

欠采样相位恢复与图像先验

Stanislas Ducotterd, Zhiyuan Hu, Michael Unser, Jonathan Dong

Biomedical Imaging Group, École polytechnique fédérale de Lausanne, Lausanne, Switzerland

ABSTRACT

相位检索旨在从仅幅度的测量中恢复复信号，这是一个具有挑战性的非线性逆问题。当前的理论和算法通常忽略信号先验。相比之下，我们在结构化随机傅里叶测量下的严重欠采样背景下评估了各种图像先验。我们的结果显示这些先验显著提高了重建效果，甚至在弱恢复阈值以下也能实现准确重建。

Index Terms— 反问题，相位检索，图像先验，优化，正则化

1. 介绍

在许多科学和工程领域，人们需要从间接测量中恢复感兴趣的对象。在这个背景下，相位检索是一个特别具有挑战性的非线性问题，旨在从仅幅度的测量 [1] 重建复信号 $\mathbf{x}$ 。

这种表述出现在各种应用中，包括 X 射线晶体学、天文成像 [2]、ptychography [3, 4] 和傅里叶 ptychography [5, 6]。

随机相位检索考虑具有独立随机条目的测量矩阵，这使得精确恢复和算法性能的研究在理论上可行 [7, 8]。最近的工作保持了这些保证，同时减少了计算成本。它们通过用结构化变换结合随机对角矩阵替换密集的随机性来实现这一点 [9]。然而，这一系列工作很大程度上忽略了关于自然信号的先验知识，这可以显著提高恢复效果。

图像先验从基于稀疏性的显式正则化器（如总变差 TV [10]）到嵌入强大去噪器中的隐式先验（例如

DRUNet [11]），并通过即插即用 (PnP) 框架集成到迭代方案中。近期的研究试图通过学习显式的正则化器 [12, 13, 14] 来结合前者的理论性质和后者的经验性能。随机测量模型为在相位检索的背景下评估这些策略提供了一个特征明确的环境。

超出本研究范围的是生成框架，包括对抗性、深度图像先验和扩散模型 [15, 16, 17, 18]。我们还参考了 [19, 20] 在标准过采样相位检索框架中使用隐式和显式正则化器的应用。

在这项工作中，我们评估了 TV、多变量领域专家 (MFoE) 和 DRUNet，从而涵盖了广泛的一系列先验。我们展示了引入先验知识带来了显著的改进，并且在极端欠采样的情况下能够实现准确的恢复。特别是，MFoE 超过了 TV 和 DRUNet，这表明原理上的显式正则化器可能特别适合处理非线性逆问题如相位检索。代码可在 Github¹ 上获取。

2. 问题定义

我们的目标是从一个通过非线性正向模型

$$\mathbf{y} = |\mathbf{SFD}\mathbf{x}| + \mathbf{n}, \quad (1)$$

获得的 m 维实测量向量 $\mathbf{y} \in \mathbb{R}^m$ 中恢复一个 n 维复信号 $\mathbf{x} \in \mathbb{C}^n$ 其中

- 已知随机对角矩阵 $\mathbf{D} \in \{\pm 1\}^{n \times n}$ 的元素独立同分布地从 $\{-1, +1\}$ 中以相等的概率抽取；
- 离散傅里叶变换矩阵是 $\mathbf{F} \in \mathbb{C}^{n \times n}$ ；
- 子采样掩码 $\mathbf{S} \in \{0, 1\}^{n \times n}$ 是对角的，其中每个对角线元素独立同分布地来自参数为 α 的伯努利分布。

Stanislas Ducotterd acknowledges funding from the European Research Council (Grant Agreement - Project No 101020573 FunLearn). Zhiyuan Hu and Jonathan Dong acknowledge funding from the Swiss National Science Foundation (Grant PZ00P2_216211).

