

Proximal Algorithm Unrolling: Flexible and Efficient Reconstruction Networks for Single-Pixel Imaging -Supplementary Material-

Ping Wang^{1,2*} Lishun Wang^{1*} Gang Qu¹ Xiaodong Wang^{1,2} Yulun Zhang³ Xin Yuan^{1†}
¹Westlake University ²Zhejiang University ³Shanghai Jiao Tong University

1. Convergence Analysis

As mentioned in main paper, the Kronecker equation $\Phi = \mathbf{W} \otimes \mathbf{H}$ builds an equivalent transformation from $\mathbf{Y} = \mathbf{H}\mathbf{X}\mathbf{W}^\top + \mathbf{E}$ to $\mathbf{y} = \Phi\mathbf{x} + \epsilon$, where $\mathbf{y} = \text{vec}(\mathbf{Y})$, $\mathbf{x} = \text{vec}(\mathbf{X})$, $\epsilon = \text{vec}(\mathbf{E})$. Accordingly, the objective function $F(\mathbf{X}) = f(\mathbf{X}) + \lambda g(\mathbf{X})$ is equivalent to its vectorized format $F(\mathbf{x}) = f(\mathbf{x}) + \lambda g(\mathbf{x})$. The following analyses are based on the vectorized objective function for clarity.

In single-pixel imaging (SPI) paradigm, $\text{Prox}_f(\mathbf{x}) = \mathbf{x} + \frac{\Phi^\top(\mathbf{y} - \Phi\mathbf{x})}{1+\rho}$ is the proximal operator of data fidelity function $f = \frac{1}{2}\|\mathbf{y} - \Phi\mathbf{x}\|_2^2$ for HQS and ADMM. We propose a learned proximal restorer (LPR) \mathcal{R}_θ to approximate the proximal operator $\text{Prox}_{g^*}(\mathbf{x}) = \frac{\rho\mathbf{x} + \lambda\mathbf{x}^*}{\rho + \lambda}$ of an explicit regularization function $g^* = \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}^*\|_2^2$ (\mathbf{x}^* is the ground truth).

1.1. Convergence proof of PnP-HQS

$\text{Prox}_f(\mathbf{x}) = \mathbf{x} + \frac{\Phi^\top(\mathbf{y} - \Phi\mathbf{x})}{1+\rho}$ is also a gradient descent mapping $(\text{Id} - \tau\nabla f)(\mathbf{x})$ with $\tau = (1+\rho)^{-1}$ and thus we use the convergence analysis of the proximal gradient descent (PGD) algorithm [1] for PnP-HQS convergence.

Proof. We denote the composition proximal operator of HQS as $T = \text{Prox}_{g^*} \circ \text{Prox}_f$, the objective function $F = f + \lambda g^*$, and we introduce

$$Q(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}) + \langle \mathbf{v} - \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \frac{1}{2}\|\mathbf{v} - \mathbf{x}\|_2^2 + \lambda g^*(\mathbf{v}). \quad (1)$$

We have

$$Q(\mathbf{x}, \mathbf{x}) = F(\mathbf{x}) \quad (2)$$

and

$$\begin{aligned} \underset{\mathbf{v}}{\text{argmin}} Q(\mathbf{x}, \mathbf{v}) &= \underset{\mathbf{v}}{\text{argmin}} \langle \mathbf{v} - \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \frac{1}{2}\|\mathbf{v} - \mathbf{x}\|_2^2 + \lambda g^*(\mathbf{v}) \\ &= \underset{\mathbf{v}}{\text{argmin}} \lambda g^*(\mathbf{v}) + \frac{1}{2}\|\mathbf{v} - (\mathbf{x} - \nabla f(\mathbf{x}))\|_2^2 \\ &= \text{Prox}_{g^*} \circ (\text{Id} - \nabla f)(\mathbf{x}) = T(\mathbf{x}). \end{aligned} \quad (3)$$

The above argmin is unique by definition of the argmin and the regularizer g^* , $\mathbf{x}^{k+1} = T(\mathbf{x}^k)$ implies that

$$Q(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq Q(\mathbf{x}^k, \mathbf{x}^k). \quad (4)$$

