

CSOPT : Calcul scientifique et optimisation  
Travaux pratiques  
Optimisation sans contraintes

Yassine Jamoud                    Samy Haffoudhi

2 novembre 2021

# 1 Minimisation de la fonction de Rosenbrock

On s'intéresse à la fonction de Rosenbrock définie par :

$$f_0(x) = \sum_{i=1}^{n-1} b(x_{i+1} - x_i^2)^2 + (1 - x_i)^2, \quad \forall x \in \mathbb{R}^n \text{ et } b \in \mathbb{R}^+.$$

## 1.1 Travail préliminaire

1. On a  $\forall n \geq 2, \forall x \in \mathbb{R}^n$  et  $b \in \mathbb{R}^+$  :

$$\nabla f_0(x) = \begin{bmatrix} -4bx_1(x_2 - x_1^2) - 2(1 - x_1) \\ 2b(x_2 - x_1^2) - 4bx_2(x_3 - x_2^2) - 2(1 - x_2) \\ \vdots \\ 2b(x_{n-1} - x_{n-2}^2) - 4bx_{n-1}(x_n - x_{n-1}^2) - 2(1 - x_{n-1}) \\ 2b(x_n - x_{n-1}^2) \end{bmatrix}$$

Soit pour  $n = 2, f_0(x) = b(x_2 - x_1^2)^2 + (1 - x_1)^2, \forall x \in \mathbb{R}^2$  et  $b \in \mathbb{R}^+$ .

$$\nabla f(x) = \begin{bmatrix} -4bx_1(x_2 - x_1^2) - 2(1 - x_1) \\ 2b(x_2 - x_1^2) \end{bmatrix} \quad \nabla^2 f(x) = \begin{bmatrix} -4b(x_2 - x_1^2) + 8bx_1^2 + 2 & -4bx_1 \\ -4bx_1 & 2b \end{bmatrix}$$

2. Pour trouver le minimiseur, on cherche à annuler le gradient :

$$\nabla f_0(x_0) = 0 \iff \begin{cases} -4bx_1(x_2 - x_1^2) - 2(1 - x_1) = 0 \\ 2b(x_2 - x_1^2) = 0 \end{cases} \iff \begin{cases} x_2 = x_1^2 \\ 1 - x_1 = 0 \end{cases}$$

D'où  $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  avec  $f_0(x_0) = 0$

Or,  $f_0$ , comme somme de carrés, est à valeurs positives donc,  $x_0$  est bien minimiser de cette fonction.

3. On pose  $f_0$  sous la forme d'un critère de moindres carrés non-linéaires tel que  $f_0(x) = \frac{1}{2}r_1^2(x) + \frac{1}{2}r_2^2(x)$

$$\text{On a alors } r(x) = \begin{bmatrix} \sqrt{2b}(x_2 - x_1^2) \\ \sqrt{2}(1 - x_1) \end{bmatrix}$$

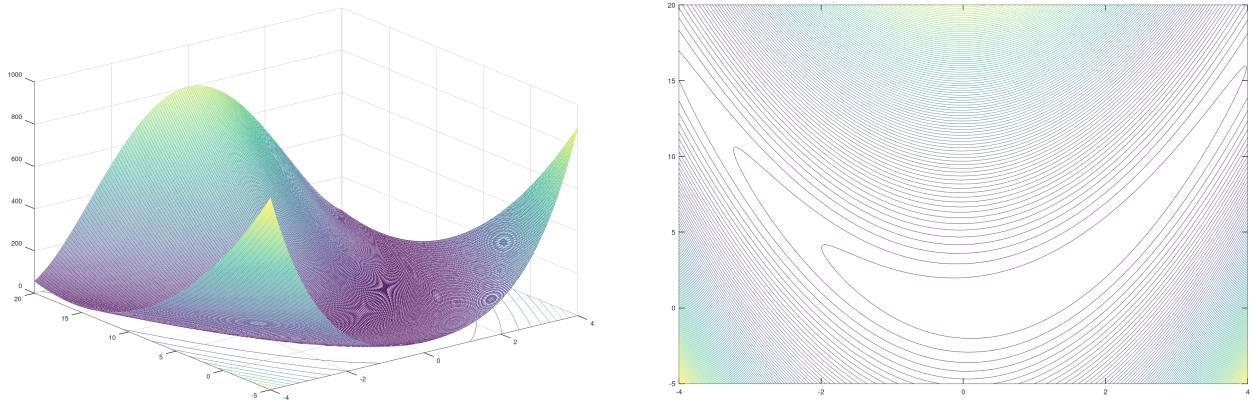
Le Jacobien s'exprime comme  $J = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} \end{bmatrix}$

