SOLVING COMPLEX QUADRATIC EQUATIONS WITH FULL-RANK RANDOM GAUSSIAN MATRICES

Shuai Huang*, Sidharth Gupta*, and Ivan Dokmanić

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign, Urbana, IL, USA
{huang816, gupta67, dokmanic}@illinois.edu

ABSTRACT

We tackle the problem of recovering a complex signal $\boldsymbol{x} \in \mathbb{C}^n$ from quadratic measurements of the form $y_i = \boldsymbol{x}^* \boldsymbol{A}_i \boldsymbol{x}$, where $\{\boldsymbol{A}_i\}_{i=1}^m$ is a set of complex iid standard Gaussian matrices. This non-convex problem is related to the well understood phase retrieval problem where $\boldsymbol{A}_i$ is a rank-1 positive semidefinite matrix. Here we study a general full-rank case which models a number of key applications such as molecular geometry recovery from distance distributions and compound measurements in phaseless diffractive imaging. Most prior work either addresses the rank-1 case or focuses on real measurements. The several papers that address the full-rank complex case adopt the semidefinite relaxation approach and are thus computationally demanding. In this paper we propose a method based on the standard framework comprising a spectral initialization followed by iterative gradient descent updates. We prove that when the number of measurements exceeds the signal's length by some constant factor, a globally optimal solution can be recovered from complex quadratic measurements with high probability. Numerical experiments on simulated data corroborate our theoretical analysis.

Index Terms— Complex quadratic equations, random Gaussian matrices, spectral initialization, phase retrieval, low rank matrix recovery.

1. INTRODUCTION

Systems of quadratic equations model many problems in applied science, including phase retrieval [1, 2, 3], unlabeled distance geometry problem (uDGP) [4, 5], blind channel estimation [6, 7]. Phase retrieval, in particular, has motivated considerable recent research on systems of quadratic equations. In phase retrieval, the measurement we work with is $y_i = |\boldsymbol{a}_i^* \boldsymbol{x}|^2 = \boldsymbol{x}^* \boldsymbol{a}_i \boldsymbol{a}_i^* \boldsymbol{x}$, where the measurement matrix $\boldsymbol{a}\boldsymbol{a}^*$ is a rank-1 positive semidefinite matrix.

We are interested in a different measurement model with high-rank measurement matrices. This measurement model arises in a number of applications such as uDGP [5] and the turnpike problem [8], as well as phaseless diffractive imaging where, either because of experimental or algorithmic design, we work with linear combinations of pixel values on the detector: $y_i = \sum_r w_r |\boldsymbol{a}_r^* \boldsymbol{x}|^2$. In this case, the measurement matrices are neither rank-1 nor real.

More concretely, in this paper, we address the related problem of recovering a complex signal $\boldsymbol{x} \in \mathbb{C}^n$ from its complex quadratic measurements $y_i \in \mathbb{C}$ of the following form:

$$y_i = \boldsymbol{x}^* \boldsymbol{A}_i \boldsymbol{x}, \quad i = 1, \ldots, m, \tag{1}$$

where $\boldsymbol{A}_i \in \mathbb{C}^{n \times n}$ is the i-th complex measurement matrix with entries whose real and imaginary coefficients are iid $\mathcal{N}(0, 1)$.

1.1. Prior art

Similar quadratic equation problems have been studied in other contexts. Candès et al. [9] cast the phase retrieval problem as a system of structured quadratic equations and solve it via gradient descent. As this is a non-convex problem, they use a suitably constructed spectral initializer [10] which can be shown to be close to a global optimum when sufficient measurements are available. To accelerate the phase retrieval process for large-scale problems, [11] computes an optimal step size for the gradient descent method at every iteration, and [12] proposes several coordinate descent algorithms with faster convergence speeds. The works of Wang, Xu, and Huang [13, 14, 15] study a generalized phase retrieval problem where $\boldsymbol{A}_i$ is a Hermitian matrix. They use algebraic methods [16] to find the number of measurements needed for a successful recovery.

