From linear to nonlinear unfolded Condat-Vũ algorithm for spectro-polarimetric hight-constrast image recovery

E. Chappon[1], N. Pustelnik[1], J. Tachella[1], L. Denneulin[2], A. Ferrari[3], M. Langlois[4]

[1]*Laboratoire de Physique, ENSL, CNRS UMR 5672, F-69342, Lyon, France*
[2]*Laboratoire de Recherche de l'EPITA, EPITA, 94270 Le Kremlin-Bicêtre, France*
[3]*Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Lab. J.-L. Lagrange, France*
[4]*Univ Lyon, Univ. Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon, UMR5574, 69230 Saint-Genis-Laval, France*

***Abstract*—Studying circumstellar environments is crucial for understanding exoplanets and stellar systems. Instruments like SPHERE can extract information about these environments by leveraging advanced image reconstruction methods, possibly based on deep learning. This work focuses on unfolded proximal neural networks based on Condat-Vũ iterations and proposes a new nonlinear formulation. To evaluate and compare the performance of the proposed reconstruction strategies, two datasets dedicated to circumstellar environments analysis in the context of high-contrast imagery have been created offering different level of complexity in the evaluation of the performance.**

***Index Terms*—High constrast image recovery, proximal algorithms, unfolded algorithms, deep learning, nonlinear algorithms.**

I. Introduction

Context – Studying circumstellar environments is essential for improving our comprehension of exoplanets and unraveling the complex mechanisms involved in the creation of stellar systems. Despite significant strides facilitated by advancements in instruments and image processing, allowing for greater resolution of these environments, the challenge of directly observing them persists. This difficulty arises from the extremely high contrast between these environments and their host stars, with the radiance of stars being much more intense, ranging from one to ten thousand times more intense than that of their surroundings.

Fortunately, certain polarimetric characteristics of the light emitted by the star and reflected in its environment enable the use of polarimetric imaging to disentangle the intertwined light of the two sources, providing a valuable tool for overcoming the limitations of direct observation.

Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE) and its instrument InfraRed Dual Imaging and Spectrograph (IRDIS) installed on the Very Large Telescope (VLT) of the European Southern Observatory (ESO), offer the possibility of acquiring high-resolution and high-contrast polarimetric images [1], [2]. During data acquisition, various factors such as geometrical transforms, instrumental blurring, or noise can lead to image degradation, requiring advanced image processing reconstruction techniques to extract the meaningful information.

State-of-the-art – The double difference and double ratio approaches [3], [4] are widely favored for deriving meaningful parameters. Nonetheless, these methods are sensitive to instrumental effects. To handle such a degradation, [5] has proposed to describe accurately the forward model which establishes the mapping between the parameters of interest and observations. Moreover, leveraging this forward model, a variational strategy, along with conventional optimization techniques, named *RHAPSODIE* for Reconstruction of High-contrAst Polarized SOurces and Deconvolution for cIrcumstellar Environments [5], has been proposed, offering a state-of-the-art method for reconstructing images of circumstellar environments.

From variational approaches to deep learning – For a long time, variational formulations and optimization techniques held the position of state-of-the-art solutions for addressing inverse problems in imaging, offering state-of-the-art reconstruction strategies in several field of astronomy including high-contrast imagery but also radio-interferometry [6]. However, the success of deep learning in recent years fundamentally changes the landscape of inverse problem solving. Consequently, in the extensive literature dedicated to this subject, three main classes of approaches can now be considered: iterative methods for solving variational problems [7] (as proposed in [5]), end-to-end ("black-box") neural networks [8], and hybrid proximal neural networks (encompassing strategies like plug-and-play and unfolded networks), capitalizing on the advantages of both worlds and providing performance, robustness, and interpretability [9]–[14].

Contributions – Our contribution aligns with the third category of methods, specifically focusing on unfolded strategies that offer a good trade-off between performance and interpretability. This approach presents an opportunity to enhance

This work is funded the Fondation Simone et Cino Del Duca - Institut de France. The authors thank the Centre Blaise Pascal of ENS Lyon for the computation facilities. The platform uses SIDUS, which was developed by Emmanuel Quemener.

the outcomes of RHAPSODIE by merging the principles of iterative methods with the progress due to deep learning. This work focuses on state-of-the-art unfolded primal-dual schemes already introduced in the literature [15], [16] and introduces a new unfolded scheme based on a nonlinear primal-dual strategy.

