NeurIPS 2021

35th Conference on Neural Information Processing System

Washington University in St. Louis
James McKelvey School of Engineering

Los Alamos NATIONAL LABORATORY

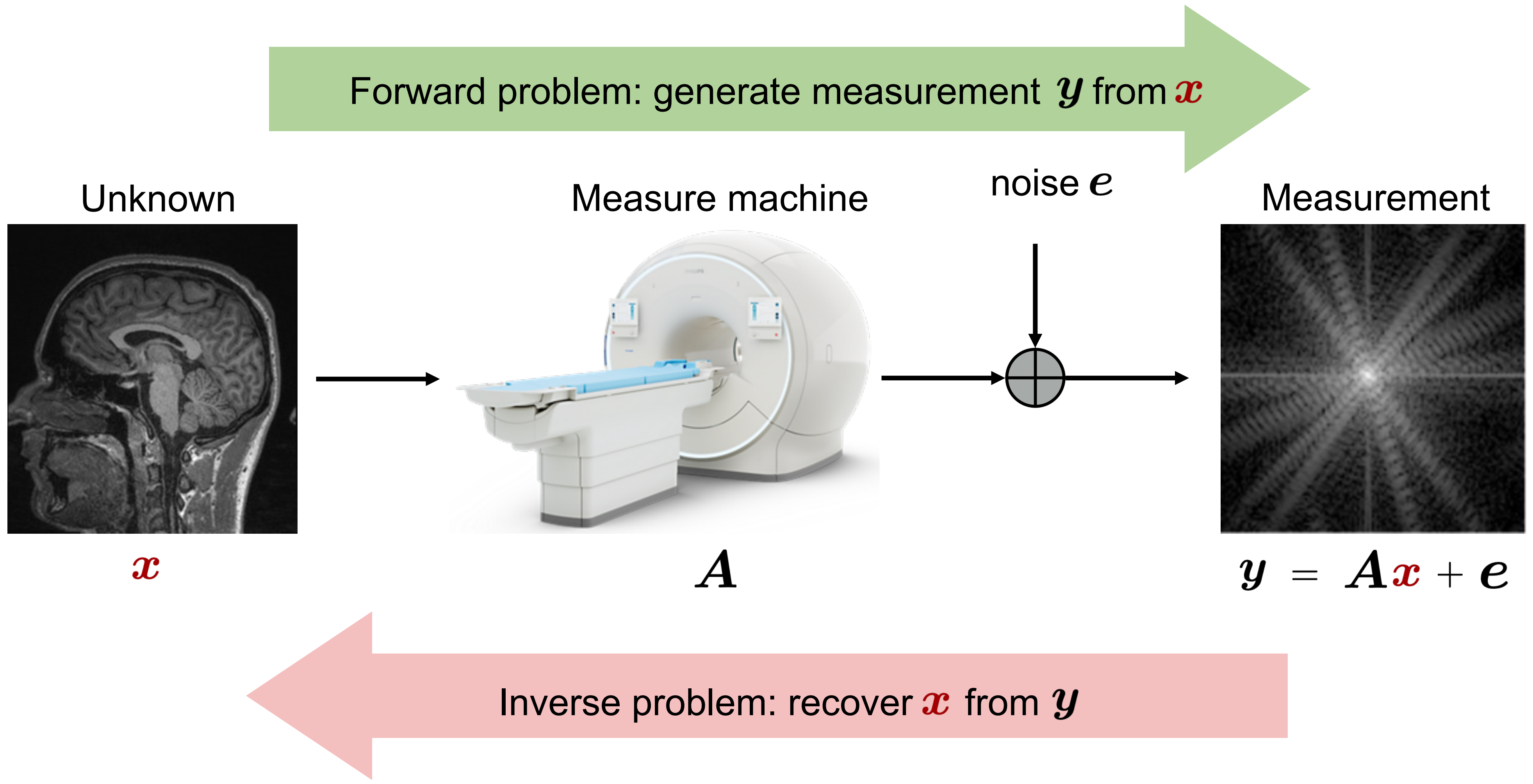
Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition

Jiaming Liu

Joint work with M. Salman Asif,
Brendt Wohlberg and Ulugbek S. Kamilov

Connect with me:
Email: jiaming.liu@wustl.edu

Most imaging problems can be formulated as inverse problems



Classic approach transforms the inverse problem to a regularized optimization

Inverse problem: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}$

Regularized optimization: $\underset{\boldsymbol{x}}{\arg\min} \; \{ \; g(\boldsymbol{x}) \; + \; h(\boldsymbol{x}) \; \}$

Data-fidelity　Regularizer

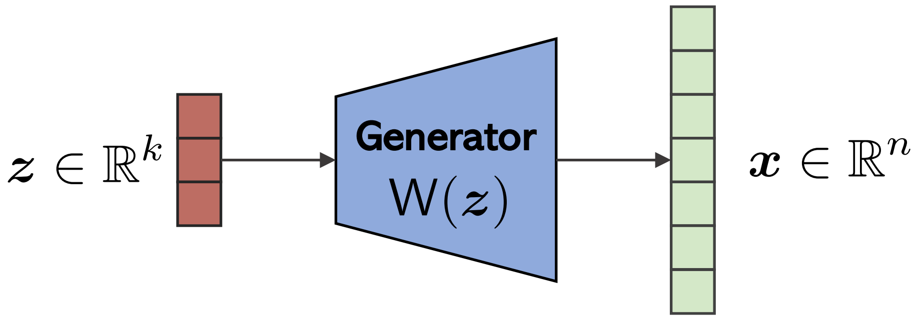
$$g(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2$$

Least-square loss

Total variation (TV)

Compressed sensing using generative models (CSGM)

Idea: **Pre-train** a generative model on a dataset of images



CSGM: Generative priors for solving inverse problems [Bora et al. 2017]

$$\min_{\boldsymbol{z}\in\mathbb{R}^k} = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\mathsf{W}(\boldsymbol{z})\|_2^2$$

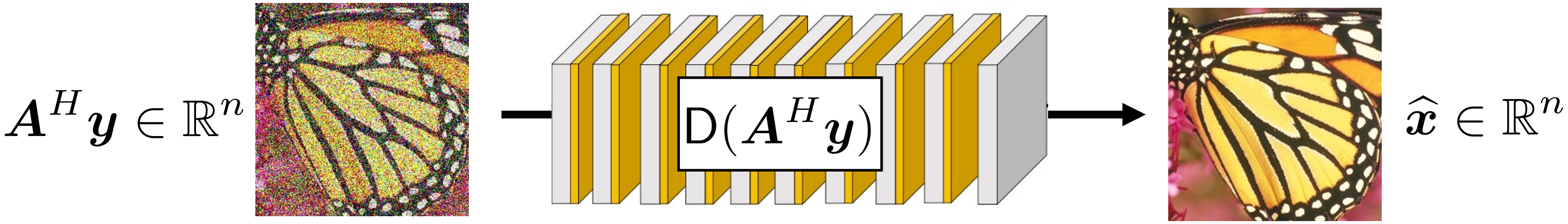
with an appropriate choice of A satisfying the *set-restricted eigenvalue condition* (S-REC) over the range of the generator

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}\boldsymbol{z}\|_2^2 \geq \mu\|\boldsymbol{x} - \boldsymbol{z}\|_2^2 - \eta \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$$

Where $\mu > 0$ and $\eta \geq 0$

Plug-and-play priors (PnP) and Regularization by denoising (RED)

Idea: **Pre-train** an artifact removing convolutional neural network (CNN) on a dataset of images [Liu et al. 2020]



PnP-PGM: replacing $\mathrm{prox}_{\gamma h}$ within the proximal gradient method (PGM) by denoiser or artifact removal CNN [Venkatakrishnan et al. 2013]

$\boldsymbol{x}^k \leftarrow \mathsf{D}(\boldsymbol{z}^k)$ Reduce image artifacts

$\boldsymbol{z}^{k-1} \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla g(\boldsymbol{x}^{k-1})$ Improve data fit

PnP-PGM can be rewritten within one line with a transfer operator T

$$\boldsymbol{x}^k = \mathsf{T}(\boldsymbol{x}^{k-1}) \qquad \text{with} \qquad \mathsf{T} := \mathsf{D}(\mathsf{I} - \gamma \nabla g)$$

