BM3D-PRGAMP: COMPRESSIVE PHASE RETRIEVAL 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

The explosion of computational imaging has seen the frontier of image processing move past linear problems, like denoising and deblurring, and towards non-linear problems such as phase retrieval. There has a been a corresponding research thrust into non-linear image recovery algorithms, but in many ways this research is stuck where linear problem research was twenty years ago: Models, if used at all, are simple designs like sparsity or smoothness.

In this paper we use denoisers to impose elaborate and accurate models in order to perform inference on generalized linear systems. More specifically, we use the state-of-the-art BM3D denoiser within the Generalized Approximate Message Passing (GAMP) framework to solve compressive phase retrieval in a variety of different contexts. Our method demonstrates recovery performance equivalent to existing techniques using *fewer than half* as many measurements. This dramatic improvement in compressive phase retrieval performance opens the door for a whole new class of imaging systems.

Index Terms— Compressive Phase Retrieval, Denoising, Generalized Approximate Message Passing

1. INTRODUCTION

Compressive phase retrieval is the problem of recovering a linearly subsampled complex signal without any phase information. That is, if $z = \mathbf{A}x_o$, with $\mathbf{A} \in \mathbb{C}^{m\times n}$, then we observe

$$y_i = |z_i + \epsilon_i|,$$

where ϵ_i is noise, and would like to estimate the vector x_o. The phase retrieval problem shows up in numerous engineering disciplines, including crystallography [1], ptychography [2], compressive imaging systems [3, 4], and more.

Phase retrieval is a relatively well studied problem and a number of methods exists to solve it. One popular technique is alternating minimization; pioneered by the Fienup [5], Gerchberg-Saxton [6], and Griffin-Lim [7] algorithms and extended in [8–12], among others. Another popular approach is convex relaxations, including PhaseLift [13], PhaseCut [14], and Douglas-Rachford methods [15]. The Wirtinger flow algorithm [16], a stochastic-gradient-descent-like algorithm, has recently earned lots of attention for its low computational cost and strong theoretical guarantees.

Starting with [17], a parallel line of research has focused on designing algorithms that, through the use of prior information, can reconstruct signals using far fewer magnitude only measurements. Such methods include modified versions of the Fienup algorithm [18], GESPAR [19], CPRL [20], TSPR [21], a low-rank and sparse method [22], and many others. Unfortunately, a large number of these algorithms do not scale well to sizes of practical interest. Of the ones that do, phase-retrieval GAMP (prGAMP) [23] has shown by far the best performance at the compressive phase retrieval task.

Thus far, compressive phase retrieval research has focused on efficiently using simple structures such as sparsity. However, many natural signals, including images, exhibit far more complicated structures. The goal of this paper is to incorporate such complicated structure priors into a phase retrieval algorithm. Toward this goal, we use the approximate message passing framework [23–25] and denoisers, which together present a flexible framework for this purpose.

To be more specific, suppose that we are interested in recovering a signal x_o of known class $\mathcal{C}$. Furthermore, assume that for this class of signals we have access to a denoising algorithm. Denoising algorithms estimate $x_o \in \mathcal{C}$ from $x_o + \sigma w$, where $w \sim N(0, I)$. We denote a denoising algorithm with $D_\sigma(\cdot)$. In this paper we focus on imaging application and pursue two goals: (i) We show how one can incorporate complicated denoisers whose explicit forms are not given into the prGAMP framework. (ii) We demonstrate that if state-of-the-art denoisers such as BM3D [26] are incorporated into the algorithm then they produce a state-of-the-art phase retrieval algorithm.

2. RELATED WORK

2.1. Denoisers as Priors

Denoisers have recently been recognized as a powerful tool for performing inference and solving inverse problems. In [27] and [28] the authors use denoisers to regularize tomographic reconstructions and perform sparse interpolation. In [29] the authors use the BM3D denoiser [26] as a regularizer in compressive sensing, super-resolution, and upsampling. In [30] the BM3D denoiser was integrated onto a GPU and used as a regularizer to simultaneously solve demosaicing, denoising, deconvolution, and inpainting.

2.2. Message Passing Algorithms Applied to Phase Retrieval

prGAMP [23] was the first message passing algorithm used for phase retrieval. We will discuss prGAMP in more detail in Section 3. Here we would like to mention two related phase retrieval message passing algorithms; prVBEM [31] and prSAMP [32].

prVBEM is a message passing algorithm derived using mean-field approximations. Unlike prGAMP, it comes with provable convergence guarantees. In testing it is shown to offer performance similar to prGAMP when the sampled signal follows a Gaussian distribution (not sparse).

prSAMP is a combination of prGAMP and swept AMP [33]. prSAMP demonstrates performance superior to prVBEM, is robust to ill-conditioned measurement matrices $\mathbf{A}$, and can incorporate sparsity priors. However, at the time of this writing, the code was

The work of C. Metzler supported by the NSF GRF Program. The work of C. Metzler and R. Baraniuk was supported in part and by the grants NSF CCF1527501, AFOSR FA9550-14-1-0088, ARO W911NF-15-1-0316, ONR N00014-12-1-0579, and DoD HM04761510007. The work of A. Maleki was supported by the grant NSF CCF-1420328.

not publicly available to compare against. While prSAMP, as an AMP-type algorithm, would seem to be a good candidate for use with a high-performance denoiser, the sequential nature of prSAMP complicates the use of a non-separable denoiser within its iterations.

3. DENOISING-BASED PHASE RETRIEVAL GAMP

As mentioned earlier, prGAMP [23] offers state-of-the-art sparse signal recovery performance. This accomplishment is achieved via the use of a simple, separable denoiser designed for Bernoulli-Gaussian distributions. The authors of prGAMP also recognized that the algorithm could support non-separable priors on x and mention turbo GAMP [34] and analysis GAMP [35] as possible extensions. In this work we seize upon this insight and apply high performance denoisers within prGAMP.

3.1. Phase Retrieval GAMP

Approximate message passing (AMP) algorithms were first introduced in [24] as fast and iterative schemes for recovering x_o from undersampled linear observations of the form $y = \mathbf{A}x_o + \epsilon$. Given an initialization of the form $x^0 = 0$ and $z^0 = y$, AMP updates its estimate, x^t, according to the following iterations:

$$\begin{aligned} x^{t+1} &= \eta(x^t + \mathbf{A}^* z^t), \\ z^t &= y - \mathbf{A}x^t + \frac{n}{m}\langle \eta'(x^{t-1} + \mathbf{A}^* z^{t-1})\rangle z^{t-1}, \end{aligned} \quad (1)$$

where η is a nonlinear function applied element-wise to $x^t + \mathbf{A}^* z^t$ and imposes the structure of the data. For instance, one may use the soft thresholding function, $\eta_\tau(u) = (|u|-\tau)\mathbb{I}(|u|-\tau \geq 0)\text{sign}(u)$, where $\mathbb{I}(\cdot)$ denotes the indicator function, to impose sparsity. $\eta'(\cdot)$ denotes the derivative of η, $< \cdot >$ denotes taking the average, and z^t may be considered as an estimate of the residual $y - \mathbf{A}x_o$ at time t. One of the main advantages of AMP is that it comes with a theoretical framework, called state evolution, that predicts its performance accurately. For the sake of brevity we do not review the state evolution here and instead refer the reader to [36] and [37].

The original AMP algorithm was designed to address independent and identically distributed additive noise on the measurements. However, in certain applications, such as quantized compressed sensing [38] and phase retrieval the algorithm must deal with more diverse distortions. To address this issue [25] proposed the generalized approximate message passing. Let $z = \mathbf{A}x_o$ denote the ideal measurements before any noise or distortion being added. Then [25] models the observations y as a sample from the distribution $y_i = p_{Y|Z}(y_i|z_i)$. The iterations of GAMP are given by

$$\begin{aligned} x^{t+1} &= \eta(x^t + \alpha^t \mathbf{A}^* s^t), \\ z^t &= y - \mathbf{A}x^t + \frac{n}{m}\langle \eta'(x^{t-1} + \alpha^{t-1}\mathbf{A}^* z^{t-1})\rangle s^{t-1}, \\ s^t &= \eta_{out}(y, z^t; \sigma_z^t). \end{aligned} \quad (2)$$

The function $\eta_{out}(y_i, z_i; \sigma)$ is given by

$$\eta_{out}(y_i, z_i; \sigma_z) \triangleq \frac{\int \tilde{z}_i p_{Y|Z}(y_i|\tilde{z}_i) N(\tilde{z}_i; z_i, \sigma_z) d\tilde{z}_i}{\int p_{Y|Z}(y_i|\tilde{z}_i) N(\tilde{z}_i; z_i, \sigma_z) d\tilde{z}_i},$$

where $N(\tilde{z}_i; z_i, \sigma_z)$ denotes the pdf of a Gaussian with mean z_i and variance σ_z.

Note that σ_z is the standard deviations of z^t and hence can be easily estimated by $\frac{1}{m}\|z^t\|_2^2$. In this work we set $\alpha^t = [\frac{1}{m}\sum_{i=1}^m \frac{\partial}{\partial z_i}\eta_{out}(y_i, z_i^t; \sigma_z)]^{-1}$, but it can be reduced to damp the algorithm. See [23] for more information. It is important to note the following features of AMP that are also inherited by GAMP.

1. For i.i.d. Gaussian random matrices, $\mathbf{A}$, $x^t + \alpha^t \mathbf{A}^* z^t$ can be modeled as $x_o + \nu^t$, where ν^t can be approximated by an i.i.d. Gaussian noise. This property has been explored extensively in the context of AMP in [37] and will be used later in our paper.
2. The main purpose of applying η is to impose the structure of x_o on the estimates of the AMP (GAMP) algorithm.

As demonstrated in [23], one can employ GAMP to solve the phase retrieval problem. By considering the model $y_i = |z_i + \epsilon_i|$, where $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, [39] has shown that

$$p_{Y|Z}(y_i|z_i; \sigma_\epsilon) = \frac{2y_i}{\sigma_\epsilon^2} \mathrm{e}^{-\frac{y_i^2 + |z_i|^2}{\sigma_\epsilon^2}} I_0\left(\frac{2y_m|z_m|}{\sigma_\epsilon^2}\right)\mathbb{I}(y_i \geq 0), \quad (3)$$

where $I_0(\cdot)$ denotes the 0^{th}-order Bessel function of the first kind and $\mathbb{I}(\cdot)$ denotes the indicator function.

In [23] the authors used this prior to calculate a function closely related to $\eta_{out}(\cdot)$.

3.2. Denoising-based Phase Retrieval GAMP

Suppose that x_o belongs to signal class $\mathcal{C}$ and that for this class we have a family of denoising algorithms D_σ indexed by standard deviation σ that are capable of estimating vectors $x_o \in \mathcal{C}$ from $x_o + \sigma\epsilon$, where $\epsilon \sim N(0, I)$. In other words, we expect the denoiser to satisfy

$$\mathbb{E}\frac{\|D_\sigma(x_o + \sigma\epsilon) - x_o\|_2^2}{\sigma^2} \ll n,$$

for all values of σ. For instance, $\mathcal{C}$ could denote the class of natural images and D_σ could be the BM3D denoiser. Our goal is to use this denoising algorithm to recover x_o. As we discussed already, in GAMP $x^t + \mathbf{A}^* z^t$ can be modeled as $x_o + \nu^t$, where ν^t has approximately a Gaussian distribution with mean zero and variance σ_z^2.[1] Hence, intuitively one may replace the η function in (2) with any denoiser that was designed for additive white Gaussian noise. However, following this replacement, we then need to determine the correction term $\frac{n}{m}\langle \eta'(x^{t-1} + \alpha^{t-1}\mathbf{A}^* z^{t-1})\rangle s^{t-1}$. Note that the explicit form of $D_\sigma(\cdot)$ is usually not known in practice.

To solve this problem, we introduce Denoising-based prGAMP (D-prGAMP) that employs the following iterations:

$$\begin{aligned} x^{t+1} &= D(x^t + \alpha^t \mathbf{A}^* s^t), \\ z^t &= y - \mathbf{A}x^t + \frac{n}{m}\langle \text{div} D(x^{t-1} + \alpha^{t-1}\mathbf{A}^* z^{t-1})\rangle s^{t-1}, \\ s^t &= \eta_{out}(y, z^t; \sigma_z^t), \end{aligned} \quad (4)$$

where the notation $\text{div}D$ denotes the divergence of D.[2] To calculate the divergence we use an approximation that was introduced in [37, 40]. We briefly review the approximation here. Given a denoiser $D_\sigma(x)$, using an i.i.d. random vector $b \sim N(0, I)$, we can estimate the divergence with

$$\begin{aligned} \text{div}D_\sigma &= \lim_{\epsilon \to 0} E_b\left\{ b^* \left(\frac{D_\sigma(x + \epsilon b) - D_\sigma(x)}{\epsilon} \right) \right\} \quad (5) \\ &\approx \mathbb{E}\left(\frac{1}{\epsilon} b^* (D_\sigma(x + \epsilon b) - D_\sigma(x)) \right), \text{ for very small } \epsilon. \end{aligned}$$

[1] The Gaussianity of the effective noise, ν^t, has been proven only for scalar denoisers. Limited experimentation suggests that ν^t in D-prGAMP is still Gaussian. Upcoming research will further investigate this claim.

[2] Although it has not been said explicitly, the denoisers will be using the standard deviation of ν^t. Hence that value should be estimated.

To compute this expected value we can use Monte Carlo sampling:

$$\text{div}\hat{D}_\sigma = \frac{1}{M}\sum_{i=1}^{M} \widehat{\text{div}}^{\,i},$$

where $\widehat{\text{div}}^{\,i}$ is the estimate of the divergence from the i^{th} Monte Carlo sample. In this work set $\epsilon = \frac{\|x\|_\infty}{1000}$ and $M = 1$. For more information on this approximation see [37].

3.3. Heurestic Adaptations

AMP and GAMP have well understood and predictable behavior only when the matrix $\mathbf{A}$ follow an i.i.d. sub-Gaussian distribution. Because practical matrices tend to have far more structure, a large amount of literature has been devoted to dealing with this problem. See, for example, [41] and [42]. We built our algorithm off of the Dec. 22, 2015 version of the original prGAMP code[3]. As a result, our algorithm takes advantage of several different heuristics, including multiple restarts, damping, and mean-removal. For more information about each of these methods see [23].

4. SIMULATION RESULTS

4.1. Setup

In this section we employ the BM3D denoiser [26] in the D-prGAMP framework (4). We call the resulting algorithms BM3D-prGAMP. We compare the performance of BM3D-prGAMP with prGAMP with and without a sparsity prior. We chose to test only these algorithms for three reasons: (1) prGAMP has demonstrated performance far superior to competing methods. (2) Competing methods do not scale well to imaging sized problems. For instance, the work [23] demonstrates that prGAMP significantly outperforms the well-known Sparse-Fienup [18], GESPAR [19], and CPRL [20] algorithms and is tens to hundreds of times faster than each. (3) Code for prSAMP [32], which may offer performance competitive with prGAMP, is not yet publicly available.

Test Setup: We tested the algorithms on the Barbara test image. Our measurements were of the form $y = |\mathbf{A}x_o + \sigma\epsilon|$, where the elements of ϵ followed and i.i.d. Gaussian distribution. The pixel values of the image x_o were in the range $[0, 255]$. Memory requirements restricted our Gaussian measurement tests to 128×128 images and time constraints meant that the tests used to generate Figures 2(a) and 2(b) and Table 1 were only 64×64. We tested at sampling rates[4] of .2, .4, .6, .8, 1, 1.5, 2, 3, 4, 5, and 6. Each test was run 3 times and the median result was recorded.

Note that the phase retrieval solution is not unique, there is a phase ambiguity: if x is a solution to the phase retrieval problem then so is $xe^{j\theta}$, for any value θ. Therefore, before we could measure the quality of our reconstruction we first had to disambiguate by rotating the phase of the estimate $\hat{x}$ so that it best aligned with x. From there, we used PSNR[5] to measure the quality of the reconstructions.

Algorithm Settings: We tested two version of prGAMP. One, denoted 'prGAMP sparse', that applied a real Gauss-Bernoulli prior on the wavelet coefficients of the image, and another, denoted 'prGAMP dense', that applied a real non-negative Guassian prior on the image pixels. Daubechies 4 wavelets were used as the sparsifying basis. We provided 'prGAMP sparse' the oracle sparsity of the signal (defined as the ratio of coefficients whose magnitude squared is 10% or more of the mean magnitude squared), but otherwise default parameters were used for both prGAMP algorithms. BM3D-prGAMP's variance parameter was set to be uniform across the image and only 3 iterations were used to learn the noise variance. Otherwise default parameters were used.

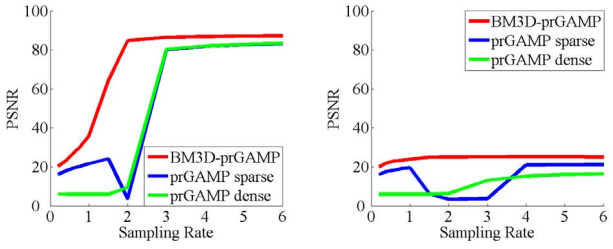
4.2. Gaussian Measurements

For our first test the elements of $\mathbf{A}$ were drawn independently from a circular, complex-valued Gaussian distribution with variance m^{-1}. While such matrices have long been of theoretical interest, recent work has shown they appear in real-world systems as well [3, 4, 43, 44]. In particular, Gaussian matrices appear whenever your measurement matrix is formed by a collection of random scatterers.

We tested at two different noise levels, first with the standard deviation of the noise $\sigma = .01$, and again at $\sigma = 20$. See Figures 2(a) and 2(b). The low noise results show that the denoising-based method dramatically outperforms the sparsity based method in the compressive regime: At sampling rates at or below 2, BM3D-prGAMP can be 10s of dB better than competing methods. The high noise results are even more striking: BM3D-prGAMP at a sampling rate above .5 outperforms prGAMP at any rate.

(a) BM3D-prGAMP (b) Wavelet Sparsity prGAMP

Fig. 1. Reconstructions of $m = .6n$ i.i.d. Gaussian sampled 128×128 Barbara image with additive white Gaussian measurement noise with standard deviation .01. Notice the reduced artifacts and finer detail in the denoising-based reconstruction.



(a) Low Noise (b) High Noise

Fig. 2. 64×64 reconstruction performance with additive white Gaussian measurement noise with standard deviation 0.01 (a) and with standard deviation 20 (b). Note that in both regimes the denoising-based method uniformly outperforms the other techniques.

Notice that when the sampling rate was 2 or 3, the wavelet spar-

[3] http://sourceforge.net/projects/gampmatlab/files/

[4] The sampling rate is the ratio $\frac{m}{n}$, where m is the dimension of the measurement vector and n is the dimension of the sampled signal.

[5] PSNR stands for peak signal-to-noise ratio and is defined as $10\log_{10}(\frac{255^2}{\text{mean}((\hat{x}-x_o)^2)})$ when the pixel range is 0 to 255. It is a measure of how closely a signal estimate $\hat{x}$ is to the true signal x_o.

sity version of prGAMP often produced very poor reconstructions of the signal. This behavior was consistent across test runs but the cause is still unknown. Fine tuning of the algorithms parameters could likely prevent this behavior, however time constraints prevented us from doing so. Note that BM3D-prGAMP exhibited no such issues.

4.3. Masked Fourier Measurements

We repeated a test performed in [23], based off of the coded diffraction pattern system proposed in [45]. In [23] the authors showed that coded masks allowed prGAMP to accurately and quickly reconstruct a Fourier sampled synthetic sparse image. We now show that, using the same setup, BM3D-prGAMP can recover natural images.

In this experiment the measurement matrix $\mathbf{A}$ is defined as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{J_1 F D_1} \\ \mathbf{J_2 F D_2} \\ \mathbf{J_3 F D_3} \\ \mathbf{J_4 F D_4} \end{bmatrix},$$

where each D_i is an $n \times n$ diagonal matrix whose entries are drawn uniformly from $\{0, 1\}$, $\mathbf{F}$ is an $n \times n$ 2D DFT matrix, and each J_i is an $\frac{n}{4} \times n$ matrix made from randomly sampled rows of an $n \times n$ identity matrix.

Figure 3 shows that through the use of a coded mask, BM3D-prGAMP can perform compressive phase retrieval using Fourier measurements and recover images of superior quality to that of wavelet sparsity-based methods.

(a) BM3D-prGAMP (b) Wavelet Sparsity prGAMP

Fig. 3. Reconstructions of $m = n$ masked Fourier sampled 128×128 Barbara image with additive white Gaussian measurement noise of standard deviation .01. The denoising-based reconstruction clearly outperforms the sparsity-based method.

4.4. Computation Times

Table 1. Average computation times, in minutes, for low noise reconstructions.

Sampling Rate (%)	**40**	**80**	**100**	**200**	**400**	**600**
prGAMP dense	15.8	27.3	28.3	44.9	7.2	7.6
prGAMP sparse	.8	1.6	2.2	23.3	8.6	7.9
BM3D-prGAMP	2.3	4.3	11.6	6.6	6.6	6.8

We include average computation times in Table 1 for the sake of completeness. These results are for the Gaussian measurement matrix reconstructions from Section 4.2. While denoising-based methods are more computationally expensive per iteration, they are not fundamentally slower. This is because using a denoiser potentially allows an algorithm to recover a signal using fewer iterations.

5. DISCUSSION

We have developed a denoising-based approximate message passing (D-prGAMP) algorithm that can recover signals from the magnitudes of their undersampled linear measurements. By employing the state-of-the-art BM3D image denoiser we have achieved state-of-the-art compressive phase retrieval of images. BM3D-prGAMP requires far fewer measurements than existing methods and is robust to noise.

6. REFERENCES

[1] R. P. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.

[2] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution fourier ptychographic microscopy," *Nature photonics*, vol. 7, no. 9, pp. 739–745, 2013.

[3] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron, "Imaging with nature: Compressive imaging using a multiply scattering medium," *Scientific reports*, vol. 4, 2014.

[4] B. Rajaei, E. W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, "Intensity-only optical compressive imaging using a multiply scattering material: a double phase retrieval system," *arXiv preprint arXiv:1510.01098*, 2015.

[5] J. R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.

[6] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, p. 237, 1972.

[7] D. W. Griffin and J. S. Lim, "Signal estimation from modified short-time fourier transform," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 2, pp. 236–243, 1984.

[8] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Hybrid projection–reflection method for phase retrieval," *JOSA A*, vol. 20, no. 6, pp. 1025–1034, 2003.

[9] A. V. Martin, F. Wang, N. Loh, T. Ekeberg, F. R. Maia, M. Hantke, G. van der Schot, C. Y. Hampton, R. G. Sierra, A. Aquila, *et al.*, "Noise-robust coherent diffractive imaging with a single diffraction pattern," *Optics Express*, vol. 20, no. 15, pp. 16650–16661, 2012.

[10] M. A. Pfeifer, G. J. Williams, I. A. Vartanyants, R. Harder, and I. K. Robinson, "Three-dimensional mapping of a deformation field inside a nanocrystal," *Nature*, vol. 442, no. 7098, pp. 63–66, 2006.

[11] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *Signal Processing, IEEE Transactions on*, vol. 63, no. 18, pp. 4814–4826, 2015.

[12] S. Marchesini, Y.-C. Tu, and H.-t. Wu, "Alternating projection, ptychographic imaging and phase synchronization," *Applied and Computational Harmonic Analysis*, 2015.

