Apprentissage auto-supervisé pour la reconstruction de phase

Victor SECHAUD[1] Patrice ABRY[1] Laurent JACQUES[2] Julián TACHELLA[1]

[1]CNRS, ENS de Lyon, LPENSL, UMR5672, 69342, Lyon Cedex 07, France

[2]UCLouvain, ICTEAM, Louvain-la-Neuve, Belgium

Résumé – Ces dernières années, les réseaux neuronaux profonds se sont imposés comme solution pour les problèmes inverses en imagerie. Ces réseaux sont généralement entraînés à l'aide de paires d'images : l'une dégradée et l'autre en haute qualité, ces dernières étant appelées « vérités terrain ». Toutefois, en imagerie médicale, le manque de données de références limite l'apprentissage supervisé. Des avancées récentes ont permis de reconstruire des images à partir des seules données de mesure, supprimant ainsi le besoin de vérités terrain. Cependant, ces méthodes restent limitées aux problèmes linéaires, excluant les processus non linéaires comme la reconstruction de phase. Nous proposons une méthode auto-supervisée qui surmonte cette difficulté dans le cas de reconstruction de phase en utilisant l'invariance naturelle des images aux translations.

Abstract – In recent years, deep neural networks have emerged as a solution for inverse imaging problems. These networks are generally trained using pairs of images: one degraded and the other of high quality, the latter being called 'ground truth'. However, in medical and scientific imaging, the lack of fully sampled data limits supervised learning. Recent advances have made it possible to reconstruct images from measurement data alone, eliminating the need for references. However, these methods remain limited to linear problems, excluding non-linear problems such as phase retrieval. We propose a self-supervised method that overcomes this limitation in the case of phase retrieval by using the natural invariance of images to translations.

1 Introduction

Les problèmes inverses jouent un rôle essentiel dans diverses applications en ingénierie et en sciences, telles que la tomographie et l'IRM [10]. Ces problèmes impliquent la reconstruction d'un signal inconnu $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^n$ à partir de mesures observées $\boldsymbol{y} \in \mathcal{Y} \subseteq \mathbb{R}^m$, modélisées par l'équation :

$$\boldsymbol{y} = h(\boldsymbol{x}) + \boldsymbol{\epsilon},$$

où $h : \mathbb{R}^n \mapsto \mathbb{R}^m$ est l'opérateur direct, et $\boldsymbol{\epsilon} \in \mathbb{R}^m$ représente le bruit. Ces problèmes sont souvent mal posés en raison de deux facteurs principaux : l'opérateur direct h peut entraîner une perte d'information (par exemple, lorsque $m < n$) et le bruit introduit une incertitude supplémentaire.

Pour y remédier, des informations préalables sur le signal sont généralement incorporées, telles que des hypothèses de constances par morceaux [14]. Si les a priori peuvent permettre des reconstructions robustes, en choisir un approprié est difficile, car un a priori inexact peut biaiser la solution et donc ne pas représenter le signal sous-jacent de manière efficace.

Une autre solution consiste à apprendre directement la fonction de reconstruction qui approxime l'opérateur inverse de h. Les méthodes d'apprentissage supervisé y parviennent en entrainant un réseau de neurones $f_{\boldsymbol{\theta}} : \mathcal{Y} \to \mathcal{X}$. Les paramètres du réseau $\boldsymbol{\theta} \in \mathbb{R}^p$ sont optimisés en minimisant l'erreur quadratique moyenne (EQM) sur un ensemble de données de signaux de vérité de terrain appariés et leurs mesures, $\{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i \in I}$:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i \in I} \|f_{\boldsymbol{\theta}}(\boldsymbol{y}_i) - \boldsymbol{x}_i\|^2. \tag{1}$$

Quoique cette approche donne fréquemment de bons résultats, elle présente deux inconvénients majeurs : *(i)* disponibilité des données d'entrainement : l'obtention d'un ensemble de données d'entraînement complet avec des vérités terrain précis peut être difficile ou peu pratique, en particulier dans des domaines tels que l'imagerie scientifique [1]. *(ii)* Décalage de la distribution : même avec un ensemble de données approprié, les performances peuvent se dégrader en cas de décalage entre la distribution des données d'apprentissage et celle des données réelles.

