

Déconvolution d'images régularisée sous contrainte de somme et de non-négativité - Illustration sur des images interférométriques

Céline THEYS¹, Henri LANTÉRI¹, Claude AIME¹

¹Laboratoire Lagrange, UMR 7293 CNRS, Université Côte d'Azur, Observatoire de la Côte d'Azur
celine.theys@unice.fr, hlanteri@unice.fr, claude.aime@unice.fr

Résumé – Le travail présenté concerne la déconvolution d'images par minimisation de divergences invariants par changement d'échelle. L'intérêt de telles divergences est d'inclure par elles-mêmes une contrainte de somme invariante sur la reconstruction. En utilisant ce type de divergence à la fois pour le terme d'attache aux données et pour le terme de régularisation, on obtient un algorithme de gradient de type point intérieur qui maintient la somme et la non-négativité. Les applications de cet algorithme très général sont nombreuses et une illustration est faite sur des images simulées en interférométrie optique.

Abstract – The paper focuses on the deconvolution of images by minimization of divergences invariant to a change of scale. The interest of such divergences for the problem of deconvolution is to intrinsically obtain a sum constraint on the reconstruction. Using this type of divergence for both the data fidelity term and the regularization term, we obtain an interior point algorithm that maintains the sum and the non-negativity. The applications of this very general algorithm are numerous and an illustration on simulated images in optical interferometry is given.

1 Introduction

On traite ici le problème de reconstruction (déconvolution) d'images sous contraintes de non-négativité et de somme des valeurs de l'objet reconstruit. On dispose de données bruitées $y_i \equiv p_i$ et d'un modèle linéaire $(H\mathbf{x})_i \equiv q_i$ traduisant l'opération de convolution où H est une matrice connue, telle que $h_{ij} \geq 0$ et $\sum_i h_{ij} = 1, \forall j$. Le problème est d'estimer \mathbf{x} sous contraintes de non négativité et de somme fixée des composantes de \mathbf{x} , le mauvais conditionnement de H lié aux conditions d'observation en fait un problème mal posé.

Dans un précédent travail, [3], on a montré que les divergences invariants par rapport à q , notées $DI(p||q)$, telles que $DI(p||\alpha q) = DI(p||q)$, possèdent une propriété spécifique qui permet de développer des algorithmes de déconvolution d'images dans lesquels la contrainte de somme constante des composantes de la solution pouvait être prise en compte de façon simple et efficace en même temps que la contrainte de non-négativité. Dans [4], on a montré comment des divergences invariants pouvaient être obtenues à partir des α et des β -divergences, en particulier les formes logarithmiques de ces divergences qui ont la particularité d'être invariants par rapport aux deux arguments p et q , c.a.d $DI(\beta p||\alpha q) = DI(p||q)$. On retrouve ainsi, entre autres, la divergence de Fujisawa-Eguchi utilisée dans [3]. Dans les algorithmes itératifs développés dans ces travaux, les itérations sont stoppées avant amplification du bruit.

Dans le présent travail, on se propose d'introduire une régularisation explicite par contrainte de douceur au sens de Tikhonov en utilisant le laplacien. Afin de satisfaire simplement

la contrainte de somme, les divergences utilisées doivent être invariants vis à vis de leurs deux arguments. Par opposition aux méthodes classiques de minimisation sous contrainte, les algorithmes proposés permettent d'éviter l'utilisation de nombreux multiplicateurs de Lagrange et leur actualisation au cours du processus itératif. On peut ainsi obtenir des algorithmes simples à mettre en oeuvre.

Les expressions de ces divergences ainsi que leurs propriétés fondamentales sont tout d'abord brièvement exposées, puis leur utilisation dans le cadre de la déconvolution régularisée est décrite et l'algorithme explicité. Enfin, ces algorithmes sont utilisés sur l'exemple de la déconvolution d'une planète extrasolaire semblable à la Terre, telle qu'elle pourrait être observée avec des interféromètres optiques mis en réseau dans l'espace.

