

Signomial programming and its applications

Philippe Ciblat

Télécom Paris, Institut Polytechnique de Paris



Outline

- 1 Refresher on convex optimization and the road to Signomial Programming
- 2 First application : power allocation for SatCom with nonlinear impairments
- 3 Second application : quantizer optimization in distributed estimation system

Joint works with Yue Bi, Arthur Louchart, Charly Poulliat and Yue Wu

Convex optimization

Optimization problem

$$\min_{\mathbf{x}} f(\mathbf{x})$$

s.t.

$$\forall \ell, g_{\ell}(\mathbf{x}) \leq 0$$

$$\forall \ell', h_{\ell'}(\mathbf{x}) = 0$$

with f and g_{ℓ} ($\forall \ell$) convex, and $h_{\ell'}$ ($\forall \ell'$) affine

Resolution tools:

- Mathematically : KKT conditions (seldom feasible)
- Numerically : algorithms such as gradient-descent, Newton, etc

Non-convex optimization : a “simple” example

Block-Coordinate Descent

If

$$f(\mathbf{x}) = f(\mathbf{x}_1, \dots, \mathbf{x}_N)$$

with $\bullet \mapsto f(\dots, \mathbf{x}_{k-1}, \bullet, \mathbf{x}_{k+1}, \dots)$ strongly convex, then

- convergence to a stationary point if constraint set is convex and **separable!**

- **Example:** $f(\mathbf{x}) = x_1^2 x_2^2$
- **Counter-example:** downlink (power constraint: $\sum_{k=1}^N x_k \leq X_{\max}$)

General case

Successive Convex Approximation (SCA)

At each iteration i , solve

$$\mathbf{x}_{i+1}^* = \arg \min_{\mathbf{x} \in \mathcal{D}} \bar{f}_i(\mathbf{x}, \mathbf{x}_i^*)$$

with \bar{f}_i an upper-bound approximating convex function of f

- $\bar{f}_i(\mathbf{x}_i^*, \mathbf{x}_i^*) = f(\mathbf{x}_i^*)$, $\nabla_{\mathbf{x}} \bar{f}_i(\mathbf{x}, \mathbf{x}_i^*)|_{\mathbf{x}=\mathbf{x}_i^*} = \nabla_{\mathbf{x}} f(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i^*}$,
- $\forall \mathbf{x} \in \mathcal{D}$, $f(\mathbf{x}) \leq \bar{f}_i(\mathbf{x}, \mathbf{x}_i^*)$.

Then SCA converges to a stationary point of f

Problem: how finding \bar{f}_i ?

Special case: Difference of Convex (DoC) \Rightarrow easy to exhibit \bar{f}_i

- $f(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x})$
- $\bar{f}_i(\mathbf{x}, \mathbf{x}_i^*) = f_1(\mathbf{x}) - f_2(\mathbf{x}_i^*) - \nabla_{\mathbf{x}} f_2(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i^*}(\mathbf{x} - \mathbf{x}_i^*)$

Geometric Programming (GP)

Let p be a posynomial, i.e.,

$$p(\mathbf{x}) = \sum_m \beta_m \prod_{n=1}^N (x_n)^{\alpha_{m,n}}$$

with $\alpha_{m,n} \in \mathbb{R}$ and $\beta_m \in \mathbb{R}_+$

Characterization

- f and g_ℓ are posynomial, and $x_n \geq 0$, $\forall n$.
- $g_\ell(\mathbf{x}) \leq 1$
- Change of variables $y_n = \log(x_n)$
- Work on $\log(f)$ and $\log(g_\ell)$
- New problem is convex

Example

$$f(\mathbf{x}) = x_1 x_2$$

- Not jointly convex:

$$\text{Hessian : } \nabla^2 f = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

It is not a positive matrix! $([1, -1] \cdot (\nabla^2 f) \cdot [1, -1]^T = -2)$

- but $j : \mathbf{y} \mapsto \log(f(e^{\mathbf{y}}))$ is **convex** since

$$\begin{aligned} j(\mathbf{y}) &= \log(e^{y_1} e^{y_2}) \\ &= y_1 + y_2 \end{aligned}$$

Complementary Geometric Programming (CGP)

f and g_ℓ ratio of posynomials

Characterization

$$\min_{\mathbf{x}} \frac{p_0(\mathbf{x})}{q_0(\mathbf{x})} \quad \text{s.t.} \quad \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \leq 1 \quad \forall i = 1, \dots, K$$

where p_i and q_i are posynomial functions $\forall i = 0, \dots, K$.

