Méthodes proximales multi-niveaux pour la restauration d'images

Guillaume Lauga, Elisa Riccietti, Nelly Pustelnik, Paulo Gonçalves

▶ **To cite this version:**

Guillaume Lauga, Elisa Riccietti, Nelly Pustelnik, Paulo Gonçalves. Méthodes proximales multi-niveaux pour la restauration d'images. GRETSI'22 - 28ème Colloque Francophone de Traitement du Signal et des Images, Sep 2022, Nancy, France. hal-03739496

HAL Id: hal-03739496
https://inria.hal.science/hal-03739496

Submitted on 28 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Méthodes proximales multi-niveaux pour la restauration d'images

Guillaume LAUGA[1], Elisa RICCIETTI[1], Nelly PUSTELNIK[2,3], Paulo GONÇALVES[1,4] *

[1]Univ Lyon, Ens de Lyon, Univ Lyon 1, CNRS, INRIA, LIP, Lyon, France.

[2]Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France.

[3]ISPGroup/ICTEAM, UCLouvain, Belgium.

[4] CPE Lyon,Villeurbanne, France.

guillaume.lauga@ens-lyon.fr, elisa.riccietti@ens-lyon.fr
nelly.pustelnik@ens-lyon.fr, paulo.goncalves@ens-lyon.fr

Résumé – Cet article présente une nouvelle méthode pour mettre en œuvre un algorithme forward-backward multi-niveau. En utilisant l'enveloppe de Moreau pour construire la correction apportée par les modèles grossiers, facilement calculable lorsque l'on connaît sous forme explicite l'opérateur proximal des fonctions considérées, nous reformulons les algorithmes proximaux multi-niveaux précédemment introduits dans la littérature sous une forme plus simple, tout en en conservant les performances. Nous montrons la convergence des itérées vers un minimum dans le cas convexe, résultat fondamental pour des problèmes mal posés. Nous validons l'approche sur des problèmes de restauration d'image de grande taille.

Abstract – This paper presents a new method for implementing a multi-level Forward-Backward algorithm. By using the Moreau envelope to build the correction brought by the coarse models, easy to compute when the explicit form of the proximal operator of the considered functions is known, we reformulate proximal multilevel algorithms in a simpler form without degrading the performance. Finally we show the convergence of the iterates to a minimum in the convex case, an important result for ill-posed problems. We evaluate this method on large scale image restoration problems.

1 Introduction

La résolution d'un grand nombre de problèmes en traitement du signal et des images fait appel à la minimisation d'une somme d'une fonction d'attache aux données f et d'une fonction de régularisation g, formellement :

$$\min_{x\in\mathcal{H}} F(x) := f(x) + g(x) \tag{1}$$

où $\mathcal{H}$ est un espace de Hilbert réel ($\mathcal{H} = \mathbb{R}^N$ dans toute la suite), $f : \mathcal{H} \to]-\infty, +\infty]$ et $g : \mathcal{H} \to]-\infty, +\infty]$ appartenant à $\Gamma_0(\mathcal{H})$ la classe des fonctions convexes, semi-continues inférieurement et propres. De plus, f est supposée différentiable de gradient L_f-Lipschitz.

Dans le cadre de la restauration d'image, on souhaite retrouver une image de bonne qualité à partir d'une image dégradée par un opérateur linéaire et/ou du bruit i.e. $z = \mathrm{A}\bar{x} + \epsilon$ où A modélise l'opérateur linéaire de dégradation, ϵ le bruit et z est l'image dégradée. Pour la résolution de ce problème mal posé, on considère généralement une résolution par moindres carrés pénalisée où on note g la fonction de pénalisation permettant de choisir les propriétés que l'on souhaite imposer à la solution. Un choix usuel est une pénalisation par norme l_1 appliquée sur des coefficients d'une transformation linéaire $\mathrm{W} \in \mathbb{R}^{K\times N}$ (ondelettes, trames, dictionnaire, ...) permettant de promouvoir la parcimonie des coefficients de la solution. Le problème de minimisation associé fait intervenir un paramètre de régularisation $\lambda > 0$ et se formule :

$$\widehat{x} \in \operatorname*{Argmin}_{x\in\mathbb{R}^N} \frac{1}{2}\|\mathrm{A}x - z\|_2^2 + \lambda\|\mathrm{W}x\|_1. \tag{2}$$

De nombreux algorithmes ont été proposés dans la littérature pour estimer $\widehat{x}$ (cf. [5]) dont la principale limitation vient de l'augmentation importante du temps de calcul avec la dimension. Des techniques de pré-conditionnement peuvent être investiguées mais requièrent généralement des hypothèses fortes sur le choix de la matrice de préconditionnement (e.g. matrice diagonale) conduisant à des gains souvent limités.

