = E\{(X - \hat{X})^2\} \leq D$
- $R(D)$ function in parametric form

$$\boxed{\begin{aligned} D(\theta) &= \frac{1}{2\pi} \int_{\omega} \min\left\{\theta, \Phi_{xx}(\omega)\right\} d\omega \\ R(\theta) &= \frac{1}{2\pi} \int_{\omega} \max\left\{0, \frac{1}{2} \log \frac{\Phi_{xx}(\omega)}{\theta}\right\} d\omega \end{aligned}}$$

- $R(D)$ for non-Gaussian sources with the same power spectral density is always lower.

$R(D)$ for Gaussian source with memory



Rate distortion function for images

- Signal model: Gaussian source with acf

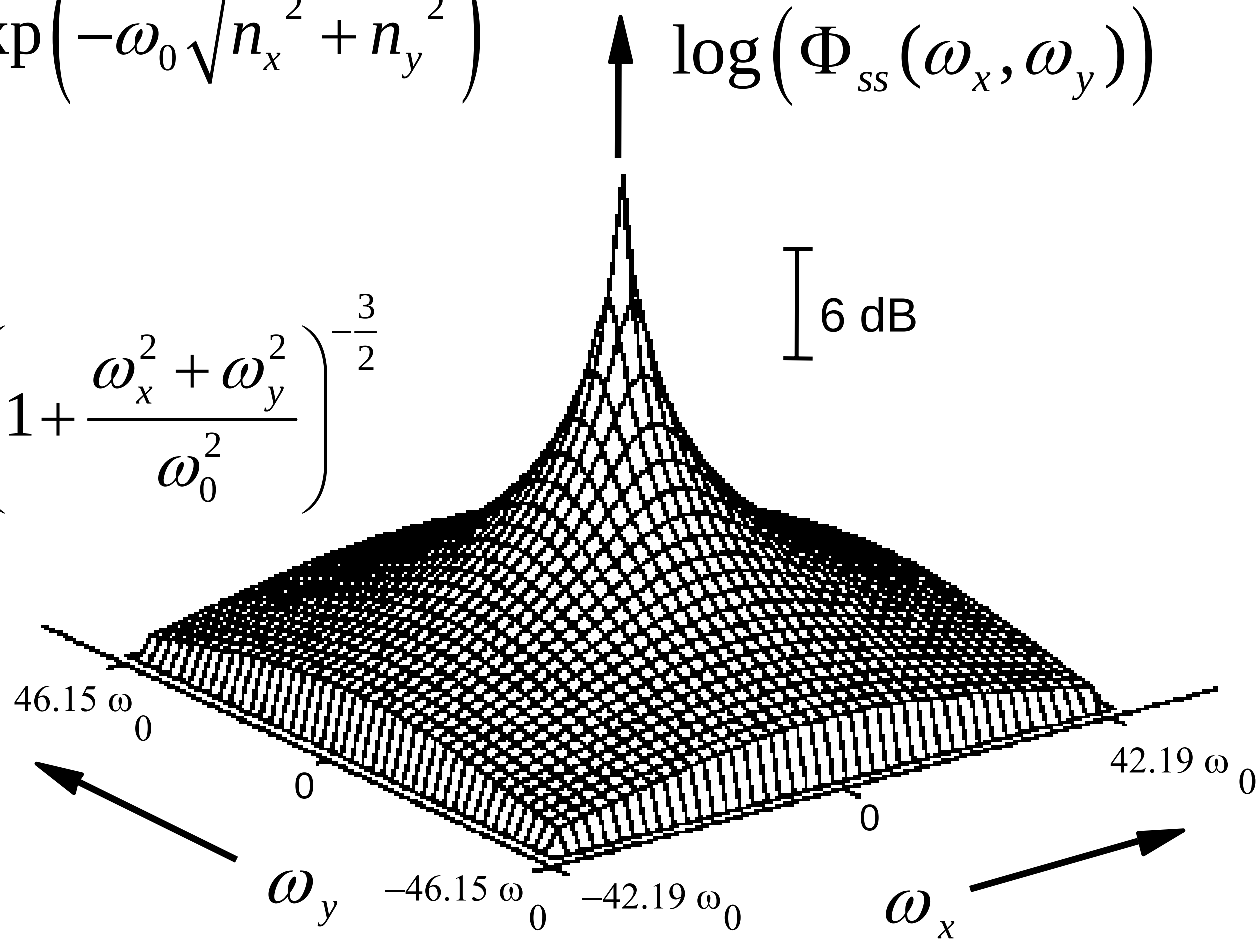
$$R_{ss}\left[n_x, n_y\right] = \exp\left(-\omega_0 \sqrt{{n_x}^2 + {n_y}^2}\right)$$