Solving the systems of quadratic equations given by (1) is closely related to low-rank matrix recovery—it is equivalent to recovering a rank-1 positive semidefinite matrix $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^*$ with $y_i = \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle$ [17]. Previous works [18, 19, 20] on low-rank matrix recovery focus on relaxing the non-convex low rank constraint to the convex minimum nuclear norm constraint, and then establishing sufficient conditions that warrant such a convex relaxation. However, recovering the relaxed $\boldsymbol{X}$ is computationally demanding even for moderate sized problems. To address this issue, [21, 22] took the low-rank matrix decomposition approach $\hat{\boldsymbol{X}} = \boldsymbol{U}\boldsymbol{V}^T$, where $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{n \times r}$ and $\hat{\boldsymbol{X}} \in \mathbb{R}^{n \times n}$ are all real matrices, and the optimization is performed directly over $\boldsymbol{U}$ and $\boldsymbol{V}$. In this work we aim to recover a complex signal $\boldsymbol{x}$ from its complex quadratic measurements.

1.2. Our contribution and paper outline

We extend the framework introduced in [9] to the case of full-rank complex random Gaussian measurement matrices. To the best of our knowledge, using this approach to solve (1) has not been addressed in the literature. In Section 2, we extend the derivations from [10] to the complex measurement case. We show that the spectral initialization concentrates around a global optimum and compute the associated concentration bounds. In Section 3, we analyze the regularity condition and derive new results for the complex measurement case. The two results are combined to give the main result. In Section 4 we show the phase transition plot via simulated experiments and demonstrate our method through an image recovery example. The detailed derivations and proofs are presented in an extended version of this paper [23].

1.3. Problem formulation

We minimize the following objective function $f(z)$ to obtain the recovered signal z:

$$f(z) = \frac{1}{m} \sum_{i=1}^{m} |z^* A_i z - y_i|^2 . \tag{2}$$

Suppose we are given a good initialization point $z^{(0)}$ (finding such a point is discussed in Section 2), the solution is updated iteratively via gradient descent:

$$z^{(t+1)} = z^{(t)} - \eta \nabla f(z) , \tag{3}$$

where $\eta > 0$ is some suitable step size. The gradient $\nabla f(z)$ can be computed as follows:

$$\nabla f(z) = \frac{1}{m} \sum_{i=1}^{m} (z^* A_i^* z - x^* A_i^* x) A_i z \\ + (z^* A_i z - x^* A_i x) A_i^* z . \tag{4}$$

If x is a global minimum of $f(z)$, then $x e^{j\phi}$ is also a global minimum for all $\phi \in [0, 2\pi)$. Consequently, it is standard to define the squared distance between the recovered solution z and the true solution x as

$$\mathrm{dist}^2(z, x) = \min_{\phi_z \in [0, 2\pi]} \left\| z - x e^{j\phi_z} \right\|_2^2 \\ = \|z\|_2^2 + \|x\|_2^2 - 2|z^* x| , \tag{5}$$

where $z^* x = |z^* x| e^{j\phi(z^* x)}$ and $\phi_z = -\phi(z^* x)$.

2. SPECTRAL INITIALIZATION

Spectral initialization is widely used to obtain an initialization that is close to a global optimum, x. Similar to [10, 9], we show that the spectral initializer concentrates around x and can be used to solve the general complex quadratic equations in (1) under the random complex Gaussian measurement model.

2.1. Initialization for signals with known norms

We first consider the case where the norm of x is known and fixed. Without loss of generality, we assume $\|x\| = 1$. Unlike in the phase retrieval problem, which uses the leading eigenvector of the Hermitian matrix $\hat{S} = \frac{1}{m} \sum_{i=1}^{m} y_i a_i a_i^*$ as the initializer, we have two possible choices here: the leading left or right singular vector v_0 of the following matrix

$$S = \frac{1}{m} \sum_{i=1}^{m} \overline{y}_i A_i . \tag{6}$$

can both be chosen as the spectral initializer $z^{(0)} = v_0$. To see why, note that the expectation of S is

$$\mathbb{E}[S] = 2 x x^* . \tag{7}$$

We prove that for sufficiently large m, the matrix S concentrates around $\mathbb{E}[S]$ in the spectral norm: $\|S - \mathbb{E}[S]\| \leq \delta$. Therefore the leading singular vectors of S are both sufficiently correlated with a global optimizer x [24].

