Distributed Video Coding

Francesca Bassi

LSS

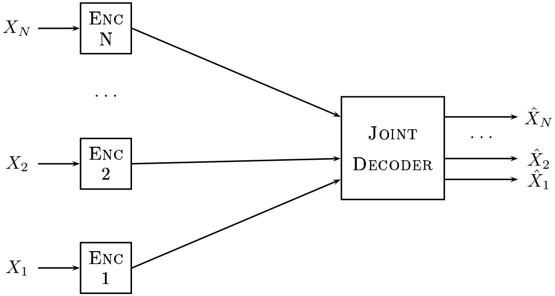
Outline

Outline

1. Introduction
   - DSC to DVC
   - Theoretical Background
2. Quadratic Gaussian Wyner-Ziv Coding
   - Limit-achieving coding
   - Finite dimensional coding
3. Non Gaussian Source-Side Information Correlation Models
   - Motivation
   - BGB, GE correlation models
   - Practical scheme GE correlation model

Outline

Distributed Source Coding

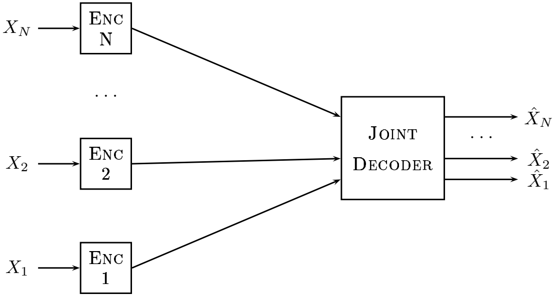


- Correlated sources.
- Independent transmission.
- Joint decoding.

Applications:

- Wireless sensor networks.
- Multiview video coding.

Distributed Source Coding

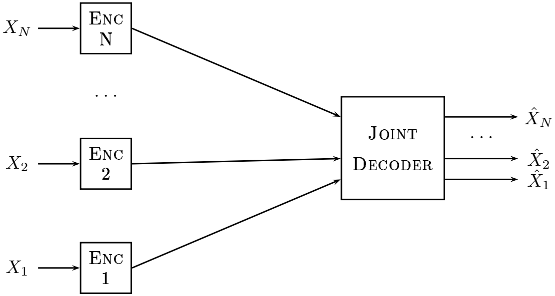


- Correlated sources.
- Independent transmission.
- Joint decoding.

Applications:

- Wireless sensor networks.
- Multiview video coding.

Distributed Source Coding



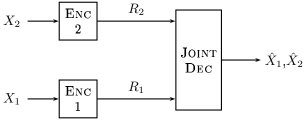
- Correlated sources.
- Independent transmission.
- Joint decoding.

Applications:

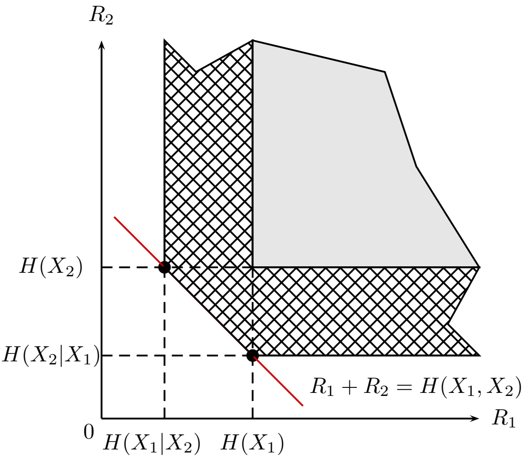
- Wireless sensor networks.
- Multiview video coding.

Slepian-Wolf Theorem

Lossless Distributed Source Coding.