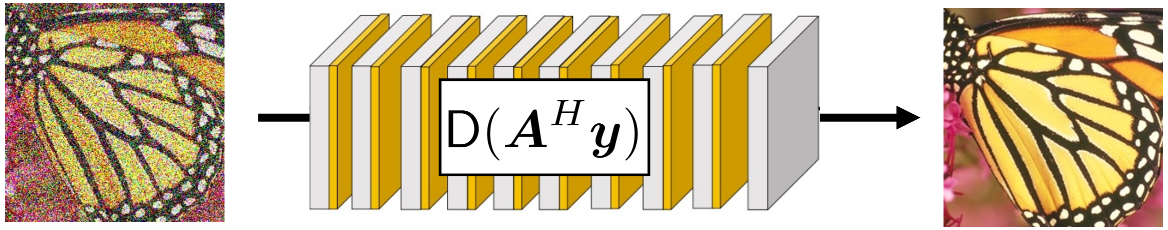
When the residual of D is Lipschitz continuous, PnP-PGM converges to a point in the fixed-point set of the operator T [Ryu et al. 2019]

$$\mathsf{Fix}(\mathsf{T}) := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \mathsf{T}(\boldsymbol{x})\}$$

Plug-and-play priors (PnP) and Regularization by denoising (RED)

Idea: **Pre-train** an artifact removing convolutional neural network (CNN) on a dataset of images [Liu et al. 2020]

$\boldsymbol{A}^H\boldsymbol{y} \in \mathbb{R}^n$



$\widehat{\boldsymbol{x}} \in \mathbb{R}^n$

SD-RED: the steepest descent variant of RED cam ne summarized as [Romano et al. 2017]

$$\boldsymbol{x}^k = \boldsymbol{x}^{k-1} - \gamma \mathsf{G}(\boldsymbol{x}^{k-1}) \qquad \text{with} \qquad \mathsf{G}(\boldsymbol{x}) = \nabla g(\boldsymbol{x}) + \tau(\boldsymbol{x} - \mathsf{D}(\boldsymbol{x}))$$

"gradient" descent

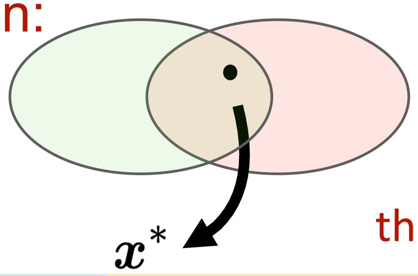
Improve data fit

Reduce image artifact

Optimal condition: suppose there exists a vector that satisfy

$$\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{D}) \quad \text{with} \quad \mathsf{Zer}(\nabla g) := \{\boldsymbol{x} \in \mathbb{R}^n : \nabla g = \boldsymbol{0}\} \quad \mathsf{Fix}(\mathsf{D}) := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \mathsf{D}(\boldsymbol{x})\}$$

Intuitive explanation:



$$\Longrightarrow \boldsymbol{0} = \nabla g(\boldsymbol{x}^*) + \tau(\boldsymbol{x}^* - \mathsf{D}(\boldsymbol{x}^*))$$

the algorithm can converge to the fixed-point $\boldsymbol{x}^*$

Recovery Analysis for PnP and RED

Latent space generative model

$$\min_{\boldsymbol{z}\in\mathbb{R}^k} = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\mathsf{W}(\boldsymbol{z})\|_2^2$$

- Pre-trained generative model.
- Domain specific, can recover images on the range of generative model.
- Nonconvex projections onto the range of a generative model.

PnP/RED

$$\boldsymbol{x}^k = \mathsf{T}(\boldsymbol{x}^{k-1}) \quad \text{with} \quad \mathsf{T} := \mathsf{D}(\mathsf{I} - \gamma\nabla g)$$

$$\boldsymbol{x}^k = \boldsymbol{x}^{k-1} - \gamma\mathsf{G}(\boldsymbol{x}^{k-1}) \quad \text{with} \quad \mathsf{G} = \nabla g + \tau(\mathsf{I} - \mathsf{D})$$

- Pre-trained denoiser or artifact remove operator.
- Leverage any off-the-shelf regularizer.
- Lack theoretical recovery guarantees available for CSGM.

Contributions of this paper:

- We first establish recovery bounds for PnP and address the relationship between the solutions of PnP and RED, under a set of sufficient conditions.
- Numerical analysis of PnP/RED and CSGM.

Recovery Analysis for PnP and RED: Numerical evaluation

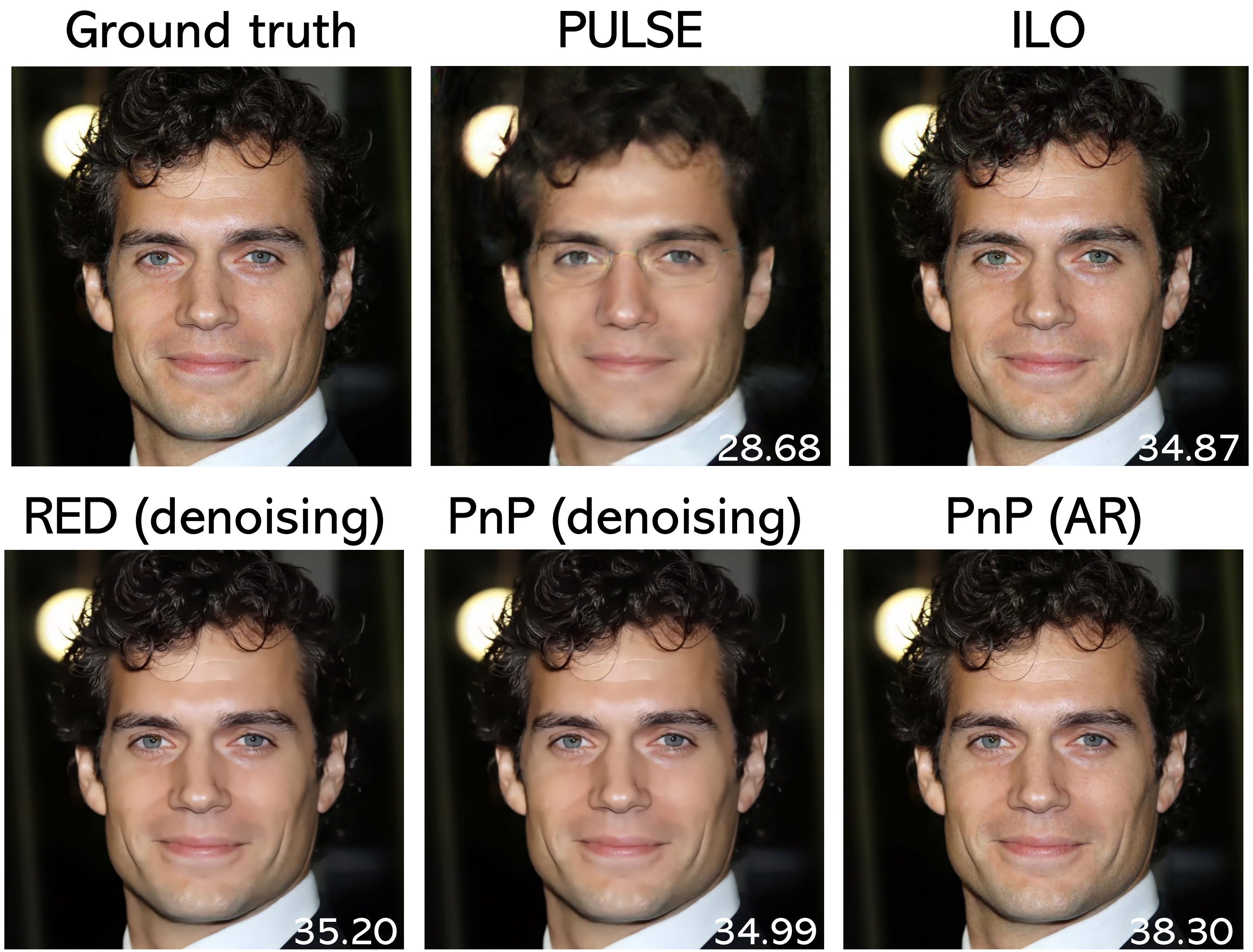


Table 3: Average PSNR (dB) values for several algorithms on test images from CelebA HQ.