The concentration proof for the matrix S hinges on the rotation invariance of iid complex standard Gaussian matrices [10]. Rotation invariance means that if we define $B = RA$, where $R \in \mathbb{C}^{n \times n}$ is a complex unitary matrix, the real and imaginary coefficients of the entries B_{ij} are still iid standard Gaussian random variables.

2.2. Initialization for signals with unknown norms

When the norm of the signal is unknown, we can estimate it from the quadratic measurements. Using (7), we can compute:

$$\mathbb{E}\left[\frac{1}{2m} \sum_{i=1}^{m} \overline{y}_i y_i\right] = \mathbb{E}\left[\frac{1}{2} x^* S x\right] = \frac{1}{2} x^* \mathbb{E}[S] x = \|x\|_2^4. \tag{8}$$

When m is sufficiently large, we can prove that $\frac{1}{2m} \sum_{i=1}^{m} \overline{y}_i y_i$ is close to its expectation $\|x\|_2^4$ with high probability. Based on this result, we can scale one of the leading singular vectors v_0 of S to get our spectral initializer $z^{(0)} = \left(\frac{1}{2m} \sum_{i=1}^{m} \overline{y}_i y_i\right)^{\frac{1}{4}} \cdot v_0$.

We have the following lemma stating that the distance between the spectral initializer $z^{(0)}$ and a global optimizer x is small with high probability when m is sufficiently large.

Lemma 1. *Under the complex Gaussian measurement model, when the number of complex quadratic measurements satisfies $m > Cn$ for some universal constant C, the distance between the spectral initializer $z^{(0)}$ and a global optimizer x is upper-bounded*

$$\mathrm{dist}^2\left(z^{(0)}, x\right) \leq \frac{11}{4} \delta \|x\|_2^2 , \tag{9}$$

with probability at least $1 - 6 \exp(-Cn \cdot C_1'(\delta))$, where $\delta > 0$ is a pre-specified constant, and $C_1'(\delta)$ is a constant depending on δ.

3. CONVERGENCE ANALYSIS

For the objective function $f(z)$ introduced in (2), there is a special neighborhood $E(\epsilon)$ around every global optimizer x:

$$E(\epsilon) = \{z \mid \mathrm{dist}(z, x) \leq \epsilon \|x\|_2\} .$$

The objective function $f(z)$ is said to satisfy the regularity condition $RC(\alpha, \beta, \epsilon)$ if the following holds for all $z \in E(\epsilon)$ [9]:

$$\mathrm{Re}\left(\left\langle \nabla f(z), z - x e^{j\phi_z}\right\rangle\right) \geq \frac{1}{\alpha} \mathrm{dist}^2(z, x) + \frac{1}{\beta} \|\nabla f(z)\|^2 , \tag{10}$$

where $\alpha > 0$, $\beta > 0$, $\epsilon > 0$ are some carefully chosen constants.

The regularity condition $RC(\alpha, \beta, \epsilon)$ ensures that the gradient descent updates (3) with a step size $\eta \in (0, \frac{2}{\beta})$ converge linearly to a global optimizer x when initialized within the neighborhood $E(\epsilon)$ [9, Lemma 7.10]:

$$\mathrm{dist}^2\left(z^{(t)}, x\right) \leq \left(1 - \frac{2\eta}{\alpha}\right)^t \mathrm{dist}^2\left(z^{(0)}, x\right) . \tag{11}$$

For the system of complex quadratic equations given by (1), we show that the objective function (2) satisfies (10) and give a set of $(\alpha, \beta, \epsilon)$-values such that spectral initialization followed by gradient descent succeeds with high probability. We shall make use of the following lemma.