Après calcul on obtient alors :

$$J = \begin{bmatrix} -2x_1\sqrt{2b} & \sqrt{2b} \\ -\sqrt{2} & 0 \end{bmatrix}$$

## 1.2 Visualisation de la fonction objectif

Après avoir complété les script `rosenbrock` et `visualisation` disponibles en Annexe, on trouve alors pour  $x_1 \in [-4, 4]$  et  $x_2 \in [-5, 20]$  :



(a) Visualisation de la fonction de Rosenbrock en 3D (b) Visualisation des lignes de niveaux de la fonction

### 1.3 Minimisation par descente itérative

- Pour commencer, on utilise la méthode de descente de plus forte pente avec un pas fixe. On obtient alors en 1000 itérations les différents figures :

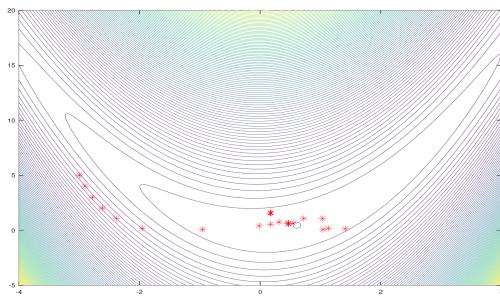


FIGURE 2 – Descente de pas 1

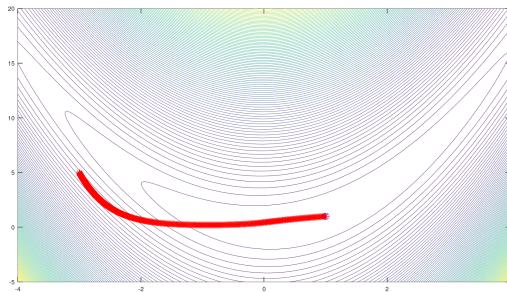


FIGURE 3 – Descente de pas  $10^{-2}$

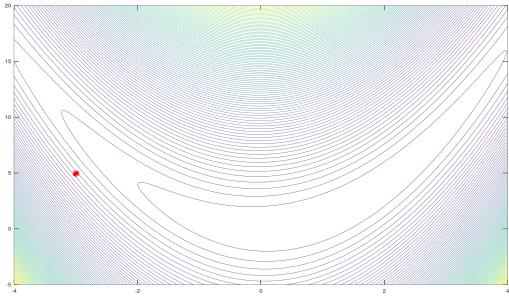


FIGURE 4 – Descente de pas  $10^{-4}$

On remarque que :

- Avec le pas de  $10^{-4}$ , les 1000 itérations ne sont pas suffisante pour tendre vers  $x_h$
- Pour les deux autres valeurs du pas, la méthode converge vers  $x_h$  mais avec le pas de 1, la méthode converge bien plus rapidement qu'avec le pas de  $10^{-2}$ .

2. On ajoute alors une recherche de pas par une technique de rebroussement avec un taux  $\beta = 0.75$  pour assurer la condition d'Armijo avec  $c = 10^{-4}$ .

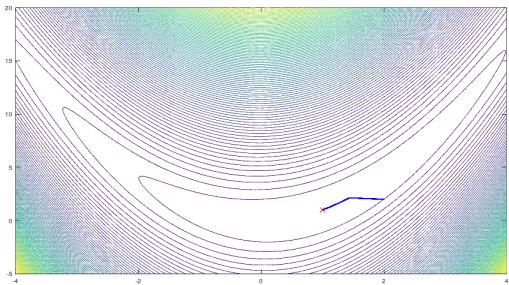


FIGURE 5 – technique de rebroussement

On observe alors que le pas varie effectivement au cours des itérations, de sorte à assurer la condition d'Armijo, ce qui garantit une décroissance suffisante et permet donc de ne pas avoir à se soucier des problèmes vus plus haut (pas trop faible ou trop élevé). Cependant, cette méthode est plus stable mais nécessite plus de calculs que pour la méthode du pas fixe, elle est par exemple bien plus lente que la méthode de pas fixe avec un pas de 1 pour cet exemple précis.

4. Pour la valeur :  $x_0 = (4, 4)$ , on obtient les tracés suivants :

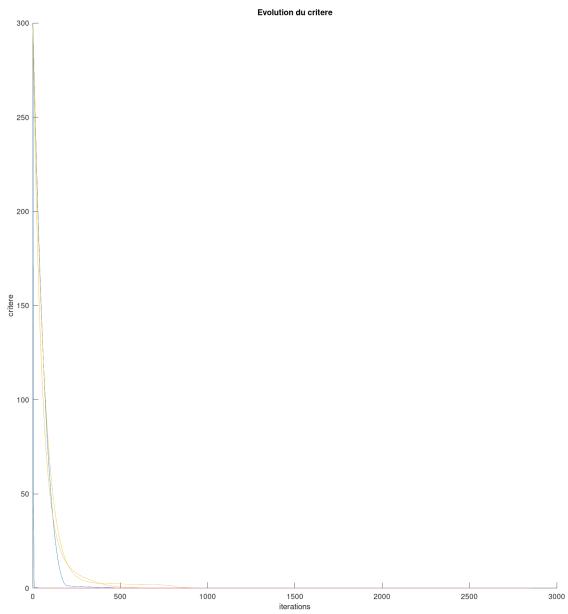


FIGURE 6 – Évolution du critère

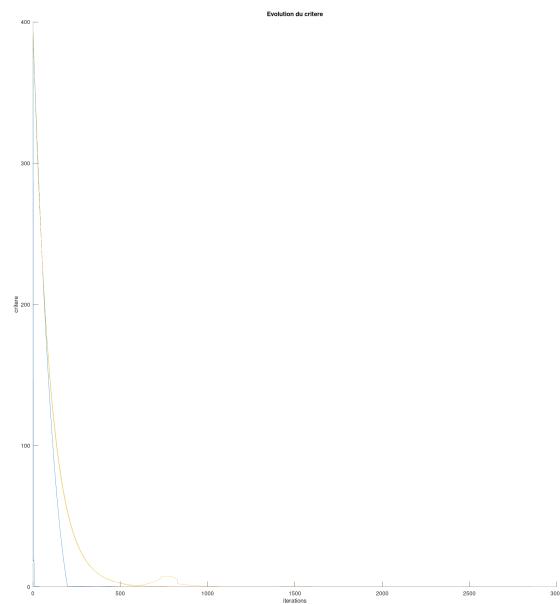


FIGURE 7 – Évolution de la norme du gradient

Pour la valeur :  $x_0 = (5, 10)$  et avec un pas variable, on obtient les tracés suivants :

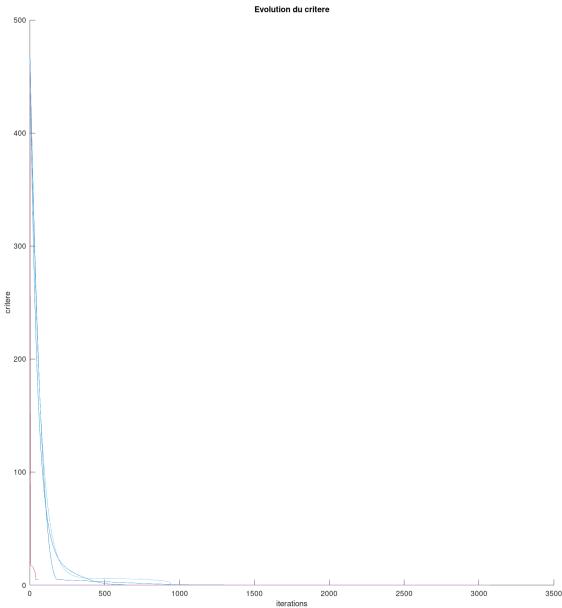


FIGURE 8 – Évolution du critère

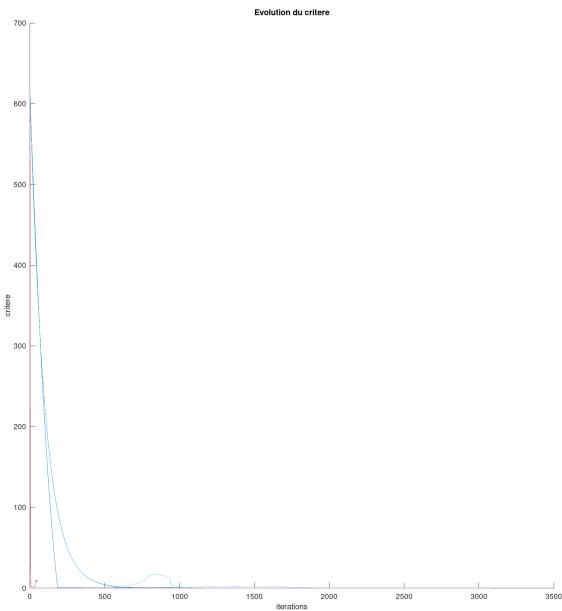


FIGURE 9 – Évolution de la norme du gradient

Nous avons cependant plusieurs problèmes d'affichage de ces graphiques sous Octave (nous n'arrivons par exemple pas à afficher la légende ou à interagir avec les tracés). On observe tout de même que pour ces 4 tracés le critère et le gradient s'annulent rapidement pour chaque méthode, il faudrait choisir des valeurs de tolérance plus adaptées pour gagner en temps d'exécution, quitte à être légèrement moins précis. Pour comparer les performances de chaque méthode représentons le nombre d'itérations et le temps d'exécution dans un tableau :

Pour  $x_0 = (4, 4)$  :

Méthode	Nombre d'itérations	Temps (s)
Descente de plus forte pente	911	0.51
Gradient conjugué	451	0.18
Newton	1989	0.37
BFGS	2862	0.75

Pour  $x_0 = (5, 10)$  :

Méthode	Nombre d'itérations	Temps (s)
Descente de plus forte pente	1526	0.70
Gradient conjugué	x	x
Newton	2043	0.64
BFGS	3000	1.30

On remarque alors que :

- Pour la première valeur de  $x_0$ , la méthode du gradient conjuguée est la plus performante que ce soit en temps de calcul ou en nombre d'itérations
- Mais la méthode du gradient conjuguée ne converge pas vers  $x_h$  pour une position initiale plus éloignée.
- Les méthodes de Newton et BFGS sont celles nécessitant le plus d'itérations
- La méthode de Newton est cependant plus rapide que celle du gradient de plus forte pente

6. Pour  $x_0 = (4, 4)$  :

Méthode	Nombre d'itérations	Temps (s)
Gauss-Newton	1603	0.31
levenberg-marquardt	1594	0.33

Pour  $x_0 = (5, 10)$  :

Méthode	Nombre d'itérations	Temps (s)
Gauss-Newton	1622	0.32
levenberg-marquardt	1607	0.35

On remarque alors que :

- Les deux méthodes convergent bien vers  $x_h$  pour les deux valeurs de  $x_0$ .
- Les performances de ces deux méthodes sont très similaires pour les deux valeurs de  $x_0$ , ce qui n'est pas le cas des autres méthodes et qui représente un avantage important.
- Ce sont les méthodes les plus rapides pour la deuxième valeur de  $x_0$  et la deuxième plus rapide pour la première valeur de  $x_0$ .
- La méthode de levenberg-marquardt nécessite le choix d'une valeur pour un paramètre. Dans ce cas ci, nous obtenons des résultats satisfaisant uniquement pour des valeurs très faibles de  $\lambda$ , la méthode est alors quasiment équivalente à celle de Gauss-Newton.

## 2 Ajustement d'une courbe non-linéaire

On cherche à minimiser la fonction objectif :

$$f_0(\alpha, \beta) = \frac{1}{2} \sum_{i=1}^N (R_i - \beta E_i^{-\alpha})^2$$

1. Après calcul, on trouve :

$$\nabla f(x) = \begin{bmatrix} \sum_{i=1}^N \beta \ln(E_i) E_i^{-\alpha} (R_i - \beta E_i^{-\alpha}) \\ - \sum_{i=1}^N E_i^{-\alpha} (R_i - \beta E_i^{-\alpha}) \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} \sum_{i=1}^N \beta E_i^{-\alpha} \ln(E_i)^2 (2\beta E_i^{-\alpha} - R_i) & \sum_{i=1}^N E_i^{-\alpha} \ln(E_i) (R_i - 2\beta E_i^{-\alpha}) \\ \sum_{i=1}^N E_i^{-\alpha} \ln(E_i) (R_i - 2\beta E_i^{-\alpha}) & \sum_{i=1}^N E_i^{-2\alpha} \end{bmatrix}$$

2. Pour les courbes de niveau, on la trace en échelle log pour  $0 < \alpha < 2$  et  $0 < \beta < 4$

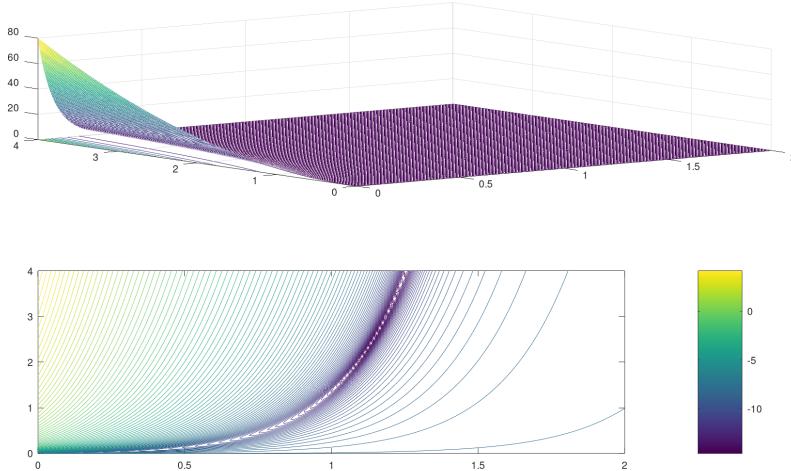


FIGURE 10 – Courbe de niveau du log-critère

3. Par la suite on applique les algorithmes vu à l'exercice un pour trouver les paramètres optimaux. On retrouve ainsi les résultats suivants :

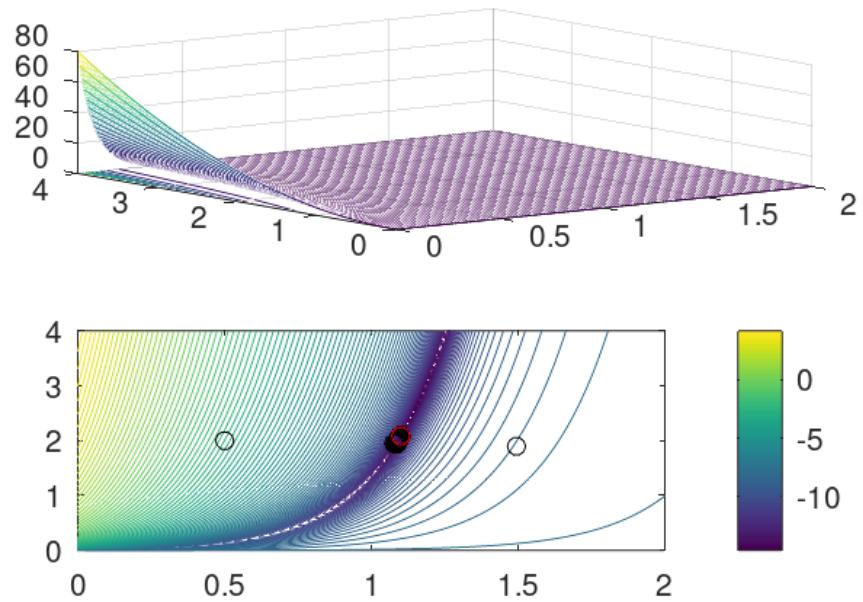


FIGURE 11 – Convergence pour une direction de plus forte pente avec pas variable

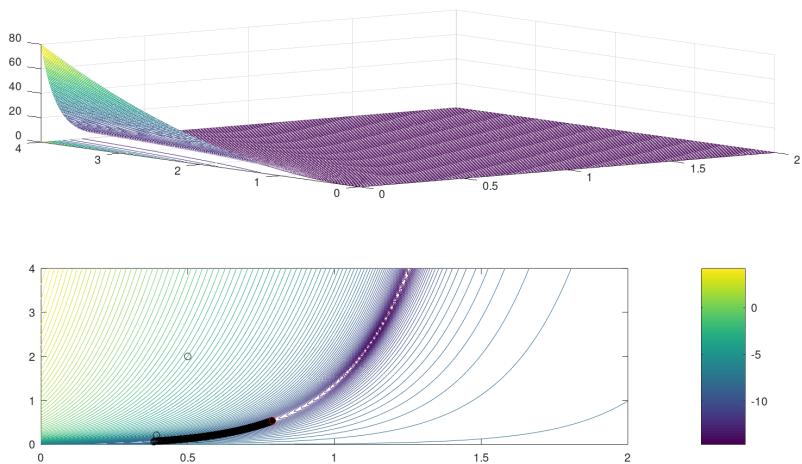


FIGURE 12 – Convergence pour une direction de Newton avec pas variable

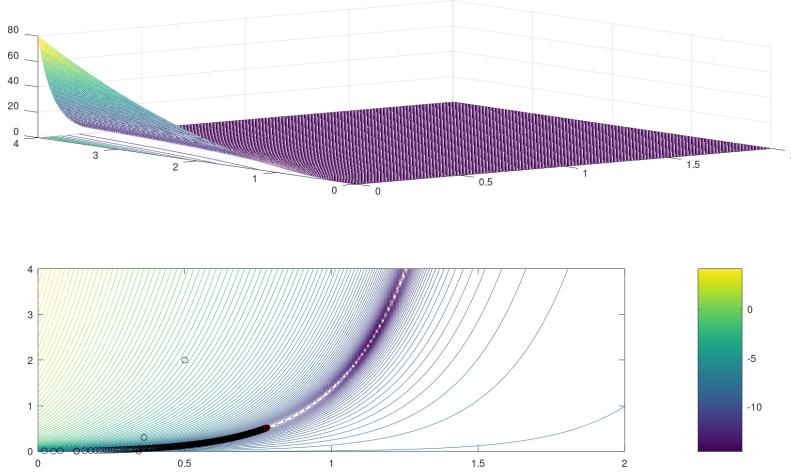


FIGURE 13 – Convergence pour une direction de Newton avec pas fixe

4. On peut reprendre les valeurs finales dans un tableau en prenant  $x_0 = [0.5; 2]$  :

Méthode	Nombre d'itérations	Temps (s)	Valeur finale
Descente de plus forte pente	1002	2	(1.10, 2.08)
Gradient conjugué	1002	10.7	(1.11, 2.15)
Newton	1002	5.9	(0.79, 0.54)
BFGS	39	0.38	(1.15, 2.55)

5. En prenant les différents  $x_h$ , on obtient la caractéristique théorique suivante en rouge par rapport à celle de mesure en bleue :

