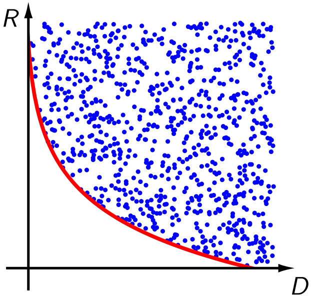
Rate-Distortion Theory III



Differential Entropy Rate for Stationary Gaussian

- Given: Autocovariance function: $\phi(k) = \phi_k = \mathrm{E}\{\,(S_n - \mu)(S_{n+k} - \mu)\,\}$

➡ Differential entropy rate for stationary Gaussian (all $\boldsymbol{C}_N$ are given by ϕ_k)

$$\begin{aligned}\bar{h}^G(\boldsymbol{S}) &= \lim_{N\to\infty} \frac{h_N^G(\boldsymbol{S})}{N} = \lim_{N\to\infty} \frac{1}{N} \cdot \frac{N}{2} \log_2 \left(2\pi e \sqrt[N]{|\boldsymbol{C}_N|} \right) \\ &= \frac{1}{2} \log_2(2\pi e) + \lim_{N\to\infty} \frac{1}{2N} \log_2 |\boldsymbol{C}_N| \end{aligned} \qquad (1)$$

- Determinate of a matrix = Product of its eigenvalues

$$\begin{aligned}\bar{h}^G(\boldsymbol{S}) &= \frac{1}{2} \log_2(2\pi e) + \lim_{N\to\infty} \frac{1}{2N} \log_2 \left(\prod_{k=0}^{N-1} \xi_k^{(N)} \right) \\ &= \frac{1}{2} \log_2(2\pi e) + \frac{1}{2} \left(\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \log_2 \xi_k^{(N)} \right) \end{aligned} \qquad (2)$$

with $\{\xi_k^{(N)}\}$ being the N eigenvalues of a matrix $\boldsymbol{C}_N$

Differential Entropy Rate for Stationary Gaussian

Theorem of Grenander and Szegö

- Given the conditions
  - $\boldsymbol{C}_N$ is a sequence of Hermitian Toeplitz matrices with elements ϕ_k on the k-th diagonal (Hermitian: Matrix is equal to its conjugate transpose)
  - The Fourier series

$$\Phi(\omega) = \sum_{k=-\infty}^{\infty} \phi_k \cdot e^{-\mathrm{i}\,\omega\, k} \tag{3}$$

    has a finite infimum $\Phi_{\inf}$ and a finite supremum $\Phi_{\sup}$
  - The function G is continuous in the interval $[\Phi_{\inf}; \Phi_{\sup}]$

➡ The following expression holds

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} G\left(\xi_k^{(N)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(\,\Phi(\omega)\,)\,\mathrm{d}\omega \tag{4}$$

where $\{\xi_k^{(N)}\}$ represent the N eigenvalues of the N-th matrix $\boldsymbol{C}_N$

Differential Entropy Rate for Stationary Gaussian

- Apply theorem of Grenander and Szegö with $G(x) = \log_2 x$

$$\bar{h}^G(\boldsymbol{S}) = \frac{1}{2}\log_2(2\pi e) + \frac{1}{2}\left(\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\log_2 \xi_k^{(N)}\right)$$

$$= \frac{1}{2}\log_2(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2 \Phi_{SS}(\omega)\,\mathrm{d}\omega \qquad (5)$$

- $\Phi_{SS}(\omega)$ = **Power Spectral Density** for zero-mean process

Differential Entropy Rate for Stationary Gaussian

$$\bar{h}^G(\boldsymbol{S}) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2\Big(2\pi e\,\Phi_{SS}(\omega)\Big)\,\mathrm{d}\omega \qquad (6)$$

with $\Phi_{SS}(\omega)$ being the Fourier series of the autocovariance function ϕ_k

$$\Phi_{SS}(\omega) = \sum_{k=-\infty}^{\infty}\phi_k \cdot e^{-\mathrm{i}\,\omega\,k} \quad \text{with} \quad \phi_k = \mathrm{E}\{\,(S_n-\mu)(S_{n+k}-\mu)\,\} \qquad (7)$$

Stationary Gauss-Markov Process: Power Spectral Density

- Autocovariance function for AR(1) processes is given by

$$\phi_k = \sigma^2 \cdot \varrho^{|k|} \qquad \text{with} \qquad \varrho = \frac{1}{\sigma^2} \mathrm{E}\{ (S_n - \mu)(S_{n+1} - \mu) \} \tag{8}$$

➡ Power spectral density for Gauss-Markov process

$$\Phi_{SS}^{GM}(\omega) = \sum_{k=-\infty}^{\infty} \phi_k \cdot e^{-\mathrm{i}\,\omega\,k} = \sigma^2 \sum_{k=-\infty}^{\infty} \varrho^{|k|} \cdot e^{-\mathrm{i}\,\omega\,k}$$

$$= \sigma^2 \left(\sum_{k=0}^{\infty} \left(\varrho e^{-\mathrm{i}\,\omega}\right)^k + \sum_{k=0}^{\infty} \left(\varrho e^{\mathrm{i}\,\omega}\right)^k - 1 \right)$$

- Next: Use formula for geometric series with $|a| < 1$

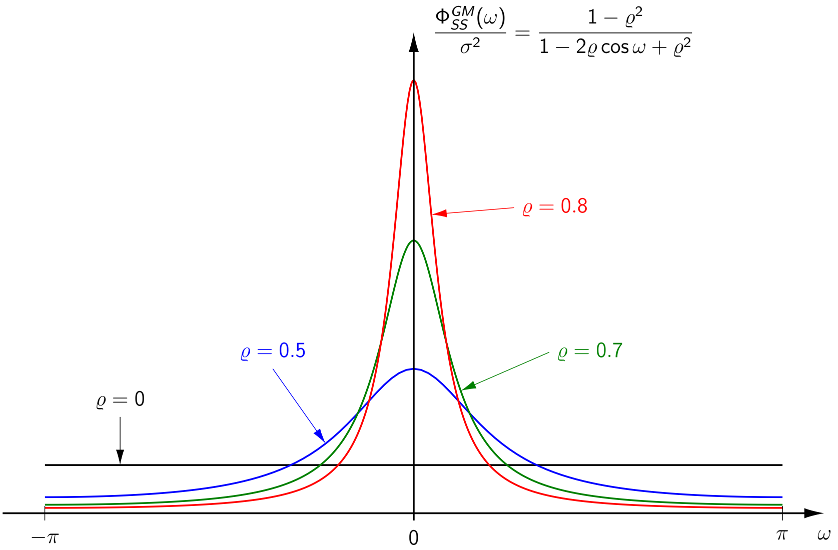
$$\sum_{k=0}^{\infty} a^k = \frac{1}{1-a} \tag{9}$$

Stationary Gauss-Markov Process: Power Spectral Density

- Power spectral density for Gauss-Markov process

$$\begin{aligned}
\Phi_{SS}^{GM}(\omega) &= \sigma^2 \left(\sum_{k=0}^{\infty} \left(\varrho e^{-\mathrm{i}\,\omega}\right)^k + \sum_{k=0}^{\infty} \left(\varrho e^{\mathrm{i}\,\omega}\right)^k - 1 \right) \qquad \left[\sum_{k=0}^{\infty} a^k = \frac{1}{1-a} \right] \\
&= \sigma^2 \left(\frac{1}{1-\varrho e^{-\mathrm{i}\,\omega}} + \frac{1}{1-\varrho e^{\mathrm{i}\,\omega}} - 1 \right) \\
&= \sigma^2 \, \frac{(1-\varrho e^{\mathrm{i}\,\omega}) + (1-\varrho e^{-\mathrm{i}\,\omega}) - (1-\varrho e^{-\mathrm{i}\,\omega})(1-\varrho e^{\mathrm{i}\,\omega})}{(1-\varrho e^{-\mathrm{i}\,\omega})(1-\varrho e^{\mathrm{i}\,\omega})} \\
&= \sigma^2 \, \frac{(1-\varrho e^{\mathrm{i}\,\omega}) + (1-\varrho e^{-\mathrm{i}\,\omega}) - (1-\varrho e^{\mathrm{i}\,\omega} - \varrho e^{-\mathrm{i}\,\omega} + \varrho^2)}{1-\varrho e^{\mathrm{i}\,\omega} - \varrho e^{-\mathrm{i}\,\omega} + \varrho^2} \\
&= \frac{\sigma^2\,(1-\varrho^2)}{1 - 2\varrho\cos\omega + \varrho^2} \qquad (10)
\end{aligned}$$