Lemma 2. *Under the complex Gaussian measurement model, when the number of complex quadratic measurements satisfies $m > Cn$ for some universal constant C, the following two inequalities hold*

with probability at least $1 - 20\exp\left(-Cn \cdot C_2'(\nu)\right)$:

$$\left|\frac{1}{m}\sum_{i=1}^m \boldsymbol{p}^* \boldsymbol{A}_i^* \boldsymbol{q} \boldsymbol{u}^* \boldsymbol{A}_i \boldsymbol{v} - 2\boldsymbol{u}^*\boldsymbol{q}\boldsymbol{p}^*\boldsymbol{v}\right| \le \nu\|\boldsymbol{p}\|_2\|\boldsymbol{q}\|_2\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2$$

$$(12)$$

$$\left|\frac{1}{m}\sum_{i=1}^m \mathrm{Re}\left(\boldsymbol{p}^* \boldsymbol{A}_i^* \boldsymbol{q} \boldsymbol{u}^* \boldsymbol{A}_i \boldsymbol{v}\right) - 2\mathrm{Re}\left(\boldsymbol{u}^*\boldsymbol{q}\boldsymbol{p}^*\boldsymbol{v}\right)\right|$$
$$\le \nu\|\boldsymbol{p}\|_2\|\boldsymbol{q}\|_2\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2\,,$$

$$(13)$$

where $\nu > 0$ is a pre-specified constant, and $C_2'(\nu)$ is some constant depending on ν.

Let $\boldsymbol{h} = \boldsymbol{z}e^{-i\phi_z} - \boldsymbol{x}$. It is easy to verify that:

$$\mathrm{Re}(\boldsymbol{h}^*\nabla f(\boldsymbol{x})) = \mathrm{Re}\left(\left\langle \nabla f(\boldsymbol{z}), \boldsymbol{z} - \boldsymbol{x}e^{j\phi_z}\right\rangle\right)\,. \qquad (14)$$

Our proofs for the regularity condition consist of three steps.

1. **Finding a lower bound on** $\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)$.

   First we find a lower bound on $\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)$ for all $\boldsymbol{h}$ satisfying $\|\boldsymbol{h}\|_2 \le \epsilon\|\boldsymbol{x}\|_2$. We proceed by showing that $\mathbb{E}\left[\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)\right] > 0$, and that $\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)$ is close to $\mathbb{E}\left[\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)\right]$. Thus a positive lower bound can be established with high probability when m is sufficiently large. We break this problem down by lower-bounding every term in $\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)$.

   Using (13) and the Cauchy-Schwartz inequality, we have:

   $$\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right)$$
   $$\ge 4\left[\left(\|\boldsymbol{h}\|_2^2 + \mathrm{Re}\left(\boldsymbol{h}^*\boldsymbol{x}\right)\right)^2 + \|\boldsymbol{h}\|_2^2\left(\|\boldsymbol{x}\|_2^2 + \mathrm{Re}(\boldsymbol{x}^*\boldsymbol{h})\right)\right]$$
   $$\quad - 2\|\boldsymbol{h}\|_2^2\nu\left(\|\boldsymbol{h}\|_2^2 + 3\|\boldsymbol{h}\|_2\|\boldsymbol{x}\|_2 + 2\|\boldsymbol{x}\|_2^2\right)$$
   $$\ge 4\left[\|\boldsymbol{h}\|_2^2\left(\|\boldsymbol{x}\|_2^2 + \mathrm{Re}(\boldsymbol{x}^*\boldsymbol{h})\right)\right]$$
   $$\quad - 2\|\boldsymbol{h}\|_2^2\nu\left(\|\boldsymbol{h}\|_2^2 + 3\|\boldsymbol{h}\|_2\|\boldsymbol{x}\|_2 + 2\|\boldsymbol{x}\|_2^2\right)$$
   $$\ge 4\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^2\left(1 - \epsilon - \frac{\nu}{2}\left(2 + 3\epsilon + \epsilon^2\right)\right)\,,$$

   $$(15)$$

   where $\epsilon = \frac{\|\boldsymbol{h}\|_2}{\|\boldsymbol{x}\|_2}$, and ν is some pre-specified constant.

