Model-based Deep Learning for Computational Imaging

Xiaojian Xu
Advisor: Ulugbek Kamilov

Department of Computer Science and Engineering
Washington University in St. Louis

PhD Dissertation Defense – July 28, 2022

Imaging is everywhere

Washington University in St. Louis

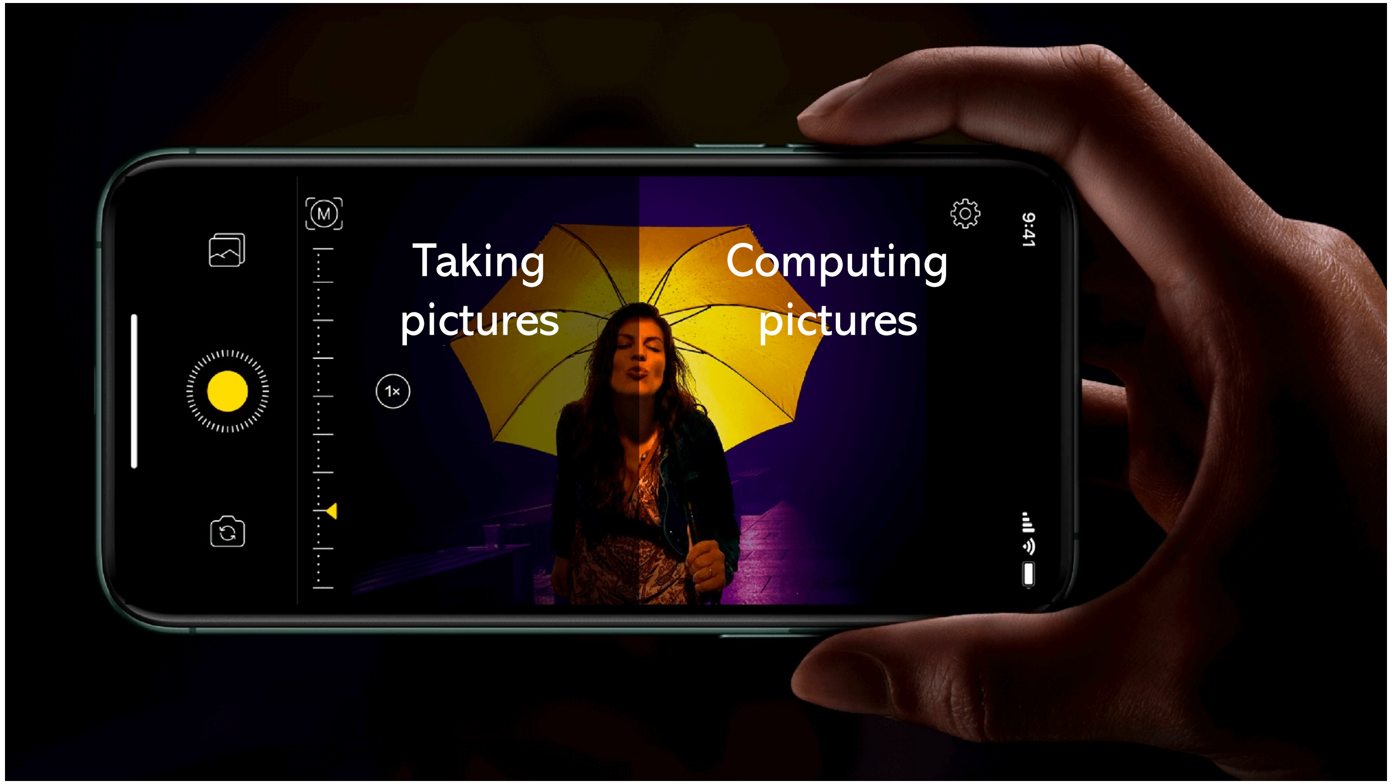
Microscopy

Photography

Astronomy

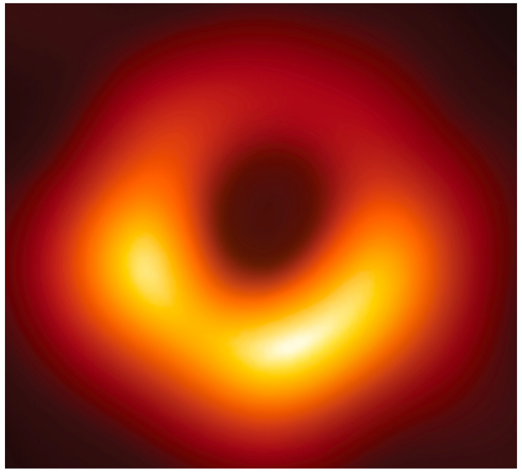
We are observing a big shift in imaging where we are no longer taking pictures but computing pictures

Washington University in St. Louis



In many computational imaging applications, we don't have direct access to the things that we want

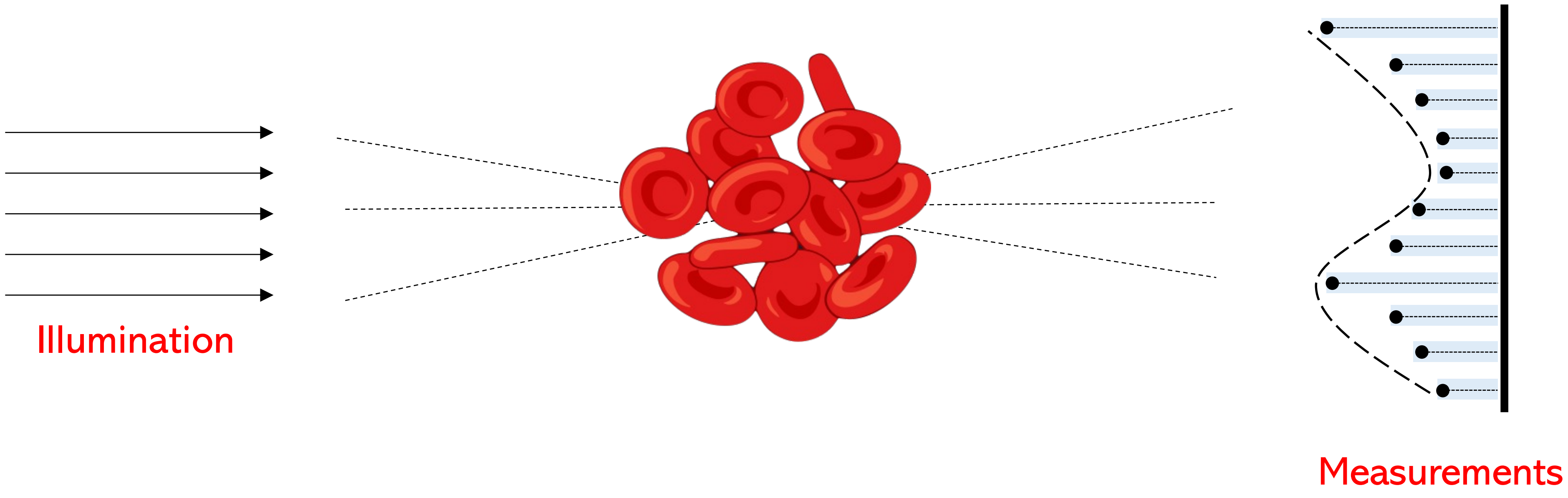
Imaging skeleton



Imaging blackhole

Imaging infant

An illustration of computational imaging in microscopy



An illustration of computational imaging in microscopy

Washington University in St.Louis

Imaging problems can be formulated as inverse problems

Signal acquisition procedure

Forward problem: generate $\boldsymbol{y}$ from $\boldsymbol{x}$

$\boldsymbol{x}$

unknown image

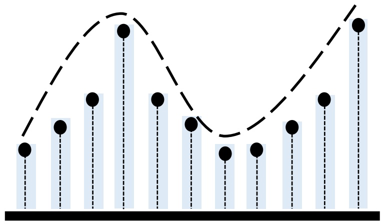
$\boldsymbol{A}$

forward model

$\boldsymbol{e}$

noise

$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}$



measurements

Inverse problem: recover $\boldsymbol{x}$ from $\boldsymbol{y}$

Image reconstruction procedure

Imaging inverse problems is challenging

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

What makes imaging inverse problems challenging?

- Solution is not unique
- Data is noisy
- Signals can be high-dimensional

Solution 1: Traditionally, inverse problems have been solved using model-based optimization approaches

Washington University in St. Louis

Inverse problem

$y = Ax + e$

- Image reconstruction problem is formulated as an optimization task

Data-fidelity Regularizer/Prior

$$\widehat{x} = \arg\min_{x} \{ d(x) + r(x) \}$$

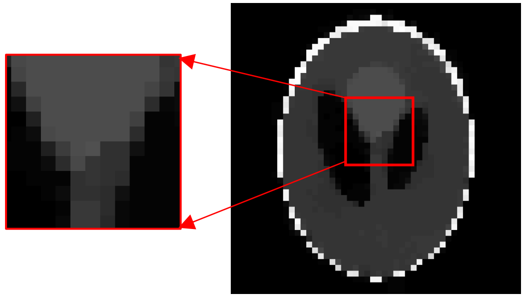
Example:

L_2 loss Total variation(TV)

$$\widehat{x} = \arg\min_{x} \{ \frac{1}{2} \|y - Ax\|_2^2 + \mathrm{TV}(x) \}$$

model-based optimization

Regularized by TV



- Pros
  - d(x): Guarantee the data-consistency

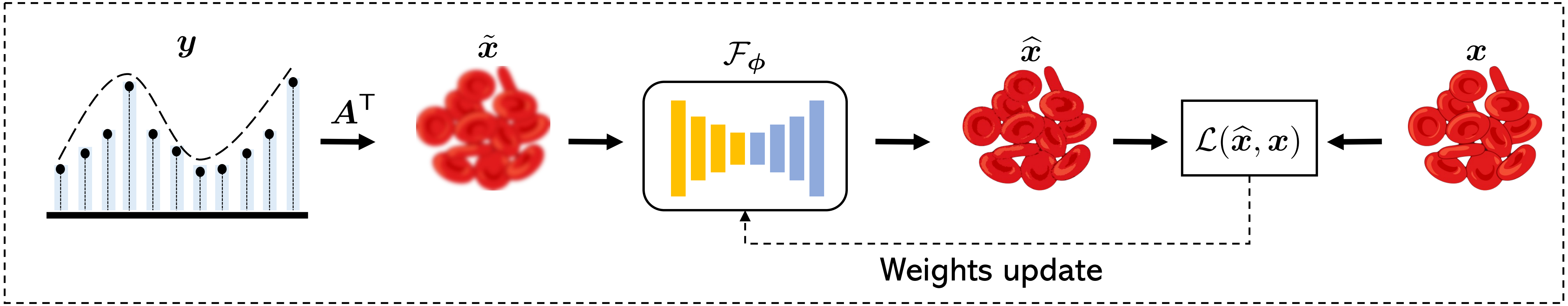
- Cons
  - r(x): Need manual design of priors

Solution 2: Deep learning is a modern approach for image reconstruction in computational imaging

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

- Deep neural networks provides a state-of-the-art tool for representing and enforcing sophisticated image priors.

End-to-end deep learning



- Pros
  - ✪ Do not need explicit priors
  - ✪ Do not need model information

- Cons
  - ✪ No data consistency guarantee
  - ✪ Sensitive to outliers

Solution 3: Modern image reconstruction problems require fast and reliable methods that can integrate model-based optimization and deep learning

Inverse problem

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}$$

Model

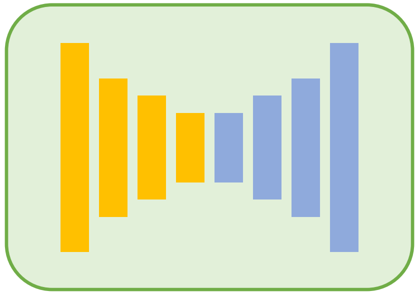
$\boldsymbol{A}$

Data consistency guarantee

$+$

Deep learning

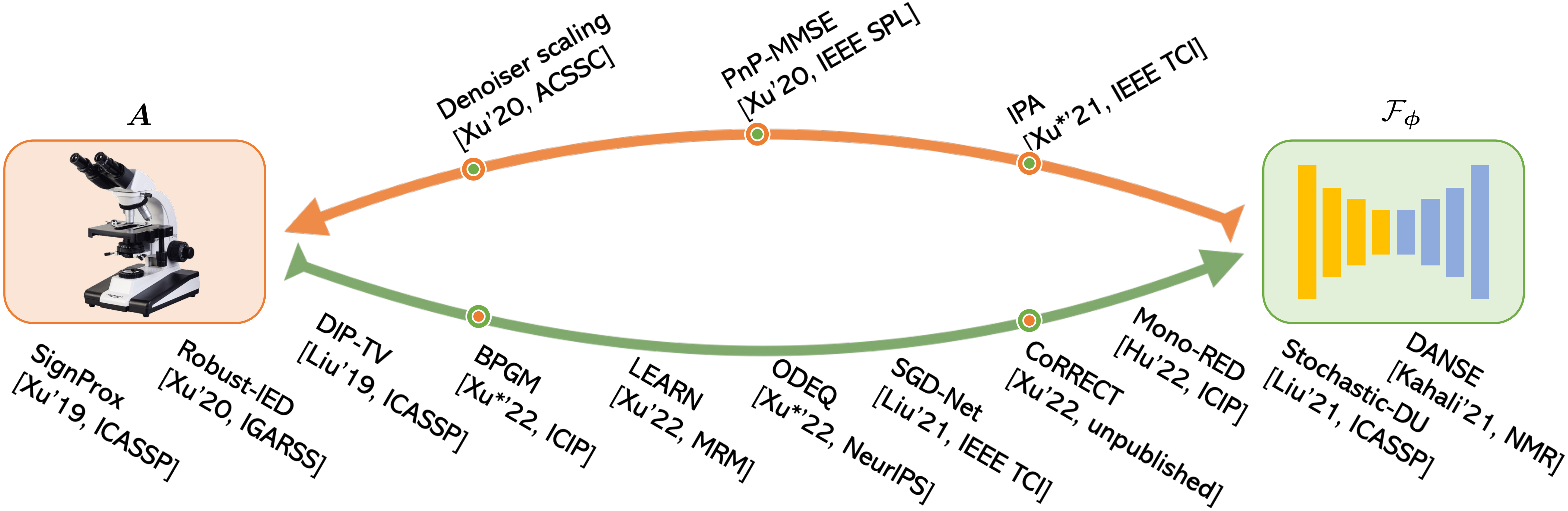
$\mathcal{F}_\phi$



Advanced prior representation

Most of the work I have collaborated and led focuses on the integration of imaging model and deep learning