Pour la résolution de problèmes en grande dimension reposant sur une fonction objectif lisse, il est possible de tirer parti d'une structure locale du problème d'optimisation (cf. VMLMB [10] ou 3MG [9]) mais on peut également envisager de jouer sur la finesse à laquelle on va résoudre le problème, une stratégie au coeur des schémas algorithmiques multi-niveaux, principalement développés pour la résolution des équations aux dérivées partielles (EDP), dans lesquels f et g sont supposées différentiables [7]. Dans ces méthodes, la fonction objectif est approchée par une suite de fonctions définies sur des domaines de dimension réduite et par conséquent moins coûteuses à optimiser. Le pas de descente est ainsi calculé à faible coût aux niveaux les plus grossiers, puis projeté vers les niveaux fins.

Récemment, cette idée a été exploitée dans [1] pour définir des algorithmes proximaux multi-niveaux applicables au problème (2) dans le cas où g est non différentiable. Dans cet article nous

*Les auteurs souhaitent remercier le GdR ISIS pour le financement du projet MOMIGS et du projet ANR-19-CE48-0009 Multisc'In.

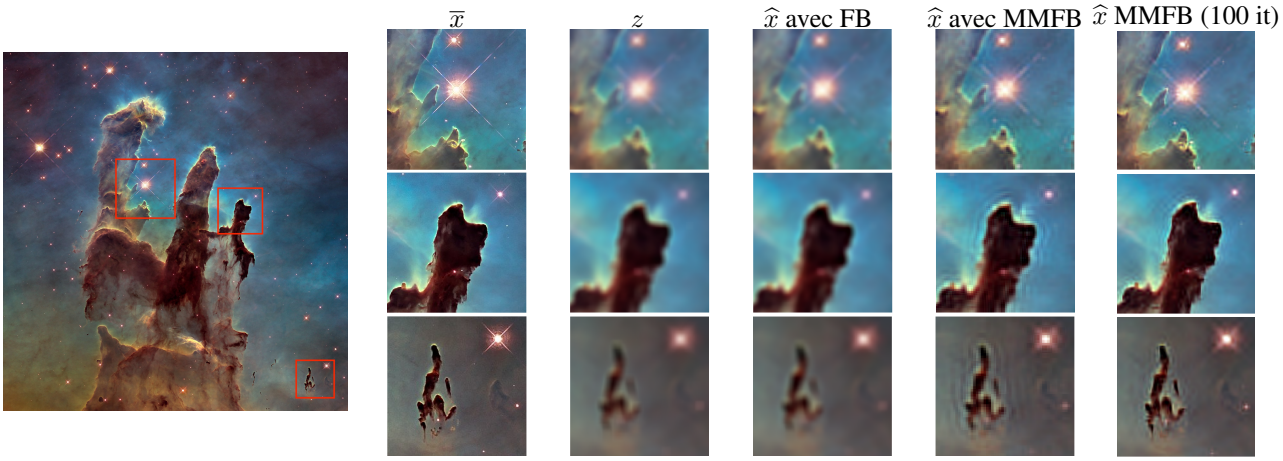

FIGURE 1 – Restauration d'une image dégradée par un flou et un bruit additif gaussien. De gauche à droite : image originale, zoom sur : image originale, image dégradée, restaurée par FB après 2 itérations, restaurée par MMFB après 2 puis 100 itérations. Paramètres utilisés : $\alpha = 1$, $\lambda = 3.5e-5$, $\gamma = 5$, $p = 1$, $m = 10$, taille(PSF)$= 23$, σ(PSF)$= 11$, et $\sigma(\epsilon) = 1e-3$. $\ll$ Pillars of Creation $\gg$. Crédit : NASA, ESA/Hubble.

proposons une variante de ces méthodes, que nous appelons MMFB pour *Moreau Multilevel Forward Backward*, qui exploite l'enveloppe de Moreau pour la définition des approximations grossières de g. Ceci permet de construire facilement une version régulière lorsque l'on connaît sous forme explicite l'opérateur proximal de g. En outre, nous montrons la convergence des itérées de MMFB, résultat important pour les problèmes mal posés et, à notre connaissance, jamais démontrée pour les méthodes proximales multi-niveaux.

