D-OAMP: A DENOISING-BASED SIGNAL RECOVERY ALGORITHM FOR COMPRESSED SENSING

Zhipeng Xue⋆ Junjie Ma† Xiaojun Yuan⋆

⋆ School of Information Science and Technology, ShanghaiTech University †Department of Electronic Engineering, City University of Hong Kong

1. Compressed Sensing

In compressed sensing (CS), we aim to learn a sparse vector $\boldsymbol{x} \in \mathbb{R}^N$ from the measurement:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}, \tag{1}$$

where

- $N > M$,
- $\boldsymbol{A} \in \mathbb{R}^{M \times N}$ is a measurement matrix that mixes the sparse vector,
- $\boldsymbol{w}$ is an additive Gaussian noise.

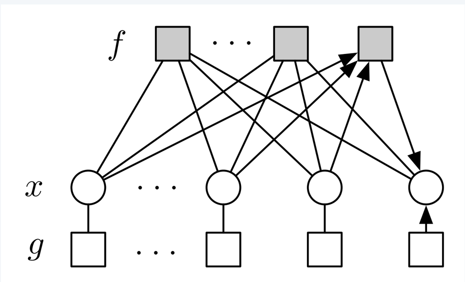
Many algorithms proposed for CS recovery including greedy algorithms and convex based algorithms. However, most of the algorithms converge slowly or have high complexity.

2. Approximate Message Passing (AMP)

AMP [1] is based on message passing. It's an efficient iterative algorithm with high converge rate and low complexity.

Algorithm 1 AMP algorithm

1: **inputs:**
  $\boldsymbol{A}, \boldsymbol{y}, \hat{\boldsymbol{x}}_0 = 0, \boldsymbol{z}_0 = \boldsymbol{y}$
2: **for** $t = 0, 1, 2, 3, \ldots, k-1$ **do**
3:   $\hat{\sigma}_t^2 = \frac{\|\boldsymbol{z}_t\|^2}{M}$
4:   $\boldsymbol{r}_t = \boldsymbol{x}_t + \boldsymbol{A}^T\boldsymbol{z}_t$
5:   $\hat{\boldsymbol{x}}_{t+1} = \eta_t(\boldsymbol{r}_t, \hat{\sigma}_t)$
6:   $\boldsymbol{z}_{t+1} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}_{t+1} + \boldsymbol{z}_t\text{div}\{D_t(\boldsymbol{r}_t)\}/M$
7: **end for**



The messages passing in a factor graph

- Intuitively, We can construct a factor graph of $\boldsymbol{x}$. By approximating messages in the factor graph under large N and dense $\boldsymbol{A}$, AMP can be derived.
- $\boldsymbol{z}_t\text{div}\{D_t(\boldsymbol{r}_t)\}/M$ is the onsager item. It's important to ensure that $\boldsymbol{r}_t$ is Gaussian distributed.
- When $\boldsymbol{A}$ is a i.i.d. Gaussian random matrix, a state evolution (SE) characterizes the MSE of AMP [2].
- **The performance of AMP can't be guaranteed for some matrixes, such as partial DCT matrix.**

3. Motivation

We want to design a CS recovery algorithm with the following characters:

- The explicit input signal prior is not needed.
- It can handle a variety of different types of random sensing matrices.

4. Orthogonal AMP (OAMP)

OAMP [3] works under various matrixes, such as i.i.d. random Gaussian matrix, random orthogonal matrix, etc.