2 Divergences invariants

Les α -divergences, [6] et les β -divergences, [8, 9], sont respectivement des divergences de Csiszär, [5] et de Bregman, [7], construites sur la fonction de base strictement convexe :

$$f(x) = \frac{1}{\gamma(\gamma-1)} [x^\gamma - \gamma x - (1-\gamma)]. \quad (1)$$

La procédure décrite dans [4] conduit à des formes de ces divergences invariants par rapport à l'argument q puis sous forme logarithmique, invariants par rapport aux deux arguments : p

et q . Dans le cas des α divergences, on obtient :

$$LAI_\gamma(p||q) = \frac{1}{\gamma} \log \sum_i q_i + \frac{1}{\gamma(\gamma-1)} \log \sum_i p_i^\gamma q_i^{1-\gamma} - \frac{1}{(\gamma-1)} \log \sum_i p_i \quad (2)$$

qui a été proposée par Cichocki et al., [10, 11]. Dans le cas des β divergences, on obtient :

$$LBI_\gamma(p||q) = \frac{1}{\gamma(\gamma-1)} \log \sum_i p_i^\gamma + \frac{1}{\gamma} \log \sum_i q_i^\gamma - \frac{1}{\gamma-1} \log \sum_i p_i q_i^{\gamma-1}. \quad (3)$$

C'est la divergence proposée par Fujisawa, [12], que nous avons utilisée dans [2]. Le cas $\gamma = 2$ correspond à l'écart quadratique moyen, rendu invariant par rapport à q , sous sa forme logarithmique.

3 Critère régularisé

3.1 Problème général

Le problème considéré est le problème classique qui consiste à exprimer la fonction objectif, ici la divergence $\mathcal{D}(\cdot)$ sous la forme d'une somme pondérée de deux termes :

$$\mathcal{D}(\mathbf{p}||\mathbf{q}) = \mathcal{D}_1(\mathbf{y}||H\mathbf{x}) + \delta \mathcal{D}_2(\mathbf{x}||\mathbf{x}_p) = J(\mathbf{x}). \quad (4)$$

Dans cette expression, $\mathcal{D}_1(\mathbf{y}||H\mathbf{x})$ est la divergence entre les données bruitées \mathbf{y} et le modèle linéaire $H\mathbf{x}$, alors que $\mathcal{D}_2(\mathbf{x}||\mathbf{x}_p)$ est un terme de régularisation qui traduit l'écart entre la solution \mathbf{x} et une solution par défaut \mathbf{x}_p ; δ est le facteur de régularisation qui pondère l'importance relative entre le terme de fidélité aux mesures et le terme de régularisation.

Le problème consiste à minimiser par rapport à \mathbf{x} , $J(\mathbf{x})$ sous les contraintes $x_i \geq 0 \forall i$ et $\sum_i x_i = K$.

3.2 Terme d'attache aux données

En ce qui concerne le terme $\mathcal{D}_1(\mathbf{y}||H\mathbf{x})$ ($\equiv \mathcal{D}_1(\mathbf{p}||\mathbf{q})$), nous présentons les résultats obtenus en utilisant pour ce terme la divergence $\mathcal{D}_1 = LBI_\gamma$. Son gradient par rapport à \mathbf{x} est dans ce cas :

$$-\left[\nabla_{\mathbf{x}} \mathcal{D}_1(\mathbf{y}||H\mathbf{x})\right]_i = \left[H^T \frac{\mathbf{y} \circ (H\mathbf{x})^{\gamma-2}}{\sum_j y_j (H\mathbf{x})_j^{\gamma-1}} \right]_i - \left[H^T \frac{(H\mathbf{x})^{\gamma-1}}{\sum_j (H\mathbf{x})_j^\gamma} \right]_i. \quad (5)$$

Dans ces expressions, le symbole " \circ " est utilisé pour le produit de Hadamard (pointwise product).

Dans la mesure où les inconnues \mathbf{x} n'interviennent que dans l'expression de \mathbf{q} , toutes les divergences invariantes, notées DI

données dans [4], incluant LBI_γ , qu'elles soient sous forme logarithmique ou non, peuvent être utilisées. En effet, elles possèdent la propriété spécifique :

$$\sum_l x_l \frac{\partial DI(\mathbf{y}||H\mathbf{x})}{\partial x_l} = 0. \quad (6)$$

Cette propriété est nécessaire et suffisante pour respecter la contrainte de somme dans l'algorithme final, $\mathcal{D}_1(\mathbf{y}||H\mathbf{x})$ satisfait cette propriété.

