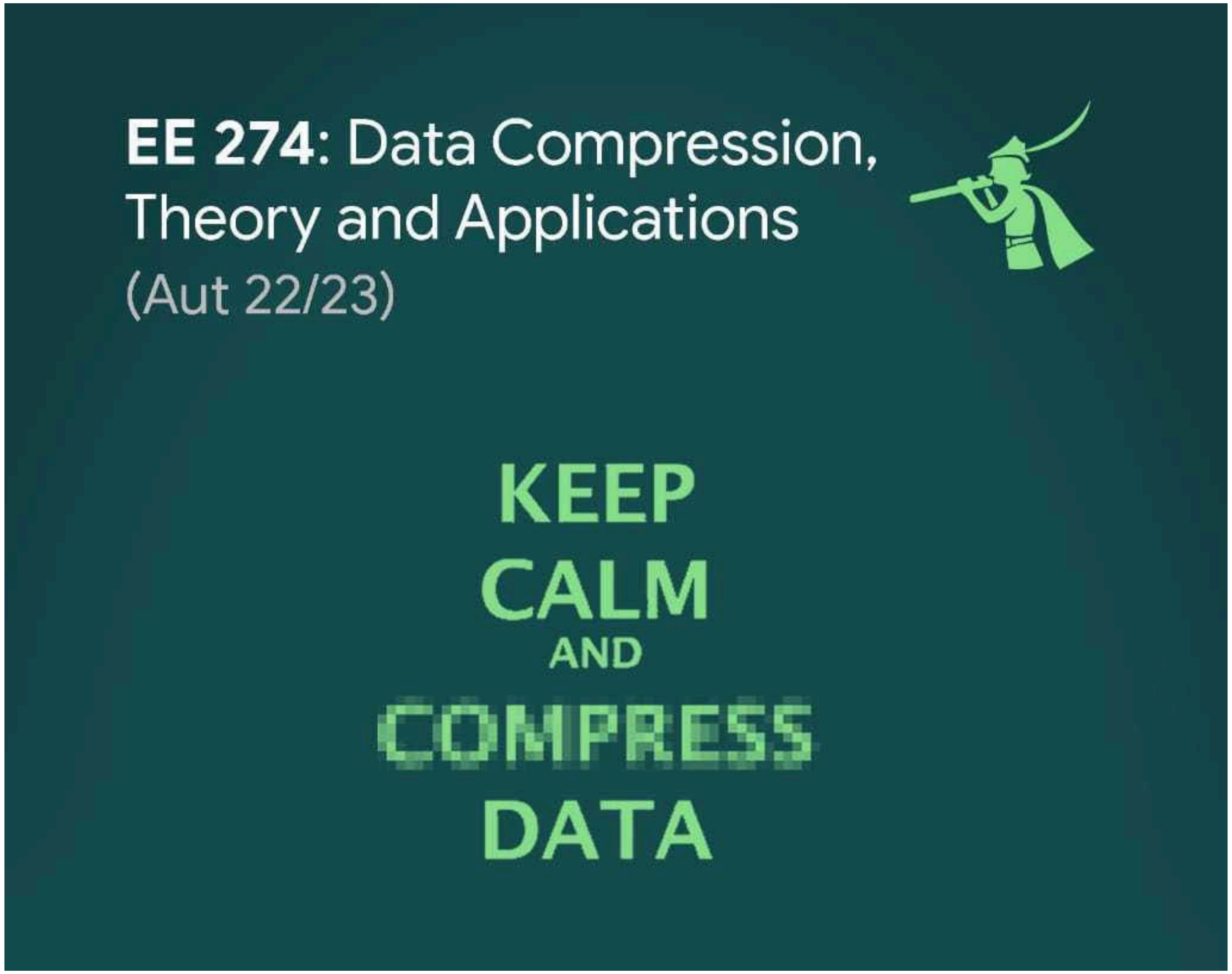


water pouring characterization of R(D) for the Gaussian source

Lossy Compression

info source: $\underbrace{X_1, \dots, X_k}_{X^k} \longrightarrow \boxed{\text{encoder}} \xrightarrow{\text{index } J \in \{1,2,\dots N\}} \boxed{\text{decoder}} \xrightarrow{Y_1, \dots, Y_k}$

$$\text{rate} = \frac{\log N}{k} \quad \frac{\text{bits}}{\text{source component}}$$

$$\text{distortion: } d(X^k, Y^k) = \frac{1}{k} \sum_{i=1}^{k} d(X_i, Y_i)$$

$R(D) =$ min rate needed to achieve (mean) distortion no more than D (optimizing across k and encoders + decoders)

Suppose $\mathbf{X} = (X_1, X_2, \ldots\ldots)$ is a stationary source.

Shannon's theorem for lossy compression:

$$R(D) = \lim_{k \to \infty} \min_{Ed(X^k, Y^k)} \frac{1}{k} I(X^k; Y^k)$$

R(D) for a Gaussian Source: I

Denote $R_G(\sigma^2, D) \triangleq \min\limits_{E[(X-Y)^2] \le D} I(X;Y)$, when $X \sim N(0, \sigma^2)$.

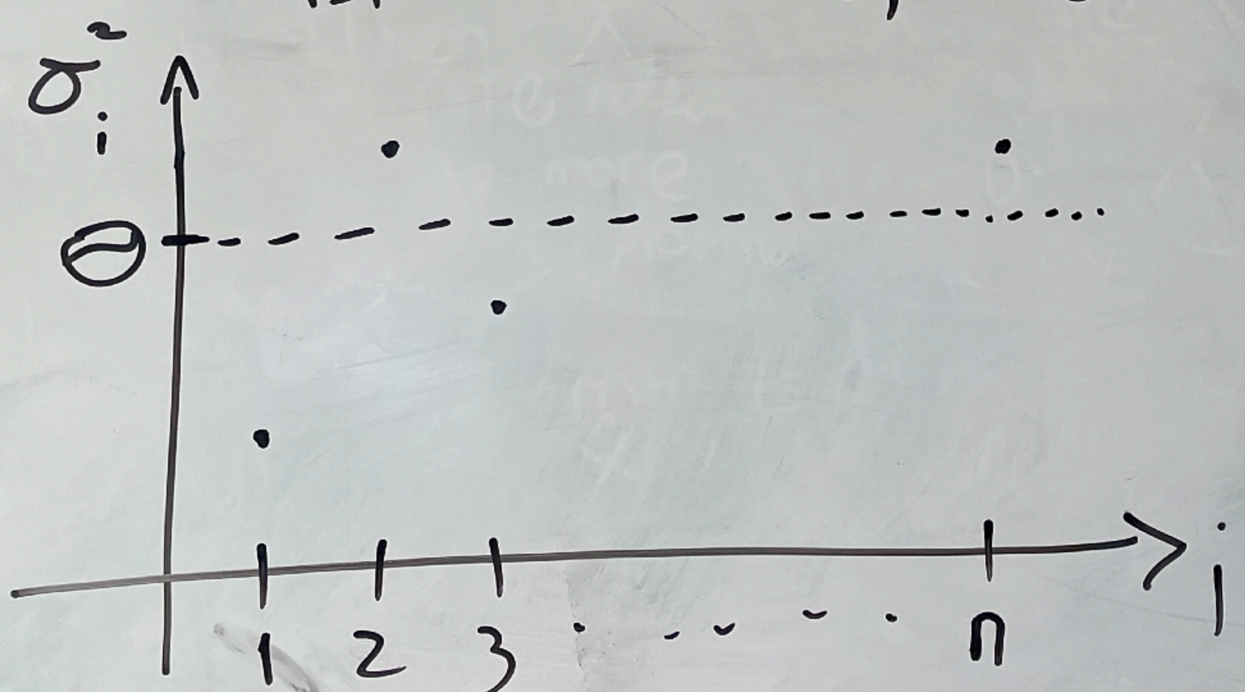
For $\vec{\sigma}^2 = \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_n^2 \end{pmatrix}$ denote $R_G(\vec{\sigma}^2, D) \triangleq \min\limits_{E[\|X^n - Y^n\|^2] \le D} \frac{1}{n} I(X^n; Y^n)$,

when $X_1, \ldots, X_n$ are independent with $X_i \sim N(0, \sigma_i^2)$.

$R(D)$ for a Gaussian Source: II

Recall: $$R_G(\vec{\sigma}^2, D) = \min \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2/D_i\right]_+$$
$$\text{s.t. } \frac{1}{n}\sum_{i=1}^{n} D_i \leq D$$

and is given parametrically by the curve:
$$D_\theta = \frac{1}{n}\sum_{i=1}^{n}\min\{\theta, \sigma_i^2\}, \quad R_\theta = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2/\theta\right]_+ \qquad (\theta > 0)$$



$R(D)$ for a Gaussian Source: III

X^n Gaussian with covariance Φ_{X^n}

Then $R(X^n, D)$ (under squared error)

is given by $R_G(\vec{\lambda}, D)$ where

$$\vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$ are the eigenvalues of Φ_{X^n}

$R(D)$ for a Gaussian Source: IV

When X^n are the first n components of a stationary Gaussian process $\mathbf{X}$ with covariance matrix $\Phi_n = \{\phi_{|i-j|}\}_{\substack{1\le i\le n\\ 1\le j\le n}}$ where $\phi_k = \mathrm{Cov}(X_i, X_{i-k})$

$$R(X^n, D) = R_G(\vec{\lambda}^{(n)}, D)$$

where $\vec{\lambda}^{(n)}$ is the vector of eigenvalues of Φ_n

R(D) for a Gaussian Source: V

Theorem (Toeplitz distribution):

Let $S(\omega) \triangleq \sum_{k=-\infty}^{\infty} \phi_k e^{-j\omega k}$ be the spectral density of $\mathbf{X}$ and $G(\cdot)$ a continuous function.

Then

$$\frac{1}{n}\sum_{i=1}^{n} G\left(\lambda_i^{(n)}\right) \xrightarrow{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} G\big(S(\omega)\big)\,d\omega$$

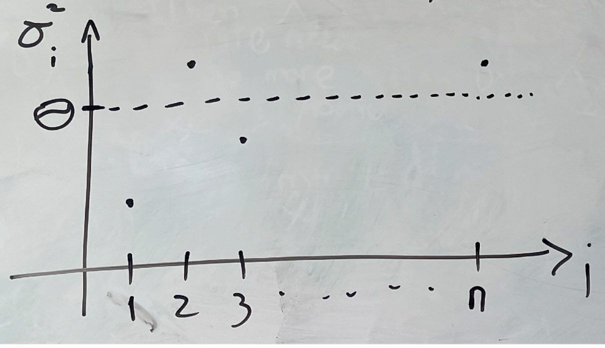
recap

R(D) for a Gaussian Source: II

Recall: $R_G(\vec{\sigma}^2, D) = \min_{\text{s.t. } \frac{1}{n}\sum_{i=1}^{n} D_i \le D} \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2 / D_i\right]_+$

and is given parametrically by the curve:

$$D_\theta = \frac{1}{n}\sum_{i=1}^{n}\min\{\theta, \sigma_i^2\}, \quad R_\theta = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2/\theta\right]_+ \qquad (\theta > 0)$$



R(D) for a Gaussian Source: III

X^n Gaussian with covariance Φ_{X^n}

Then $R(X^n, D)$ (under squared error) is given by $R_G(\vec{\lambda}, D)$ where

$$\vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$ are the eigenvalues of Φ_{X^n}

R(D) for a Gaussian Source: IV

When X^n are the first n components of a stationary Gaussian process $\mathbf{X}$ with covariance matrix $\Phi_n = \{\phi_{|i-j|}\}_{\substack{1\le i\le n \\ 1\le j\le n}}$ where $\phi_k = \mathrm{Cov}(X_i, X_{i-k})$

$$R(X^n, D) = R_G(\vec{\lambda}^{(n)}, D)$$

where $\vec{\lambda}^{(n)}$ is the vector of eigenvalues of Φ_n

R(D) for a Gaussian Source: V

Theorem (Toeplitz distribution):

Let $S(\omega) \triangleq \sum_{k=-\infty}^{\infty} \phi_k e^{-j\omega k}$ be the spectral density of $\mathbf{X}$ and $G(\cdot)$ a continuous function.

Then $$\frac{1}{n}\sum_{i=1}^{n} G(\lambda_i^{(n)}) \xrightarrow{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} G(S(\omega))\,d\omega$$

R(D) for a Gaussian Source: VI

Specializing to $G(\lambda) = \min\{\Theta, \lambda\}$ and to $G(\lambda) = \left[\frac{1}{2}\log \lambda/\Theta\right]_+$ we get:

The rate distortion function of a stationary Gaussian process with spectral density $S(\omega)$ is given parametrically by:

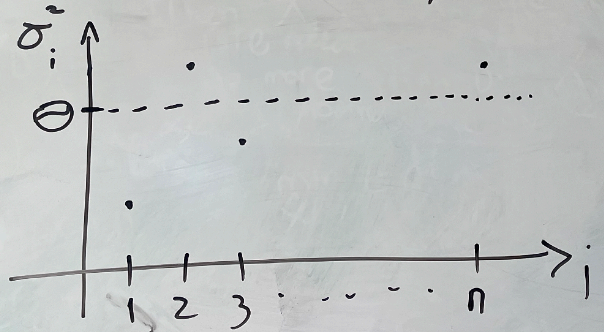
$$D_\Theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min\{\Theta, S(\omega)\}\,d\omega, \quad R_\Theta = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\log\frac{S(\omega)}{\Theta}\right]_+ d\omega$$

R(D) for a Gaussian Source: II

Recall: $R_G(\vec{\sigma}^2, D) = \min_{\text{s.t. } \frac{1}{n}\sum_{i=1}^{n} D_i \le D} \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2 / D_i\right]_+$

and is given parametrically by the curve:

$$D_\theta = \frac{1}{n}\sum_{i=1}^{n}\min\{\theta, \sigma_i^2\}, \quad R_\theta = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\log \sigma_i^2/\theta\right]_+ \qquad (\theta > 0)$$



R(D) for a Gaussian Source: III

X^n Gaussian with covariance Φ_{X^n}

Then $R(X^n, D)$ (under squared error) is given by $R_G(\vec{\lambda}, D)$ where

$\vec{\lambda} = \begin{pmatrix}\lambda_1 \\ \vdots \\ \lambda_n\end{pmatrix}$ are the eigenvalues of Φ_{X^n}

R(D) for a Gaussian Source: IV

When X^n are the first n components of a stationary Gaussian process X with covariance matrix $\Phi_n = \{\phi_{|i-j|}\}_{\substack{1\le i\le n \\ 1\le j\le n}}$ where $\phi_k = Cov(X_i, X_{i-k})$

$$R(X^n, D) = R_G(\vec{\lambda}^{(n)}, D)$$

where $\vec{\lambda}^{(n)}$ is the vector of eigenvalues of Φ_n

R(D) for a Gaussian Source: V

Theorem (Toeplitz distribution):

Let $S(\omega) \triangleq \sum_{k=-\infty}^{\infty} \phi_k e^{-j\omega k}$ be the spectral density of X and $G(\cdot)$ a continuous function.

Then $\frac{1}{n}\sum_{i=1}^{n} G(\lambda_i^{(n)}) \xrightarrow{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} G(S(\omega))\,d\omega$

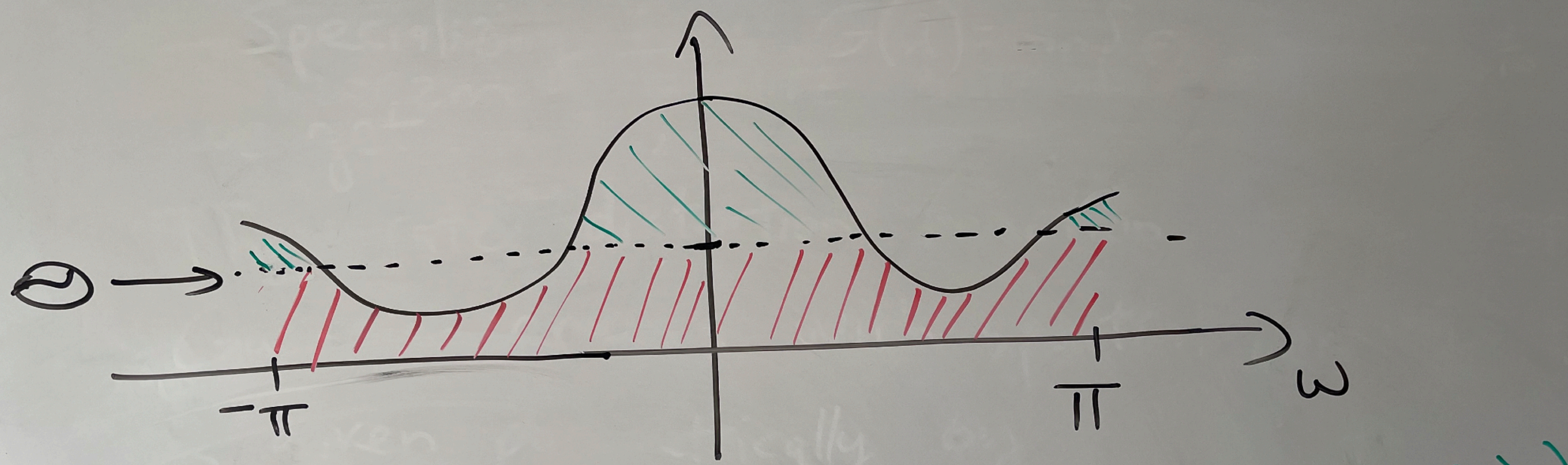
R(D) for a Gaussian Source: VI

Specializing to $G(\lambda) = \min\{\theta, \lambda\}$ and to $G(\lambda) = \left[\frac{1}{2}\log \lambda/\theta\right]_+$ we get:

The rate distortion function of a stationary Gaussian process with spectral density $S(\omega)$ is given parametrically by:

$$D_\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}\min\{\theta, S(\omega)\}\,d\omega, \quad R_\theta = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\log \frac{S(\omega)}{\theta}\right]_+ d\omega$$

R(D) for a Gaussian Source: VII



D_Θ = Area $\left(////\right)$ R_Θ = Area in log scale $\left(\backslash\backslash\backslash\right)$

R(D) for a Gaussian Source: VI

Specializing to $G(\lambda) = \min\{\Theta, \lambda\}$ and to $G(\lambda) = \left[\frac{1}{2}\log \lambda/\Theta\right]_+$ we get:

The rate distortion function of a stationary Gaussian process with spectral density $S(\omega)$ is given parametrically by:

$$D_\Theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min\{\Theta, S(\omega)\}\,d\omega, \quad R_\Theta = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\log \frac{S(\omega)}{\Theta}\right]_+ d\omega$$

R(D) for a Gaussian Source: VIII

For $D \le \min_{\omega} S(\omega)$ can verify that

$$R(D) = \left[\frac{1}{2}\log \sigma^2/D\right]_+$$

where σ^2 is the variance of the innovations of X

("kind-of" justification/inspiration for predictive coding)