- CGP are nonconvex and become GP when q_i are monomials.
- SCA by replacing posynomial denominator with approximate monomial
- Convergence to a stationary point

Monomial approximation and lower bound of posynomial

Let

- $Q_m(\mathbf{x}) = \beta_m \prod_{n=1}^N x_n^{\alpha_{m,n}}$ be a monomial (with $\beta_m > 0$)
- $Q(\mathbf{x}) := \sum_m Q_m(\mathbf{x})$ be a posynomial

Result

$$Q(\mathbf{x}) \geq \tilde{Q}(\mathbf{x}) := \prod_m \left(\frac{Q_m(\mathbf{x})}{\delta_m} \right)^{\delta_m}$$

In addition, if $\delta_m = Q_m(\mathbf{x}_0)/Q(\mathbf{x}_0)$, then

- $Q(\mathbf{x}_0) = \tilde{Q}(\mathbf{x}_0)$,
- and

$$\frac{\partial Q}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}_0} = \frac{\partial \tilde{Q}}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}_0}$$

Sketch of proof

- Comparison between arithmetic mean and geometric mean:

$$\sum_m \delta_m z_m \geq \prod_m z_m^{\delta_m}$$

with $\delta_m \geq 0$ and $\sum_m \delta_m = 1$.

- Consider $z_m = Q_m(\mathbf{x})/\delta_m$ and $\delta_m = Q_m(\mathbf{x}_0)/Q(\mathbf{x}_0)$
- $\tilde{Q}(\mathbf{x}_0) = \prod_m (Q(\mathbf{x}_0))^{Q_m(\mathbf{x}_0)/Q(\mathbf{x}_0)} = Q(\mathbf{x}_0)^{\sum_m Q_m(\mathbf{x}_0)/Q(\mathbf{x}_0)} = Q(\mathbf{x}_0)$
- $\frac{\partial \log Q}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}_0} = \frac{\sum_m \partial Q_m / \partial \mathbf{x} |_{\mathbf{x}=\mathbf{x}_0}}{\sum_m Q_m(\mathbf{x}_0)}$
- $\log \tilde{Q}(\mathbf{x}) = \sum_m \delta_m \log_2(Q_m(\mathbf{x})/\delta_m)$
 $\frac{\partial \log \tilde{Q}}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}_0} = \sum_m \delta_m \frac{\partial Q_m / \partial \mathbf{x} |_{\mathbf{x}=\mathbf{x}_0}}{Q_m(\mathbf{x}_0)} = \sum_m \frac{\partial Q_m / \partial \mathbf{x} |_{\mathbf{x}=\mathbf{x}_0}}{\sum_m Q_m(\mathbf{x}_0)}$

Signomial Programming (SP)

As for CGP but $\beta_m \in \mathbb{R}$

Characterization

$$\min_{\mathbf{x}} \frac{a_0(\mathbf{x}) - b_0(\mathbf{x})}{c_0(\mathbf{x}) - d_0(\mathbf{x})} \quad \text{s.t.} \quad \frac{a_i(\mathbf{x}) - b_i(\mathbf{x})}{c_i(\mathbf{x}) - d_i(\mathbf{x})} \leq 1 \quad \forall i = 1, \dots, K$$

where a_i, b_i, c_i and d_i are posynomial functions and $c_i - d_i > 0$, $\forall i = 0, \dots, K$.

- SP problems are nonconvex and can be converted into CGP

$$\min_{\mathbf{x}, t} t \quad \text{s.t.} \quad \frac{a_i(\mathbf{x}) - b_i(\mathbf{x})}{c_i(\mathbf{x}) - d_i(\mathbf{x})} \leq t^{\delta_{0,i}} \rightarrow \frac{a_i(\mathbf{x}) + t^{\delta_{0,i}} d_i(\mathbf{x})}{b_i(\mathbf{x}) + t^{\delta_{0,i}} c_i(\mathbf{x})} \leq 1.$$

- SCA and converges to a stationary point

Example 1

$$\min_{\mathbf{P}} \sum_k P_k \text{ s.t. } R_k \geq R_k^t, \forall k$$

Non convex optimization due to the constraint

Rewrite the constraint

$$R_k \geq R_k^t \Leftrightarrow \log_2 \left(1 + \frac{G_k P_k}{\sum_{m \neq k} G_m P_m + P_w} \right) \geq R_k^t$$

$$\Leftrightarrow \sum_{m \neq k} G_m G_k^{-1} P_m P_k^{-1} + G_k^{-1} P_w P_k^{-1} \leq \frac{1}{2^{R_k^t} - 1}$$