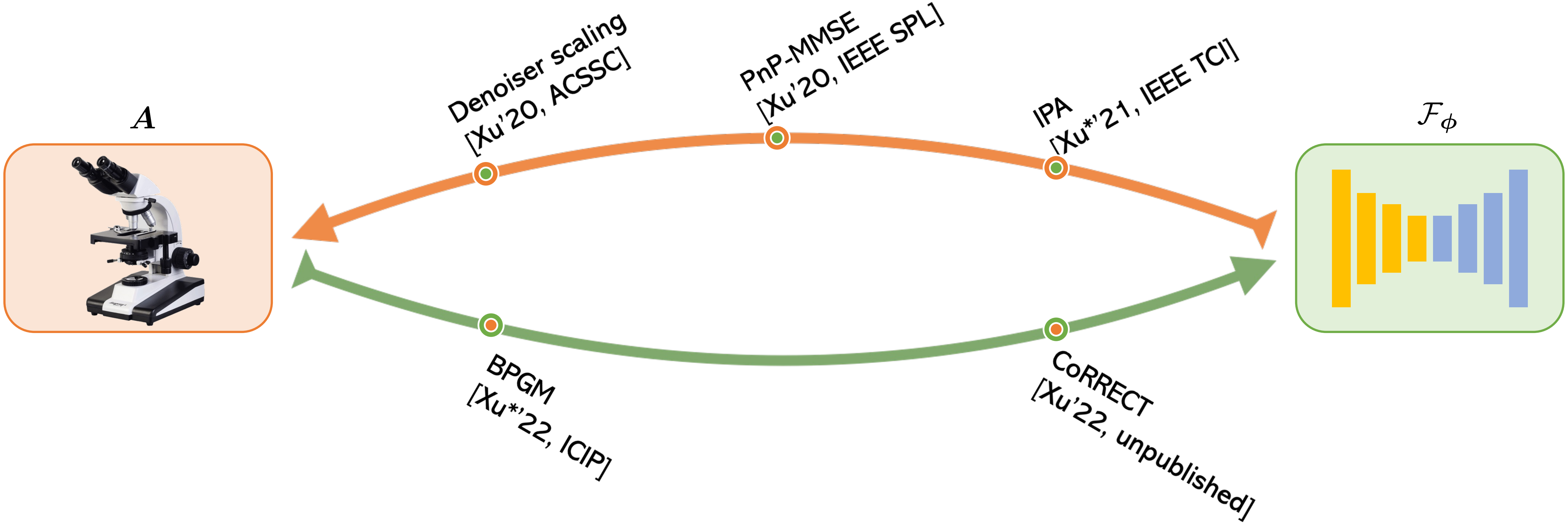
Inverse problem
$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}$$



* Co-first author

The dissertation is constituted by the following work

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



* Co-first author

My dissertation is constituted by the following work

Understanding statistical interpretation of plug-and-play priors

- Boosting the Performance of Plug-and-Play Priors via Denoiser Scaling. ACSSC, 2020
- Provable Convergence of Plug-and-Play Priors with MMSE Denoisers. IEEE SPL, 2020

Adapting plug-and-play priors to large-scale problems

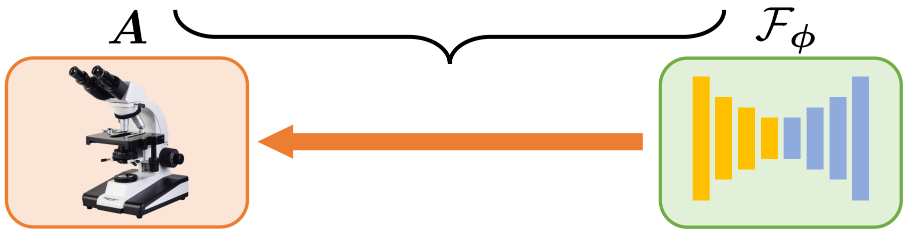
- Scalable Plug-and-Play ADMM With Convergence Guarantees. IEEE TCI, 2021

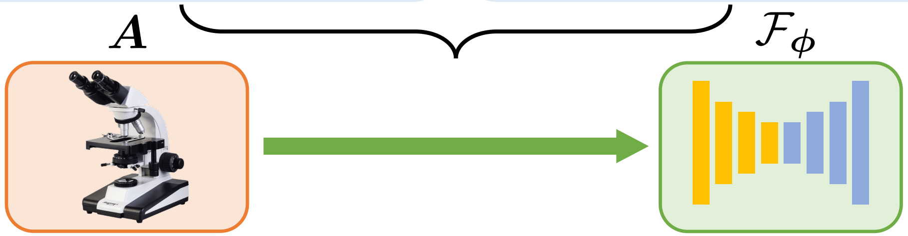
Extending plug-and-play priors to the non-Euclidean setting

- Bregman Plug-and-Play Priors. ICIP, 2022

Applying model-based deep learning algorithms to MRI

- CoRRECT: A Deep Unfolding Framework for Motion-Corrected Quantitative R_2^* Recovery, unpublished, 2022





My dissertation is constituted by the following work

Washington University in St. Louis

Understanding statistical interpretation of plug-and-play priors

- Boosting the Performance of Plug-and-Play Priors via Denoiser Scaling. ACSSC, 2020
- Provable Convergence of Plug-and-Play Priors with MMSE Denoisers. IEEE SPL, 2020

Adapting plug-and-play priors to large-scale problems

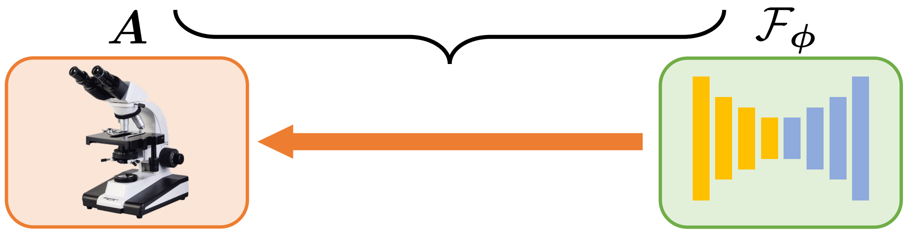
- Scalable Plug-and-Play ADMM With Convergence Guarantees. IEEE TCI, 2021

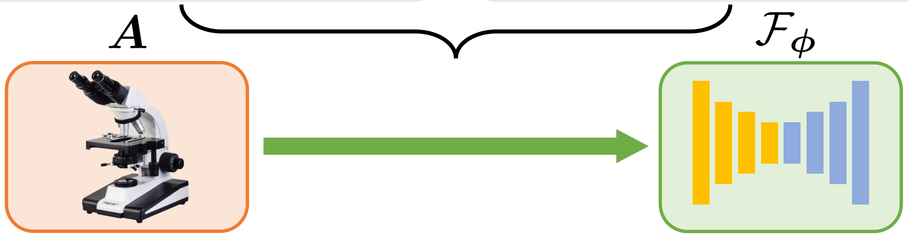
Extending plug-and-play priors to the non-Euclidean setting

- Bregman Plug-and-Play Priors. ICIP, 2022

Applying model-based deep learning algorithms to MRI

- CoRRECT: A Deep Unfolding Framework for Motion-Corrected Quantitative R_2^* Recovery, unpublished, 2022





Proximal methods use proximal operators to solve nonsmooth problems

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

- Recall our regularized optimization task

Data-fidelity Regularizer/Prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{ \ d(\boldsymbol{x}) \ + \ r(\boldsymbol{x}) \ \}$$

- Solutions: proximal methods
  - (Accelerated) Proximal gradient method (PGM/APGM) [Bruck'77, Beck'09, Nesterov'13]
  - Alternating direction method of multipliers (ADMM) [Boyd'10]

PGM/APGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{s}^{k-1} - \gamma \nabla d(\boldsymbol{s}^{k-1})$$
$$\boldsymbol{x}^k \leftarrow \mathrm{prox}_{\gamma r}(\boldsymbol{z}^k)$$
$$\boldsymbol{s}^k \leftarrow \boldsymbol{x}^k + ((q_{k-1} - 1)/q_k)(\boldsymbol{x}^k - \boldsymbol{x}^{k-1})$$

ADMM

$$\boldsymbol{z}^k \leftarrow \mathrm{prox}_{\gamma d}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$$
$$\boldsymbol{x}^k \leftarrow \mathrm{prox}_{\gamma r}(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$$
$$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$$

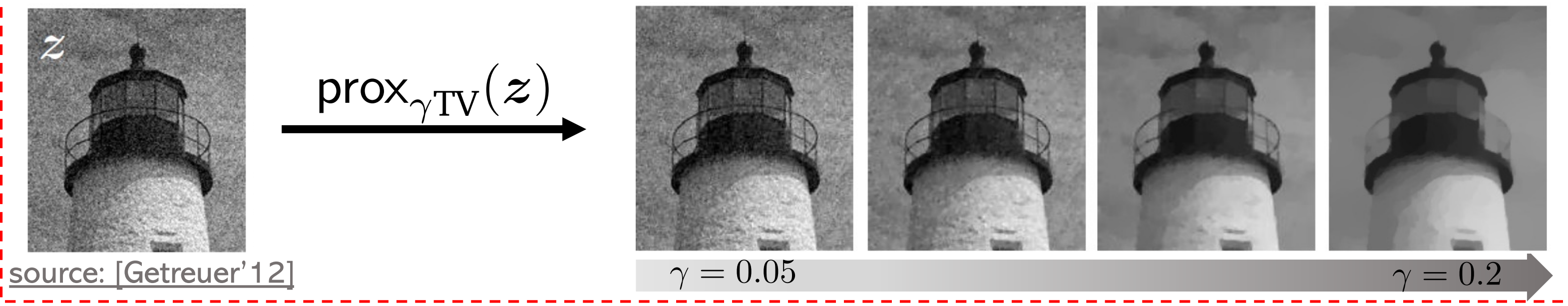
PGM and ADMM use proximal operators to solve nonsmooth problems

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{ \underbrace{d(\boldsymbol{x})}_{\text{data-fidelity}} + \underbrace{r(\boldsymbol{x})}_{\text{prior}} \}$$

PGM/APGM		ADMM
$\boldsymbol{z}^k \leftarrow \boldsymbol{s}^{k-1} - \gamma \nabla d(\boldsymbol{s}^{k-1})$	increase data consistency	$\boldsymbol{z}^k \leftarrow \mathsf{prox}_{\gamma d}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$
$\boldsymbol{x}^k \leftarrow \mathsf{prox}_{\gamma r}(\boldsymbol{z}^k)$	denoising	$\boldsymbol{x}^k \leftarrow \mathsf{prox}_{\gamma r}(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$
$\boldsymbol{s}^k \leftarrow \boldsymbol{x}^k + ((q_{k-1} - 1)/q_k)(\boldsymbol{x}^k - \boldsymbol{x}^{k-1})$		$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$

$$\mathsf{prox}_{\gamma r}(\boldsymbol{z}) = \arg\min_{\boldsymbol{x}} \left\{ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|^2 + \gamma r(\boldsymbol{x}) \right\}$$

Prox: Proximal map is an image denoiser for additive white Gaussian noise (AWGN)



source: [Getreuer'12]

PGM and ADMM use proximal operators to solve nonsmooth problems

Washington University in St. Louis

data-fidelity prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}}\{d(\boldsymbol{x}) + r(\boldsymbol{x})\}$$

PGM/APGM		ADMM
$\boldsymbol{z}^k \leftarrow \boldsymbol{s}^{k-1} - \gamma \nabla d(\boldsymbol{s}^{k-1})$	increase data consistency	$\boldsymbol{z}^k \leftarrow \mathsf{prox}_{\gamma d}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$
$\boldsymbol{x}^k \leftarrow \mathsf{prox}_{\gamma r}(\boldsymbol{z}^k)$	denoising	$\boldsymbol{x}^k \leftarrow \mathsf{prox}_{\gamma r}(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$
$\boldsymbol{s}^k \leftarrow \boldsymbol{x}^k + ((q_{k-1} - 1)/q_k)(\boldsymbol{x}^k - \boldsymbol{x}^{k-1})$		$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$

Plug-and-Play Prior (PnP) [Venkat, Bouman, Wohlberg'13]: replace proximal map with AWGN denoisers

$$\mathsf{prox}_{\gamma r} \Rightarrow \mathsf{D}_\sigma$$

where $\sigma \geq 0$ refers to denoising strength.