To highlight the advantages of unfolded architectures and the significance of the nonlinear unfolded strategy, a second contribution is the generation of a realistic dataset to simulate data encountered in SPHERE. This dataset serves as the basis for training and evaluating networks performance.

The study aims to present a spectrum of models ranging from standard techniques to advanced ones, illustrating varying levels of complexity.

Outline – This paper is structured as follows. In Section II, we delve into the explanation of the forward model encountered on SPHERE-IRDIS. Section III outlines generic primal-dual iterations, facilitating the transition from standard variational procedures to more powerfull unfolded approaches. This section also introduces a novel unfolded architecture based on a non-linear primal-dual scheme. The third section provides numerical experiments on a realistic database created specifically for this study, whose details are thoroughly described. It demonstrates the performance of a spectrum of unfolded primal-dual architectures, encompassing both standard techniques and advanced ones, showcasing varying levels of complexity. Finally the fifth section is devoted to conclusions and perspectives.

II. Acquisition model and pre-process data

Direct observation of the circumstellar environment is challenging because the light of the environment that we observe is a mixture of light from the star itself I^u_{star} and light that has been scattered by the star's accretion disk $I^u_{\text{disk}} + I^p$. This scattered light is polarized, unlike the light coming directly from the star. The objective of polarimetric imaging (as in SPHERE), when considering Jones formalism, is to reach the parameters (I^u, I^p, θ) formed by:

- $I^u = I^u_{\text{star}} + I^u_{\text{disk}}$: Intensity of the unpolarised light, which is the light that directly comes to us from the star combined with unpolarized disk light.
- I^p: Intensity of the polarized light, which is the light from the star scattered on the disk of dust. This is a useful parameter for studying the formation processes of a stellar environment. This parameter will be at the core of this study.
- θ: Linear polarisation angle associated with I^p, depending on the scattering angle of the star's light on the dust in the disk.

As described in [5], the observations are composed of K acquisitions with different modulations of the polarisation, induced by a half-wave plate's (HWP) angle which takes values $\alpha_\ell \in \{0°, 22.5°, 45°, 77.5°\}$ and a filtering through two analyzers, whose angles are $\psi_j = \{0°, 90°\}$. The number of acquisitions for one stellar environment can vary between $L = 32$ and $L = 512$. The forward model proposed in [5] reads, for every (j, ℓ),

$$z_{j,\ell} = T_{j,\ell} A \, I^{\text{det}}_{j,\ell} + \varepsilon_{j,\ell}, \tag{1}$$

which establishes the relation between the observation $z_{j,\ell}$ and

$$I^{\text{det}}_{j,\ell} = \frac{1}{2} I^u + I^p \cos^2(\theta - 2\alpha_\ell - \psi_j) \tag{2}$$

and A denotes a convolution operator associated with instrumental PSF, $T_{j,\ell}$ models geometrical transforms which map the space of parameters to the space of data, and $\varepsilon_{j,\ell}$ denotes a heteroscedastic Gaussian noise.

This model is not linear with respect to the Jones parameter θ. To overcome this difficulty, it is quite standard in astronomy, to use Stokes parameters (I, Q, U) , which can be expressed in terms of Jones parameters and vice versa.

$$\begin{cases} I^p = \sqrt{Q^2 + U^2} \\ \theta = (1/2) \arctan(U/Q) \mod \pi \\ I^u = I - \sqrt{Q^2 + U^2}. \end{cases}$$

The forward model becomes

$$z_{j,\ell} = \sum_{m=1}^{3} \nu_{j,\ell,m} \, T_{j,\ell} A \, S_m + \varepsilon_{j,\ell}, \tag{3}$$

with $S_1 = I$, $S_2 = Q$ and $S_3 = U$ and

$$\begin{cases} \nu_{j,\ell,1} = \frac{1}{2} \\ \nu_{j,\ell,2} = \frac{1}{2} \cos\big(4\,\alpha_\ell + 2\,\psi_j\big) \\ \nu_{j,\ell,3} = \frac{1}{2} \sin\big(4\,\alpha_\ell + 2\,\psi_j\big) \end{cases}$$

or equivalently. Among all these parameters $(I^u, I^p, \theta, I, Q, U)$ the principal parameter of interest is I^p. However, (Q, U) that corresponds to projection of I^p in two distinct planes are also of interest. So in the rest of the present work we will focus on the estimation of either I^p, Q, or U.

