AMP Methods for Fourier-Structured Operators and Signals

Saurav K. Shastri, Rizwan Ahmad, Christopher A. Metzler, Philip Schniter

The Ohio State University · University of Maryland

2022 Asilomar Conf. on Signals, Systems, and Computers
Paper ID: 1261
Supported in part by NSF CCF-1955587 and NIH R01-EB029957

Plug-and-Play (PnP) image recovery

- Goal: Recover an unknown image $\boldsymbol{x}_0$ from noisy parallel MRI measurements $\boldsymbol{y}$

$$\boldsymbol{y} = \underbrace{\begin{bmatrix} \boldsymbol{MF}\,\mathrm{Diag}(\boldsymbol{s}_1) \\ \vdots \\ \boldsymbol{MF}\,\mathrm{Diag}(\boldsymbol{s}_C) \end{bmatrix}}_{\triangleq \boldsymbol{A}} \boldsymbol{x}_0 + \boldsymbol{w}, \text{ with } \begin{cases} \boldsymbol{M}\text{: sampling mask} \\ \boldsymbol{F}\text{: 2D Fourier transform} \\ \boldsymbol{s}_c\text{: estimated coil map} \end{cases}$$

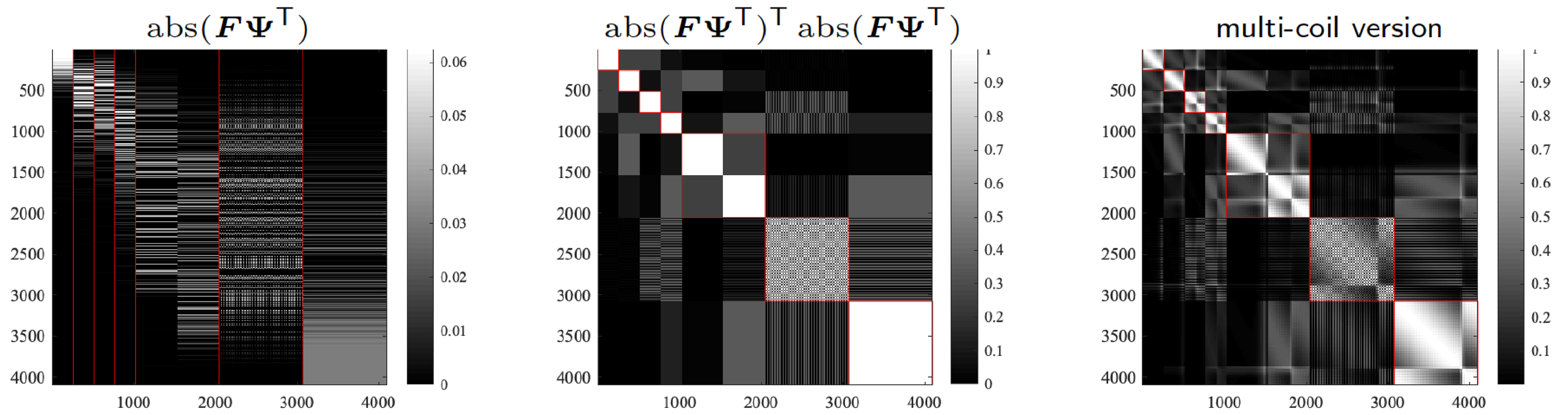
- Plug-and-play (PnP) algorithms iteratively call a deep-net image denoiser, which can be trained . . .
  - from very few images, using patches
  - independently of $\boldsymbol{A}$, facilitating generalization to any $\boldsymbol{A}$
- But there are some downsides to PnP:
  - PnP algs require careful tuning of parameters and early stopping
  - The denoiser input-error is non-white and non-Gaussian, and difficult to characterize, so it's unclear how to optimally train the denoiser

Approximate message passing (AMP) algorithms

- AMP is a family of autotuning PnP algorithms that have remarkable properties for large random $\boldsymbol{A}$:
  - The denoiser input-error is AWGN with predictable variance
  - With an MMSE denoiser, AMP algs converge to the MMSE estimate of $\boldsymbol{x}_0$ given $\boldsymbol{y}$
- Challenge: In most signal recovery problems, $\boldsymbol{A}$ does not satisfy AMP's randomness assumptions!

