IMAGE PROCESSING ON LINE

Published in Image Processing On Line on 2017–01–26.
Submitted on 2016–05–25, accepted on 2016–12–28.
ISSN 2105–1232 © 2017 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials, software, datasets and online demo at
https://doi.org/10.5201/ipol.2017.178

Robust Phase Retrieval with the Swept Approximate Message Passing (prSAMP) Algorithm

Boshra Rajaei[1,5,7], Sylvain Gigan[2,4,5], Florent Krzakala[3,4,5], Laurent Daudet[1,5,6]

[1] Institut Langevin, ESPCI and CNRS UMR 7587, Paris, France
[2] Laboratoire Kastler Brossel, CNRS UMR 8552 and École Normale Supérieure, Paris, France.
[3] Laboratoire de Physique Statistique, CNRS UMR 8550 and École Normale Supérieure, Paris, France.
[4] Sorbonne Universités, UPMC Universit Paris 06, Paris, France
[5] PSL Research University, Paris, France
[6] Paris Diderot University, Sorbonne Paris Cité, Paris, France
[7] Sadjad University of Technology, Mashhad, Iran

Communicated by Ives Rey-Otero

Abstract

In phase retrieval, the goal is to recover a complex signal from the magnitude of its linear measurements. While many well-known algorithms guarantee deterministic recovery of the unknown signal using i.i.d. random measurement matrices, they suffer serious convergence issues for some ill-conditioned measurement matrices. As an example, this happens in optical imagers using binary intensity-only spatial light modulators to shape the input wavefront. The problem of ill-conditioned measurement matrices has also been a topic of interest for compressed sensing researchers during the past decade. In this paper, using recent advances in generic compressed sensing, we propose a new phase retrieval algorithm that well-behaves for a large class of measurement matrices, including Gaussian and Bernoulli binary i.i.d. random matrices, using both sparse and dense input signals. This algorithm is also robust to the strong noise levels found in some imaging applications.

Source Code

The C source code of the algorithm described in this article is accessible at the IPOL web page of the article[1].

Keywords: phase retrieval; compressed sensing; phaseless imaging; approximate message passing

[1] https://doi.org/10.5201/ipol.2017.178

1 Introduction

This paper considers the fundamental problem of recovering a complex signal, $\mathbf{x}$, from the magnitude of its linear projections. This problem is called *phase retrieval* (PR). Indeed, in many imaging setups, detectors (for instance, CCD cameras) are fundamentally intensity-only. Getting access to phase measurements may not be possible, or may involve a significantly more complex physical setup, e.g. with interferometry. Some of these applications include X-ray crystallography [12], X-ray diffraction imaging [4], optical imagers [26, 14] and astronomical imaging [9]. PR problems in the presence of additive noise may be formulated [8] as

$$\mathbf{y} = |\mathbf{H}\mathbf{x} + \mathbf{w}|^2, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}_+^M$ is the known (measured) output, $\mathbf{H}$ is the $M \times N$ known complex projection matrix, $\mathbf{x} \in \mathbb{C}^N$ is the unknown input, and $\mathbf{w} \in \mathbb{C}^M$ is the "noise" term – upon which some statistical assumptions are made. Classic phase retrieval techniques look for solutions of Equation (1), with $\mathbf{H} \in \mathbb{C}^{M\times N}$. Many methods have been reported in phase retrieval, where this measurement matrix $\mathbf{H}$ is the discrete Fourier transform, or a random matrix with i.i.d. Gaussian coefficients. These methods include, but are not limited to, convex relaxation algorithms such as phaseLift [7] and phaseCut [25], error reduction algorithms such as Gerchberg and Saxton [11] and Fienup [10] and several variants of them [17, 18] and spectral recovery methods [1].

Here we are interested in the more challenging problem of recovering $\mathbf{x} \in \mathbb{C}^N$ using more general structures of measurement matrices. Possibly these matrices can be ill-conditioned, such as for instance with random Bernoulli binary projection matrices $\mathbf{H} \in \{0, 1\}^{M\times N}$. This is the situation we face in real imaging applications using binary intensity spatial light modulators (SLM) such as digital micromirror devices (DMD) [8]. Using these ill-conditioned matrices one is often faced with convergence issues with most of the aforementioned algorithms.

Signal recovery using ill-conditioned projection matrices is also a challenging problem in other signal processing fields, e.g. compressed sensing (CS). In compressed sensing an unknown signal $\mathbf{x}$ is reconstructed by finding solutions to an underdetermined ($M \ll N$) linear system, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$. Recently, in compressed sensing, there have been attempts to reconstruct a sparse signal using generic matrices [24, 5, 16]. Some of the most efficient methods, introduced in the next section, combine a Bayesian approach with a well-known compressed sensing algorithm, the approximate message passing algorithm (AMP) [15]. In the context of CS, AMP is an iterative algorithm that originates from loopy belief propagation, although with a reduced computational complexity. AMP has been shown to be effective with a minimal number of measurements, while being efficient in terms of computational complexity. In particular, one of its extensions, called swept AMP (SwAMP), demonstrates good convergence properties over ill-conditioned noisy matrices [16].

Generalized AMP (GAMP) is an extension of AMP for arbitrary output channels, i.e. $\mathbf{y} = q(\mathbf{H}\mathbf{x} + \mathbf{w})$. Following the Bayesian method, GAMP has been extended to the phase retrieval problem, in an algorithm called phase retrieval generalized AMP (prGAMP) [22]. By utilizing a magnitude-only output channel over $\mathbf{y}$ prGAMP reaches near-optimal results to the classic PR problem, with a smaller number of measurements. Another instance of Bayesian-based PR approach is the phase retrieval variational Bayes expectation maximization (prVBEM) [8] approach, based on a mean-field variational Bayes technique. prVBEM was originally developed for the task of calibrating the transmission matrix of a strongly scattering material, using binary measurements [8]. Although prVBEM has both small complexity and robustness to strong noise, its application has only been demonstrated in the context of light focusing [20].

In this paper we mix the idea of SwAMP with phase retrieval strategies, in order to solve Equation (1) over a wide class of measurement matrices, such as random i.i.d. Bernoulli binary matrices.

This new algorithm is called prSAMP, which was briefly introduced in [20]. In the context of compressive imaging though scattering material [14], we here show that prSAMP can effectively deal with the phase-less recovery problem (1) using intensity-only SLM. This yields to these two different problems in the calibration and recovery steps: 1) complex input and binary measurement matrix, i.e. $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{H} \in \{0,1\}^{M\times N}$, and 2) binary input and complex measurement matrix, i.e. $\mathbf{x} \in \{0,1\}^N$ and $\mathbf{H} \in \mathbb{C}^{M\times N}$. Obviously, the special case of input and matrix both being complex, as addressed by most PR algorithms, is also solvable by the algorithm.

2 Notation

In this section a brief summary of the notations that is used throughout the paper is provided. As usual, scalars, vectors and matrices are written in small regular-face, small bold-face and capital bold-face letters, respectively. The ith entry of a vector $\mathbf{x}$ is denoted by $x[i]$ and the ith column of matrix $\mathbf{H}$ by $\mathbf{h}[i]$. The $\times$ and $\odot$ operators stand for vector and element-wise multiplication. We also use $(.)^{\circ 2}$ and $\oslash$ for element-wise square power and division. In algorithms, a function p is represented as $@p$, that is defined either in text or in another algorithm. Additionally, in assignment statements, when the left hand side is a vector, it implicitly denotes parallel computation of all vector entries. Otherwise, in sequential calculations we have distinct entries assignments.

3 prSAMP Algorithm

In this paper we propose a new phase retrieval method that is a mixture of two recent CS and PR ideas, in addition to some modifications in order to work for 2D image recovery. The first part is swept approximate message passing algorithm (SwAMP) [16], which is one of the many variants of approximate message passing (AMP) [15] for compressed sensing.

AMP

In the context of CS, AMP is an iterative algorithm that reconstructs a sparse signal $\mathbf{x}$ from a set of under-determined linear noisy measurements, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where $\mathbf{w} \sim \mathcal{N}(0, \sigma^2)$. Figure 1 shows the statistical approach in the AMP method [13], where the algorithm starts from initial posterior estimates of signal average and variance, $\mathbf{x}_a^0$ and $\mathbf{x}_v^0$. It then follows three main steps iteratively: 1) calculate output mean and variance variables, $\boldsymbol{\omega}$ and $\mathbf{v}$; 2) calculate input maximum likelihood terms, $\mathbf{r}$ and $\mathbf{s}$, which are also called AMP Gaussian fields; and 3) use AMP denoisers to update input signal mean and variance, $\mathbf{x}_a$ and $\mathbf{x}_v$. The AMP denoisers carry prior knowledge of the input unknown signal. Later in this paper we define two denoisers for binary and Gaussian random signals.

These steps of the main iteration loop of AMP can be formulated as follows,

$$\mathbf{v}^t = |\mathbf{H}|^2 \mathbf{x}_{\mathrm{v}}^{t-1}, \tag{2}$$
$$\boldsymbol{\omega}^t = \mathbf{H}\mathbf{x}_{\mathrm{a}}^{t-1} - (\mathbf{y} - \boldsymbol{\omega}^{t-1}) \odot \mathbf{v}^t \odot (\mathbf{v}^{t-1} + \sigma^2)^{-1}, \tag{3}$$
$$\mathbf{s}^t = [|\mathbf{H}^*|^2 (\mathbf{v}^t + \sigma^2)^{-1}]^{-1}, \tag{4}$$
$$\mathbf{r}^t = \mathbf{x}_{\mathrm{a}}^t + \mathbf{s}^t \odot \mathbf{H}^*[(\mathbf{y} - \boldsymbol{\omega}^t) \odot (\mathbf{v}^t + \sigma^2)^{-1}], \tag{5}$$
$$[\mathbf{x}_{\mathrm{a}}^t,\ \mathbf{x}_{\mathrm{v}}^t] = p_{\mathrm{in}}(\mathbf{r}^t, \mathbf{s}^t), \tag{6}$$

where $\odot$ and $(\cdot)^{-1}$ respectively denote element-wise product and inversion, $(\cdot)^t$ is a time index, $(\cdot)^*$ is the conjugate-transpose, and p_{in} denotes the AMP denoiser based on the desired signal prior, which returns both the mean and variance estimate of the unknown signal.

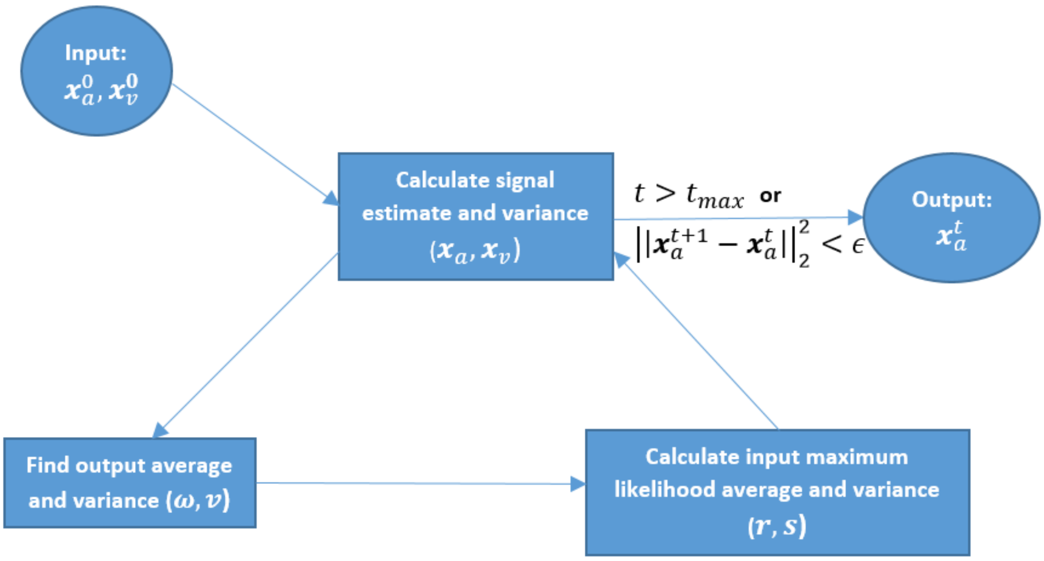

Figure 1: Approximate message passing algorithm using a Bayesian statistical approach.

AMP has been shown to converge to the optimal solution, while working with i.i.d. Gaussian measurement matrices [2]. However, it does not necessarily converge for generic possibly ill-conditioned measurement matrices [6]. This is where the idea of SwAMP brings up.

SwAMP

SwAMP is a simple change in step 2 of AMP. Instead of standard parallel calculation, a sequential, or *swept*, random update of AMP maximum likelihood variables is suggested which shows significant stabilization of the AMP loop while working with different non i.i.d. and/or ill-conditioned measurement matrices [16].

To extend the AMP framework to our PR problem of Equation (1), we first use the generalized AMP (GAMP) [21] which is an extension of AMP for arbitrary output channels, i.e. $\mathbf{y} = q(\mathbf{Hx}+\mathbf{w})$. This adds an output function, $@p_{out}$, which is dependent on the stochastic description of $q(.)$. Normal CS problems follow an additive white Gaussian noise (AWGN) channel as output prior. For the PR case, we follow what is proposed in a GAMP-based phase retrieval algorithm, which is called prGAMP [22] and formulates $@p_{out}$ as $q(|z|) = |z|$ and $q(|z|) = |z| + w$.

prSAMP

Algorithm 1 describes the phase retrieval version of SwAMP, denoted as prSAMP, which combines the swept update ordering and the phase retrieval output channel in the AMP iteration. Beside the intensity measurements, $\mathbf{y}$, and the measurement matrix, $\mathbf{H}$, the algorithm has a few other input parameters. These include the two stopping parameters, maximum number of iterations, t_{max}, and the precision threshold, ϵ. The algorithm stops if it reaches t_{max} iterations or if the difference between two successive estimations is less than ϵ, $\|\mathbf{x}_a^t - \mathbf{x}_a^{t-1}\|_2^2 < \epsilon$. The other parameter is v_0. During prSAMP iterations, variance terms may become negative or very small. This prevents the algorithm to improve its current estimation which happens often during the first iterations. In these cases the bad variance values replace by v_0. There are also two damping parameters, α and α_{2d}. Damping is necessary in case of ill-conditioned matrices. We use the first damp factor for the $\mathbf{s}$ and $\mathbf{r}$ variables in step 2 and the second for 2D signals. If the input signal is actually the vectorized version of a 2D image, after step 3 we add one step to take into account the 2D relation between $\mathbf{x}_a$ elements. Here, we employ a simple damping with respect to $\mathbf{o}^t$ which is the local median in iteration t, but more sophisticated 2D priors may be established for better smoothness in the recovered signal. Finally, we have input and output prior functions and their associated parameters, which are defined in separate algorithms. In this paper we are interested in two different scenarios: 1) binary projection matrix and complex input signal; and 2) complex projection matrix and binary input. The two associated input priors are explained in Algorithms 3 and 4.

In the main loop, the algorithm starts by estimating the output average and variance terms, $\boldsymbol{\omega}^t$ and $\mathbf{v}^t$. Then the output prior is applied over these variables to calculate the $\mathbf{g}^t$ and $\mathbf{dg}^t$ mean and variance terms. In the case of phase retrieval, these are defined distinctly in Algorithm 2, but in normal compressed sensing with additive white Gaussian noise (AWGN) the calculation is straightforward. The AWGN output prior indicates $(\mathbf{y}-\boldsymbol{\omega})/(\mathbf{v}+\Delta)$ and $-1//(\mathbf{v}+\Delta)$ for variables $\mathbf{g}$ and $\mathbf{dg}$, respectively.

In the second step, we have the sequential swept iteration for maximum likelihood terms, $\mathbf{s}$ and $\mathbf{r}$. It has been claimed in the SwAMP original paper that random computation of involved variables results in better convergence, therefore we also follow the same method. After each index i is calculated from the input signal, the updates should apply over the output channel variables. Finally, the estimate of the unknown input signal is returned as the $\mathbf{x}_a^t$ variable.

Considering a circular Gaussian additive white noise in measurements, $|\mathbf{y}|$ follows a Rician probability density function which is the basis for a PR output channel derivation in the prGAMP paper [22]. We also follow the same formulation. Algorithm 2 explains the PR output prior. Here, $@I_0(.)$ and $@I_1(.)$ functions are respectively 0^{th} and 1^{st}-order modified Bessel functions of first kind.

As we mentioned earlier, in this paper we are interested in solving the PR problem in two cases: 1) a calibration step with $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{H} \in \{0,1\}^{M\times N}$ and 2) a recovery step with $\mathbf{x} \in \{0,1\}^N$ and $\mathbf{H} \in \mathbb{C}^{M\times N}$. Therefore, a Gaussian input prior for the calibration phase is a reasonable choice, as described in Algorithm 3. Furthermore, a possible binary prior is given in Algorithm 4 for the reconstruction step.

4 Implementation

To reconstruct the complex signal $\mathbf{x}$ (up to a global phase) using its intensity-only projections, the size M of the measurement vector should be at least $2N$ – it has been established recently that, in a generic case, $M \leq 4N$ measurements are required [3] to recover a unique $\mathbf{x}$. This means that prSAMP follows a computational complexity of $O(N^3)$ which is a bottleneck for real-time imaging. Due to the sequential nature of the swept loop we can not solve this scaling issue directly but there are two possibilities to alleviate it: 1) in the calibration phase, since different rows of the measurement matrix are inherently independent, the algorithm is fully parallel. In the supplementary files, two extensions of prSAMP using OMP and MPI parallel tools, are provided; 2) the other enhancement option is an idea we call block-based phase retrieval [19]. This block-based PR method starts by splitting the $M \times N$ input problem into K, $m_i \times n_i$ sub-problems, where $\sum_{i=0}^{K-1} n_i = N$, $m_i = \lceil \alpha n_i \rceil$ and $\alpha = M/N$. The K sub-problems are then solved in parallel. Finally, all the partial results are merged with a few extra global measurements, by applying a low-dimension global phase tuning step. In this way the order of the prSAMP algorithm breaks down into $O(N^3/K^2)$. This comes at a price of being able to design the measurement matrix in a general block-diagonal manner which is the case in any physical systems where one can probe the whole object by parts. Block-based prSAMP is extended in the supplementary files using Matlab.

Algorithm 1: prSAMP

input : Measurements $\mathbf{y} \in \mathbb{R}_+^M$, projection matrix $\mathbf{H} \in \{0,1\}^{M \times N}$ or $\mathbf{H} \in \mathbb{C}^{M \times N}$, initial unknown input signal mean $\mathbf{x}_a^0 \in \mathbb{C}^N$ or $\mathbf{x}_a^0 \in \{0,1\}^N$, initial input signal variance $\mathbf{x}_v^0 \in \mathbb{R}_+^N$, maximum number of iterations t_{max}, precision threshold ϵ, negative variance substitute v_o, damping factor α, 2D signal damping factor α_{2d}, input prior function $@p_{in}$, output prior function $@p_{out}$, input prior parameters $\boldsymbol{\theta}_{in}$, output prior parameters $\boldsymbol{\theta}_{out}$

output: Final input signal mean $\mathbf{x}_a^t$, final input signal variance $\mathbf{x}_v^t$

$N \leftarrow length(\mathbf{x}_a^0)$

$\mathbf{g}^0 \leftarrow 0$, $\mathbf{s} \leftarrow 0$ $\mathbf{r} \leftarrow 0$

for $t = 1$ **to** t_{max} **do**

- $\mathbf{v}^{(t,0)} = \mathbf{H}^{\circ 2} \times \mathbf{x}_v^{t-1}$ — *Output variance estimate*
- $\boldsymbol{\omega}^{(t,0)} = \mathbf{H} \times \mathbf{x}_a^{t-1} - \mathbf{v}^{(t,0)} \odot \mathbf{g}^{(t-1,N)}$ — *Output average estimate*
- $[\mathbf{g}^{(t,0)}, \mathbf{dg}^{(t,0)}] = @p_{out}(\mathbf{y}, \mathbf{v}^{(t,0)}, \boldsymbol{\omega}^{(t,0)}, \boldsymbol{\theta}_{out}, v_o)$ — *Output channel prior*
- $\tau \leftarrow$ Random-Permutation$([1, \ldots, N])$ — *Random permutation to calculate AMP Gaussian fields*
- **for** $k = 1$ **to** N **do**
  - $i = \tau[k]$
  - $s[i] = \alpha s[i] + (1-\alpha)(-\mathbf{dg}^{(t,k-1)^T} \times \mathbf{h}[i]^{\circ 2})^{-1}$ — *Maximum likelihood of signal variance*
  - **if** $s[i] < 0$ **then**
    - $s[i] \leftarrow v_o$ — *Correct negative variance values*
  - $r[i] = \alpha r[i] + (1-\alpha)(\mathbf{x}_a^{t-1}[i] + s[i](\mathbf{g}^{(t,k-1)^T} \times \mathbf{h}[i])$ — *Maximum likelihood of signal average*
  - $[x_a^t[i], x_v^t[i]] = @p_{in}(r[i], s[i], \boldsymbol{\theta}_{in}, v_o)$ — *Calculate input prior function*
  - *Apply update of i-th input element over output channel variables.*
  - $\mathbf{v}^{(t,k)} = \mathbf{v}^{(t,k-1)} + \mathbf{h}[i]^{\circ 2} \odot (x_v^t[i] - x_v^{t-1}[i])$
  - $\boldsymbol{\omega}^{(t,k)} = \mathbf{h}[i] \odot (x_a^t[i] - x_a^{t-1}[i]) - (\mathbf{v}^{(t,k)} - \mathbf{v}^{(t,k-1)}) \odot \mathbf{g}^{(t,0)}$
  - $[\mathbf{g}^{(t,k)}, \mathbf{dg}^{(t,k)}] = @p_{out}(\mathbf{y}, \mathbf{v}^{(t,k)}, \boldsymbol{\omega}^{(t,k)}, \boldsymbol{\theta}_{out}, v_o)$
- **if** $\alpha_{2d} > 0$ **then**
  - $\mathbf{x}_a^t = \alpha_{2d}\mathbf{x}_a^t + (1-\alpha_{2d})\mathbf{o}^t$ — *Damping of 2D signal according to the local median.*
- **if** $|\mathbf{x}_a^t - \mathbf{x}_a^{t-1}|/N < \epsilon$ **then**
  - **break** — *Convergence control*

return $\mathbf{x}_a^t$, $\mathbf{x}_v^t$

5 Parameter Study

There are two groups of parameters: first, the main prSAMP parameters and second, priors parameters. Depending on prior knowledge of input and output signals, $\mathbf{x}$ and $\mathbf{y}$, we may require to provide parameters like, noise variance in measurements, Δ, an estimation of input sparsity level, ρ, or input mean and variance, m and σ, in case of Gaussian input prior. Beside these obvious prior-dependent parameters, there are a few main parameters that play an important role in the algorithm convergence. The initial estimation of the unknown input signal, $\mathbf{x}^0$, and the damping factor, α, are the two more important ones.

As it is well-known the compressive phase retrieval problem generally suffers from convergence to local minima [24]. Empirical studies show that the situation is worse while working with ill-conditioned non-Gaussian i.i.d. random matrices [16]. In case of Gaussian input signals, like what we have in the calibration phase, using a pseudo random generator to initialize $\mathbf{x}_a^0$ seems a reasonable choice. Afterwards if the algorithm diverges, a complete restart with a new random initial vector is

Algorithm 2: Phase retrieval output prior ($@p_{out}$)

input : $\mathbf{y}, \mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\theta}, v_o$
output: $\mathbf{g}, \mathbf{dg}$
$\Delta \leftarrow \theta[0]$ *An initial estimation of variance of noise in output channel*
$\epsilon \leftarrow \theta[1]$ *A parameter that controls small values of average and variance terms*
$\boldsymbol{\omega}[\boldsymbol{\omega} < \epsilon] \leftarrow \epsilon$ *Prevent appearance of very large values in* **g** *variable.*
$\boldsymbol{\phi} = 2\mathbf{y} \odot |\boldsymbol{\omega}| \oslash (\Delta + \mathbf{v})$
$\mathbf{R}_0 = @I_1(\boldsymbol{\phi}) \oslash @I_0(\boldsymbol{\phi})$
$\mathbf{R}_0[\boldsymbol{\phi} = 0] \leftarrow 1$
$\mathbf{g} = \boldsymbol{\omega} \oslash (\Delta + \mathbf{v}) \odot (\mathbf{y} \oslash |\boldsymbol{\omega}| \odot \mathbf{R}_0 - 1)$ *Estimate of average output signal*
$\tilde{\mathbf{v}} = \mathbf{y}^{\circ 2} \odot (1 - \mathbf{R}_0^{\circ 2}) \oslash (1 + \Delta \oslash \mathbf{v})^{\circ 2} + \Delta.\mathbf{v} \oslash (\Delta + \mathbf{v})$
$\tilde{\mathbf{v}}[\tilde{\mathbf{v}} < \epsilon] \leftarrow v_0$
$\mathbf{dg} = 1 \oslash \mathbf{v} \odot (\tilde{\mathbf{v}} \oslash \mathbf{v} - 1)$ *Estimate of variance of output channel*
return $\mathbf{g}, \mathbf{dg}$

Algorithm 3: Gaussian input prior ($@p_{in}$)

input : $r, s, \boldsymbol{\theta}, v_o$
output: x_a, x_v
$\rho \leftarrow \theta[0]$, $m \leftarrow \theta[1]$,$v \leftarrow \theta[2]$, $\epsilon \leftarrow \theta[3]$ *Gaussian input parameters: input sparsity level, ρ, average and variance estimations, m and v, and ϵ to control small values of variance terms.*
$\sigma = v.s/(s + v)$
$M = (v.r + s.m)/(s + v)$
$\beta = |m|^2/v - |M|^2/\sigma$
$z = (1 - \rho).e^{\beta/2} + \rho.s/(s + v)$
if $z < \epsilon$ **then**
  $z \leftarrow v_0$ *Prevent appearance of very large values in subsequent variables.*
$x_a = \rho.s.M/(z.(s + v))$ *Estimate of input unknown signal*
$x_v = 0.5.\rho.s/(s + v).(2s + |M|^2)/z - 0.5|x_a|^2$ *Estimate of input variance*
if $x_v < \epsilon$ ***or*** $x_v = \infty$ **then**
  $x_v \leftarrow v_o$
return x_a, x_v

Algorithm 4: Binary input prior ($@p_{in}$)

input : $r, s, \boldsymbol{\theta}, v_o$
output: x_a, x_v
$\rho \leftarrow \theta[0]$, $\epsilon \leftarrow \theta[1]$ *Binary input parameters: input sparsity level, ρ, and ϵ to control small values of variance terms.*
$z = (1 - \rho) \exp(-|r|^2/(2s)) + \rho \exp(-|1 - r|^2/(2s))$
if $z < \epsilon$ **then**
  $z \leftarrow v_0$ *Prevent appearance of very large values in subsequent variables.*
$x_a = z^{-1}\rho \exp(-|1 - r|^2/(2s))$ *Estimate of input unknown signal*
$x_v = x_a - x_a^2$ *Estimate of input variance*
if $x_v < \epsilon$ **then**
  $x_v \leftarrow v_o$
return x_a, x_v

necessary. Multiple restarts was first suggested in the prGAMP paper [22]. The solution that yields the lowest normalized residual, $NR = \|\mathbf{y} - |\mathbf{A}\hat{\mathbf{x}}|\|_2^2/\|\mathbf{y}\|_2^2$, is selected as the algorithm output. In case of other types of input signals, a good initial point would guarantee the algorithm convergence. For example, in the recovery phase of our optical imager, we employ a low resolution (LR) version of the input image. This LR signal may be gathered from negative outputs of the DMD array or numerically estimated based on a specific image database. For signals of length $N = 2^{14}$, a LR version of 2^6 is a good starting point. Depending on the initial point confidence, the $\mathbf{x}_v^0$ variance vector is selected from the $(0, 1]$ interval. In the calibration and recovery steps we empirically set $\mathbf{x}_v^0$ values to 0.5 and 0.1, respectively.

The other important parameter is the damping factor, α. Damping slows down the convergence rate of the algorithm, and hence prevents being stuck into a possibly wrong local minima, while still keeping information from previous iterations. Here, α is a scalar from the $[0, 1)$ interval where 0 indicates no damping situation. In case of ill-conditioned measurement matrices we need more damping. In our experiments we use 0.9 and 0.2 values for the calibration and recovery steps, respectively.

In case of 2D input signals we have another damping parameter, α_{2d}. In this paper as a 2D prior we used a simple damping step which mixes the current solution, $x[i]_a^t$, with a representative of its neighborhood, $o[i]^t$. The representative is the median over a 5×5 block centered at element i. This may improve by taking into account learned priors like the RBM prior as it is proposed in [23]. The more sophisticated priors usually come with the price of an offline learning step. Hence, since our simple damping prior provides satisfactory results, as it is shown in the next section, we left further improvements to the interested reader.

The other parameters include: maximum number of iterations, t_{max}, precision factor, ϵ, and negative variance factor, v_0. The number of iterations is usually a factor of the number of nonzero elements in the input unknown signal, ρN. In the calibration step, with a full rank input vector, we set t_{max} to $N/4$ empirically. But for small N, it is necessary to let the algorithm pass the initial oscillations. With small N, we may use $t_{max} \leftarrow N$.

The precision factor, ϵ, is another measure of convergence which ensures a minimum difference between two successive solutions. A difference less than ϵ indicates that the algorithm is iterating around a local minimum and, hence, there is no progress.

Finally, a negative variance factor, v_0, is employed in case of resulting a negative variance term. There are various variance variables in the prSAMP algorithm like: s and ω variables in the main algorithm and $\tilde{v}$, z and x_v variables in priors. These terms have to be positive and not extremely small. Therefore, in case of negative or very small variance terms, v_0 is used as a replacement value. This parameter should be sufficiently large and in the range of x_v^0 because negative variance indicates a bad situation in the prSAMP iteration and we should set the variance to a large value to let the algorithm converge to another mean point.

6 Experimental Results

In this section we investigate the application of using the prSAMP algorithm to solve the phase retrieval problem (1) in two different situations; first, $\mathbf{H} \in \{0, 1\}^{M \times N}$, and $\mathbf{x} \in \mathbb{C}^N$ and second, $\mathbf{H} \in \mathbb{C}^{M \times N}$ and $\mathbf{x} \in \{0, 1\}^N$. Using a binary transition matrix to recover the complex input signal, Figure 2 shows the phase transition plot for $N = 256$ and SNR equal to 30 dB ($\Delta = 10^{-3}$). The error is measured in terms of normalized mean square error (NMSE) between original and recovered signals after compensating the global phase shift. Each point in the plot is the lowest NMSE obtained by prSAMP in 50 distinct trials. As a result of the ill-conditioned binary measurement matrix the damping factor, α, is set to 0.9. A phase transition curve is generated by applying a NMSE threshold of 0.2. The plot confirms the recently established rate of $M \geq 4N$ to reconstruct a dense signal $\mathbf{x}$ in phase retrieval regime. The effect of increasing the number of measurements on the recovered signal is shown in Figure 3, using the same settings as in the previous experiment and a random input. Here, we have $K = N$.

For $\rho = 1$, the reconstruction performance of prSAMP is studied at four different noise levels (SNR equal to 30, 20, 10 and 5 dB) and $0.2 \leq \delta \leq 4.0$ in Figure 4. According to this experiment, in case of strong noisy measurements after $\delta = 1$ adding more samples does not improve the results significantly.

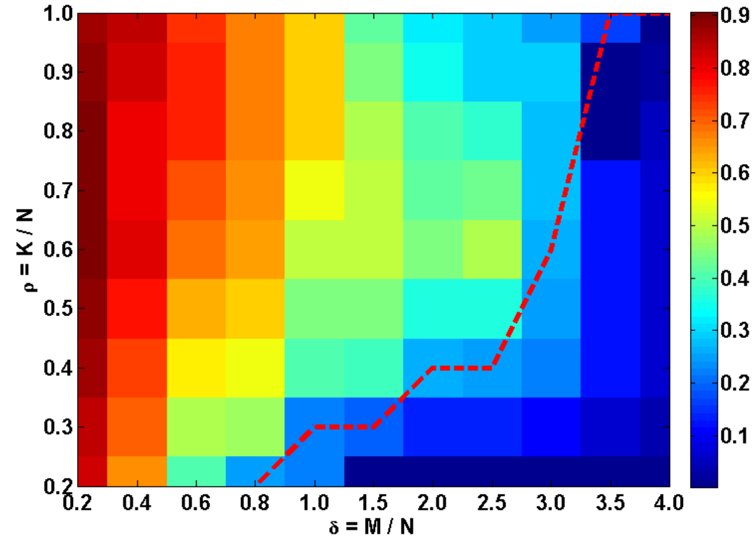

Figure 2: prSAMP phase transition plot for solving the phase retrieval problem using a binary measurement matrix. Here, $N = 256$ and SNR of 30 dB are considered in all the experiments. The performance criterion is NMSE which is selected out of 50 independent trials. The red dashed line represents a transition from failure to success by applying a threshold of 0.2.

In the second experiment, prSAMP is applied to the problem of reconstructing binary random input signals. Figure 5 shows the corresponding phase transition plot. Except α, which is set to 0.2, the other parameters are similar to the first experiment. The recovery error is measured in terms of best correlation out of 50 runs. Here, the number of necessary measurements for complete recovery is decreased significantly at different sparsity levels probably due to the binary input prior. This fact also has been shown in Figure 6 which plots a random binary signal with $K = 50$ and its two reconstructions at $M = 200$ and $M = 250$. Comparing to Figure 3, a complete recovery is obtained, using significantly less measurements ($M = N$).

7 Computational Complexity

Finally, it would be interesting to have a brief discussion on the computational complexity of the prSAMP algorithm. Even though, as our experiments show, the algorithm performs well for ill-conditioned matrices and strong noise situations, it does not scale well as the size of the input signal increases. In Algorithm 1 the number of iterations, t_{max}, and measurements, M, grow linearly with the input size N. Therefore, prSAMP has a cubic $O(N^3)$ computational complexity. In addition to this, the amount of data that the algorithm has to handle at least scales with $O(N^2)$. This is challenging with large inputs. In [19] a block-based version of prSAMP has been proposed that can reduce the computational cost and also the memory requirements of the original prSAMP by several orders of magnitude.

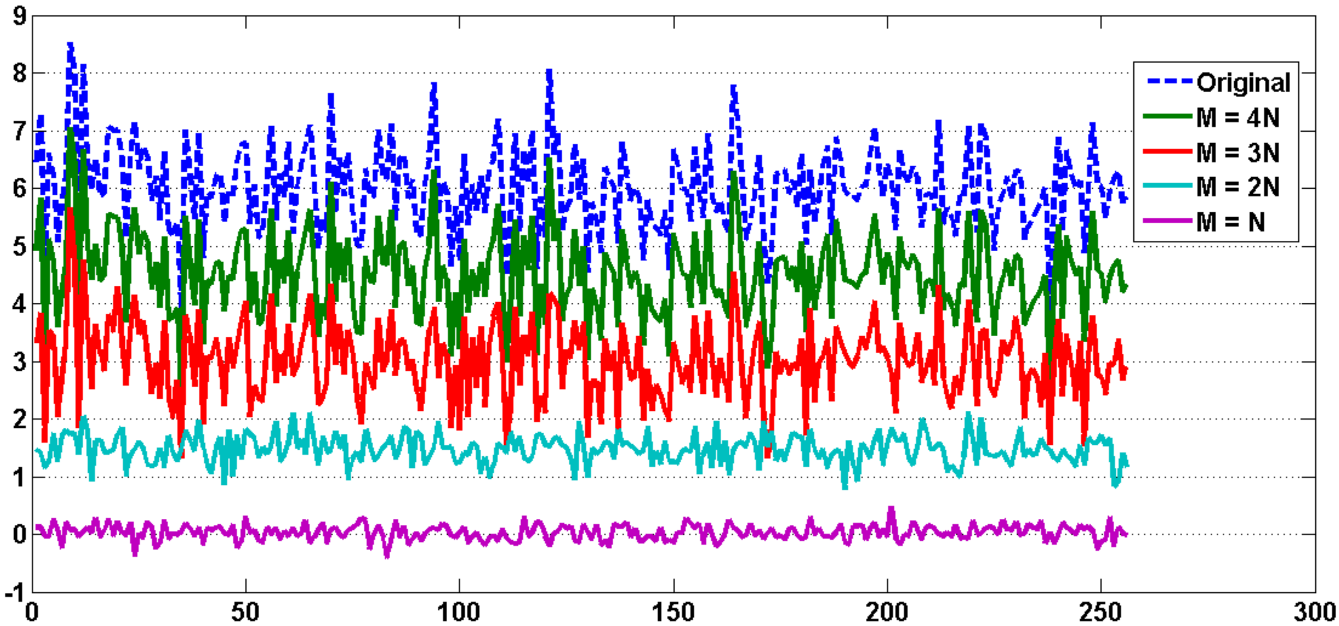

Figure 3: An instance signal $\mathbf{x} \in \mathbb{C}^N$ ($\rho = 1$) and its four prSAMP reconstructions at $\delta = \{1, 2, 3$ and $4\}$ using a binary measurement matrix (different offsets are applied for presentation purposes). The real part is plotted and the imaginary part has similar behavior. Complete recovery happens at $\delta = 4$.

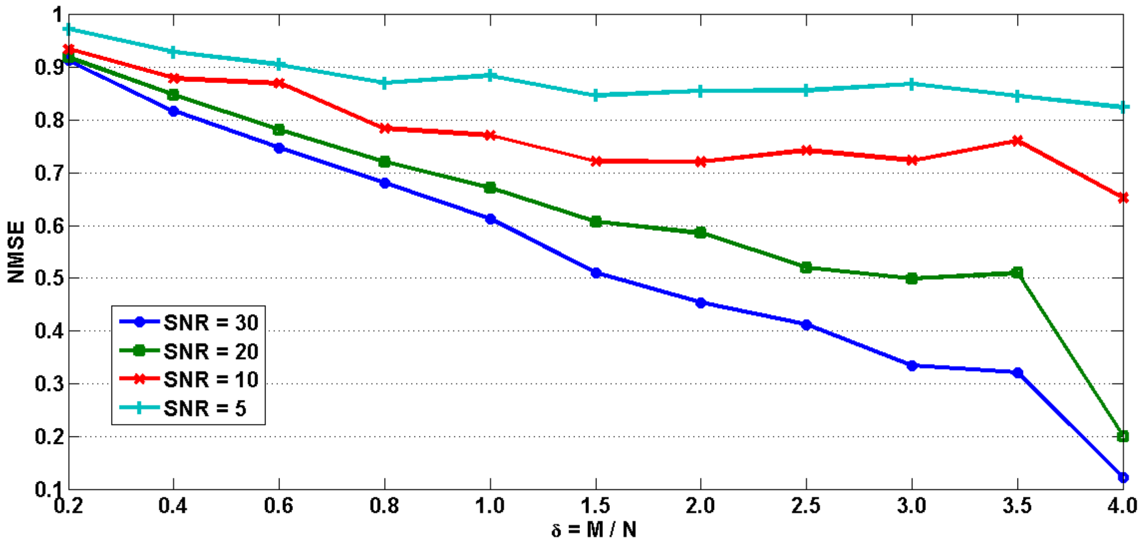

Figure 4: The effect of noise on prSAMP performance, as a function of the measurement sampling factor $\delta = M/N$, using a binary measurement matrix.

8 Conclusion

In this study, a new phase retrieval algorithm has been proposed, called phase retrieval swept AMP (prSAMP). prSWAMP is here numerically evaluated in two situations inspired by real imaging setups. In particular, prSWAMP solves the challenging problem of estimating a complex input signal using binary input patterns. In reverse, we also show that prSAMP accurately estimates a binary unknown signal using a complex transmission matrix.

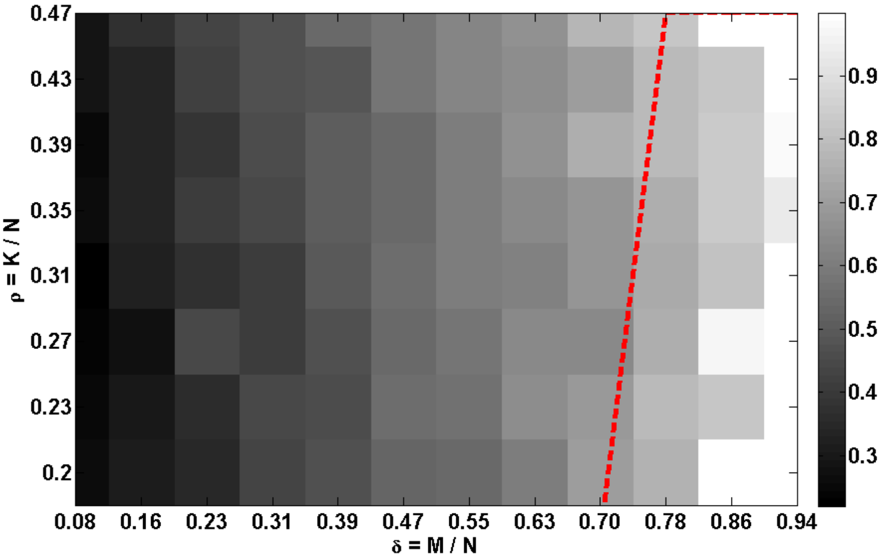

Figure 5: prSAMP phase transition plot for solving the phase retrieval problem, with a binary input, using a complex measurement matrix. Here, $N = 256$ and SNR of 30 dB are considered in all the experiments. The performance criterion is the correlation, which is selected as the best out of 50 independent trials. The red dashed line represents a transition from failure to success by applying a threshold of 0.8.

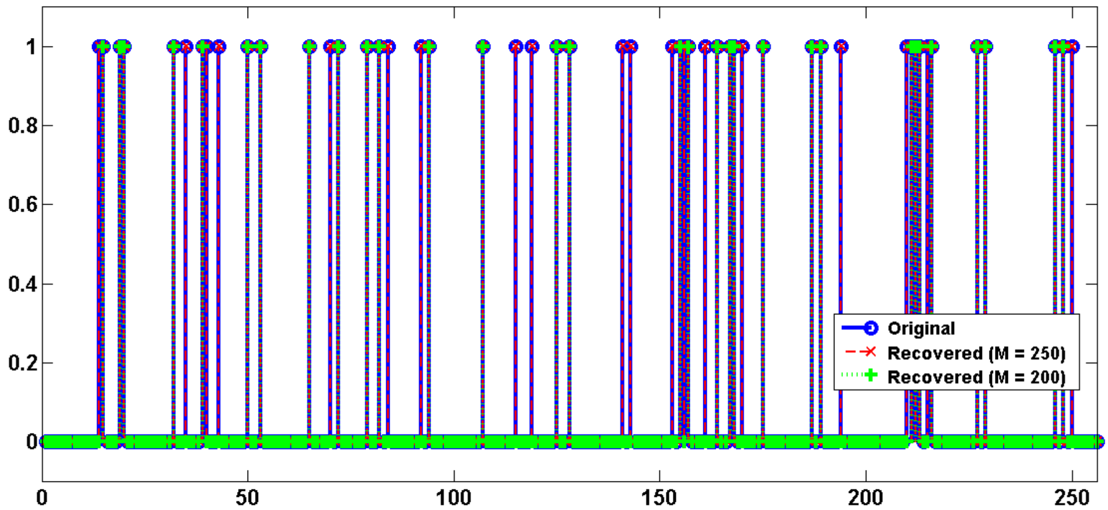

Figure 6: An instance signal $\mathbf{x} \in \{0, 1\}^N$ ($K = 50$) and its two prSAMP reconstructions at $M = \{200 \text{ and } 250\}$ after applying a threshold of 0.5.

Acknowledgment

This research has received funding from PSL Research University under contract CSI:PSL, from the European Research Council under the EU's 7th Framework Programme (FP/2007- 2013/ERC Grant Agreement 307087-SPARCS and 278025- COMEDIA); and from LABEX WIFI under references ANR-10-LABX-24 and ANR-10-IDEX-0001-02-PSL⋆.

This work was also granted access to the HPC resources of MesoPSL financed by the Region Ile de France and the project Equip@Meso (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche.

References

[1] B. Alexeev, A.S. Bandeira, M. Fickus, and D.G. Mixon, *Phase retrieval with polarization*, SIAM Journal on Imaging Sciences, 7 (2014), pp. 35–66. https://doi.org/10.1137/12089939X.

[2] M. Bayati and A. Montanari, *The dynamics of message passing on dense graphs, with applications to compressed sensing*, IEEE Transactions on Information Theory, 57 (2011), pp. 764–785. https://doi.org/10.1109/TIT.2010.2094817.

[3] B.G. Bodmann and N. Hammen, *Stable phase retrieval with low-redundancy frames*, Advances in computational mathematics, 41 (2015), pp. 317–331. https://doi.org/10.1007/s10444-014-9359-y.

[4] O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D. K. Satapathy, and J. F. Veen, *Diffractive imaging for periodic samples: Retrieving one-dimensional concentration profiles across microfluidic channels*, Acta Crystallographica Section A, 63 (2007), pp. 306–314. https://doi.org/10.1107/S0108767307021903.

[5] B. Çakmak, O. Winther, and B. Fleury, *S-AMP: Approximate message passing for general matrix ensembles*, in IEEE Information Theory Workshop, 2014, pp. 192–196. https://doi.org/10.1109/ITW.2014.6970819.

[6] F. Caltagirone, L. Zdeborová, and F. Krzakala, *On convergence of approximate message passing*, in IEEE International Symposium on Information Theory, 2014, pp. 1812–1816. https://doi.org/10.1109/ISIT.2014.6875146.

[7] E.J. Candes, T. Strohmer, and V. Voroninski, *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, 66 (2013), pp. 1241–1274. https://doi.org/10.1002/cpa.21432.

[8] 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, 23 (2015), pp. 11898–11911. https://doi.org/10.1364/OE.23.011898.

[9] C. Fienup and J. Dainty, *Phase retrieval and image reconstruction for astronomy*, Image Recovery: Theory and Application, (1987), pp. 231–275.

[10] J.R. Fienup, *Reconstruction of an object from the modulus of its Fourier transform*, Optics letters, 3 (1978), pp. 27–29. https://doi.org/10.1364/OL.3.000027.

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

[12] R.W. Harrison, *Phase problem in crystallography*, Journal of the Optical Society of America A, 10 (1993), pp. 1046–1055. https://doi.org/10.1364/JOSAA.10.001046.

[13] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, *Probabilistic reconstruction in compressed sensing: Algorithms, phase diagrams, and threshold achieving matrices*, Journal of Statistical Mechanics: Theory and Experiment, 2012 (2012), p. P08009. https://doi.org/10.1088/1742-5468/2012/08/P08009.

[14] 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, 4 (2014). https://doi.org/10.1038/srep05552.

[15] A. Maleki, *Approximate Message Passing Algorithms for Compressed Sensing*, PhD thesis, Stanford University, 2010.

[16] A. Manoel, F. Krzakala, E.W. Tramel, and L. Zdeborová, *Swept approximate message passing for sparse estimation*, in International Conference on Machine Learning, 2014.

[17] S. Marchesini, *Invited article: A unified evaluation of iterative projection algorithms for phase retrieval*, Review of scientific instruments, 78 (2007), p. 011301. https://doi.org/10.1063/1.2403783.

[18] P. Netrapalli, P. Jain, and S. Sanghavi, *Phase retrieval using alternating minimization*, arXiv preprint arXiv:1306.0160, (2013).

[19] B. Rajaei, S. Gigan, F. Krzakala, and L. Daudet, *Fast phase retrieval for high dimensions: A block-based approach*, IEEE Signal Processing Letters, 23 (2016), pp. 1179–1182. https://doi.org/10.1109/LSP.2016.2587618.

[20] B. Rajaei, E.W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, *Intensity-only optical compressive imaging using a multiply scattering material and a double phase retrieval approach*, in IEEE International Conference on Acoustics, Speech, & Signal Processing, 2016, pp. 4054–4058. https://doi.org/10.1109/ICASSP.2016.7472439.

[21] S. Rangan, *Generalized approximate message passing for estimation with random linear mixing*, in IEEE International Symposium on Information Theory, 2011, pp. 2168–2172.

[22] P. Schniter and S. Rangan, *Compressive phase retrieval via generalized approximate message passing*, IEEE Transactions on Signal Processing, 63 (2015), pp. 1043–1055. https://doi.org/10.1109/TSP.2014.2386294.

[23] E.W. Tramel, A. Drémeau, and F. Krzakala, *Approximate message passing with restricted Boltzmann machine priors*, Journal of Statistical Mechanics: Theory and Experiment, 2016 (2016), p. 073401. https://doi.org/10.1088/1742-5468/2016/07/073401.

[24] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborová, *Adaptive damping and mean removal for the generalized approximate message passing algorithm*, in IEEE International Conference on Acoustics, Speech, & Signal Processing, 2015, pp. 2021–2025. https://doi.org/10.1109/ICASSP.2015.7178325.

[25] I. Waldspurger, A. dAspremont, and S. Mallat, *Phase recovery, maxcut and complex semidefinite programming*, Mathematical Programming, 149 (2015), pp. 47–81. https://doi.org/10.1007/s10107-013-0738-9.

[26] A. Walther, *The question of phase retrieval in optics*, Optica Acta: International Journal of Optics, 10 (1963), pp. 41–49.