Algorithm 2 OAMP algorithm

1: **inputs:**
  $\boldsymbol{A}, \boldsymbol{y}, \sigma^2, \hat{\boldsymbol{x}}_0 = 0, \boldsymbol{z}_0 = \boldsymbol{y}$
2: **for** $t = 0, 1, 2, 3 \ldots, k$ **do**
3:   $\hat{v}_t = \frac{\|\boldsymbol{z}_t\|^2 - M\sigma^2}{tr(\boldsymbol{A}^T\boldsymbol{A})}$
4:   $\boldsymbol{r}_t = \hat{\boldsymbol{x}}_t + \boldsymbol{W}_t\boldsymbol{z}_t$
5:   $\hat{\sigma}_t^2 = \Phi(\hat{v}_t^2)$
6:   $\hat{\boldsymbol{x}}_{t+1} = D_t(\boldsymbol{r}_t, \hat{\sigma}_t)$
7:   $\boldsymbol{z}_{t+1} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}_{t+1}$
8: **end for**

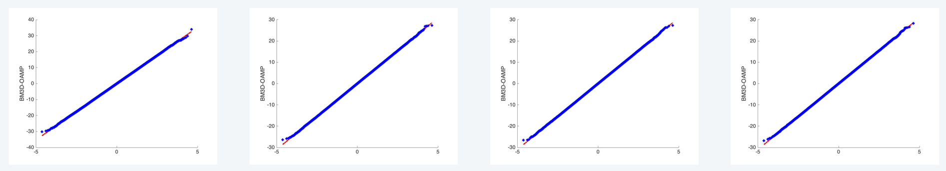
- The ideal of OAMP comes form Turbo principle [4]. Turbo principle is a kind of message passing principle.
- Extrinsic informations pass between two decoders like the messages passing through the factor graph.
- In OAMP the onsager item vanished, and a divergence free (divergence equals to 0) constraint is imposed on the nonlinear function $\eta_t(r_t)$.

5. A Key Observation

Similar to AMP, in each iteration of OAMP, $\boldsymbol{r}_t = \hat{\boldsymbol{x}}_t + \boldsymbol{W}_t\boldsymbol{z}_t$ can be modeled as a white Gaussian noise additive signal:

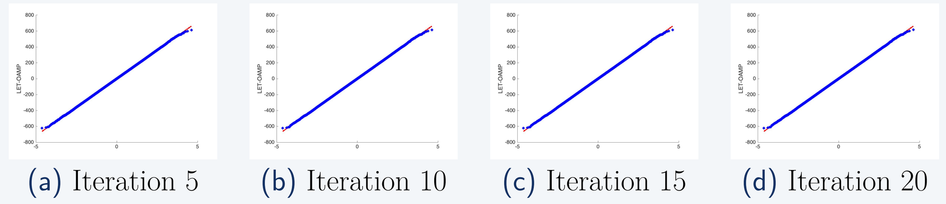
$$\boldsymbol{r}_t = \boldsymbol{x}_0 + \tau_t\boldsymbol{e}, \tag{2}$$

where $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$.
The following figures show the distribution of the Gaussian noise part of $\boldsymbol{r}_t$



(a) Iteration 5 (b) Iteration 10 (c) Iteration 15 (d) Iteration 20

QQplots of $\boldsymbol{r}$ at various iterations of BM3D-OAMP



(a) Iteration 5 (b) Iteration 10 (c) Iteration 15 (d) Iteration 20

QQplots of $\boldsymbol{r}$ at various iterations of LET-OAMP

6. State Evolution of OAMP

We use the SE below to characterize OAMP

$$\begin{aligned} \tau_t^2 &= \Phi(v_t^2), \\ v_{t+1}^2 &= \frac{1}{N}E\left\{\|D_t\left(\boldsymbol{x}_0 + \tau_t\boldsymbol{e}\right) - \boldsymbol{x}_0\|^2\right\}, \end{aligned} \tag{3}$$

where the expectation is taken over $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, and $v_0 = \|\boldsymbol{x}_0\|^2/N$. For partial orthogonal sensing matrix,

$$\Phi(v_t^2) = \frac{N-M}{M}v_t^2 + \sigma^2. \tag{4}$$

7. Denoising-based OAMP (D-OAMP)

D-OAMP is based on OAMP. We construct a divergence free denoiser from any chosen denoiser $\hat{D}(\boldsymbol{r})$:

$$D(\boldsymbol{r}) = C\left(\hat{D}(\boldsymbol{r}) - \frac{1}{N}\text{div}\left\{\hat{D}(\boldsymbol{r})\right\}\boldsymbol{r}\right). \tag{5}$$