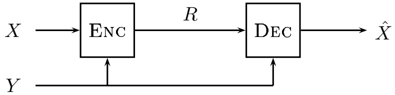
Achievable rate region for i.i.d. discrete sources X_1 and X_2. [Slepian-Wolf theorem, 1973]



Distributed Video Coding

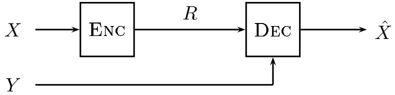
Video coding under DSC paradigm:
one source (side information Y) is available uncoded at the joint decoder.

Classical video coding:



Side information built by the encoder. Well suited for broadcasting scenarios: complex encoder.

Distributed video coding:

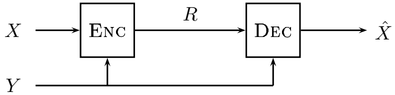


Side information built by the decoder. Well suited for mobile devices video upload: simple encoder.

Distributed Video Coding

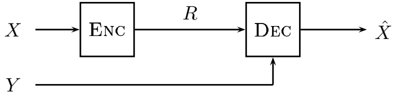
Video coding under DSC paradigm:
one source (side information Y) is available uncoded at the joint decoder.

Classical video coding:



Side information built by the encoder. Well suited for broadcasting scenarios: complex encoder.

Distributed video coding:

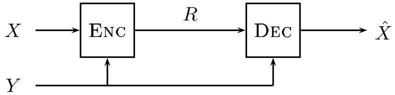


Side information built by the decoder. Well suited for mobile devices video upload: simple encoder.

Distributed Video Coding

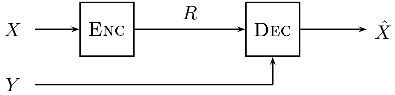
Video coding under DSC paradigm:
one source (side information Y) is available uncoded at the joint decoder.

Classical video coding:



Side information built by the encoder. Well suited for broadcasting scenarios: complex encoder.

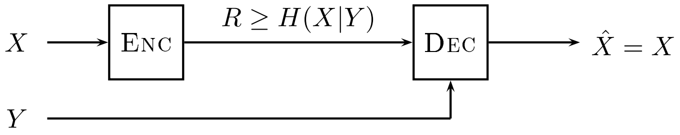
Distributed video coding:



Side information built by the decoder. Well suited for mobile devices video upload: simple encoder.

Lossless Source Coding with Side Info at the Decoder

Slepian-Wolf rate region corner points:
no rate loss with respect to two-sided side information.



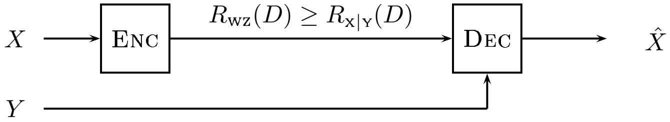
Correlation model $Y = X \oplus \Delta$

Asymmetric coding:

- Systematic code parity bits.
- Syndrome coding.

Lossy Source Coding With Side Info at the Decoder

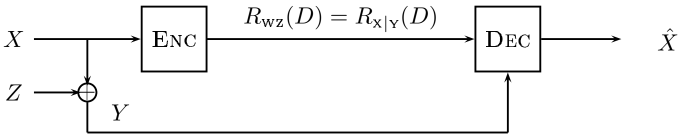
Wyner-Ziv rate distortion function.



- Correlation model $Y = X + Z$.
- Distortion $D = \frac{1}{N} E[d(X, \hat{X})]$.
- $\mathcal{U}$ auxiliary codebook, $Y \longrightarrow X \longrightarrow U$.
- $f_E : \mathcal{X} \longrightarrow \mathcal{U}$ encoding mapping, $f_D : \mathcal{U} \times \mathcal{Y} \longrightarrow \hat{\mathcal{X}}$ decoding mapping.
- $$R(D) = \min_{p(u|x)p(\hat{x}|u,y) : E[d(X,\hat{X})] \leq D} \left[I(X|U) - I(Y|U) \right].$$

Wyner-Ziv theorem: Quadratic Gaussian Case

- X, Z Gaussian and independent.
- $D = \frac{1}{N} E[\| X - \hat{X} \|^2]$.