Geometric programming

Remark: SINR_k^{-1} is a posynomial but $(1 + \text{SINR}_k)^{-1}$ is just a ratio of posynomials

Example 2

$$\max_{\mathbf{P}} \sum_k \log_2 \left(1 + \frac{G_k P_k}{\sum_{m \neq k} G_m P_m + P_w} \right) \quad \text{s.t.} \quad \sum_k P_k \leq P_{\max}$$

\Leftrightarrow

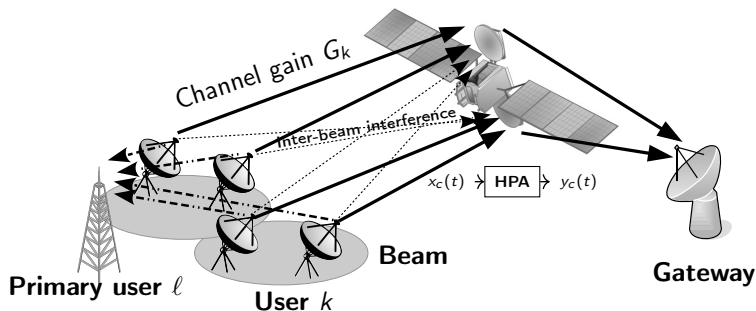
$$\min_{\mathbf{P}} \prod_k \left(\frac{\sum_{m \neq k} G_m P_m + P_w}{\sum_m G_m P_m + P_w} \right) \quad \text{s.t.} \quad \sum_k P_k \leq P_{\max}$$

\Leftrightarrow

$$\min_{\mathbf{P}} \frac{\prod_k (\sum_{m \neq k} G_m P_m + P_w)}{\prod_k (\sum_m G_m P_m + P_w)} \quad \text{s.t.} \quad \sum_k P_k \leq P_{\max}$$

Complementary Geometric programming

Application 1: power allocation with nonlinear impairments



- Multi-beam return link from terrestrial users to satellite (gain G_k for user k) with full frequency reuse → **inter-beam interference**
- OFDMA intra-beam scheduling
- Terrestrial primary system → **interference constraint**
- Nonlinear amplifier on satellite board → **non-linear interference**

Signal model

Volterra series model for High Power Amplifier (HPA)

$$y_c(t) = \gamma_1 x_c(t) + \gamma_3 x_c(t) \overline{x_c(t)} + w_c(t)$$

For user k , we get

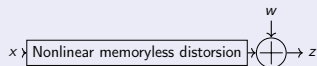
$$z_k = z_k^L + z_k^{\text{NL}} + w_k$$

Let

- $\mathcal{P}_L(k) = \mathbb{E}[|z_k^L|^2]$ be the auto-correlation of the linear part,
- $\mathcal{P}_{\text{NL}}(k) = \mathbb{E}[|z_k^{\text{NL}}|^2]$ be the auto-correlation of the nonlinear part,
- $\mathcal{P}_{\text{LNL}}(k) = \mathbb{E}[z_k^L \overline{z_k^{\text{NL}}}]$ be the cross-correlation between the linear and nonlinear parts.

Capacity expressions

Pinsker's formula (Gaussian codebooks)



$$C = \log_2 \left(1 + \frac{\mathbb{E}[x\bar{z}] \mathbb{E}[z\bar{x}]}{\mathbb{E}[|x|^2] \mathbb{E}[z\bar{z}] - \mathbb{E}[x\bar{z}] \mathbb{E}[z\bar{x}]} \right)$$

- Assuming optimal decoder,

$$C(k) = \log_2 (1 + Q(k))$$

with

$$Q(k) = \frac{\mathcal{P}_L^2(k) + 2\mathcal{P}_L(k)\Re\{\mathcal{P}_{LNL}(k)\} + |\mathcal{P}_{LNL}(k)|^2}{\mathcal{P}_L(k)\mathcal{P}_{NL}(k) + \mathcal{P}_L(k)\mathcal{P}_W - |\mathcal{P}_{LNL}(k)|^2}$$

- Assuming nonlinear part as noise,

$$\underline{C}(k) = \log_2 (1 + \underline{Q}(k))$$

with

$$\underline{Q}(k) = \frac{\mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W}$$