3.3 Terme de régularisation.

Pour le terme de régularisation, $\mathcal{D}_2(\mathbf{x}||\mathbf{x}_p)$, nous nous focalisons sur une régularisation classique par contrainte de douceur de type Tikhonov et plus spécifiquement au cas où on utilise le laplacien de la solution. Dans ce cas, on exprime le laplacien Lx sous la forme $(I - A)\mathbf{x} = \mathbf{x} - A\mathbf{x}$. Dans cette expression, $A\mathbf{x} \equiv \mathbf{x}_p$ est une forme lissée de la solution et on a $\sum_i (A\mathbf{x})_i = \sum_i x_i$. Notons qu'on peut utiliser pour A , n'importe quelle forme d'opérateur passe bas. L'expression classique du terme de pénalité $\|Lx\|^2$ exprime alors la divergence entre \mathbf{x} et $A\mathbf{x}$ sous forme d'une norme euclidienne. Dans cette étude, nous proposons d'exprimer l'écart entre ces deux quantités en utilisant la divergence $\mathcal{D}_2 = LBI_\gamma(\mathbf{x}||A\mathbf{x}) \equiv LBI_\gamma(x||x_p)$. On a alors pour le gradient par rapport à \mathbf{x} :

$$\left[\nabla_{\mathbf{x}} \mathcal{D}_2(\mathbf{x}||A\mathbf{x})\right]_i = \frac{1}{(\gamma-1)} \frac{x_i^{\gamma-1}}{\sum_j x_j^\gamma} + \frac{[A^T(A\mathbf{x})^{\gamma-1}]_i}{\sum_j (A\mathbf{x})_j^\gamma} - \frac{[A^T(\mathbf{x} \circ (A\mathbf{x})^{\gamma-2})]_i}{\sum_j x_j (A\mathbf{x})_j^{\gamma-1}} - \frac{1}{(\gamma-1)} \frac{[(A\mathbf{x})^{\gamma-1}]_i}{\sum_j x_j (A\mathbf{x})_j^{\gamma-1}}. \quad (7)$$

La divergence $\mathcal{D}_2(\mathbf{x}||A\mathbf{x})$ qui est invariante par rapport aux deux arguments, possède une propriété analogue à (6) :

$$\sum_l x_l \frac{\partial LBI_\gamma(x||Ax)}{\partial x_l} = 0. \quad (8)$$

On peut noter que la divergence $LAI_\gamma(x||Ax)$ possède la même propriété et peut, par conséquent, aussi être utilisée pour \mathcal{D}_2 . Remarquons également que la propriété (8) est spécifiquement liée aux divergences logarithmiques, qui sont invariantes par rapport aux deux arguments.

De la même façon que pour le terme d'attache aux données, cette propriété permet de satisfaire la contrainte de somme dans l'algorithme final.

Si l'on considère le cas d'une régularisation au sens de Tikhonov avec une solution a priori connue \mathbf{x}_c , c.a.d ne dépendant pas de \mathbf{x} , toutes les divergences invariantes proposées dans [4], peuvent être utilisées. Si $\mathcal{D}_2 = LBI_\gamma(\mathbf{x}||\mathbf{x}_c)$, le gradient par rapport à \mathbf{x} est égal à :

$$-\left[\nabla_{\mathbf{x}} \mathcal{D}_2(\mathbf{x}||\mathbf{x}_c)\right]_i = \frac{1}{\gamma-1} \left(\frac{(x_c)_i^{\gamma-1}}{\sum_j x_j (x_c)_j^{\gamma-1}} - \frac{x_i^{\gamma-1}}{\sum_j x_j^\gamma} \right). \quad (9)$$

Dans le cas d'une solution a priori constante, $(x_c)_i = K, \forall i$, l'expression (9) se simplifie. Dans nos simulations, nous considérons la régularisation laplacienne, (7).

4 Algorithmique

Les algorithmes de reconstruction sous contrainte de non-négativité et de somme constante sont développés suivant la méthode décrite dans [1, 3]. Ce sont des algorithmes de points intérieurs, fondés sur les conditions de Karush Kuhn Tucker et qui s'écrivent dans le cas de la minimisation du critère (4) :

$$x_i^{k+1} = x_i^k + \alpha_i^k x_i^k [-\nabla_{\mathbf{x}} \mathcal{D}_1(\mathbf{y} \| H\mathbf{x}) - \delta \nabla_{\mathbf{x}} \mathcal{D}_2(\mathbf{x} \| A\mathbf{x})]_i. \quad (10)$$

L'intervalle de recherche du pas de descente assurant la non-négativité de x_i^{k+1} , pour $x_i^k \geq 0$, $\forall i$ est tout d'abord calculé. Le pas de descente assurant la convergence est ensuite calculé dans cet intervalle, par une méthode de recherche de pas unidimensionnelle simplifiée de type Armijo.

