CONVERGENCE OF NONCONVEX PNP-ADMM WITH MMSE DENOISERS

Chicago Park[*,†], *Shirin Shoushtari*[*,‡], *Weijie Gan*[†], *and Ulugbek S. Kamilov*[†,‡]

[†]Department of Computer Science & Engineering, Washington University in St. Louis, USA
[‡]Department of Electrical & Systems Engineering, Washington University in St. Louis, USA

ABSTRACT

Plug-and-Play Alternating Direction Method of Multipliers (PnP-ADMM) is a widely-used algorithm for solving inverse problems by integrating physical measurement models and convolutional neural network (CNN) priors. PnP-ADMM has been theoretically proven to converge for convex data-fidelity terms and nonexpansive CNNs. It has however been observed that PnP-ADMM often empirically converges even for expansive CNNs. This paper presents a theoretical explanation for the observed stability of PnP-ADMM based on the interpretation of the CNN prior as a minimum mean-squared error (MMSE) denoiser. Our explanation parallels a similar argument recently made for the iterative shrinkage/thresholding algorithm variant of PnP (PnP-ISTA) and relies on the connection between MMSE denoisers and proximal operators. We also numerically evaluate the performance gap between PnP-ADMM using a nonexpansive DnCNN denoiser and expansive DRUNet denoiser, thus motivating the use of expansive CNNs.

Index Terms— Computational imaging, inverse problems, plug-and-play priors, nonconvex optimization.

1. INTRODUCTION

A fundamental problem in computational imaging is the recovery of an unknown image $x \in \mathbb{R}^n$ from noisy measurements

$$y = Ax + e,$$

where $A \in \mathbb{R}^{m \times n}$ is the measurement operator and $e \in \mathbb{R}^n$ is the measurements noise. It is common to formulate the recovery as a composite optimization problem

$$\widehat{x} = \arg\min_{x \in \mathbb{R}^n} f(x) \quad \text{with} \quad f(x) = g(x) + h(x), \quad (1)$$

where the function g represents the data-fidelity term and h denotes the regularizer or the prior term.

Proximal algorithms are commonly used to solve the optimization problems in eq. (1) when the functions g or h are nonsmooth. For example, iterative shrinkage/thresholding algorithm (ISTA) [1, 2] and alternating direction method of multipliers (ADMM) [3,4] have been widely-used in the context of imaging inverse problems. The iterations of ADMM can be expressed as

$$x^k \leftarrow \text{prox}_{\gamma g}(z^{k-1} - s^{k-1}) \quad (2a)$$

$$z^k \leftarrow \text{prox}_{\gamma h}(x^k + s^{k-1}) \quad (2b)$$

$$s^k \leftarrow s^{k-1} + x^k - z^k. \quad (2c)$$

The key operation within ADMM is the *proximal operator*

$$\text{prox}_{\gamma h}(z) := \arg\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \|x - z\|_2^2 + \gamma h(x) \right\}, \quad (3)$$

where the parameter $\gamma > 0$ controls the influence of h. When $h(x) = -\log(p_x(x))$, the proximal operator can be interpreted as a maximum a posteriori (MAP) estimator for the additive white Gaussian noise (AWGN) denoising problem

$$z = x + n \quad \text{with} \quad x \sim p_x, \quad n \sim \mathcal{N}(0, \sigma^2 I). \quad (4)$$

The statistical interpretation of the proximal operator as a denoiser has led to the development of the plug-and-play ADMM (PnP-ADMM) [5], where $\text{prox}_{\gamma h}$ is replaced with a more general image denoiser D_σ

$$x^k \leftarrow \text{prox}_{\gamma g}(z^{k-1} - s^{k-1}) \quad (5a)$$

$$z^k \leftarrow D_\sigma(x^k + s^{k-1}) \quad (5b)$$

$$s^k \leftarrow s^{k-1} + x^k - z^k, \quad (5c)$$

where $\sigma > 0$ controls the relative strength of the denoiser. The popularity of deep learning has led to a wide adoption of PnP for exploiting learned priors specified through convolutional neural networks (CNNs), leading to its state-of-the-art performance in a variety of applications [6–8].