Capacity expressions (cont'd)

$$\mathcal{P}_L(k) = \gamma_1^2 G_k P_k,$$

$$\mathcal{P}_{NL}(k) = 4\gamma_3^2 \alpha^{(1)} G_k P_k \sum_{k', k''} G_{k'} G_{k''} P_{k'} P_{k''}$$

$$+ 2\gamma_3^2 \alpha^{(2)} \sum_{k_1, k_2, k_3 | k = k_1 + k_2 - k_3} G_{k_1} G_{k_2} G_{k_3} P_{k_1} P_{k_2} P_{k_3}$$

$$+ 4\gamma_3^2 \beta^{(1)} (\delta_{k,1}^c G_{k-1} P_{k-1} + \delta_{k,K}^c G_{k+1} P_{k+1}) \sum_{k', k''=1}^K G_{k'} G_{k''} P_{k'} P_{k''}$$

$$+ 2\gamma_3^2 \beta^{(2)} \sum_{k_1, k_2, k_3 | k = k_1 + k_2 - k_3 \pm 1} G_{k_1} G_{k_2} G_{k_3} P_{k_1} P_{k_2} P_{k_3}$$

$$\mathcal{P}_{LNL}(k) = 2\gamma_1 \gamma_3 \lambda G_k P_k \sum_{k'} G_{k'} P_{k'}$$

- $\alpha^{(1)}$, $\alpha^{(2)}$, $\beta^{(1)}$, $\beta^{(2)}$, λ are positive.
- $\mathcal{P}_L(k)$, $\mathcal{P}_{NL}(k)$, $\mathcal{P}_{LNL}(k)$ are posynomials

Power minimization (with \underline{C})

$$\min_{\mathbf{P}} \sum_{k=1}^K P_k \quad \text{s.t.} \quad \log_2 \left(1 + \frac{\mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W} \right) \geq R_k \quad \forall k$$

which is equivalent to

$$\min_{\mathbf{P}} \sum_{k=1}^K P_k$$

s.t.

$$\mathcal{P}_L(k)^{-1} (\mathcal{P}_{NL}(k) + \mathcal{P}_W) \leq \frac{1}{2^{R_k} - 1} \quad \forall k = 1, \dots, K$$

Geometric Programming

Maxmin data rate (with \underline{C})

$$\max_{\mathbf{P}} \min_k \frac{\mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W}$$

which is equivalent to

$$\begin{array}{ll} \max_{\mathbf{P}, t} & \min_{\mathbf{P}, t} t^{-1} \\ \text{s.t.} & \text{s.t.} \\ \frac{\mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W} \geq t \quad \forall k & t \mathcal{P}_L(k)^{-1} (\mathcal{P}_{NL}(k) + \mathcal{P}_W) \leq 1 \quad \forall k \end{array}$$

Geometric Programming

Sum-rate maximization (with \underline{C})

$$\begin{aligned}\max_{\mathbf{P}} \sum_{k=1}^K \log_2 \left(1 + \frac{\mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W} \right) &= \max_{\mathbf{P}} \prod_{k=1}^K \frac{\mathcal{P}_{NL}(k) + \mathcal{P}_W + \mathcal{P}_L(k)}{\mathcal{P}_{NL}(k) + \mathcal{P}_W} \\ &= \min_{\mathbf{P}} \prod_{k=1}^K \frac{\mathcal{P}_{NL}(k) + \mathcal{P}_W}{\mathcal{P}_{NL}(k) + \mathcal{P}_W + \mathcal{P}_L(k)} \\ &= \min \frac{\text{posynomial}}{\text{posynomial}}\end{aligned}$$

Complementary Geometric Programming

Sum-rate maximization (with C)

Due to sign $-$ in $Q(k)$, we have

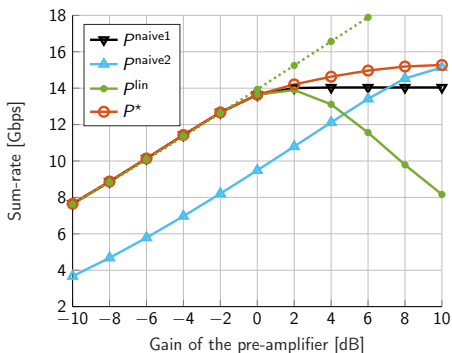
$$\min \frac{\text{signomial}}{\text{signomial}}$$

under ratio of signomials.