En section 2, nous rappelons les grands principes des méthodes multi-niveaux. En section 3, nous décrivons MMFB et les garanties de convergence associées. Enfin en section 4, des résultats expérimentaux permettent de valider le bon comportement de MMFB dans un contexte de restauration d'image.

2 Les méthodes de gradient multi-niveau

Sans perte de généralité, nous considérons le cas à deux niveaux : on indexe par h (resp. H) toutes les quantités définies au niveau fin (resp. grossier). On définit ainsi $F_h = F : \mathbb{R}^N_h \to]-\infty, +\infty]$ la fonction objectif au niveau fin, $F_H : \mathbb{R}^N_H \to]-\infty, +\infty]$ avec $N_H < N_h$ son approximation au niveau grossier, $\mathrm{I}^H_h : \mathbb{R}^{N_h} \to \mathbb{R}^{N_H}$, un opérateur linéaire qui envoie l'information du niveau fin au niveau grossier, et réciproquement $\mathrm{I}^h_H : \mathbb{R}^{N_H} \to \mathbb{R}^{N_h}$ celui qui envoie l'information du niveau grossier au niveau fin. Il est classique de choisir $\mathrm{I}^h_H = \eta(\mathrm{I}^H_h)^T$, avec $\eta > 0$. La méthode de plus forte pente construit la suite d'itérées suivante, pour tout $\tau_{h,k} > 0$,

$$x_{h,k+1} = x_{h,k} + p_{h,k} \qquad \text{avec} \qquad p_{h,k} = -\tau_{h,k}\nabla F_h(x_{h,k}).$$

Dans un schéma multi-niveau, à la place de la mise à jour $p_{h,k}$ précédente, on va projeter $x_{h,k}$ au niveau grossier : $x_{H,k,0} := \mathrm{I}^H_h x_{h,k}$ et définir une suite $(x_{H,k,\ell})_{\ell\in\mathbb{N}}$ (où k représente l'itération courante au niveau fin et ℓ l'indice des itérations au niveau grossier) telle que :

$$x_{H,k,\ell+1} = x_{H,k,\ell} + d_{H,k,\ell} \quad \text{avec} \quad d_{H,k,\ell} = -\alpha_{H,k,\ell} D_{H,k,\ell}$$

permettant de définir, après $m \geq 0$ itérations au niveau grossier, le pas $p_{h,k}$ à effectuer au niveau fin comme la prolongation de la somme de m itérations du gradient au niveau grossier :

$$p_{h,k} = I^h_H(x_{H,m} - x_{H,0}) = I^h_H \sum_{\ell=1}^{m} d_{H,k,\ell}.$$

Pour garantir la cohérence au premier ordre entre les deux niveaux, et assurer que la direction prise au niveau grossier est bien une direction de descente pour F_h, la direction au niveau grossier est modifiée comme suit :

$$D_{H,k,\ell} = \nabla F_H(x_{H,k,\ell}) + v_{H,k},$$

avec :

$$v_{H,k} = I^H_h \nabla F_h(x_{h,k}) - \nabla F_H(x_{H,k,0}).$$

Ce choix pour $v_{H,k}$ garantit que dans un voisinage de $x_{H,k,0}$, $D_{H,k,\ell} \sim I^H_h \nabla F_h(x_{h,k})$ et si $p_{h,k} = -I^h_H D_{H,k,\ell}$ alors $p_{h,k}$ est une direction de descente pour F_h :

$$\begin{aligned}\nabla F_h(x_{h,k})^T p_{h,k} &= -\nabla F_h(x_{h,k})^T I^h_H D_{H,k,\ell} \\ &= -\eta(I^H_h \nabla F_h(x_{h,k})^T D_{H,k,\ell} \\ &\sim -\eta \|I^H_h \nabla F_h(x_{h,k})\|^2 < 0.\end{aligned}$$

On va descendre au niveau grossier seulement si on estime que la direction donnée par le gradient grossier a le potentiel de minimiser F_h de manière importante : sa norme est grande relativement à la direction qu'on aurait sinon pris au niveau fin [8] :

$$\|I^H_h \nabla F_h(x_{h,k})\| > \kappa \|\nabla F_h(x_{h,k})\|. \tag{3}$$

où $\kappa \in]0, \min\{1, \|I^H_h\|\}[$. Si cette condition n'est pas satisfaite on effectue un pas de gradient classique au niveau fin.

