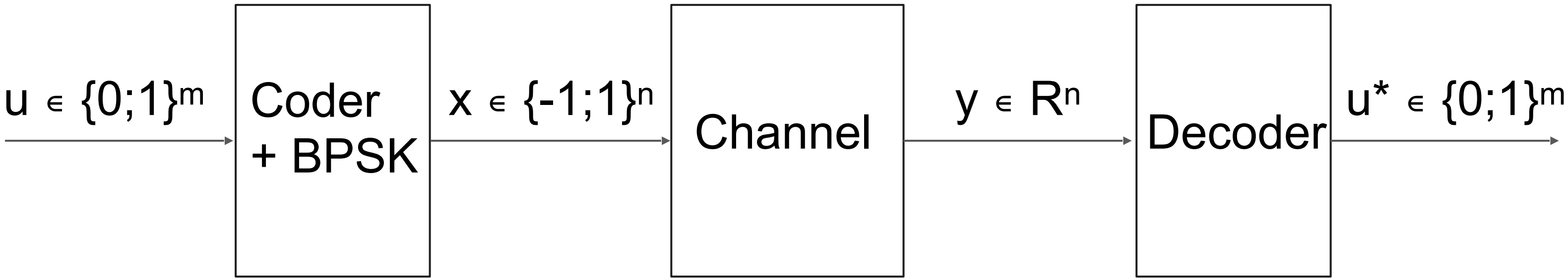
LDPC error floor estimation using Importance Sampling

Student: Uglovskii Artem

Agenda

- Introduction to LDPC codes
- Problem formulation
- Importance Sampling
- Quality metrics
- Numerical results
- Conclusion

Intro to coding theory



Simple codes

$$1010 \longrightarrow \underbrace{11\ldots1}_{n\ \text{times}}00\ldots011\ldots100\ldots0$$