2. **Finding an upper bound on** $\|\nabla f(\boldsymbol{z})\|_2^2$.

   We next compute an upper bound on $\|\nabla f(\boldsymbol{z})\|_2^2$. This is equivalent to finding the supremum of $|\boldsymbol{u}^*\nabla f(\boldsymbol{z})|^2$ over all $\boldsymbol{u} \in \mathbb{C}^N$ with $\|\boldsymbol{u}\|_2 = 1$. We can then use (12) to obtain an upper bound on $|\boldsymbol{u}^*\nabla f(\boldsymbol{z})|$ for all $\boldsymbol{u} \in \mathbb{C}^n$ satisfying $\|\boldsymbol{u}\|_2 = 1$:

   $$|\boldsymbol{u}^*\nabla f(\boldsymbol{z})| \le 2(2+\nu)\|\boldsymbol{h}\|_2\|\boldsymbol{x}\|_2^2\left(\epsilon^2 + 3\epsilon + 2\right)\,. \qquad (16)$$

   We then have:

   $$\|\nabla f(\boldsymbol{z})\|_2^2 = \sup_{\|\boldsymbol{u}\|_2=1}|\boldsymbol{u}^*\nabla f(\boldsymbol{z})|^2$$
   $$\le 4(2+\nu)^2\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^4\left(\epsilon^2 + 3\epsilon + 2\right)^2\,.$$

   $$(17)$$

3. **Choosing** $\{\alpha, \beta, \epsilon\}$**-values for the regularity condition.**

   We can now combine the theory we have developed so far to determine the values of $\{\alpha, \beta, \epsilon\}$. The are multiple choices for their values to make sure the initializer $\boldsymbol{z}^{(0)}$ stays in the neighborhood

of $\boldsymbol{x}$ and the objective function $f(\boldsymbol{x})$ satisfies the regularity condition (10). For example, if we choose $\delta = 0.01$, using lemma 1, with high probability we can get that:

$$\epsilon = \frac{\|\boldsymbol{h}\|_2}{\|\boldsymbol{x}\|_2} \le \sqrt{\frac{11}{400}} \le 0.2\,. \qquad (18)$$

If we choose $\epsilon \le 0.2$ and $\nu = 0.01$, it is easy to verify that

$$\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right) > 3\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^2 \qquad (19)$$
$$\|\nabla f(\boldsymbol{z})\|_2^2 < 120\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^4\,. \qquad (20)$$

We can choose $\alpha = \frac{2}{3} \cdot \frac{1}{\|\boldsymbol{x}\|_2^2}$ and $\beta = \frac{2}{3} \cdot 120\|\boldsymbol{x}\|_2^2$ to obtain:

$$\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right) > \frac{3}{2}\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^2 + \frac{3}{2}\|\boldsymbol{h}\|_2^2\|\boldsymbol{x}\|_2^2$$
$$> \frac{1}{\alpha}\|\boldsymbol{h}\|_2^2 + \frac{1}{\beta}\|\nabla f(\boldsymbol{z})\|_2^2\,.$$

$$(21)$$

When m is sufficiently large, the regularity condition holds for all $\boldsymbol{h}$ satisfying $\epsilon = \frac{\|\boldsymbol{h}\|_2}{\|\boldsymbol{x}\|_2} \le 0.2$ with high probability. We can show that the following update converges linearly to a global optimizer:

$$\boldsymbol{z}^{(t+1)} = \boldsymbol{z}^{(t)} - \eta\nabla f(\boldsymbol{z})\,, \qquad \text{(3 revisited)}$$

where $0 < \eta < \frac{2}{\beta}$.