AMP for parallel magnetic resonance (MR) imaging

- For parallel MRI, damped AMP has been proposed [Sarkar et al'21]
  - but it is heuristic and doesn't appear to follow a state evolution
- For MRI with 2D point masks, modified VAMP algs were proposed: VDAMP [Millard et al'20] and P-VDAMP [Millard et al'22]
  - but 2D point masks are impractical and uncommon in 2D MRI
- Proposed approach: Recover the wavelet coefficients $\boldsymbol{c}_0$, not pixels $\boldsymbol{x}_0$.
  - This gives $\boldsymbol{y} = \boldsymbol{B}\boldsymbol{c}_0 + \boldsymbol{w}$ with masked Fourier-wavelet matrix $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{\Psi}^\mathsf{T}$
- For AMP algorithms, $\boldsymbol{B}$ has desirable behavior:
  - columns of different subbands are relatively decoupled from eachother
  - columns of each subband have a randomizing effect on that subband



Proposed algorithm: Denoising GEC (D-GEC)

We build upon the generalized expectation consistent (GEC) algorithm from Fletcher et al'16:

require: $\boldsymbol{f}_1(\cdot)$, $\boldsymbol{f}_2(\cdot)$, and $\mathrm{gdiag}(\cdot)$
initialize: $\boldsymbol{r}_1, \boldsymbol{\gamma}_1$
for $t = 0, 1, 2, \dots$

$\widehat{\boldsymbol{x}}_1 \leftarrow \boldsymbol{f}_1(\boldsymbol{r}_1, \boldsymbol{\gamma}_1)$ linear estimation
$\boldsymbol{\eta}_1 \leftarrow \mathrm{Diag}(\mathrm{gdiag}(\nabla \boldsymbol{f}_1(\boldsymbol{r}_1, \boldsymbol{\gamma}_1)))^{-1}\boldsymbol{\gamma}_1$
$\boldsymbol{\gamma}_2 \leftarrow \boldsymbol{\eta}_1 - \boldsymbol{\gamma}_1$
$\boldsymbol{r}_2 \leftarrow \mathrm{Diag}(\boldsymbol{\gamma}_2)^{-1}(\mathrm{Diag}(\boldsymbol{\eta}_1)\widehat{\boldsymbol{x}}_1 - \mathrm{Diag}(\boldsymbol{\gamma}_1)\boldsymbol{r}_1)$ Onsager

$\widehat{\boldsymbol{x}}_2 \leftarrow \boldsymbol{f}_2(\boldsymbol{r}_2, \boldsymbol{\gamma}_2)$ denoising
$\boldsymbol{\eta}_2 \leftarrow \mathrm{Diag}(\mathrm{gdiag}(\nabla \boldsymbol{f}_2(\boldsymbol{r}_2, \boldsymbol{\gamma}_2)))^{-1}\boldsymbol{\gamma}_2$
$\boldsymbol{\gamma}_1 \leftarrow \boldsymbol{\eta}_2 - \boldsymbol{\gamma}_2$
$\boldsymbol{r}_1 \leftarrow \mathrm{Diag}(\boldsymbol{\gamma}_1)^{-1}(\mathrm{Diag}(\boldsymbol{\eta}_2)\widehat{\boldsymbol{x}}_2 - \mathrm{Diag}(\boldsymbol{\gamma}_2)\boldsymbol{r}_2)$ Onsager

- GEC is essentially Peaceman-Rachford ADMM with adaptive vector-valued stepsizes $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$
- The GEC linear estimation stage is preconditioned LS:

$$\boldsymbol{f}_1(\boldsymbol{r}, \boldsymbol{\gamma}) = \big(\gamma_w \boldsymbol{B}^\mathsf{H}\boldsymbol{B} + \mathrm{Diag}(\boldsymbol{\gamma})\big)^{-1}\big(\gamma_w \boldsymbol{B}^\mathsf{H}\boldsymbol{y} + \mathrm{Diag}(\boldsymbol{\gamma})\boldsymbol{r}\big)$$

which can be implemented using the conjugate gradient method

- For $\boldsymbol{f}_2$, we propose to "plug in" a deep denoiser
- $\nabla \boldsymbol{f}_i$ denotes the Jacobian, and $\mathrm{gdiag}(\cdot)$ averages its diagonal across different wavelet subbands. D-GEC approximates the Jacobian using a Monte-Carlo approach [Ramani et al'08]