PnP methods leverage the power of deep learning by using the deep-learned denoisers inside the model-based optimization

Washington University in St. Louis

PnP-PGM/PnP-APGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{s}^{k-1} - \gamma \nabla d(\boldsymbol{s}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k)$$

$$\boldsymbol{s}^k \leftarrow \boldsymbol{x}^k + ((q_{k-1} - 1)/q_k)(\boldsymbol{x}^k - \boldsymbol{x}^{k-1})$$

any off-the-shelf AWGN image denoiser

PnP-ADMM

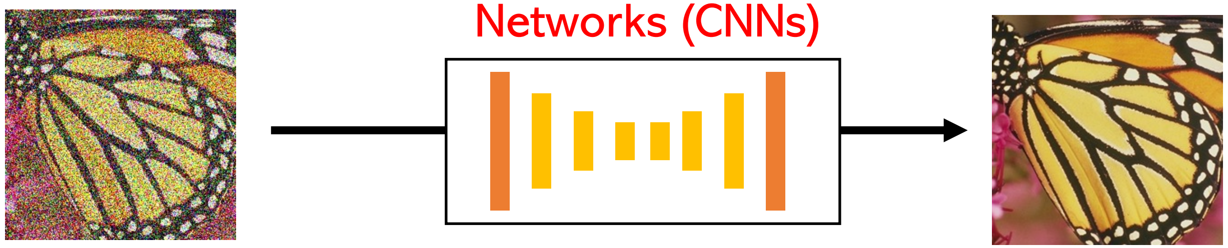
$$\boldsymbol{z}^k \leftarrow \mathsf{prox}_{\gamma d}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$$

$$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$$

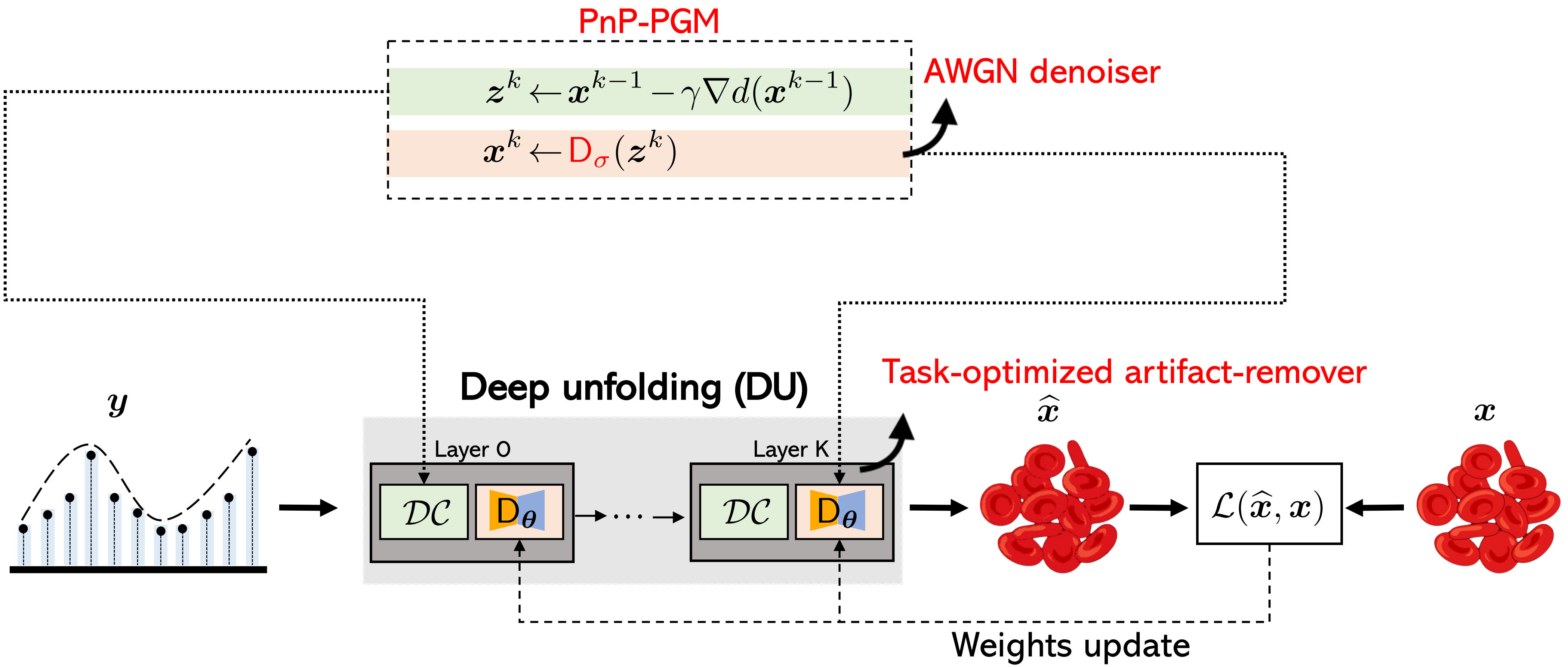
Example: D_σ could be a neural network

Convolutional Neural Networks (CNNs)



PnP therefore leverages the power of deep learning inside the model-based optimization!

PnP can be extended to deep unfolding framework that interprets the iterations of an image recovery algorithm as layers of a neural network



$\mathcal{DC}$: Data consistency module

D_θ : Deep learning module

Work 1:
Understanding statistical interpretation of plug-and-play priors

- Provable Convergence of Plug-and-Play Priors with MMSE Denoisers. IEEE SPL, 2020

Convergence analysis of PnP approaches is challenging because they lost the interpretation as an optimization problem for an arbitrary denoiser

PGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathrm{prox}_{\gamma r}(\boldsymbol{z}^k)$$

objective function

prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{d(\boldsymbol{x}) + r(\boldsymbol{x})\}$$

proximal map — prox $\Rightarrow$ arbitrary denoiser — D_σ

PnP-PGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k)$$

objective function

prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{d(\boldsymbol{x}) + \ ? \ \}$$

For most off-the-shelf denoisers, it is impossible to write them as explicit regularizers!

Convergence analysis of PnP approaches is challenging because they lost the interpretation as an optimization problem for an arbitrary denoiser

PGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathsf{prox}_{\gamma r}(\boldsymbol{z}^k)$$

objective function

prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{d(\boldsymbol{x}) + r(\boldsymbol{x})\}$$

proximal map → arbitrary denoiser

prox $\Rightarrow$ D_σ

PnP-PGM

$$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$$

$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k) \qquad \mathsf{D}_\sigma = \text{MMSE}$$

objective function

prior

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{d(\boldsymbol{x}) + \ ? \ \}$$

- We established the first convergence guarantee of PnP-PGM algorithm for MMSE denoisers to give insights for other denoisers

Convergence analysis of PnP approaches is challenging because they lost the interpretation as an optimization problem for an arbitrary denoiser

Washington University in St. Louis

PGM

$$z^k \leftarrow x^{k-1} - \gamma \nabla d(x^{k-1})$$
$$x^k \leftarrow \mathrm{prox}_{\gamma r}(z^k)$$

proximal map — prox $\Rightarrow$ arbitrary denoiser — D_σ

PnP-PGM

$$z^k \leftarrow x^{k-1} - \gamma \nabla d(x^{k-1})$$
$$x^k \leftarrow \mathsf{D}_\sigma(z^k) \qquad \mathsf{D}_\sigma = \text{MMSE}$$

objective function (PGM):

$$\widehat{x} = \arg\min_{x} \{ d(x) + \underbrace{r(x)}_{\text{prior}} \}$$

objective function (PnP-PGM):

$$\widehat{x} = \arg\min_{x} \{ \underbrace{d(x) + r_{\text{mmse}}(x)}_{f(x)} \}$$

prior: $r_{\text{mmse}}(x)$

- We established the first convergence guarantee of PnP-PGM algorithm for MMSE denoisers to give insights for other denoisers

Insights into the performance of PnP for MMSE denoisers are meaningful as many denoisers are interpreted as approximate MMSE denoisers

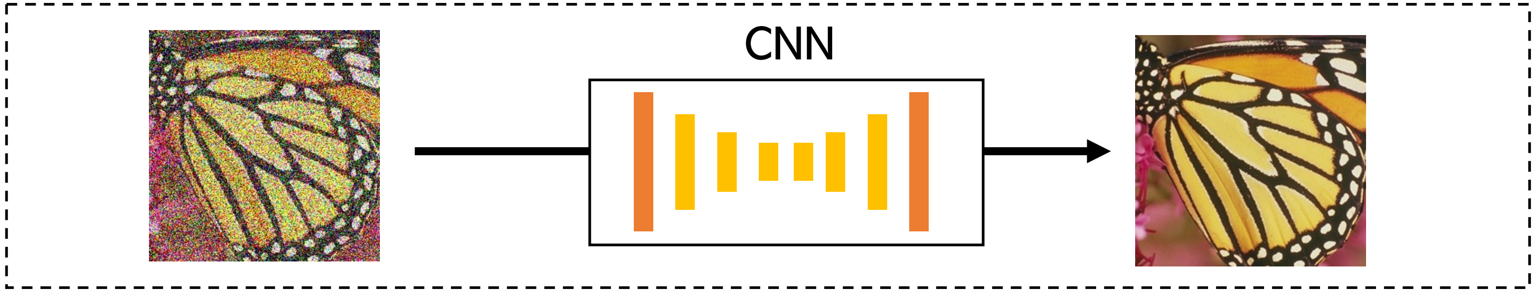
- Noise corruption scenario

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n} \quad \text{where} \quad \boldsymbol{x} \sim p_{\boldsymbol{x}}, \quad \boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \tau \mathbf{I})\,.$$

- MMSE denoiser

$$\mathrm{D}_{\sigma}(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}] = \int_{\mathbb{R}^n} \boldsymbol{x} p_{\boldsymbol{x}|\boldsymbol{z}}(\boldsymbol{x}|\boldsymbol{z}) \mathrm{d}\boldsymbol{x}$$

- Many denoisers (pre-trained CNNs, NLM, BM3D) can be interpreted as approximate or empirical MMSE denoisers



A MMSE denoiser can be established as a proximal map of some regularizer function r_{mmse}

MMSE denoiser

$D_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}]$

- Assumptions
  - The function d is continuously differentiable and has a Lipschitz continuous gradient with constant $L>0$
  - The prior p_x is non-degenerate over R^n
  - The function f has a finite infimum $f^* > -\infty$

A MMSE denoiser can be established as a proximal operator of some regularizer function r_{mmse}

MMSE denoiser

$\mathsf{D}_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}]$

MMSE $\mathsf{D}_\sigma = \text{prox}_{\gamma r_{\text{mmse}}}$

$$r_{\text{mmse}}(\boldsymbol{x}) := \begin{cases} -\frac{1}{2\gamma}\|\boldsymbol{x} - \mathsf{D}_\sigma^{-1}(\boldsymbol{x})\|^2 + \frac{\sigma^2}{\gamma} r_\sigma(\mathsf{D}_\sigma^{-1}(\boldsymbol{x})) & \text{for } \boldsymbol{x} \in \text{Im}(\mathsf{D}_\sigma) \\ +\infty & \text{for } \boldsymbol{x} \notin \text{Im}(\mathsf{D}_\sigma) \end{cases} \quad \text{where} \quad r_\sigma(\cdot) := -\log(p_{\boldsymbol{z}}(\cdot))$$

PnP-PGM		PGM
$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$		$\boldsymbol{z}^k \leftarrow \boldsymbol{x}^{k-1} - \gamma \nabla d(\boldsymbol{x}^{k-1})$
$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k)$	MMSE $\mathsf{D}_\sigma = \text{prox}_{\gamma r_{\text{mmse}}}$ →	$\boldsymbol{x}^k \leftarrow \text{prox}_{\gamma r_{\text{mmse}}}(\boldsymbol{z}^k)$

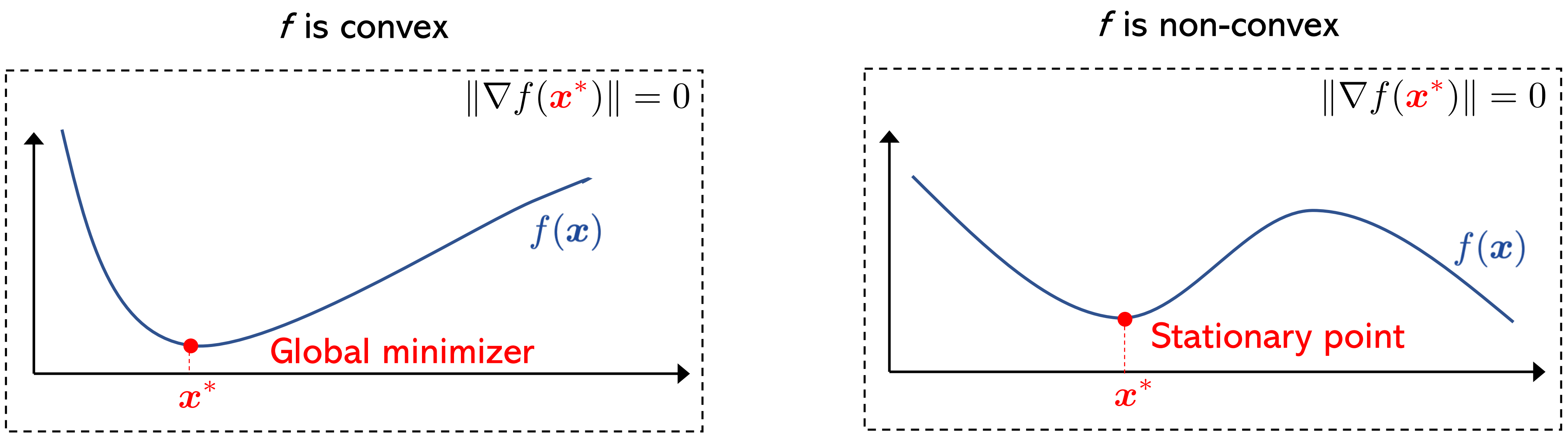
- PnP-PGM with MMSE denoiser is linked to the minimization of some global objective function $f(x)$

Objective function: $f(\boldsymbol{x}) = d(\boldsymbol{x}) + r_{\text{mmse}}(\boldsymbol{x})$

The iterates produced by PnP-PGM algorithm for MMSE denoisers converge to the stationary point of some possibly non-convex function

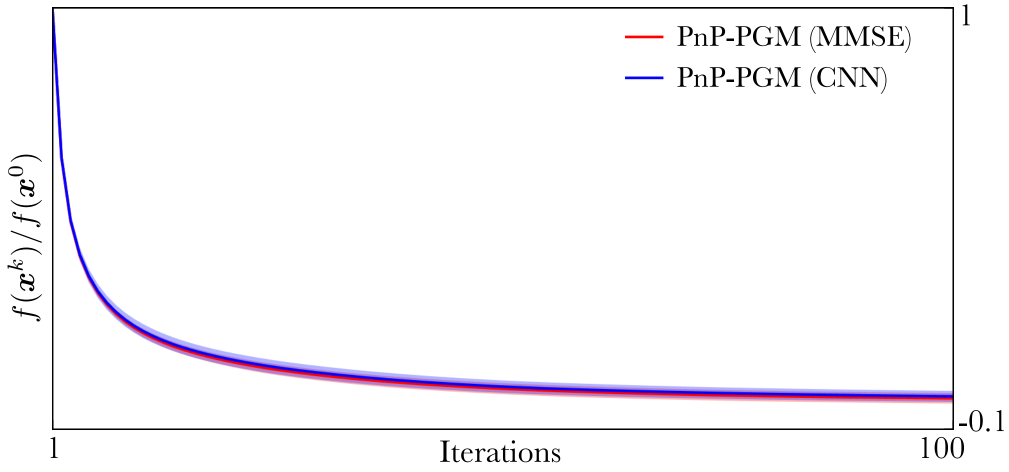
Objective function
$f(\boldsymbol{x}) = d(\boldsymbol{x}) + r_{\text{mmse}}(\boldsymbol{x})$