Double-difference and double-ratio are state-of-the-art methods allowing to prereconstruct the data and thus only focus on removing blur and residual noise leading to the degradation model of the form:

$$y = A\overline{x} + \epsilon \quad \text{with} \quad \overline{x} = \{I^p, Q, U\} \tag{4}$$

where ϵ denotes some additive noise whose covariance matrix is modified from (1) due to the prereconstruction.

III. From variational approaches to unfolded

The variational formulation for the resolution of (4) generally relies on a minimization problem of the form:

$$\widehat{x} \in \underset{x}{\text{Argmin}} \; f(Ax, y) + \lambda g(x) \tag{5}$$

where the first term corresponds to a "distance" between the observations y and the acquisition model while the second term g enables to incorporate prior on the solution. The regularization parameter $\lambda > 0$ controls the trade-off between adequacy to the data fidelity term and the prior.

Over the past three decades, there has been a progression in the complexity of priors and their corresponding inversion performance. This evolution began with the use of nonsmooth priors and has advanced to the adoption of implicit priors involving neural networks in more recent times [13]. These evolution are deeply intricated with advances in optimization and essentially whose iterations efficiency relies on having an explicit form of the proximal operator [17], [18].

In this work, we focus on primal-dual strategies allowing to handle more complex priors of the form $g = r(D\cdot)$ where D model some sparsifying transform and r inforcing sparsity prior (i.e. ℓ_1-norm).

The general Condat-Vũ iterations [15], [16] handled in this work adapted to quadratic data fidelity term (i.e. $f(A\cdot, z) = \frac{1}{2}\|A\cdot - z\|^2$) take the form:

for $k = 0, 1, \dots$

$$\left\lfloor \begin{array}{l} x^{[k+1]} = x^{[k]} - \gamma A^\top (Ax^{[k]} - y) - \gamma D_k^* u^{[k]}) \\ u^{[k+1]} = \mathrm{prox}_{\tau r^*}\left(u^{[k]} + \tau D_k\left(2x^{[k+1]} - x^{[k]}\right)\right), \end{array} \right.$$

where r^* denotes the Fenchel conjugate of r.

Several variations have been considered in the literature:

- **Standard Condat-Vũ** (CV) iterations are obtained when $D_k = D$ at each iteration. The convergence of the sequence $(x^{[k]})_{k\in\mathbb{N}}$ to $\widehat{x}$ is guaranteed when $\frac{1}{\gamma} - \tau\|D\|^2 \geq \frac{\|A\|^2}{2}$. It generally requires a large number of iterations [15], [16].
- **Truncated CV** – Truncated CV iterations with fixed $D_k = D$ mimicking a deep learning architecture combining linear steps and non-linear ones over the primal-dual variables [19]. This configuration is very close to the Standard CV except that the number of iterations is chosen to be small ($\sim$ 20), as with an early stopping strategy.
- **Unfolded CV** corresponds to truncated CV iterations with D_k varying from an iteration to another one [19].

If the Standard CV framework aligns with the conventional variational approach, wherein the dictionary is fixed, leading to a TV penalization when D models horizontal/vertical differences, the other two configurations are suited for the unfolded framework. These configurations relies on truncated iterations, but either the dictionary D or $(D_k)_{1\leq k\leq K}$ is learned.

IV. Unfolded non-linear Condat-Vũ

In this work we go beyond the work in [19] in order to unfold a formulation of the CV algorithm involving a nonlinear mapping D. The pioneer work on non-linear formulation of primal-dual algorithm traced back to Valkonen's work [20] dedicated to the resolution of

$$\min_x \max_u \frac{1}{2}\|Ax - z\|_2^2 + \langle D(x), u\rangle - r^*(u)$$

When D is linear, the problem fits (5) with $g = r(D\cdot)$. The non-linear formulation of CV algorithm reads:

for $k = 0, 1, \dots, K$

$$\left\lfloor \begin{array}{l} x^{[k+1]} = x^{[k]} - \gamma A^\top (Ax^{[k]} - y) - \gamma[\nabla D_k(x^{[k]})]^* u^{[k]}) \\ u^{[k+1]} = \mathrm{prox}_{\tau r^*}\left(u^{[k]} + \tau D_k\left(2x^{[k+1]} - x^{[k]}\right)\right). \end{array} \right.$$