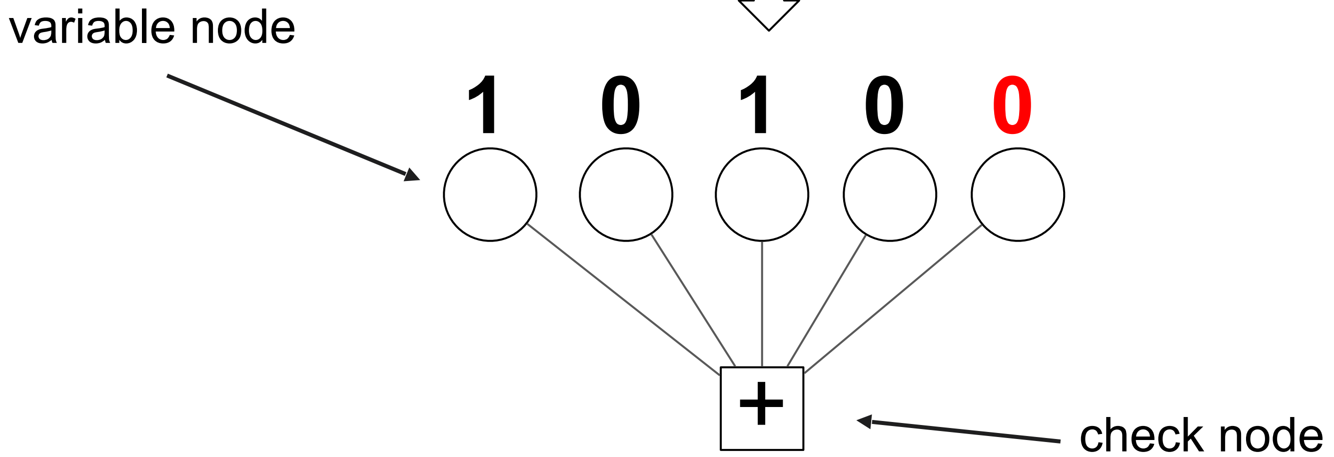
$$1010 \longrightarrow 10100$$

parity bit

Parity check codes

10100

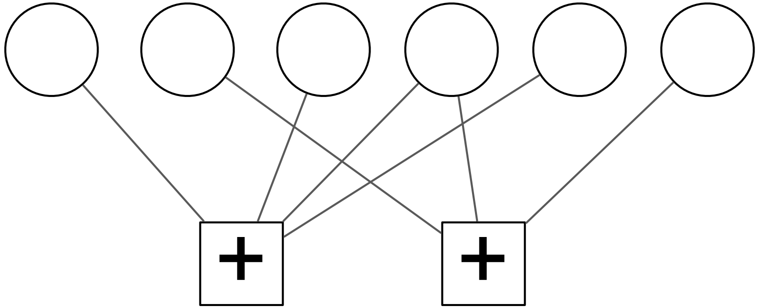
$\Rightarrow$ $x_1+x_2+x_3+x_4+x_5=0$



Parity check codes

$$\begin{cases} x_1+x_3+x_4+\color{red}{x_5}=0 \\ x_2+x_4+\color{red}{x_6}=0 \end{cases}$$

$$H = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$



LDPC codes

$$
H = \begin{bmatrix}
1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1\\
1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0\\
0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&1&0&0&0\\
0&0&1&0&0&0&1&0&0&0&0&0&0&1&0&0&0&1&0&0\\
0&0&0&1&0&0&0&0&0&0&1&0&0&0&1&0&0&0&1&0\\
0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1\\
1&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0&0&1&0&0\\
0&1&0&0&0&0&1&0&0&0&1&0&0&0&0&1&0&0&0&0\\
0&0&1&0&0&0&0&1&0&0&0&0&1&0&0&0&0&0&1&0\\
0&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0&1&0&0&0\\
0&0&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0&0&0&1
\end{bmatrix}
$$

LDPC codes

$H_c = [\ -P^T{}_{m \times (n-m)} \mid E_{m \times m}\ ]$

$G_c = [\ E_{(n-m) \times (n-m)} \mid P_{(n-m) \times m}\ ]$

A block code *C* with parity check matrix H can be defined in two ways:

1) $C = \{\ c \in \{0; 1\}^n \mid c * H^T = 0\ \}$

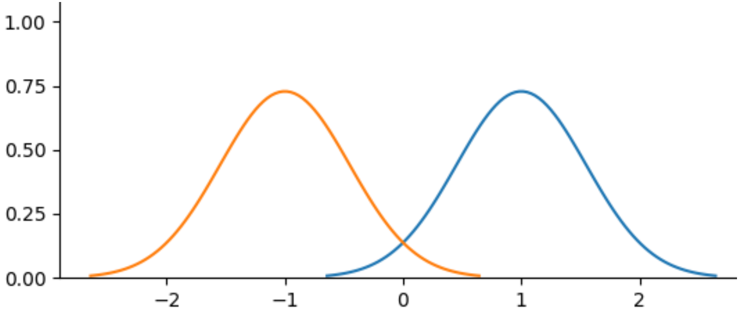
1) $C = \{\ u * G$ for every $u \in \{0; 1\}^{n-m}\ \}$

girafe ai

Decoding of LDPC codes

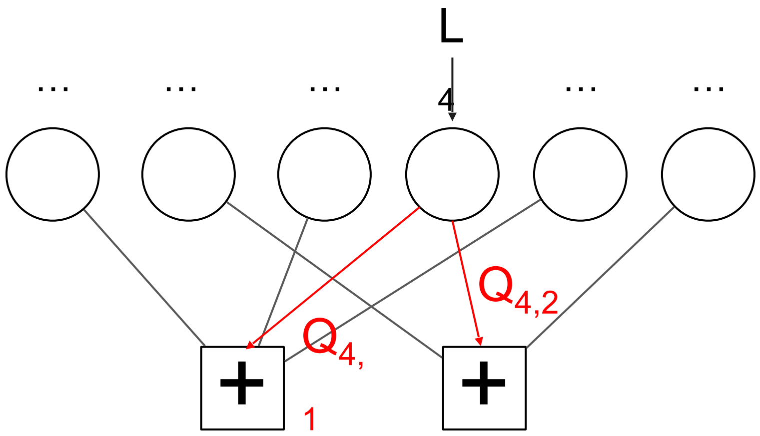
AWGN channel: noise ~ $N(0, \sigma)$

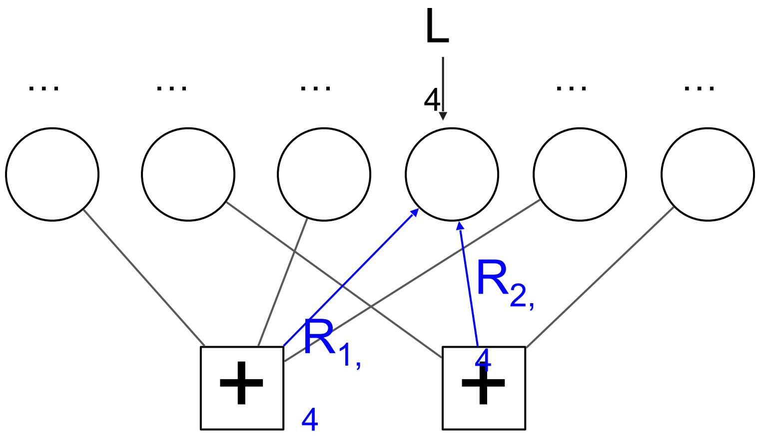
y = ±1 + noise



$L_j = \log [\, \Pr(x_j=0 \mid y_j) / \Pr(x_j=1 \mid y_j) \,] = 2y_j / \sigma^2$

Decoding of LDPC codes: MinSum





$$R_{j,i} = \prod_{k \in N(j) \setminus i} \operatorname{sign}(Q_{k,j}) * \min_{k \in N(j) \setminus i} |Q_{k,j}|$$

$$\boxed{Q_i = L_i + \sum_{k \in N(i)} R_{k,i}}$$

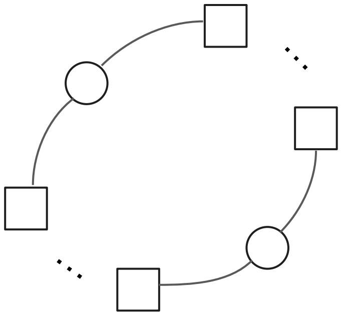
$$Q_{i,j} = L_i + \sum_{k \in N(i) \setminus j} R_{k,i}$$

Trapping Sets

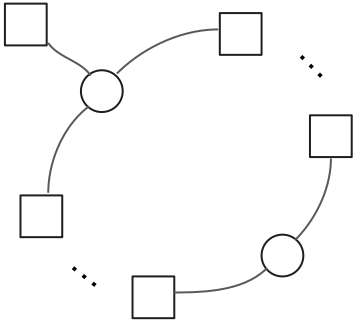
TS, denoted as ***(a, b)***, is the set of ***a*** variable nodes and adjacent check nodes, ***b*** of which have an odd number of occurrences in the given TS.

Trapping Sets

TS, denoted as ***(a, b)***, is the set of ***a*** variable nodes and adjacent check nodes, ***b*** of which have an odd number of occurrences in the given TS.



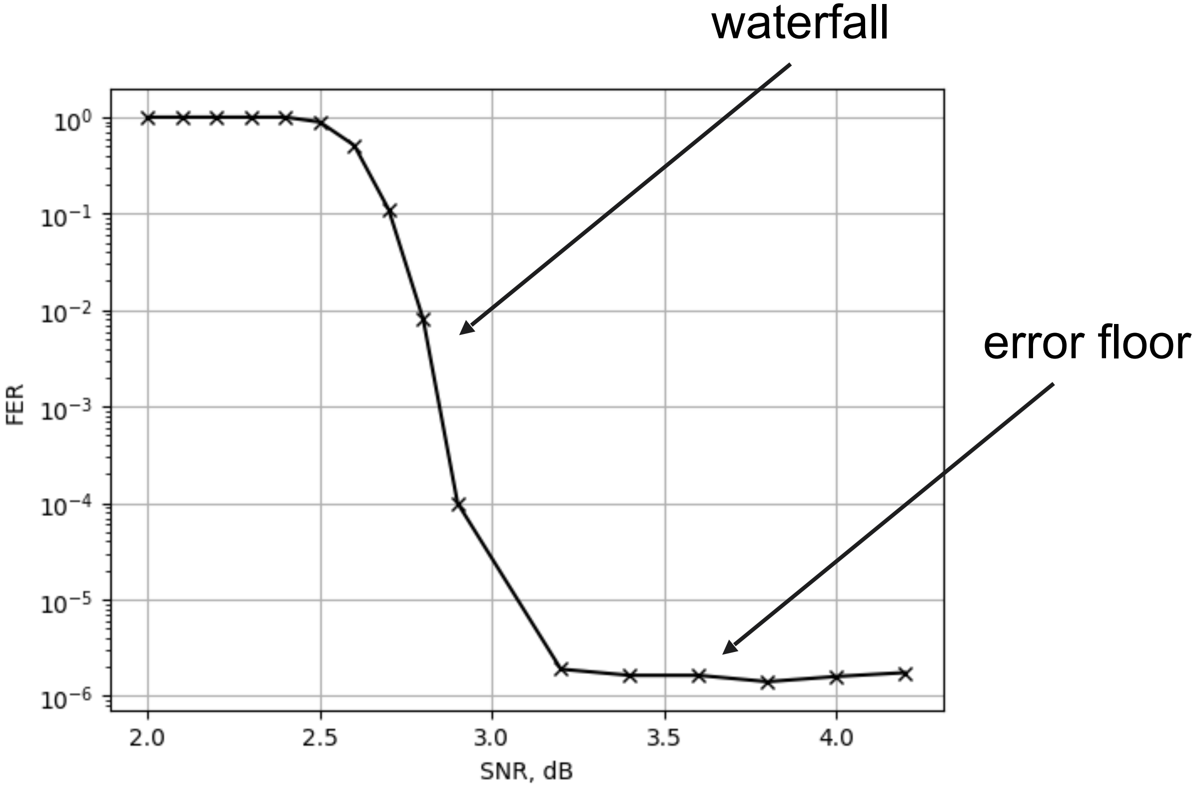
(a, 0)