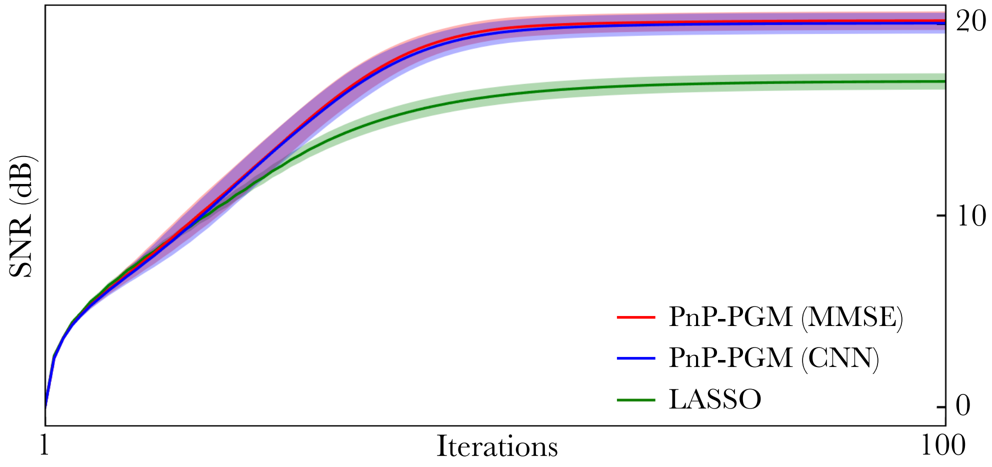
- **Theorem**: Run PnP-PGM with the MMSE denoiser with step size $0 < \gamma \leq 1/L$, then
  - $\{f(\boldsymbol{x}^k)\}_{k \geq 0}$ monotonically decreases
  - $\|\nabla f(\boldsymbol{x}^k)\| \to 0$ as $k \to \infty$



Performance evaluation on compressive sensing validates our theoretical analysis and shows excellent agreement between MMSE and CNN denoisers

- Performance evaluation on compressive sensing





- Highlights
  - The monotonic convergence of objective function matches the prediction of our analysis
  - The excellent agreement between two variants of PnP-PGM for MMSE and CNN denoisers

Related work in my dissertation that also investigates statistical interpretation of Plug-and-Play priors

Washington University in St.Louis

Understanding statistical interpretation of Plug-and-Play priors

- Boosting the Performance of Plug-and-Play Priors via Denoiser Scaling. ACSSC. 2020

- Provable Convergence of Plug-and-Play Priors with MMSE Denoisers. In: IEEE SPL. 2020

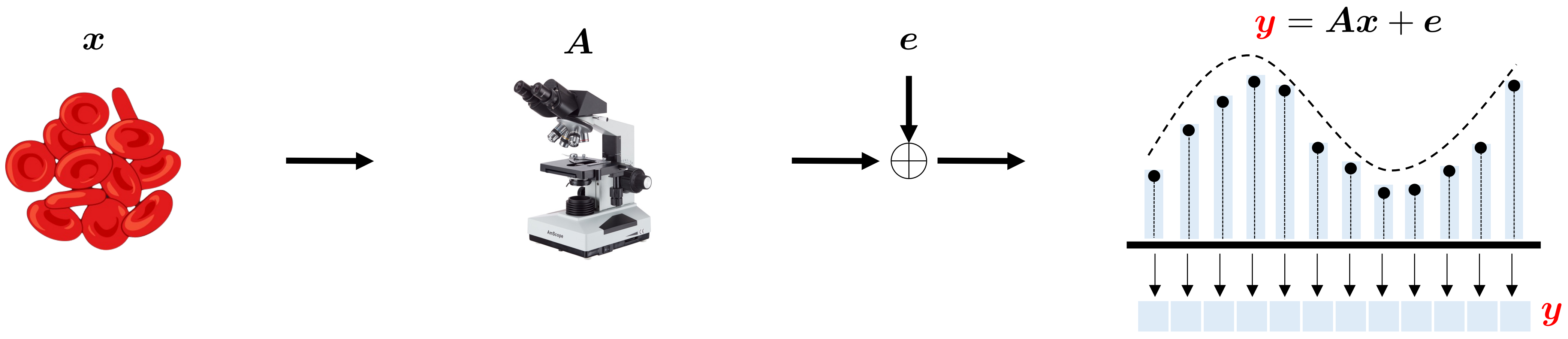
Interpretation of PnP with MMSE denoiser

A denoiser scaling technique that allow us to control the influence of denoisers with PnP.

Work 2:
Adapting plug-and-play priors to large-scale problems

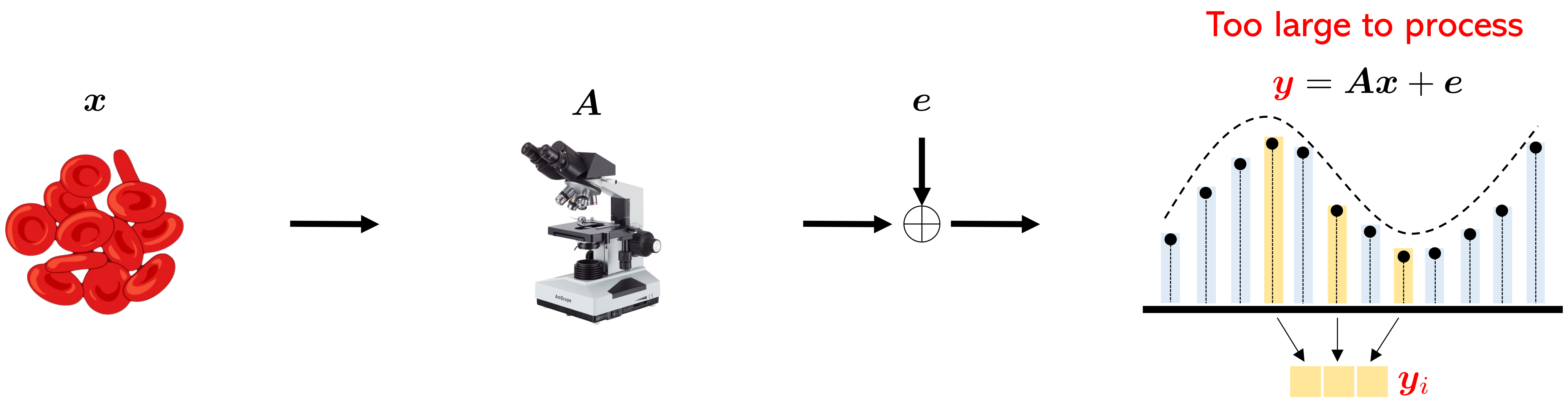
- Scalable Plug-and-Play ADMM With Convergence Guarantees. IEEE TCI, 2021

Current PnP-ADMM algorithms are not practical for addressing large-scale problems



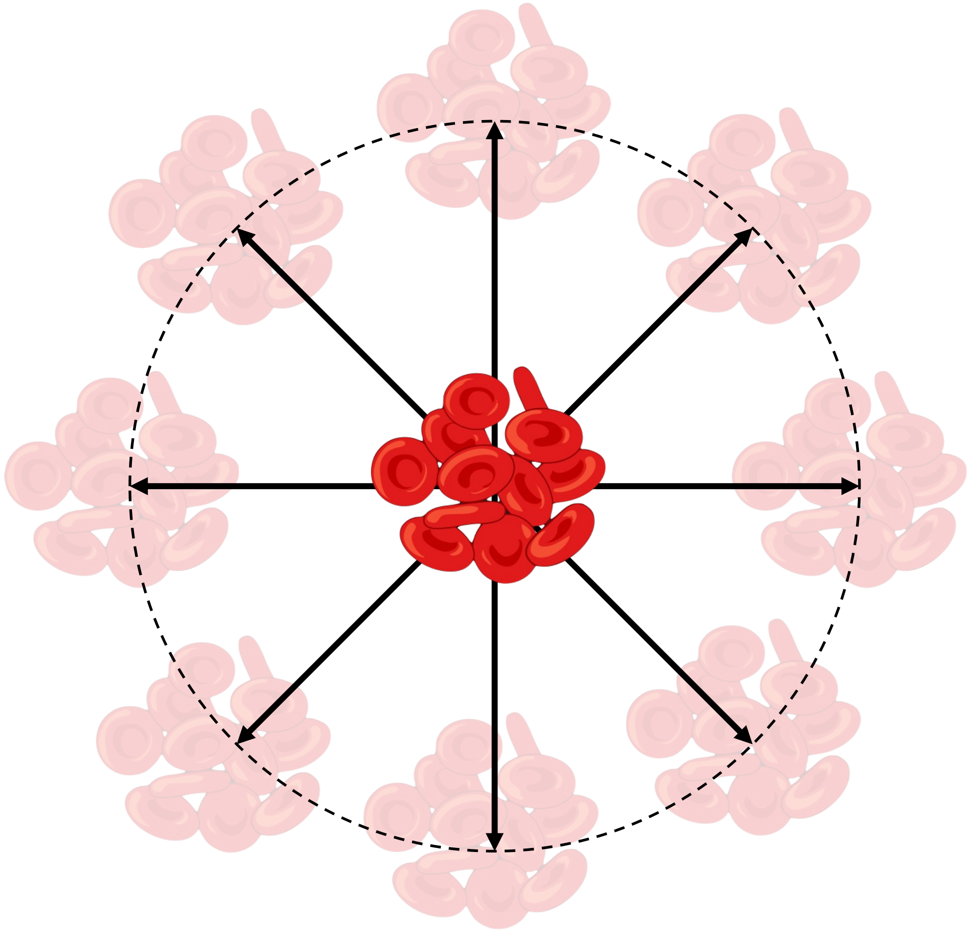
Too large to process

We propose an incremental variant of the widely used PnP-ADMM algorithm to make it scalable to problems involving large-scale measurements

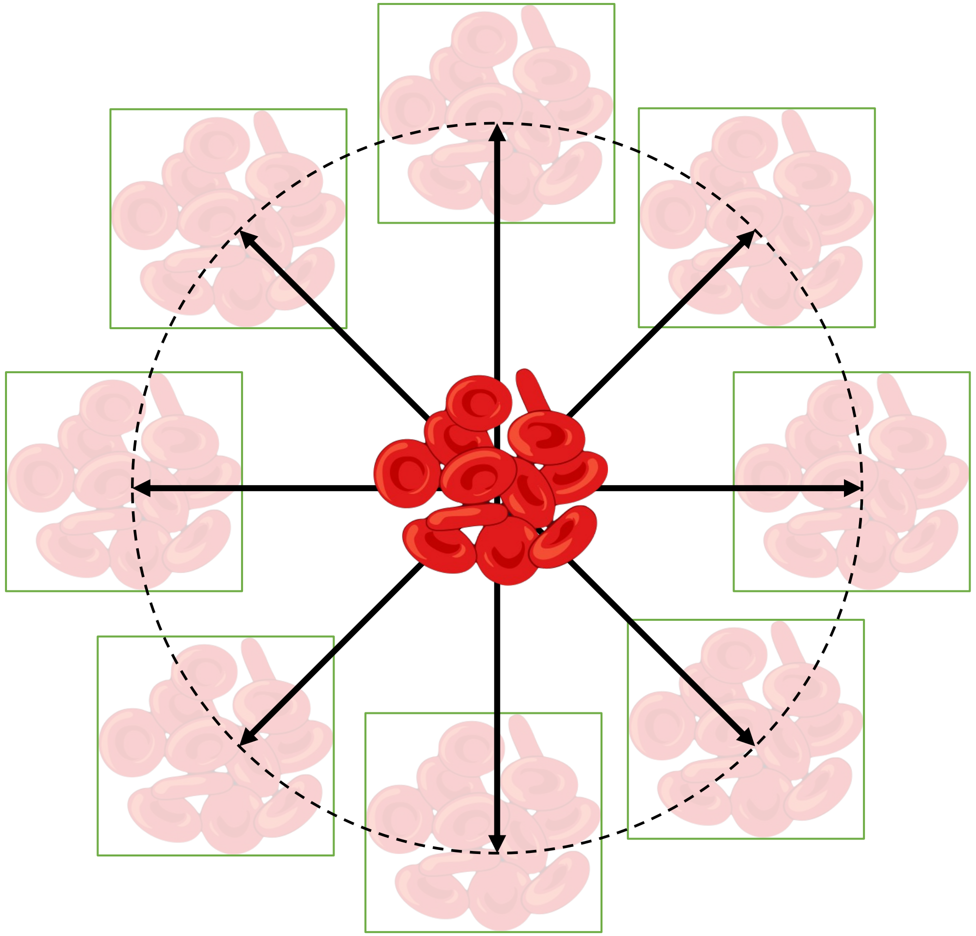


We propose an incremental variant of the widely used PnP-ADMM algorithm to make it scalable to problems involving large-scale measurements