La contrainte de somme est assurée quel que soit le pas grâce aux propriétés (6) et (8). On peut facilement vérifier que nous avons $\sum_i x_i^{k+1} = \sum_i x_i^k$ dans l'algorithme (10), cela signifie que la contrainte de somme est satisfaite pour tous les estimés successifs et il suffit de fixer la somme des valeurs initiales à la constante voulue. Il s'agit de la propriété principale connectée à des divergences invariantes par changement d'échelle et par conséquent, la procédure proposée est une méthode de point intérieur, c'est-à-dire que toutes les estimations successives remplissent les contraintes, propriété particulièrement intéressante dans le contexte des problèmes inverses.

5 Illustration

5.1 Interférométrie optique

L'application que nous considérons ici est l'imagerie par interférométrie spatiale d'exoplanètes. Des petits télescopes mis en réseau cohérent peuvent donner la même résolution angulaire qu'un très grand télescope. Ces réseaux sont couramment utilisés en radioastronomie. Pour l'observation des exoplanètes, le domaine optique garde tout son intérêt (voir exoplanet.eu), et une flottille de petits télescopes dans l'espace, distants entre eux de quelques centaines de mètres, pourrait donner dans le futur les premières images d'une exo-Terre. L'utilisation de ces réseaux de télescopes fournira des images qu'il faudra améliorer par le traitement d'image.

C'est le problème que nous considérons ici pour un réseau de 16 petits télescopes opérant dans la configuration non redondante sur réseau hexagonal proposée par Kopilovitch et Sodin (K&S), [13], et pour laquelle nous donnons la couverture spectrale et la réponse impulsionnelle (PSF) dans la figure 1.

5.2 Simulations

L'objet choisi pour l'exoplanète \mathbf{x} est une image basse résolution de la terre, donnée dans la figure 3(e). En raison de la structure de la PSF, l'image $H\mathbf{x}$ présente des répliques de \mathbf{x} sur le maillage hexagonal, figure 3(a).

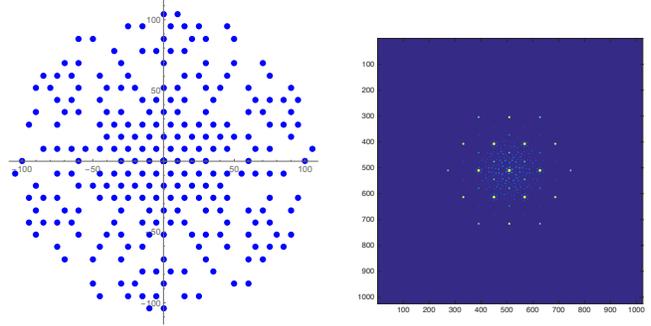


FIGURE 1 – Couverture du plan des fréquences spatiales et réponse impulsionnelle (H) pour K&S.

Un capteur réalise une photodétection, que l'on modélise par une transformation de Poisson de moyenne $H\mathbf{x}$, et auquel on ajoute un bruit Gaussien pouvant venir d'un bruit de lecture, de la fluctuation d'un fond continu, ou de toute autre source donnant un bruit additif indépendant du processus de Poisson. Les données bruitées \mathbf{y} s'écrivent alors suivant le modèle :

$$\mathbf{y} = \mathcal{P}(H\mathbf{x}) + \mathbf{n} \quad (11)$$

où \mathcal{P} est la transformation de Poisson et $\mathbf{n} \sim \mathcal{N}(0, \sigma^2)$. L'image résultante est représentée dans la figure 3(b) pour un total de 3×10^6 photons dans l'image et $\sigma^2 = 3$, le zoom sur la partie centrale de l'image est donné sur la figure 3(c). Toutes les images sont de taille 1024×1024 points.

L'algorithme (10) a été implémenté sur les données précédentes pour $\gamma = 2$, valeur particulière, comme indiqué précédemment. Une première étude dans [4] montre un minimum de l'erreur quadratique pour cette valeur de γ .