(a, 1)

Paper: https://ieeexplore.ieee.org/abstract/document/7177241

Error floor



girafe ai

Problem Statement

- Very low probability can hardly be estimated with Monte Carlo method.
- There is a need in the fast and accurate error floor estimation for LDPC code construction.
- There can be several different estimations, which one is better? There is a need in the universal criterion.

Importance Sampling

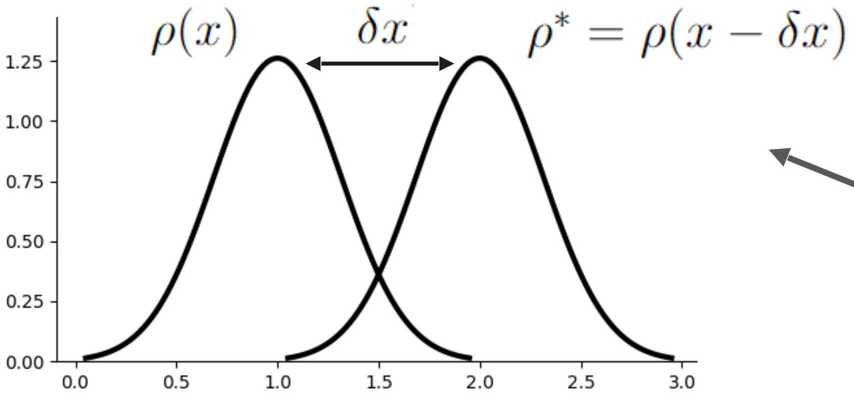
Let X be a random variable from R^n with pdf ρ and let Y be a random variable with pdf ρ^*, such that $\rho^* \neq 0$, then

$$\mathbb{E}_\rho f(X) = \int_{\mathbb{R}^n} f(\vec{x})\rho(\vec{x})d\vec{x} = \int_{\mathbb{R}^n} f(\vec{y})\frac{\rho(\vec{y})}{\rho^*(\vec{y})}\rho^*(\vec{y})d\vec{y} = \mathbb{E}_{\rho^*} f(Y)w(Y)$$

Importance Sampling

Let X be a random variable from R^n with pdf ρ and let Y be a random variable with pdf ρ^*, such that $\rho^* \neq 0$, then

$$\mathbb{E}_\rho f(X) = \int_{\mathbb{R}^n} f(\vec{x})\rho(\vec{x})d\vec{x} = \int_{\mathbb{R}^n} f(\vec{y})\frac{\rho(\vec{y})}{\rho^*(\vec{y})}\rho^*(\vec{y})d\vec{y} = \mathbb{E}_{\rho^*} f(Y)w(Y)$$



Mean Shift IS

Goals and Tasks

Goal:

Develop a criterion for comparing the results of different estimates of the LDPC code error floor and compare the IS estimates with uniform and shifted normal distributions.