Moreover, since f is L -smooth, we have by the descent lemma, for any $t \leq \frac{1}{L}$ and any $\mathbf{x}, \mathbf{v} \in \mathbb{R}^n$,

$$f(\mathbf{v}) \leq f(\mathbf{x}) + \langle \mathbf{v} - \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \frac{1}{2t}\|\mathbf{v} - \mathbf{x}\|_2^2. \quad (5)$$

By combining (1) with (5), for every $\mathbf{x}, \mathbf{v} \in \mathbb{R}^n$ and taking $t = 1 < \frac{1}{L}$, we have

$$F(\mathbf{x}) \leq Q(\mathbf{x}, \mathbf{v}). \quad (6)$$

Therefore, by combining (2), (4) and (6), we get at iteration k ,

$$F(\mathbf{x}^{k+1}) \leq Q(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq Q(\mathbf{x}^k, \mathbf{x}^k) = F(\mathbf{x}^k). \quad (7)$$

The sequence $\{F(\mathbf{x}^k)\}_{k=1}^\infty$ is thus non-increasing and lower-bounded. $\{F(\mathbf{x}^k)\}_{k=1}^\infty$ thus converges to a limit F^* . Note that $Q(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq Q(\mathbf{x}^k, \mathbf{x}^k)$ in (4) implies

$$\lambda g^*(\mathbf{x}^{k+1}) \leq \lambda g^*(\mathbf{x}^k) - \langle \mathbf{x}^{k+1} - \mathbf{x}^k, \nabla f(\mathbf{x}^k) \rangle - \frac{1}{2}\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2. \quad (8)$$

By combining (5) with step size $t = \frac{1}{L}$, we have

$$\begin{aligned} F(\mathbf{x}^{k+1}) &= f(\mathbf{x}^{k+1}) + \lambda g^*(\mathbf{x}^{k+1}) \\ &\leq \lambda g^*(\mathbf{x}^k) - \langle \mathbf{x}^{k+1} - \mathbf{x}^k, \nabla f(\mathbf{x}^k) \rangle - \frac{1}{2}\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2 \\ &\quad + f(\mathbf{x}^k) + \langle \mathbf{x}^{k+1} - \mathbf{x}^k, \nabla f(\mathbf{x}^k) \rangle + \frac{L}{2}\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2 \\ &= F(\mathbf{x}^k) - \frac{1}{2}(1-L)\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2. \end{aligned} \quad (9)$$

Summing over $k = 0, 1, \dots, K$ gives

$$\begin{aligned} \sum_{k=0}^K \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2 &\leq \frac{2}{1-L} [F(\mathbf{x}_0) - F(\mathbf{x}_{K+1})] \\ &\leq \frac{2}{1-L} [F(\mathbf{x}_0) - F^*]. \end{aligned} \quad (10)$$

*Equal contribution.

†Corresponding author: Xin Yuan, xyuan@westlake.edu.cn

Therefore, $\lim_{k \rightarrow \infty} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2 = 0$ with the convergence rate $\gamma^k = \min_{0 \leq i \leq k} \|\mathbf{x}_{i+1} - \mathbf{x}_i\|_2^2 \leq \frac{2}{k} \frac{F(\mathbf{x}_0) - \lim F(\mathbf{x}^k)}{1-L}$. \square

1.2. Convergence proof of PnP-ADMM

We use the convergence analysis of typical ADMM in Theorem 1.

Theorem 1 (Convergence of ADMM [2]). *If the functions $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ are proper, closed, and convex, there exist a saddle point $(\mathbf{x}^*, \mathbf{v}^*, \mathbf{u}^*)$, the ADMM iterates satisfy:*

- *Residual convergence.* $\lim_{k \rightarrow \infty} (\mathbf{x}^k - \mathbf{v}^k) = 0$, i.e., the iterates approach feasibility.
- *Objective convergence.* $\lim_{k \rightarrow \infty} F = \lim_{k \rightarrow \infty} f(\mathbf{x}^k) + g(\mathbf{v}^k) = f(\mathbf{x}^*) + g(\mathbf{v}^*) = F^*$, i.e., the objective function of the iterates approaches the optimal value F^* .
- *Dual variable convergence.* $\lim_{k \rightarrow \infty} \mathbf{u}^k = \mathbf{u}^*$, where \mathbf{u}^* is a dual optimal point.

