MRI Image Recovery Using Damped Denoising Vector AMP

Subrata Sarkar
Rizwan Ahmad
Philip Schniter

THE OHIO STATE UNIVERSITY

ICASSP 2021 – Paper ID: 3966
(Supported by NSF 1955587, NIH 135489)

Magnetic Resonance Imaging (MRI)

- Magnetic resonance imaging (MRI) is a non-invasive diagnostic tool that provides excellent soft-tissue contrast without using ionizing radiation.
- The measurements $\boldsymbol{y}$ are in the spatial Fourier domain, called k-space:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{w}, \text{ with } \boldsymbol{A} = \boldsymbol{M}\boldsymbol{F}.$$

Above, $\boldsymbol{x}_0$ is the image, $\boldsymbol{F} \in \mathbb{C}^{N\times N}$ is the 2D DFT matrix, $\boldsymbol{M} \in \mathbb{C}^{M\times N}$ is a sampling mask and $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \gamma_w^{-1}\boldsymbol{I})$ is AWGN.

- The primary drawback of MRI time needed to collect the measurements.
- To accelerate MRI, one collects only a few k-space samples: $M \ll N$.

Goal: Recover the unknown image $\boldsymbol{x}_0 \in \mathbb{C}^N$ from $\boldsymbol{y} \in \mathbb{C}^M$.

Approach: Plug-and-Play recovery using Damped Denoising Vector-AMP.

Plug-and-Play (PnP) Image Recovery

- The classical approach to image recovery is optimization:

$$\arg\min_{\boldsymbol{x}} \left\{ \frac{\gamma_w}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \phi(\boldsymbol{x}) \right\}, \quad (1)$$

where the regularizer $\phi(\cdot)$ penalizes $\boldsymbol{x}$ that are atypical for images.

- ADMM is a popular algorithm to solve this optimization problem:

$$\begin{aligned} \boldsymbol{x}^{t+1} &= \arg\min_{\boldsymbol{x}} \tfrac{\gamma_w}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \tfrac{\gamma}{2}\|\boldsymbol{x} - \boldsymbol{v}^t + \boldsymbol{u}^t\|^2 \\ \boldsymbol{v}^{t+1} &= \text{prox}_{\gamma^{-1}\phi}(\boldsymbol{x}^{t+1} + \boldsymbol{u}^t) \\ \boldsymbol{u}^{t+1} &= \boldsymbol{u}^t + (\boldsymbol{x}^{t+1} - \boldsymbol{v}^{t+1}), \end{aligned} \quad (2)$$

where $\text{prox}_\rho(\boldsymbol{r}) \triangleq \arg\min_{\boldsymbol{x}}\{\rho(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{r}\|^2\}$.

- The prox operation (2) can be interpreted as MAP denoising of the AWGN-corrupted image $\boldsymbol{r} = \boldsymbol{x} + \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}/\gamma)$ under prior $\boldsymbol{x} \sim \frac{1}{Z}e^{-\phi(\boldsymbol{x})}$.
- To improve performance, PnP-ADMM [1] replaces the prox operator with a sophisticated image denoiser $\boldsymbol{f}(\cdot)$ like BM3D or DnCNN.
- PnP can be generalized to other algorithms like FISTA, PDS, etc.

Approximate Message Passing (AMP)

- AMP [2] is a computationally efficient iterative algorithm for solving (1) that yields optimal recovery under large random $\boldsymbol{A}$.
- When $\boldsymbol{A}$ is large, i.i.d., and sub-Gaussian, ...
  - AMP's macroscopic behavior is rigorously characterized by state-evolution (SE) [3].
  - AMP converges very quickly, e.g., 10-20 iterations.
  - When $\boldsymbol{f}$ is the MMSE denoiser and the SE has a unique fixed-point, AMP provably converges to the MMSE $\widehat{\boldsymbol{x}}$ [3].
- When used with an image denoiser $\boldsymbol{f}$ like BM3D or DnCNN, AMP is called "denoising-AMP" (D-AMP) [4].

$$\begin{aligned} \boldsymbol{v}^{t+1} &= \beta \cdot \left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^t + \tfrac{1}{M}\boldsymbol{v}^t \,\text{tr}\{\nabla \boldsymbol{f}(\boldsymbol{x}^{t-1} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{v}^t; 1/\tau^t)\}\right) \\ \tau^{t+1} &= \tfrac{1}{M}\|\boldsymbol{v}^{t+1}\|^2 \\ \boldsymbol{x}^{t+1} &= \boldsymbol{f}(\boldsymbol{x}^t + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{v}^{t+1}; 1/\tau^{t+1}) \end{aligned}$$

where $\beta = N/\|\boldsymbol{A}\|_F^2$. The quantity $\text{tr}\{\nabla \boldsymbol{f}(\cdot; 1/\tau)\}/N$ is known as the divergence, and is approximated using Monte Carlo [4] in practice.