Tasks:

1. Identify the limitations of the existing criteria.
2. Develop a new criterion that eliminates these limitations.
3. Compare the estimates of IS with uniformly distribution and with the mean shifted normal one.

Novelty

Most of paper use Mean-Shift IS (MS-IS) and the only criterion, which has no need in the MC values is gamma-criterion:

- M. Zhu, M. Jiang, and C. Zhao, *“Error floor estimation of qc-ldpc coded modulation with importance sampling,”* IEEE Communications Letters, vol. 25, no. 1, pp. 28–32, 2021
- Neshaastegaran, Peyman, Amir H. Banihashemi и Ramy H. Gohary (май 2021). *“Error Floor Estimation of LDPC Coded Modulation Systems Using Importance Sampling”*. В: IEEE Transactions on Communications 69.5, с. 2784—2799. doi: 10.1109/TCOMM.2021.3057625

girafe ai

IS applied to LDPC codes

- $f(\vec{x}) = \mathbb{I}_{\mathbb{A}}(\vec{x})$, where *A* is a set of inputs causing the decoder failure

  $\Pr(\vec{x} \in \mathbb{A}) \ll 1$

$\rho^*(\vec{y})$ should satisfy both:

- $\Pr(\vec{y} \in \mathbb{A}) \sim 1$

  $\mathbb{E}\, \mathbb{I}^2_{\mathbb{A}}(Y) w^2(Y)$ is small enough

Error floor estimation

Searching & choosing the Trapping Sets

↓

Choosing the distribution parameters

↓

Importance Sampling modulation

Choosing Trapping Sets

Searching all the possible TSs:

- Find all the cycles in Tanner graph
- Find unions of cycles

Choosing the dangerous ones:

- Apply Mean Shift IS with shifting E
- Choose those, which Bit Error Rate (BER) > ber threshold

Choosing the Distribution parameters

For each Signal-to-Noise Ratio (SNR) and each pdf:

- Iterate through distribution parameters
- Choose the first one, which satisfy the condition: $\Pr(\vec{y} \in \mathbb{A}) > \theta$ where θ is some predefined parameter ∊ (0, 1).

girafe ai

Importance Sampling Modulation

$$\rho^*(\vec{y}) = \frac{1}{|T|} \sum_{t \in T} \prod_{i \in t} p_{IS}(\vec{y}_i) \prod_{j \notin t} p_{MC}(\vec{y}_j)$$

T – set of chosen TSs.

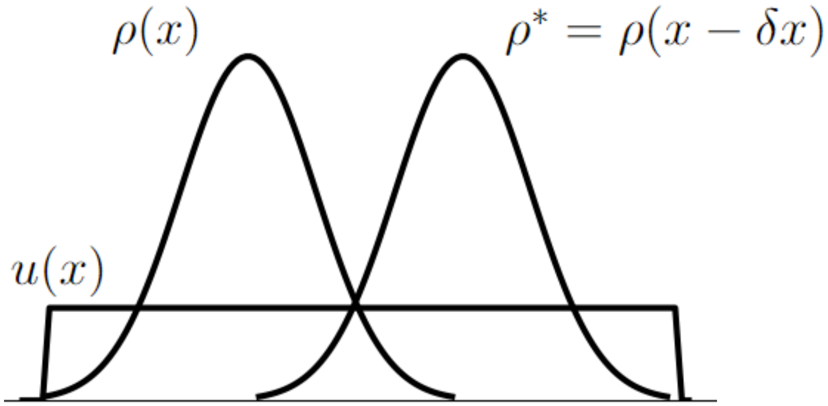
p_{IS} – sampling pdf (shifted normal or uniform).

p_{MC} – pdf of AWGN channel.

i, j – indexes of variable nodes.

girafe ai

Uniform vs Normal



If shifted normal pdf is used for sampling, then weight function ω has log-normal distribution with large variance. It yields a slow convergence.

Uniform vs Normal

For unlimited domain (e.g. normal pdf):

$$\mathbb{E}_{\rho^*} w(Y) = 1 \quad \Longrightarrow \quad \widehat{\mathbb{E}}_{\rho^*} \mathbb{I}_{\mathbb{A}}(X) = \sum_{i=1}^{N} \mathbb{I}_{\mathbb{A}}(\vec{y}_i) w(\vec{y}_i)$$

For limited domain (e.g. uniform pdf):

$$\mathbb{E}_{\rho^*} w(Y) < 1 \quad \Longrightarrow \quad \widetilde{\mathbb{E}}_{\rho^*} \mathbb{I}_{\mathbb{A}}(X) = \frac{\sum_{i=1}^{N} \mathbb{I}_{\mathbb{A}}(\vec{y}_i) w(\vec{y}_i)}{\sum_{i=1}^{N} w(\vec{y}_i)}$$

Metrics: MSE

$$\mathrm{MSE} = \sum_{\mathrm{SNR}} \left(\log \mathbb{E}_{\rho^*}^{IS} \, \mathbb{I}_{\mathbb{A}}(X) - \log \mathbb{E}_{\rho}^{MC} \, \mathbb{I}_{\mathbb{A}}(X) \right)^2$$

Pros:

+ independence from the estimator

Cons:

- need to calculate MC values

Metrics: gamma

$$\gamma(\rho^*) = \frac{1}{N}\left(\frac{\widehat{\mathbb{E}}_{\rho^*}\,\mathbb{I}^2(X)}{\widehat{\mathbb{E}}^2_{\rho^*}\,\mathbb{I}(X)} - 1\right)$$

Pros:

+ no need to calculate MC values

Cons:

- dependence from the estimator

$$\mathbb{E}_\rho f(X) \in \left(\widehat{\mathbb{E}}_\rho f(X) - \frac{C}{\sqrt{N}}, \widehat{\mathbb{E}}_\rho f(X) + \frac{C}{\sqrt{N}}\right)$$

New criterion

Most of the confidence intervals are as follows:

$$\left(p - CN^{-\alpha}, p + CN^{-\alpha}\right)$$

$C > 0$,
$\alpha > 0$,
p – predicted value
N – number of trials

New criterion

Suppose, we have two independent estimations of *N* variables *F* and *F’*, than

$$\alpha = -\frac{\ln|F - F'|}{\ln N}$$

$$\boxed{r = \frac{\alpha(\text{uniform pdf})}{\alpha(\text{normal pdf})}}$$

girafe ai

New criterion

Suppose, we have two independent estimations of *N* variables *F* and *F'*, than

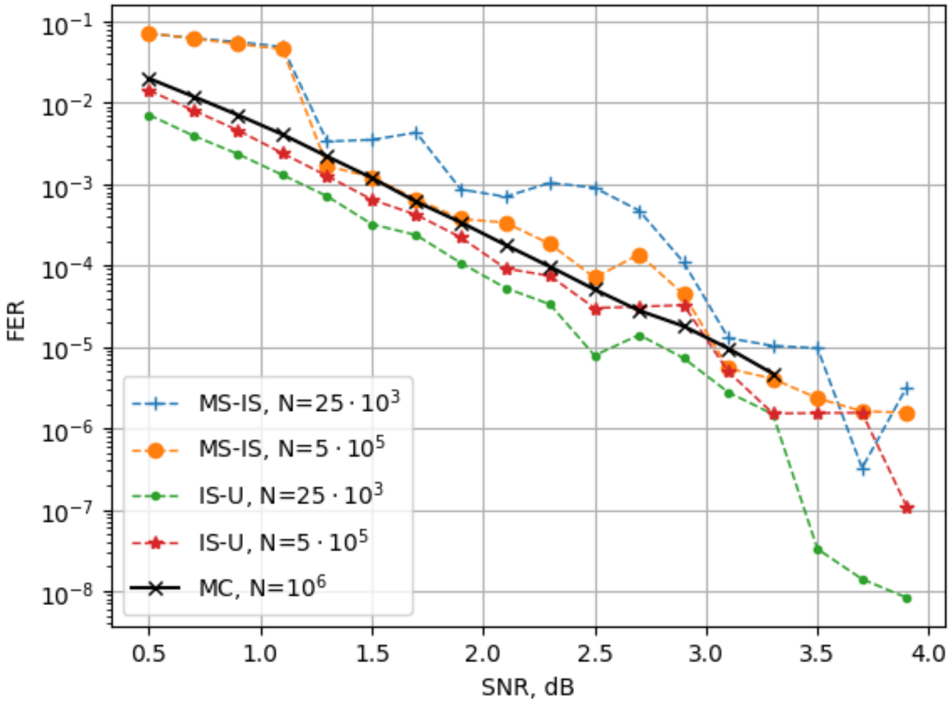
$$\alpha = -\frac{\ln|F - F'|}{\ln N}$$

$$\boxed{r = \frac{\alpha(\text{uniform pdf})}{\alpha(\text{normal pdf})}}$$

P.S.: Actually, *F* must satisfy the following condition

$$N^{\alpha}\left(F(X_1, \ldots, X_N) - p\right) \xrightarrow[N \to \infty]{d} \eta$$