[13] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

[14] I. Waldspurger, A. dâĂŹAspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[15] L. Demanet and P. Hand, "Stable optimizationless recovery from phaseless linear measurements," *Journal of Fourier Analysis and Applications*, vol. 20, no. 1, pp. 199–221, 2014.

[16] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *Information Theory, IEEE Transactions on*, vol. 61, no. 4, pp. 1985–2007, 2015.

[17] M. L. Moravec, J. K. Romberg, and R. G. Baraniuk, "Compressive phase retrieval," in *Optical Engineering+ Applications*, pp. 670120–670120, International Society for Optics and Photonics, 2007.

[18] S. Mukherjee and C. S. Seelamantula, "Fienup algorithm with sparsity constraints: application to frequency-domain optical-coherence tomography," *Signal Processing, IEEE Transactions on*, vol. 62, no. 18, pp. 4659–4672, 2014.

[19] Y. Shechtman, A. Beck, and Y. C. Eldar, "Gespar: Efficient phase retrieval of sparse signals," *Signal Processing, IEEE Transactions on*, vol. 62, no. 4, pp. 928–938, 2014.

[20] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Cprl–an extension of compressive sensing to the phase retrieval problem," in *Advances in Neural Information Processing Systems*, pp. 1367–1375, 2012.

[21] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Uniqueness guarantees and recovery algorithms," *arXiv preprint arXiv:1311.2745*, 2013.

[22] S. Bahmani and J. Romberg, "Efficient compressive phase retrieval with constrained sensing vectors," in *Advances in Neural Information Processing Systems*, pp. 523–531, 2015.

[23] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *Signal Processing, IEEE Transactions on*, vol. 63, no. 4, pp. 1043–1055, 2015.

[24] 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.

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

[26] 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.

[27] S. Venkatakrishnan, C. Bouman, B. Wohlberg, *et al.*, "Plug-and-play priors for model based reconstruction," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, pp. 945–948, IEEE, 2013.

[28] S. Sreehari, S. Venkatakrishnan, B. Wohlberg, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *arXiv preprint arXiv:1512.07331*, 2015.

[29] A. Danielyan, A. Foi, V. Katkovnik, and K. Egiazarian, "Spatially adaptive filtering as regularization in inverse imaging: Compressive sensing, super-resolution, and upsampling," *Super-Resolution Imaging (P. Milanfar, ed.), CRC Press/Taylor*, pp. 123–153, 2010.

[30] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pająk, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, *et al.*, "Flexisp: A flexible camera image processing framework," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 6, p. 231, 2014.

[31] A. Drémeau and F. Krzakala, "Phase recovery from a bayesian point of view: the variational approach," *arXiv preprint arXiv:1410.1368*, 2014.

[32] B. Rajaei, E. W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, "Intensity-only optical compressive imaging using a multiply scattering material: a double phase retrieval system," *arXiv preprint arXiv:1510.01098*, 2015.

[33] A. Manoel, F. Krzakala, E. W. Tramel, and L. Zdeborová, "Sparse estimation with the swept approximated message-passing algorithm," *arXiv preprint arXiv:1406.4311*, 2014.

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

[35] M. Borgerding and P. Schniter, "Generalized approximate message passing for the cosparse analysis model," *CoRR*, vol. abs/1312.3968, 2013.

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

[38] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *Signal Processing Letters, IEEE*, vol. 17, no. 2, pp. 149–152, 2010.

[39] S. O. Rice, "Mathematical analysis of random noise," *Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.

[40] 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.

[41] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborová, "Adaptive damping and mean removal for the generalized approximate message passing algorithm," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, pp. 2021–2025, IEEE, 2015.

[42] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *Information Theory (ISIT), 2014 IEEE International Symposium on*, pp. 236–240, IEEE, 2014.

[43] S. Popoff, G. Lerosey, R. Carminati, M. Fink, A. Boccara, and S. Gigan, "Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media," *Physical review letters*, vol. 104, no. 10, p. 100601, 2010.

[44] A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a dmd and phase retrieval techniques," *Optics express*, vol. 23, no. 9, pp. 11898–11911, 2015.

[45] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 39, no. 2, pp. 277–299, 2015.