In the specific case where $D_k(\cdot) \equiv D(\cdot)$, we obtain non-linear CV iterations which is the conterpart of Valkonen work in the context of Condat-Vũ iterations [15], [16] rather than Chambolle-Pock ones [21]. In our experiments, we will consider the case where the non-linear operator vary through the iterations considering $D_k(\cdot) = W_k \mathrm{Relu}(V_k\cdot)$.

V. Numerical results

In order to train, and then evaluate and compare the performance of these CV architectures for studying circumstellar environment observed from SPERE-IRDIS we have created two realistic datasets and then trained the unfolded schemes with Pytorch. All the strategies have been implemented using the DeepInverse open-source library [22].

A. Datasets

The considered datasets are generated from Debris DIsks Tools (DDIT) library [23] that produces synthetic images of $(I^u_{\text{disk}}, I^p, \theta)$. I^u_{star} has been obtained from real observational high-contrast coronagraphic data from the SPHERE instrument at the VLT/ESO [1]. (I, Q, U) are deduced from the resulting (I^u, I^p, θ).

In order to make the dataset as realistic as possible, the generated images are generated using several values for the semi-major axis of the disk, inclination, eccentricity, and ratio between the star and disk intensity, leading to a large panel of images (examples are displayed in Figure 1). The synthetic images are of size 244×254 pixels.

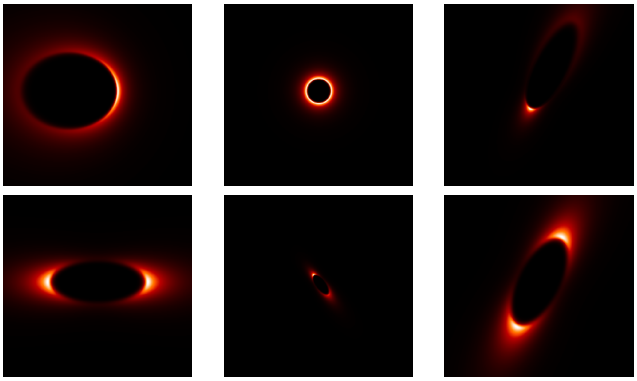

Fig. 1. Examples of the polarized disk intensities I^p of the DDiT dataset used in this work.

Synthetic dataset – The first dataset consists in considering (I^p, Q, U) and to degrade it with a point spread function (PSF) extracted from real observational data taken with the SPHERE instrument at the VLT/ESO using its apodized aperture [1]. We add a Gaussian noise with a standard deviation of 0.1. This dataset has been created in order to provide an ideal dataset allowing to particularly fit the data-term chosen as $f(A\cdot, z) = \frac{1}{2}\|A\cdot - z\|^2$, which fits white Gaussian noise assumption when considering a *maximum a posterior* interpretation.

More realistic dataset – The second dataset has been created combining I^p, Q, U to generate the dataset, A as described previously, and combined with the RHAPSODIE forward model to create observation z obtained from (1).

We then remove geometrical transforms to only focus on debluring/denoising part. Pre-reconstructed data are derived from observations z with a pipeline based on the double difference algorithm leading to (4) where the distribution of the noise is now far from being white Gaussian noise.

B. Architecture and training

The datasets were partitioned into a training set, encompassing 85% of the data (573 images), and a test set, comprising the remaining 102 images. The Standard CV has been run with anisotropic TV penalisation for 500 iterations and the results are displayed for several values of λ. The unfolded neural networks have been trained using the ADAM optimizer until the training loss reached stabilization. The unfolded schemes were trained with $(P, K) = (16, 80), (32, 40), (64, 20)$, manipulating either the number of filters P or the depth of the unfolded neural network. For each (P, K), the sensitivity to the initialization has been evaluated with three different initialization realization. The unfolded version where $D_k = W_k V_k$ has been created in order to increase significantly the number of parameters and better mimic deep architectures. V_k is thus mimicking D_k as previously described while W_k is linear operator whose input/output space is the same and has a number of channel $P = \{16, 32, 64\}$. For the nonlinear unfolded CV, $[\nabla D_k(x^{[k]})]^* \approx V_k^* \text{relu}(W_k^*)$.

C. Results

Figure 2 summarizes the performance obtained with the different strategies. We illustrate the performance w.r.t the number of parameters involved in each recovery strategy. On the synthetic dataset where the distribution of the noise is simply a white Gaussian noise, we can observe that unfolded CV with learned D_k at each layer provides a good compromise between performance and number of parameters while for more complex and realistic dataset, considering a larger number of parameters and handle with nonlinear CV provides much better performance. The variability in the same method/architecture with different configurations (number of parameters/number of layers) is illustrated showing the interest in comparing the results for different network complexities.

In Figure 3 (resp. 4), we display an example of original (I^p, Q, U) as well as the degraded counterpart and the reconstruction with the different CV strategies. We can clearly observe that unfolded strategies outperform the standard variational procedure and that integrating nonlinearity leads to very realistic results both on the synthetic dataset and the more realistic one.

VI. Conclusion

This work is the first study devoted to unfolded deep learning architectures and, more generally, to deep learning strategies in the context of high-contrast spectro-polarimetric image retrieval. In order to evaluate performance, two datasets were created and several unfolded schemes were developed. We highlight the need to use more complex unfolded architectures, such as the nonlinear ones, on more realistic datasets.

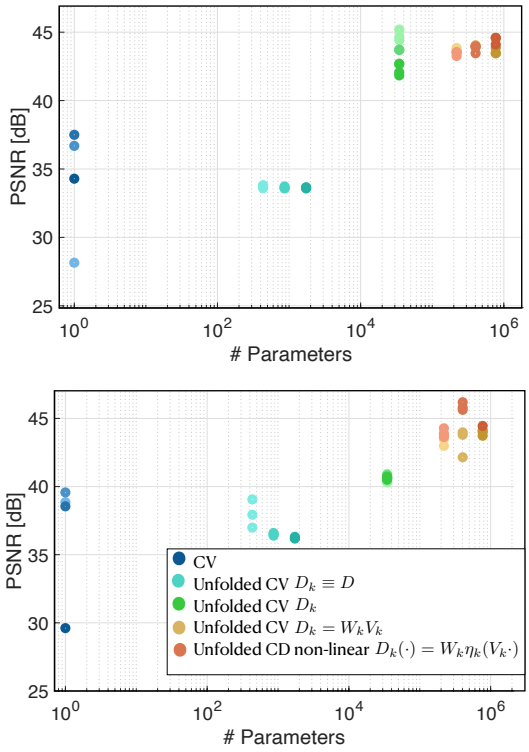

Fig. 2. Performance in terms of PSNR w.r.t parameter numbers for different strategies for (top) the synthetic dataset and (bottom) the pre-reconstructed one. Each configuration contains the results for different initialization of the training and different set of parameters. The dots with the same color correspond to several trials while the different parameter values are displayed from light to darker color. For instance for Standard CV, the darker is the color, the larger is λ. Similar for unfolded strategies, the darker is the color, the larger is P.

References

[1] J.-L. Beuzit, A. Vigan, D. Mouillet, K. Dohlen, R. Gratton, A. Boccaletti, J.-F. Sauvage, H. Martin Schmid, M. Langlois, C. Petit, et al., "SPHERE: the exoplanet imager for the Very Large Telescope," Astronomy & Astrophysics, vol. 631, pp. A155, 2019.

[2] K. Dohlen, M. Langlois, M. Saisse, L. Hill, A. Origne, M. Jacquet, C. Fabron, J.C. Blanc, M. Llored, M. Carle, et al., "Ground-based and Airborne Instrumentation for Astronomy II," in Proc. Spie, 2008, vol. 7014, p. 70143L.

[3] J. Tinbergen, Astronomical Polarimetry, Cambridge University Press, 2005.

[4] H. Avenhaus, S. P. Quanz, H.M. Schmid, M. R Meyer, A. Garufi, S. Wolf, and C. Dominik, "Structures in the protoplanetary disk of HD142527 seen in polarized scattered light," The Astrophysical Journal, vol. 781, no. 2, pp. 87, 2014.

[5] L. Denneulin, M. Langlois, E. Thiébaut, and N. Pustelnik, "RHAPSODIE: Reconstruction of high-contrast polarized sources and deconvolution for circumstellar environments," Astronomy & Astrophysics, vol. 653, pp. A138, 2021.

[6] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (SARA): a novel algorithm for radio-interferometric imaging," Monthly Notices of the Royal Astronomical Society, vol. 426, no. 2, pp. 1223–1234, 2012.

[7] N. Pustelnik, A. Benazza-Benhayia, Y. Zheng, and J.-C. Pesquet, "Wavelet-based image deconvolution and reconstruction," Wiley Encyclopedia of EEE, 2016.

[8] C. Dong, C.C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13. Springer, 2014, pp. 184–199.

[9] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in Proc. 27th International Conference on Machine Learning, 2010, pp. 399–406.

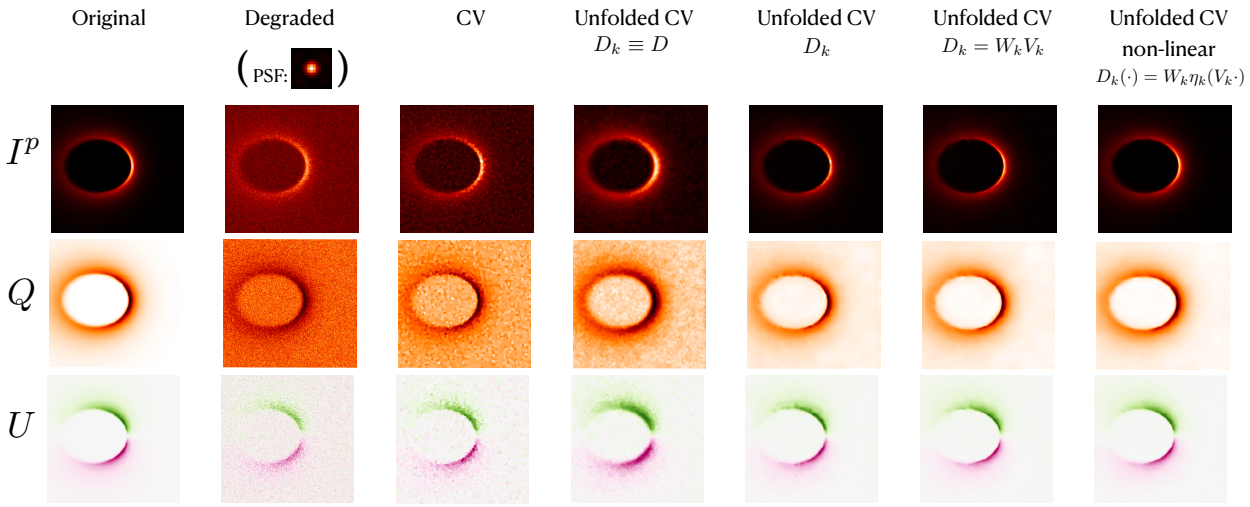

Fig. 3. Visualisation of the reconstructed images for the synthetic dataset. Each method is illustrated with the configuration leading to a good trade-off performance from Figure 2. Condat-Vũ consider finite difference operator with $\lambda = 0.1$. Unfolded CV Fix trained for $K = 80$ and $P = 16$. Unfolded CV not Fix is trained for $K = 80$ and $P = 16$. Unfolded CV deep $D_k = W_k V_k$ is trained for $K = 80$ and $P = 16$. Unfolded CV non linear $D_k(\cdot) = W_k \eta_k(V_k \cdot)$ is trained for $K = 32$ and $P = 40$.

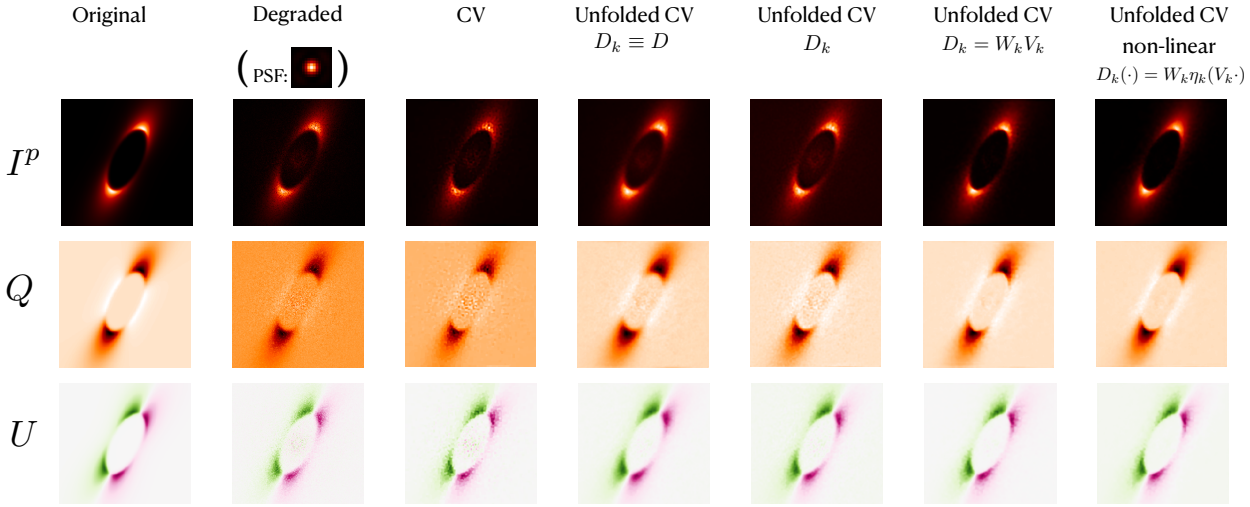

Fig. 4. Visualisation of the reconstructed images for the synthetic dataset. Each method is illustrated with the configuration leading to the best performance from Figure 2. Condat-Vũ consider finite difference operator with $\lambda = 0.1$. Unfolded CV Fix trained for $K = 80$ and $P = 16$. Unfolded CV not Fix is trained for $K = 80$ and $P = 16$. Unfolded CV deep $D_k = W_k V_k$ is trained for $K = 80$ and $P = 16$. Unfolded CV non linear $D_k(\cdot) = W_k \eta_k(V_k \cdot)$ is trained for $K = 32$ and $P = 40$.

[10] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," IEEE J. J. Sel. Areas Inf. Theory, vol. 1, no. 1, pp. 39–56, 2020.

[11] J.-C. Pesquet, A. Repetti, M. Terris, and Y. Wiaux, "Learning maximally monotone operators for image recovery," SIAM J. Imaging Sci., vol. 14, no. 3, pp. 1206–1237, 2021.

[12] S. Hurault, A. Leclaire, and N. Papadakis, "Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization," in International Conference on Machine Learning. PMLR, 2022, pp. 9483–9505.

[13] U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications," IEEE Signal Process. Mag., vol. 40, no. 1, pp. 85–97, 2023.

[14] H. T. V. Le, A. Repetti, and N. Pustelnik, "PNN: From proximal algorithms to robust unfolded image denoising networks and plug-and-play methods," preprint arXiv:2308.03139, 2023.

[15] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," J. Optim. Theory Appl., vol. 158, no. 2, pp. 460–479, 2013.

[16] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," Advances in Computational Mathematics, vol. 38, pp. 667–681, 2013.

[17] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," Fixed-point algorithms for inverse problems in science and engineering, pp. 185–212, 2011.

[18] G. Chierchia, E. Chouzenoux, P.L. Combettes, and J.-C. Pesquet, "The proximity operator repository," http://proximity-operator.net/.

[19] M. Jiu and N. Pustelnik, "A deep primal-dual proximal network for image restoration," IEEE J. Sel. Top. Signal Process., vol. 15, no. 2, pp. 190–203, 2021.

[20] T. Valkonen, "A primal–dual hybrid gradient method for nonlinear operators with applications to MRI," Inverse Problems, vol. 30, no. 5, pp. 055012, 2014.

[21] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," J. Math. Imaging Vis., vol. 40, pp. 120–145, 2011.

[22] J. Tachella, D. Chen, S. Hurault, and M. Terris, "DeepInverse: A deep learning framework for inverse problems in imaging," June 2023.

[23] J. Olofsson, J. Milli, A. Bayo, T. Henning, and N. Engler, "The challenge of measuring the phase function of debris discs-application to HR 4796 A," Astronomy & Astrophysics, vol. 640, pp. A12, 2020.