Stationary Gauss-Markov Process: Power Spectral Density



Stationary Gauss-Markov Process: Differential Entropy Rate

- Differential entropy rate for stationary Gauss-Markov process

$$
\begin{aligned}
\bar{h}^{GM}(\boldsymbol{S}) &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \log_2\left(2\pi e \cdot \Phi_{SS}^{GM}(\omega)\right)\,\mathrm{d}\omega \\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \log_2 \frac{2\pi e\,\sigma^2\,(1-\varrho^2)}{1-2\varrho\cos\omega+\varrho^2}\,\mathrm{d}\omega \\
&= \frac{1}{4\pi}\log_2(\,2\pi e\,\sigma^2\,(1-\varrho^2)\,)\int_{-\pi}^{\pi}\mathrm{d}\omega \\
&\qquad - \frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2(1-2\varrho\cos\omega+\varrho^2)\,\mathrm{d}\omega \\
&= \frac{1}{2}\log_2(\,2\pi e\,\sigma^2\,(1-\varrho^2)\,) \\
&\qquad - \frac{1}{2\pi}\int_{0}^{\pi}\log_2(1-2\varrho\cos\omega+\varrho^2)\,\mathrm{d}\omega
\end{aligned}
$$

Stationary Gauss-Markov Process: Differential Entropy Rate

- Continue: Differential entropy rate for stationary Gauss-Markov

$$\bar{h}^{GM}(\boldsymbol{S}) = \frac{1}{2}\log_2\left(2\pi e\,\sigma^2\,(1-\varrho^2)\right) - \frac{1}{2\pi\,\ln 2}\int_0^{\pi} \ln(1 - 2\varrho\,\cos\omega + \varrho^2)\,\mathrm{d}\omega$$

- Dini's integral (see, for example, P. Nahin, "Inside Interesting Integrals")

$$\int_0^{\pi} \ln(1 - 2\alpha\,\cos\omega + \alpha^2)\,\mathrm{d}\omega = \begin{cases} 0 & : \quad \alpha^2 \leq 1 \\ \pi\ln\alpha^2 & : \quad \alpha^2 > 1 \end{cases} \tag{11}$$

Differential Entropy Rate for Gauss-Markov

- Stationary Gauss-Markov with variance σ^2 and correlation coefficient ϱ

$$\bar{h}^{GM}(\boldsymbol{S}) = \frac{1}{2}\log_2\left(2\pi e\,\sigma^2\,(1-\varrho^2)\right) \tag{12}$$

Summary: Differential Entropy Rate for Gaussians

Stationary Gaussian processes

- Statistical properties are given by auto-covariance function

$$\phi_k = \phi(k) = \mathrm{E}\{\, (S_n - \mu)(S_{n+k} - \mu) \,\}$$

➡ Differential entropy rate for stationary Gaussian processes

$$\bar{h}^G(\boldsymbol{S}) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2\Big(2\pi e\, \Phi_{SS}(\omega)\Big)\, \mathrm{d}\omega, \qquad \Phi_{SS}(\omega) = \sum_{k=-\infty}^{\infty} \phi_k e^{-\mathrm{i}\omega k}$$

- Special stationary Gaussian processes:

$$\text{Gauss-Markov:} \quad \bar{h}^{GM}(\boldsymbol{S}) \;=\; \frac{1}{2} \log_2\left(\, 2\pi e\, \sigma^2 \left(1 - \varrho^2\right)\right)$$

$$\text{Gaussian iid:} \quad \bar{h}^{Giid}(\boldsymbol{S}) \;=\; \frac{1}{2} \log_2\left(\, 2\pi e\, \sigma^2 \,\right)$$

- Inequalities for differential entropy rate

$$\text{Given } \phi_k \text{ or } \Phi_{SS}(\omega): \qquad \bar{h}(\boldsymbol{S}) \;\le\; \bar{h}^G(\boldsymbol{S})$$

$$\text{Given variance } \sigma^2: \qquad \bar{h}(\boldsymbol{S}) \;\le\; \bar{h}^{Giid}(\boldsymbol{S})$$

Shannon Lower Bound

- Lower bound for rate-distortion function $R(D)$

$$
\begin{aligned}
R(D) &= \lim_{N\to\infty} \inf_{g_N:\,\delta_N(g_N)\le D} \frac{I_N(\boldsymbol{S};\boldsymbol{S}')}{N} \\
&= \lim_{N\to\infty} \inf_{g_N:\,\delta_N(g_N)\le D} \frac{h_N(\boldsymbol{S}) - h_N(\boldsymbol{S}\,|\,\boldsymbol{S}')}{N} \\
&= \lim_{N\to\infty} \frac{h_N(\boldsymbol{S})}{N} - \lim_{N\to\infty} \sup_{g_N:\,\delta_N(g_N)\le D} \frac{h_N(\boldsymbol{S}\,|\,\boldsymbol{S}')}{N} \\
\text{(a)} \quad &= \bar{h}(\boldsymbol{S}) - \lim_{N\to\infty} \sup_{g_N:\,\delta_N(g_N)\le D} \frac{h_N(\boldsymbol{S}-\boldsymbol{S}'\,|\,\boldsymbol{S}')}{N} \\
\text{(b)} \quad &\ge \bar{h}(\boldsymbol{S}) - \lim_{N\to\infty} \sup_{g_N:\,\delta_N(g_N)\le D} \frac{h_N(\boldsymbol{S}-\boldsymbol{S}')}{N} \qquad (13)
\end{aligned}
$$

(a) Modification of the mean does not change differential entropy

(b) Conditioning does not increase differential entropy

Shannon Lower Bound

Lower bound of rate-distortion function $R(D)$: **Shannon lower bound** $R_L(D)$

$$R(D) \geq R_L(D), \qquad \boxed{R_{\mathrm{L}}(D) = \bar{h}(\boldsymbol{S}) - \lim_{N\to\infty} \sup_{g_N:\delta_N(g_N)\leq D} \frac{h_N(\boldsymbol{S}-\boldsymbol{S}')}{N}} \qquad (14)$$

- When does it coincide with rate-distortion function?
  - In the derivation, we used the inequality: $h_N(\boldsymbol{S}-\boldsymbol{S}' \,|\, \boldsymbol{S}') \leq h_N(\boldsymbol{S}-\boldsymbol{S}')$
  - ➡ Lower bound $R_L(D)$ is equal to rate-distortion function $R(D)$ if and only if the approximation error $\boldsymbol{Z} = \boldsymbol{S} - \boldsymbol{S}'$ is independent of the reconstruction $\boldsymbol{S}'$

- Can we calculate it?
  - For many distortion measures $d_N(\boldsymbol{s}, \boldsymbol{s}')$, it can be calculated
  - ➡ Particularly, it can be calculated for all p-norm distortion measures

$$d_N(\boldsymbol{s}, \boldsymbol{s}') = ||\boldsymbol{s} - \boldsymbol{s}'||_p^p = \frac{1}{N} \sum_{k=0}^{N-1} |s_k - s_k'|^p$$

Shannon Lower Bound for MSE Distortion

- Consider random process for approximation error: $\boldsymbol{Z} = \boldsymbol{S} - \boldsymbol{S'}$

➡ MSE distortion D is given by

$$D = \mathrm{E}\{\, (S - S')^2 \,\} = \mathrm{E}\{\, Z^2 \,\} = \sigma_Z^2 + \mu_Z^2 \tag{15}$$