L'apprentissage auto-supervisé offre une alternative qui surmonte ces défis [1]. Contrairement aux approches supervisées, il ne nécessite pas d'ensembles de données contenant des signaux de référence. Au lieu de cela, les méthodes auto-supervisées peuvent être entrainées directement sur les données de mesure, ce qui est particulièrement avantageux lorsque les données de vérité terrain sont coûteuses ou indisponibles.

En outre, en évitant de s'appuyer sur les vérités terrain, l'apprentissage auto-supervisé réduit le risque de décalage de la distribution.

Cependant, l'apprentissage uniquement à partir de mesures présente des difficultés particulières, notamment lorsque l'opérateur direct n'est pas inversible. Pour les opérateurs linéaires, l'apprentissage auto-supervisé devient possible en supposant que l'espace de signal sous-jacent présente une invariance par rapport à des transformations spécifiques, telles que des rotations ou des translations [3]. Cette approche peut également être étendue à certains problèmes non linéaires, tels que la reconstruction de signaux fortement quantifiés, par exemple, des mesures à 1-bit [20] ou le declipping [16].

Dans ce travail, nous proposons un cadre d'apprentissage auto-supervisé qui applique l'équivariance pour les translations. Ce cadre est appliqué aux problèmes de reconstruction de phase, démontrant par des expériences qu'il permet d'obtenir des performances comparables aux méthodes entièrement supervisées.

Nos principales contributions sont les suivantes :

1. Fonction de perte auto-supervisée : Nous introduisons une fonction de perte pour l'entraînement des réseaux neuronaux dans les tâches de reconstruction de phase qui ne nécessite pas la connaissance de vérité terrain.

2. Nous montrons que la méthode auto-supervisée proposée peut être aussi performante que l'approche supervisée.

2 État de l'art

Apprentissage auto-supervisé - Dans le contexte de la restauration de données (comme le débruitage ou le défloutage), les méthodes auto-supervisées permettent de former des réseaux de reconstruction en utilisant uniquement des données de mesure [3]. Les méthodes existantes peuvent être classées en deux catégories : celles qui traitent le bruit du problème, par exemple les méthodes Noise2X [7, 8], et celles qui traitent les opérateurs non inversibles [19]. Le dernier ensemble de méthodes se concentre principalement sur le cas de l'opérateur de mesure linéaire où la perte d'information est associée à son espace nul. Il est possible de remédier à ce manque d'information en *(i)* en accédant aux mesures de plusieurs opérateurs ayant des espaces nuls différents [19], ou *(ii)* en supposant que la distribution du signal soit invariante par rapport à un ensemble de transformations telles que les translations ou les rotations, une méthode connue sous le nom d'imagerie équivariante (EI) [3]. Le cadre de l'EI est bien étudié pour les problèmes linéaires inverses, à la fois en termes d'applications [3, 15] et de théorie [19]. À notre connaissance, les seuls cas de problèmes inverses non linéaires dans le cadre de l'EI sont l'acquisition compressive à un bit [20] et la désaturation [16].

Reconstruction de phase - Les méthodes classiques de reconstruction de phase, telles que l'algorithme de Gerchberg-Saxton [5], reposent sur un raffinement itératif, mais sont limitées par une convergence lente et une sensibilité au bruit. Les approches basées sur des modèles répondent à ces limitations en incorporant des techniques de régularisation telles que la sparsité [11] pour améliorer la qualité de la reconstruction ; cependant, leurs performances dépendent fortement de la précision des a priori supposés. Les techniques d'optimisation convexe, notamment PhaseLift [2] et PhaseMax [6], reformulent la reconstruction de phase comme un problème convexe en se plaçant dans l'espace de matrice ou en utilisant la programmation semi-définie. Bien que ces méthodes offrent des garanties de récupération sous certaines conditions, elles nécessitent des calculs intensifs et des ratios d'échantillonnage élevés. Enfin, les méthodes d'apprentissage, principalement basées sur des réseaux de neurones, ont récemment émergé comme une alternative efficace [4].

