

# Reconstruction de phase exploitant des propriétés de parcimonie et application à la reconstruction de séquences vidéo de microscopie

Yoann LE MONTAGNER<sup>1,2</sup>, Elsa ANGELINI<sup>2</sup>, Jean-Christophe OLIVO-MARIN<sup>1</sup>

<sup>1</sup>Institut Pasteur, Unité d'Analyse d'Images Quantitative, CNRS URA 2582, F-75015 Paris

<sup>2</sup>Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, F-75013 Paris

lemontag@telecom-paristech.fr, elsa.angelini@telecom-paristech.fr, jcolivo@pasteur.fr

**Résumé** – La théorie de l'échantillonnage compressé (*compressed sensing*, CS) prédit que les images présentant une certaine structure peuvent être échantillonnées de façon compressée à partir d'un faible nombre de projection linéaire non-adaptatives. Néanmoins, est-il possible d'étendre cette théorie à des mesures non-linéaires, telles que des mesures de module de transformée de Fourier optique (donc dépourvues d'information de phase) ? Dans ce papier, nous étudions ce problème dans le cadre de la reconstruction de vidéos de microscopie. Nous proposons un algorithme de reconstruction exploitant une contrainte de régularisation similaire à la variation totale (TV), et présentons des résultats préliminaires obtenus sur des vidéos de test réelles.

**Abstract** – The theory of compressed sensing (CS) predicts that structured images can be sampled in a compressive manner with very few non-adaptive linear measurements. However, is such a recovery still possible with non-linear measurements, such as optical-based Fourier modulus (therefore deprived of phase information)? In this paper, we study this problem in the context of microscopy video reconstructions. We propose a recovery algorithm exploiting a specific TV-based regularization constraint, and we present some preliminary results obtained on real test sequences.

## 1 Introduction

### 1.1 Principe du *compressed sensing*

Le problème inverse visé par le CS peut se formuler de la façon suivante : étant donné un signal d'intérêt  $x \in \mathbb{R}^N$  observé via un opérateur linéaire *aléatoire*  $\Phi$  générant un vecteur de mesures  $y \in \mathbb{R}^M$  avec  $M \ll N$ , est-il possible de reconstruire  $x$  à partir de  $y$  ? Le caractère aléatoire de  $\Phi$  ne doit pas être compris dans un sens strict, mais comme le fait que  $\Phi$  doit diffuser l'information contenue dans  $x$  sur l'ensemble du vecteur  $y$ . Parmi les opérateurs de ce type, citons les matrices aléatoires gaussiennes ou de Bernouilli [3], les transformées de Fourier ou Hadamard sous-échantillonnées aléatoirement [2], ou les matrices unitaires construites en vue d'applications particulières [4].

Plusieurs travaux ([5, 1]) montrent que, sous certaines conditions techniques,  $x$  peut être reconstruit à partir de  $y$  s'il peut être représenté comme une combinaison linéaire parcimonieuse de vecteurs d'une base connue (ou d'un dictionnaire). Alors, l'étape de reconstruction consiste à résoudre un problème d'optimisation convexe comprenant typiquement une contrainte sur la norme  $l_1$  du jeu de coefficients de la représentation de  $x$  dans cette base.

### 1.2 Mesures de Fourier optiques

Dans cet article, nous nous intéressons au cas où l'opérateur de mesure n'acquiert que le module de la transformée

de Fourier, et où seule une sélection aléatoire uniforme des coefficients de Fourier est effectivement échantillonné. La raison pour laquelle on s'intéresse à des procédures d'échantillonnage telles que celle-ci est qu'il est possible d'implémenter des opérateurs de transformée de Fourier au moyen de montages optiques. Cette propriété peut être mise à profit pour concevoir des systèmes d'acquisition réalisant une transformée de Fourier optique en amont des capteurs photo-électriques, permettant ainsi de bénéficier des avantages de l'échantillonnage compressé.

Néanmoins, la mesure du signal complexe issu de la transformée de Fourier demeure ardue, dans la mesure les transducteurs photo-électriques tels que les capteurs CCD ou CMOS produisent des signaux correspondant à l'énergie des photons incidents, laquelle est indépendante de la phase de l'onde électro-magnétique correspondante. Mesurer la phase nécessite des montages optiques plus complexes, qui ne sont pas toujours compatibles avec les impératifs expérimentaux du moment.