Au cours des itérations, l'effet de l'algorithme est de réduire progressivement les répliques de l'objet jusqu'à les supprimer dans la meilleure reconstruction au sens de l'erreur quadratique moyenne $\epsilon^k = \|\mathbf{x}^k - \mathbf{x}^*\| / \|\mathbf{x}^*\|$.

Les meilleures reconstructions de l'objet sont données dans la figure 3, colonne de droite, rangées de haut en bas, dans l'ordre croissant de la valeur de l'erreur de reconstruction. On a alors la meilleure reconstruction pour, une régularisation laplacienne, fig. 3(f), pas de régularisation ($\delta = 0$) et arrêt au minimum de l'erreur, fig. 3(g) et régularisation par une constante, fig. 3(h). Pour les algorithmes régularisés, la valeur du paramètre δ a été choisie comme celle stabilisant l'erreur minimale de reconstruction, le réseau de courbes de l'erreur ϵ^k en fonction de δ est donnée sur l'exemple de la régularisation laplacienne dans la figure 2. Les objets reconstruits sont donnés pour $k = 1000$. La meilleure reconstruction est obtenue dans l'exemple particulier de ces données avec la régularisation laplacienne. On peut vérifier l'intérêt de la régularisation en comparant les résultats régularisés avec le résultat non régularisé au même nombre d'itérations, $k = 1000$, figure 3(d).

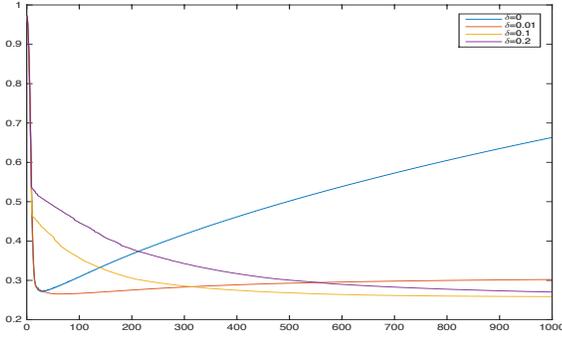


FIGURE 2 – Erreur de reconstruction au cours des itérations, e^k en fonction de δ pour une régularisation laplacienne, $\delta = 0.1$ valeur optimale.

6 Conclusion

Nous avons proposé l'utilisation de divergences invariantes par changement d'échelle pour construire un algorithme régularisé de déconvolution. L'algorithme proposé est un algorithme très général de type point intérieur, qui satisfait une contrainte de non-négativité et de somme invariante sur l'objet reconstruit. Une application sur des images simulées en interférométrie optique permet de valider l'algorithme.

Références

- [1] H. Lantéri, M. Roche, and C. Aime, "Penalized maximum likelihood image restoration with positivity constraints- multiplicative algorithms," *Inverse problems*, vol. 18, pp. 1397–1419, 2002.
- [2] H. Lantéri and C. Theys, *La gamma divergence de Fujisawa-Eguchi, une alternative pour imposer la contrainte de somme. Application au démélange linéaire*, in *GRETSI*, 2013.
- [3] H. Lantéri, C. Theys and C. Aime, *Divergences invariantes par changement d'échelle pour la reconstruction de signaux ou d'images Application à l'interférométrie optique*, in *GRETSI*, 2015.
- [4] H. Lantéri, C. Theys and C. Aime, *Scale invariant divergences for signal and image reconstruction*, in *EUSIPCO*, 2015.
- [5] I. Csiszár, "Information measures : a critical survey," in *Transactions of the 7th Prague conference*, 1974.
- [6] S. I. Amari and H. Nagaoka, *Methods of information geometry*, Oxford University Press, New york, 2000.
- [7] L. M. Bregman, "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 200–217, 1967.
- [8] A. Basu, I. R. Harris, N. L. Hjort, and M. C. Jones, "Robust and efficient estimation by minimizing a density power divergence," *Biometrika*, vol. 85, no. 3, pp. 549–559, 1998.
- [9] S. Eguchi and Y. Kano, "Robustifying maximum likelihood estimation," Tech. Rep., Institute of statistical mathematics, 2001.
- [10] A. Cichocki and S. Amari, "Families of alpha-beta and gamma divergences : flexible and robust measures of similarities," *Entropy*, vol. 12, 2010.
- [11] A. H. Phan A. Cichocki, R. Zdunek and S.I. Amari, *Non negative matrix and tensor factorization*, J. Wiley, 2010.
- [12] H. Fujisawa and S. Eguchi, "Robust parameter estimation with small bias against heavy contamination," *Multivariate Analysis*, vol. 99, no. 9, 2008.
- [13] L. E. Kopilovich and L. G. Sodin, Eds., *Multielement system design in astronomy and radio science*, vol. 268 of *Astrophysics and Space Science Library*, 2001.