3 Reconstruction de phase

3.1 Formulation du problème

Soit $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ une matrice dont les lignes sont notées $\{\boldsymbol{a}_j\}_{1\le j\le m}$. Nous définissons la formulation directe $\boldsymbol{y} = h(\boldsymbol{x})$ comme suit :

$$y_j = |\boldsymbol{a}_j^\top \boldsymbol{x}|^2, \quad \forall j \in \{1,\dots,m\},$$

ou sous forme matricielle :

$$\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}|^2,$$

pour tout $\boldsymbol{x} \in \mathcal{X}$. Le modèle $\mathcal{X} \in \mathbb{C}^n$ correspond à l'ensemble des signaux que nous cherchons à reconstruire, généralement bien plus petit que $\mathbb{C}^n$. L'opérateur $\boldsymbol{A}$ représente la composante linéaire de la physique d'acquisition et est souvent une matrice aléatoire ou une matrice de transformation de Fourier [17]. Le rapport $\alpha = m/n$ représente le taux d'échantillonnage. Un α petit signifie une compression élevée des mesures, rendant la reconstruction plus difficile, mais l'acquisition moins couteuse.

3.2 Hypothèse d'invariance

Dans notre travail, nous considérons un modèle $\mathcal{X}$ invariant à certaines transformations, comme les translations ou les rotations. Mathématiquement, pour un groupe de transformations G, cette invariance s'écrit :

$$\boldsymbol{T}_g\mathcal{X} = \mathcal{X} \quad \text{pour tout } g \in G, \tag{2}$$

où $\boldsymbol{T}_g$ est l'opérateur de transformation linéaire associé à $g \in G$. Cette hypothèse est naturelle pour différentes transformations et nombreux signaux : l'ensemble des images naturelles est invariant aux rotations et translations. La raison de considérer cette hypothèse est que l'invariance aux transformations $\boldsymbol{T}_g$ nous donne accès aux ensembles $\mathcal{Y}_g = h(\boldsymbol{T}_g\mathcal{X})$ pour tout $g \in G$. En effet, nous pouvons voir $\mathcal{Y}$ comme un ensemble de mesures associées au modèle $\mathcal{X}$ pour différents opérateurs directs $h_g(\cdot) := h(\boldsymbol{T}_g\cdot)$ grâce au fait que

$$\mathcal{Y}_g = |\boldsymbol{A}\boldsymbol{T}_g\mathcal{X}|^2 = |\boldsymbol{A}\mathcal{X}|^2 \quad \text{si} \quad \boldsymbol{T}_g\mathcal{X} = \mathcal{X}.$$

Cela peut fournir des informations supplémentaires au-delà du noyau de $\boldsymbol{A}$, nous permettant d'apprendre l'ensemble des signaux $\mathcal{X}$ à partir des ensembles de mesures $\mathcal{Y}_g$. Toutefois, pour que cette approche soit efficace, la transformation $\boldsymbol{T}_g$ doit modifier le noyau de $\boldsymbol{A}$, c'est-à-dire que $\ker(\boldsymbol{A}) \neq \ker(\boldsymbol{A}\boldsymbol{T}_g)$ pour tout $g \in G$. Une condition nécessaire est que l'opérateur $\boldsymbol{A}$ ne commute pas avec les transformations du groupe G [19].

3.3 Metrique

La reconstruction du signal $\boldsymbol{x} \in \mathbb{C}^n$ à partir des observations $\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}|^2$ n'est pas unique. En effet, toute transformation de la forme $\boldsymbol{x} \mapsto e^{i\varphi}\boldsymbol{x}$, où φ est un réel, préserve les mesures d'intensité :

$$|\boldsymbol{A}(e^{i\varphi}\boldsymbol{x})|^2 = |\boldsymbol{A}\boldsymbol{x}|^2.$$

Par conséquent, toute méthode de reconstruction ne peut estimer $\boldsymbol{x}$ qu'à une phase globale près. Pour évaluer la qualité de reconstruction, nous utilisons donc la similarité cosinus, notée $\mathrm{CS}(\cdot,\cdot)$, qui est insensible à ces variations de phase. Cette métrique est définie par :

$$\mathrm{CS}(\boldsymbol{x},\hat{\boldsymbol{x}}) = \frac{|\boldsymbol{x}^*\hat{\boldsymbol{x}}|}{\|\boldsymbol{x}\|_2\|\hat{\boldsymbol{x}}\|_2}, \tag{3}$$

où $\boldsymbol{x}^*$ désigne l'adjoint de $\boldsymbol{x}$. Une similarité cosinus de 1 indique indique que la reconstruction est parfaite, c'est-à-dire qu'il existe une phase (et éventuellement un facteur d'échelle) tel que

$$\hat{\mathbf{x}} = r\, e^{i\varphi}\mathbf{x}, \quad \text{avec } r \in \mathbb{R} \text{ et } \varphi \in \mathbb{R}.$$