Method \ CS Ratio	10%	20%	30%
TV	32.13	35.24	37.41
PULSE [34]	27.45	29.98	33.06
ILO [35]	36.15	40.98	43.46
RED (denoising)	35.46	41.59	45.65
PnP (denoising)	35.61	41.51	45.71
PnP (AR)	**39.19**	**44.20**	**48.66**

- AR is better than AWGN (expected!)
- RED is nearly equivalent of PnP (somehow surprising!)
- PnP(AR) is competitive with PULSE and ILO using StyleGan2 (surprising!)

Scaling strategy for PnP is adopted from [Xiao et al. 20]

Recovery Analysis for PnP and RED: Main results

Theorem 1. *Run PnP-PGM for $t \geq 1$ iterations under Assumptions 1-2 for the regularized problem with no noise and $\boldsymbol{x}^* \in \mathsf{Zer}(\mathsf{R})$. Then, the sequence $\boldsymbol{x}^t$ generated by PnP-PGM satisfies*

$$\|\boldsymbol{x}^t - \boldsymbol{x}^*\|_2 \leq c\|\boldsymbol{x}^{t-1} - \boldsymbol{x}^*\|_2 \leq c^t\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2 \, , \tag{1}$$

where $\boldsymbol{x}^0 \in \mathbb{R}^n$ and $c := (1+\alpha)\max\{|1-\gamma\mu|, |1-\gamma\lambda|\}$ with $\lambda := \lambda_{\mathsf{max}}(\boldsymbol{A}^\mathsf{T}\boldsymbol{A})$.

Theorem 1 extends the theoretical analysis of PnP in [Ryu et al. 2019] by showing convergence to the true solution x∗ instead of the fixed points Fix(T) of T. Note that the condition x0 in the range of D can be easily enforced by simply passing any initial image through the operator D.

Recovery Analysis for PnP and RED: Main results

Theorem 2. *Run PnP-PGM for $t \geq 1$ iterations under Assumptions 1-2 for the regularized problem with $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{e} \in \mathbb{R}^m$. Then, the sequence $\boldsymbol{x}^t$ generated by PnP-PGM satisfies*

$$\|\boldsymbol{x}^t - \boldsymbol{x}^*\|_2 \leq c\|\boldsymbol{x}^{t-1} - \boldsymbol{x}^*\|_2 + \varepsilon \leq c^t\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2 + \frac{\varepsilon(1-c^t)}{(1-c)}\,, \tag{2}$$

where

$$\varepsilon := (1+c)\left[\left(1+2\sqrt{\lambda/\mu}\right)\|\boldsymbol{x}^* - \mathsf{proj}_{\mathsf{Zer(R)}}(\boldsymbol{x}^*)\|_2 + 2/\sqrt{\mu}\|\boldsymbol{e}\|_2 + \delta(1+1/\alpha)\right] \tag{3}$$

and $c := (1+\alpha)\,\mathsf{max}\{|1-\gamma\mu|, |1-\gamma\lambda|\}$ *with* $\lambda := \lambda_{\mathsf{max}}(\boldsymbol{A}^\mathsf{T}\boldsymbol{A})$.

Theorem 2 extends Theorem 1 by allowing x∗ to be outside of Fix(D) and extends the analysis in [Bora et al. 2017] by considering operators D that do not necessarily project onto the range of a generative model.

Recovery Analysis for PnP and RED: Main results

Theorem 2. *Run PnP-PGM for $t \geq 1$ iterations under Assumptions 1-2 for the regularized problem with $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{e} \in \mathbb{R}^m$. Then, the sequence $\boldsymbol{x}^t$ generated by PnP-PGM satisfies*

$$\|\boldsymbol{x}^t - \boldsymbol{x}^*\|_2 \leq c\|\boldsymbol{x}^{t-1} - \boldsymbol{x}^*\|_2 + \varepsilon \leq c^t\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2 + \frac{\varepsilon(1-c^t)}{(1-c)}\ , \tag{2}$$

where

$$\varepsilon := (1+c)\left[\left(1+2\sqrt{\lambda/\mu}\right)\|\boldsymbol{x}^* - \mathsf{proj}_{\mathsf{Zer(R)}}(\boldsymbol{x}^*)\|_2 + 2/\sqrt{\mu}\|\boldsymbol{e}\|_2 + \delta(1+1/\alpha)\right] \tag{3}$$

and $c := (1+\alpha)\,\mathsf{max}\{|1-\gamma\mu|, |1-\gamma\lambda|\}$ *with* $\lambda := \lambda_{\mathsf{max}}(\boldsymbol{A}^\mathsf{T}\boldsymbol{A})$.