There has been significant interest in theoretically understanding the convergence behavior of PnP-ADMM [9–14] (see
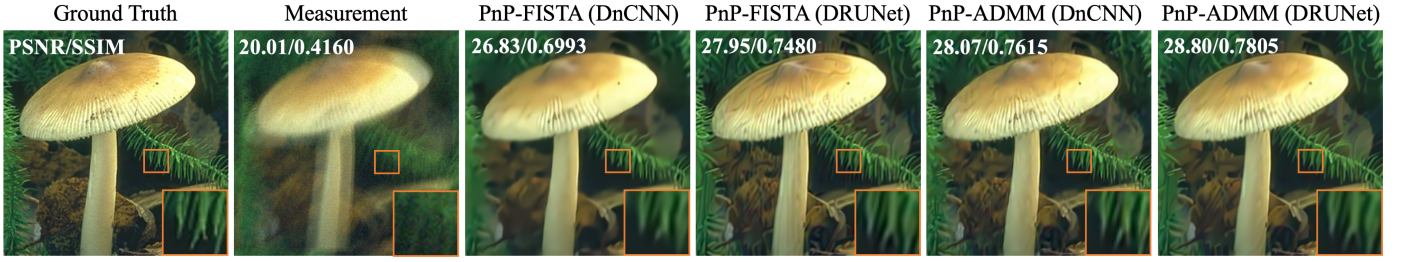
Fig. 1. Comparison of four different methods for deblurring a color image with a noise level of 0.03. The reconstruction performance is quantified using PSNR and SSIM in the top-left corner of each image. Note the improved performance of PnP-ADMM using an expansive DRUNet denoiser compared to nonexpansive DnCNN denoiser.

also the recent review [15]). The most well-known theoretical convergence results for PnP-ADMM are based on monotone operator theory, and require for g to be convex and D_σ or its residual to be nonexpansive [9, 11, 13]. Another well-known result does not assume nonexpansiveness by instead requiring bounded ∇g and D_σ [10]. Recently PnP-ADMM was also analyzed for proximal operators associated with denoisers trained as gradients of explicitly specified regularizers [14].

While the current theory provides useful insights into the stability of PnP-ADMM iterations, it has been observed that PnP-ADMM often converges for denoisers that are expansive, unbounded, and/or are not trained using any explicitly specified function h. The goal of this paper is to offer an explanation for this stability by building on the recent analysis of PnP-ISTA for denoisers that perform minimum mean squared error (MMSE) estimation [16]. The analysis of PnP-ISTA relies on an elegant connection between MMSE denoisers and proximal operators established by Gribonval [17]. What makes this connection pertinent is that the MSE is frequently used as a loss function for training stete-of-the-art image denoisers. We explicitly relate the statistical interpretation of CNNs trained as MMSE denoisers with the nonconvex analysis of the traditional ADMM algorithm [18–22]. We numerically motivate our theoretical exposition by comparing PnP-ADMM using two different pretrained CNN denoisers, a nonexpansive DnCNN and an expansive DRUNet. The numerical results show the convergence of PnP-ADMM with an expansive denoiser and highlight the limitations of nonexpansive denoisers as priors within the PnP framework.

2. CONVERGENCE THEORY

We state three assumptions required to establish the convergence of PnP-ADMM under a nonconvex data-fidelity term g and a MMSE denoiser D_σ.

Assumption 1. *The prior p_x is non-degenerate over $\mathbb{R}^n$.*

We call the probability distribution p_x a degenerate distribution

over $\mathbb{R}^n$, if its support lies on a space with lower dimension than n. For the MMSE denoiser D_σ with the image set of $\mathcal{X} := \mathsf{Im}(D_\sigma)$, Assumption 1 is required to establish an explicit link between D_σ and the following regularizer [17]

$$h_{\mathsf{mmse}}(\boldsymbol{x}) := \qquad\qquad (6)$$
$$\begin{cases} -\frac{1}{2\gamma}\|\boldsymbol{x} - D_\sigma^{-1}(\boldsymbol{x})\|_2^2 + \frac{\sigma^2}{\gamma} h_\sigma(D_\sigma^{-1}(\boldsymbol{x})) & \text{for } \boldsymbol{x} \in \mathcal{X} \\ +\infty & \text{for } \boldsymbol{x} \notin \mathcal{X}, \end{cases}$$

where $\gamma > 0$ is the step-size, D_σ^{-1} is a smooth inverse mapping over $\mathcal{X}$, and $h_\sigma(\cdot) = -\log(p_z(\cdot))$, where p_z os the probability distribution in (4). Due to smoothness of both D_σ^{-1} and h_σ, the function h_{mmse} is smooth for all $\boldsymbol{x} \in \mathcal{X}$. For additional discussion see [16, 17].