AMP for MRI

- In MRI, the measurement matrix $\boldsymbol{A}$ is not i.i.d., and so AMP tends to perform poorly or even diverge.
- Several MRI-specific variations of AMP have been proposed:
  - BM3D-AMP-MRI [5]: uses $\beta = 1$ in D-AMP, which stabilizes the algorithm but degrades the fixed points.
  - Variable-density AMP (VD-AMP) [6] is a wavelet-denoiser-based AMP/VAMP hybrid. It works well with the point-sampling mask, but fails for other masks like Cartesian.
  - De-biased D-AMP (DD-AMP) [7] uses a diagonal-matrix β in D-AMP, and works well (empirically) with a wide range of masks, e.g., Cartesian.

Vector Approximate Message Passing (VAMP)

- Vector AMP (VAMP) [8] has similar properties to AMP but holds for the larger class of right-orthogonally invariant (ROI) random matrices.
- When $\boldsymbol{A}$ is ROI, i.e., has SVD $\boldsymbol{USV}^{\mathsf{H}}$ with large random unitary $\boldsymbol{V}$, ...
  - VAMP's macroscopic behavior is rigorously characterized by state-evolution (SE) [8].
  - VAMP converges very quickly, e.g., 5-15 iterations.
  - With MMSE $\boldsymbol{f}$ and unique SE fixed-point, VAMP yields MMSE $\widehat{\boldsymbol{x}}$ [8,9].
- When used with an image denoiser $\boldsymbol{f}$ like BM3D or DnCNN, VAMP is called "denoising-VAMP" (D-VAMP) [10].

Damped Denoising VAMP (DD-VAMP)

- In MRI, the measurement matrix $\boldsymbol{A}$ is *not* ROI, and so VAMP tends to perform poorly or even diverge.
- We propose carefully chosen damping to alleviate these issues:
  - We propose to damp α_1 to reduce its approximation error due to Monte Carlo
  - We propose to transform the variance α_1 and the precision γ_2 to amplitudes for damping, and then transform them back.
- Note that DD-VAMP reduces to D-VAMP when $\theta = 1 = \zeta$.

initialize: $\boldsymbol{r}_2^0,\ \gamma_2^0,\ \theta, \zeta \in (0,1],\ \boldsymbol{q} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$
for $t = 0, 1, 2, \dots$

$\boldsymbol{x}_2^t = \boldsymbol{g}(\boldsymbol{r}_2^t; \gamma_2^t)$	linear estimation
$\alpha_2^t = \text{tr}\{\nabla \boldsymbol{g}(\boldsymbol{r}_2^t; \gamma_2^t)\}/N$	divergence
$\boldsymbol{r}_1^t = (\boldsymbol{x}_2^t - \alpha_2^t \boldsymbol{r}_2^t)/(1 - \alpha_2^t),\ \gamma_1^t = \gamma_2^t(1 - \alpha_2^t)/\alpha_2^t$	Onsager correction
$\boldsymbol{x}_1^t = \boldsymbol{f}(\boldsymbol{r}_1^t; \gamma_1^t)$	denoising
$\overline{\alpha}_1^t = \epsilon^{-1}\boldsymbol{q}^{\mathsf{H}}[\boldsymbol{f}(\boldsymbol{r}_1^t + \epsilon\boldsymbol{q}; \gamma_1^t) - \boldsymbol{f}(\boldsymbol{r}_1^t; \gamma_1^t)]$	Monte-Carlo divrgnce
$\alpha_1^t = \left[\theta(\overline{\alpha}_1^t)^{\frac{1}{2}} + (1-\theta)(\alpha_1^{t-1})^{\frac{1}{2}}\right]^2$	damping
$\overline{\boldsymbol{r}}_2^{t+1} = (\boldsymbol{x}_1^t - \alpha_1^t \boldsymbol{r}_1^t)/(1 - \alpha_1^t),\ \overline{\gamma}_2^{t+1} = \gamma_1^t(1 - \alpha_1^t)/\alpha_1^t$	Onsager correction
$\boldsymbol{r}_2^{t+1} = \zeta\overline{\boldsymbol{r}}_2^{t+1} + (1-\zeta)\boldsymbol{r}_2^t$	damping
$\gamma_2^{t+1} = \left[\zeta(\overline{\gamma}_2^{t+1})^{-\frac{1}{2}} + (1-\zeta)(\gamma_2^t)^{-\frac{1}{2}}\right]^{-2}$	damping