$\text{div}\left\{\hat{D}(\boldsymbol{r})\right\}$ is the divergence of $\hat{D}(\boldsymbol{r})$. It's clear the divergence of (5) is zero.
Based on the model (2), the tuning of parameter C is based on stein's unbiased risk estimate (SURE) [5], we choose C by minimizing the SURE:

$$\begin{aligned} \widehat{\text{MSE}} &= \frac{1}{N}\|D(\boldsymbol{r}) - \boldsymbol{r}\|^2 + \frac{2\tau^2}{N}\text{div}\{D(\boldsymbol{r})\} - \tau^2 \\ &= \frac{1}{N}\|D(\boldsymbol{r}) - \boldsymbol{r}\|^2 - \tau^2 \\ &= \frac{1}{N}\|C\left(\hat{D}(\boldsymbol{r}) - \text{div}\{\hat{D}(\boldsymbol{r})\}\boldsymbol{r}\right) - \boldsymbol{r}\|^2 - \tau^2, \end{aligned} \tag{6}$$

where the second equality is due to the fact that $D(\boldsymbol{r})$ is divergence free. The optimal C that minimizes (6) (which is a quadratic function of C) is given by

$$C^\star = \frac{\boldsymbol{r}^T\left(\hat{D}(\boldsymbol{r}) - \text{div}\{\hat{D}(\boldsymbol{r})\}\boldsymbol{r}\right)}{\|\hat{D}(\boldsymbol{r}) - \text{div}\{\hat{D}(\boldsymbol{r})\}\boldsymbol{r}\|^2}. \tag{7}$$

8. Denoisers for D-OAMP

To apply D-OAMP in image applications, we exploit the rich literature on image denoising.

- SURE-LET denoiser
  Choosing denoiser $\hat{D}(\boldsymbol{r})$ as a SURE-LET denoiser [6]. A SURE LET denoiser is a linear combination of multiple elementary denoisers. Thus, the constructed divergence free denoiser has the following form:

$$D(\boldsymbol{r}) = \sum_{k=1}^{K} C_k\left(\hat{D}_k(\boldsymbol{r}) - \frac{1}{N}\text{div}\{\hat{D}_k(\boldsymbol{r})\}\boldsymbol{r}\right). \tag{8}$$

  Denote $\boldsymbol{C}^\star \equiv [C_1^\star, C_2^\star, \ldots, C_K^\star]^T$ where $\{C_k^\star\}$ is the optimal value of $\{C_k\}$ that minimize SURE. Let $\boldsymbol{G}_k \equiv \hat{D}_k(\boldsymbol{r}) - \text{div}\{\hat{D}_k(\boldsymbol{r})\}\boldsymbol{r}$, and define $M_{i,j} \equiv \boldsymbol{G}_i^T\boldsymbol{G}_j$, $\boldsymbol{b} \equiv [\boldsymbol{G}_1^T\boldsymbol{r}, \boldsymbol{G}_2^T\boldsymbol{r}, \ldots, \boldsymbol{G}_K^T\boldsymbol{r}]^T$. we obtain the following optimal combining coefficients:

$$\boldsymbol{C}^\star = \boldsymbol{M}^{-1}\boldsymbol{b}. \tag{9}$$

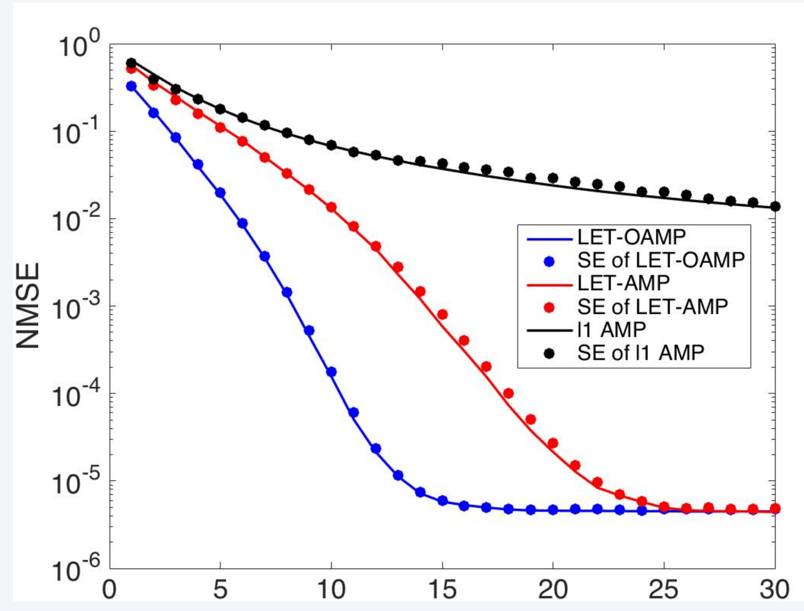
  In the numerical results, we choose $\hat{D}(\boldsymbol{r})$ as the the piecewise linear kernel in [7] and denote the algorithm as LET-OAMP.
- BM3D denoiser
  We can choose $\hat{D}(\boldsymbol{r})$ as a BM3D denoiser [8]. The divergence of BM3D denoiser has no explicit expression, we can use Monte Carlo method [9] to compute the divergence. Let $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ be an i.i.d. random Gaussian vector. The divergence of $D(\boldsymbol{r})$ can be estimated as:

$$\text{div}\left\{D(\boldsymbol{r}^t)\right\} \approx E_{\boldsymbol{e}}\left\{\boldsymbol{e}^T\left(\frac{D(\boldsymbol{r}^t + \delta\boldsymbol{e}) - D(\boldsymbol{r}^t)}{\delta}\right)\right\}, \tag{10}$$

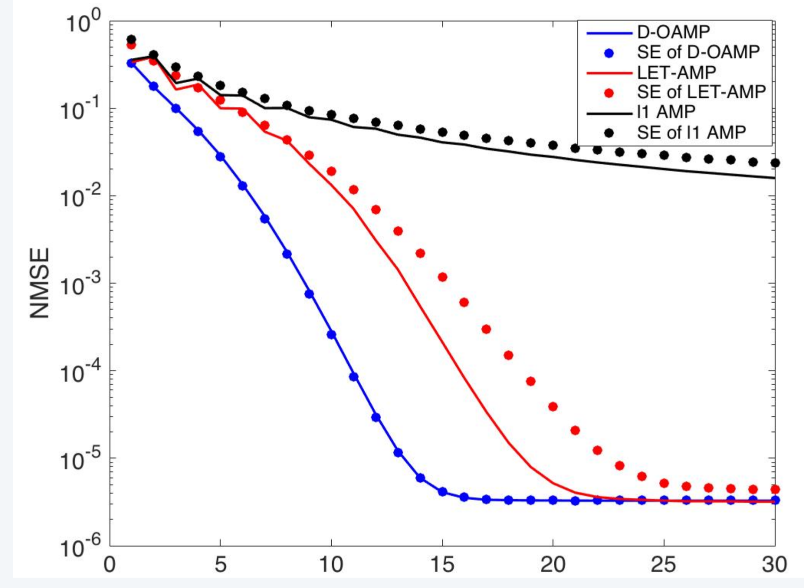
  where δ is a small constant. The expectation in (10) can be approximated by sample average. It is observed in [9] that one sample is good enough for high-dimensional problems.

9. Sparse Signal Recovery

We choose LET-OAMP to recover sparse signals. The performance characterized by normalized mean square error (NMSE) is shown in the following figures (noise level: 50dB, sparsity rate: 0.2, measurement rate: 0.5).



Gaussian random sensing matrix.