Au contraire, une valeur proche de 0 indique une reconstruction aléatoire. Le facteur d'échelle r n'est pas problématique car il peut facilement être connue en remarquant que

$$|\boldsymbol{A}\hat{\boldsymbol{x}}|^2 = |\boldsymbol{A}re^{i\varphi}\boldsymbol{x}|^2 = r^2|\boldsymbol{A}\boldsymbol{x}|^2 = r^2\boldsymbol{y}.$$

Dans la suite, pour parler d'égalité à une phase globale près, nous utiliserons le symbole $\overset{\varphi}{=}$. Ainsi, $\hat{\boldsymbol{x}} \overset{\varphi}{=} \boldsymbol{x}$ signifie

$$\exists \varphi \in \mathbb{R}, \text{ tel que } \hat{\boldsymbol{x}} = e^{i\varphi}\boldsymbol{x}.$$

4 Apprentissage auto-supervisé

Fonction de perte de consistence des mesures (CM) - Une méthode de reconstruction efficace doit respecter la consistence des mesures, c'est-à-dire que pour toute mesure $\boldsymbol{y}$, nous avons

$$h(f_{\boldsymbol{\theta}}(\boldsymbol{y})) = \boldsymbol{y}.$$

Garantir la consistence de mesure est essentiel, car cela assure que les estimations restent fidèles aux observations. Cependant, cela n'implique pas nécessairement l'unicité de la solution, car plusieurs solutions peuvent satisfaire les mêmes mesures. Une fonction de perte naturelle pour imposer cette consistence est :

$$\mathcal{L}_{\text{CM}}(\boldsymbol{\theta}) = \sum_{i \in I} \| \boldsymbol{y}_i - h\left(f_{\boldsymbol{\theta}}(\boldsymbol{y}_i)\right) \|^2. \tag{4}$$

Cette fonction de perte est aussi connue dans le domaine de reconstruction de phase comme la perte d'intensité. Nous pouvons également considérer une variante de cette perte, la perte d'amplitude :

$$\mathcal{L}_{\text{A}}(\boldsymbol{\theta}) = \sum_{i \in I} \left\| \sqrt{\boldsymbol{y}_i} - \sqrt{h\left(f_{\boldsymbol{\theta}}(\boldsymbol{y}_i)\right)} \right\|^2. \tag{5}$$

Bien que ces deux fonctions aient le même minimum global pour un f_θ suffisamment flexible, des observations empiriques suggèrent qu'il est parfois préférable d'utiliser la perte d'amplitude, particulièrement dans le cas de signaux bruités [21].

Fonction de perte d'équivariance - Pour apprendre efficacement la phase, nous utilisons l'hypothèse d'invariance du modèle définie dans la section 3.2, et allons ainsi au-delà des limitations imposées par la consistence des mesures seule. Pour exploiter sur le réseau de reconstruction cette propriété d'invariance, nous remarquons que, puisque $\boldsymbol{T}_g\boldsymbol{x} \in \mathcal{X}$, la fonction f doit être capable de reconstruire à la fois $\boldsymbol{x}$ et $\boldsymbol{T}_g\boldsymbol{x}$ pour tout $\boldsymbol{x} \in \mathcal{X}$ et $g \in G$, c'est-à-dire :

$$f\left(h(\boldsymbol{T}_g\boldsymbol{x})\right) \overset{\varphi}{=} \boldsymbol{T}_g\boldsymbol{x} \quad \text{et} \quad f(h(\boldsymbol{x})) \overset{\varphi}{=} \boldsymbol{x}. \tag{6}$$

Ainsi, nous voulons avoir :

$$f(h(\boldsymbol{T}_g\boldsymbol{x})) \overset{\varphi}{=} \boldsymbol{T}_g f(h(\boldsymbol{x})) \text{ pour tout } \boldsymbol{x} \in \mathcal{X}, \text{ et } g \in G. \tag{7}$$