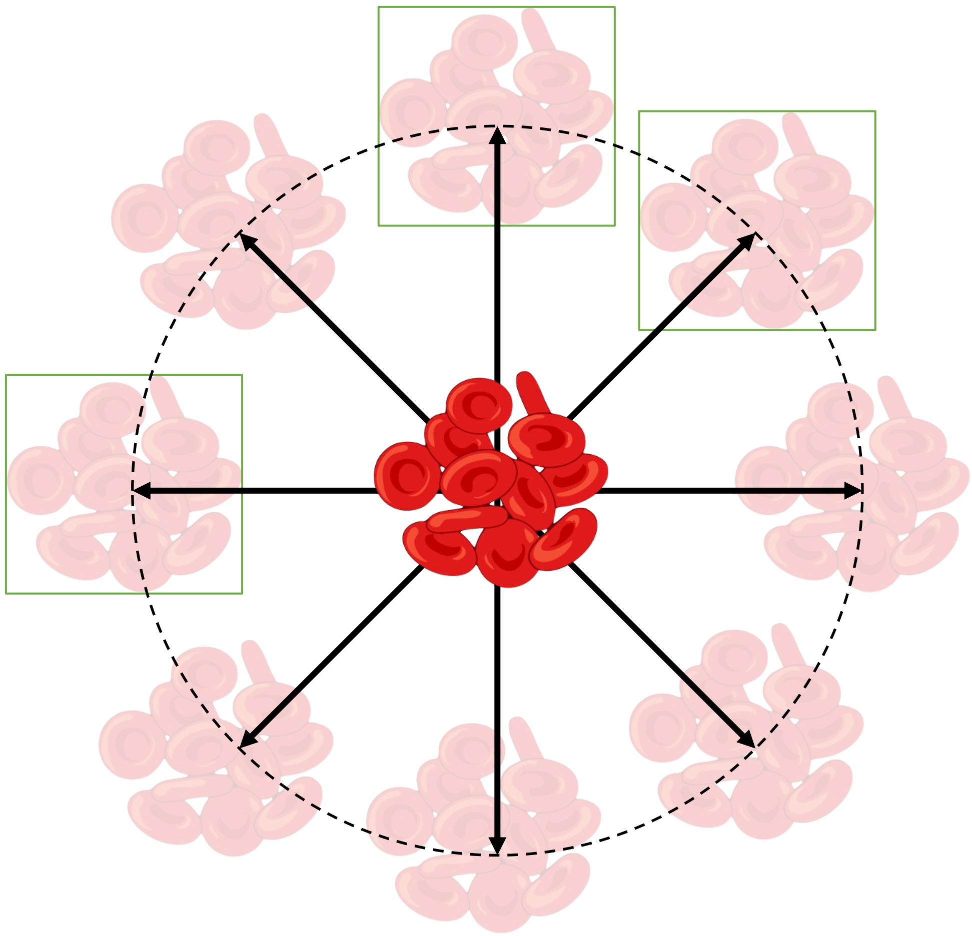
Washington University in St. Louis



We propose an incremental variant of the widely used PnP-ADMM algorithm to make it scalable to problems involving large-scale measurements

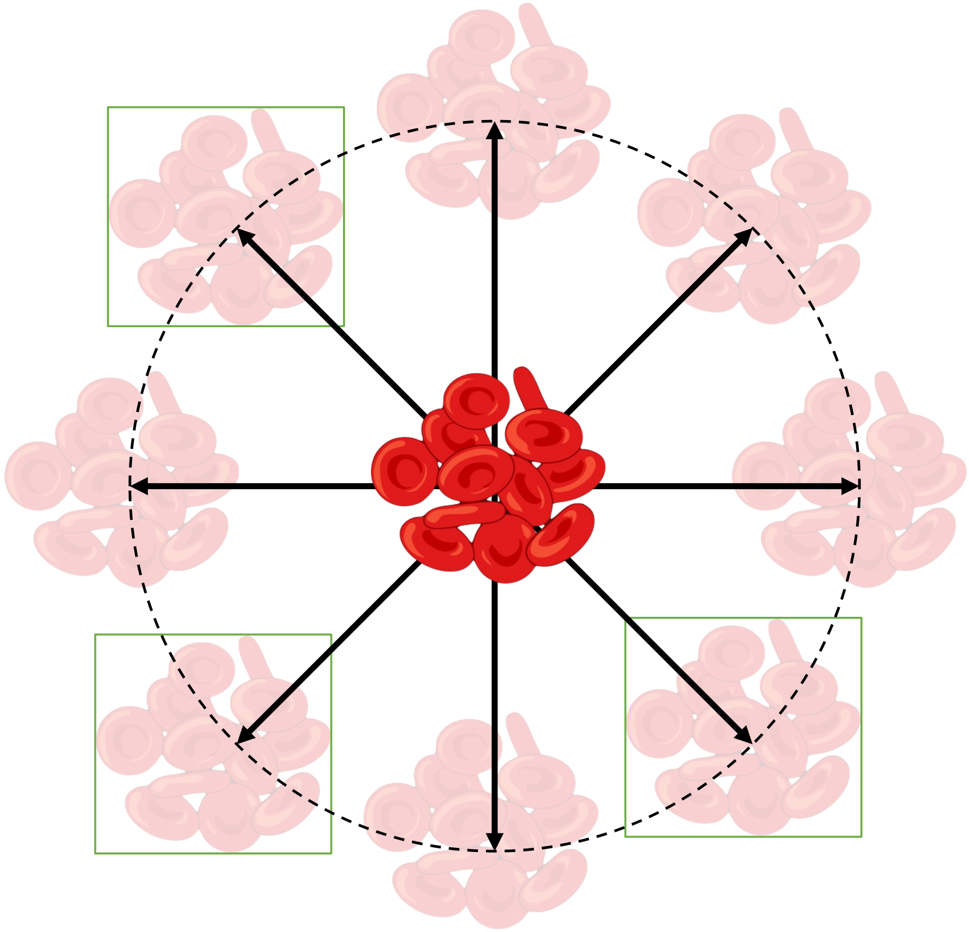


We propose an incremental variant of the widely used PnP-ADMM algorithm to make it scalable to problems involving large-scale measurements

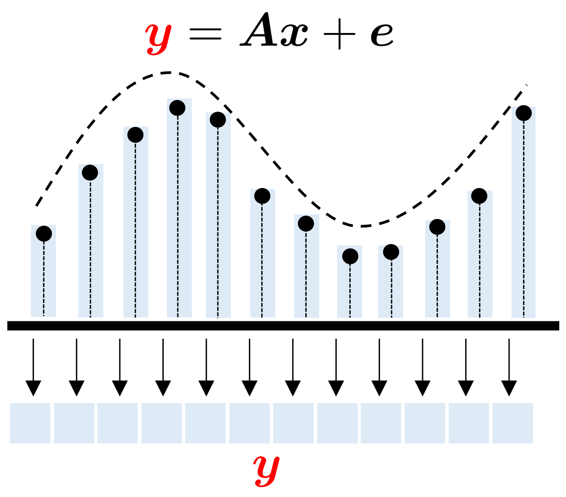


We propose an incremental variant of the widely used PnP-ADMM algorithm to make it scalable to problems involving large-scale measurements

Washington University in St. Louis



The efficiency of PnP can be improved by processing partial measurements at each iteration

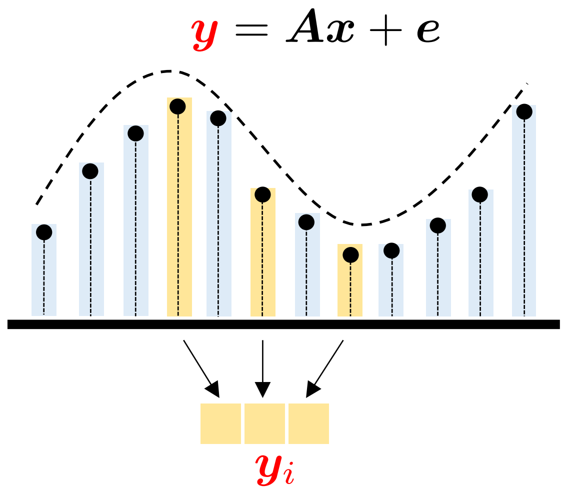


$$d(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$$

Full-batch PnP-ADMM

$$\boldsymbol{z}^k \leftarrow \mathsf{prox}_{\gamma d}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$$
$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$$
$$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$$

Reduce memory cost
Improve efficiency



$$d_i(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y}_i - \boldsymbol{A}_i\boldsymbol{x}\|_2^2$$

Incremental PnP-ADMM (**IPA**)

$$\boldsymbol{z}^k \leftarrow \mathsf{prox}_{\gamma d_i}(\boldsymbol{x}^{k-1} - \boldsymbol{s}^{k-1})$$
$$\boldsymbol{x}^k \leftarrow \mathsf{D}_\sigma(\boldsymbol{z}^k + \boldsymbol{s}^{k-1})$$
$$\boldsymbol{s}^k \leftarrow \boldsymbol{s}^{k-1} + (\boldsymbol{z}^k - \boldsymbol{x}^k)$$

IPA has a similar convergence behavior with full-batch PnP-ADMM

- **Assumptions:**
  - Each d_i is proper, closed , convex and Lipschitz continuous with constant $L_i > 0$, We define the largest Lipschitz constant as $L = \max\{L_1 , \ldots L_b\}$
  - The residual $R_\sigma := I - D_\sigma$ of the D_σ is firmly nonexpensive.
  - $\mathrm{zer}(T) \neq \emptyset$ with $\mathsf{T} := \gamma\partial d + (\mathsf{D}_\sigma^{-1} - \mathsf{I})$. There also exits $R < \infty$ such that $\|\boldsymbol{x}^k - \boldsymbol{x}^*\|_2 \leq R$ for all $\boldsymbol{x}^* \in \mathsf{zer}(\mathsf{T})$.

IPA has a similar convergence behavior with full-batch PnP-ADMM

- Theoretical convergence results: IPA vs. full-batch PnP-ADMM with $\mathsf{S} := \mathsf{D}_\sigma - \mathsf{G}(2\mathsf{D}_\sigma - \mathsf{I})$

IPA (ours):

Error term

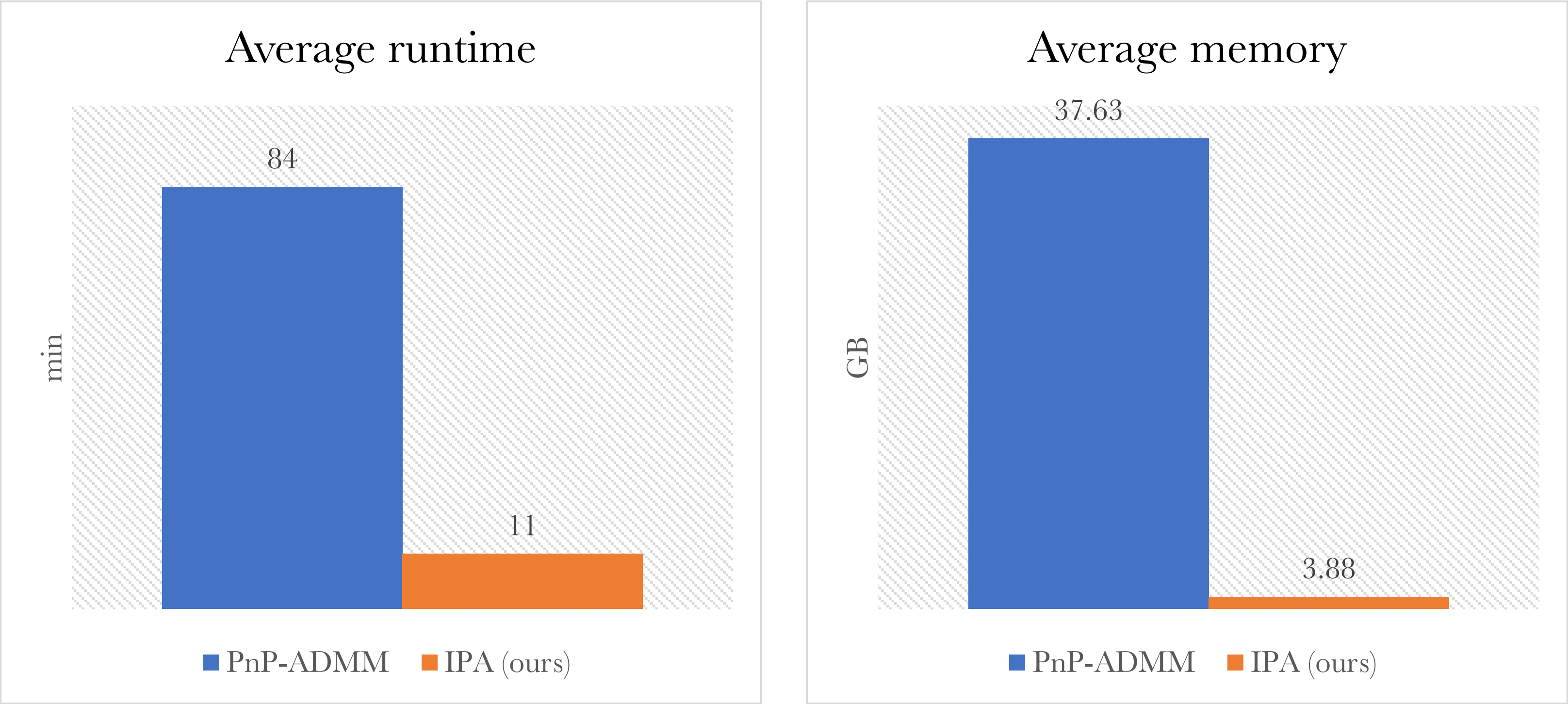
$$\mathbb{E}\left[\frac{1}{t}\sum_{k=1}^{t}\|\mathsf{S}(\boldsymbol{v}^k)\|_2^2\right] \leq \frac{(R+2\gamma L)^2}{t} + \boxed{\max\{\gamma,\gamma^2\}C}$$

PnP-ADMM:

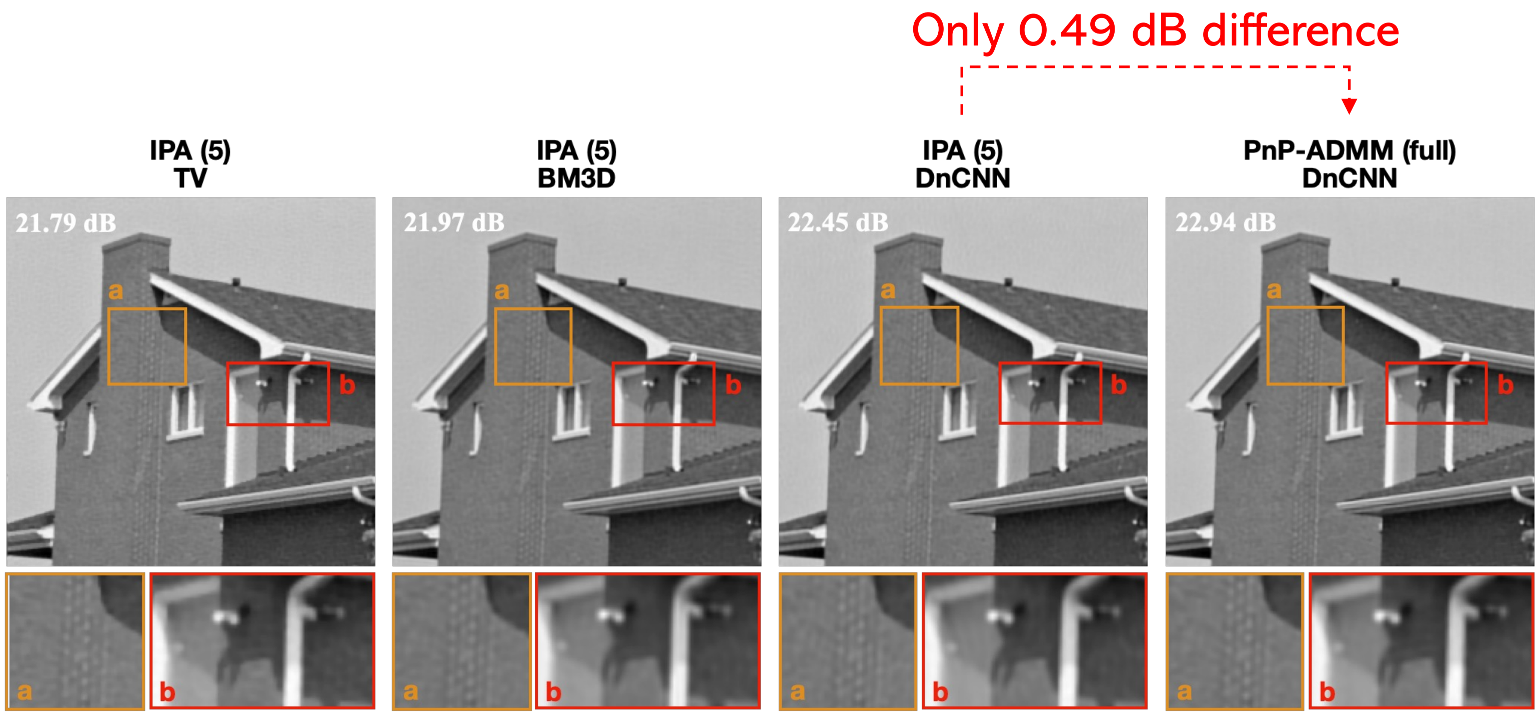
$$\frac{1}{t}\sum_{k=1}^{t}\|\mathsf{S}(\boldsymbol{v}^k)\|_2^2 \leq \frac{(R+2\gamma L)^2}{t}$$

Our experiments show that IPA runs much faster and use much less memory than the traditional full-batch PnP-ADMM

PnP-ADMM (# measurement = 600/iteration) vs. IPA (# measurement = 60/iteration)



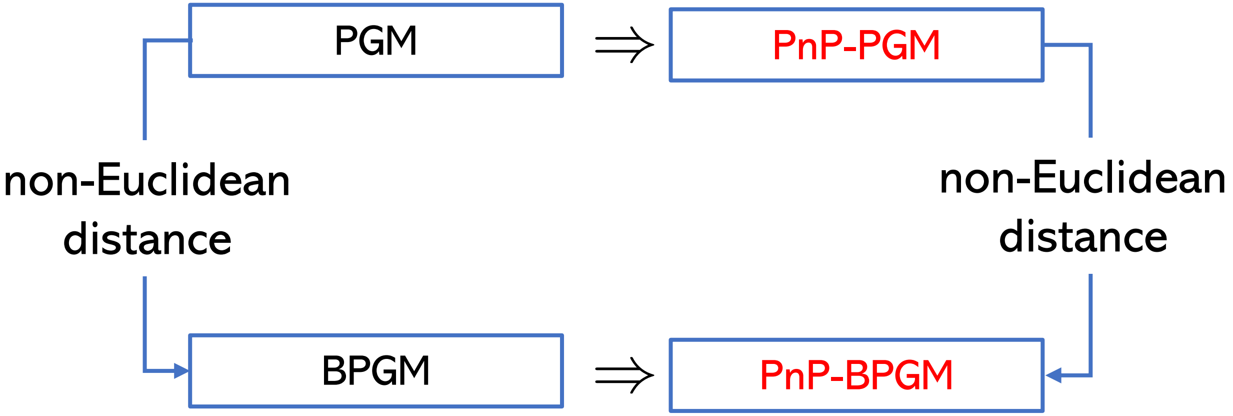
Our experiments show that IPA achieves almost the the same high-quality results as the traditional full-batch PnP-ADMM



Related work in my dissertation that extends PnP to the non-Euclidean setting

Extending Plug-and-Play Priors to the non-Euclidean setting

- Bregman Plug-and-Play Priors. ICIP, 2022

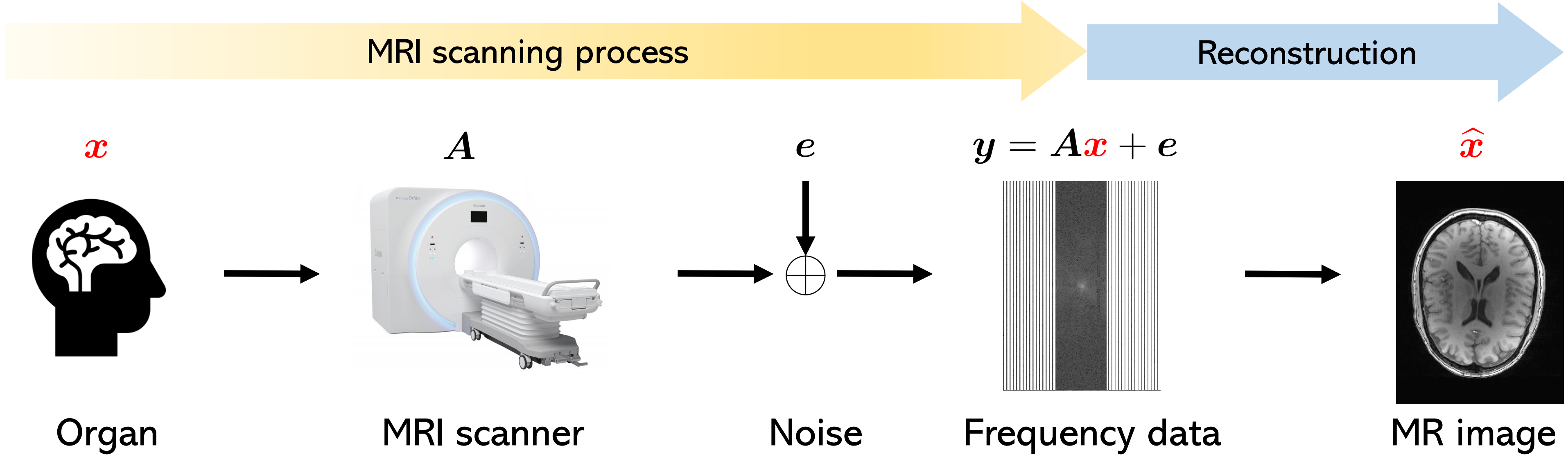


Work 3:
Applying model-based deep learning algorithms to MRI

- CoRRECT: A Deep Unfolding Framework for Motion-Corrected Quantitative R_2^* Recovery, unpublished, 2022

Magnetic resonance imaging (MRI) is an important medical technique that involves body scanning

The recovery of MRI images from subsampled k-space measurements is a typical imaging problem



$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{F}\boldsymbol{S}$$

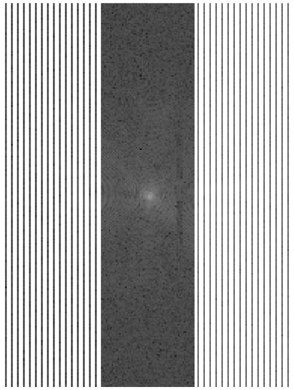
P: Frequency sampling operator

F: Fourier transform operator

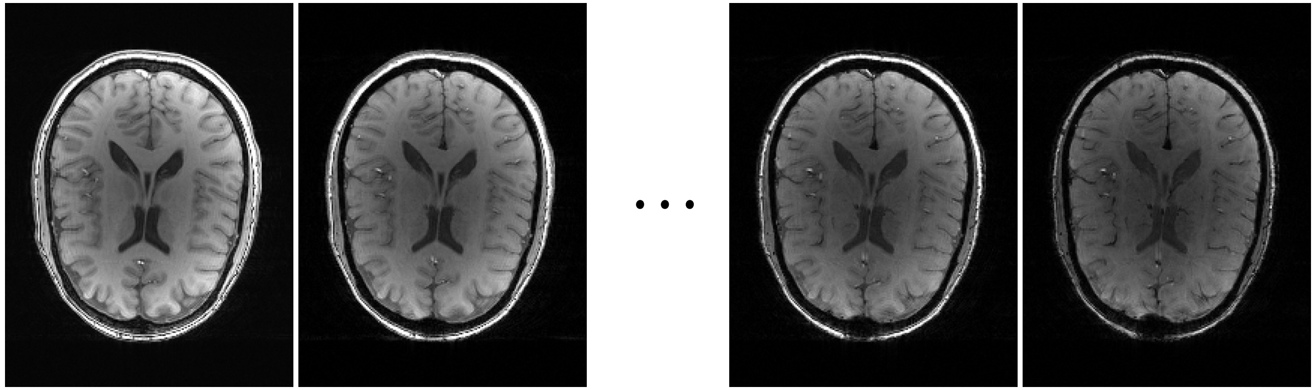
S: Pixel-wise coil sensitivity map

The reconstructed MRI images can be used to compute quantitative MRI maps

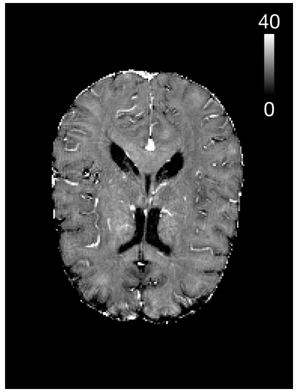
Frequency data



multi-Gradient-Recalled-Echo (mGRE) images $X(t)$



Quantitative map R_2^*



t = 4ms → t = 40ms

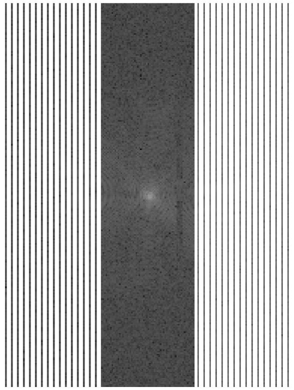
$$X(t) = X_0 \cdot \exp(-R_2^* \cdot t) \cdot F(t)$$

Biophysical model (voxel-wise, intensity-only)

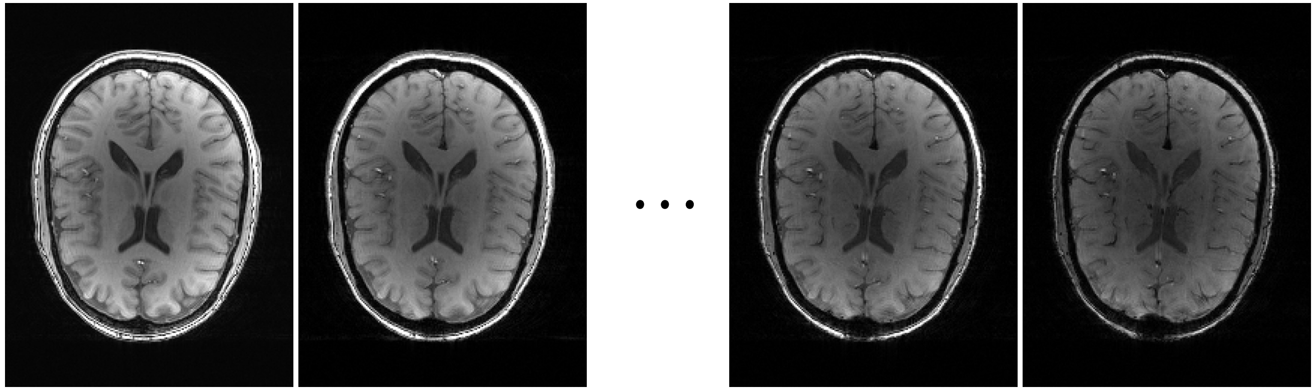
The reconstructed MRI images can be used to compute quantitative MRI maps

Washington University in St. Louis

Frequency data



multi-Gradient-Recalled-Echo (mGRE) images X(t)



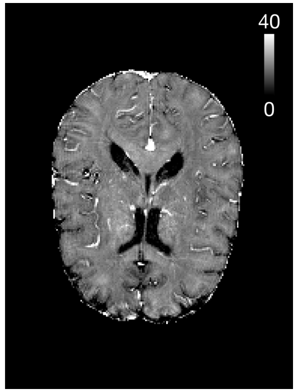
t = 4ms

t = 40ms

$$X(t) = X_0 \cdot \exp(-R_2^* \cdot t) \cdot F(t)$$

Biophysical model (voxel-wise, intensity-only)

Quantitative map R_2^*



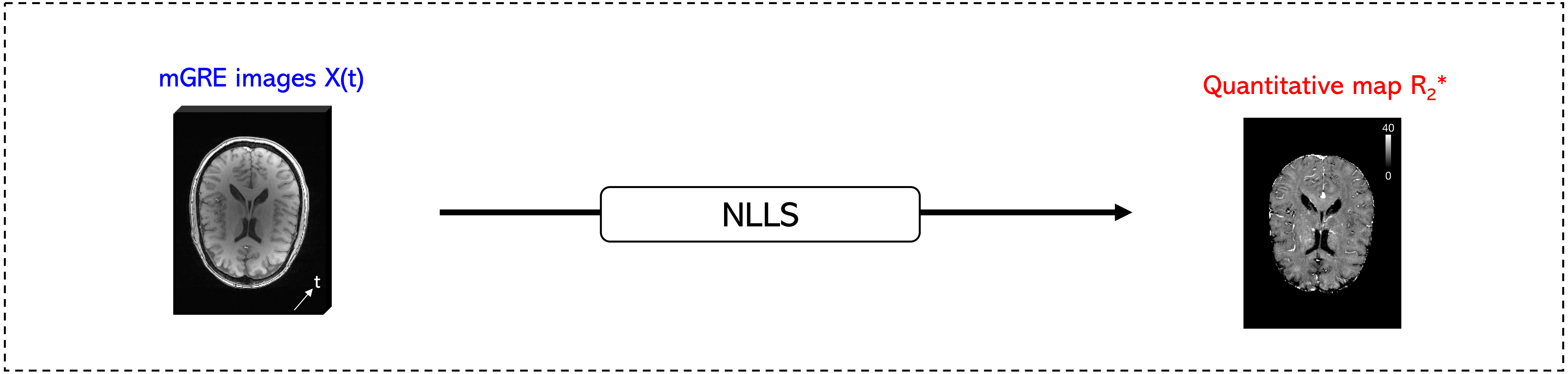
Quantitative R_2^* images can be computed from mGRE images X(t) using NLLS method