3 MMFB : Moreau Multilevel FB

Nous proposons un schéma de forward-backward multi-niveau convergeant reposant sur des itérations de gradient au niveau grossier, construit avec l'enveloppe de Moreau de g_h et des

Algorithm 1: Multilevel Gradient Method (MGM)

Data: $x_{h,0}$, $\epsilon_h > 0$, $\kappa \in]0, \min\{1, \|I_h^H\|\}[$, $k = 0$
while $\|x_{h,k+1} - x_{h,k}\| > \epsilon_h$ **do**
  if $\|I_h^H \nabla F_h(x_{h,k})\| > \kappa \|\nabla F_h(x_{h,k})\|$ **then**
    $x_{H,k,0} = I_h^H x_{h,k}$
    $v_{H,k} = I_h^H \nabla F_h(x_{h,k}) - \nabla F_H(x_{H,k,0})$
    for $\ell = 0, \ldots, m-1$ **do**
      Choisir $\alpha_{H,k,\ell}$
      $d_{H,k,\ell} = -\alpha_{H,k,\ell} (\nabla F_H(x_{H,k,\ell}) + v_{H,k})$
      $x_{H,k,\ell+1} = x_{H,k,\ell} + d_{H,k,\ell}$
    end
    $x_{h,k+1} = x_{h,k} + I_H^h(x_{H,k,m} - x_{H,k,0})$
  else
    Choisir $\tau_{h,k}$
    $x_{h,k+1} = x_{h,k} - \tau_{h,k} \nabla F_h(x_{h,k})$
  end
end

itérations de FB au niveau fin. Ces deux concepts sont rappelés avant d'introduire la méthode proposée.

Forward-Backward (FB) Pour résoudre un problème d'optimisation du type (1) où la fonction g_h n'est pas différentiable, une stratégie standard repose sur des itérations du schéma forward-backward qui au niveau fin s'écrit pour tous $\tau_{h,k} > 0$:

$$x_{h,k+1} = \text{prox}_{\tau_{h,k} g_h}(x_{h,k} - \tau_{h,k} \nabla f_h(x_{h,k})),$$

et pour lequel la convergence de la suite $x_{h,k+1}$ vers un minimiseur de $F_h = f_h + g_h$ est assurée pour $\tau_{h,k} \in]0, 2L_{f_h}^{-1}[$.

Enveloppe de Moreau Les itérations au niveau grossier vont reposer sur l'enveloppe de Moreau des fonctions g_h et g_H (la version réduite de g_h de taille N_H). Soient $\gamma > 0$ et g une fonction convexe, semi-continue inférieurement et propre de $\mathcal{H}$ dans $]-\infty, +\infty]$, son enveloppe de Moreau, notée ${}^{\gamma}g$ est la fonction convexe, continue, à valeurs réelles, définie par

$${}^{\gamma}g = \inf_{y \in \mathcal{H}} g(y) + (1/2\gamma)\| \cdot - y\|^2 \tag{4}$$

ou également, pour tout $x \in \mathcal{H}$, par

$${}^{\gamma}g(x) = g\left(\text{prox}_{\gamma g}(x)\right) + \frac{1}{2\gamma}\|x - \text{prox}_{\gamma g}(x)\|^2. \tag{5}$$

Le gradient de ${}^{\gamma}g_H$ est γ^{-1}-lipschitzien et tel que (Prop. 12.30 [4]) :

$$\nabla {}^{\gamma}g_H = \gamma^{-1}(\text{Id} - \text{prox}_{\gamma g_H}). \tag{6}$$

${}^{\gamma}g_H$ fournit donc un choix naturel d'approximation qui est lisse et permet l'application de la méthode du gradient classique au niveau grossier. Le choix de γ sera discuté en section 4.