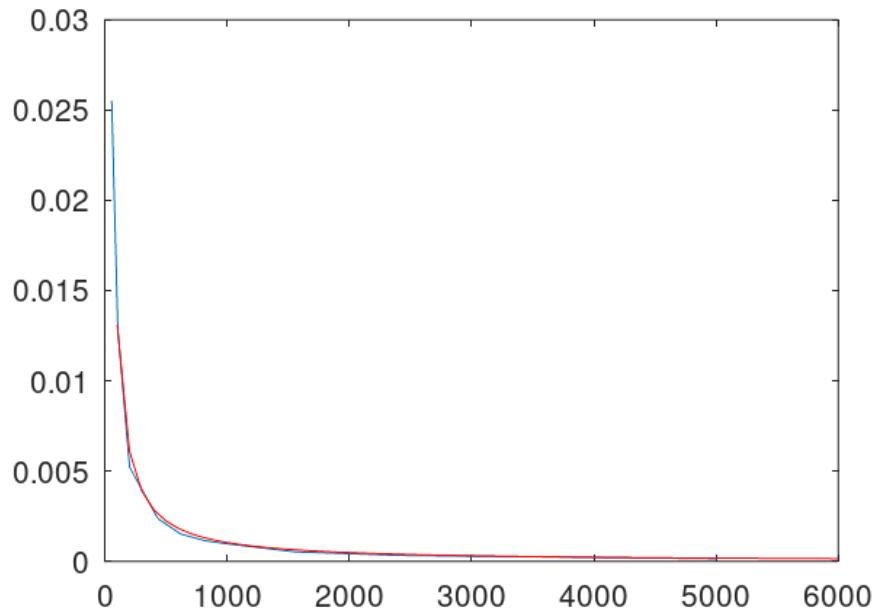


FIGURE 14 – Caractéristiques théorique en et expérimentale pour  $xh = [1.10, 2.08]$

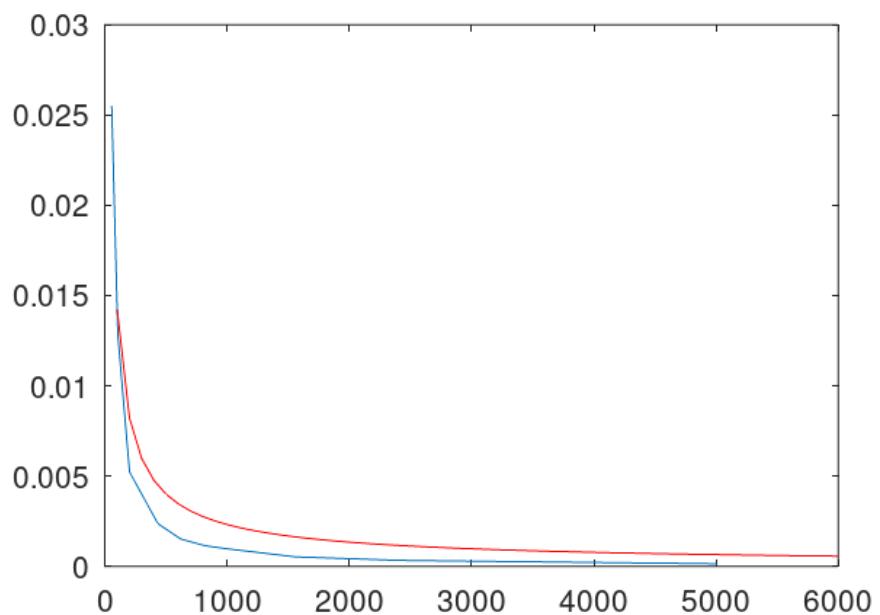


FIGURE 15 – Caractéristiques théorique en et expérimentale pour  $xh = [0.79, 0.54]$

### 3 Débruitage d'un signal par minimisation d'un critère composite

## 4 Inversion numérique d'une transformée de Laplace par optimisation sous contraintes

## A Minimisation de la fonction de Rosenbrock