Recall that $f = \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2$ and $g^* = \frac{\lambda}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2$. The epigraph of $h \in \{f, g^*\}$

$$\text{epi } h = \{(\mathbf{x}, r) \in \mathbb{R}^n \times \mathbb{R} | h(\mathbf{x}) < r\} \quad (11)$$

is a closed non-empty convex set. It means that both f and g^* are proper, closed, and convex, thus PnP-ADMM converge to the saddle point $(\mathbf{x}^*, \mathbf{v}^*, \mathbf{u}^*)$.

2. Optimization Trajectories of Unrolling Networks

As analyzed in main paper, the optimization trajectories of unrolling networks are usually zigzagged since their MSE loss function could just impose that the final result is good. Besides, the artifacts in intermediate results may be the manifestations of informative representations in image domain. To demystify this, we take SAUNet, a representative unrolling network, as an example. As shown in Fig. 1, we visualize the optimization trajectories of SAUNet [4] at compressive ration 0.25 in different conditions, which are composed of the input \mathbf{X}^0 , the intermediate results $\{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^6\}$, the final output \mathbf{X}^7 (i.e., 8 images in total), and we use PSNR (dB) to measure the image quality. At the original optimization trajectory, \mathbf{X}^3 and \mathbf{X}^5 appears to be full of artifacts, and \mathbf{X}^7 is the best. To understand what the visual artifacts are, we add Gaussian noise with 0.1 variance into \mathbf{X}^3 or reduce/remove the artifacts by $f(\mathbf{X}^3) = \frac{\mu \mathbf{X}^3 + \lambda \mathbf{X}}{\mu + \lambda}$ (\mathbf{X} is the known ground truth). Clearly, any operations over the artifacts lead to a image quality decrease over the final results \mathbf{X}^7 . Hence, there is a seemingly counterintuitive fact that the visual artifacts in intermediate results are not useless in unrolling networks.

In typical unrolling networks, several subnets are separated by a proximal operator and thus useful representations cannot be propagated between subnets. We think that such artifacts may be the incarnations of useful representations for implicit information propagation.

3. Accuracy & Flexibility Comparison

As mentioned in main paper, PnP methods perform well on generalization but poorly on accuracy, and unrolling methods are the opposite. Most previous unrolling methods must be retrained or fine-tuned once the degradation matrix $\Phi \in \mathbb{R}^{m \times n}$ ($m \ll n$) changes, i.e., there is an one-to-one relationship between a well-trained model and a specific compressive ratio (CR). PnP methods are generic for different tasks and the degrees of degradation and thus are flexible for varying CRs.

Proximal unrolling (ProxUnroll) is proposed to integrate the merits of PnP and unrolling methods for high accuracy, high efficiency, strong flexibility, and fast convergence. When CR changes from 0.05 to 0.25, Fig. 2 visualizes the real experiment results of PnP-DRUNet [5] (PnP method), SAUNet-0.25 [4] (trained on CR=0.25), SAUNet-0.15 (trained for CR=0.15), HATNet-0.25 [3] (trained for CR=0.25), HATNet-0.15 (trained for CR=0.15), and our HQS-ProxUnroll. Through the comparison on the marked regions, it is clear that our single model HQS-ProxUnroll can recover best details. PnP-DRUNet is highly flexible for varying CRs but its performance is average due to the lack of domain-specific knowledge. As two representative unrolling networks, SAUNet and HATNet perform well just on the trained CR or close CRs.