¹<https://github.com/StanislasDucotterd/欠采样相位检索>

- 高斯噪声的噪声水平是 σ_n , $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ 。

算子 $|\cdot|$ 表示逐元素模。形式上, $\mathbf{y}$ 在 $\mathbb{R}^n$ 中有 m 个非零项, 但我们通过丢弃零分量将其识别为 $\mathbb{R}^m$ 。算子 $\mathbf{D}$ 和 $\mathbf{F}$ 准确地模拟了光学系统 [1, 9] 中扩束器和透镜的组合。

一个关键的超参数是采样比率 α 。较小的 α 值对应于严重的欠采样, 使得恢复问题变得越来越难。在这项工作中, 我们研究了噪声 less 测量中的采样比率 $\alpha \in \{k/10\}_{k=1}^{10}$ 以及没有欠采样的含噪声测量 $\sigma_n \in \{0.001, 0.01, 0.1, 1\}$ 。请注意, 在弱恢复阈值 $\alpha_{\text{WR}} = 1$ 以下, 如果没有先验信息 [7], 则无法恢复关于 $\mathbf{x}$ 的任何信息。

结构化随机模型通常对应于任意深度的傅里叶矩阵和对角矩阵级联。先前的工作表明, 至少需要两层 (即 $\mathbf{FD}_2\mathbf{FD}_1$) 来模拟密集模型而无需先验知识 [9]。尽管如此, 我们发现即使是在 (1) 中的更简单的单层设置也显著受益于正则化。

3. 方法

3.1. 数据拟合项及其近似算子

为了处理相位检索问题, 我们采用交替减少数据保真项

$$f(\mathbf{x}) = \frac{1}{2} \|\mathbf{SFD}\mathbf{x} - \mathbf{y}\|_2^2 \quad (2)$$

并强制解决方案先验信息的迭代方案。

我们考虑的测量算子具有一般形式 $\mathbf{SU}$, 其中 $\mathbf{S}$ 是一个二进制采样掩码 (第 2 节), $\mathbf{U}$ 是幺正的。在这种设置下, 可以使用 f 的近端算子来制定迭代方法, 该近端算子定义为 $\text{prox}_{\lambda f}(\mathbf{x}) = \arg \min_{\mathbf{z} \in \mathbb{C}^n} \frac{1}{2} \|\mathbf{z} - \mathbf{x}\|_2^2 + \lambda f(\mathbf{z})$ 。

定理 1. 数据拟合项 $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{SU}\mathbf{x} - \mathbf{y}\|_2^2$ 的近端算子可以表示为

$$\text{prox}_{\lambda f}(\mathbf{x}) = \mathbf{U}^H \mathbf{S} \left(\frac{1}{1+\lambda} \mathbf{U}\mathbf{x} + \frac{\lambda}{1+\lambda} \mathbf{y} e^{j \arg(\mathbf{U}\mathbf{x})} \right) + \mathbf{U}^H (\mathbf{I} - \mathbf{S}) \mathbf{U}\mathbf{x}. \quad (3)$$

Proof. 对于标量问题,

$$\text{prox}_{\frac{\lambda}{2}(|\cdot| - \mathbf{y})^2}(\mathbf{x}) = \arg \min_{z \in \mathbb{C}} \frac{1}{2} |z - \mathbf{x}|^2 + \frac{\lambda}{2} (|z| - \mathbf{y})^2 \quad (4)$$

$$= \arg \min_{z \in \mathbb{C}} |z|^2 + |\mathbf{x}|^2 + \lambda (|z| - \mathbf{y})^2 - 2|z||\mathbf{x}| \cos(\arg z - \arg \mathbf{x}) \quad (5)$$

以及唯一依赖于 $\arg z$ 的项通过设置 $\arg z = \arg \mathbf{x}$ 来最小化。代入相位后, 可以计算最优幅度, 得到

$$\text{prox}_{\frac{\lambda}{2}(|\cdot| - \mathbf{y})^2}(\mathbf{x}) = \frac{1}{1+\lambda} \mathbf{x} + \frac{\lambda}{1+\lambda} \mathbf{y} e^{j \arg \mathbf{x}}. \quad (6)$$

多维情况下使用二值掩码给出

$$\arg \min_{\mathbf{z} \in \mathbb{C}^n} \frac{1}{2} \|\mathbf{z} - \mathbf{x}\|_2^2 + \frac{\lambda}{2} \|\mathbf{S}\mathbf{z} - \mathbf{y}\|_2^2 \quad (7)$$

$$= \mathbf{S} \left(\frac{1}{1+\lambda} \mathbf{x} + \frac{\lambda}{1+\lambda} \mathbf{y} e^{j \arg \mathbf{x}} \right) + (\mathbf{I} - \mathbf{S}) \mathbf{x}. \quad (8)$$