- Biophysical model

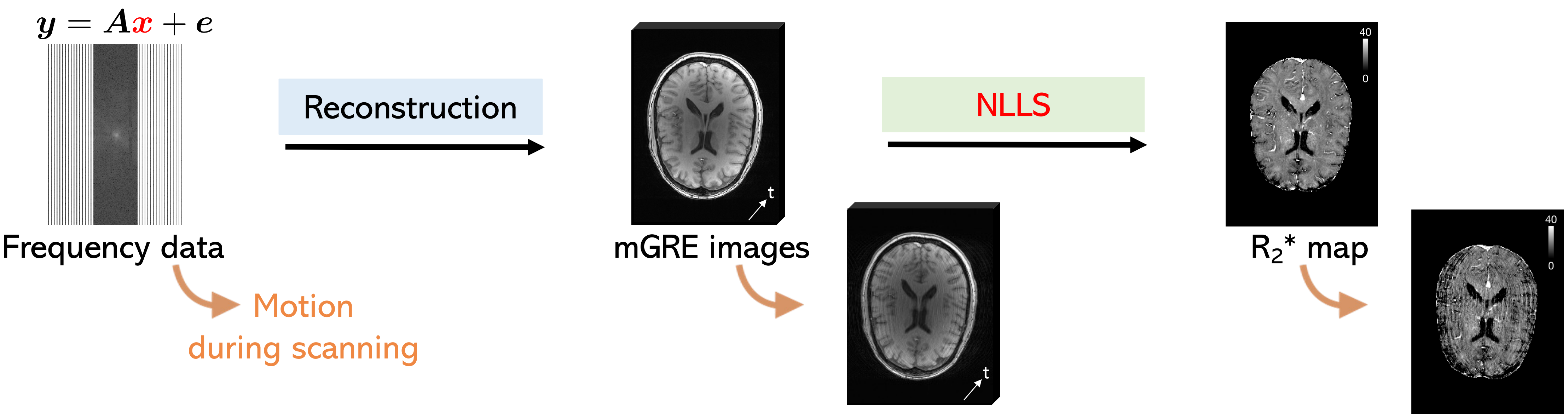
$$X(t) = X_0 \cdot \exp(-R_2^* \cdot t) \cdot F(t)$$

- Non-linear least square (NLLS)

$$\widehat{X_0}, \widehat{R_2^*} = \underset{X_0, R_2^*}{\arg\min} \sum_t \|X(t) - X_0 \cdot \exp(-R_2^* \cdot t) \cdot F(t)\|^2$$



However, NLLS suffers from many issues



- Issues with NLLS
  - It is limited by the reconstruction step
  - It is sensitive to artifacts
  - It is very time-consuming

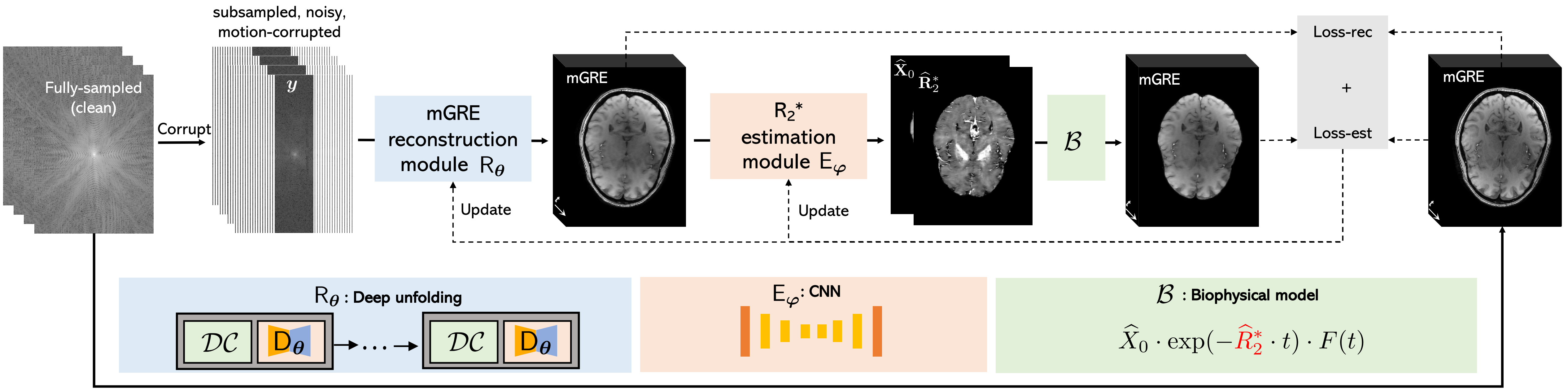
However, NLLS suffers from many issues

Co-design of MRI reconstruction and R_2^* estimation with correction for motion (CoRRECT)

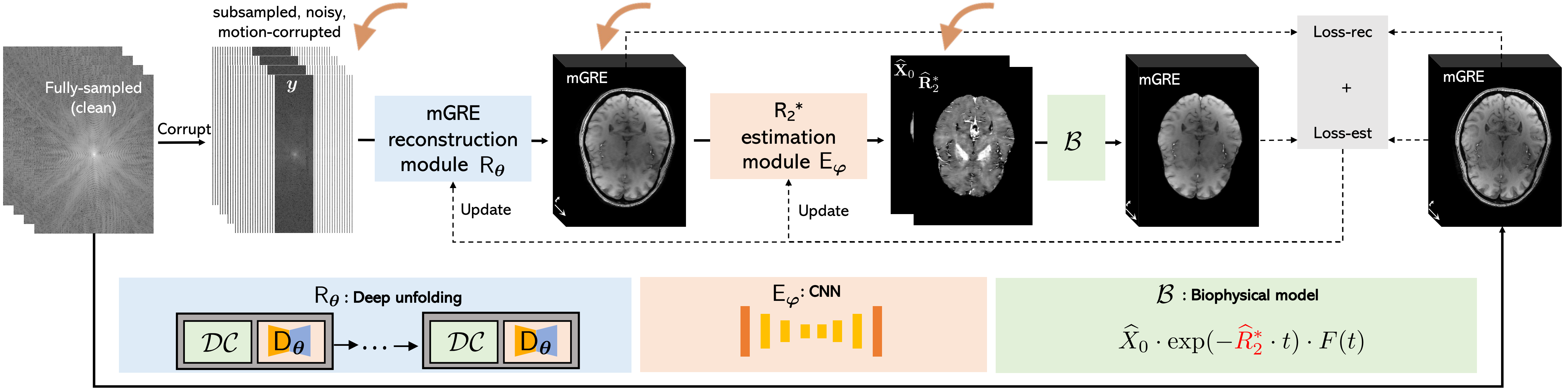
- Issues with NLLS
  - It is limited by the reconstruction step
  - It is sensitive to artifacts
  - It is very time-consuming

- Our proposal
  - Jointly reconstruct mGRE images & estimate R_2^* maps directly from the subsampled, noisy and motion-corrupted frequency data

Our proposal: co-design of MRI reconstruction and R_2* estimation with correction for motion (CoRRECT)



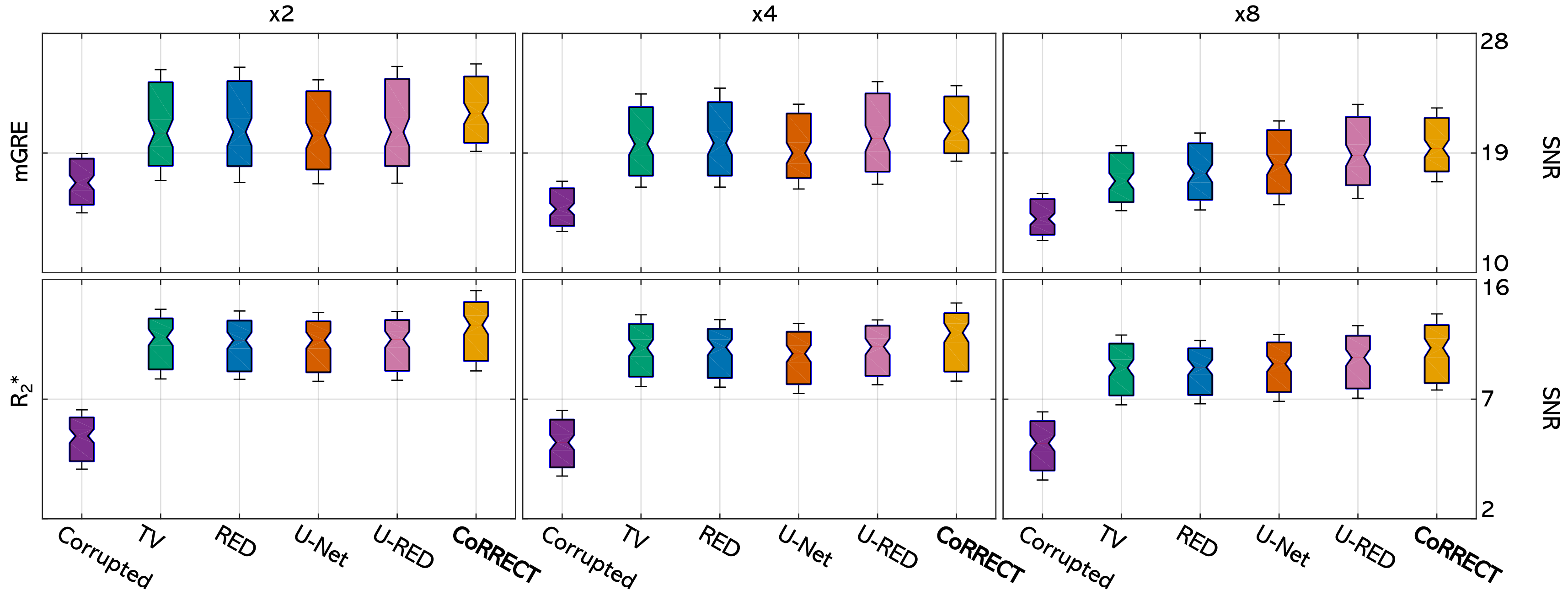
Our proposal: co-design of MRI reconstruction and R_2^* estimation with correction for motion (CoRRECT)



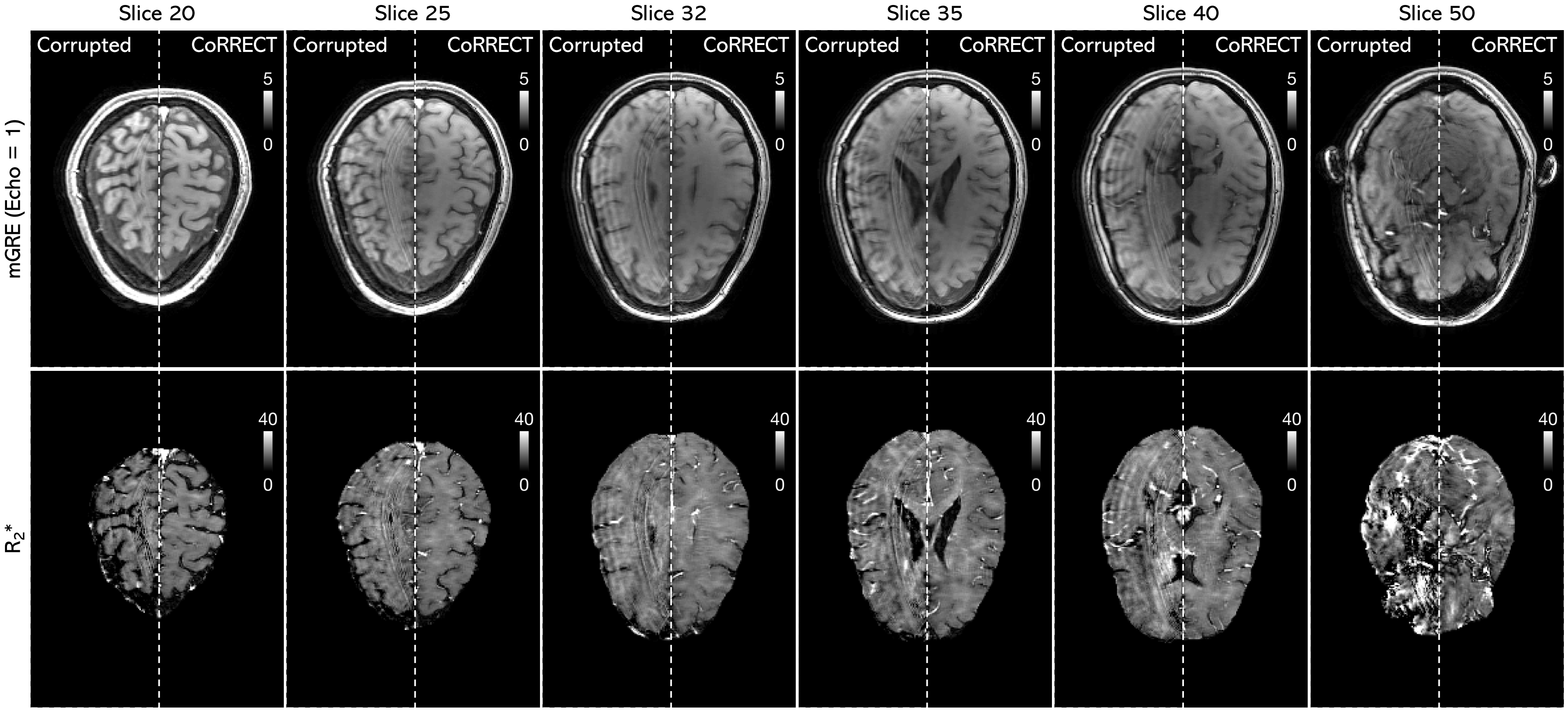
- Highlights
  - Fast end-to-end estimation of R_2^* from raw k-space data without ground truth R_2^* maps
  - Joint model-based mGRE reconstruction and model-based R_2^* estimation
  - Multitask handling: motion correction, noise removal, magnetic field inhomogeneity correction

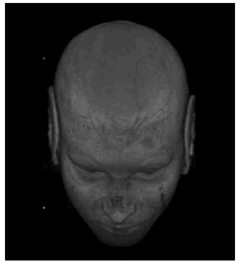
Our method shows great performance in producing high-quality mGRE image and R_2* maps on the synthetic data both qualitatively and quantitatively