La composition de f avec h doit alors être équivariante par rapport au groupe G. Pour garantir cela, nous proposons une fonction de perte qui impose cette propriété d'équivariance :

$$\mathcal{L}_{\text{EI}}(\boldsymbol{\theta}) = -\sum_{i \in I} \sum_{g \in G} \text{CS}\Big(\boldsymbol{T}_g f_{\boldsymbol{\theta}}\left(\boldsymbol{y}_i\right), f_{\boldsymbol{\theta}}\big(h(\boldsymbol{T}_g f_{\boldsymbol{\theta}}\left(\boldsymbol{y}_i\right))\big)\Big). \tag{8}$$

L'utilisation de la CS dans la fonction de perte ne garantit pas des facteurs d'échelle r corrects entre $\boldsymbol{T}_g f_{\boldsymbol{\theta}}\left(\boldsymbol{y}_i\right)$ et $f_{\boldsymbol{\theta}}\big(h(\boldsymbol{T}_g f_{\boldsymbol{\theta}}\left(\boldsymbol{y}_i\right))\big)$; c'est pourquoi il est nécessaire de l'associer à une fonction de perte de consistence des mesures.

La fonction de perte finale est donc une combinaison des pertes de consistence des mesures et d'équivariance :

$$\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_{\text{A}}(\boldsymbol{\theta}) + \lambda \mathcal{L}_{\text{E}}(\boldsymbol{\theta}). \tag{9}$$

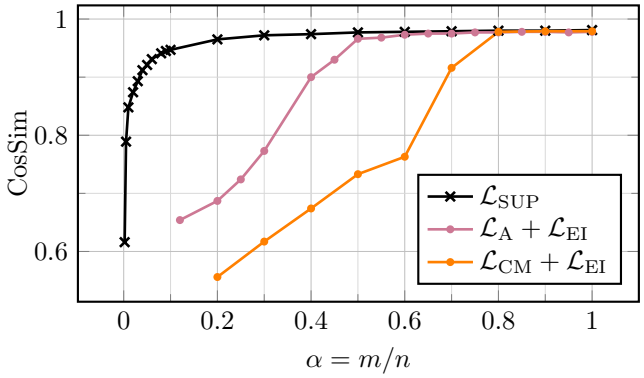

FIGURE 1 : Cosinus similarité en fonction du taux d'échantillonnage α pour la méthode supervisée et les méthodes auto-supervisées.

Le paramètre λ est un hyperparamètre positif qui contrôle le compromis entre les deux pertes. La fonction de reconstruction peut être basée sur n'importe quelle architecture de réseau de neurones. Pour nos expériences, nous avons choisi d'utiliser un U-Net [13].

5 Expériences

Dans cette section, nous décrivons la méthodologie adoptée pour tester et comparer notre approche. L'ensemble des expériences ont été réalisées à l'aide de la bibliothèque open-source DeepInv [18] sur le cluster de machine du Centre Blaise Pascal à l'ENS de Lyon [12]. Tout d'abord, nous construisons un jeu de données en utilisant le dataset MNIST. À partir des images initiales $\boldsymbol{x}_0 \in \mathbb{R}^n$, où $n = 28 \times 28$, nous générons unes images complexes $\boldsymbol{x} = e^{i\boldsymbol{x}_0} \in \mathbb{C}^n$ puis nous calculons les mesures $\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}|^2$. La matrice complexe $\boldsymbol{A}$ est une matrice aléatoire de dimensions $m \times n$, dont chaque élément $a_{k,l}$ suit une distribution normale centrée, donnée par :

$$a_{k,l} \sim \mathcal{N}(0, 1/2m) + i\mathcal{N}(0, 1/2m).$$

Ces matrices servent à modéliser des milieux diffusants, utile en imagerie optique [9]. Nous évaluons notre méthode pour plusieurs valeurs du ratio d'échantillonnage α pour le groupe de transformations des translations, c'est à dire $\boldsymbol{T}_g$ agit sur $\boldsymbol{x}$ comme un shift des pixels. Les méthodes classiques de reconstruction de phase, vus en section 2 comme PhaseLift et PhaseMax nécessitent un taux d'échantillonnage élevé, $\alpha > 2$. Cependant elles ne sont pas restreintes sur un modèle $\mathcal{X}$ et peuvent reconstruire sur l'espace $\mathbb{C}^n$ tout entier. Une comparaison avec ces méthodes ne serait donc pas pertinente ; c'est pourquoi nous nous limitons à une comparaison avec la méthode supervisée. Nous modifions la fonction de perte 1 pour l'adaptée à la reconstruction de phase :

$$\mathcal{L}_{\text{SUP}}(\boldsymbol{\theta}) = -\sum_{i \in I} \text{CS}\big(\boldsymbol{x}_i, f_{\boldsymbol{\theta}}(\boldsymbol{y}_i)\big),$$