Signomial Programming

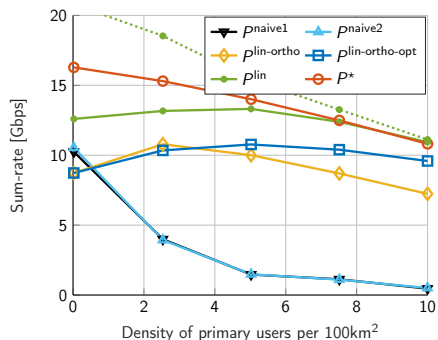
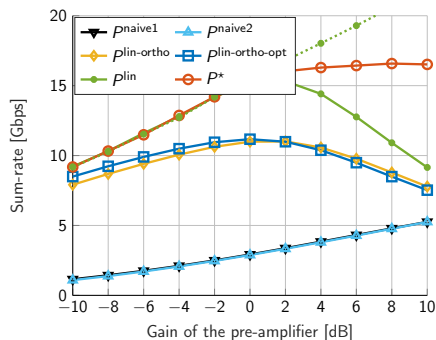
Numerical illustrations - 1

- Single-beam, no primary users : only **non-linear interference**
- $K = 6$ users, $P_{\max} = 50\text{W}$ (47dBm), $\gamma_3 = 0.05$
- Rainy weather (G_k strongly different between users)
- Sota
 - ▶ Naive1: $P_k = P \forall k$ and line search to find optimal P
 - ▶ Naive2: $G_k P_k = P \forall k$ and line search to find optimal P
 - ▶ Lin: non-linear interference not taken into account in C



Numerical illustrations - 2

- Multi-beam, primary users
- Additional Sota
 - ▶ Lin-ortho: Lagunas's algo for orthogonal beam and non-linear interference not taken into account in C
 - ▶ Lin-ortho-opt: optimal algo for orthogonal beam and non-linear interference not taken into account in C



Application 2: quantizer optimization in distributed system

- K sensors collecting a noisy sample y_k at sensor k

$$y_k = \theta + w_k$$

where $w_k \sim \mathcal{N}(0, \sigma_w^2)$ and $\theta \in [-1, 1]$

- Communication of y_k to Fusion Center with a B bits quantizer Q_k

$$q_k = Q_k(y_k)$$

where

$$Q_k(u) \triangleq \begin{cases} 0 & \text{if } u < \tau_{k,1}, \\ L & \text{if } u \geq \tau_{k,L}, \\ i & \text{if } \tau_{k,i} \leq u < \tau_{k,i+1}, \text{ for } i \in \{1, \dots, L-1\}, \end{cases}$$

with $L = 2^B - 1$ and $\{\tau_{k,i}\}_{i \in \{1, \dots, L\}}$ increasing threshold's sequence

Goal

Build in a relevant way the threshold sequences $\{\tau_{k,i}\}_{i \in \{1, \dots, L\}}$

Cramer-Rao Bound (CRB)

$$\mathbb{E}[(\hat{\theta} - \theta)^2] \geq \text{CRB}_{\theta}(\boldsymbol{\tau}) = \frac{1}{\sum_{k=1}^K F_{\theta}^{(k)}(\boldsymbol{\tau})},$$

with $F_{\theta}^{(k)}(\boldsymbol{\tau})$ the Fisher information for one sample of the k -th sensor, and

$$F_{\theta}^{(k)}(\boldsymbol{\tau}) = \eta_1(\tau_{k,1}, \theta) + \eta_L(\tau_{k,L}, \theta) + \sum_{i=1}^{L-1} \eta(\tau_{k,i}, \tau_{k,i+1}, \theta)$$

with $\eta_1(t, \theta) = \eta(-\infty, t, \theta)$, $\eta_L(t, \theta) \triangleq \eta(t, +\infty, \theta)$,

$$\eta(t, t', \theta) = \frac{(\psi(t' - \theta) - \psi(t - \theta))^2}{\Psi(t' - \theta) - \Psi(t - \theta)},$$

and $\Psi(\bullet) = \Phi(\bullet/\sigma_w)$ and $\psi(\bullet) = \phi(\bullet/\sigma_w)/\sigma_w$ where Φ and ϕ are the cdf and pdf for standard normal distribution

Asymptotic CRB ($K \rightarrow \infty$)

- Assume the thresholds $\{\tau_{k,i}\}_{k,i}$ correspond to a realization of a random variable driven by the pdf g and the cdf G .
- Let $\lambda_{1/L}$ be the marginal pdf of $\tau_{k,1/L}$ (the first/last threshold), and λ_i the joint pdf two consecutive thresholds.