➡ Consider supremum of N-th order differential entropy

$$\begin{aligned} \sup_{g_N:\, \delta_N(g_N) \le D} h_N(\boldsymbol{S} - \boldsymbol{S'}) &= \sup_{f_{\boldsymbol{Z}}:\, \sigma_Z^2 + \mu_Z^2 \le D} h_N(\boldsymbol{Z}) \\ \text{(a)} \quad &= \sup_{f_{\boldsymbol{Z}}:\, \sigma_Z^2 = D} h_N(\boldsymbol{Z}) \\ \text{(b)} \quad &= h_N^{Giid}(\boldsymbol{Z} \,|\, \sigma_Z^2 = D) \\ &= \frac{N}{2} \log_2(2\pi e\, D) \end{aligned} \tag{16}$$

(a) Modification of the mean does not change differential entropy

(b) For given variance, Gaussian iid has maximum differential entropy

Shannon Lower Bound for MSE Distortion

- Using derived supremum of N-th order differential entropy

$$R_{\mathrm{L}}(D) = \bar{h}(\boldsymbol{S}) - \lim_{N\to\infty} \sup_{g_N:\,\delta_N(g_N)\le D} \frac{h_N(\boldsymbol{S}-\boldsymbol{S}')}{N}$$

$$= \bar{h}(\boldsymbol{S}) - \lim_{N\to\infty} \frac{1}{N}\cdot\frac{N}{2}\log_2(2\pi e\, D) \qquad (17)$$

Shannon Lower Bound: MSE Distortion

- Shannon lower bound as rate-distortion function

$$\boxed{R_{\mathrm{L}}(D) = \bar{h}(\boldsymbol{S}) - \frac{1}{2}\log_2(2\pi e\, D)} \qquad (18)$$

- Shannon lower bound as distortion-rate function

$$\boxed{D_{\mathrm{L}}(R) = \frac{1}{2\pi e}\cdot 2^{2\bar{h}(\boldsymbol{S})}\cdot 2^{-2R}} \qquad (19)$$

General Form of Shannon Lower Bound (MSE)

Differential entropy of scaled random process

- Given: Random process $\boldsymbol{X}$ with N-th order differential entropy $h_N(\boldsymbol{X})$
- Consider scaled random process $\boldsymbol{Y} = a\boldsymbol{X}$ with $a > 0$

➡ Cumulative distribution and probability density function (N-th order)

$$F_{\boldsymbol{Y}}(\boldsymbol{y}) = \mathrm{P}(\boldsymbol{Y} < \boldsymbol{y}) = \mathrm{P}(a\boldsymbol{X} < \boldsymbol{y}) = \mathrm{P}\left(\boldsymbol{X} < a^{-1}\boldsymbol{y}\right) = F_{\boldsymbol{X}}(a^{-1}\boldsymbol{y})$$

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{\partial}{\partial \boldsymbol{y}} F_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{\partial}{\partial \boldsymbol{y}} F_{\boldsymbol{X}}(a^{-1}\boldsymbol{y}) = a^{-1} \cdot f_{\boldsymbol{X}}(a^{-1}\boldsymbol{y})$$

➡ N-th order differential entropy rate

$$\begin{aligned} h_N(\boldsymbol{Y}) &= \mathrm{E}\{-\log_2 f_{\boldsymbol{Y}}(\boldsymbol{Y})\} = \mathrm{E}\left\{-\log_2\left(a^{-1} \cdot f_{\boldsymbol{X}}(a^{-1}\boldsymbol{Y})\right)\right\} \\ &= \mathrm{E}\{-\log_2 f_{\boldsymbol{X}}(\boldsymbol{X}) + \log_2 a\} \\ &= h_N(\boldsymbol{X}) + \log_2 a \end{aligned}$$

➡ N-th order differential entropy rate

$$\bar{h}(\boldsymbol{S}) = \bar{h}(\boldsymbol{S}/\sigma) + \frac{1}{2}\log_2 \sigma^2 \qquad \text{[note: } \boldsymbol{S}/\sigma \text{ has unit variance]} \qquad (20)$$

General Form of Shannon Lower Bound (MSE)

- Shannon Lower Bound as distortion-rate function

$$
\begin{aligned}
D_L(R) &= \frac{1}{2\pi e} \cdot 2^{2\bar{h}(\boldsymbol{S})} \cdot 2^{-2R} \\
&= \frac{1}{2\pi e} \cdot 2^{2\bar{h}(\boldsymbol{S}/\sigma)+\log_2 \sigma^2} \cdot 2^{-2R} \\
&= \frac{1}{2\pi e} \cdot 2^{2\bar{h}(\boldsymbol{S}/\sigma)} \cdot 2^{\log_2 \sigma^2} \cdot 2^{-2R} \\
&= \varepsilon^2 \cdot \sigma^2 \cdot 2^{-2R} \qquad \text{with} \qquad \varepsilon^2 = \frac{1}{2\pi e} \cdot 2^{2\bar{h}(\boldsymbol{S}/\sigma)} \qquad (21)
\end{aligned}
$$

➡ The factor ε^2 does only depend on the shape of the pdf, not on its variance

➡ General form of Shannon lower bound for MSE distortion

$$
\begin{aligned}
D_L(R) &= \varepsilon^2 \cdot \sigma^2 \cdot 2^{-2R} \\
R_L(D) &= \frac{1}{2} \log_2 \left(\frac{\varepsilon^2 \cdot \sigma^2}{D} \right)
\end{aligned}
\qquad \text{with} \qquad \varepsilon^2 = \frac{1}{2\pi e} \cdot 2^{2\bar{h}(\boldsymbol{S}/\sigma)}
$$

Shannon Lower Bound (MSE): Stationary Gaussian

- Differential entropy rate for stationary Gaussian sources

$$\bar{h}^G(\boldsymbol{S}) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2\Big(2\pi e\, \Phi_{SS}(\omega)\Big)\, \mathrm{d}\omega \qquad (22)$$

➡ **Shannon lower bound for stationary Gaussian sources** (MSE distortion)

$$\begin{aligned} R_{\mathrm{L}}^G(D) &= \bar{h}^G(\boldsymbol{S}) - \frac{1}{2}\log_2(2\pi e\, D) \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2\Big(2\pi e\, \Phi_{SS}(\omega)\Big)\, \mathrm{d}\omega - \frac{1}{2}\log_2(2\pi e\, D) \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2\Big(2\pi e\, \Phi_{SS}(\omega)\Big)\, \mathrm{d}\omega - \frac{1}{4\pi}\log_2(2\pi e\, D) \int_{-\pi}^{\pi} \mathrm{d}\omega \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2 \frac{\Phi_{SS}(\omega)}{D}\, \mathrm{d}\omega \qquad (23) \end{aligned}$$

Shannon Lower Bound (MSE): Stationary Gauss-Markov

- Differential entropy rate for stationary Gauss-Markov sources

$$\bar{h}^{GM}(\boldsymbol{S}) = \frac{1}{2}\log_2\left(\, 2\pi e\,\sigma^2\left(1-\varrho^2\right)\right) \tag{24}$$

➡ **Shannon lower bound for stationary Gauss-Markov sources** (MSE)

$$\begin{aligned} R_{\mathrm{L}}^{GM}(D) &= \bar{h}^{GM}(\boldsymbol{S}) - \frac{1}{2}\log_2(2\pi e\, D) \\ &= \frac{1}{2}\log_2\left(2\pi e\,\sigma^2\left(1-\varrho^2\right)\right) - \frac{1}{2}\log_2(2\pi e\, D) \\ &= \frac{1}{2}\log_2\frac{\sigma^2\left(1-\varrho^2\right)}{D} \end{aligned} \tag{25}$$

➡ **Shannon lower bound as distortion-rate function**

$$D_{\mathrm{L}}^{GM}(R) = \left(1-\varrho^2\right)\sigma^2\, 2^{-2R} \tag{26}$$

Shannon Lower Bound (MSE) for IID Sources

- For iid sources, we have

$$\bar{h}(\boldsymbol{S}) = h(S)$$

Shannon Lower Bound for MSE distortion and IID sources

- Rate-distortion function

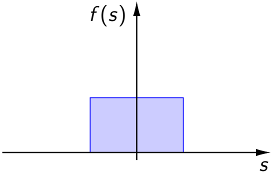
$$R_{\mathrm{L}}(D) = h(S) - \frac{1}{2}\log_2(2\pi e\,D) \tag{27}$$