Washington University in St. Louis

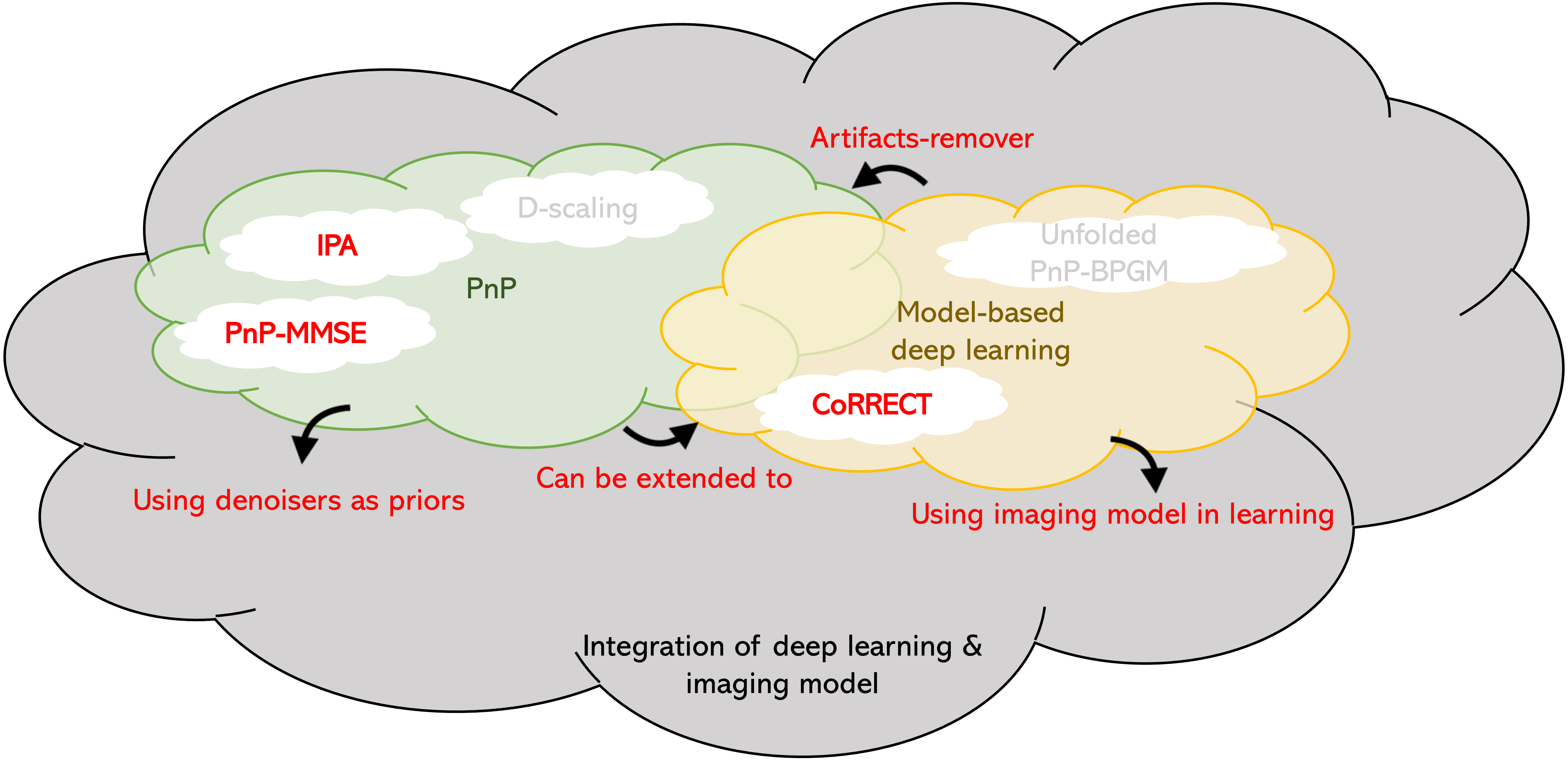


Our models trained on the synthetic data show great performance in producing high-quality mGRE images and R_2* maps on the experimental data





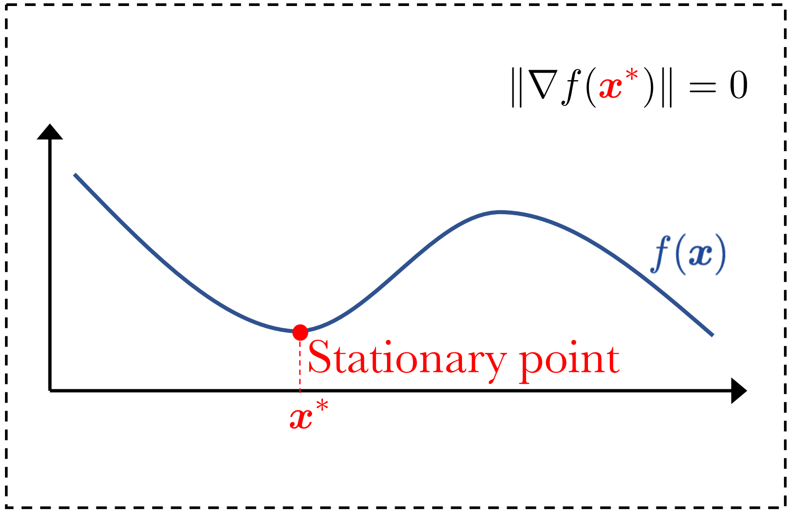
Summary



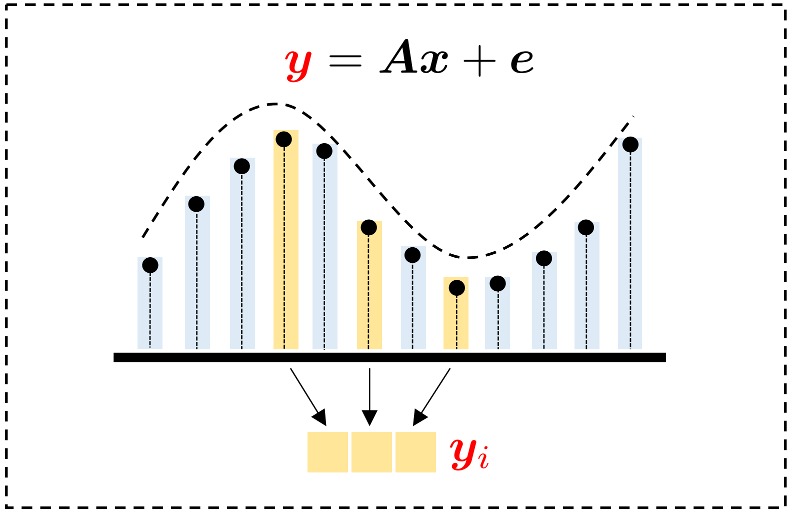
D-scaling
IPA
PnP
PnP-MMSE
Artifacts-remover
Unfolded
PnP-BPGM
Model-based
deep learning
CoRRECT
Using denoisers as priors
Can be extended to
Using imaging model in learning
Integration of deep learning &
imaging model

Summary

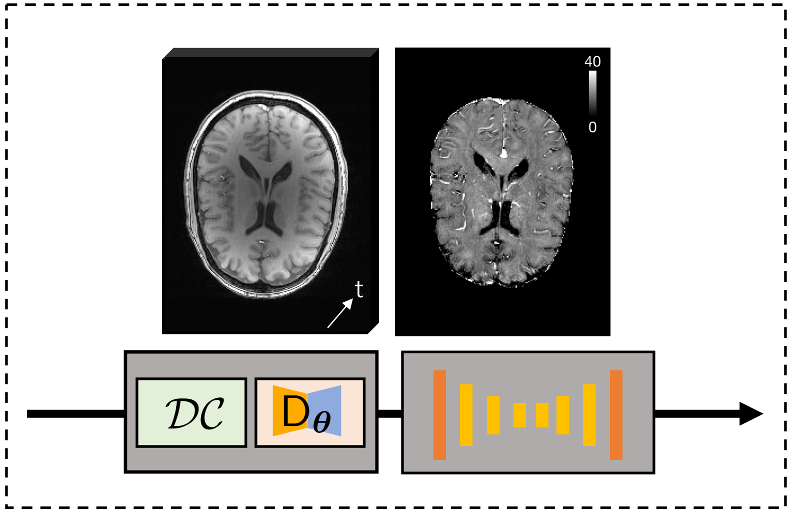
PnP(MMSE)



IPA



CoRRECT



Theory of imaging algorithms

Design of imaging algorithms

Application of imaging algorithms

The story behind my research journey

- 2017 - 08: To WashU
- 2018 - 02: Join Lab
- 2020 - 04: Oral exam
- 2021 - 05: Proposal defense
- 2022 - 07: Thesis defense

Acknowledgement

Washington University in St. Louis

CSE Department Chair

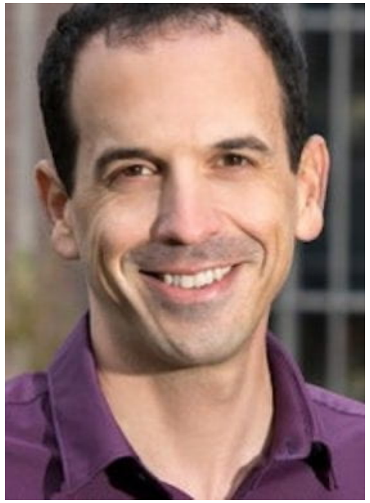
Prof.
Roch Guérin

Acknowledgement

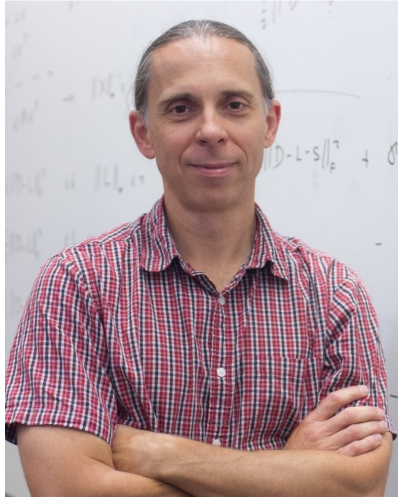
Committee Members

Prof.
Ulugbek S. Kamilov

Prof.
Tao Ju



Prof.
Netanel Raviv



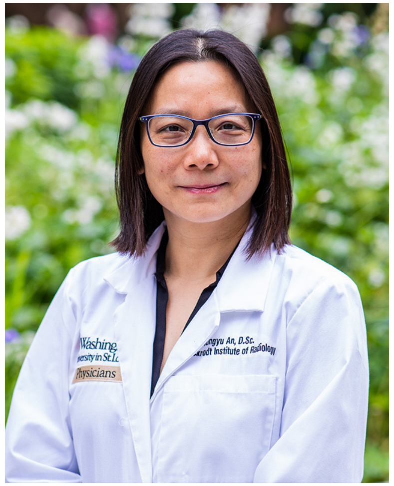
Dr.
Brendt Wohlberg

Prof.
Dmitriy A. Yablonskiy

Prof.
William Yeoh

Acknowledgement

Collator/Intern mentor



Prof.
Hongyu An
WashU

Dr.
Hassan Mansour
MERL

Dr.
Petros T. Boufounos
MERL

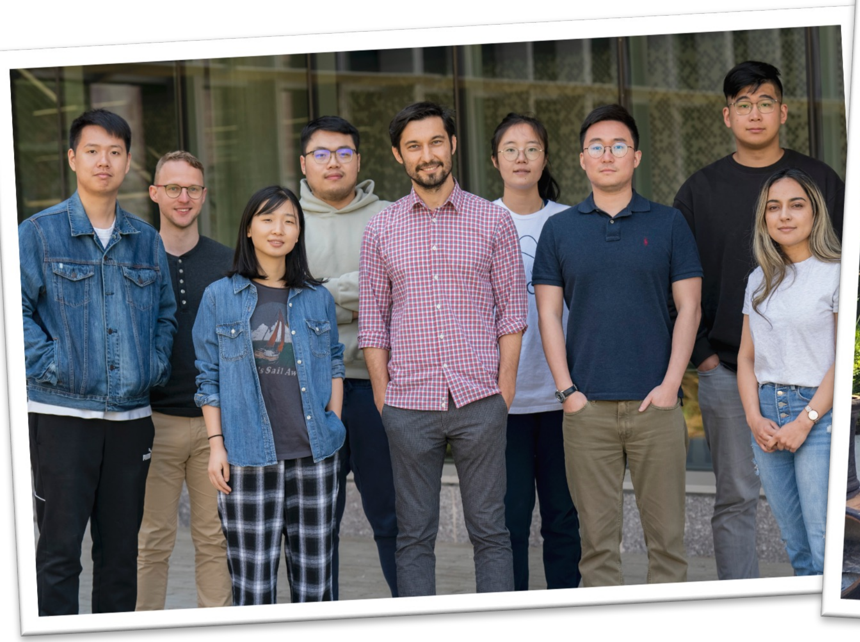
Dr.
Brain Wheelwright
Meta

Dr.
Mellissa Geng
Meta

Washington University in St.Louis

Acknowledgement

CIG labmates: Yu Sun, Jiaming Liu, Weijie Gan, Yuyang Hu, Zhixin Sun, Shirin Shoushtari, Tomas Kerepecky, and Guangxiao Song, Shiqi Xu, Zihui Wu, Max Troop, Long Fa, Weijie Gan, Jiarui Xing, Hao Tang, Ryogo Suzuki, Mingyang Xie, YuKun Li, Eddie Chandler, Julia Zeng and Yixuan Luo, EddieChandler, Zhixin Sun, Julia Zeng and Yixuan Luo, …



Acknowledgement

My friends: Zhihao Xia, Bei Wu, Wei Tang, Rusi Yan, Jeffery Jung, Aaron Park, Yifan Xu, Tiantain Zhu, Ruixuan Dai, Qianyu Liu, Dan Zeng, Chenfeng Zhao, Shenghua He, Chunyu Song, Jian Wang, Zhenzhang, …

Acknowledgement

My boyfriend Kyle

Acknowledgement

My boyfriend Kyle & his family: David, Gail, Shannon, Kenta, Blu and Fi

Acknowledgement

My family: Dad, mom, sister and brother

Acknowledgement

My advisor: Ulugbek Kamilov