这源于目标函数的可分离性; 对于含有 $S_{kk} = 1$ 的项使用标量近似算子 (4), 而对于 $S_{kk} = 0$ 的项则使用恒等映射。最后, 通过 [21],

$$\text{prox}_{g(\mathbf{U} \cdot)}(\mathbf{x}) = \mathbf{U} \text{operatorname{prox}_g}(\mathbf{U}\mathbf{x}), \quad (9)$$

得出 (3)。 $\square$

3.2. 变分方法

我们最小化数据保真项与 TV 或 MFoE 正则化的和。TV 是一种已建立的手工设计正则化技术。它促进了原始图像梯度的稀疏性。相比之下, MFoE 学习了噪声相关的正则化器 $R_\sigma: \mathbb{R}^n \rightarrow \mathbb{R}_+$, 使得对于任意图像 $\mathbf{x} \in \mathbb{R}^n$ 和 $\mathbf{y} \sim \mathcal{N}(\mathbf{x}, \sigma^2 \mathbf{I})$, 有 $\text{prox}_{R_\sigma}(\mathbf{y}) \approx \mathbf{x}$ 。参数 σ 使正则化器适应噪声水平。

为了将 TV 和 MFoE 扩展到复杂数据, 我们分别对幅度和归一化相位进行正则化, 如

$$\mathcal{R}_\sigma(\mathbf{x}) = R_\sigma(|\mathbf{x}|) + R_\sigma\left(\frac{\arg(\mathbf{x})}{2\pi}\right). \quad (10)$$

所示。为了获得重建, 我们求解

$$\arg \min_{\mathbf{x} \in \mathbb{C}^n} \frac{1}{2\sigma_n} \|\mathbf{SFD}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \mathcal{R}_\sigma(\mathbf{x}), \quad (11)$$

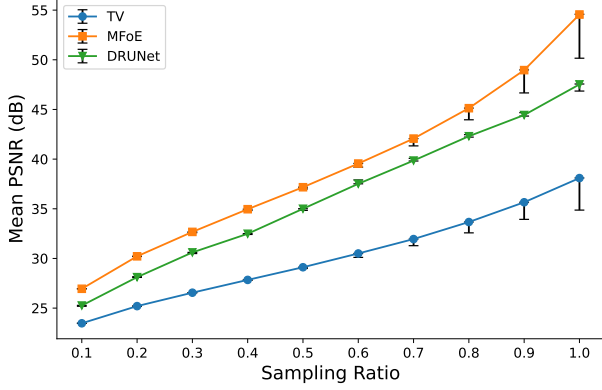


Fig. 1. 峰值信噪比 (dB) 对于每个模型和采样比率在无噪声情况下的表现。误差条显示了在 BSD68 上的三次随机实现中的最佳、中位数和最差性能。

Algorithm 1 加速近端梯度下降法带重启机制

输入: 初始化 $\mathbf{x}_0 \in \mathbb{C}^n$, 步长 γ , 测量噪声水平 σ_n , 噪声水平 σ , 以及容差 $\varepsilon > 0$

- 1: 设置 $t_0 = 1, k = 0, r_0 = \infty, s_0 = \infty, \mathbf{z}_0 = \mathbf{x}_0$
- 2: while $r_k > \varepsilon$ do
- 3: $\mathbf{x}_{k+1} = \text{prox}_{\gamma f / \sigma_n}(\mathbf{z}_k - \gamma \nabla R_\sigma(\mathbf{z}_k))$
- 4: $s_{k+1} = R_\sigma(\mathbf{z}_k)$
- 5: $t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$
- 6: $\mathbf{z}_{k+1} = \mathbf{x}_{k+1} + \frac{t_k - 1}{t_{k+1}}(\mathbf{x}_{k+1} - \mathbf{x}_k)$
- 7: if $s_{k+1} > s_k$ then
- 8: $\mathbf{z}_{k+1} = \mathbf{x}_{k+1}$
- 9: $t_{k+1} = 1$
- 10: end if
- 11: $r_{k+1} = \|\mathbf{x}_{k+1} - \mathbf{x}_k\| / \|\mathbf{x}_{k+1}\|$
- 12: $k \leftarrow k + 1$
- 13: end while

输出: 近似解 $\mathbf{x}_{k+1}$

通过带有重启的加速近端梯度下降 (算法 1), 使用 $\gamma = 1/\text{Lip}(\nabla R_\sigma), \varepsilon = 10^{-5}, \lambda = 10^3$ 用于 TV, 并使用 $\lambda = 10^4$ 用于 MFoE。