Pour contourner ce problème, nous proposons ici de reconstruire directement l'image sans connaître la phase des coefficients de Fourier mesurés. Formellement, le système à inverser devient alors :

$$y = |\Phi x| + \text{bruit} \quad (1)$$

où  $x$  est le signal d'intérêt à reconstruire (une image 2D dans notre cas),  $y$  le vecteur de mesure,  $\Phi$  une transformée de Fourier partielle (i.e. dont on a supprimé un grand

nombre de lignes), et  $|\cdot|$  l'opération de passage au module sur chacune des composantes du vecteur en argument. Ce problème, bien connu en l'absence de ce passage au module, nécessite de développer de nouvelles méthodes pour pallier l'absence d'information de phase.

### 1.3 Phase retrieval

La reconstruction d'un signal à partir du module de sa transformée de Fourier est un problème de traitement du signal classique, dénommé reconstruction de phase (*phase retrieval*) : cette technique de reconstruction a été utilisée par exemple dans le cadre de modalités d'imagerie par rayons X, pour des applications en cristallographie (voir [7, 11]). Pour reconstruire un signal  $x \in \mathbb{R}^N$  à partir d'un vecteur de mesure  $y$  vérifiant (1), la méthode proposée par [7] distingue tout d'abord deux sous-ensembles de  $\mathbb{R}^N$  :

- Le sous-ensemble  $\mathcal{D}_{y,\epsilon}$  (*data set*), qui caractérise l'ensemble des images compatibles avec le vecteur de mesures  $y$  :  $\mathcal{D}_{y,\epsilon} = \{x \in \mathbb{R}^N \text{ t.q. } \|y - |\Phi x|\|_2 \leq \epsilon\}$ . Cet ensemble dépend d'un paramètre  $\epsilon$ , lié au niveau de bruit affectant les mesures.
- Un sous-ensemble  $\mathcal{R}$  (*regularization set*), qui correspond à l'ensemble des signaux qui vérifient un ou plusieurs critères de régularité définis *a priori*. Dans le cadre d'applications de cristallographie,  $\mathcal{R}$  consiste typiquement en l'ensemble des images 2D dont le support est restreint à un sous-ensemble de pixels.

L'estimateur  $\hat{x}$  de l'image à reconstruire est alors calculé par cet algorithme comme étant la limite d'une suite de projections alternées sur les deux ensembles  $\mathcal{D}_{y,\epsilon}$  et  $\mathcal{R}$  :

$$\hat{x} = (\Pi_{\mathcal{R}} \circ \Pi_{\mathcal{D}_{y,\epsilon}} \circ \Pi_{\mathcal{R}} \circ \dots \circ \Pi_{\mathcal{D}_{y,\epsilon}})(x_0) \quad (2)$$

où  $x_0$  est une estimation initiale de la solution, et  $\Pi_{\mathcal{D}_{y,\epsilon}}$  et  $\Pi_{\mathcal{R}}$  désignent les projecteurs sur  $\mathcal{D}_{y,\epsilon}$  et  $\mathcal{R}$  :

$$\Pi_{\mathcal{D}_{y,\epsilon}}(x) = \arg \min_{z \in \mathcal{D}_{y,\epsilon}} \|z - x\|_2 \quad (3)$$

et  $\Pi_{\mathcal{R}}$  défini de façon analogue. [7] montre que la suite d'estimateurs (2) converge vers l'intersection de  $\mathcal{D}_{y,\epsilon}$  et  $\mathcal{R}$ . Une propriété remarquable de l'ensemble  $\Pi_{\mathcal{D}_{y,\epsilon}}$  est que, bien qu'il ne soit pas convexe, le problème de projection associé (3) est soluble de façon exacte<sup>1</sup>, et que l'évaluation de la solution est rapide en termes de temps de calcul.