Above, $\boldsymbol{g}(\cdot; \gamma)$ is the linear MMSE estimator under prior signal precision γ:

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{r}; \gamma) &\triangleq \arg\min_{\boldsymbol{x}} \left\{ \frac{\gamma_w}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \frac{\gamma}{2}\|\boldsymbol{x} - \boldsymbol{r}\|^2 \right\} \\ &= \boldsymbol{F}^{\mathsf{H}}(\gamma_w \boldsymbol{M}^{\mathsf{T}}\boldsymbol{M} + \gamma \boldsymbol{I})^{-1}(\gamma \boldsymbol{F}\boldsymbol{r} + \gamma_w \boldsymbol{M}^{\mathsf{T}}\boldsymbol{y}) \end{aligned}$$

and $\text{tr}\{\nabla \boldsymbol{g}(\boldsymbol{r}; \gamma)\}/N = ((1 - M/N)\gamma_w + \gamma)/(\gamma_w + \gamma)$.

DD-VAMP++

- Empirically, the fixed points of DD-VAMP are similar or better than those of PnP-ADMM. However, damping slows DD-VAMP's convergence.
- Importantly, VAMP reduces to the Peaceman-Rachford variant of ADMM (ADMM-PR) when the precisions are fixed, i.e., $\gamma_1^t = \gamma_2^t = \gamma, \forall t$.
- We propose to initialize DD-VAMP using ADMM-PR:
  - First run PnP-ADMM-PR for T_{swi} iterations at precision γ, then switch to DD-VAMP.
  - Tune the parameters T_{swi} and γ using training data.

We call this method "DD-VAMP++."

References

[1] S. V. Venkatakrishnan, C. A. Bouman, B. Wohlberg, "Plug-and-play priors for model based reconstruction," *GlobalSIP*'13.
[2] D. L. Donoho, A. Maleki, A. Montanari, "Message passing algorithms for compressed sensing," *PNAS*'09.
[3] M. Bayati, A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Info. Thy*'11.
[4] C. A. Metzler, A. Maleki, R. G. Baraniuk, "BM3D-AMP: A new image recovery algorithm based on BM3D denoising," *ICIP*'15. (http://dsp.rice.edu/software/DAMP-toolbox)
[5] E. M. Eksioglu, A. K. Tanc, "Denoising AMP for MRI reconstruction: BM3D-AMP-MRI," *SIAM JIS*'18.
[6] C. Millard, A. T. Hess, B. Mailhé, and J. Tanner, "Approximate message passing with a colored aliasing model for variable density Fourier sampled images," *arXiv:2003.02701*, 2020.