- Distortion-rate function

$$D_{\mathrm{L}}(R) = \frac{1}{2\pi e} \cdot 2^{2h(S)} \cdot 2^{-2R} \tag{28}$$

Shannon Lower Bound (MSE) for Selected IID Sources

Uniform

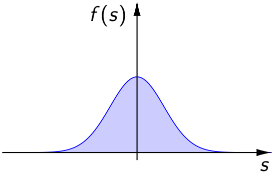


$$f(s) = \begin{cases} \frac{1}{2a} & : \quad |s| \leq a \\ 0 & : \quad \text{otherwise} \end{cases}$$

$$h(S) = \frac{1}{2} \log_2(12\, \sigma_S^2)$$

$$D_{\mathrm{L}}(R) = \underbrace{\frac{6}{\pi e}}_{\approx 0.7} \sigma_S^2 \cdot 2^{-2R}$$

Gaussian

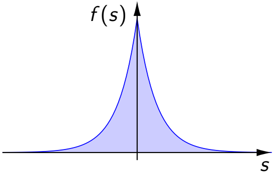


$$f(s) = \frac{1}{\sigma_S \sqrt{2\pi}}\, e^{-\frac{s^2}{2\sigma_S^2}}$$

$$h(S) = \frac{1}{2} \log_2(2\pi e\, \sigma_S^2)$$

$$D_{\mathrm{L}}(R) = \sigma_S^2 \cdot 2^{-2R}$$

Laplacian



$$f(s) = \frac{1}{\sigma_S \sqrt{2}}\, e^{-\frac{\sqrt{2}}{\sigma_S}|s|}$$

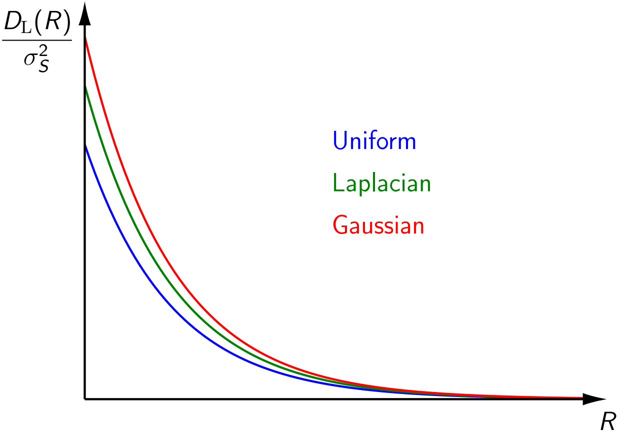
$$h(S) = \frac{1}{2} \log_2(2\, e^2\, \sigma_S^2)$$

$$D_{\mathrm{L}}(R) = \underbrace{\frac{e}{\pi}}_{\approx 0.865} \sigma_S^2 \cdot 2^{-2R}$$

Shannon Lower Bound (MSE) for Selected IID Sources

- Shannon lower bound for MSE and IID sources

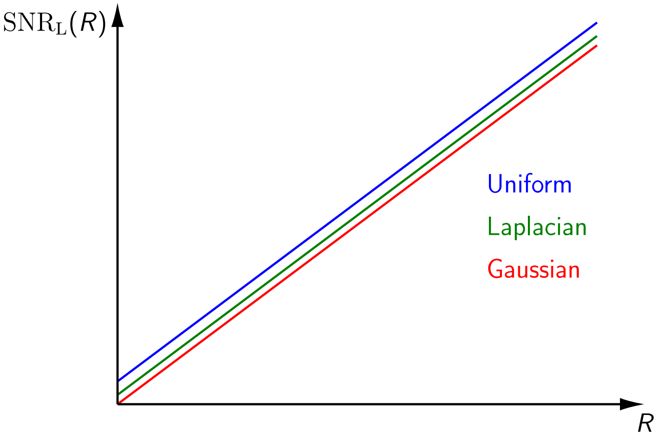
$$D_{\mathrm{L}}(R) = \varepsilon^2 \cdot \sigma_S^2 \cdot 2^{-2R}$$



Shannon Lower Bound (MSE) for Selected IID Sources

- Shannon lower bound as signal-to-noise ratio (SNR)

$$\mathrm{SNR} = 10\,\log_{10}\frac{\sigma_S^2}{D} \quad\Longrightarrow\quad \mathrm{SNR_L}(R) = \underbrace{(20\log_{10}2)}_{\approx 6.02}\cdot R - (10\log_{10}\varepsilon^2)$$



Asymptotic Tightness of Shannon Lower Bound

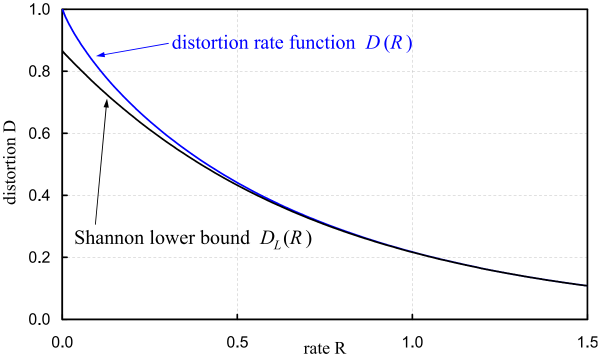
- Shannon lower bound $R_{\mathrm{L}}(D)$ approaches rate-distortion function $R(D)$ for high rates R (or small distortions D)

$$\lim_{D\to 0} R(D) - R_{\mathrm{L}}(D) = 0, \qquad \lim_{R\to\infty} D(R) - D_{\mathrm{L}}(R) = 0 \qquad (29)$$

- Often: Very good approximation of $R(D)$ for bit rates $R > 1$ bit per sample

Example:

Laplacian IID



High-Rate Approximations of Rate-Distortion Function

- IID sources

$$\text{Gaussian:} \quad D_L(R) = \sigma^2 \cdot 2^{-2R}$$

$$\text{Laplacian:} \quad D_L(R) = \frac{e}{\pi} \cdot \sigma^2 \cdot 2^{-2R}$$

$$\text{Uniform:} \quad D_L(R) = \frac{6}{\pi e} \cdot \sigma^2 \cdot 2^{-2R}$$

- Gaussian sources

$$\text{Gauss-Markov:} \quad D_L(R) = (1 - \varrho^2) \cdot \sigma^2 \cdot 2^{-2R}$$

$$\text{Stationary Gauss:} \quad D_L(R) = \frac{1}{2\pi e} \cdot 2^{\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \Phi_{SS}(\omega)\, \mathrm{d}\omega \right)} \cdot 2^{-2R}$$

➡ **Can use these approximations for evaluating the coding efficiency of lossy coding techniques for high rates**

Rate-Distortion Function for Gaussian IID Sources and MSE

- N-th order probability density function for Gaussian iid sources

$$f_{\boldsymbol{S}}^{(N)}(\boldsymbol{s}) = \prod_{k=0}^{N-1} f_S(s_k) \qquad \text{with} \qquad f_S(s) = \frac{1}{\sqrt{2\pi\,\sigma_S^2}}\, e^{-\frac{(s-\mu_S)^2}{2\,\sigma_S^2}} \tag{30}$$

Theorem

For MSE distortion and Gaussian iid sources:

➡ Rate-distortion function $R(D)$ is equal to Shannon lower bound $R_L(D)$

$$R(D) = R_L(D) \qquad \Longrightarrow \qquad R(D) = \inf_{g:\,\delta(g)\le D} I(g) = \frac{1}{2}\log_2 \frac{\sigma^2}{D} \tag{31}$$

- How can we proof it?
  - ➡ Need to show that Shannon lower bound is achievable
  - ➡ Need to find a conditional pdf $g_{S'|S}(s' \,|\, s)$ for which the Shannon lower bound is achieved

Proof: $R(D) = R_L(D)$ for Gaussian and MSE Distortion

What do we know?