Assumption 2. *Function g is continuously differentiable. The function h in (6) associated with D_σ has a Lipschitz continuous gradients with constants $M > 0$ over the set $\mathcal{X}$.*

The smoothness of g is a mild and commonly used assumption [21]. Since it is known (see [16, 17]) that the implicitly defined function h is infinitely differentiable over $\mathcal{X}$, our assumption is a mild extension that additionally requires Lipschitz continuity of the gradient.

Assumption 3. *The explicit data-fidelity term and the implicit regularizer are bounded from below*

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} g(\boldsymbol{x}) > -\infty, \qquad \inf_{\boldsymbol{x}\in\mathbb{R}^n} h(\boldsymbol{x}) > -\infty.$$

Assumption 3 implies that there exists $f^* > -\infty$ such that $f(\boldsymbol{x}) \geq f^*$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Theorem 1. *Run PnP-ADMM with a MMSE denoiser under Assumptions 1-3 using a fixed step-size $0 < \gamma \leq 1/(2M)$. Then, then the iterates satisfy $\|\nabla f(\boldsymbol{x}^k)\|_2 \to \boldsymbol{0}$ as $k \to 0$, where $f = g + h$ with h defined in (6).*

The proof is provided in the appendix. Theorem 1 establishes the convergence of PnP-ADMM with MMSE denoisers to a
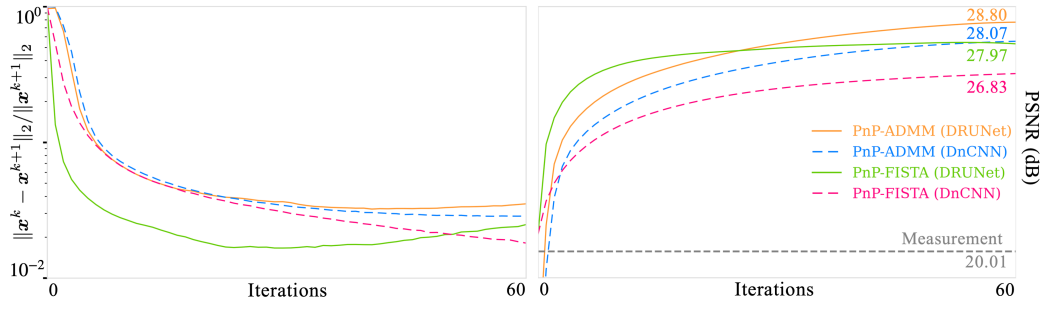
Fig. 2. Comparison of PnP-ADMM and PnP-FISTA, each using a non-expansive DnCNN denoiser and an expansive DRUNet denoiser. The figure plots the evolution of $\|\boldsymbol{x}^k - \boldsymbol{x}^{k+1}\|_2/\|\boldsymbol{x}^{k+1}\|_2$, while the right one that of PSNR (dB).

critical point of the problem (1) where h is defined in (6). It is important to note that the proof does not require the convexity of g or h, or nonexpansiveness of the denoiser. This result suggests the stability of PnP-ADMM under possibly expansive denoisers trained to minimize MSE. The convexity of h_{mmse} requires for D_σ to be monotone and nonexpansive, which does not hold for many CNNs.

3. NUMERICAL EVALUATION

We explore the convergence and performance of two PnP algorithms, PnP-ADMM and PnP-FISTA, on the problem of image deblurring using two popular CNN priors, DnCNN and DRUNet, trained to minimize mean squared error. The DnCNN is trained to be nonexpansive by removing its batch normalization layers and using spectral normalization to bound its Lipschitz constant. The evaluation of these methods is conducted using four different blur kernels and four CBSD68 test images at three distinct noise levels (0.01, 0.02, and 0.03). The performance is quantified using peak signal-to-noise ratio (PSNR) in dB and structural similarity index (SSIM) metrics.

Figure 1 compares all the four methods applied to a color image with a noise level of 0.03. It is observed that PnP-ADMM converges for the expansive denoiser and achieves significantly better results in terms of both PSNR and SSIM. PnP-FISTA with DRUNet outperforms the one using a nonexpansive DnCNN. Figure 2 presents the convergence profiles of PnP-ADMM and PnP-FISTA with DnCNN and DRUNet as image priors. Theoretical support for the convergence of PnP-ADMM is established through Theorem 1, which does not impose any assumptions on the expansiveness of the image prior. The results in Figure 2 plot the evolution of $\|\boldsymbol{x}^k - \boldsymbol{x}^{k+1}\|_2/\|\boldsymbol{x}^{k+1}\|_2$ and PSNR for the image presented in Figure 1, supporting the theoretical observation that PnP-ADMM can converge for expansive CNNs.