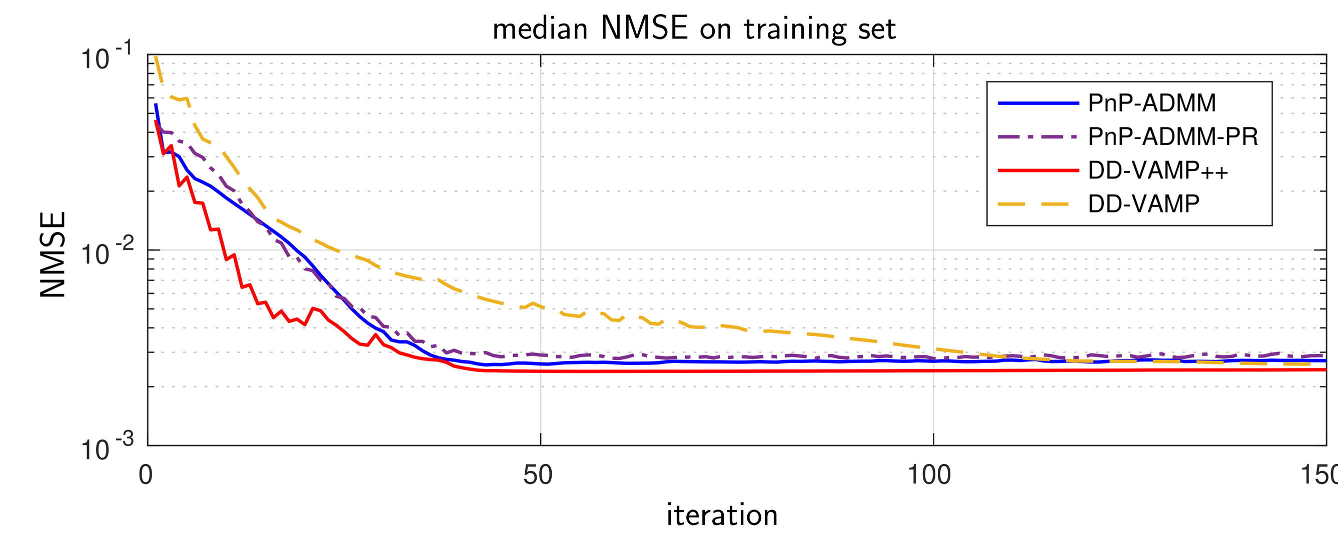
Image Recovery in MRI

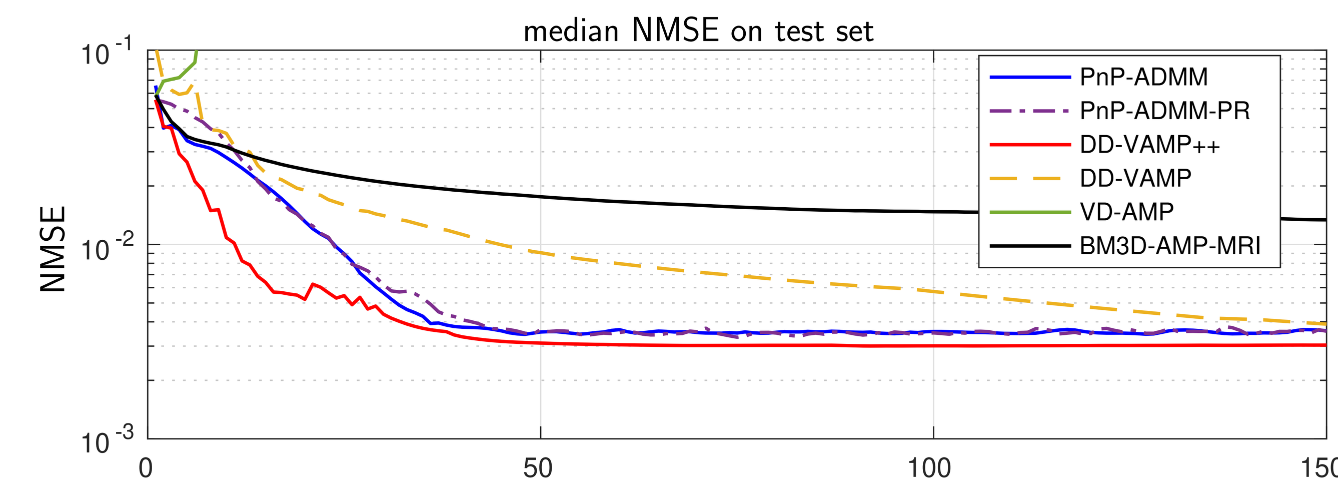
Experiment Setup

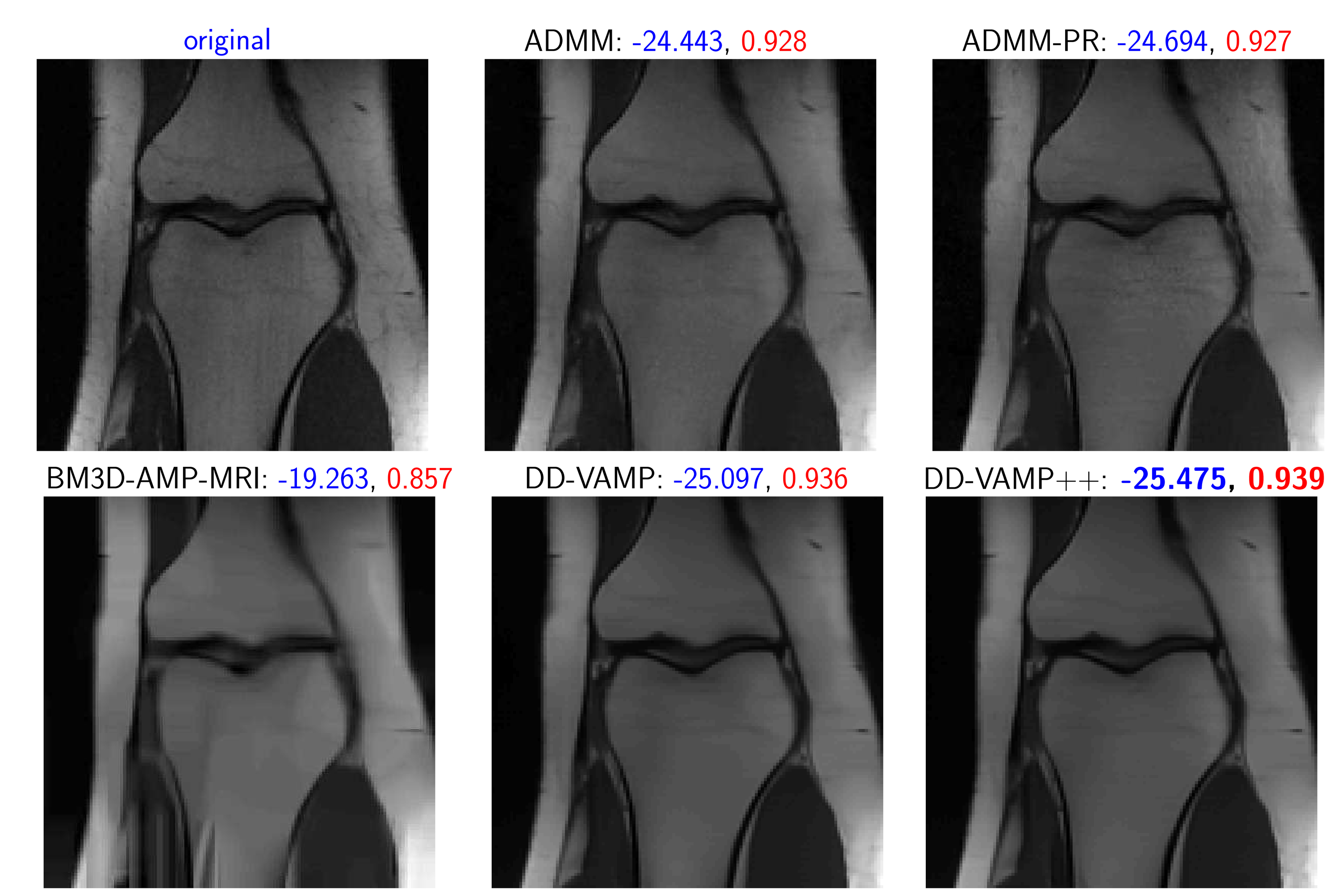
- Cartesian sampling mask $\boldsymbol{M}$ with acceleration $R = N/M = 4$.
- 128×128 mid-slice, non-fat-suppressed fastMRI knee images [11].
- DnCNN denoiser [12] used unless otherwise noted.

Training

- The dataset was randomly split into 30 training and 19 testing images.
- We tuned all algorithmic parameters to minimize NMSE averaged over iterations $t = 30 \dots 150$ and medianed over the training images.







Captions: NMSE (dB) and SSIM of example recovery after 150 iterations

References (cont.)

[7] S. Sarkar, "Solving linear and bilinear inverse problems using approximate message passing methods," Ph.D. dissertation, Ohio State, 2020.
[8] S. Rangan, P. Schniter, A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Info. Thy*'19.
[9] A. K. Fletcher, P. Pandit, S. Rangan, S. Sarkar, and P. Schniter, "Plug-in estimation in high-dimensional linear inverse problems: A rigorous analysis," *Proc. NeurIPS*, 2018.
[10] P. Schniter, S. Rangan, A. K. Fletcher, "Denoising based vector approximate message passing," *BASP*'17.
[11] J. Zbontar, F. Knoll, A. Sriram, et al., "fastMRI: An open dataset and benchmarks for accelerated MRI," *arXiv:1811.08839*, 2018.
[12] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," TIP'17.