Proposed denoiser: Corr+Corr

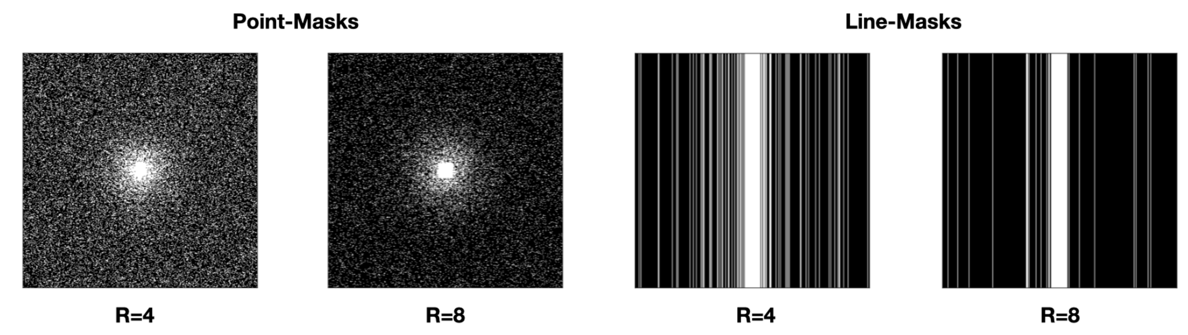
- GEC yields denoiser input-error that is AWGN with known iteration- and subband-dependent precisions $\boldsymbol{\gamma}$ *in each wavelet subband*
  - In the *pixel* domain, the error is correlated Gaussian with known covariance matrix $\boldsymbol{\Psi}\,\mathrm{Diag}(\boldsymbol{\gamma})^{-1}\boldsymbol{\Psi}^\mathsf{T}$
  - How should we inform the denoiser about $(\boldsymbol{\Psi}, \boldsymbol{\gamma})$?
- We take an arbitrary existing denoiser (e.g., DnCNN) and feed independent realizations of $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Psi}\,\mathrm{Diag}(\boldsymbol{\gamma})^{-1}\boldsymbol{\Psi}^\mathsf{T})$ into extra channels
  - The denoiser learns to extract the error statistics
  - We call it "corr+corr"
- Example PSNRs for depth-1 2D wavelet transform:

$\sqrt{\gamma^{-1}}$	white DnCNN	Metzler et al'21 DnCNN	corr+corr DnCNN	genie DnCNN
[48,47,6,19]	25.36	31.23	31.69	32.12
[10,40,23,14]	32.44	34.87	35.24	35.54
[13,7,8,10]	36.50	31.03	37.02	37.41
[10,10,10,10]	37.41	31.94	37.31	37.63
uniform [0-50,0-50,0-50,0-50]	31.07	33.24	34.08	—

white DnCNN trained unif [0-50]; Metzler et al'21 DnCNN & corr+corr DnCNN trained unif [0-50,0-50,0-50,0-50]

MR image recovery experiments

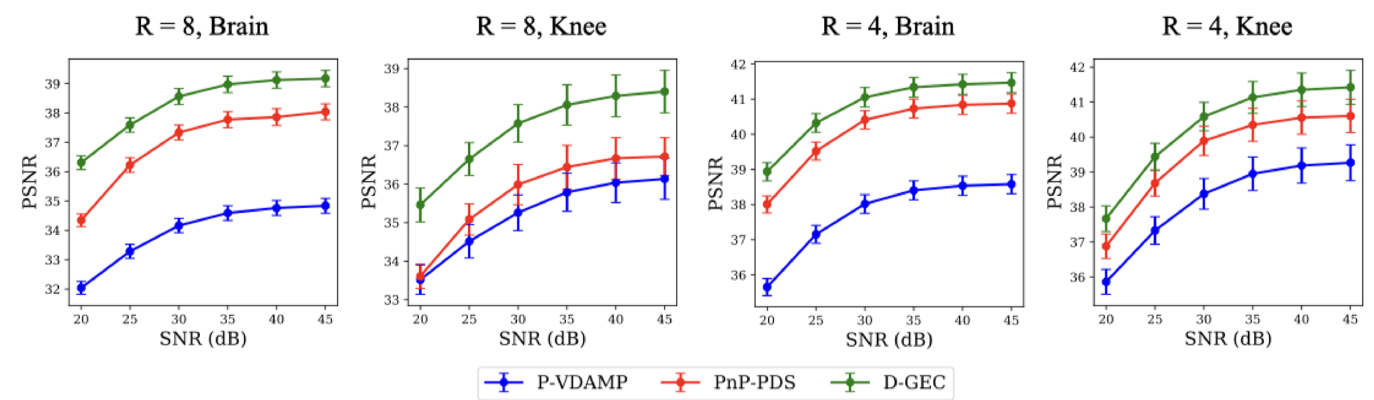
- Setup:
  - multi-coil fastMRI [Zbontar et al'18] brain and knee data
  - 8 virtual coils; acceleration $R = N/M = 4$ & 8
  - additional AWGN $\boldsymbol{w}$ for noise-robustness study
  - variable-density 2D point- and line-masks:



- 2D line-mask results averaged over 16 test images:

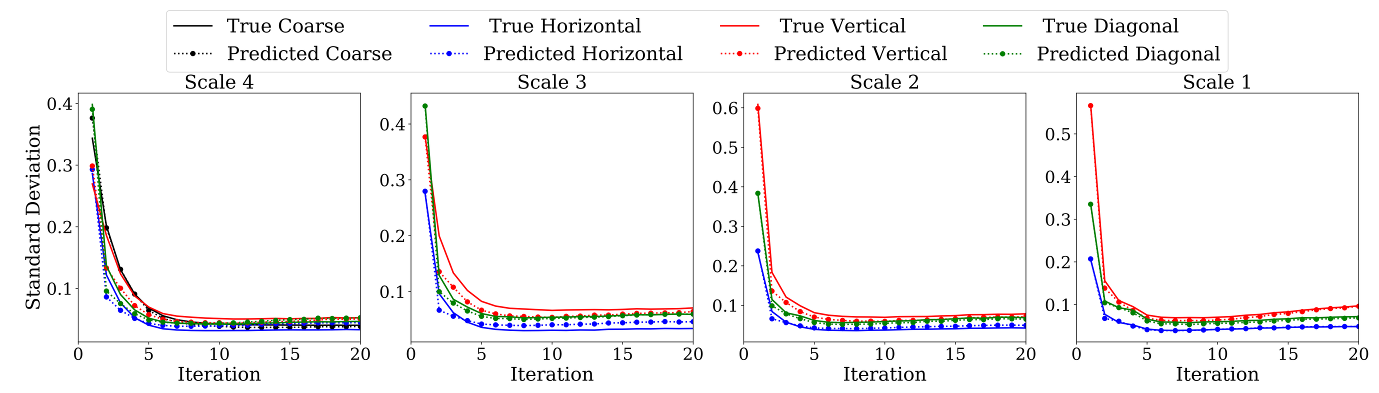
	Knee				Brain			
	$R=4$		$R=8$		$R=4$		$R=8$	
method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
P-VDAMP	33.84	0.9018	20.34	0.5614	30.30	0.8847	13.51	0.4763
PnP-PDS	36.28	0.9204	32.34	0.8556	38.07	0.9501	28.97	0.8269
D-GEC	**38.82**	**0.9504**	**33.66**	**0.8893**	**39.04**	**0.9631**	**30.61**	**0.9015**

- Average PSNR versus measurement SNR with 2D point-mask:

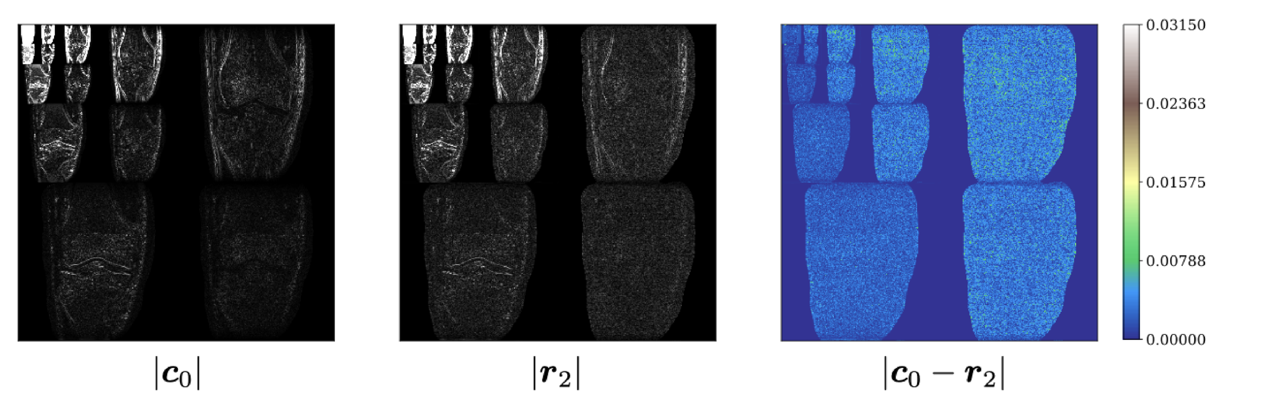


Note: PnP-PDS penalty and stopping iteration tuned for every (SNR,R,dataset)

- Predictability of D-GEC error variance vs iteration with 2D line-mask at R = 4:



- Example wavelet error at iteration 10 with 2D line-mask at R = 4:



- Example wavelet-error QQ plots at iteration 10 with 2D line-mask at R = 4: