On the rate distortion function of Bernoulli Gaussian sequences

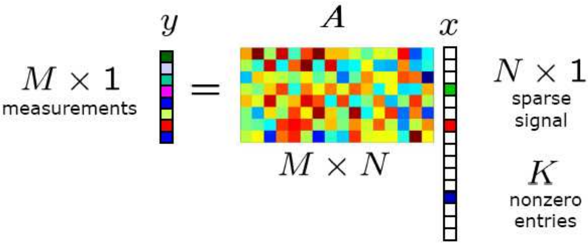
Cheng Chang

HP Labs, Palo Alto

01-30-2009

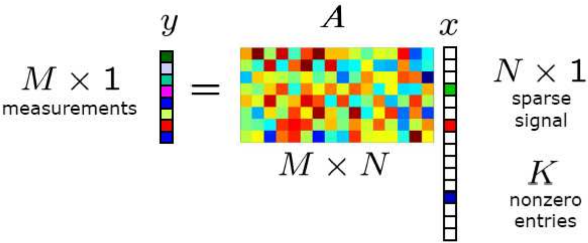
Motivation

- Measurement of sparse signals
  - Compressive sensing (Donoho 2006 , Candes and Tao 2006)
  - $K < M \ll N$, no source model



Motivation

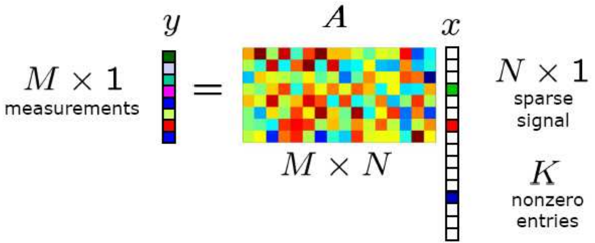
- Measurement of sparse signals
  - Compressive sensing (Donoho 2006 , Candes and Tao 2006)
  - $K < M \ll N$, no source model



- Compression of sparse signals (some source model)?

Motivation

- Measurement of sparse signals
  - Compressive sensing (Donoho 2006 , Candes and Tao 2006)
  - $K < M \ll N$, no source model



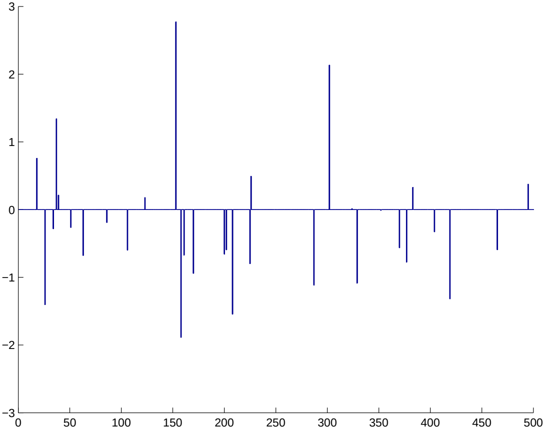
- Compression of sparse signals (some source model)?
  - Compression of x not y (Fletcher, Rangan and Goyal 2007)

Outline

1. ## Introduction
   - Bernoulli-Gaussian sequences and sparsity
   - Lossy source coding and rate distortion function $R(D)$
2. ## Main Result
   - Some obvious bounds on $R(D)$ of Bernoulli-Gaussian
   - An improved lower bound on $R(D)$
   - Proof techniques
   - Implications on coding sparse signals with high fidelity
3. ## Conclusions and Future Work

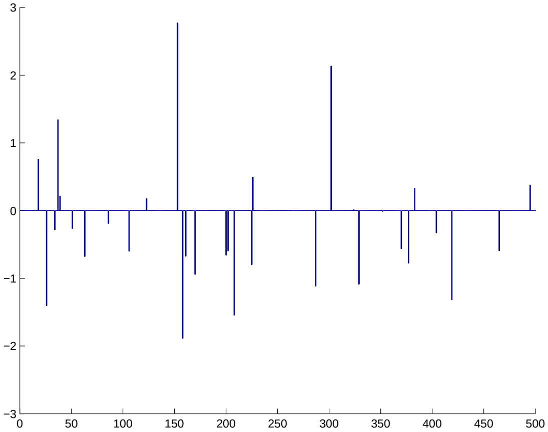
Source model for sparse signals

- $x^n \in \mathcal{R}^n$, $x_i = 0$ most of the time (6%)



Source model for sparse signals

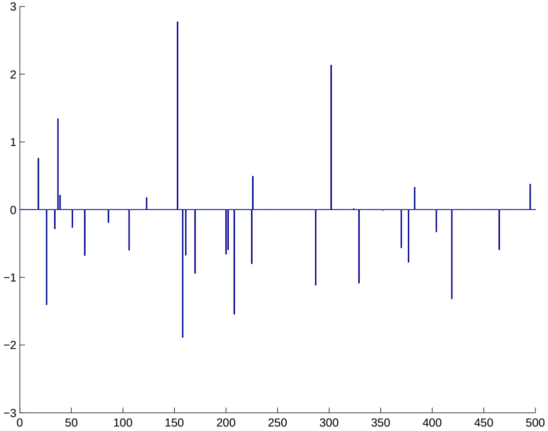
- $x^n \in \mathcal{R}^n$, $x_i = 0$ most of the time (6%)



- Bernoulli-Gaussian random variable $x = b \times s$
  - $b \sim$ Bernoulli-p, $s \sim N(0, \sigma^2)$, $x \sim \Xi(p, \sigma^2)$

Source model for sparse signals

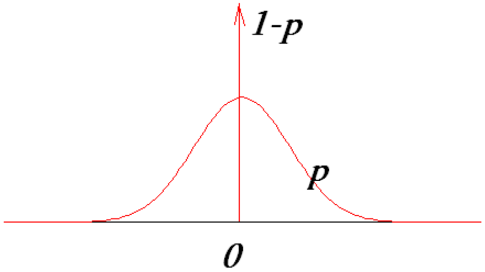
- $x^n \in \mathcal{R}^n$, $x_i = 0$ most of the time (6%)



- Bernoulli-Gaussian random variable $x = b \times s$
  - $b \sim$ Bernoulli-p, $s \sim N(0, \sigma^2)$, $x \sim \Xi(p, \sigma^2)$
  - $p = 0.05$, $\sigma^2 = 1$

Bernoulli-Gaussian random variable $x \sim \Xi(p, \sigma^2)$

- x is not a continuous random variable (no PDF), infinite entropy



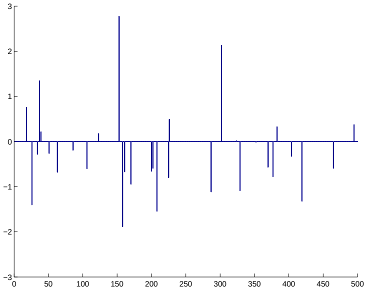
- Some classical results for continuous r.v. might need re-examination

Compressing $x^n \sim \Xi(p, \sigma^2)$

- Lossy source coding: representing real sequence by bits
  - Encoder f_n and decoder g_n: $f_n : \mathcal{R}^n \longrightarrow \{0,1\}^{nR}$ *and* $g_n : \{0,1\}^{nR} \longrightarrow \mathcal{R}^n$,

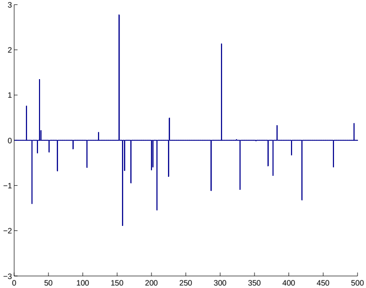
Compressing $x^n \sim \Xi(p, \sigma^2)$

- Lossy source coding: representing real sequence by bits
  - Encoder f_n and decoder g_n: $f_n : \mathcal{R}^n \longrightarrow \{0,1\}^{nR}$ *and* $g_n : \{0,1\}^{nR} \longrightarrow \mathcal{R}^n$,
  - $x^n \longrightarrow f_n \longrightarrow a^{nR} \longrightarrow g_n \longrightarrow \hat{x}^n$
  - $\hat{x}^n = g_n(f_n(x^n))$



Compressing $x^n \sim \Xi(p, \sigma^2)$

- Lossy source coding: representing real sequence by bits
  - Encoder f_n and decoder g_n: $f_n : \mathcal{R}^n \longrightarrow \{0,1\}^{nR}$ *and* $g_n : \{0,1\}^{nR} \longrightarrow \mathcal{R}^n$,
  - $x^n \longrightarrow f_n \longrightarrow a^{nR} \longrightarrow g_n \longrightarrow \hat{x}^n$
  - $\hat{x}^n = g_n(f_n(x^n))$



- Average distortion $d(x^n, \hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$

Rate distortion function $R(D)$ (1)

- $R(D)$ minimum rate such that $E(d(x^n, \hat{x}^n)) \leq D$
  - $R(D) = \min_{p_{\hat{x}|x} : \int_{x,\hat{x}} p_x(x) p_{\hat{x}|x}(\hat{x}|x) d(x,\hat{x}) \leq D} I(x;\hat{x})$ (Shannon)
  - x: discrete or continuous
  - Computation of $R(D)$ for *discrete* random variables (Blahut)

Rate distortion function $R(D)$ (1)

- $R(D)$ minimum rate such that $E(d(x^n, \hat{x}^n)) \leq D$
  - $R(D) = \min_{p_{\hat{x}|x} : \int_{x,\hat{x}} p_x(x) p_{\hat{x}|x}(\hat{x}|x) d(x,\hat{x}) \leq D} I(x; \hat{x})$ (Shannon)
  - x: discrete or continuous
  - Computation of $R(D)$ for *discrete* random variables (Blahut)
- Gaussian random variables $x \sim N(0, \sigma^2)$: $R(D) = \frac{1}{2} \log_2 \frac{\sigma^2}{D}$

Rate distortion function $R(D)$ (1)

- $R(D)$ minimum rate such that $E(d(x^n, \hat{x}^n)) \leq D$
  - $R(D) = \min_{p_{\hat{x}|x}: \int_{x,\hat{x}} p_x(x) p_{\hat{x}|x}(\hat{x}|x) d(x,\hat{x}) \leq D} I(x;\hat{x})$ (Shannon)
  - x: discrete or continuous
  - Computation of $R(D)$ for *discrete* random variables (Blahut)
- Gaussian random variables $x \sim N(0, \sigma^2)$: $R(D) = \frac{1}{2} \log_2 \frac{\sigma^2}{D}$
- Lower and upper bound of $R(D)$ for $x \sim \Xi(p, \sigma^2)$
  - Construction of a encoder-decoder pair: upper bound on $R(D)$
  - Under some rate R such that the distortion constraint cannot be satisfied: R is a lower bound

Rate distortion function $R(D)$ (2)

- Average sense, $E(d(x^n, g_n(f_n(x^n)))) \leq D$, $R(D)$

Rate distortion function $R(D)$ (2)

- Average sense, $E(d(x^n, g_n(f_n(x^n)))) \leq D$, $R(D)$
- Strong sense $\Pr(d(x^n, g_n(f_n(x^n))) > D) \rightarrow 0$, $R'(D)$

Rate distortion function $R(D)$ (2)

- Average sense, $E(d(x^n, g_n(f_n(x^n)))) \leq D$, $R(D)$
- Strong sense $\Pr(d(x^n, g_n(f_n(x^n))) > D) \to 0$, $R'(D)$
- $R(D) = R'(D)$, upper bound $R(D)$ and lower bound $R'(D)$
    - Discrete random variable: covering lemma (Csiszar)
    - Continuous random variable: quantization of the PDF (our paper) or use Kiminori 2005
    - (f_n, g_n) good in strong sense $\Longrightarrow (f'_n, g'_n)$ good in both

Rate distortion function $R(D)$ (2)

- Average sense, $E(d(x^n, g_n(f_n(x^n)))) \leq D$, $R(D)$
- Strong sense $\Pr(d(x^n, g_n(f_n(x^n))) > D) \to 0$, $R'(D)$
- $R(D) = R'(D)$, upper bound $R(D)$ and lower bound $R'(D)$
  - Discrete random variable: covering lemma (Csiszar)
  - Continuous random variable: quantization of the PDF (our paper) or use Kiminori 2005
  - (f_n, g_n) good in strong sense $\Longrightarrow (f'_n, g'_n)$ good in both
- Typical Gaussian sequence, continuous r.v. $\Longleftrightarrow$ discrete r.v.

Outline

1. Introduction
   - Bernoulli-Gaussian sequences and sparsity
   - Lossy source coding and rate distortion function $R(D)$
2. Main Result
   - Some obvious bounds on $R(D)$ of Bernoulli-Gaussian
   - An improved lower bound on $R(D)$
   - Proof techniques
   - Implications on coding sparse signals with high fidelity
3. Conclusions and Future Work

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (1)

- $x = b \times s \sim (1-p, p) \times N(0, \sigma^2)$
  - $R(D)$ for $\Xi(p, \sigma^2) = R(\frac{D}{\sigma^2})$ for $\Xi(p, 1)$, so from now on $x \sim (1-p, p) \times N(0, 1)$

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (1)

- $x = b \times s \sim (1-p,p) \times N(0,\sigma^2)$
  - $R(D)$ for $\Xi(p,\sigma^2) = R(\frac{D}{\sigma^2})$ for $\Xi(p,1)$, so from now on $x \sim (1-p,p) \times N(0,1)$
- Gaussian r.v. has the highest $R(D)$ among r.v. with the same variance
- Upper bound 1: $VAR(x) = p$
  - $R(D) \leq \frac{1}{2} \log \frac{p}{D}$

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy
  - To encode b^n, $nH(p)$ bits needed, typically np 1's in b^n
  - For the none-zero entries x^{np}, $\sum_1^{np}(x_i - \hat{x}_i)^2 \leq nD \Longrightarrow \frac{1}{np}\sum_1^{np}(x_i - \hat{x}_i)^2 \leq \frac{D}{p}$

    Gaussian: $np\frac{1}{2}\log\frac{p}{D}$ bits

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy
  - To encode b^n, $nH(p)$ bits needed, typically np 1's in b^n
  - For the none-zero entries x^{np}, $\sum_1^{np}(x_i - \hat{x}_i)^2 \leq nD \Longrightarrow \frac{1}{np}\sum_1^{np}(x_i - \hat{x}_i)^2 \leq \frac{D}{p}$

    Gaussian: $np\frac{1}{2}\log\frac{p}{D}$ bits
  - Total number of bits $n(H(p) + \frac{p}{2}\log\frac{p}{D})$
  - Rate: $R(D) \leq H(p) + \frac{p}{2}\log\frac{p}{D}$

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy
  - To encode b^n, $nH(p)$ bits needed, typically np 1's in b^n
  - For the none-zero entries x^{np}, $\sum_1^{np}(x_i - \hat{x}_i)^2 \leq nD \Longrightarrow \frac{1}{np}\sum_1^{np}(x_i - \hat{x}_i)^2 \leq \frac{D}{p}$

    Gaussian: $np\frac{1}{2}\log\frac{p}{D}$ bits
  - Total number of bits $n(H(p) + \frac{p}{2}\log\frac{p}{D})$
  - Rate: $R(D) \leq H(p) + \frac{p}{2}\log\frac{p}{D}$
- Genie gives b^n to the decoder, encoder still needs some bits to describe the non-zero part of x^n
- Lower bound: $R(D) \geq \frac{p}{2}\log\frac{p}{D}$

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy
  - To encode b^n, $nH(p)$ bits needed, typically np 1's in b^n
  - For the none-zero entries x^{np}, $\sum_1^{np}(x_i - \hat{x}_i)^2 \leq nD \Longrightarrow \frac{1}{np}\sum_1^{np}(x_i - \hat{x}_i)^2 \leq \frac{D}{p}$

    Gaussian: $np\frac{1}{2}\log\frac{p}{D}$ bits
  - Total number of bits $n(H(p) + \frac{p}{2}\log\frac{p}{D})$
  - Rate: $R(D) \leq H(p) + \frac{p}{2}\log\frac{p}{D}$
- Genie gives b^n to the decoder, encoder still needs some bits to describe the non-zero part of x^n
- Lower bound: $R(D) \geq \frac{p}{2}\log\frac{p}{D}$
- Shannon lower bound does not apply $(H(x) - H(D))$

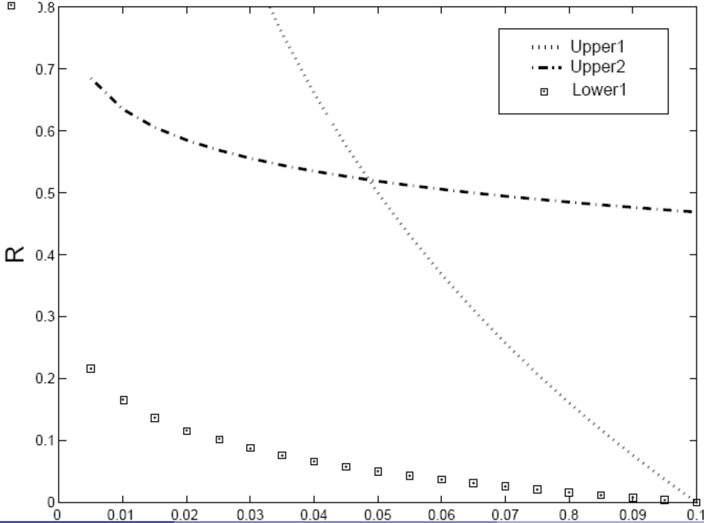
Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (2)

- Encoder observes x^n, $x_i = b_i \times s_i \longrightarrow (b_i, s_i)$
- Upper bound 2: b^n *lossless*, none-zero part of x^n lossy
  - To encode b^n, $nH(p)$ bits needed, typically np 1's in b^n
  - For the none-zero entries x^{np}, $\sum_{1}^{np}(x_i - \hat{x}_i)^2 \leq nD \Longrightarrow \frac{1}{np}\sum_{1}^{np}(x_i - \hat{x}_i)^2 \leq \frac{D}{p}$

    Gaussian: $np\frac{1}{2}\log\frac{p}{D}$ bits
  - Total number of bits $n(H(p) + \frac{p}{2}\log\frac{p}{D})$
  - Rate: $R(D) \leq H(p) + \frac{p}{2}\log\frac{p}{D}$
- Genie gives b^n to the decoder, encoder still needs some bits to describe the non-zero part of x^n
- Lower bound: $R(D) \geq \frac{p}{2}\log\frac{p}{D}$
- Shannon lower bound does not apply $(H(x) - H(D))$
- Improvement on the lower bound?

Some obvious bounds on $R(D)$ of Bernoulli-Gaussian (3)

- $\frac{p}{2}\log\frac{p}{D} \le R(D) \le \min\{H(p) + \frac{p}{2}\log\frac{p}{D}, \frac{1}{2}\log\frac{p}{D}\}, p = 0.1$



Outline

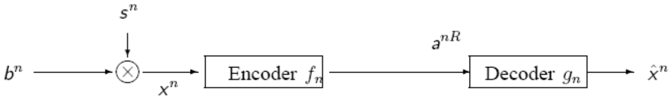
1. Introduction
   - Bernoulli-Gaussian sequences and sparsity
   - Lossy source coding and rate distortion function $R(D)$

2. Main Result
   - Some obvious bounds on $R(D)$ of Bernoulli-Gaussian
   - An improved lower bound on $R(D)$
   - Proof techniques
   - Implications on coding sparse signals with high fidelity

3. Conclusions and Future Work

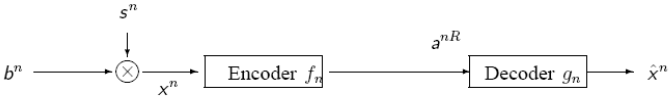
Improvement on the lower bound

- (f_n, g_n) good in both average and strong sense



Improvement on the lower bound

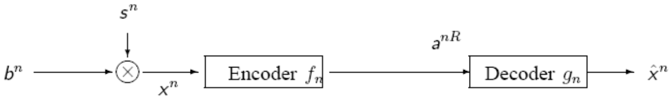
- (f_n, g_n) good in both average and strong sense



- A three step approach
- Step 1: $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n | b^n)$

Improvement on the lower bound

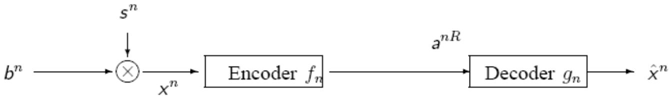
- (f_n, g_n) good in both average and strong sense



- A three step approach
- Step 1: $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n | b^n)$
  - Standard information theory (even for hybrid sources), $b \perp s$
  - Has nothing to do with $\hat{x}^n$
  - Step 2 and 3 bounding the two mutual information separately

Improvement on the lower bound

- (f_n, g_n) good in both average and strong sense



- A three step approach
- Step 1: $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n | b^n)$
  - Standard information theory (even for hybrid sources), $b \perp s$
  - Has nothing to do with $\hat{x}^n$
  - Step 2 and 3 bounding the two mutual information separately
  - $nR \geq I(a^{nR}; s^n) + I(a^{nR}; b^n | s^n)$

Step 2: bounding $I(a^{nR}; s^n | b^n)$

- $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n | b^n)$
- Genie gives b^n to the decoder, minimum "rate" needed, such that, (f_n, g_n) is good (in average sense), but (f_n, g_n) is designed in the way that genie does not exist.

Step 2: bounding $I(a^{nR}; s^n|b^n)$

- $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n|b^n)$
- Genie gives b^n to the decoder, minimum "rate" needed, such that, (f_n, g_n) is good (in average sense), but (f_n, g_n) is designed in the way that genie does not exist.
  - With genie, lower bound 1: $I(a^{nR}; s^n|b^n) \geq \frac{p}{2} \log \frac{p}{D}$

Step 2: bounding $I(a^{nR}; s^n|b^n)$

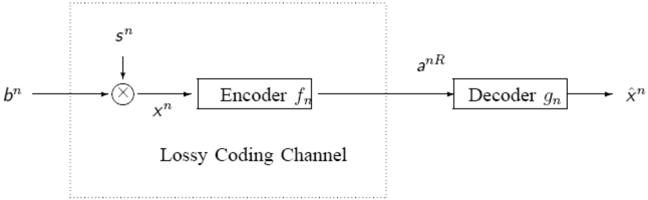
- $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n|b^n)$
- Genie gives b^n to the decoder, minimum "rate" needed, such that, (f_n, g_n) is good (in average sense), but (f_n, g_n) is designed in the way that genie does not exist.
  - With genie, lower bound 1: $I(a^{nR}; s^n|b^n) \geq \frac{p}{2} \log \frac{p}{D}$
  - Without: $I(a^{nR}; s^n|b^n) \geq \frac{p}{2} \log \frac{p}{D-(1-p)E[\hat{x}^2|b=0]}$
  - Proof: (f_n, g_n) good in average sense

Step 2: bounding $I(a^{nR}; s^n | b^n)$

- $nR \geq I(a^{nR}; b^n) + I(a^{nR}; s^n | b^n)$
- Genie gives b^n to the decoder, minimum "rate" needed, such that, (f_n, g_n) is good (in average sense), but (f_n, g_n) is designed in the way that genie does not exist.
  - With genie, lower bound 1: $I(a^{nR}; s^n | b^n) \geq \frac{p}{2} \log \frac{p}{D}$
  - Without: $I(a^{nR}; s^n | b^n) \geq \frac{p}{2} \log \frac{p}{D-(1-p)E[\hat{x}^2 | b=0]}$
  - Proof: (f_n, g_n) good in average sense
- $I(a^{nR}; s^n | b^n) \geq \frac{p}{2} \log \frac{p}{D}$, possible further improvement 1

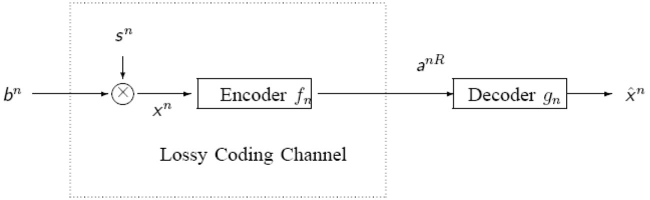
Step 3: bounding $I(a^{nR}, b^n)$ (1)

- (f_n, g_n) good in strong sense

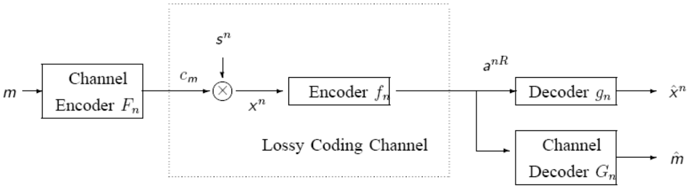


Step 3: bounding $I(a^{nR}, b^n)$ (1)

- (f_n, g_n) good in strong sense



- A new look



  - $m \in \{1, 2, \ldots 2^{n\underline{R}}\}$, $\underline{R}$ determined later
  - Standard Gallager random channel coding, $c_m \sim p$

Step 3: bounding $I(a^{nR}, b^n)$ (2)

- The "randomized capacity" for the "lossy coding" channel
  - Code book selection $\Pr(C_p = C) = \prod_{m=1}^{2^{n\underline{R}}} p^{1(c_m)}(1-p)^{n-1(c_m)}$
  - Code book is then shared between channel coders (F_n, G_n)
  - Average error: $Error_n = \sum_{C \in \mathcal{B}_n^{2^{n\underline{R}}}} \Pr(C_p = C)\left(\Pr(m \neq \hat{m}(a^{nR}) | C_p = C)\right)$

Step 3: bounding $I(a^{nR}, b^n)$ (2)

- The "randomized capacity" for the "lossy coding" channel
  - Code book selection $\Pr(C_p = C) = \prod_{m=1}^{2^{n\underline{R}}} p^{1(c_m)}(1-p)^{n-1(c_m)}$
  - Code book is then shared between channel coders (F_n, G_n)
  - Average error: $Error_n = \sum_{C \in \mathcal{B}_n^{2^{n\underline{R}}}} \Pr(C_p = C)\left(\Pr(m \neq \hat{m}(a^{nR})|C_p = C)\right)$
- $\underline{R}$ achievable if $\lim_{n\to\infty} Error_n = 0$

Step 3: bounding $I(a^{nR}, b^n)$ (2)

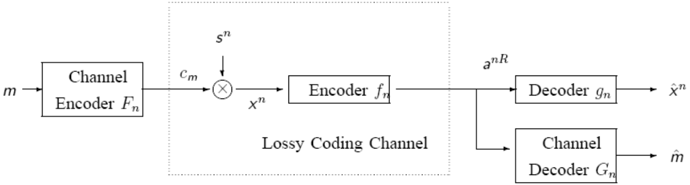
- The "randomized capacity" for the "lossy coding" channel
  - Code book selection $\Pr(C_p = C) = \prod_{m=1}^{2^{n\underline{R}}} p^{1(c_m)}(1-p)^{n-1(c_m)}$
  - Code book is then shared between channel coders (F_n, G_n)
  - Average error: $Error_n = \sum_{C \in \mathcal{B}_n^{2^{n\underline{R}}}} \Pr(C_p = C)\left(\Pr(m \neq \hat{m}(a^{nR}) | C_p = C)\right)$
- $\underline{R}$ achievable if $\lim_{n\to\infty} Error_n = 0$
- Duality: $I(a^{nR}, b^n) \geq \sup \underline{R}$
- Need to design G_n at some rate $\underline{R} > 0$, such that $Error_n \to 0$

Step 3: bounding $I(a^{nR}, b^n)$ (2)

- The "randomized capacity" for the "lossy coding" channel
  - Code book selection $\Pr(C_p = C) = \prod_{m=1}^{2^{n\underline{R}}} p^{1(c_m)}(1-p)^{n-1(c_m)}$
  - Code book is then shared between channel coders (F_n, G_n)
  - Average error: $Error_n = \sum_{C \in \mathcal{B}_n^{2^{n\underline{R}}}} \Pr(C_p = C)\left(\Pr(m \neq \hat{m}(a^{nR}) | C_p = C)\right)$
- $\underline{R}$ achievable if $\lim_{n \to \infty} Error_n = 0$
- Duality: $I(a^{nR}, b^n) \geq \sup \underline{R}$
- Need to design G_n at some rate $\underline{R} > 0$, such that $Error_n \to 0$
- $R(D) \geq \frac{1}{n}(I(a^{nR}; s^n | b^n) + I(a^{nR}, b^n)) \geq \frac{p}{2} \log \frac{p}{D} + \underline{R}$
- Previous lower bound $\frac{p}{2} \log \frac{p}{D}$

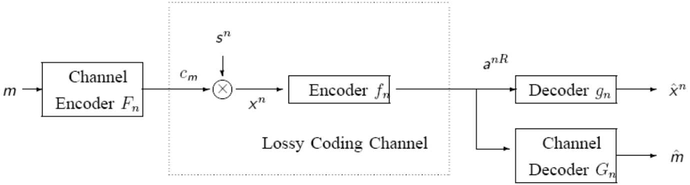
Step 3: bounding $I(a^{nR}, b^n)$ (3)

- Construction of a decoder G_n at a positive $\underline{R}$ such that $Error_n \to 0$

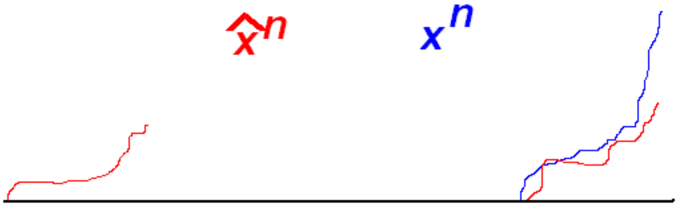


Step 3: bounding $I(a^{nR}, b^n)$ (3)

- Construction of a decoder G_n at a positive $\underline{R}$ such that $Error_n \to 0$

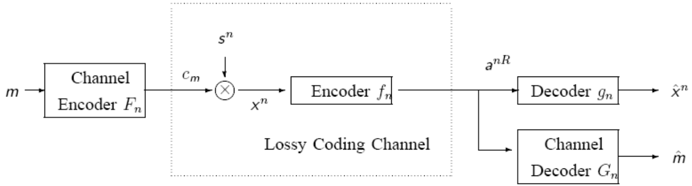


- G_n guesses m from $\hat{x}^n$: $\hat{m} = \arg\max_i \sum_{k=1}^{n} 1(|c_i(k)\hat{x}_k)| \geq L)$
  - Count the number of positions k of a codeword c_i such that $|\hat{x}_k| > L$

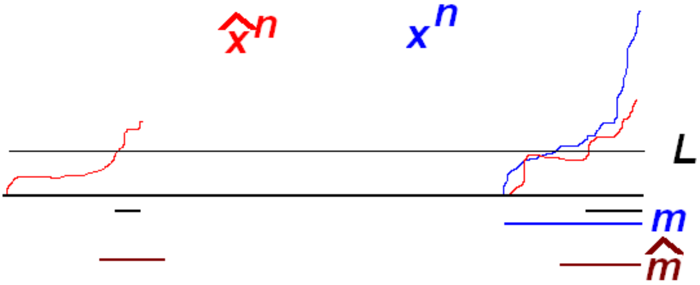


Step 3: bounding $I(a^{nR}, b^n)$ (3)

- Construction of a decoder G_n at a positive $\underline{R}$ such that $Error_n \to 0$



- G_n guesses m from $\hat{x}^n$: $\hat{m} = \arg\max_i \sum_{k=1}^{n} 1(|c_i(k)\hat{x}_k)| \geq L)$
  - Count the number of positions k of a codeword c_i such that $|\hat{x}_k| > L$

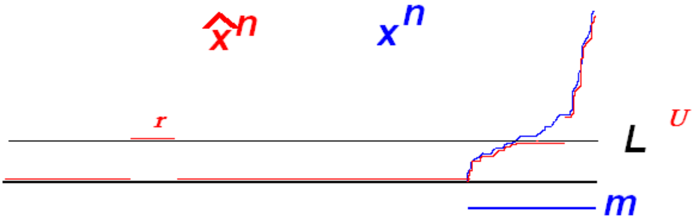


Step 3: bounding $I(a^{nR}, b^n)$ (4)

- Arbitrarily varying channel, constraints on g_n, f_n: $d(x^n, \hat{x}^n) \leq D$ for most x^n (ignore the rate constraint)

Step 3: bounding $I(a^{nR}, b^n)$ (4)

- Arbitrarily varying channel, constraints on g_n, f_n: $d(x^n, \hat{x}^n) \leq D$ for most x^n (ignore the rate constraint)
- (f_n, g_n) is the adversary. What would they do knowing the channel coding scheme G_n, L (without knowing the codebook)?
  - With a budget $\sum (x_i - \hat{x}_i)^2 < nD$
  - Minimize the hits and maximize the misses for the true codeword m



  - A zero-sum game with payoff $\underline{R}$, player 1: L, player 2: (r, U)

Step 3: bounding $I(a^{nR}, b^n)$ (5)

- The adversary's strategy (r, U), (typical behavior of x^n)

$$\underline{R} = \max_{L \geq 0}\{ \min_{U \geq L, r \in [0,1-p] : T_1(L,U,r) \leq D} h(L,U,r)\}$$

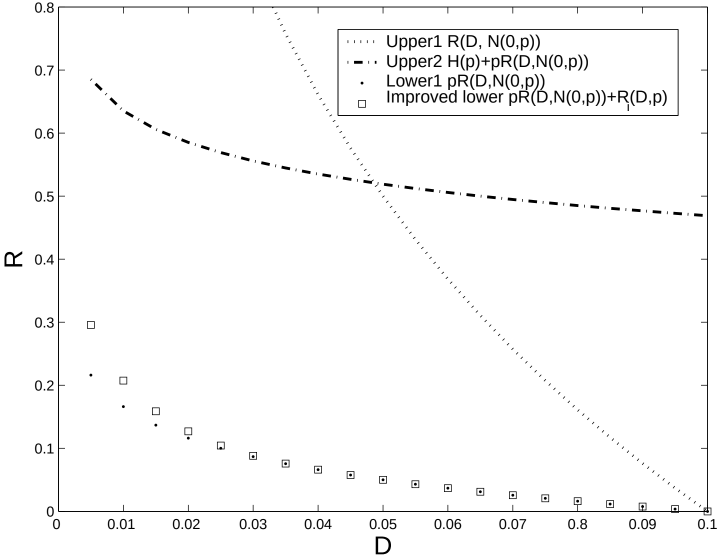
$$h(L,U,r) =$$

$$\begin{cases} (p \times \Pr(|s| > U) + r) D(\frac{p \times \Pr(|s|>U)}{p \times \Pr(|s|>U)+r} \| p) & , \frac{p \times \Pr(|s|>U)}{p \times \Pr(|s|>U)+r} \geq p \\ 0 & , \frac{p \times \Pr(|s|>U)}{p \times \Pr(|s|>U)+r} < p \end{cases}$$

$$s \sim N(0,1) \text{ and } T_1(L,U,r) = rL^2 + 2p \int_L^U (s-L)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds$$

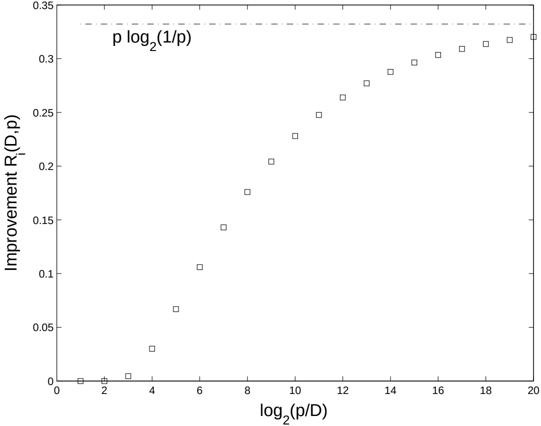
Bounds on $R(D)$ of Bernoulli-Gaussian (1)

- Improvement $\underline{R}$ increases as D decreases



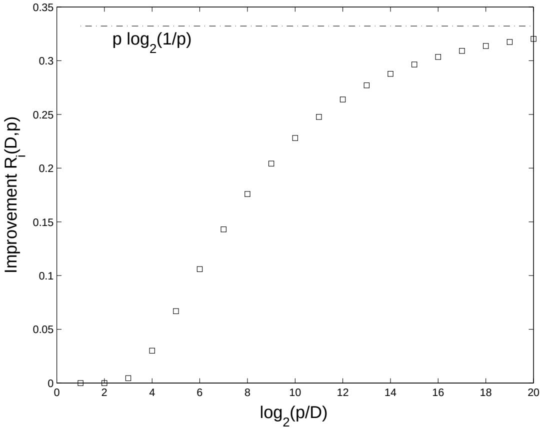
Bounds on $R(D)$ of Bernoulli-Gaussian (2)

- $\underline{R}(D) \rightarrow p \log \frac{1}{p}$



Bounds on $R(D)$ of Bernoulli-Gaussian (2)

- $\underline{R}(D) \rightarrow p \log \frac{1}{p}$



- Sparse signal $p \rightarrow 0$, the known gap $H(p)$ is almost eliminated $\frac{p \log 1p}{H(p)} \rightarrow 1$

Conclusions

- Source model for sparse signals: Bernoulli Gaussian

Conclusions

- Source model for sparse signals: Bernoulli Gaussian
- Some obvious lower and upper bounds with gap at most $H(p)$
- Improvement on the lower bound: $\sim p \log \frac{1}{p}$ for small D
- Duality between rate distortion function and the capacity of the lossy source coding channel (arbitrarily varying channel)

Conclusions

- Source model for sparse signals: Bernoulli Gaussian
- Some obvious lower and upper bounds with gap at most $H(p)$
- Improvement on the lower bound: $\sim p \log \frac{1}{p}$ for small D
- Duality between rate distortion function and the capacity of the lossy source coding channel (arbitrarily varying channel)
- Future work