MMFB On définit

$$F_{H,\gamma}(x_H) = f_H(x_H) + {}^{\gamma}g_H(x_H) \tag{7}$$

puis, uniquement pour définir $v_{H,k}$, on reformule le pas classique du forward-backward en utilisant cette fois l'enveloppe de Moreau de g_h de paramètre $\tau_{h,k}$:

$$x_{h,k+1} = x_{h,k} - \tau_{h,k} D_{h,k} \tag{8}$$

avec

Algorithm 2: MMFB

Data: $x_{h,0}$, $\epsilon_h > 0$, $\kappa \in]0, \min\{1, \|I_h^H\|\}[$, $\gamma > 0$, $m > 0$, $p > 0$, $k = 0$, $r = 0$
while $\|x_{h,k+1} - x_{h,k}\| > \epsilon_h$ **do**
  if $\|I_h^H D_{h,k}\| > \kappa \|D_{h,k}\|$ *and* $r < p$ **then**
    $r = r + 1$
    $x_{H,k,0} = I_h^H x_{h,k}$
    $v_{H,k} = I_h^H D_{h,k} - \nabla F_{H,\gamma}(x_{H,k,0})$
    for $\ell = 0 \ldots m-1$ **do**
      Choisir $\alpha_{H,k,\ell}$
      $d_{H,k,\ell} = -\alpha_{H,k,\ell} (\nabla F_{H,\gamma}(x_{H,k,\ell}) + v_{H,k})$
      $x_{H,k,\ell+1} = x_{H,k,\ell} + d_{H,k,\ell}$
    end
    Choisir $\bar{\tau}_{h,k}$, et calculer :
    $\bar{x}_{h,k} = x_{h,k} + \bar{\tau}_{h,k} I_H^h(x_{H,k,m} - x_{H,k,0})$
    Choisir $\tau_{h,k}$ et calculer :
    $x_{h,k+1} = \text{prox}_{\tau_{h,k} g_h}(\bar{x}_{h,k} - \tau_{h,k} \nabla f_h(\bar{x}_{h,k}))$
  else
    Choisir $\tau_{h,k}$ et calculer :
    $x_{h,k+1} = \text{prox}_{\tau_{h,k} g_h}(x_{h,k} - \tau_{h,k} \nabla f_h(x_{h,k}))$
  end
end

$$D_{h,k} = \nabla f_h(x_{h,k}) + \nabla^{\tau_{h,k}} g_h(x_{h,k}) - (\tau_{h,k}) \nabla^{\tau_{h,k}} g_h(\nabla f(x_{h,k})) \tag{9}$$

où $D_{h,k}$ joue le rôle de $\nabla F_h(x)$ dans les méthodes en section 2 et il permet une bonne définition du terme de correction $v_{H,k}$:

$$v_{H,k} = I_h^H D_{h,k} - (\nabla f_H(x_{H,k,0}) + \nabla^{\gamma} g_H(x_{H,k,0}))$$

qui garantit la cohérence du premier ordre : $D_{H,k,0} = I_h^H D_{h,k}$ où $D_{H,k,0}$ joue le même rôle que pour la méthode du gradient multi-niveau. Par rapport aux algorithmes développés jusqu'ici [1], cette version permet une définition claire et une construction simple de g_H et $v_{H,k}$ dès lors que le calcul de l'opérateur proximal associé possède une forme explicite : une hypothèse assez raisonnable ici. Notre méthode est résumée dans l'algorithme 2. Les paramètres liés à la longueur de pas au niveau fin et au niveau grossier peuvent être sélectionnés en fixant une taille sous le seuil garantissant la convergence, défini par les constantes de Lipschitz associées aux fonctions considérées, ou par une recherche linéaire. La seconde solution est généralement plus coûteuse, mais peut apporter dans certains cas une convergence plus rapide.

Pour assurer la convergence des itérées, on impose au plus p utilisations des modèles grossiers F_H, ce qui est recommandé pour obtenir un bon temps de calcul (cf section 4).

Convergence des itérées Dès lors que l'on utilise les modèles grossiers un nombre fini de fois, on peut prouver la convergence des itérées vers un minimiseur de $F = F_h$. On définit d'abord la suite de corrections issues des modèles grossiers.