$$R_{WZ}(D) = \begin{cases} \frac{1}{2}\log_2\left(\frac{\sigma_z^2\sigma_x^2}{(\sigma_z^2+\sigma_x^2)D}\right), & 0 < D \leq \frac{\sigma_z^2\sigma_x^2}{\sigma_z^2+\sigma_x^2} \\ 0, & D > \frac{\sigma_z^2\sigma_x^2}{\sigma_z^2+\sigma_x^2} \end{cases}$$

How Is Compression Achieved? Binning Principle

- Encoder: maps the source sequence into the auxiliary codebook word. Codebook divided into 2^{NR} equivalence classes (bins).
- Channel: index of the bin the active codeword belongs to.
- Decoder: minimum distance decoding of the side information over the active bin.

Design requirements

- Codebook: good *source code*.
- Bin: good *channel code*.

How Is Compression Achieved? Binning Principle

- Encoder: maps the source sequence into the auxiliary codebook word. Codebook divided into 2^{NR} equivalence classes (bins).
- Channel: index of the bin the active codeword belongs to.
- Decoder: minimum distance decoding of the side information over the active bin.

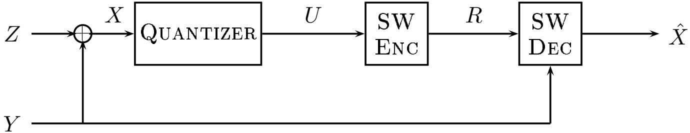
Design requirements

- Codebook: good *source code*.
- Bin: good *channel code*.

Outline

Quantization and asymmetric coding (1)

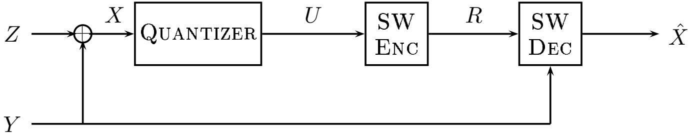
- $X = Y + Z$, with Y and Z independent and Gaussian.
- $R_{WZ}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_Z^2}{D}\right)$.



- High resolution assumption.
- Ideal asymmetric coding.
- $R = \frac{1}{N}H(U|Y)$.
- $\hat{X} = \hat{U}$.

Quantization and asymmetric coding (1)

- $X = Y + Z$, with Y and Z independent and Gaussian.
- $R_{WZ}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_Z^2}{D}\right)$.



- High resolution assumption.
- Ideal asymmetric coding.
- $R = \frac{1}{N}H(U|Y)$.
- $\hat{X} = \hat{U}$.

Quantization and asymmetric coding (2)

Asymptotical behavior ($N \longrightarrow \infty$):

- Vector quantizer: uniform.
- Entropy coding of U.

$$R = \frac{1}{N}H(U|Y) = \frac{1}{N}h(X|Y) - \frac{1}{N}\log(V(\Lambda)).$$

- Granular distortion $D_g = V(\Lambda)^{\frac{2}{N}}G(\Lambda)$.
- $G(\Lambda)$ normalized second moment of the quantizer.

$$R = \frac{1}{2}\log_2\left(\frac{\sigma_z^2}{D}\right) + \frac{1}{2}\log_2\left(2\pi e \cdot G(\Lambda)\right).$$

- $\lim_{N\longrightarrow\infty} G(\Lambda) = 1/2\pi e$.