Theorem 3. *Suppose that Assumptions 1-3 are satisfied and that* $\mathsf{Zer}(\nabla g) \cap \mathsf{Zer}(\mathsf{R}) \neq \varnothing$, *then PnP and RED have the same set of solutions:* $\mathsf{Fix}(\mathsf{T}) = \mathsf{Zer}(\mathsf{G})$.

Proof Technology

Assumption 1. *The residual* $\mathsf{R} := \mathsf{I} - \mathsf{D}$ *of the operator* D *is bounded by* δ *and Lipschitz continuous with constant* $\alpha > 0$.

$$\|\mathsf{R}(\boldsymbol{x})\|_2 \leq \delta \quad \textit{and} \quad \|\mathsf{R}(\boldsymbol{x}) - \mathsf{R}(\boldsymbol{z})\|_2 \leq \alpha\|\boldsymbol{x} - \boldsymbol{z}\|_2, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n \ .$$

Assumption 2. *The measurement operator* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *satisfies the set-restricted eigenvalue condition (S-REC) over* $\mathsf{Im}(\mathsf{D}) \subseteq \mathbb{R}^n$ *with* $\mu > 0$, *which can be written as*

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}\boldsymbol{z}\|_2^2 \geq \mu\|\boldsymbol{x} - \boldsymbol{z}\|_2^2, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathsf{Im}(\mathsf{D}) \ .$$

Assumption 3. *The denoiser* D *is nonexpansive*

$$\|\mathsf{D}(\boldsymbol{x}) - \mathsf{D}(\boldsymbol{z})\|_2 \leq \|\boldsymbol{x} - \boldsymbol{z}\|_2 \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n \ .$$

- The α-Lipschitz continuity of R can be enforced by using any of the recent techniques for training Lipschitz constrained deep neural nets [Ryu et al. 2019]

- The S-REC in Assumption 2 was adopted from the corresponding assumption for CSGM [Bora et al. 2017]

Conclusion and Broader impact

- ✓ We address the theoretical gap between PnP/RED and CSGM for solving inverse problems
- ✓ Our theoretical results provide a new type of convergence for PnP-PGM by showing convergence relative to the true solution.
- ✓ We show full equivalence of PnP and RED under some explicit conditions on the inverse problem
- ✓ We provide additional evidence on the suboptimality of AWGN denoisers compared to artifact-removal operators
- ✓ Potential applications (e.g., computational microscopy, computerized tomography, medical imaging, and image restoration).

References

1. Bora, A., Jalal, A., Price, E. and Dimakis, A.G., 2017, July. Compressed sensing using generative models. In *International Conference on Machine Learning* (pp. 537-546). PMLR.
2. Liu, J., Sun, Y., Eldeniz, C., Gan, W., An, H. and Kamilov, U.S., 2020. Rare: Image reconstruction using deep priors learned without groundtruth. IEEE Journal of Selected Topics in Signal Processing, 14(6), pp.1088-1099.
3. Venkatakrishnan, S.V., Bouman, C.A. and Wohlberg, B., 2013, December. Plug-and-play priors for model based reconstruction. In 2013 IEEE Global Conference on Signal and Information Processing (pp. 945-948). IEEE.
4. Ryu, E., Liu, J., Wang, S., Chen, X., Wang, Z. and Yin, W., 2019, May. Plug-and-play methods provably converge with properly trained denoisers. In International Conference on Machine Learning (pp. 5546-5557). PMLR.
5. Romano, Y., Elad, M. and Milanfar, P., 2017. The little engine that could: Regularization by denoising (RED). SIAM Journal on Imaging Sciences, 10(4), pp.1804-1844.
6. Xu, X., Liu, J., Sun, Y., Wohlberg, B. and Kamilov, U.S., 2020, November. Boosting the Performance of Plug-and-Play Priors via Denoiser Scaling. In 2020 54th Asilomar Conference on Signals, Systems, and Computers (pp. 1305-1312). IEEE.

Thanks & Follow us!

Paper

SCAN ME

Code

SCAN ME

Twitter

SCAN ME