References

- [1] Amir Beck. *First-order methods in optimization*. SIAM, 2017. 1
- [2] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011. 2
- [3] Gang Qu, Ping Wang, and Xin Yuan. Dual-scale transformer for large-scale single-pixel imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25327–25337, 2024. 2
- [4] Ping Wang and Xin Yuan. Saunet: Spatial-attention unfolding network for image compressive sensing. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 5099–5108, 2023. 2
- [5] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021. 2

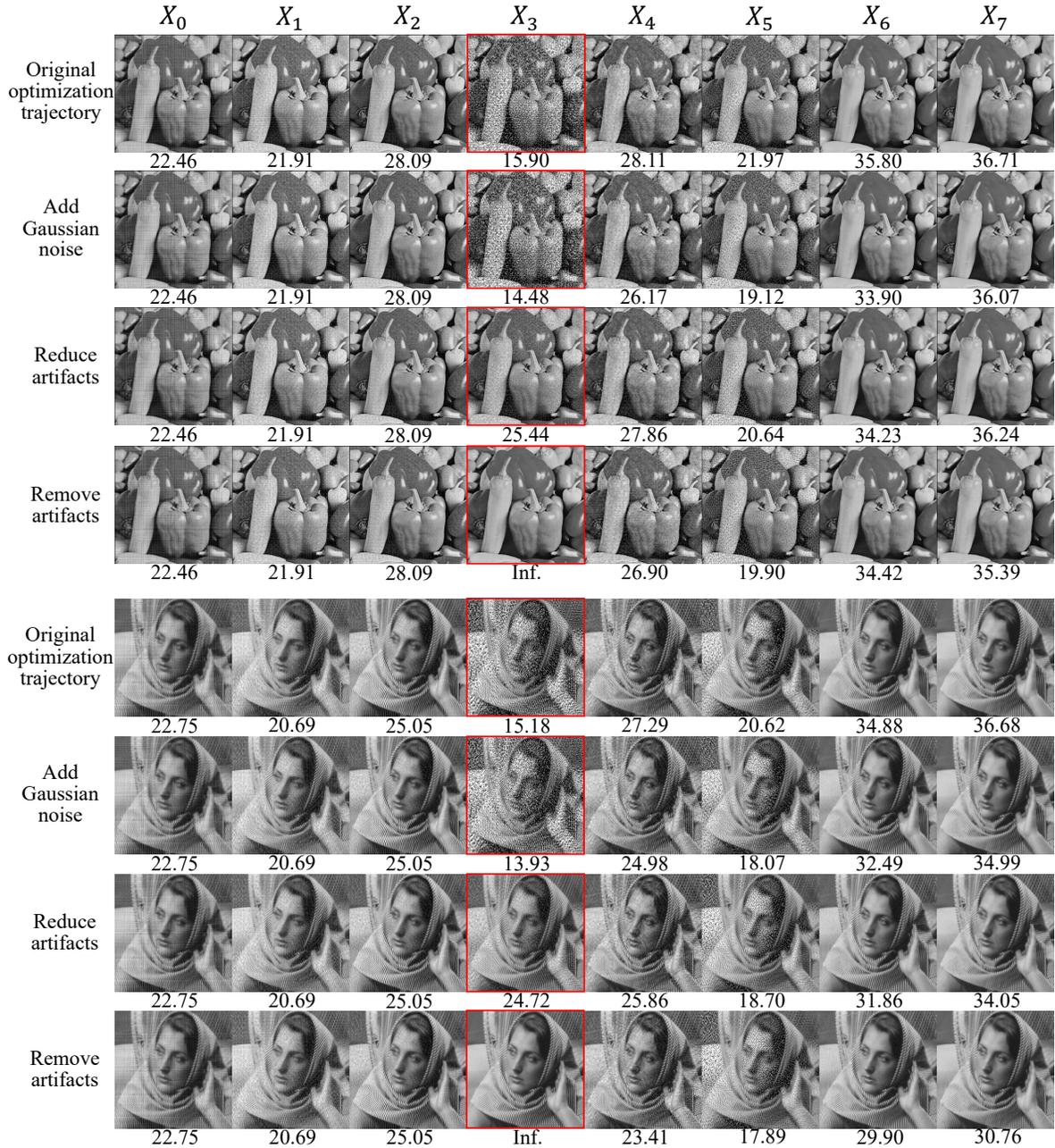


Figure 1. Optimization trajectories of a representative unrolling network (SAUNet) by adding Gaussian noise, reducing artifacts, removing artifacts on one intermediate result highlighted by red box. The image quality is measured by PSNR (dB).

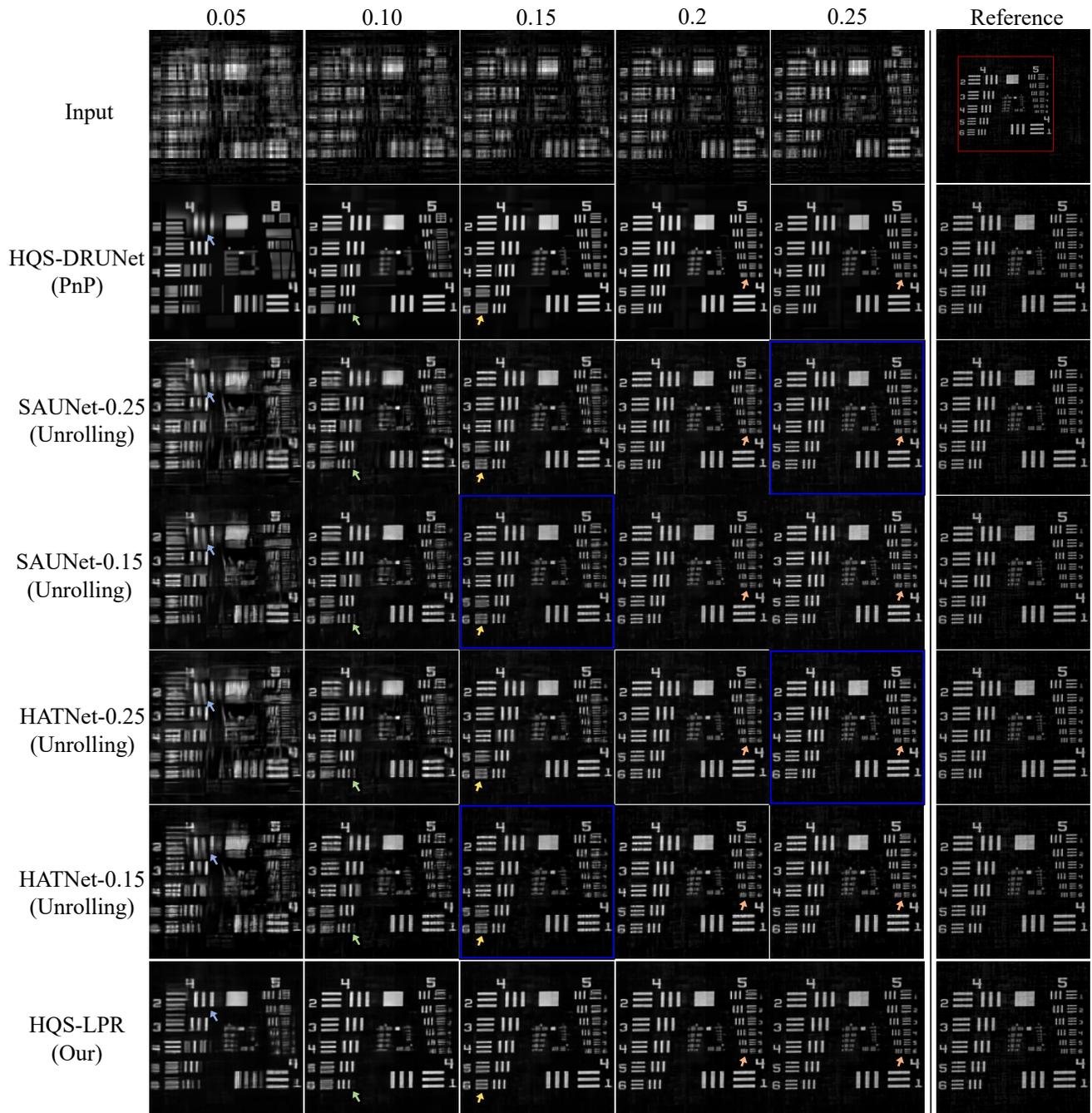


Figure 2. The testing results of the real captured “resolution target” with 256×256 pixels. Unrolling networks are trained on a specific CR marked by blue box.