对于 TV, 它不依赖于 σ , 我们从复高斯噪声 $\mathbf{x}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ 中求解一次 (11)。对于 MFoE, 我们通过连续方案利用噪声参数: 我们将问题求解三次, 首先使用 $\sigma = 1$, 然后是 $\sigma = 1/4$, 最后是 $\sigma = 1/16$ 。第一次

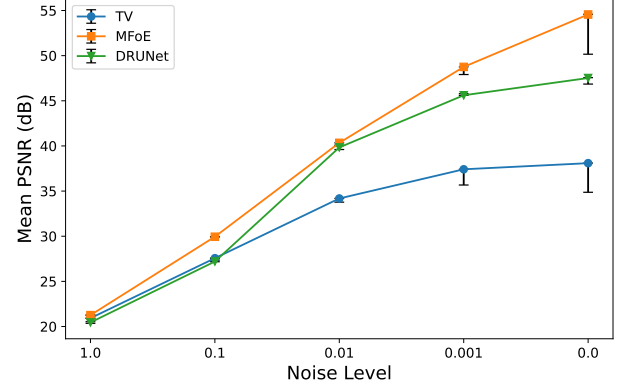


Fig. 2. 各模型和噪声水平在无亚采样情况下的 PSNR (dB)。误差条显示了在 BSD68 上的三次随机实现中的最佳、中位数和最差性能。

Algorithm 2 即插即用与 DRUNet

输入: 初始化 $\mathbf{x}_0 \in \mathbb{C}^n$, 步骤数量 K , 正则化参数 λ , 测量噪声水平 σ_n , 噪声计划 σ_k

- 1: 设置 $k = 0$
- 2: while $k < K$ do
- 3: $\beta = \lambda \sigma_n^2 / \sigma_k^2$
- 4: $\mathbf{x}_{k+1} = \text{prox}_{f/\beta}(D_{\sigma_k}(\mathbf{x}_k))$
- 5: $k \leftarrow k + 1$
- 6: end while

输出: 近似解 $\mathbf{x}_{k+1}$

运行初始化为 $\mathbf{x}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, 后续运行则采用前一次的解决方案。这逐步从粗到细地改进了重建。此外, 在第一次 MFoE 运行中, 我们降低数据拟合项的权重以避免较差的局部最小值——每当 $\alpha > 0.25$ 时, 我们将 $\mathbf{S}$ 子采样为其条目数量的四分之一。

3.3. 即插即用方法

PnP 方法用现成的去噪器替换了迭代重建方案中的显式正则化项。我们遵循 [11] 并应用 Algorithm 2, 使用 DRUNet_{D_σ} 。为了处理复杂数据, 我们将信号分为幅度和相位两部分, 对两者都应用 D_σ , 然后重新组合。

算法初始化为 $\mathbf{x}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, 并使用 $\lambda = 0.23$, 如 [11] 所示。噪声调度从 $\sigma_0 = 1$ 开始并几何递减到

$\sigma_K = 1/(1000\alpha)$, 其中 α 是采样比率。与原始工作使用 $K = 8, 24$, 或 40 不同, 我们设置 $K = 1000$ 。这种步骤数量的增加是实现良好重建所必需的事实可以归因于我们的前向模型的非线性性质, 这与在 [11] 中考虑的线性算子形成对比。我们没有对 DRUNet 应用 MFoE 的数据拟合降权策略, 因为它不会带来性能提升。

4. 实验

为了构建复值信号, 我们从实数灰度图像开始, 将像素值在 $[0, 1]$ 线性映射到 $[-\pi/2, \pi/2]$ 的相位上, 并分配单位幅度。为了评估对正向模型的鲁棒性, 我们使用了三个独立绘制的扩散器, 并报告了所有这些扩散器的表现情况。算法 1 和 2 的超参数是在 Set12 数据集 [22] 上调整的, 而性能则在 BSD68[23] 上报告。

无噪声采样情况和有噪声情况的结果分别显示在图 1 和图 2 中。视觉重建结果展示在图 3 中。用于评估重建性能的指标是广泛使用的峰值信噪比 (PSNR, 值越高越好)。有趣的是, 尽管 DRUNet 在去噪和多个线性逆问题中表现出最先进的性能, 但在所有设置下它都输给了 MFoE。这表明 PnP 框架可能在这种具有挑战性的非线性逆问题背景下表现不佳。