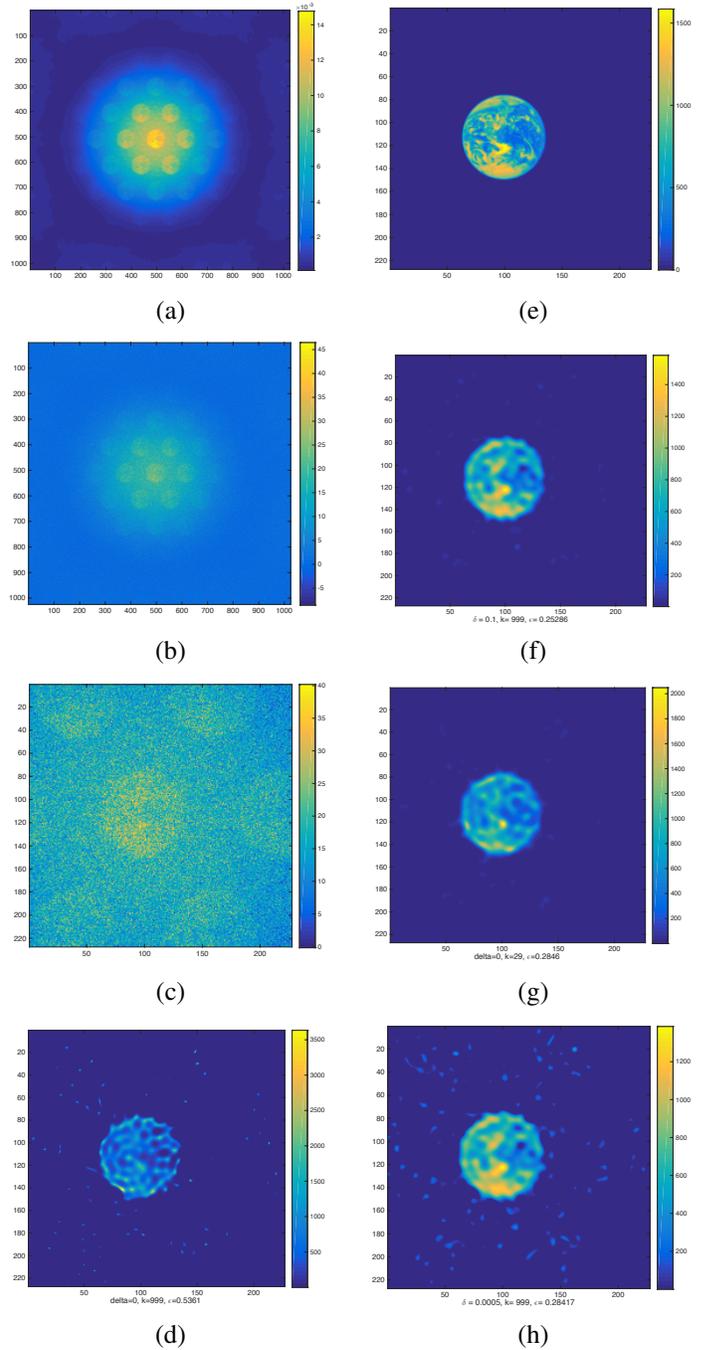


FIGURE 3 – (a) Image convoluée Hx , (b) Image convoluée bruitée y , (c) Zoom au centre de (b), (d) Objet reconstruit à $k = 999$, $\delta = 0$, (e) Objet x , (f) Objet reconstruit, régularisation laplacienne, $\delta = 0.1$, (g) Early stopping à $k = 29$, $\delta = 0$, (h) Objet reconstruit, régularisation par une constante, $\delta = 5e - 4$, $(x_c)_i = 1, \forall i$.