Lemme 1. *Supposons que* $\sup_{k,\ell} \alpha_{H,k,\ell} < 2(L_{f,H} + \gamma)^{-1}$, *avec* $L_{f,H}$ *la constante de Lipschitz de* f_H, *et que la suite des* $\tau_{h,k} \in]0, +\infty[$ *est telle que* $\sup_{k \in \mathbb{N}} \tau_{h,k} < 2L_{f_h}^{-1}$. *En notant* $\bar{\tau} = \sup_k \bar{\tau}_{h,k}$, *la suite* $(b_{h,k})_{k \in \mathbb{N}}$ *dans* $\mathcal{H}$ *issue de l'algorithme,*

définie comme

$$b_{h,k} = (\tau_{h,k})^{-1}\bar{\tau}(Id - \tau_{h,k}\nabla f_h)\left(I_H^h(x_{H,k,m} - x_{H,k,0})\right)$$

si on utilise les modèles grossiers et 0 *sinon ; est telle que* $\sum_{k\in\mathbb{N}} \|b_{h,k}\| < +\infty$.

Démonstration. La suite admet au plus p termes non nuls, qui sont tous majorés en norme par $\bar{\tau}(1/\tau_{h,k} + L_h)\|I_H^h\|C_H$ avec $C_H = \sup\{\|x_{H,k,0} - x_{H,k,*}\|\} < +\infty$ où $x_{H,k,*}$ est un minimum du modèle grossier défini à l'itération k. Par l'hypothèse sur les $\alpha_{H,k,\ell}$, les suites $(x_{H,k,\ell})_{l\in\mathbb{N}}$ convergent toutes et sont monotones, on peut donc majorer par C_H. □

Théorème 1. *Sous l'hypothèse que* F_h *admette au moins un minimum, et les mêmes hypothèses que celles du Lemme 1, alors si on fixe* $x_{h,0} \in \mathcal{H}$*, et :*

$$x_{h,k+1} = prox_{\tau_{h,k}g_h}(x_{h,k} - \tau_{h,k}(\nabla f_{h,k} + b_{h,k})) \qquad (10)$$

Alors $x_{h,k}$ *converge vers un minimiseur de* F_h.

La preuve découle du Lemme 1 et du Théorème 3.4 dans [6].

4 Résultats numériques

Nous illustrons numériquement les performances de notre algorithme dans le cadre de la restauration d'image en le comparant à l'algorithme forward-backward (FB) avec ou sans recherche de pas linéaire, implémentée via une stratégie de backtracking. Nous proposons de résoudre le problème (2) afin de restaurer des images de grande taille (e.g. 2048×2048, i.e., $N \simeq 4.10^6$). Ces images sont dégradées par un flou représenté par une matrice A_h construite avec HNO [2] sous la forme d'un produit de Kroenecker avec des conditions aux bords symétriques et par un bruit gaussien (voir légende Fig. 1). Pour le modèle grossier, on utilise I_h^H comme la matrice décimée issue de l'ondelette de Haar $[1\ 1]$ et $I_H^h = \frac{1}{4}(I_h^H)^T$. On construit ensuite f_H avec la matrice de flou $A_H = I_h^H A_h I_H^h$ (qui n'est jamais utilisée explicitement grâce aux propriétés du produit de Kroenecker [1, 2]). Ainsi $f_h = \frac{1}{2}\|A_h x_h - z_h\|^2$ et $f_H = \frac{1}{2}\|A_H x_H - I_h^H z_h\|^2$. Le terme de pénalisation $g_h = \|W_h x_h\|_1$ est défini avec une décomposition maximale avec les ondelettes à 11 niveaux, on construit $g_H = \|W_H x_H\|_1$ avec une décomposition à 10 niveaux, puis 9 au modèle grossier suivant etc. Sur la figure 2 sont représentées les courbes de convergence en fonction des itérations effectuées ou du temps de calcul CPU pour l'image ≪ Lunch atop a Skyscraper ≫ en noir et blanc avec une hiérarchie à 3 niveaux : 2048×2048, 1024×1024, et 512×512. Le choix fait ici basé sur nos expériences numériques est de n'autoriser qu'une seule utilisation des modèles grossiers, faite au début du processus itératif et qui permet de trouver un bon point de départ pour le niveau fin.