D'autres travaux ont amélioré l'algorithme (2), principalement dans le but d'accélérer la vitesse de convergence (voir par exemple [10] et références incluses). Cependant, à notre connaissance, tous les algorithmes de *phase retrieval* existants reposent sur la construction d'une suite d'estimateurs qui convergent vers l'intersection de deux

1. Pour  $\epsilon = 0$ ,  $\Pi_{\mathcal{D}_{y,\epsilon}}(x)$  est calculé de la façon suivante : 1) on calcule la transformée de Fourier  $X$  de  $x$ ; 2) pour toutes les fréquences spatiales  $k$  pour lesquelles une mesure  $y(k)$  est disponible, on remplace le module calculé  $|X(k)|$  par  $y(k)$ ; 3) pour finir, inverser la transformée de Fourier. Pour  $\epsilon > 0$ , la seconde étape est plus complexe, mais peut toujours être réalisée en  $\mathcal{O}(N)$  opérations.

ensembles, l'un de caractérisant les signaux compatibles avec les mesures, l'autre les informations disponibles *a priori* sur la solution. Dans notre cas, ces informations sont de nature différente de celles disponibles dans le cas d'applications de cristallographie (i.e. contraintes sur le support du signal). Nous introduisons donc dans la section suivante une approche de reconstruction générale adaptée pour la reconstruction de vidéos de microscopie, qui exploite un schéma itératif de projections alternées similaire à celui utilisé en *phase retrieval*, mais qui autorise l'utilisation d'autres types d'informations *a priori*.

## 2 Reconstruction vidéo

### 2.1 Formulation du problème

Nous formulons notre problème de reconstruction comme suit : comment reconstruire une image 2D  $x \in \mathbb{R}^N$  issue d'une séquence vidéo, en connaissant d'une part un vecteur  $y \in \mathbb{R}_+^M$  de mesures vérifiant (1), et d'autre part l'image  $a$  précédant  $x$  dans la vidéo ?

Il est théoriquement possible, en itérant ce processus de reconstruction, de reconstituer une vidéo complète à partir d'une image initiale (*key-frame*) et de mesures partielles du module de la transformée de Fourier de chaque image de la séquence. En pratique, la propagation et l'amplification des erreurs de reconstruction sont toutefois susceptibles de limiter le nombre d'images pouvant être reconstituées à partir d'une *key-frame* initiale.

### 2.2 Variation totale hybride

Pour résoudre ce problème en utilisant les outils présentés en section 1.3, on doit préciser ce qu'est dans notre cas l'ensemble  $\mathcal{R}$ . Dans le cadre de vidéo de microscopie, inspiré de [8], on suppose que l'image  $x$  vérifie les propriétés de parcimonie suivantes :

- la carte de gradient spatial  $\nabla x$  est parcimonieuse,
- la carte de différence  $(x - a)$  entre l'image à reconstruire  $x$  et sa prédécesseuse  $a$  est parcimonieuse,
- les coefficients non-nuls dans la carte  $\nabla x$ , qui correspondent principalement aux contours des objets dans l'image  $x$ , sont situés près des coefficients non-nuls de la carte  $(x - a)$ , i.e. les fronts de déplacement.

Pour promouvoir ces propriétés, on introduit la fonctionnelle suivante – la *variation totale hybride* – définie sur l'espace  $\mathbb{R}^N$  des images 2D :

$$\|x\|_{\text{TV},w,a} = \sum_P w(P) \sqrt{\|\nabla x(P)\|_2^2 + |x(P) - a(P)|^2} \quad (4)$$

où  $P$  parcourt l'ensemble des pixels de l'image, et  $w \in \mathbb{R}^N$  est une carte de pondération vérifiant  $0 < w(P) < +\infty$ , et fixée *a priori* telle que  $w(P)$  soit petit au niveau des pixels  $P$  où l'on s'attend à observer des contours dans l'image à

```

fonction RECONSTRUCTION IMAGE( $x_0, \tau_0, \Delta_{\text{tolvar}}, \alpha$ )
 $k \leftarrow 0$ 
 $\tau \leftarrow \tau_0$ 
boucle
   $k \leftarrow k + 1$ 
   $x_k \leftarrow \Pi_{w,a,\tau} \circ \Pi_{y,\epsilon}(x_{k-1})$ 
   $\delta_k \leftarrow \|x_k - x_{k-1}\|_2 / \|x_{k-1}\|_2$ 
  si  $\delta_k < \Delta_{\text{tolvar}}$  alors
     $x_{\text{candidat}} \leftarrow x_k$ 
     $\tau \leftarrow \alpha\tau$ 
  sinon si non-convergence détectée alors
    retourner  $x_{\text{candidat}}$ 
  fin si
fin boucle
fin fonction