Table 1 presents the performance comparison of PnP-ADMM and PnP-FISTA for both image priors. The reported results are

| | Noise Level | | | |
PnP Methods	0.01	0.02	0.03	avg
ADMM (DRUNet)	**30.67**	**29.48**	**28.03**	**29.39**
ADMM (DnCNN)	28.84	28.01	26.25	27.70
FISTA (DRUNet)	27.30	27.20	26.84	27.11
FISTA (DnCNN)	26.57	26.40	25.90	26.29

Table 1. Performance of PnP-ADMM and PnP-FISTA using two priors on image deblurring at different levels of noise.
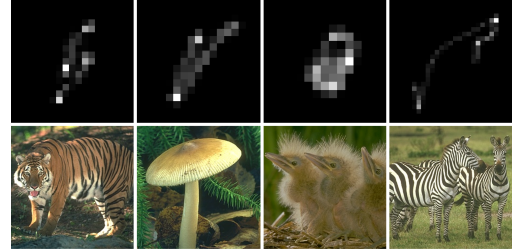

Fig. 3. Four blur kernels [23] and test color images [24] used for the numerical evaluation.

obtained by averaging PSNR values over all the test images and kernels, illustrated in Figure 3. Note how PnP-ADMM outperforms PnP-FISTA in terms of PSNR for both priors. Additionally, the use of the expansive denoiser, DRUNet, results in better performance compared to the nonexpansive denoiser, DnCNN, regardless of the algorithm used. The presented findings emphasize the notable advantage that PnP methods can leverage advanced denoisers, resulting in enhanced performance.

4. CONCLUSION

This paper presents new insights to the well-established PnP methodology by revisiting the convergence of PnP-ADMM un-

der MMSE denoisers. We present convergence analysis of PnP-ADMM for expansive image denoisers and possibly nonconvex data-fidelity terms. We present numerical results highlighting the potential benefit of using expansive denoisers corresponding to CNNs trained to minimize MSE. Our findings emphasize that superior performance of PnP-ADMM under expansive denoisers can still come with stable convergence. This highlights the potential of combining state-of-the-art denoising techniques with the flexibility of PnP.

A. APPENDIX

We now present the convergence analysis of PnP-ADMM without any convexity or nonexpansiveness assumptions. Consider the augmented Lagrangian of the objective function $f(\boldsymbol{x}) = g(\boldsymbol{x}) + h(\boldsymbol{x})$

$$\mu(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{s}) \tag{7}$$
$$= g(\boldsymbol{x}) + h(\boldsymbol{z}) + \frac{1}{\gamma}\boldsymbol{s}^{\mathsf{T}}(\boldsymbol{x} - \boldsymbol{z}) + \frac{1}{2\gamma}\|\boldsymbol{x} - \boldsymbol{z}\|_2^2.$$

For iteration $k \geq 1$ we have

$$\mu(\boldsymbol{x}^k, \boldsymbol{z}^k, \boldsymbol{s}^k) - \mu(\boldsymbol{x}^k, \boldsymbol{z}^k, \boldsymbol{s}^{k-1}) = \frac{1}{\gamma}\|\boldsymbol{s}^k - \boldsymbol{s}^{k-1}\|_2^2, \tag{8}$$

where we used the definition of the augmented Lagrangian in (7) and the ADMM update (5c). The results from [17] imply that for a MMSE denoiser we have $\mathsf{D}_\sigma = \mathsf{prox}_{\gamma h}$. Therefore, from the optimality of proximal operator and the ADMM update (5c), we have $\boldsymbol{s}^k = \gamma \nabla h(\boldsymbol{z}^k)$. Using the Lipschitz continuity of ∇h in Assumption 2, we have

$$h(\boldsymbol{z}^k) - h(\boldsymbol{z}^{k-1}) \tag{9}$$
$$\leq \frac{1}{\gamma}(\boldsymbol{s}^k)^{\mathsf{T}}(\boldsymbol{z}^k - \boldsymbol{z}^{k-1}) + \frac{M}{2}\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2^2.$$