Pour $k \in \{1, \dots, m\}$ et $l \in \{1, \dots, n\}$. Les résultats sont représenté dans la figure 1. Enfin, nous corrigeons la phase globale de $f(\boldsymbol{y})$ pour l'aligner avec $\boldsymbol{x}$ afin d'assurer une représentation visuelle cohérente. Un exemple de reconstruction pour différente valeur de α est donné dans la figure 2. On remarque que pour des valeurs de α supérieures à 0.5, la méthode auto-supervisée donne des résultats comparables à la méthode

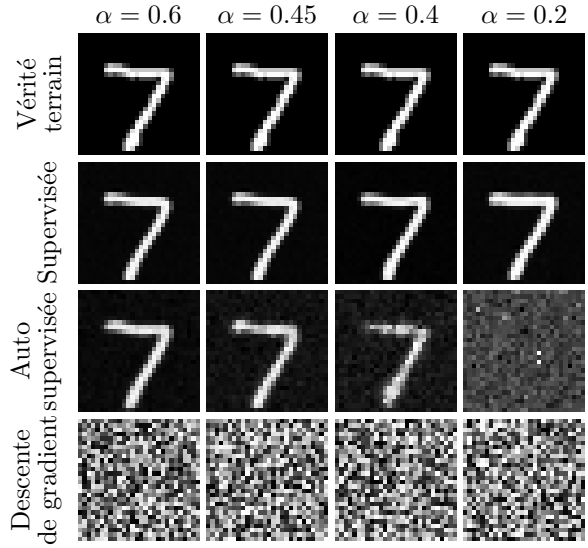

FIGURE 2 : Images reconstruites pour les méthodes supervisée, auto-supervisée et pour une résolution par descente de gradient.

supervisée. En revanche, cette dernière performe nettement mieux pour des valeurs de α faible ($\alpha < 0.3$). On observe également que méthode entrainé avec la perte d'intensité 4 en tant que CM permet une compression moindre que celle entrainé avec la perte d'amplitude 5.

Paramètre d'apprentissage - Nous avons utilisé un réseau U-Net comportant 4 niveaux de descente et 4 niveaux de montée. Le paramètre λ de la fonction de perte 9 est fixé à 1 dans toutes les expériences. Pour des raisons computationnelles lors du calcul de la fonction de perte EI 8, nous sommons sur deux translations par image, plutôt que sur l'ensemble des transformations du groupe G. L'optimisation a été effectuée à l'aide de l'algorithme Adam avec un taux d'apprentissage de $5e^{-5}$. Le modèle a été entraîné sur 15 epochs avec une taille de batch de 5, sur un tier des images de MNIST.

6 Conclusion

Nous avons présenté une approche auto-supervisée qui tire parti de l'invariance pour résoudre le problème de la reconstruction de phase. Nos expériences préliminaires montrent que notre méthode atteint des performances comparables à celles de l'approche supervisée dans le cadre de l'invariance par translation. Ce travail met en évidence de nouvelles possibilités de solutions basées sur l'apprentissage pour les problèmes inverses non linéaires. D'autres travaux peuvent être réalisés pour étendre cette approche à divers groupes de transformations ou d'autre type de signaux, tout en proposant un cadre théorique assurant la faisabilité de la reconstruction.

7 Remerciements

Ce travail a été financé par le projet ANR-20-CE40-0001-01. La recherche de Laurent Jacques est en partie financée par le FRS-FNRS (QuadSense, T.0160.24).

Références

[1] Chinmay BELTHANGADY et Loic A ROYER : Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nature methods*, 16(12):1215–1225, 2019.

[2] Emmanuel J. CANDÈS, Yonina C. ELDAR, Thomas STROHMER et Vladislav VORONINSKI : Phase Retrieval via Matrix Completion. *SIAM Review*, 57(2):225–251, janvier 2015.

[3] Dongdong CHEN, Julián TACHELLA et Mike E. DAVIES : Equivariant Imaging : Learning Beyond the Range Space. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4379–4388, août 2021.

[4] Jonathan DONG, Lorenzo VALZANIA, Antoine MAILLARD, Thanh-an PHAM, Sylvain GIGAN et Michael UNSER : Phase Retrieval : From Computational Imaging to Machine Learning. *IEEE Signal Processing Magazine*, 40(1):45–57, janvier 2023. arXiv :2204.03554 [physics].

[5] R. W. GERCHBERG : A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–246, 1972.

[6] Tom GOLDSTEIN et Christoph STUDER : PhaseMax : Convex Phase Retrieval via Basis Pursuit. *IEEE Transactions on Information Theory*, 64(4):2675–2689, avril 2018.

[7] Alexander KRULL, Tim-Oliver BUCHHOLZ et Florian JUG : Noise2void-learning denoising from single noisy images. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2129–2137, 2019.

[8] Jaakko LEHTINEN, Jacob MUNKBERG, Jon HASSELGREN, Samuli LAINE, Tero KARRAS, Miika AITTALA et Timo AILA : Noise2Noise : Learning image restoration without clean data. *International Conference on Machine Learning*, 2018.

[9] Antoine LIUTKUS, David MARTINA, Sébastien POPOFF, Gilles CHARDON, Ori KATZ, Geoffroy LEROSEY, Sylvain GIGAN, Laurent DAUDET et Igor CARRON : Imaging with nature : Compressive imaging using a multiply scattering medium. *Scientific reports*, 4(1):5552, 2014.

[10] Michael LUSTIG, David DONOHO et John M PAULY : Sparse MRI : The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine : An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007. Publisher : Wiley Online Library.

[11] Henrik OHLSSON et Yonina C ELDAR : On conditions for uniqueness in sparse phase retrieval. *In 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1841–1845. IEEE, 2014.

[12] Emmanuel QUEMENER et Marianne CORVELLEC : Sidus—the solution for extreme deduplication of an operating system. *Linux Journal*, 2013(235):3, 2013.

[13] Olaf RONNEBERGER, Philipp FISCHER et Thomas BROX : U-net : Convolutional networks for biomedical image segmentation. *In MICCAI 2015 : 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.

[14] Leonid I RUDIN, Stanley OSHER et Emad FATEMI : Nonlinear total variation based noise removal algorithms. *Physica D : nonlinear phenomena*, 60(1-4):259–268, 1992. Publisher : Elsevier.

[15] Jérémy SCANVIC, Mike DAVIES, Patrice ABRY et Julián TACHELLA : Self-Supervised Learning for Image Super-Resolution and Deblurring, mars 2024. arXiv :2312.11232 [cs, eess].

[16] Victor SECHAUD, Laurent JACQUES, Patrice ABRY et Julián TACHELLA : Equivariance-based self-supervised learning for audio signal recovery from clipped measurements. *In 2024 32nd European Signal Processing Conference (EUSIPCO)*, pages 852–856. IEEE, 2024.

[17] Yoav SHECHTMAN, Yonina C. ELDAR, Oren COHEN, Henry Nicholas CHAPMAN, Jianwei MIAO et Mordechai SEGEV : Phase Retrieval with Application to Optical Imaging : A contemporary overview. *IEEE Signal Processing Magazine*, 32(3):87–109, mai 2015.

[18] J TACHELLA, D CHEN, S HURAULT, M TERRIS et A WANG : Deepinverse : A deep learning framework for inverse problems in imaging. *URL : https ://deepinv. github. io/deepinv*, 2023.

[19] Julián TACHELLA, Dongdong CHEN et Mike DAVIES : Sensing theorems for unsupervised learning in linear inverse problems. *Journal of Machine Learning Research*, 24(39):1–45, 2023.

[20] Julián TACHELLA et Laurent JACQUES : Learning to reconstruct signals from binary measurements alone. *Transactions on Machine Learning Research*, 2023.

[21] Li-Hao YEH, Jonathan DONG, Jingshan ZHONG, Lei TIAN, Michael CHEN, Gongguo TANG, Mahdi SOLTANOLKOTABI et Laura WALLER : Experimental robustness of fourier ptychography phase retrieval algorithms. *Optics express*, 23(26):33214–33240, 2015.