Partial DCT random sensing matrix

10. Natural Image Reconstruction

For natural image reconstruction, we adopt BM3D denoiser to D-OAMP (BM3D-OAMP). Under partial DCT random sensing matrix, the constructed images using BM3D-OAMP are shown in the following figures.

Original Barbara.

(a) BM3D-AMP, PSNR: 23.5dB. (b) BM3D-OAMP, PSNR: 27.1dB.

Reconstruction Barbara using D-AMP and D-OAMP, the measurement rate is 0.05.

When measurement rate is very low (up to 0.05), BM3D-OAMP can also reconstruct Barbara with high PSNR (27.1dB).

11. More Numerical Results

More detailed numerical results about natural image reconstruction using D-OAMP and D-AMP are listed below.

Image name	**Lena**			**Boat**			**Barbara**			**Fingerprint**		
Measurement rate	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**
EM-GM-GAMP	26.89	29.50	32.38	24.70	27.78	30.95	24.47	27.69	32.14	22.51	26.04	29.58
LET-AMP	22.53	30.78	33.74	21.51	29.00	32.42	19.69	28.92	32.46	17.48	27.25	31.11
LET-OAMP	**27.72**	**31.12**	**34.80**	**25.70**	**29.44**	**33.54**	**25.49**	**29.36**	**33.70**	**23.60**	**27.82**	**32.48**
BM3D-AMP	34.72	38.20	38.98	34.09	34.70	37.92	35.15	38.24	41.39	28.97	33.18	37.00
BM3D-OAMP	**35.50**	**38.26**	**42.36**	**34.59**	**37.60**	**41.28**	**36.51**	**39.83**	**43.35**	**31.01**	**35.24**	**39.79**

PSNR of reconstructed images under orthogonal random sensing matrix

Image name	**Lena**			**Boat**			**Barbara**			**Fingerprint**		
Measurement rate	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**	**30%**	**50%**	**70%**
LET-AMP	12.68	12.47	12.29	12.48	12.7	12.75	12.49	12.45	12.83	12.74	12.41	13.57
LET-OAMP	**4.60**	**3.97**	**2.72**	**4.39**	**3.86**	**3.33**	**4.32**	**3.72**	**3.45**	**7.95**	**5.1**	**4.92**

Reconstruction time of images under orthogonal random sensing matrix

D-OAMP outperform D-AMP in both reconstruction PSNR and time.

12. Conclusions

- The recovery quality of D-OAMP is the best of all existing CS recovery algorithms under partial orthogonal sensing matrixes.
- D-OAMP converges faster comparing to other existing algorithms under i.i.d. Gaussian and partial orthogonal matrix.
- By adopting different denoisers, D-OAMP can be used in many scenarios, such as image reconstruction and matrix recovery.

References

[1] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18914–18919, 2009.

[2] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," vol. 57, no. 2, pp. 764–785, Feb. 2011.

[3] J. Ma and L. Ping, "Orthogonal AMP," *arXiv preprint arXiv:1602.06509*, 2016.

[4] J. Ma, X. Yuan, and L. Ping, "Turbo compressed sensing with partial DFT sensing matrix," vol. 22, no. 2, pp. 158–161, Feb 2015.

[5] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *The annals of Statistics*, pp. 1135–1151, 1981.

[6] T. Blu and F. Luisier, "The SURE-LET approach to image denoising," vol. 16, no. 11, pp. 2778–2786, 2007.

[7] C. Guo and M. E. Davies, "Near optimal compressed sensing without priors: Parametric SURE approximate message passing," vol. 63, no. 8, pp. 2130–2141, 2015.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 606414–606414.

[9] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *arXiv preprint arXiv:1406.4175*, 2014.

Contact Information

Emails:

- Zhipeng Xue: xuezhp@shanghaitech.edu.cn
- Junjie Ma: hemajunjie@gmail.com
- Xiaojun Yuan: yuanxj@shanghaitech.edu.cn