From the fact that $\boldsymbol{x}^k = \mathsf{prox}_{\gamma g}\left(\boldsymbol{z}^{k-1} - \boldsymbol{s}^{k-1}\right)$, we have

$$\mu(\boldsymbol{x}^k, \boldsymbol{z}^{k-1}, \boldsymbol{s}^{k-1}) \leq \mu(\boldsymbol{x}^{k-1}, \boldsymbol{z}^{k-1}, \boldsymbol{s}^{k-1}). \tag{10}$$

By combining equations (8), (9) , and (10), we get

$$\mu(\boldsymbol{x}^k, \boldsymbol{z}^k, \boldsymbol{s}^k) - \mu(\boldsymbol{x}^{k-1}, \boldsymbol{z}^{k-1}, \boldsymbol{s}^{k-1})$$
$$\leq \frac{1}{\gamma}\|\boldsymbol{s}^k - \boldsymbol{s}^{k-1}\|_2^2 - (\frac{1 - \gamma M}{2\gamma})\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2^2$$
$$\leq -\eta\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2^2, \tag{11}$$

where $\eta := (1 - \gamma M - 2\gamma^2 M^2)/(2\gamma)$, and in the last inequality, we used

$$\|\boldsymbol{s}^k - \boldsymbol{s}^{k-1}\|_2 = \gamma\|\nabla h(\boldsymbol{z}^k) - \nabla h(\boldsymbol{z}^{k-1})\|_2$$
$$\leq \gamma M\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2. \tag{12}$$

Since $0 < \gamma \leq 1/(2M)$, we have $\eta > 0$, which implies that the augmented Lagrangian is monotonically decreasing. Using the fact that $\boldsymbol{s}^k = \gamma \nabla h(\boldsymbol{z}^k)$, we can write

$$\mu(\boldsymbol{x}^k, \boldsymbol{z}^k, \boldsymbol{s}^k) = g(\boldsymbol{x}^k) + h(\boldsymbol{z}^k) + \nabla h(\boldsymbol{z}^k)^{\mathsf{T}}(\boldsymbol{x}^k - \boldsymbol{z}^k)$$
$$+ \frac{1}{2\gamma}\|\boldsymbol{x}^k - \boldsymbol{z}^k\|_2^2 \geq g(\boldsymbol{x}^k) + h(\boldsymbol{x}^k),$$

where we used Lipschitz continuity of ∇h and the fact that $\gamma M \leq 1$. This inequality establishes that augmented Lagrangian is bounded from below due to the fact that both function g and h are bounded from below. This implies that there exists $\mu^* > -\infty$ such that almost surely $\mu^* \leq \mu(\boldsymbol{x}^k, \boldsymbol{z}^k, \boldsymbol{s}^k), k \geq 1$. Thus, the augmented Lagrangian converges due to monotocity. By summing both sides of eq. (11) over $t \geq 1$ iteration, we have

$$\sum_{k=1}^{t}\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2^2 \leq \frac{\mu(\boldsymbol{x}^0, \boldsymbol{z}^0, \boldsymbol{s}^0) - \mu(\boldsymbol{x}^t, \boldsymbol{z}^t, \boldsymbol{s}^t)}{\eta}$$
$$\leq \frac{\mu(\boldsymbol{x}^0, \boldsymbol{z}^0, \boldsymbol{s}^0) - \mu^*}{\eta}, \tag{13}$$

which implies that $\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2^2 \to 0$ as $k \to \infty$. Eq. (12) ensures convergence of $\|\boldsymbol{s}^k - \boldsymbol{s}^{k-1}\|_2$ and $\|\boldsymbol{x}^k - \boldsymbol{z}^k\|_2$ to 0 as $k \to \infty$. For the objective function in (1), by adding and subtracting required terms, we have

$$\|\nabla f(\boldsymbol{x}^k)\|_2 = \|\nabla g(\boldsymbol{x}^k) + \nabla h(\boldsymbol{x}^k)\|_2$$
$$= \|\nabla g(\boldsymbol{x}^k) + \frac{1}{\gamma}\left(\boldsymbol{x}^k - \boldsymbol{z}^{k-1} + \boldsymbol{s}^{k-1}\right)$$
$$+ \nabla h(\boldsymbol{x}^k) + \frac{1}{\gamma}\left(\boldsymbol{z}^k - \boldsymbol{x}^k - \boldsymbol{s}^{k-1}\right) + \frac{1}{\gamma}(\boldsymbol{z}^{k-1} - \boldsymbol{z}^k)\|_2$$
$$= \|\frac{1}{\gamma}(\boldsymbol{z}^{k-1} - \boldsymbol{z}^k) + \nabla h(\boldsymbol{x}^k) - \nabla h(\boldsymbol{z}^k)\|_2$$
$$\leq \frac{1}{\gamma}\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2 + M\|\boldsymbol{x}^k - \boldsymbol{z}^k\|_2 \tag{14}$$