Numerical Results: matrix (96, 48)

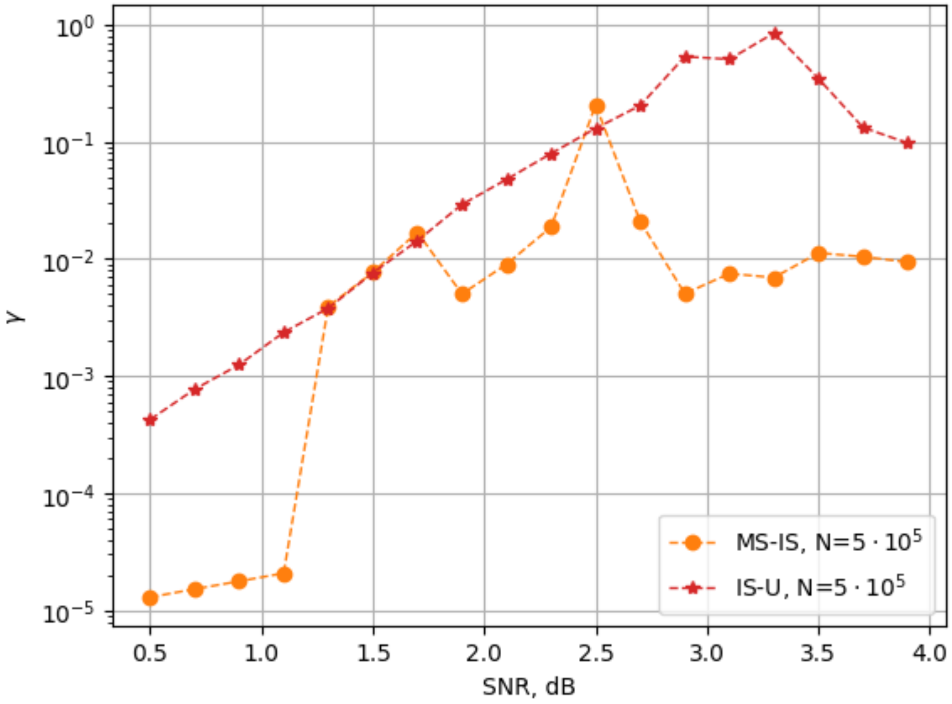


girafe ai

MSE: matrix (96, 48)

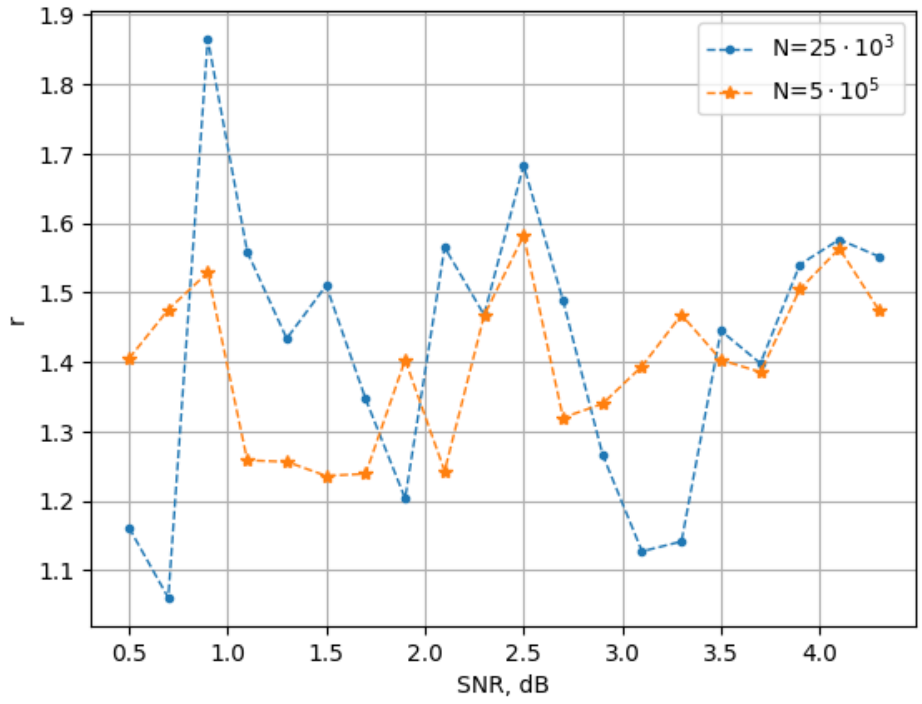
Estimation	MS-IS, $N = 25 \cdot 10^3$	IS-U, $N = 25 \cdot 10^3$	MS-IS, $N = 5 \cdot 10^5$	IS-U, $N = 5 \cdot 10^5$
MSE from 0.5 dB	9.15	3.83	3.56	0.87
MSE from 1.3 dB	6.37	2.92	0.78	0.64

Gamma: matrix (96, 48)

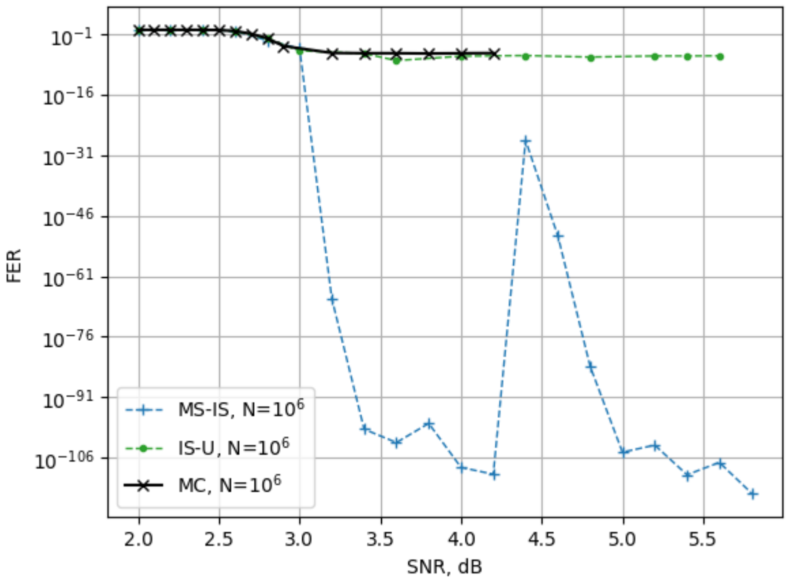


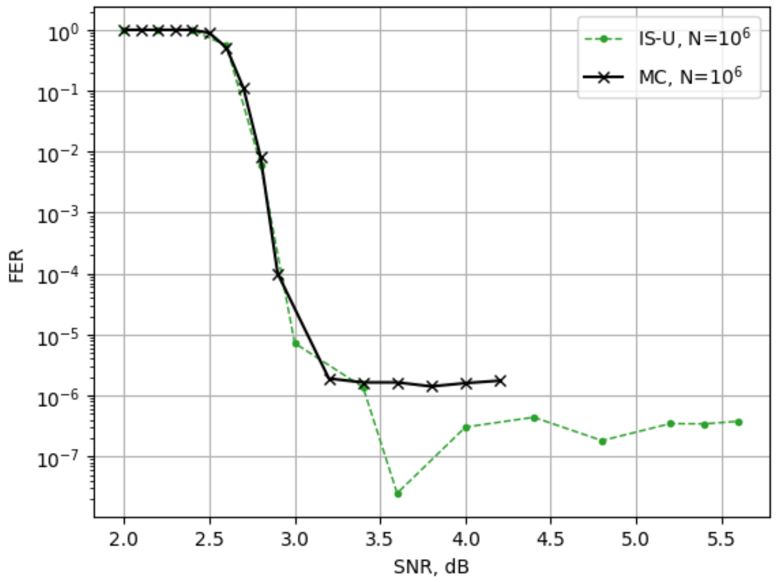
girafe ai

New criterion: matrix (96, 48)



Numerical Results: matrix (17k, 3k)





girafe ai

Conclusion

- The comparison showed that IS-U provide more accurate estimations than MS-IS
- New criterion allows us to compare estimations without MC values and regardless of the estimator’s structure

Q/A

girafe ai