$$\text{CRB}_{\theta}(\boldsymbol{\tau}) \approx \frac{1}{K \cdot \underline{F}_{\theta}(g)}, \text{ with } \underline{F}_{\theta}(g) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K F_{\theta}^{(k)}(\boldsymbol{\tau})$$

Then

$$\begin{aligned} \underline{F}_{\theta}(g) &= \int_{-\infty}^{\infty} \eta_1(t, \theta) \lambda_1(t) dt + \int_{-\infty}^{\infty} \eta_L(t, \theta) \lambda_L(t) dt \\ &+ \sum_{i=1}^{L-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\tau'} \eta(t, t', \theta) \lambda_i(t, t') dt dt' \end{aligned}$$

where $\lambda_1(t) = c[1 - G(t)]^{L-1}g(t)$, $\lambda_L(t) = c[G(t)]^{L-1}g(t)$, and $\lambda_i(t, t') = c_i[G(t)]^{i-1}[1 - G(t')]^{L-i}g(t)g(t')$

Optimization problem : worst case

$$\max_g \min_{\theta \in [-1,1]} \underline{F}_\theta(g)$$

- Discretization of θ with M points : $\theta_j, \forall j \in \{1, \dots, M\}$
- Discretization of g with N points on $\{p_1, \dots, p_N\}$.
 - ▶ Variable of optimization: $\mathbf{a} = \{a_m\}_{m \in \{1, \dots, N\}}$ with

$$a_m = \frac{g(p_m)}{\sum_{m'=1}^N g(p_{m'})} \quad \forall m \in \{1, \dots, N\}$$

- ▶ $\underline{F}_\theta(g)$ leads to $f_j(\mathbf{a})$

$$\begin{aligned} f_j(\mathbf{a}) &= \sum_{\ell=1}^N d_{\ell,j}^{(1)} (1 - \sum_{j=1}^{\ell} a_j)^{L-1} a_\ell + \sum_{\ell=1}^N d_{\ell,j}^{(L)} (\sum_{j=1}^{\ell} a_j)^{L-1} a_\ell \\ &+ \sum_{i=1}^{L-1} \sum_{\ell_2=1}^N \sum_{\ell_1=1}^{\ell_2} d_{\ell_1, \ell_2, j}^{(i)} (\sum_{j=1}^{\ell_1} a_j)^{i-1} (1 - \sum_{j=1}^{\ell_2} a_j)^{L-i} a_{\ell_1} a_{\ell_2} \end{aligned}$$

Algorithm

Consequently,

$$\max_{\mathbf{a}} \min_{j \in \{1, \dots, M\}} f_j(\mathbf{a}) \text{ s.t. } \sum_{m=1}^N a_m = 1$$

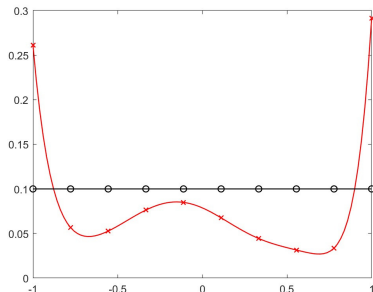
Optimization problem

$$\begin{aligned} & \max_{\mathbf{a}, x} x \\ & \text{s.t. } f_j(\mathbf{a}) \geq x, \forall j \in \{1, \dots, M\}, \text{ and } \sum_{m=1}^N a_m = 1 \end{aligned}$$

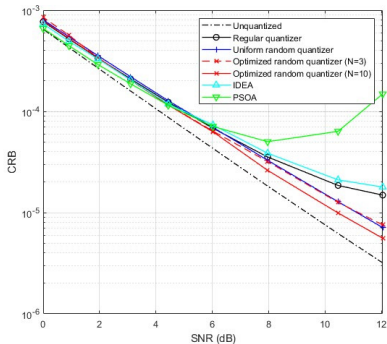
Signomial Programming

Numerical illustrations

$K = 2000$, $L = 4$, $M = 8$, and $N = 10$



Optimal (red) g



MSE versus SNR

Conclusion

- Signomial programming is useful as soon as nonlinearity comes from multivariate polynomials
- Other applications : DAC/ADC optimization

Related publications:

- ▶ A. Louchart et al., *Some Power Allocation Algorithms for Cognitive Uplink Satellite Systems*, Eurasip Journal of Wireless Communications and Networking, 2023.
- ▶ Y. Bi et al., *Multi-bit Quantizer Design for Distributed Parameter Estimation*, IEEE Statistical Signal Processing Workshop, 2025.