where we used optimality of proximal operator from eq. (5a) and (5b), and in the last line, we used triangle inequality and Lipschitz continuity of ∇h. From convergence of $\|\boldsymbol{z}^k - \boldsymbol{z}^{k-1}\|_2$ and $\|\boldsymbol{x}^k - \boldsymbol{z}^k\|_2$ to 0 as $k \to \infty$, we get

$$\|\nabla f(\boldsymbol{x}^k)\|_2 \to 0,$$

x as $k \to 0$. This shows that PnP-ADMM with MMSE denoiser converges to a critical point of the original objective function $f(\boldsymbol{x})$.

B. REFERENCES

[1] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, 2003.

[2] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imag. Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[3] M. V. Afonso, J. M.Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, September 2010.

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[5] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," *IEEE Glob. Conf. Signal Inf. Process.*, pp. 945–948, 2013.

[6] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Computer Vision and Pattern Recog. (CVPR)*, Honolulu, HI, USA, July 2017, pp. 2808–2817.

[7] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, S. Liu, E. T. Reehorst, and P. Schniter, "Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 105–116, 2020.

[8] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Trans. Pattern Anal. and Machine Intell.*, 2022.

[9] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 408–423, December 2016.

[10] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play admm for image restoration: Fixed-point convergence and applications," *IEEE Trans. on Comput. Imag.*, vol. 3, no. 1, pp. 84–98, 2016.

[11] E. K. Ryu, J. Liu, S. Wnag, X. Chen, Z. Wang, and W. Yin, "Plug-and-play methods provably converge with properly trained denoisers," in *Proc. 36th Int. Conf. Machine Learning (ICML)*, June 2019, pp. 5546–5557.

[12] E. T. Reehorst and P. Schniter, "Regularization by denoising: Clarifications and new interpretations," *IEEE Trans. Comput. Imag.*, vol. 5, no. 1, pp. 52–67, Mar. 2019.

[13] Y. Sun and et al., "Scalable plug-and-play admm with convergence guarantees," *IEEE Trans. on Comput. Imag.*, vol. 7, pp. 849–863, 2021.

[14] S. Hurault, A. Leclaire, and N. Papadakis, "Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization," in *Proc. 39th Int. Conf. Machine Learning (ICML)*, Baltimore, MD, July 17-23, 2022, pp. 9483–9505.

[15] 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," *IEEE Signal Process. Mag.*, vol. 40, no. 1, pp. 85–97, Jan. 2023.

[16] X. Xu, Y. Sun, J. Liu, B. Wohlberg, and U. S. Kamilov, "Provable convergence of plug-and-play priors with MMSE denoisers," *IEEE Signal Process. Letters*, vol. 27, pp. 1280–1284, 2020.

[17] R. Gribonval, "Should penalized least squares regression be interpreted as maximum a posteriori estimation?," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2405–2410, May 2011.

[18] Z. Li and J. Li, "A simple proximal stochastic gradient method for nonsmooth nonconvex optimization," *Adv. in Neural Inf. Process. Syst.*, vol. 31, 2018.

[19] F. Wang, W. Cao, and Z. Xu, "Convergence of multi-block Bregman ADMM for nonconvex composite problems," *Sci. China Inf. Sciences*, vol. 61, 2018.

[20] K. Guo, DR. Han, and T. Wu, "Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints," *Int. Journal of Computer Math.*, vol. 94, no. 8, 2017.

[21] B. Jiang, T. Lin, S. Ma, and S. Zhang, "Structured nonconvex and nonsmooth optimization: algorithms and iteration complexity analysis," *Comput. Optim. and Appl.*, vol. 72, no. 1, pp. 115–157, 2019.

[22] M. Yashtini, "Multi-block nonconvex nonsmooth proximal admm: Convergence and rates under kurdyka–łojasiewicz property," *Journal of Optim. Theory and Appl.*, vol. 190, no. 3, pp. 966–998, 2021.

[23] A. Levin and et al., "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Computer Vision and Pattern Recog (CVPR)*, 2009.

[24] 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 *Proc. 8th Int. Conf. Computer Vision*, July 2001, vol. 2, pp. 416–423.