Combining the spectral initialization analysis and convergence analysis, we can finally state our main results as follows:

Theorem 1. *Under the complex Gaussian measurement model given by (1), when the number of complex quadratic measurements $m > Cn$ for some universal constant C, and the step size $0 < \eta \le \frac{2}{\beta}$, the gradient descent update (3) initialized with the spectral initializer $\boldsymbol{z}^{(0)}$ obeying (9) converges linearly to a global optimizer $\boldsymbol{x}$:*

$$\mathrm{dist}^2\left(\boldsymbol{z}^{(t)}, \boldsymbol{x}\right) \le \left(1 - \frac{2\eta}{\alpha}\right)^t \cdot \frac{11}{4}\delta\|\boldsymbol{x}\|_2^2\,, \qquad (22)$$

with probability at least

$$1 - 26\exp\left(-Cn \cdot C_3'(\delta, \nu)\right)\,,$$

where $\delta > 0, \nu > 0$ are some pre-specified constants chosen to ensure $\mathrm{Re}\left(\langle \boldsymbol{h}, \nabla f(\boldsymbol{z})\rangle\right) > 0$, $\{\alpha, \beta\}$ are constants depending on $\{\delta, \nu\}$ and $C_3'(\delta, \nu)$ is some constant depending on $\{\delta, \nu\}$.

4. EXPERIMENTAL RESULTS

We perform numerical experiments to empirically evaluate the performance of our approach. In all experiments, the iterative gradient descent reconstruction is terminated if the distance between successive iterations is less than 10^{-6} or if 2500 iterations are completed.[1]

Theorem 1 states that the step size is upper-bounded by $\frac{2}{\beta}$ where β is one of the regularity condition parameters in (10). To show (21), β is proportional to the squared norm of the signal. Hence, in all experiments the step size is chosen to be $\frac{0.1}{\|\boldsymbol{x}\|_2^2}$ where the signal norm, $\|\boldsymbol{x}\|_2^2$, is estimated using (8).

---

[1]Code available at https://github.com/swing-research/random_quadratic_equations under the MIT License.
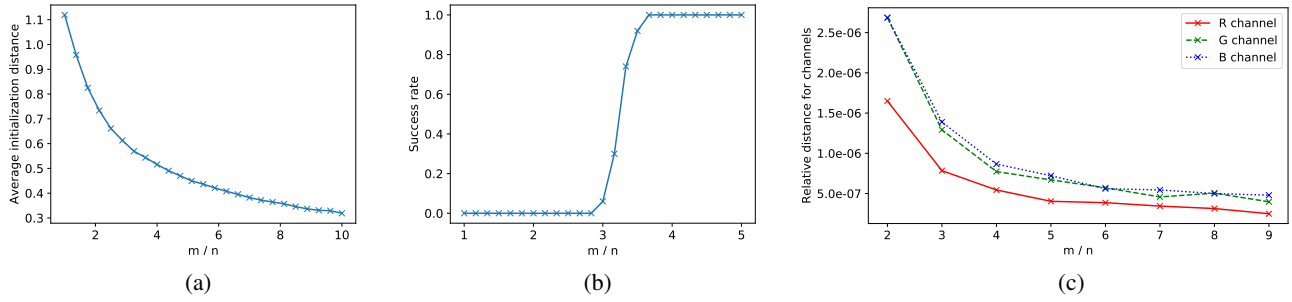
Fig. 1: (a) Closeness of spectral initialization with varying number of measurements; (b) Phase transition plot showing the empirical probability of success based on 100 trials with varying number of measurements; (c) Reconstruction performance of the image from Fig. 2 with varying number of measurements.

4.1. Closeness of spectral initializer

In this experiment we monitor how the distance between the initialization and the true solution varies with the number of measurements. We fix $n = 100$ and try different values of m with $\frac{m}{n}$ uniformly sampled between 1 and 10. We run 100 random trials for each $\frac{m}{n}$ value and calculate the average distance between the initialization and a global optimizer. In each trial we generate a random signal $\boldsymbol{x} \in \mathbb{C}^n$ with $\|\boldsymbol{x}\| = 1$, and m complex random Gaussian matrices, producing m complex quadratic measurements. In Fig. 1a we can see that the spectral initializer becomes increasingly closer to a global optimizer as $\frac{m}{n}$ increases.