```

FIGURE 1 – Pseudo-code de l’algorithme de reconstruction proposé. Quatre arguments sont à définir en entrée :  $x_0$  et  $\tau_0$ , qui sont les valeurs initiales de l’image reconstruite et de la borne sur la TV hybride, ainsi que  $\Delta_{\text{tolvar}}$  et  $\alpha$ , qui contrôlent le comportement général de l’algorithme. Lorsque la variation relative  $\delta_k$  entre les itérations successives devient inférieure à  $\Delta_{\text{tolvar}}$ , la borne  $\tau$  est réduite d’un facteur  $\alpha$  ( $0 < \alpha < 1$ ), jusqu’à ce que l’algorithme devienne non-convergent.

reconstruire  $x$ . La façon dont ce paramètre  $w$  est fixé sera discutée plus avant en section 3.2. Enfin, on définit l’ensemble  $\mathcal{R}$  comme étant une ligne de niveau de cette fonctionnelle (4) :  $\mathcal{R}_{w,a,\tau} = \{x \in \mathbb{R}^N \text{ t.q. } \|x\|_{\text{TV},w,a} \leq \tau\}$ . Le nouveau paramètre  $\tau > 0$  sera par la suite ajusté dynamiquement lors du processus de reconstruction.

### 2.3 Opérateur de projection $\Pi_{\mathcal{R}_{w,a,\tau}}$

Utiliser  $\mathcal{R}_{w,a,\tau}$  dans le schéma de reconstruction (2) nécessite de pouvoir évaluer efficacement  $\Pi_{\mathcal{R}_{w,a,\tau}}(x)$  :

$$\Pi_{\mathcal{R}_{w,a,\tau}}(x) = \arg \min_z \|z - x\|_2 \text{ t.q. } \|z\|_{\text{TV},w,a} \leq \tau \quad (5)$$

Pour résoudre ce problème d’optimisation convexe, nous utilisons un algorithme inspiré de la méthode de projection proposée par [6]. Le principe sur lequel repose cet algorithme consiste à reformuler le problème contraint (5) en un problème non-contraint équivalent :

$$\arg \min_u \tau \|u\|_{\infty, \frac{1}{w}} + \left\langle u^{(1)}, a \right\rangle + \frac{1}{2} \left\| x - u^{(1)} + \text{Div } u^{(2,3)} \right\|_2^2 \quad (6)$$

Ici,  $u$  représente un champ de vecteurs 3D (i.e. une image à 3 canaux) défini sur le même domaine que  $x$ ,  $u^{(1)}$  et  $u^{(2,3)}$  désignent respectivement la première et la concaténation des deuxième et troisième composantes de  $u$ ,  $\text{Div}$  est l’opérateur adjoint de  $-\nabla$ , et  $\|u\|_{\infty, \frac{1}{w}} = \max_P \frac{1}{w(P)} \|u(P)\|_2$  où  $P$  parcourt l’ensemble des pixels. Sans rentrer dans les détails, on peut vérifier que (5) et (6) sont effectivement équivalents, et que les points optimaux correspondants  $z^*$  et  $u^*$  sont liés par la relation  $z^* = x - u^{*(1)} + \text{Div } u^{*(2,3)}$ ; (6) est quant à lui résolu via une descente de gradient incorporant une accélération type Nesterov (voir [6, 12, 13]).

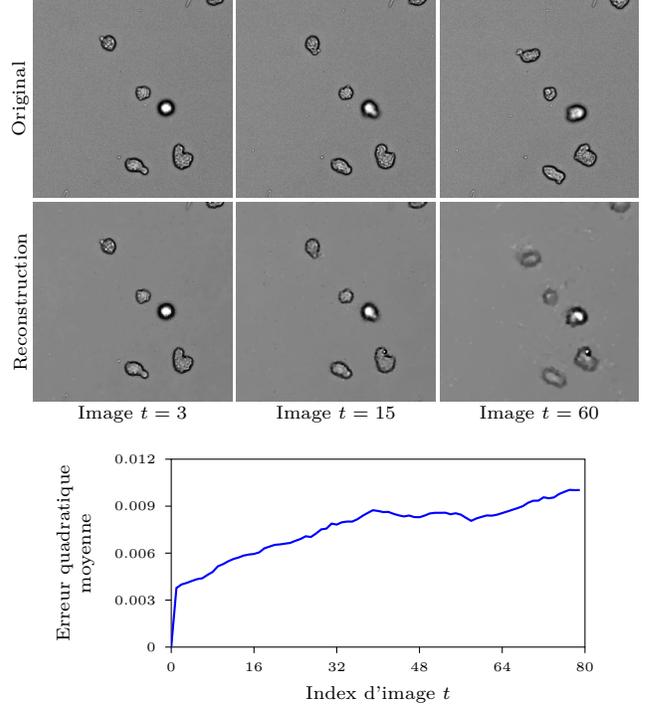


FIGURE 2 – Reconstruction de la séquence de test *Amibes*, à partir de l’image initiale  $t = 0$  et des valeurs de module de 25% des coefficients de Fourier de chaque image de la séquence. La carte de pondération est initialisée de façon uniforme ( $w = 1$ ). Est également présentée l’évolution en fonction du temps de l’erreur quadratique moyenne calculée image-par-image entre la vidéo originale et le résultat.

### 2.4 Algorithme complet

L’algorithme de reconstruction proposé est basé sur des projections alternées sur  $\mathcal{D}_{y,\epsilon}$  et  $\mathcal{R}_{w,a,\tau}$ , dont les définitions font intervenir deux paramètres scalaires  $\epsilon$  et  $\tau$ .  $\epsilon$  contrôle la taille de l’ensemble  $\mathcal{D}_{y,\epsilon}$ , et sa valeur est définie proportionnellement à l’écart-type du bruit affectant les mesures; définir la valeur de  $\tau$  est en revanche moins immédiat, aussi nous avons développé une heuristique adaptative pour ajuster dynamiquement ce paramètre lors du processus de reconstruction itératif.

Ce processus d’ajustement dynamique repose sur l’observation suivante. Le système de projection alternée (2) produit une suite d’estimateurs qui convergent vers l’intersection  $\mathcal{R}_{w,a,\tau} \cap \mathcal{D}_{y,\epsilon}$ , mais cette intersection est vide lorsque  $\tau$  est inférieur à un certain seuil  $\tau^*$  (si l’image  $a$  n’est pas constante, l’ensemble  $\mathcal{R}_{w,a,\tau}$  est lui-même vide lorsque  $\tau = 0$ ): par conséquent, l’algorithme devient non-convergent si  $\tau < \tau^*$ . À partir de cette remarque, nous proposons un algorithme de reconstruction où  $\tau$  est initialisé à une valeur élevée arbitraire  $\tau_0$ , puis réduite jusqu’à ce que l’algorithme devienne non convergent (voir le pseudo-code dans la Fig. 1). L’algorithme retourne le résultat (noté  $x_{\text{candidat}}$ ) obtenu avec la plus petite valeur de

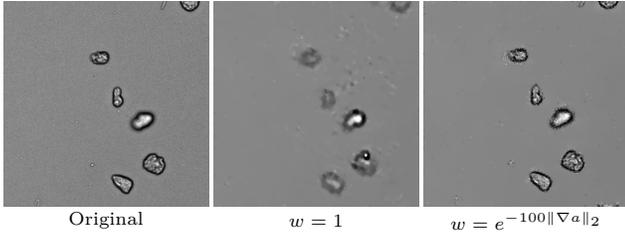


FIGURE 3 – Image  $t = 70$  de la séquence *Amibes* reconstruite avec une carte de pondération  $w$  uniforme, et une carte non-uniforme adaptée.

$\tau$  qui conduit à la convergence.

Cependant, détecter que la suite d’estimateurs  $x_k$  ne va pas converger pour une valeur donnée du paramètre  $\tau$  est un problème difficile, dans la mesure où nous ne disposons pas de résultat théorique quant à la vitesse de convergence de cette suite. Pour résoudre cette question, nous avons mis en place un critère empirique basé sur les propriétés de la suite  $\delta_k$ , qui mesure les variations relatives entre deux itérations successives (voir [9]). Les résultats obtenus valident cette approche, même si elle n’est vraisemblablement pas optimale en termes de temps de calcul.

## 3 Résultats

### 3.1 Exemple de reconstruction

La figure 2 présente un résultat de reconstruction obtenu avec une séquence de test de microscopie (cellules d’*amibes*). Pour cette simulation, 25% des coefficients de Fourier (choisis de façon aléatoire uniforme) ont été utilisés ; la première image de la séquence (correspondant à  $t = 0$ ) est en outre utilisée en tant que *key-frame* (cf. Sec. 2.1). La carte de pondération  $w$  est fixée – dans un premier temps – à une valeur uniforme ( $w(P) = 1$ ).

Le résultat obtenu pour les premières images de la séquence (i.e. celles correspondantes aux petites valeurs de  $t$ , typiquement  $t < 10$ ) présentent un bon niveau de fidélité à l’original, cette fidélité se dégradant au cours du temps : on note en particulier un accroissement de l’effet de flou au niveau des contours des objets, et une augmentation de l’erreur quadratique moyenne calculée sur chaque image entre l’original et la reconstruction. Ces phénomènes résultent de l’accumulation d’erreurs de reconstruction sur chaque image, l’estimateur de l’image  $t$  étant défini à partir de l’estimateur de l’image  $t - 1$ .

### 3.2 Carte de pondération $w$

Le résultat de la figure 2 a été obtenu avec une carte de pondération  $w$  uniforme. Toutefois, l’utilisation d’une carte de pondération non-uniforme présentant les propriétés décrites en Sec. 2.2 permet de contenir l’apparition des distorsions mentionnées ci-dessus, et d’améliorer ainsi

nettement la qualité de la vidéo reconstruite.

Plusieurs cartes de pondération ont été testées, notamment  $w(P) = e^{-\kappa\|\nabla a(P)\|_2}$  et  $w(P) = (\eta + \|\nabla a(P)\|_2)^{-1}$ , avec différentes valeurs des paramètres  $\kappa > 0$  et  $\eta > 0$ . La figure 3 présente un exemple d’amélioration obtenue grâce à ces cartes non-uniformes : l’effet de flou est en particulier complètement supprimé. Néanmoins, le choix de la carte de pondération ainsi que la calibration automatique des paramètres  $\kappa$  ou  $\eta$  demeurent problématique, cette calibration étant réalisée de façon manuelle pour le moment.

## 4 Conclusion

Dans cet article, nous avons présenté une nouvelle méthode de reconstruction de vidéos à partir de mesures de module de Fourier partielles, inspirée de la théorie du CS. Cette méthode de reconstruction est inspirée de techniques de *phase retrieval*, et introduit une énergie de régularisation originale, la *variation totale hybride*. Les résultats préliminaires démontrent la faisabilité de la reconstruction vidéo réalisée à partir de mesures de Fourier partielles, et privées d’information de phase. Plusieurs problèmes restent néanmoins à résoudre quant à la détermination automatique des différents paramètres utilisés dans la contrainte de régularisation, afin d’améliorer la robustesse de la méthode en vue de son implémentation au sein d’un dispositif d’imagerie réel utilisant le CS comme paradigme d’échantillonnage.

## Références

- [1] E. Candès, Y. C. Eldar, D. Needell, and P. Randall. Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, 31(1) :59–73, 2010.
- [2] E. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse Problems*, 23(3) :969–985, 2007.
- [3] E. Candès and T. Tao. Near-optimal signal recovery from random projections : universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12) :5406–5425, 2006.
- [4] T. T. Do, L. Gan, N. H. Nguyen, and T. D. Tran. Fast and Efficient Compressive Sensing Using Structurally Random Matrices. *IEEE Transactions on Signal Processing*, 60(1) :139–154, 2012.
- [5] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4) :1289–1306, 2006.
- [6] J. M. Fadili and G. Peyré. Total variation projection with first order schemes. *IEEE Trans. Image Processing*, 20(3) :657–669, 2011.
- [7] J. R. Fienup. Phase retrieval algorithms : a comparison. *Applied optics*, 21(15) :2758–2769, Aug. 1982.
- [8] Y. Le Montagner, E. Angelini, and J.-C. Olivo-Marin. Video reconstruction using compressed sensing measurements and 3D total variation regularization for bio-imaging applications. In *ICIP*, pages 917–920. IEEE, 2012.
- [9] Y. Le Montagner, E. Angelini, and J.-C. Olivo-Marin. Phase retrieval with sparsity priors and application to microscopy video reconstruction. In *ISBI*. IEEE, 2013.
- [10] S. Marchesini. A unified evaluation of iterative projection algorithms for phase retrieval. *Rev. scientific instruments*, 78(1), 2007.
- [11] J. Miao, P. Charalambous, and J. Kirz. Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature*, 400(July) :342–344, 1999.
- [12] Y. Nesterov. Gradient methods for minimizing composite objective function. Technical report, Université Catholique de Louvain, 2007.
- [13] P. Weiss. *Algorithmes rapides d’optimisation convexe : application à la restauration d’images et à la détection de changements*. PhD thesis, Université de Nice - Sophia Antipolis, 2008.