Pour l'enveloppe de Moreau de g_H, on a choisi $\gamma = 5$, qui semble être un bon compromis entre vitesse de convergence du modèle grossier (γ grand) et régularité de l'approximation (γ petit pour être proche de g_H). Les résultats de notre approche peuvent varier si le backtracking et utilisé ou non au niveau grossier comme illustré Fig. 2 ; mais dans tous les cas notre méthode est plus rapide que FB. Par ailleurs, le backtracking au niveau fin n'améliorant pas les performances par rapport à un pas fixe, les résultats associés ne sont pas présentés. La convergence plus rapide observée en figure 2 se traduit en une reconstruction plus rapidement comme illustré sur l'image de la figure 1 après 2 itérations.

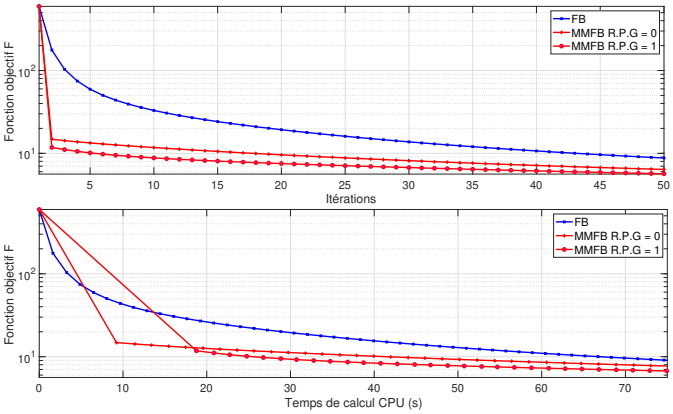

FIGURE 2 – Fonction objectif pour ≪ Lunch atop a Skyscraper ≫ avec ou sans backtracking (R.P.) au niveau grossier (G.) en fonction du nombre d'itérations (haut) et du temps de calcul (bas).

5 Conclusion

Nous avons proposé une méthode FB multi-niveau convergente pour la restauration d'images permettant d'approcher la solution optimale en un nombre d'itérations bien plus faible et d'envisager l'analyse d'images très haute-résolution. Une perspective de recherche future est d'étendre cette approche à d'autres schémas algorithmiques proximaux ainsi que d'étudier les taux de convergence associés.

Références

[1] P. Parpas, *A Multilevel Proximal Gradient Algorithm for a Class of Composite Optimization Problems*. SIAM Journal on Scientific Computing, 2017.

[2] P. Hansen, J. Nagy, et D. O'Leary, *Deblurring Images : Matrices, Spectra, and Filtering*. Fundam. Algorithms 3, SIAM, Philadelphia, 2006.

[3] A. Beck, et M. Teboulle, *A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems*. SIAM J. Imaging Sciences, Vol. 2, No. 1, pp 183-202, 2009.

[4] H. Bauschke, P. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics, 2017.

[5] P. L. Combettes and J.-C. Pesquet *Proximal splitting methods in signal processing*. In Fixed-point algorithms for inverse problems in science and engineering, Springer, New York, NY, pp. 185-212, 2011.

[6] P. Combettes, V. Wajs, *Signal recovery by proximal forward-backward splitting*. SIAM Multiscale Model. Simul., Vol. 4, No. 4, pp 1168-1200, 2005.

[7] S.G. Nash, *A multigrid approach to discretized optimization problems*. Optim. Methods Softw. 14,99–116 2000.

[8] H. Calandra, S. Gratton, E. Riccietti, X. Vasseur, *On high-order multilevel optimization strategies*, SIAM Journal on Optimization, 31(1) :307-330, 2020.

[9] E. Chouzenoux, J.-C. Pesquet, A. Florescu, *A Stochastic 3MG Algorithm with Application to 2D Filter Identification*, In Proceedings of the 22nd European Signal Processing Conference (EUSIPCO 2014), pages 1587-1591, Lisbon, Portugal, 1-5 september 2014.

[10] É. Thiébaut, *Optimization issues in blind deconvolution algorithms*, SPIE Conf. Astronomical Data Analysis II, Vol. 4847, 174-183, 2002.