# BM3D-AMP: A NEW IMAGE RECOVERY ALGORITHM BASED ON BM3D DENOISING

*Christopher A. Metzler*[⋆] *Arian Maleki*[†] *Richard G. Baraniuk*[⋆]

[⋆] Department of Electrical and Computer Engineering at Rice University
[†] Department of Statistics at Columbia University

## ABSTRACT

A denoising algorithm seeks to remove perturbations or errors from a signal. The last three decades have seen extensive research devoted to this arena, and as a result, today's denoisers are highly optimized algorithms that effectively remove large amounts of additive white Gaussian noise. A compressive sensing (CS) reconstruction algorithm seeks to recover a structured signal acquired from a small number of randomized measurements. Typical CS reconstruction algorithms can be cast as iteratively estimating a signal from a perturbed observation. This paper answers a natural question: How can one effectively employ a generic denoiser in a CS reconstruction algorithm? In response, we develop a denoising-based approximate message passing (D-AMP) algorithm that is capable of high-performance reconstruction. We demonstrate using the high performance BM3D denoiser that D-AMP offers state-of-the-art CS recovery performance for natural images (on average 9dB better than sparsity-based algorithms), while operating tens of times faster than the only competitive method. A critical insight in our approach is the use of an appropriate Onsager correction term in the D-AMP iterations, which coerces the signal perturbation at each iteration to be very close to the white Gaussian noise that denoisers are typically designed to remove. On the analytical side, we develop a new state evolution framework for deterministic signals that accurately predicts the performance of D-AMP and enables us to derive several useful theoretical features.

***Index Terms***— Compressive Sensing, Denoising, Approximate Message Passing, Onsager

## 1. INTRODUCTION

The fundamental problem in computational imaging is to reconstruct a signal from a series of measurements. When the measurements are linear, we can model the measurement process via the equation $y = \mathbf{A}x_o + w$, where $y$ is a length $m$ vector representing the measurements, $\mathbf{A}$ is an $m \times n$ matrix that models the measurement process, $x_o$ is a length $n$ vector representing our signal of interest, and $w$ is a vector representing measurement noise. A computational imaging system then seeks to use the measurements $y$, along with knowledge of the system matrix $\mathbf{A}$, to recover $x_o$. Recovering $x_o$ when the problem is underdetermined is known as compressive sensing (CS) and has been the focus of a vast amount of research over the past decade [1–18].

The work of C. Metzler supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. The work of C. Metzler and R. Baraniuk was supported in part and by the grants NSF CCF-0926127, NSF CCF-1117939, ONR N00014-12-1-0579, ONR N00014-11-1-0714, and ARO MURI W911NF-09-1-0383. The work of A. Maleki was supproted by the grant NSF CCF-1420328.

The majority of CS research has recovered $x_o$ by assuming that it is sparse in some transform basis. Unfortunately, sparsity-based methods are inappropriate for many imaging applications. The reason for this failure is that natural images do not have an exactly sparse representation in any known basis (DCT, wavelet, curvelet, etc.). As a result, algorithms that seek only wavelet-sparsity do a poor job of recovering the signal.

In response to this failure, researchers have considered more elaborate structures for CS recovery. These include minimal total variation [6, 7], block sparsity [8], wavelet tree sparsity [9, 10], hidden Markov mixture models [11–13], non-local self-similarity [14–16], and simple representations in adaptive bases [17,18]. Many of these approaches have led to significant improvements in imaging tasks.

In this paper, we take a complementary approach to enhancing the performance of CS recovery with natural images [19–21]. Rather than focusing on developing new signal models, we demonstrate how the existing rich literature on *signal denoising* can be leveraged for enhanced CS recovery. The idea is simple: Image denoising algorithms (whether based on an explicit or implicit model) have been developed and optimized for decades. Hence, any CS recovery scheme that employs a denoising algorithm should be able to capture elaborate structures that have heretofore not been captured by existing CS recovery schemes.

Inspired by this philosophy, we design a new CS recovery framework, called denoising-based approximate message passing D-AMP, that uses a denoising algorithm to recover signals from compressive measurements. D-AMP has several advantages over existing CS recovery algorithms: (i) It can be easily applied to many different signal classes. (ii) It outperforms existing CS recovery algorithms and is extremely robust to measurement noise. (iii) It comes with an analysis framework that not only characterizes the fundamental limits of the algorithm, but also suggests how we can best use D-AMP in practice.

Our work is not the first to use advanced denoisers to solve the CS recovery problem. Egiazarian et al. developed a denoising-based CS recovery algorithm [22] that uses the same authors' BM3D denoising algorithm [23] to impose a non-parametric model on the reconstructed signal. This method solves the CS problem when the measurement matrix is an undersampled linear transform. D-AMP is a fundamental improvement on this scheme. By Guassianizing the effective noise at each iteration our algorithm enables the true performance potential of the BM3D denoiser. In [16] it was shown that the BM3D-based algorithm performed considerably worse than NLR-CS. Therefore we compare our algorithm only with NLR-CS.

D-AMP is based on the AMP framework. AMP has been the subject of extensive research in the field of compressed sensing [3, 13, 21, 24–38]. Most previous work consider a Bayesian framework in which a signal prior $p_x$ is defined on the class of signals $C$, to which $x_o$ belongs. AMP is then considered as a heuristic approach

to calculate the posterior mean $\mathbb{E}(x_o \mid y, A)$. In the development of D-AMP, we are not concerned with whether the algorithm is approximating a posterior distribution for a certain prior or not. Instead, we rely on one important feature of AMP—that at every iteration AMP denoises a signal that is effectively the true signal $x_o$ plus i.i.d. Guassian noise. This important feature will be clarified later.

## 2. DENOISING-BASED APPROXIMATE MESSAGE PASSING

D-AMP assumes that $x_o$ belongs to a class of signals $C \subset \mathbb{R}^n$, such as the class of natural images of a certain size, for which a family of denoisers, $\{D_\sigma \ : \ \sigma > 0\}$ exists (the subscript $\sigma$ indicates that our denoiser is parameterized according to the standard deviation of the noise). Each denoiser $D_\sigma$ can be applied to $x_o + \sigma z$ with $z \sim N(0, I)$ and returns an estimate of $x_o$ that is hopefully closer to $x_o$ than $x_o + \sigma z$. In this paper we treat each denoiser as a black-box; it receives a signal with additive noise and returns an estimate of $x_o$. Hence, we do not assume any knowledge of the signal structure/information the denoising algorithm is employing to achieve its goal. This makes our derivations applicable to a wide variety of signal classes and a wide variety of denoisers.

D-AMP employs a denoiser in the following iteration:

$$\begin{aligned} x^{t+1} &= D_{\hat{\sigma}^t}(x^t + \mathbf{A}^* z^t), \\ z^t &= y - \mathbf{A}x^t + z^{t-1} D'_{\hat{\sigma}^{t-1}}(x^{t-1} + \mathbf{A}^* z^{t-1})/m. \end{aligned}$$

Here, $x^t$ and $z^t$ are the estimates of $x_o$ and the residue $y - Ax_o$ at iteration $t$ respectively. The term $\mathbf{A}^*$ denotes the conjugate transpose of $\mathbf{A}$ and $D'_{\hat{\sigma}^{t-1}}$ denotes the divergence of the denoiser. The term $z^{t-1} D'_{\hat{\sigma}^{t-1}}(x^{t-1} + \mathbf{A}^* z^{t-1})/m$ is known as the *Onsager correction term*. As in the original AMP algorithm, $x^t + \mathbf{A}^* z^t$ can be written as $x_o + v^t$, where $v^t$ is known as the effective noise. We can estimate the variance of $v^t$ with $(\hat{\sigma}^t)^2 = \frac{\|z^t\|_2^2}{m}$ [24]. D-AMP works well because the Onsager correction term Guassianizes $v^t$ and because typical denoisers are designed to handle additive white Gaussian noise. At this point, we are unable to prove $v^t$ is exactly Gaussian, but comprehensive simulations provided in the extended version of this paper [20] demonstrate its Gaussianity.

Note that the Onsager correction term contains the divergence of the denoiser. While simple denoisers yield an exact solution for their divergence, high-performance denoisers are often data dependent; making it very difficult to characterize their input-output relation explicitly. Therefore, we approximated the divergence using a Monte Carlo technique [39]: given a denoiser $D_\sigma(x)$, using an i.i.d. random vector $b \sim N(0, I)$, we can estimate the divergence with

$$\begin{aligned} D'_\sigma &= \lim_{\epsilon \to 0} E_b \left\{ b^* \left( \frac{D_\sigma(x + \epsilon b) - D_\sigma(x)}{\epsilon} \right) \right\} \qquad (1) \\ &\approx \mathbb{E}\left( \frac{1}{\epsilon} b^* (D_\sigma(x + \epsilon b) - D_\sigma(x)) \right), \text{ for very small } \epsilon. \end{aligned}$$

We used Monte Carlo sampling to approximate the expected value using the average of a series of random samples. We generate $M$ i.i.d., $N(0, I)$ vectors $b^1, b^2, \ldots, b^M$. For each vector $b^i$ we obtain a rough estimate of the divergence. We then obtain an improved estimate of the divergence by averaging. According to the weak law of large numbers, as $M \to \infty$ this estimate converges to $\mathbb{E}\left(\frac{1}{\epsilon} b^* (D^t_{\sigma,\tau}(x + \epsilon b) - D^t_{\sigma,\tau}(x))\right)$. Due to the high-dimensionality of the signal, we can accurately approximate the expected value using only a single random sample. As a result, calculation of the Onsager correction term is quite efficient and requires only one additional application of the denoising algorithm.

## 3. THEORETICAL ANALYSIS OF D-AMP

In this section we theoretically analyze the performance of D-AMP. This task is eased by the fact that our signal estimate at each iteration of the algorithm is computed by denoising a signal that is effectively $x_o + v^t$. Because the elements of $v^t$ follow an i.i.d. Gaussian distribution, the standard deviation $\sigma^t$ of $v^t$ completely determines the peformance of the algorithm.

### 3.1. State Evolution

We present a framework, called *state evolution*, to theoretically track the standard deviation of the noise, $\sigma^t$ at each iteration of D-AMP. Our framework extends the state evolution framework proposed in [3, 24]. Through extensive simulations we show that, in high-dimensional settings (for the denoisers that we consider in this paper), our state evolution predicts the mean square error (MSE) of D-AMP accurately.

Starting from $\theta^0 = \frac{\|x_o\|_2^2}{n}$ the state evolution generates a sequence of numbers through the following iterations:

$$\theta^{t+1}(x_o, \delta, \sigma_w^2) = \frac{1}{n} \mathbb{E} \| D_{\sigma^t}(x_o + \sigma^t z) - x_o \|_2^2, \qquad (2)$$

where $\sigma_w$ is the standard deviation of the measurment noise, $\delta := m/n$ is the under-determinacy of the problem, $(\sigma^t)^2 := \frac{\theta^t}{\delta}(x_o, \delta, \sigma_w^2) + \sigma_w^2$, and the expectation is with respect to $z \sim N(0, I)$. Note that our notation $\theta^{t+1}(x_o, \delta, \sigma_w^2)$ is set to emphasize that $\theta^t$ may depend on $x_o$, the under-determinacy $\delta$, and the measurement noise. Consider the iterations of D-AMP, and let $x^t$ denote its estimate at iteration $t$. Our empirical findings show that, in all the cases we studied, the MSE of D-AMP is predicted accurately by the state evolution. We formally state our finding.

**Finding 1.** *Assume the elements of $\mathbf{A}$ are i.i.d. and follow a sub-Gaussian distribution[1] with the following properties $\mathbb{E}(\mathbf{A}_{ij}) = 0$, $\mathbb{E}(\mathbf{A}_{ij}^2) = 1/n$ and $\mathbb{E}(\mathbf{A}_{ij}^4) = O(1/n^2)$. Under these conditions, if the D-AMP algorithm starts from $x^0 = 0$, then for large values of $m$ and $n$, state evolution predicts the mean square error of D-AMP.*

$$\theta^t(x_o, \delta, \sigma_w^2) \approx \frac{1}{n} \|x^t - x_o\|_2^2.$$

For more information on the simulations that lead to this finding and how it differs from similar findings in [24] and [34], please see Section III.C of [20]. In the next two sections we explore the implications of this finding on the performance and usage of D-AMP.

### 3.2. Recovery Performance

The following result shows that D-AMP is robust to measurement noise and can perfectly recover $x_o$ when no noise is present. This result requires formalizing our notion of denoiser slightly.

**Definition 1.** *$D_\sigma$ is called a proper family of denoisers of level $\kappa$ ($\kappa \in (0, 1)$) for the class of signals $C$ if and only if*

$$\sup_{x_o \in C} \frac{\mathbb{E} \| D_\sigma(x_o + \sigma \epsilon) - x_o \|_2^2}{n} < \kappa \sigma^2, \qquad (3)$$

*for every $\sigma > 0$. Note that the expectation is with respect to $\epsilon \sim N(0, I)$.*

[1] A random variable $X$ is said to be sub-Gaussian if there exists some real constant $c$ such that $\mathbb{E}(e^{tX}) \leq e^{ct^2}$ for any $t \in \mathbb{R}$. [40]

*It turns out that many of the denoisers that are popular in CS, including soft-thresholding and block soft-thresholding, are proper at certain levels. For more information see [20].*

Assume that the denoiser is proper at level $\kappa$. Let $\theta^{\infty}(x_o, \sigma_w^2, \delta)$ denote the fixed point of the state evolution equation. Define the noise sensitivity of D-AMP as

$$NS(\sigma_w^2, \delta) = \sup_{x_o \in C} \theta^{\infty}(x_o, \delta, \sigma_w^2).$$

The following proposition provides an upper bound for the noise sensitivity as a function of the number of measurements and the variance of the measurement noise.

**Proposition 1.** *Let $D_\sigma$ denote a proper family of denoisers at level $\kappa$. Then, for $\delta > \kappa$, the noise sensitivity of D-AMP satisfies*

$$NS(\sigma_w^2, \delta) \leq \frac{\kappa \sigma_w^2}{1 - \frac{\kappa}{\delta}}.$$

See [20] for proof.

**Remark 1.** *According to Finding 1, $\theta^{\infty}(x_o, \delta, \sigma_w^2)$ is equivalent to the MSE of the D-AMP algorithm. Therefore, Proposition 1 shows that the reconstruction MSE is proportional to the variance of the measurement noise. As the number of measurements decreases, the D-AMP reconstruction MSE increases. Fortunately, the increase is mild as long as $\delta$ is not too close to $\kappa$.*

**Remark 2.** *If $\delta > \kappa$, then as $\sigma_w^2 \to 0$, $\sup_{x_o \in C} \theta^{\infty}(x_o, \delta, \sigma_w^2) \to 0$. Combined with Finding 1, this implies D-AMP will have perfect recovery in the noiseless setting.*

### 3.3. Parameter Tuning

In this section we demonstrate how our theoretical framework simplifies the problem of parameter tuning. At first glance, the problem of tuning the D-AMP parameters appears daunting. To state this challenge, we overload our notation of a denoiser to $D_{\sigma,\tau}$, where $\tau$ denotes the set of parameters of the denoiser. According to this notation, the state evolution is given by

$$o^{t+1}(\tau^0, \tau^1, \ldots, \tau^t) = \frac{1}{n} \mathbb{E} \| D_{\sigma^t, \tau^t}(x_o + \sigma^t z) - x_o \|_2^2,$$

where $(\sigma^t)^2 = \frac{o^t(\tau^0, \tau^1, \ldots, \tau^{t-1})}{\delta} + \sigma_w^2$. Note that we have changed our notation for the state evolution variables to emphasize the dependence of $o^{t+1}$ on the choice of the parameters we pick at the previous iterations.

**Definition 2.** *A sequence of parameters $\tau_*^1, \ldots, \tau_*^t$ is called optimal at iteration $t+1$ if*

$$o^{t+1}(\tau_*^0, \ldots, \tau_*^t) = \min_{\tau^0, \tau^1, \ldots, \tau^t} o^{t+1}(\tau^0, \tau^1, \ldots, \tau^t).$$

Note that $\tau_*^0, \ldots, \tau_*^t$ produce the smallest mean square error D-AMP can achieve after $t$ iterations.

It seems from our formulation that we should solve a joint optimization on $\tau^1, \ldots, \tau^t$ to obtain the optimal values of these parameters. However, it turns out that in D-AMP the optimal parameters can be found much more easily. Consider the following greedy parameter tuning algorithm:

1. Tune $\tau^0$ such that $o^1(\tau^0)$ is minimized. Call the optimal value $\tau_*^0$.
2. If $\tau^0, \ldots, \tau^{t-1}$ are set to $\tau_*^0, \ldots, \tau_*^{t-1}$, then set $\tau^t$ such that it minimizes $o^{t+1}(\tau_*^0, \ldots, \tau_*^{t-1}, \tau^t)$.

This optimization is similar to the optimization of the parameters in the denoising literature [39, 41]. It turns out that in the context of D-AMP this strategy is optimal. The following result proves the optimality of this greedy strategy.

**Proposition 2.** *Suppose that the denoiser $D_{\sigma,\tau}$ is monotone in the sense that*

$$\inf_{\tau} \mathbb{E} \| D_{\sigma,\tau}(x_o + \sigma\epsilon) - x_o \|_2^2$$

*is a non-decreasing function of $\sigma$. If $\tau_*^0, \ldots, \tau_*^t$ is generated according to the greedy tuning algorithm described above, then $o^{t+1}(\tau_*^0, \ldots, \tau_*^t) \leq o^{t+1}(\tau^0, \ldots, \tau^t), \forall \tau^0, \ldots, \tau^t$.*

See [20] for proof.

To summarize this result, to tune D-AMP's parameters over many iterations simply tune a denoiser once at each iteration. There exists an extensive literature on tuning the free parameters of denoisers [41]. Diverse and powerful algorithms such as SURE (Stein's Unbiased Risk Estimation) have been proposed for this purpose [39]. D-AMP can employ any of these tuning schemes.

## 4. SIMULATION RESULTS

In this section, we compare the performance of D-AMP, based on the BM3D [23] and BM3D-SAPCA [42] denoisers, with other CS reconstruction algorithms.[2] (BM3D-SAPCA is an extension to the popular BM3D denoiser. It offers incrementally better performance than BM3D but is hundreds of times slower). In particular, we compare the performance of these D-AMP algorithms, called BM3D-AMP and BM3D-SAPCA-AMP, with the original AMP algorithm [3], turbo-AMP [13], and NLR-CS [16]. AMP uses a wavelet sparsity model (Daubechies-4 wavelets), turbo-AMP uses a hidden Markov tree model, and NLR-CS uses a low-rank self-similarity matrix model. NLR-CS represents the current state-of-the-art in CS reconstruction algorithms. It was shown in [16] and [20] that NLR-CS dramatically outperforms TV sparsity methods [43], the original BM3D CS recovery algorithm [22], and Model-CoSaMP [11]. Therefore results for these methods are not presented here.

The parameters of D-AMP were set following the methods described in Section 3.3. Both D-AMP algorithms were run for 30 iterations. AMP was run for 30 iterations as well. Turbo-AMP was run for 10 iterations. (We experimented with running turbo-AMP for 30 iterations but found that this yielded no improvement in performance while nearly tripling the computation time). Because the DCT iterative soft-thresholding method used to generate an intitial estimate in NLR-CS's provided source code[3] failed for Gaussian measurement matrices, we generated the initial estimates used by NLR-CS by running BM3D-AMP for 8 iterations for noiseless tests and 4 iterations for noisy tests. Only 4 iterations of BM3D-AMP were used during noisy tests because, if 8 iterations were used, the initial estimates from BM3D-AMP were often better than the final estimates from NLR-CS. Turbo-AMP and NLR-CS were otherwise tested under their default settings.

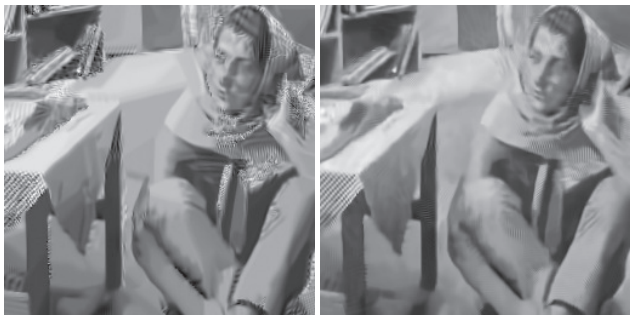
We report the average recovery performance, in terms of PSNR[4], across 6 standard test images[5]: Lena, Barbara, Boat, Fingerprint,

[2] Our code is available at http://dsp.rice.edu/software/DAMP-toolbox

[3] http://see.xidian.edu.cn/faculty/wsdong/NLR_Exps.htm

[4] PSNR stands for peak signal-to-noise ratio and is defined as $20 \log_{10}(\frac{255^2}{\text{mean}(\hat{x}^2 - x_o^2)})$ when the pixel range is 0 to 255.

[5] http://decsai.ugr.es/~javier/denoise/test_images/



(a) NLR-CS recovery (b) BM3D-SAPCA-AMP recovery

**Fig. 1**. Reconstructions of 10% sampled $256 \times 256$ Barbara image with additive white Gaussian measurement noise of standard deviation 30. Notice BM3D-SAPCA-AMP exhibits far fewer artifacts than NLR-CS.

House, and Peppers. All images have pixel values between 0 and 255 and were rescaled to $128 \times 128$ (with the exception of the higher resolution examples displayed above). All tests used an $m \times n$ measurement matrix that was generated by first using Matlab's `randn(m,n)` command and then normalizing the columns. All simulations were conducted on a 3.16 GHz Xeon quad-core processor with 32GB of memory.

Tables 1-3 summarize the results of our experiments. Table 1 compares the performance of different recovery algorithms at various sampling rates[6] when no measurement noise is present. BM3D-SAPCA-AMP outperforms all the other algorithms in a majority of the tests. In Table 2 we compare the performance of BM3D-AMP to NLR-CS when measurement noise of standard deviation 20 is present. As one might expect, D-AMP is exceptionally robust to noise. Table 3 demonstrates that, depending on the denoiser in use, D-AMP can be quite computationally efficient: The BM3D variant of D-AMP is dramatically faster than NLR-CS. The table also illustrates how using different denoisers within D-AMP presents not only a means of capturing different signal models, but also a way to balance performance and run times. For more detailed results please see [20].

**Table 1**. Average PSNR of reconstructions without noise

| **Sampling Rate (%)** | **10** | **20** | **30** | **40** | **50** |
|---|---|---|---|---|---|
| AMP | 18.07 | 20.51 | 22.85 | 24.96 | 27.07 |
| Turbo-AMP | 18.25 | 22.43 | 24.69 | 26.64 | 28.85 |
| NLR-CS | **25.57** | 29.89 | 33.15 | 35.78 | 38.28 |
| BM3D-AMP | 25.18 | 29.66 | 32.72 | 35.18 | 37.41 |
| BM3D-SAPCA-AMP | 22.81 | **30.33** | **33.63** | **36.15** | **38.45** |

**Table 2**. Average PSNR of reconstructions with noise

| **Sampling Rate (%)** | **10** | **20** | **30** | **40** | **50** |
|---|---|---|---|---|---|
| NLR-CS | 23.20 | 24.77 | 25.35 | 25.05 | 23.74 |
| BM3D-AMP | 24.14 | 26.34 | 27.39 | 27.76 | 27.93 |

[6]Sampling rate is defined to be $m/n \times 100$.

**Table 3**. Average computation times, in minutes

| **Sampling Rate (%)** | **10** | **20** | **30** | **40** | **50** |
|---|---|---|---|---|---|
| AMP | 0.3 | 0.6 | 1.0 | 1.3 | 1.6 |
| Turbo-AMP | 1.4 | 2.4 | 3.4 | 4.5 | 5.5 |
| NLR-CS | 31.6 | 60.6 | 88.1 | 122.8 | 152.2 |
| BM3D-AMP | 1.0 | 1.3 | 1.5 | 1.7 | 2.2 |
| BM3D-SAPCA-AMP | 318.3 | 328.7 | 345.1 | 362.1 | 378.0 |

## 5. DISCUSSION

Through extensive testing we have demonstrated that the approximate message passing (AMP) CS recovery algorithm can be extended to use general denoisers to great effect. Variations of this denoising-based AMP (D-AMP) algorithm deliver state-of-the-art compressively sampled image recovery performance while maintaining a low computational footprint. Our theoretical results and computational simulations show that the performance of D-AMP can be predicted accurately by state evolution. We have also proven that the problem of tuning the parameters of D-AMP is no more difficult than the tuning of the denoiser that is used in the algorithm. Finally, we have shown that D-AMP is robust to measurement noise.

D-AMP represents a "plug and play" method to recover compressively sampled signals; simply choose a denoiser well-matched to the signal model and plug it in the AMP framework. Since designing a denoising algorithm is easier than designing a recovery algorithm, D-AMP can benefit many different application areas.

## 6. REFERENCES


[1] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203 – 4215, Dec. 2005.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[3] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proc. Natl. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, 2009.

[4] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Select. Top. Signal Processing*, vol. 1, no. 4, pp. 586–598, 2007.

[5] M. Elad, B. Matalon, J. Shtok, and M. Zibulevsky, "A wide-angle view at iterated shrinkage algorithms," *Proc. SPIE (Wavelet XII)*, Aug. 2007.

[6] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[7] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint," *Magn. Reson. Med.*, vol. 57, pp. 1086–1098, Jun. 2007.

[8] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Processing*, vol. 58, pp. 3042–3054, Jun. 2010.

[9] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inform. Theory*, vol. 56, pp. 1982 –2001, Apr. 2010.

[10] C. Hedge, P. Indyk, and L. Schmidt, "A fast approximation algorithm for tree-sparse recovery," *Preprint*, 2014.

[11] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden markov tree model," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, pp. 5137–5140, Mar. 2008.

[12] Y. Kim, M. S. Nadar, and A. Bilgin, "Compressed sensing using a gaussian scale mixtures model in wavelet domain," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, pp. 3365–3368, Sept. 2010.

[13] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Trans. Signal Processing*, vol. 60, no. 7, pp. 3439–3448, 2012.

[14] G. Peyre, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Computer Vision ECCV 2008* (D. Forsyth, P. Torr, and A. Zisserman, eds.), vol. 5304 of *Lecture Notes in Computer Science*, pp. 57–68, Springer Berlin Heidelberg, 2008.

[15] J. Zhang, S. Liu, and D. Zhao, "Improved total variation based image compressive sensing recovery by nonlocal regularization," *Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS)*, pp. 2135–2138, Mar. 2013.

[16] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," to appear in *IEEE Trans. Image Processing*, 2014.

[17] C. Chen, E. W. Tramel, and J. E. Fowler, "Compressed-sensing recovery of images and video using multihypothesis predictions," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, pp. 1193–1198, Nov. 2011.

[18] "Image compressive sensing recovery using adaptively learned sparsifying basis via $\ell_0$ minimization," *Signal Processing*, vol. 103, no. 0, pp. 114 – 126, 2014.

[19] C. Metzler, A. Maleki, and R. Baraniuk, "From denoising to compressed sensing," *ELEC599 Project Report*, Apr. 2014.

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

[21] J. Tan, Y. Ma, and D. Baron, "Compressive imaging via approximate message passing with image denoising," *arXiv preprint:1405.4429*, 2014.

[22] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, vol. 1, pp. I – 549–I – 552, Sept. 2007.

[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. Image Processing*, vol. 16, pp. 2080–2095, Aug. 2007.

[24] A. Maleki, "Approximate message passing algorithm for compressed sensing," *Stanford University PhD Thesis*, Nov. 2010.

[25] D. L. Donoho, A. Maleki, and A. Montanari, "Noise sensitivity phase transition," *IEEE Trans. Inform. Theory*, vol. 57, Oct. 2011.

[26] A. Maleki, L. Anitori, Z. Yang, and R. G. Baraniuk, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," *IEEE Trans. Inform. Theory*, vol. 59, pp. 4290–4308, Jul. 2013.

[27] D. L. Donoho, A. Maleki, and A. Montanari, "Construction of message passing algorithms for compressed sensing," *Preprint*, 2010.

[28] P. Schniter, "Turbo reconstruction of structured sparse signals," in *Proc. IEEE Conf. Inform. Science and Systems (CISS)*, Mar. 2010.

[29] S. Kudekar and H. D. Pfister, "The effect of spatial coupling on compressive sensing," in *Proc. Allerton Conf. Communication, Control, and Computing*, pp. 347–353, 2010.

[30] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2168–2172, 2011.

[31] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing," *arXiv preprint:1310.2632*, 2013.

[32] F. Krzakala, M. Mézard, F. Sausset, Y. F. Sun, and L. Zdeborová, "Statistical-physics-based reconstruction in compressed sensing," *Physical Review X*, vol. 2, no. 2, p. 021005, 2012.

[33] D. L. Donoho, A. Javanmard, and A. Montanari, "Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 1231–1235, 2012.

[34] D. L. Donoho, I. Johnstone, and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising," *IEEE Trans. Inform. Theory*, vol. 59, no. 6, pp. 3396–3433, 2013.

[35] A. Maleki and A. Montanari, "Analysis of approximate message passing algorithm," in *Proc. IEEE Conf. Inform. Science and Systems (CISS)*, pp. 1–7, Mar. 2010.

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

[37] J. Barbier, F. Krzakala, M. Mézard, and L. Zdeborová, "Compressed sensing of approximately-sparse signals: Phase transitions and optimal reconstruction," in *Proc. Allerton Conf. Communication, Control, and Computing*, pp. 800–807, 2012.

[38] M. Borgerding and P. Schniter, "Generalized approximate message passing for the cosparse analysis model," *arXiv preprint:1312.3968*, 2013.

[39] S. Ramani, T. Blu, and M. Unser, "Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms," *IEEE Trans. Image Processing*, pp. 1540–1554, 2008.

[40] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *Arxiv preprint:1011.3027v7*, 2011.

[41] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Processing*, vol. 1, pp. 322–336, Jul. 1992.

[42] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Bm3d image denoising with shape-adaptive principal component analysis," in *Proc. Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS09)*, 2009.

[43] "http://users.ece.gatech.edu/~justin/l1magic/,"