1 Independence of approximation error from reconstruction

- In derivation for general Shannon lower bound, we use the inequality

$$h_N(\boldsymbol{S} - \boldsymbol{S}' \,|\, \boldsymbol{S}') \le h_N(\boldsymbol{S} - \boldsymbol{S}')$$

➡ Equality: **Approximation error $\boldsymbol{Z} = \boldsymbol{S} - \boldsymbol{S}'$ must be independent of $\boldsymbol{S}'$**

2 Pdf of approximation error

- Supremum of differential entropy for given MSE D is given by Gaussian iid

$$\sup_{\mathrm{E}\{Z^2\} \le D} h_N(\boldsymbol{Z}) = h_N^{Giid}(\boldsymbol{Z} \,|\, \sigma_Z^2 = D) = \frac{N}{2} \log_2(2\pi e D)$$

➡ **Approximation error $\boldsymbol{Z} = \boldsymbol{S} - \boldsymbol{S}'$ must be Gaussian iid with $\sigma_Z^2 = D$**

➡ **Conditional pdf $g_{Z|S'}(z \,|\, s')$ must have the form**

$$g_{Z|S'}(z \,|\, s') = f_Z(z) = \frac{1}{\sqrt{2\pi D}}\, e^{-\frac{z^2}{2D}} \tag{32}$$

Proof: $R(D) = R_L(D)$ for Gaussian and MSE Distortion

- **We know:** Shannon lower bound can only be achieved if

$$g_{Z|S'}(z\,|\,s') = f_Z(z) = \frac{1}{\sqrt{2\pi D}}\, e^{-\frac{z^2}{2D}} \qquad \text{where} \qquad Z = S - S'$$

- Given $g_{Z|S'}(z\,|\,s')$, we also know $g_{S|S'}(s\,|\,s')$

$$Z = S - S' \quad \Longrightarrow \quad g_{S|S'}(s\,|\,s') = g_{Z|S}(s - s'\,|\,s') = \frac{1}{\sqrt{2\pi D}}\, e^{-\frac{(s-s')^2}{2D}}$$

- Sought pdf $g_{S'|S}(s'\,|\,s)$ can be determined using Bayes rule

$$g_{S'|S}(s'\,|\,s) = g_{S|S'}(s\,|\,s') \cdot \frac{f_{S'}(s')}{f_S(s)}$$

- If $g_{S|S'}(s|s')$, $f_{S'}(s')$, and $f_S(s)$ are valid pdf's, $g_{S'|S}(s'|s)$ is also a valid pdf

➡ **Need to verify that our choice for $g_{Z|S'}(z\,|\,s')$ yields a valid pdf $f_{S'}(s')$**

Proof: $R(D) = R_L(D)$ for Gaussian and MSE Distortion

- Random variables S can be represented as

$$S = Z + S' \qquad \Longleftrightarrow \qquad Z = S - S' \tag{33}$$

➡ Have shown: Random variables Z and S' must be independent of each other

- Assume $f_Z(z)$ and $f_{S'}(s')$ are given

  ➡ $f_S(s)$ = Pdf for sum of two independent random variables Z and S'

$$f_S(s) = \int_{-\infty}^{\infty} f_Z(z)\, f_{S'}(s-z)\,\mathrm{d}z \;=\; (f_Z * f_{S'})(s) \tag{34}$$

  ➡ $f_S(s)$ = Convolution of the source pdfs $f_Z(z)$ and $f_{S'}(s')$

- Known: Convolution of two Gaussians $\mathcal{N}(\mu_1, \sigma_1^2)$ and $\mathcal{N}(\mu_2, \sigma_2^2)$

  ➡ Another Gaussian $\mathcal{N}(\mu, \sigma^2)$ with

$$\mu = \mu_1 + \mu_2 \qquad \text{and} \qquad \sigma^2 = \sigma_1^2 + \sigma_2^2$$

Proof: $R(D) = R_L(D)$ for Gaussian and MSE Distortion

- Given: Gaussian pdfs for source S and difference signal $Z = S - S'$

$$f_S(s) = \frac{1}{\sqrt{2\pi\,\sigma_S^2}}\, e^{-\frac{(s-\mu_S)^2}{2\,\sigma_S^2}} \qquad \text{and} \qquad f_Z(z) = \frac{1}{\sqrt{2\pi\,D}}\, e^{-\frac{z^2}{2\,D}}$$

➡ Pdf $f_{S'}(s')$ of reconstruction S' is also Gaussian with $\mu_{S'}$ and $\sigma_{S'}^2$ given by

$$\begin{aligned} \mu_S &= \mu_Z + \mu_{S'} = \mu_{S'} & \Longrightarrow \quad \mu_{S'} &= \mu_S \\ \sigma_S^2 &= \sigma_Z^2 + \sigma_{S'}^2 = D + \sigma_{S'}^2 & \Longrightarrow \quad \sigma_{S'}^2 &= \sigma_S^2 - D \end{aligned}$$

➡ Pdf $f_{S'}(s')$ of reconstruction S'

$$f_{S'}(s') = \frac{1}{\sqrt{2\pi\,(\sigma_S^2 - D)}}\, e^{-\frac{(s'-\mu_S)^2}{2\,(\sigma_S^2 - D)}} \tag{35}$$

➡ Valid pdf (no negative values): **Choice $g_{Z|S'}(z\,|\,s')$ is valid**

Rate-Distortion Function (MSE) for Gaussian IID Sources

Double check rate-distortion function $R(D)$ for our choice

$$g_{S-S'|S'}(s - s' \mid s') = f_Z(z) = \frac{1}{\sqrt{2\pi D}}\, e^{-\frac{z^2}{2D}}$$

➡ Distortion D

$$\delta(g) = \mathrm{E}\{\, (S - S')\,\} = \mathrm{E}\{\, Z^2 \,\} = \sigma_Z^2 = D \tag{36}$$

➡ Mutual information $I(S;S')$

$$\begin{aligned} I(g) &= h(S) - h(S \mid S') = h(S) - h(S - S' \mid S') = h(S) - h(Z) \\ &= \frac{1}{2}\log_2(2\pi e\, \sigma_S^2) - \frac{1}{2}\log_2(2\pi e\, D) = \frac{1}{2}\log_2 \frac{\sigma_S^2}{D} \end{aligned} \tag{37}$$

➡ **For Gaussian IID sources and MSE distortion, the Shannon lower bound coincides with the rate-distortion function**

Rate-Distortion Function (MSE) for Gaussian IID Sources

Gaussian IID sources and MSE distortion

- Distortion-rate and rate-distortion function coincide Shannon lower bound

$$D(R) = \sigma^2 \cdot 2^{-2R}, \qquad R(D) = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2}{D} & : \quad D \leq \sigma^2 \\ 0 & : \quad D > \sigma^2 \end{cases} \tag{38}$$

- Distortion-rate function as signal-to-noise ratio (SNR)

$$\mathrm{SNR}(R) = 10 \log_{10} \frac{\sigma^2}{D(R)} = (20 \log_{10} 2)\, R \approx 6.02\, R \qquad [\mathrm{dB}] \tag{39}$$

For MSE distortion and given signal variance σ^2

- $R(D)$ is maximized for Gaussian IID sources (without proof)

➡ Gaussian IID sources are the hardest to code

Rate-Distortion Function (MSE) for Stationary Gaussian

- Consider: General stationary Gaussian with given autocovariance function

$$\phi_k = \phi(k) = \mathrm{E}\{\, (S_n - \mu)(S_{n+k} - \mu) \,\}$$

- N-th order pdf for stationary Gaussian

$$f_G(\boldsymbol{s}) = \frac{1}{\sqrt{(2\pi)^N \, |\boldsymbol{C}_N|}} \, e^{-\frac{1}{2} (\boldsymbol{s}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{C}_N^{-1} (\boldsymbol{s}-\boldsymbol{\mu})} \tag{40}$$

where covariance matrix $\boldsymbol{C}_N$ is given by autocovariance function ϕ_k

- Derivation of rate-distortion function $R(D)$:
  - Consider vectors $\boldsymbol{S}$ of N consecutive samples
  - Transform vectors $\boldsymbol{S}$: N mutually independent Gaussian RV
    - ➡ Re-use result for Gaussian iid
  - Consider optimal rate distribution between the N Gaussian RV
  - Derive N-th order rate-distortion function $R_N(D)$
  - Consider limit for $N \to \infty$

Eigendecomposition of Covariance Matrix

Eigendecomposition of real symmetric covariance matrix $\boldsymbol{C}_N$

$$\boldsymbol{C}_N = \boldsymbol{A}_N \cdot \boldsymbol{\Lambda}_N \cdot \boldsymbol{A}_N^{\mathrm{T}} \qquad \Longleftrightarrow \qquad \boldsymbol{C}_N^{-1} = \boldsymbol{A}_N \cdot \boldsymbol{\Lambda}_N^{-1} \cdot \boldsymbol{A}_N^{\mathrm{T}}$$

➡ $\boldsymbol{A}_N$: Orthogonal matrix of N unit-norm eigenvectors $\{\boldsymbol{a}_k\}$

$$\boldsymbol{A}_N = \begin{bmatrix} | & | & | & & | \\ \boldsymbol{a}_0 & \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_{N-1} \\ | & | & | & & | \end{bmatrix} \qquad \text{with} \qquad \boldsymbol{A}_N^{-1} = \boldsymbol{A}_N^{\mathrm{T}}$$

➡ $\boldsymbol{\Lambda}_N$: Diagonal matrix with eigenvalues $\{\xi_k\}$ of $\boldsymbol{C}_N$ on its main diagonal

$$\boldsymbol{\Lambda}_N = \begin{bmatrix} \xi_0 & 0 & 0 & \dots & 0 \\ 0 & \xi_1 & 0 & \dots & 0 \\ 0 & 0 & \xi_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \xi_{N-1} \end{bmatrix}, \qquad \boldsymbol{\Lambda}_N^{-1} = \begin{bmatrix} 1/\xi_0 & 0 & 0 & \dots & 0 \\ 0 & 1/\xi_1 & 0 & \dots & 0 \\ 0 & 0 & 1/\xi_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1/\xi_{N-1} \end{bmatrix}$$

Orthogonal Transform of Signal Vectors with N Elements

Concept of transforming signal vectors

- Partition messages into blocks $\boldsymbol{s} = (s_0, s_1, \cdots, s_{N-1})$ of N samples
- Each signal vector $\boldsymbol{s}$ with N samples is transformed according to

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix} = \begin{bmatrix} & & \\ & \boldsymbol{A}_N & \\ & & \end{bmatrix} \cdot \left(\begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_{N-1} \end{bmatrix} - \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} \right)$$

- Code vectors of N transform coefficients $\boldsymbol{u} = (u_0, u_1, \cdots, u_{N-1})$

$$(u_0, u_1, \cdots, u_{N-1}) \quad \mapsto \quad (u'_0, u'_1, \cdots, u'_{N-1})$$

- Get reconstructed signal $\boldsymbol{s}'$ by inverse transform

$$\begin{bmatrix} s'_0 \\ s'_1 \\ \vdots \\ s'_{N-1} \end{bmatrix} = \begin{bmatrix} & & \\ & \boldsymbol{A}_N^{\mathrm{T}} & \\ & & \end{bmatrix} \cdot \begin{bmatrix} u'_0 \\ u'_1 \\ \vdots \\ u'_{N-1} \end{bmatrix} + \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix}$$

Effect of Transform on MSE Distortion

- Transform and inverse transform

$$\text{before coding:} \quad \boldsymbol{u} = \boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{s} - \boldsymbol{\mu}) \quad \Longleftrightarrow \quad \boldsymbol{s} = \boldsymbol{A}_N \, \boldsymbol{u} + \boldsymbol{\mu}$$

$$\text{after coding:} \quad \boldsymbol{s'} = \boldsymbol{A}_N \, \boldsymbol{u'} + \boldsymbol{\mu} \quad \Longleftrightarrow \quad \boldsymbol{u'} = \boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{s'} - \boldsymbol{\mu})$$

➡ Effect on reconstruction error $\boldsymbol{s} - \boldsymbol{s'}$

$$\boldsymbol{u} - \boldsymbol{u'} = \boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{s} - \boldsymbol{s'}) \quad \Longleftrightarrow \quad \boldsymbol{s} - \boldsymbol{s'} = \boldsymbol{A}_N \, (\boldsymbol{u} - \boldsymbol{u'})$$

➡ Mean squared error distortion for N samples

$$\begin{aligned} d_N(\boldsymbol{s}, \boldsymbol{s'}) &= \frac{1}{N} \sum_{k=0}^{N-1} (s_k - s'_k)^2 = \frac{1}{N} (\boldsymbol{s} - \boldsymbol{s'})^{\mathrm{T}} (\boldsymbol{s} - \boldsymbol{s'}) \\ &= \frac{1}{N} (\boldsymbol{u} - \boldsymbol{u'})^{\mathrm{T}} \boldsymbol{A}_N^{\mathrm{T}} \boldsymbol{A}_N \, (\boldsymbol{u} - \boldsymbol{u'}) = \frac{1}{N} (\boldsymbol{u} - \boldsymbol{u'})^{\mathrm{T}} (\boldsymbol{u} - \boldsymbol{u'}) \\ &= \frac{1}{N} \sum_{k=0}^{N-1} (u_k - u'_k)^2 = d_N(\boldsymbol{u}; \boldsymbol{u'}) \end{aligned}$$

➡ Same distortion: **Can consider coding in transform domain**

N-th Order Pdf of Transform Coefficients

Linear transform of Gaussian random vectors $\boldsymbol{S}$ (with N samples)

$$\boldsymbol{U} = \boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{S} - \boldsymbol{\mu})$$

- Property: Any linear combination of Gaussian random variables yields another Gaussian random variable

➡ Vectors of transform coefficients $\boldsymbol{U}$ have multivariate Gaussian distribution

➡ Mean of transform coefficient vectors

$$\boldsymbol{\mu}_U = \mathrm{E}\{\,\boldsymbol{U}\,\} = \mathrm{E}\Big\{\,\boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{S} - \boldsymbol{\mu})\,\Big\} = \boldsymbol{A}_N^{\mathrm{T}} \Big(\,\mathrm{E}\{\,\boldsymbol{S}\,\} - \boldsymbol{\mu}\Big) = \boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{\mu} - \boldsymbol{\mu}) = 0$$

➡ N-th order covariance matrix of transform coefficient vectors $\boldsymbol{U}$

$$\begin{aligned}
\boldsymbol{C}_N^{UU} &= \mathrm{E}\{\,(\boldsymbol{U} - \boldsymbol{\mu}_u)(\boldsymbol{U} - \boldsymbol{\mu}_u)^{\mathrm{T}}\,\} = \mathrm{E}\Big\{\,\boldsymbol{A}_N^{\mathrm{T}} (\boldsymbol{S} - \boldsymbol{\mu})(\boldsymbol{S} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{A}_N\,\Big\} \\
&= \boldsymbol{A}_N^{\mathrm{T}}\, \mathrm{E}\{\,(\boldsymbol{S} - \boldsymbol{\mu})(\boldsymbol{S} - \boldsymbol{\mu})^{\mathrm{T}}\,\}\, \boldsymbol{A}_N = \boldsymbol{A}_N^{\mathrm{T}}\, \boldsymbol{C}_N\, \boldsymbol{A}_N \\
&= \boldsymbol{A}_N^{\mathrm{T}} \left(\boldsymbol{A}_N \cdot \boldsymbol{\Lambda}_N \cdot \boldsymbol{A}_N^{\mathrm{T}}\right) \boldsymbol{A}_N = \boldsymbol{\Lambda}_N
\end{aligned}$$

N-th Order Pdf of Transform Coefficients

N-th order pdf of transform coefficient vectors $\boldsymbol{u}$

- Multivariate Gaussian with $\mu = 0$ and $\boldsymbol{C}_N = \boldsymbol{\Lambda}_N$

$$f(\boldsymbol{u}) = \frac{1}{\sqrt{(2\pi)^N \, |\boldsymbol{\Lambda}_N|}} \, e^{-\frac{1}{2}\, \boldsymbol{u}^{\mathrm{T}} \, \boldsymbol{\Lambda}_N^{-1} \, \boldsymbol{u}}$$

- $\boldsymbol{\Lambda}_N$ = diagonal matrix with eigenvalues $\{\xi_k\}$ of $\boldsymbol{C}_N$ on its main diagonal

$$\boldsymbol{\Lambda}_N^{-1} = \begin{bmatrix} 1/\xi_0 & 0 & 0 & \dots & 0 \\ 0 & 1/\xi_1 & 0 & \dots & 0 \\ 0 & 0 & 1/\xi_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1/\xi_{N-1} \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \boldsymbol{u}^{\mathrm{T}} \, \boldsymbol{\Lambda}_N^{-1} \boldsymbol{u} &= \sum_{k=0}^{N-1} \frac{u_k^2}{\xi_k} \\ \left|\boldsymbol{\Lambda}_N\right| &= \prod_{k=0}^{N-1} \xi_k \end{aligned}$$

➡ N-th order pdf of transform coefficient vectors $\boldsymbol{u}$

$$f(\boldsymbol{u}) = \frac{1}{\sqrt{(2\pi)^N \prod_k \xi_k}} \, e^{-\frac{1}{2} \sum_k \frac{u_k^2}{\xi_k}} = \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi\, \xi_k}} \, e^{-\frac{u_k^2}{2\,\xi_k}} = \prod_{k=0}^{N-1} f_k(u_k)$$

N-th Order Rate-Distortion Function

- N-th order pdf of transform coefficient vectors $\boldsymbol{u}$

$$f(\boldsymbol{u}) = \prod_{k=0}^{N-1} f_k(u_k) \qquad \text{with} \qquad f_k(u_k) = \frac{1}{\sqrt{2\pi\xi_k}} e^{-\frac{u_k^2}{2\xi_k}}$$

➡ Random variables U_k in vector $\boldsymbol{U}$ are mutually independent

➡ Each random variable U_k is Gaussian with $\mu_k = 0$ and $\sigma_k^2 = \xi_k$

- Recall: First-order rate-distortion function for Gaussian components U_k

$$D_k(R_k) = \sigma_k^2\, 2^{-2R_k} = \xi_k\, 2^{-2R_k} \qquad (\xi_k\,:\; k\text{-th eigenvalue of } \boldsymbol{C}_N)$$

➡ N-th order rate-distortion function

$$D_N(R) = \frac{1}{N} \sum_{k=0}^{N-1} \xi_k\, 2^{-2R_k} \qquad \text{with} \qquad R = \frac{1}{N} \sum_{k=0}^{N-1} R_k$$

➡ Next problem: **How to distribute rate among components?**

Optimal Bit Allocation among Mut. Independent Gaussian

- Optimization problem

$$\min_{R_0, R_1, \ldots, R_{N-1}} D_N(R) = \frac{1}{N} \sum_{k=0}^{N-1} \xi_k \, 2^{-2R_k} \qquad \text{such that} \qquad R = \frac{1}{N} \sum_{k=0}^{N-1} R_k$$

- Using inequality for arithmetic and geometric means

$$D_N(R) = \frac{1}{N} \sum_{k=0}^{N-1} \xi_k \, 2^{-2R_k} \geq \left(\prod_{k=0}^{N-1} \xi_k \, 2^{-2R_k} \right)^{\frac{1}{N}} = \left(\prod_{k=0}^{N-1} \xi_k \right)^{\frac{1}{N}} \left(\prod_{k=0}^{N-1} 2^{-2R_k} \right)^{\frac{1}{N}}$$

$$= \left(\prod_{k=0}^{N-1} \xi_k \right)^{\frac{1}{N}} 2^{-2\left(\frac{1}{N} \sum_{k=0}^{N-1} R_k \right)} = \tilde{\xi} \cdot 2^{-2R}$$

- Equality if and only if

$$\forall k: \quad \xi_k \cdot 2^{-2R_k} = \tilde{\xi} \cdot 2^{-2R}$$

Optimal Bit Allocation among Mut. Independent Gaussians

- We have shown

$$D_N(R) \geq \tilde{\xi} \cdot 2^{-2R} \qquad \text{with} \qquad \tilde{\xi} = \left(\prod_{k=0}^{N-1} \xi_k \right)^{\frac{1}{N}} = \sqrt[N]{|\boldsymbol{C}_N|}$$

- Optimal bit allocation (equality in above inequality) if and only if

$$\xi_k \cdot 2^{-2R_k} \;=\; \tilde{\xi} \cdot 2^{-2R}$$

$$D_k(R_k) = \xi_k \cdot 2^{-2R_k} \;=\; D$$

➡ All component distortions D_k are equal to overall distortion D, yielding

$$R_k = \frac{1}{2} \log_2 \frac{\xi_k}{D}$$

- But: Component rates R_k cannot become negative

➡ We require

$$R_k \geq 0 \quad \Longleftrightarrow \quad \xi_k \geq D$$

Optimal Bit Allocation among Mut. Independent Gaussians

- Optimal bit allocation
  - Do not assign any bits for components with $\xi_k \leq D$
  - Distribute rate R among remaining components

➡ Can be elegantly specified using parameter $\theta > 0$

$$D_k = \min\left(\theta,\, \xi_k\right)$$

➡ Associated rate allocation (using $D_k = \xi_k\, 2^{-2R_k}$)

$$D_k = \xi_k\, 2^{-2R_k} \;=\; \min\left(\theta,\, \xi_k\right)$$

$$R_k \;=\; -\frac{1}{2}\log_2 \frac{\min\left(\theta,\, \xi_k\right)}{\xi_k} = \max\left(0,\, \frac{1}{2}\log_2 \frac{\xi_k}{\theta}\right)$$

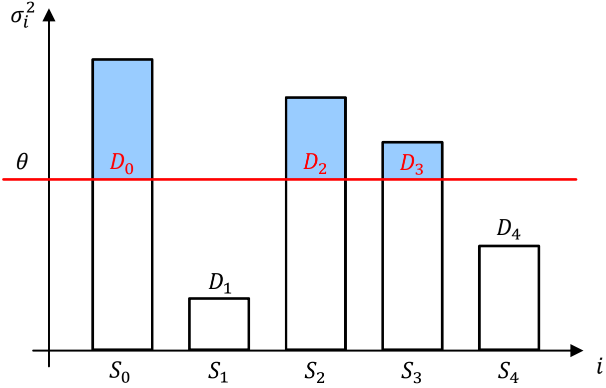
➡ Referred to as reverse water filling

Optimal Bit Allocation: Reverse Water Filling

Optimal bit allocation: Independent Gaussian RV and MSE distortion

- Do not assign any rate to random variables S_i with $\sigma_i^2 \leq \theta$
- Code other S_i so that component distortions $D_i = \theta$ are obtained

$$D_i = \min\left(\sigma_i^2, \theta\right)$$



N-th Order Rate-Distortion Function

- Parametric formulation for component distortions and rates

$$D_k(\theta) = \min\left(\theta,\, \xi_k\right) \qquad \text{and} \qquad R_k = \max\left(0,\, \frac{1}{2}\log_2 \frac{\xi_k}{\theta}\right) \tag{41}$$

- Parametric formulation of N-th order rate-distortion function

$$D_N(\theta) \;=\; \frac{1}{N}\sum_{k=0}^{N-1} D_k = \frac{1}{N}\sum_{k=0}^{N-1} \min\left(\xi_k, \theta\right) \tag{42}$$

$$R_N(\theta) \;=\; \frac{1}{N}\sum_{k=0}^{N-1} R_k = \frac{1}{N}\sum_{k=0}^{N-1} \max\left(0, \frac{1}{2}\log_2 \frac{\xi_k}{\theta}\right) \tag{43}$$