Nested Lattice Quantization (1)

Limit-achieving pdf $p(x,y,u,\hat{x}) = p(x,y)p(u|x)p(\hat{x}|u,y)$:

$$\begin{cases} X = Y + Z \\ U = \alpha X + Q \\ \hat{X} = (1-\alpha)U + Y \end{cases}$$

$$\begin{cases} Q \sim \mathcal{N}(0,\sigma_q^2) \\ \alpha = 1 - \frac{\sigma_z^2}{D} \end{cases}$$

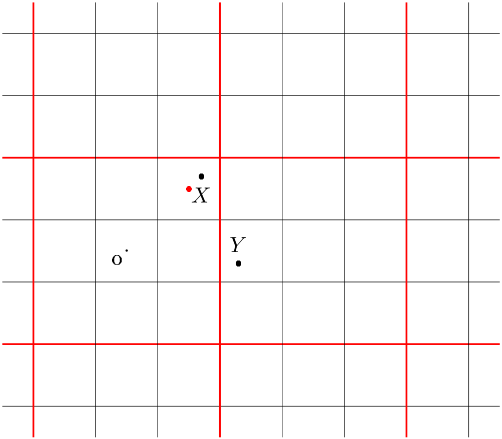
$\alpha \in [0,1)$
$\lim_{R \longrightarrow \infty} \alpha = 1$

Goal: definition of

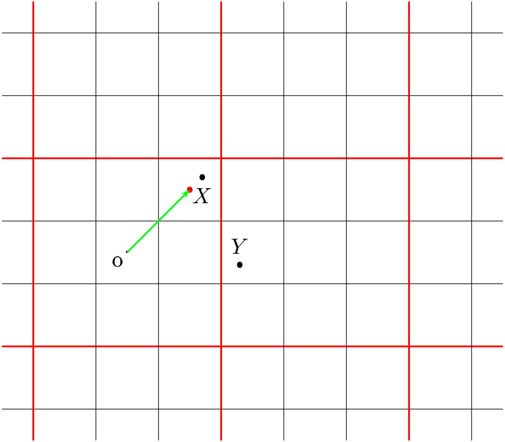
- $f_q : \mathcal{X}^N \longrightarrow \mathcal{U}^N$
- $f_e : \mathcal{Y}^N \times \mathcal{U}^N \longrightarrow \hat{\mathcal{X}}^N$

to (asymptotically) match the pdf.

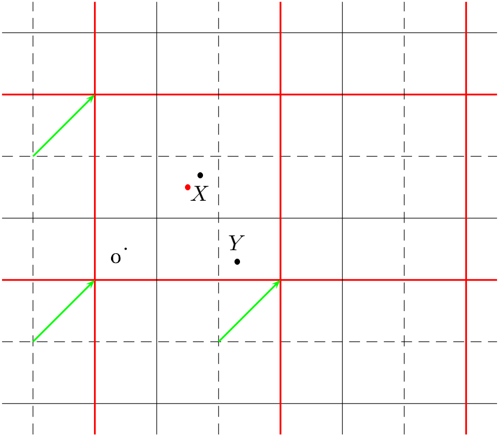
Nested Lattice Quantization (2)



Nested Lattice Quantization (2)

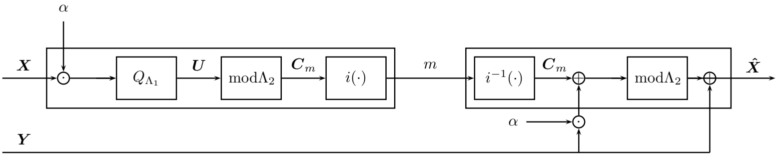


Nested Lattice Quantization (2)



Nested Lattice Quantization (3)

Asymptotical behavior ($N \longrightarrow \infty$):



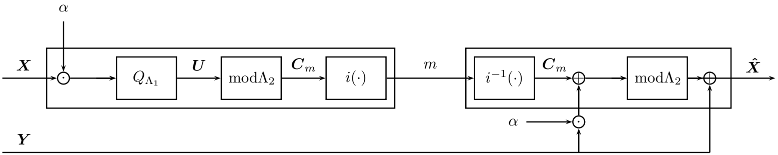
$$R = \frac{1}{N}\log_2\left(\frac{V(\Lambda_2)}{V(\Lambda_1)}\right) = \frac{1}{2}\log_2\left(\frac{\sigma^2(\Lambda_2)}{D_{g_1}}\right)$$

$$\sigma^2(\Lambda_2) = G(\Lambda_2)\cdot\sigma_z^2$$

$$\lim_{N\longrightarrow\infty} R = \frac{1}{2}\log_2\left(\frac{\sigma_z^2}{D}\right)$$

Nested Lattice Quantization (3)

Asymptotical behavior ($N \longrightarrow \infty$):



$$R = \frac{1}{N}\log_2\left(\frac{V(\Lambda_2)}{V(\Lambda_1)}\right) = \frac{1}{2}\log_2\left(\frac{\sigma^2(\Lambda_2)}{D_{g_1}}\right)$$

$$\sigma^2(\Lambda_2) = G(\Lambda_2)\cdot\sigma_z^2$$

$$\lim_{N\longrightarrow\infty} R = \frac{1}{2}\log_2\left(\frac{\sigma_z^2}{D}\right)$$

Finite Dimensional Nested Lattice Quantization

- Self noise phenomenon: $N_{eq} = (1-\alpha)Z + Q$.
- Decoding error event.

$$D = \frac{V(\Lambda_2)G(\Lambda_2)}{2^{2R}}(1-P_e) + D_e \cdot P_e$$

$$P_e = 1 - \Pr(N_{eq} \in \mathcal{V}_0(\Lambda_2))$$

- Trade off source coding and channel coding requirements.
- Inflating parameter to tune the design.

Finite Dimensional Nested Lattice Quantization

- Self noise phenomenon: $N_{eq} = (1-\alpha)Z + Q$.
- Decoding error event.

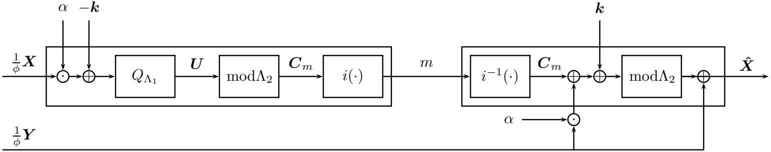
$$D = \frac{V(\Lambda_2)G(\Lambda_2)}{2^{2R}}(1-P_e) + D_e \cdot P_e$$

$$P_e = 1 - \Pr(N_{eq} \in \mathcal{V}_0(\Lambda_2))$$

- Trade off source coding and channel coding requirements.
- Inflating parameter to tune the design.

Finite Dimensional Nested Lattice Quantization

$$\phi \in (0, 1]$$

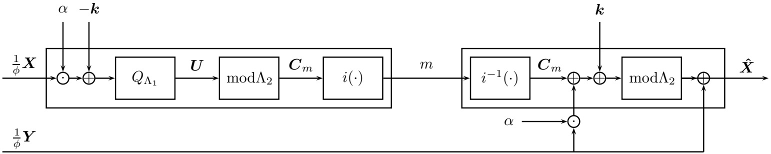


$$D = \min_{\alpha,\phi} \left((1-\alpha)^2 \sigma_z^2 + \frac{\alpha \sigma_z^2}{2^{2R}} \frac{G(\Lambda_2)}{\phi^2} + D_e \cdot P_e \right)$$

- $D_e \cdot P_e$ vanishes only if $N \longrightarrow \infty$.
- Finite dimension $\longrightarrow$ coupling of the source and channel coding problems.

Finite Dimensional Nested Lattice Quantization

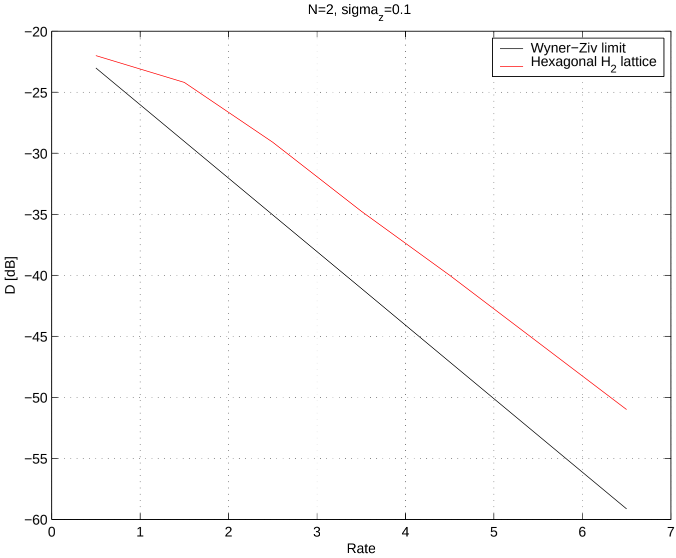
$$\phi \in (0, 1]$$



$$D = \min_{\alpha,\phi}\left((1-\alpha)^2\sigma_z^2 + \frac{\alpha\sigma_z^2}{2^{2R}}\frac{G(\Lambda_2)}{\phi^2} + D_e \cdot P_e\right)$$

- $D_e \cdot P_e$ vanishes only if $N \longrightarrow \infty$.
- Finite dimension $\longrightarrow$ coupling of the source and channel coding problems.

Finite Dimensional Nested Lattice Quantization



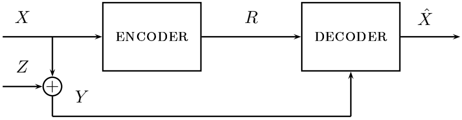
Finite Dimensional Quantization and Asymmetric Coding

$$D = D_g(1 - P_{ol}) + D_{ol} \cdot P_{ol}$$

$$P_{ol} = 1 - \Pr\left(X \in \mathcal{V}_0(\Lambda)\right)$$

Outline

Building the Side Information



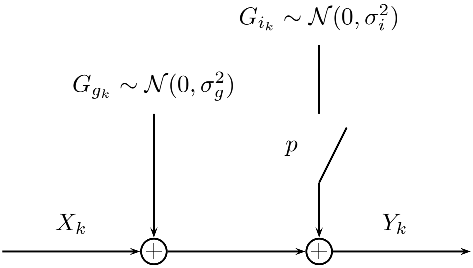
Classical prediction techniques:

- Interpolation of key frames.
- Motion compensation of previous frames.

Correlation model: non stationary!

- Gaussian Bernoulli Gaussian correlation model.
- Gaussian Erasure correlation model.

Gaussian Bernoulli Gaussian correlation model

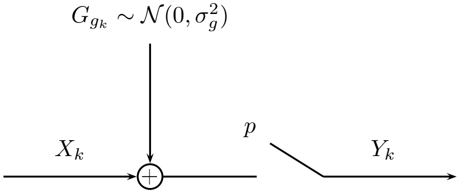


Interpolation noise.

$$\sigma_i^2 \gg \sigma_g^2$$

$$Y_k = \begin{cases} X_k + G_{g_k} & \text{prob} \quad (1-p) \\ X_k + (G_{g_k} + G_{i_k}) & \text{prob} \quad p \end{cases}$$

Gaussian Erasure correlation model



Motion estimation.

$$Y_k = \begin{cases} X_k + G_{g_k} & \text{prob} \;\; (1-p) \\ \Delta & \text{prob} \;\; p \end{cases}$$

The decoder is informed of the position of the erased symbols.

Lower Bounds for the Rate Distortion Function

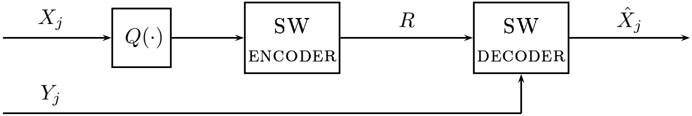
- Genie-aided: the encoder and the decoder know the positions of impulses (GBG) or erasures (GE).
- Time-division regime governed by the Bernoulli process.

$$\underline{R}_{\text{GBG}}(D) = (1-p)\, R_{X|Y,\sigma_g^2}(D) + p\, R_{X|Y,\sigma_g^2+\sigma_i^2}(D).$$

$$\underline{R}_{\text{GE}}(D) = (1-p)\, R_{X|Y,\sigma_g^2}(D) + p\, R_{X}(D).$$

$$D < \sigma_g^2$$

Upper Bounds for the Rate Distortion Function



ψ quantization step.

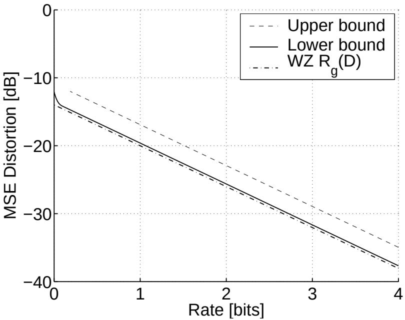
$$\begin{aligned} \overline{R}_{\text{GBG}} &= H(U|Y) = h(X|Y) - \log_2(\psi) \\ &= (1-p) \cdot h(X|Y, \sigma_g^2) + p \cdot h(X|Y, \sigma_g^2 + \sigma_i^2) - \log_2(\psi). \end{aligned}$$

$$\begin{aligned} \overline{R}_{\text{GE}} &= H(U|Y) = h(X|Y) - \log_2(\psi) \\ &= (1-p) \cdot h(X|Y \neq \Delta) + p \cdot h(X|Y = \Delta) - \log_2(\psi). \end{aligned}$$

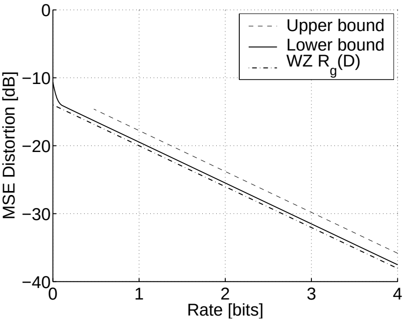
Bounds for the Rate Distortion Function

Parameters:

$$p = 0.05,\ \sigma_x^2 = 1,\ \sigma_g^2 = 0.04,\ \sigma_i^2 = 1$$



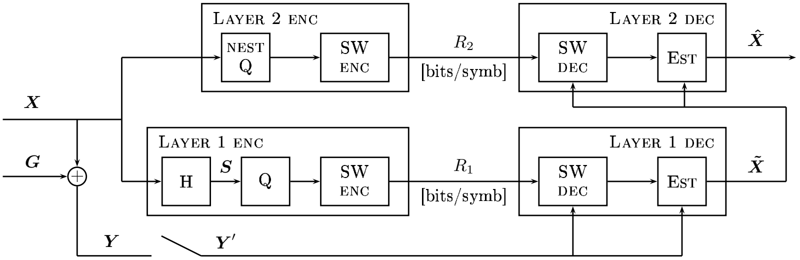
GBG correlation model



GE correlation model

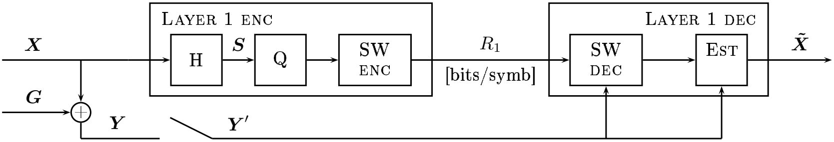
GE correlation model: two-layer scheme

- $\boldsymbol{X} \sim \mathcal{N}(0, \sigma_x^2 I_N)$
- $\boldsymbol{G} \sim \mathcal{N}(0, \sigma_g^2 I_N)$
- D distortion target
- $R = R_1 + R_2$



Layer 1

Purpose of Layer 1: fill in the blanks!



- $H_{[N-K\times N]}$ (from orthonormal basis)
- $H \cdot \boldsymbol{X} = \boldsymbol{S} \sim \mathcal{N}(0, \sigma_x^2 I_{N-K})$
- $V \cdot \boldsymbol{Y} = \boldsymbol{Y}'$

Estimation:

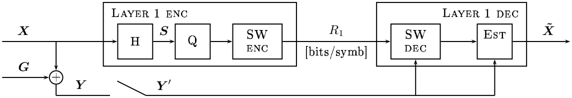
$$\begin{bmatrix} \boldsymbol{S} \\ V \cdot \boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} H \\ V \end{bmatrix} \cdot \boldsymbol{X} + \begin{bmatrix} \boldsymbol{Q} \\ V \cdot \boldsymbol{G} \end{bmatrix}$$

Layer 1: High Rate

$$R_1(D_q) = \frac{N-K}{N}\left(\frac{1}{2}\log_2\left(2\pi e \sigma_x^2\right) - \frac{1}{2}\log_2\left(12 D_q\right)\right)$$

$$D_1 = \frac{1}{N} E[\| \boldsymbol{X} - \tilde{\boldsymbol{X}} \|^2] = \frac{1}{N}\text{trace}\left(\Gamma_{\text{MAP}}\right)$$

$$\Gamma_{\text{MAP}} = \Delta^{-1} - \Delta^{-1}H^T\left(\alpha 2^{-\beta R_1} I + H\Delta^{-1}H^T\right)^{-1} H\Delta^{-1}$$

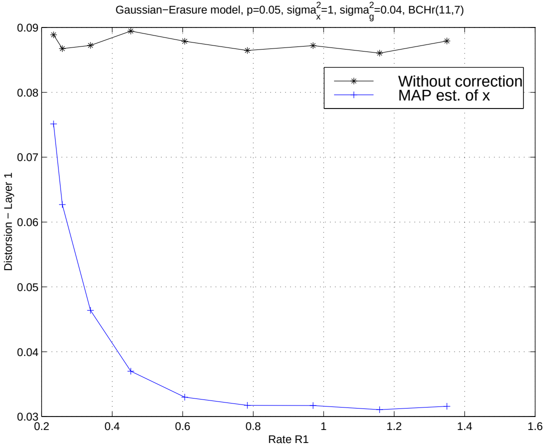


with

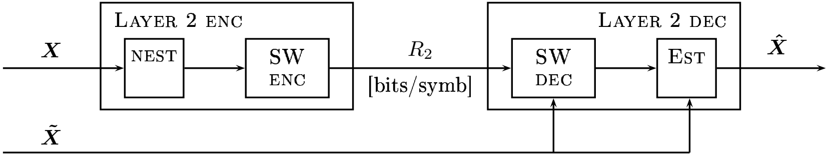
- $\Gamma_{\text{MAP}} = E[(\boldsymbol{X} - \tilde{\boldsymbol{X}})(\boldsymbol{X} - \tilde{\boldsymbol{X}})^T]$
- $\Delta = \sigma_g^{-2} V^T V + \sigma_x^{-2} I_N$
- $\alpha = \frac{2\pi e \sigma_x^2}{12}$
- $\beta = 2\frac{N}{N-K} > 0$

Layer 1: High Rate

$H_{[4\times11]}$ from DFT matrix



Layer 2



$$R_2 = \frac{1}{2}\log_2(2\pi e D_1) - \frac{1}{2}\log_2(12D)$$

$$D(R_1, R_2) = \frac{1}{N}\frac{2\pi e}{12}\,\text{trace}\left(\Gamma_{\text{MAP}}(R_1)\right) 2^{-2R_2}.$$

Simulations

$H_{[N-K \times N]}$:

- rows from DFT matrix
- rows i and $N-K+1-i$ Hermitian symmetric

Parameters:

- $N = 11$, $K = 7$
- $p = 0.05$, $\sigma_g^2 = 0.04$, $\sigma_x^2 = 1$
- Coarse step = 3

Simulation results