作为参考, 在所有场景中简单地应用梯度下降仅能获得大约 4dB 的 PSNR, 这突显了图像先验的重要性。即使是像总变差 (TV) 这样的简单先验也能显著提升性能。

5. 结论

我们研究了图像先验在噪声或极端欠采样随机相位检索中的益处。我们考虑了显式正则化器 (TV)、学习的显式正则化器 (MFoE) 和基于深度学习的隐式先验 (DRUNet)。我们的实验结果表明, 引入先验显著改善了重建效果, 即使在远低于弱恢复阈值的情况下也是如此。MFoE 一致优于 TV 和 DRUNet, 这突显了学习的显式正则化器对具有挑战性的非线性逆问题的有效性。

6. REFERENCES

- [1] J. Dong, L. Valzania, A. Maillard, T.-a. Pham, S. Gigan, and M. Unser, “Phase retrieval: From computational imaging to machine learning: A tutorial,” *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 45–57, 2023.
- [2] C. Fienup and J. Dainty, “Phase retrieval and image reconstruction for astronomy,” *Image Recovery: Theory and Application*, vol. 231, p. 275, 1987.
- [3] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, “Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens,” *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.
- [4] L. Valzania, J. Dong, and S. Gigan, “Accelerating ptychographic reconstructions using spectral initializations,” *Optics Letters*, vol. 46, no. 6, pp. 1357–1360, 2021.
- [5] G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution Fourier ptychographic microscopy,” *Nature Photonics*, vol. 7, no. 9, pp. 739–745, 2013.
- [6] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, “Experimental robustness of Fourier ptychography phase retrieval algorithms,” *Optics Express*, vol. 23, no. 26, pp. 33 214–33 240, 2015.
- [7] M. Mondelli and A. Montanari, “Fundamental limits of weak recovery with applications to phase retrieval,” in *Conference on Learning Theory*, 2018, pp. 1445–1450.
- [8] A. Maillard, B. Loureiro, F. Krzakala, and L. Zdeborová, “Phase retrieval in high dimensions: Statistical and computational phase tran-



Fig. 3. 重建结果来自 TV、MFoE 和 DRUNet。报告的指标是（余弦相似度，PSNR）。

- sitions,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 11 071–11 082, 2020.
- [9] Z. Hu, J. Tachella, M. Unser, and J. Dong, “Structured random model for fast and robust phase retrieval,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2025, pp. 1–5.
- [10] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
- [11] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, “Plug-and-play image restoration with deep denoiser prior,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6360–6376, Oct. 2022.
- [12] S. Hurault, A. Leclaire, and N. Papadakis, “Gradient step denoiser for convergent plug-and-play,” in *International Conference on Learning Representations*, 2022.
- [13] A. Goujon, S. Neumayer, P. Bohra, S. Ducotterd, and M. Unser, “A neural-network-based convex regularizer for inverse problems,” *IEEE Transactions on Computational Imaging*, vol. 9, pp. 781–795, 2023.
- [14] S. Ducotterd and M. Unser, “Multivariate fields of experts,” *arXiv:2508.06490*, 2025.
- [15] P. Hand, O. Leong, and V. Voroninski, “Phase retrieval under a generative prior,” *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [16] P. Bohra, T.-a. Pham, J. Dong, and M. Unser, “Bayesian inversion for nonlinear imaging models using deep generative priors,” *IEEE Transactions on Computational Imaging*, vol. 8, pp. 1237–1249, 2023.
- [17] E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, “Deep phase decoder: Self-calibrating phase microscopy with an untrained deep neural network,” *Optica*, vol. 7, no. 6, pp. 559–562, 2020.
- [18] B. Zhang, W. Chu, J. Berner, C. Meng, A. Anandkumar, and Y. Song, “Improving diffusion inverse problem solving with decoupled noise annealing,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 20 895–20 905.
- [19] Çağatay Işıl and F. S. Oktem, “Deep plug-and-play HIO approach for phase retrieval,” *Applied Optics*, vol. 64, no. 5, pp. 84–94, Feb 2025.
- [20] A. Denker, J. Hertrich, Z. Kereta, S. Cipiccia, E. Erin, and S. Arridge, “Plug-and-play half-quadratic splitting for ptychography,” in *Scale Space and Variational Methods in Computer Vision*, 2025, pp. 269–281.
- [21] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, 2011, pp. 185–212.
- [22] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [23] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *IEEE International Conference on Computer Vision*, vol. 2, 2001, pp. 416–423.