- Rate-distortion function: Limit for $N \to \infty$

Rate-Distortion Function for Stationary Gaussian and MSE

- Rate-distortion function: Limit for $N \to \infty$

$$D(\theta) = \lim_{N\to\infty} D_N(\theta) \;=\; \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \min\left(\xi_k^{(N)}, \theta\right) \tag{44}$$

$$D(\theta) = \lim_{N\to\infty} R_N(\theta) \;=\; \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \max\left(0, \frac{1}{2}\log_2 \frac{\xi_k^{(N)}}{\theta}\right) \tag{45}$$

with $\{\xi_k^{(N)}\}$ denoting the eigenvalues of the N-th order covariance matrix $\boldsymbol{C}_N$

- Recall: Theorem of Grenander and Szegö for series of Toeplitz matrices

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} G\left(\xi_k^{(N)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G\left(\,\Phi(\omega)\,\right) d\omega \tag{46}$$

with $\Phi(\omega)$ being the Fourier series of the autocovariance function ϕ_k (values on k-th diagonal of the series of Toeplitz matrices)

Rate-Distortion Function for Stationary Gaussian and MSE

Rate-Distortion Function for Stationary Gaussian Sources

- Parametric formulation with $\theta > 0$

$$D(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min\big(\Phi_{SS}(\omega), \theta\big)\, \mathrm{d}\omega \tag{47}$$

$$R(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \max\left(0, \frac{1}{2} \log_2 \frac{\Phi_{SS}(\omega)}{\theta}\right) \mathrm{d}\omega \tag{48}$$

with $\Phi_{SS}(\omega)$ being the Fourier series of the autocovariance function ϕ_k

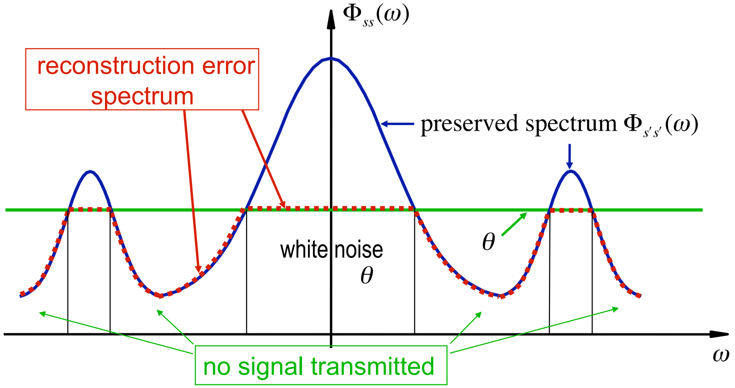
$$\Phi_{SS}(\omega) = \sum_{k=-\infty}^{\infty} \phi_k \cdot e^{-\mathrm{i}\,\omega\,k} \quad \text{with} \quad \phi_k = \mathrm{E}\{\,(S_n - \mu)(S_{n+k} - \mu)\,\} \tag{49}$$

For MSE distortion and given autocovariance function ϕ_k

- $R(D)$ is maximized for Gaussian sources (without proof)

➡ Gaussian sources are the hardest to code

Illustration of Parametric Rate-Distortion Function



Rate-Distortion Function for Stationary Gauss-Markov

- Autocovariance function and its Fourier series

$$\phi_k = \sigma^2 \, \rho^{|k|} \quad \Longleftrightarrow \quad \Phi_{SS}(\omega) = \frac{\sigma^2 \, (1-\varrho^2)}{1 - 2\varrho \, \cos\omega + \varrho^2} \tag{50}$$

- For $\varrho \geq 0$: All frequency components are coded if we choose

$$\theta \leq \min_{\forall\omega} \Phi_{SS}(\omega) = \Phi_{SS}(\pi) = \sigma^2 \, \frac{1-\varrho^2}{1+2\varrho+\varrho^2} = \sigma^2 \, \frac{1-\varrho}{1+\varrho} \tag{51}$$

- For this range, rate-distortion function becomes

$$D(\theta) \;=\; \frac{1}{2\pi} \int_{-\pi}^{\pi} \min\Big(\Phi_{SS}(\omega), \theta\Big) \, \mathrm{d}\omega = \frac{\theta}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}\omega = \theta$$

$$R(\theta) \;=\; \frac{1}{2\pi} \int_{-\pi}^{\pi} \max\left(0, \frac{1}{2} \log_2 \frac{\Phi_{SS}(\omega)}{\theta}\right) \, \mathrm{d}\omega = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2 \frac{\Phi_{SS}(\omega)}{D} \, \mathrm{d}\omega$$

➡ **Same as Shannon lower bound in this range!**

Rate-Distortion Function for Stationary Gauss-Markov

Stationary Gauss-Markov with $\varrho \geq 0$

- If we choose

$$\theta \leq \sigma^2 \, \frac{1-\varrho}{1+\varrho}$$

we get $D = \theta$ and

$$R(D) = R_L(D) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2 \frac{\Phi_{SS}(\omega)}{D} \, \mathrm{d}\omega = \frac{1}{2} \log_2 \frac{\sigma^2 \, (1-\varrho^2)}{D}$$

➡ Rate-distortion / distortion-rate function for this range

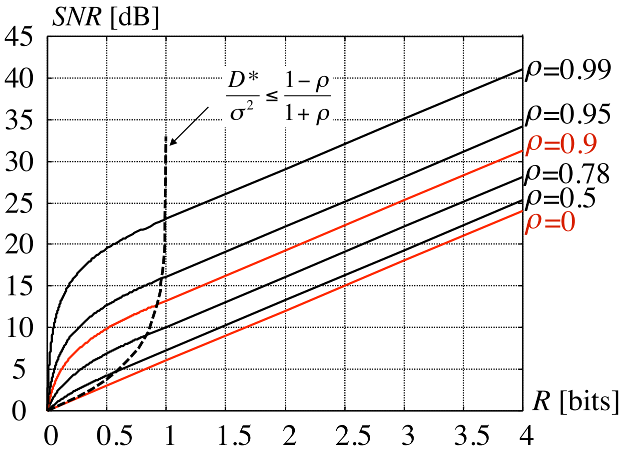
$$R(D) \;=\; \frac{1}{2} \log_2 \frac{\sigma^2 \, (1-\varrho^2)}{D} \qquad \text{for } D \leq \sigma^2 \, \frac{1-\varrho}{1+\varrho}$$

$$D(R) \;=\; (1-\varrho^2) \, \sigma^2 \, 2^{-2R} \qquad \text{for } R \geq \log_2(1+\varrho)$$

➡ Guaranteed to be equal to Shannon lower bound for $R \geq 1$ bit/sample

Rate-Distortion Function for Stationary Gauss-Markov

- Distortion-rate function $D(R)$ as signal-to-noise ratio (SNR)



Part Summary

Mutual Information

- Information that a random variables carries about another random variable
- Discrete random variables: $I(A;B) = H(A) - H(A\,|\,B)$
- Continuous random variables: $I(A;B) = h(A) - h(A\,|\,B)$

Differential Entropy

- Definition for continuous random variables
- N-th order differential entropy and differential entropy rate
- Gaussian sources maximize differential entropy rate for given autocovariance

Rate-Distortion Function

- Operational and information rate-distortion function $R(D)$
- Fundamental bound for source coding (special case: lossless coding)
- Discrete sources: Convex function with $R(0) = \bar{H}(\boldsymbol{S})$
- Continuous sources: Convex function with $R(0) \to \infty$
- MSE distortion: $D(0) = \sigma^2$

Part Summary

Shannon Lower Bound

- Lower bound for rate-distortion function
- Asymptotically tight for high rates / small distortions
- Suitable reference for performance evaluation at high rates
- Shannon lower bound for IID sources and MSE distortion
- Shannon lower bound for Gaussian sources and MSE distortion

Rate-Distortion Function for Gaussian Sources and MSE distortion

Gaussian IID sources:	Coincides with Shannon lower bound
Stationary Gaussian:	Parametric formulation
Stationary Gauss-Markov:	Coincides with SLB for $R \geq \log_2(1 + \varrho)$
$R(D)$ for stat. Gaussian:	Upper bound of $R(D)$ for all other sources with same autocovariance function

➡ Gaussian sources are the most difficult to code (for MSE distortion)