- Power spectral d
(neglecting alias

$$\Phi_{ss}(\omega_x, \omega_y) = \frac{2\pi}{\omega_0^2}\left(1 + \frac{\omega_x^2 + \omega_y^2}{\omega_0^2}\right)^{-\frac{3}{2}}$$

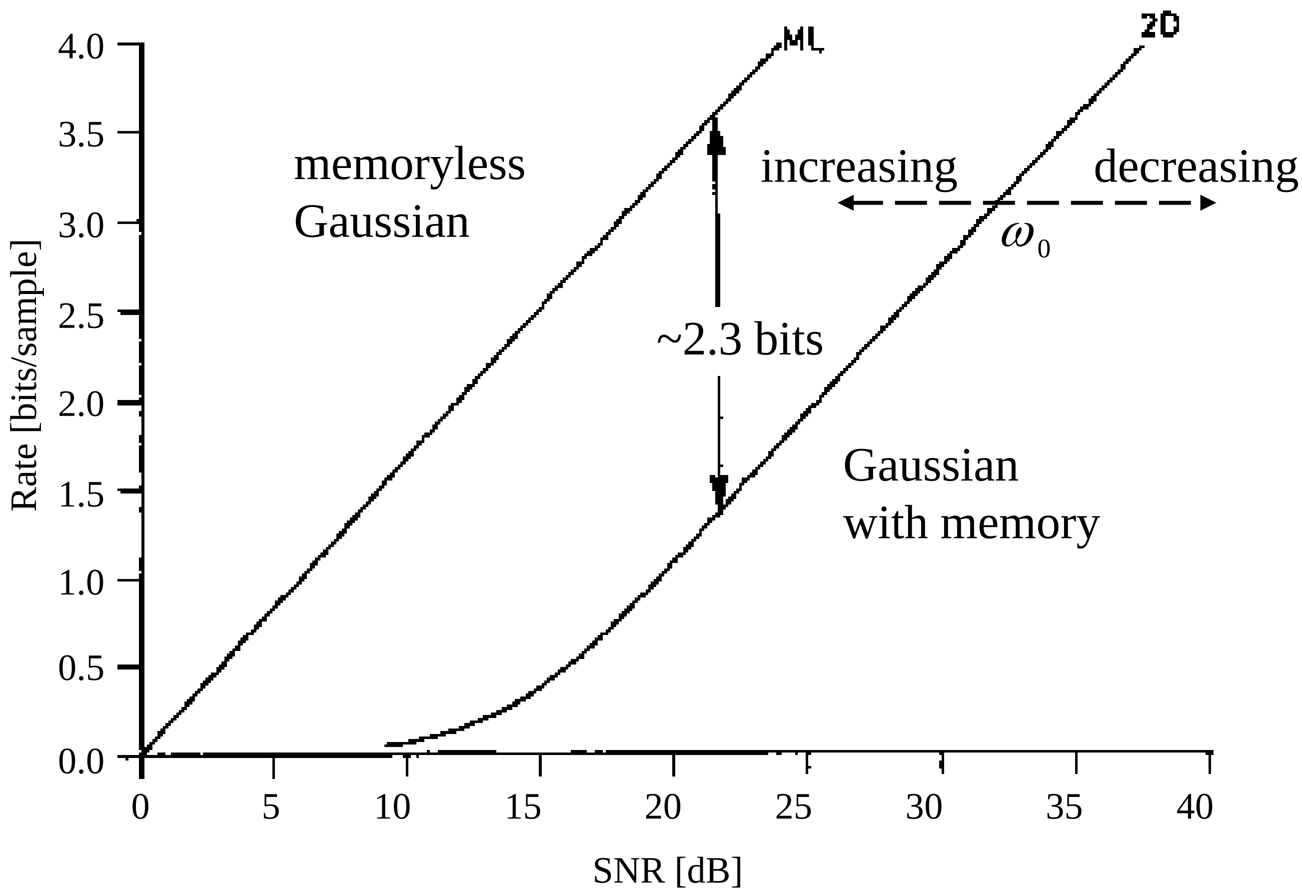
$$\omega_0 = -\ln(0.93)$$

correlation between
adjacent pixels



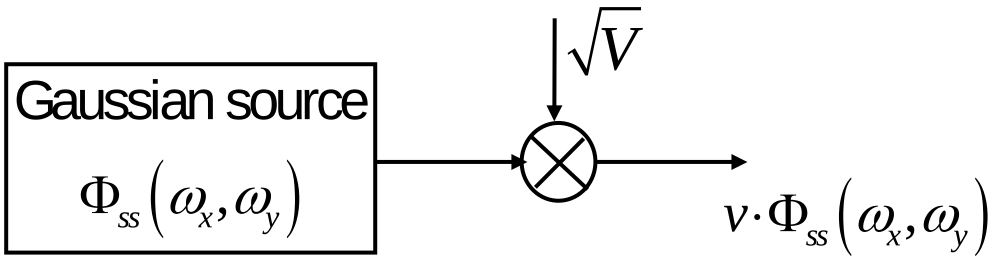
Rate distortion function for images (cont.)

- Mean squared error criterion: $D = E\{(S - \hat{S})^2\}$
- After numerical integration:



Infinite Gaussian mixture modeling

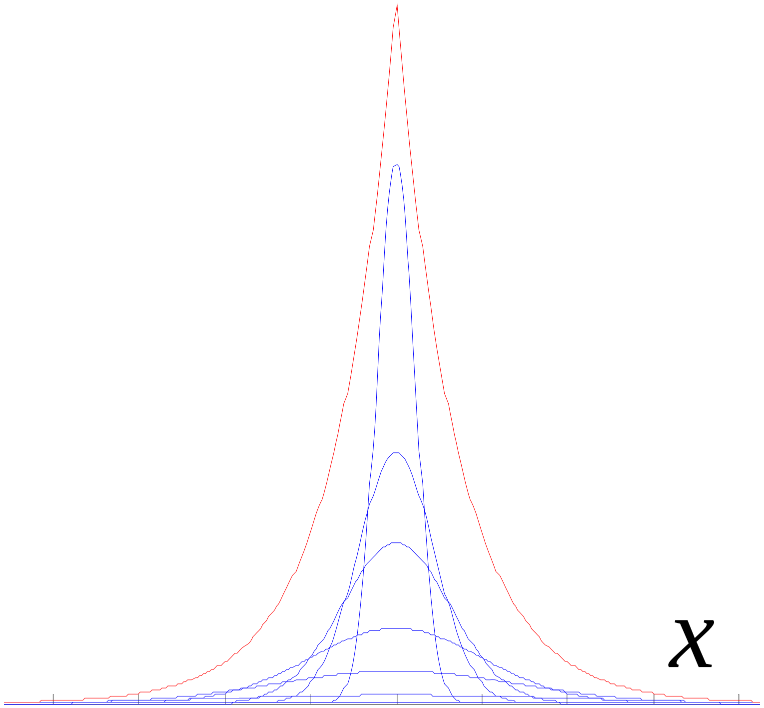
- Images are not Gaussian
  - Local pixel differences or prediction errors possess Laplacian pdf
  - Transform coefficients (DCT, DWT, . . .) possess Laplacian pdf
- Infinite Gaussian mixture model



- Exponentially distributed variance V, mean $E[V]=\sigma^2$

$$p_V(v) = \begin{cases} \frac{1}{\sigma^2} e^{-v/\sigma^2}; & v \geq 0 \\ 0 & else \end{cases}$$

Infinite Gaussian mixture modeling



$$p_X(x) = \int_0^\infty \frac{1}{\sqrt{2\pi v}} \cdot e^{-x^2/2v} \, \frac{1}{\sigma^2} e^{-v/\sigma^2} dv$$

$$= \sqrt{\frac{1}{2\sigma^2}} \cdot e^{-\sqrt{2} \cdot |x| / \sigma}$$

- Infinite mixture of Gaussians with zero mean and exponential variance distribution yields a Laplacian!
- Elegant explanation of ubiquitous Laplacian pdfs in images

Rate reduction due to source splitting

- Same distortion $D \le v$ for each sub-source (implies high-rate assumption)
- Rate-distortion function, if σ^2 is known

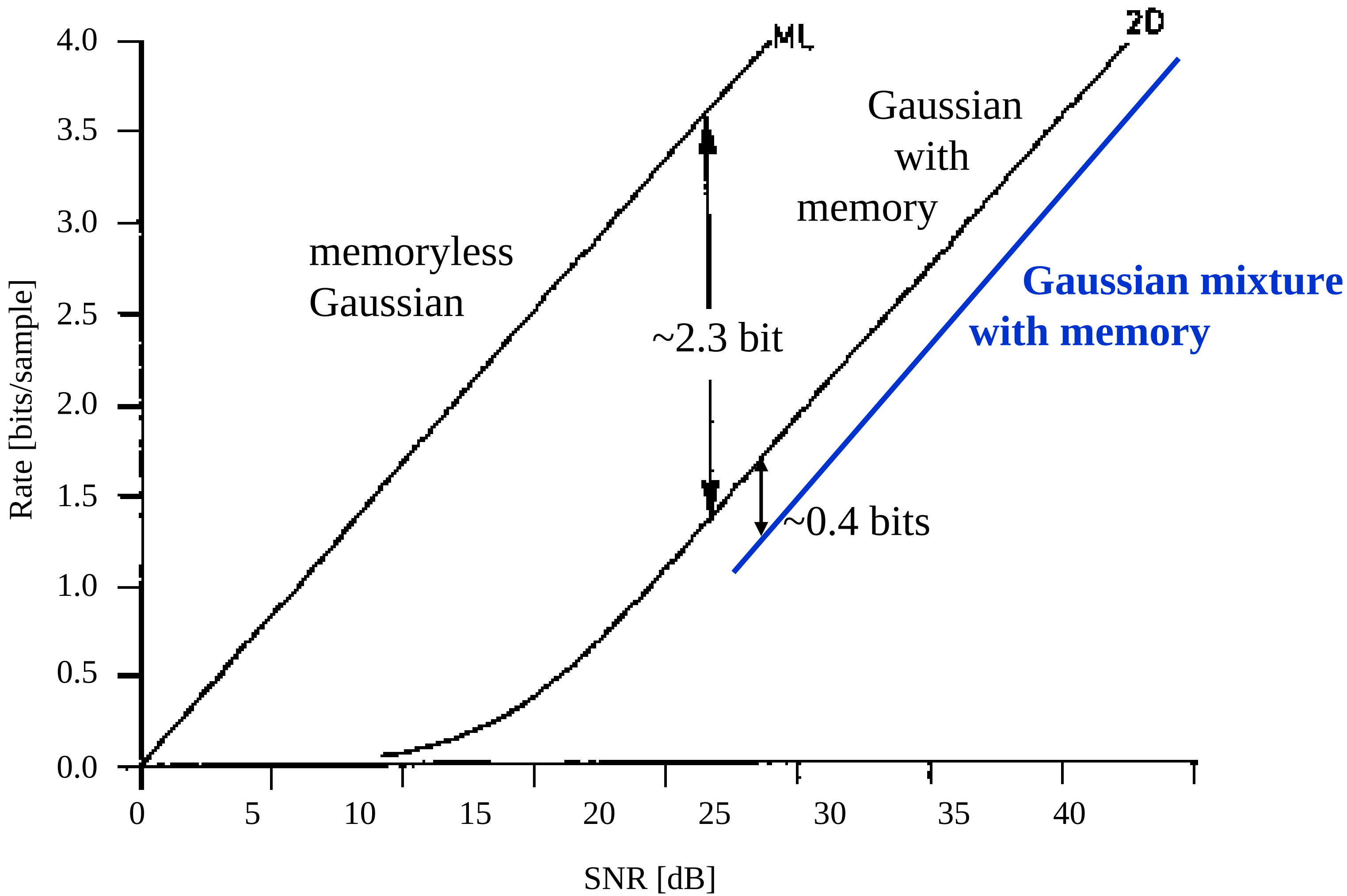
$$R(D) = \int_0^\infty \frac{1}{2}\log_2\left(\frac{v}{D}\right)\frac{1}{\sigma^2}e^{-v/\sigma^2}dv = -\frac{\gamma}{2\ln 2} + \frac{1}{2}\log_2\left(\frac{\sigma^2}{D}\right)$$

Euler's constant $\gamma = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k} - \ln n \simeq 0.577216$

- Rate reduction relative to Gaussian with variance σ^2

$$\frac{\gamma}{2\ln 2} \simeq 0.416373 \text{ bits}$$

Rate distortion function for images



Summary: rate distortion theory

- Rate-distortion theory: minimum transmission bit-rate for given distortion
- Shannon Lower Bound assumes statistical independence between distortion and reconstructed signal
- *R(D)* for memoryless Gaussian source and MSE: 6 dB/bit
- *R(D)* for Gaussian source with memory and MSE: encode spectral components independently, introduce white noise, suppress small spectral components
- Theoretical gain ~2.3 bits/sample by exploiting spatial redundancy in the video signal
- Additional theoretical source splitting gain of ~0.4 bits/sample with Gaussian mixture model

Reading

- Taubman, Marcellin, Chapter 3.1, “Rate-Distortion Theory”
- T. M. Cover, J. A. Thomas, “Elements of Information Theory,” Wiley 1991, Chapter 13: Rate Distortion Theory”