4.2. Phase transition behavior

In this experiment we evaluate how the proposed approach transits from a failure phase to a success phase as we increase the number of measurements. We fix $n = 100$ and try different values of m with $\frac{m}{n}$ sampled uniformly sampled between 1 and 5. We again run 100 random trials for each $\frac{m}{n}$ value and calculate the success rate. A success is declared if the distance (5) between the recovered and true signal is less than 10^{-5}. Again, in each trial a random signal $\boldsymbol{x} \in \mathbb{C}^n$ with $\|\boldsymbol{x}\| = 1$ is generated and reconstructed. Fig. 1b shows that approximately $4n$ measurements are needed to successfully recover the signal.[2]

4.3. Reconstruction of an image

In this experiment we reconstruct an image via its complex quadratic measurements given by (1). We reconstruct the three color channels of an image separately. We choose $\frac{m}{n} = 4$. Fig. 2 shows the absolute value of the spectral initialization and the corresponding reconstruction. We define the relative error as $\frac{\|\hat{\boldsymbol{x}}|-\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$, where $|\hat{\boldsymbol{x}}|$ is the absolute value of the recovered image and $\boldsymbol{x}$ is the original image. As the channels can have non-unit norm, we define relative distance as $\frac{\text{dist}(\boldsymbol{x}, \hat{\boldsymbol{x}})}{\|\boldsymbol{x}\|}$.

The spectral initialization relative error is 0.34. The relative distances between the three channels of the original and their spectral initializations are 0.53, 0.49 and 0.51. The reconstruction relative error is 4.78×10^{-7}. The relative distances between the three channels of the original and their reconstructions are 5.45×10^{-7}, 7.73×10^{-7} and 8.66×10^{-7}. Additionally, we also reconstruct the
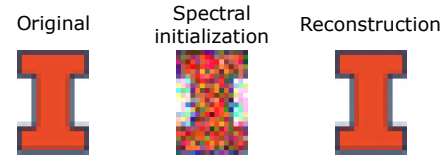

Fig. 2: Spectral initialization and reconstruction of the University of Illinois at Urbana-Champaign logo from its complex random quadratic Gaussian measurements.

same image with varying number of measurements. Fig. 1c shows the relative distances of the recovered images for each channel.

5. CONCLUSION AND FUTURE WORK

In this paper we addressed the problem of recovering a signal $\boldsymbol{x} \in \mathbb{C}^n$ in a system of random complex quadratic equations $y_i = \boldsymbol{x}^* \boldsymbol{A}_i \boldsymbol{x}$, where $\{\boldsymbol{A}_i\}_{i=1}^m$ are iid complex standard Gaussian matrices. When the number of complex measurements $m > Cn$ for some universal constant C, we can prove that with high probability: 1) the spectral initializer $\boldsymbol{z}^{(0)}$ is close to a global optimizer, 2) the gradient descent update initialized with $\boldsymbol{z}^{(0)}$ converges linearly to a global optimizer. Numerical experiments corroborate the theoretical analysis and show that a global optimum can be successfully recovered when m is sufficiently large. Additionally, our analysis complements the existing results for real measurements and rank-1 positive semidefinite measurement matrices.

Recent phase retrieval works show that a regularized spectral initialization and gradient descent update can improve the robustness and performance of the recovery algorithm [25, 26]. Chen, et al., [27] further proved that vanilla gradient descent with random initialization enjoys favorable convergence guarantees in solving the phase retrieval problem. Our future work involves analyzing similar algorithmic enhancements for the random complex Gaussian measurement model as well as extending our approach to recovering structured high-rank complex matrices that arise in key applications.

---

[2]We note that changing the minimum distance required for success or the algorithm stopping criteria can shift the curve.

6. REFERENCES

[1] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, August 1982.

[2] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, May 2015.

[3] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," *CoRR*, vol. abs/1510.07713, October 2015.

[4] P. Duxbury, L. Granlund, S. Gujarathi, P. Juhas, and S. Billinge, "The unassigned distance geometry problem," *Discrete Applied Mathematics*, vol. 204, no. Supplement C, pp. 117 – 132, May 2016.

[5] S. Huang and I. Dokmanić, "Reconstructing point sets from distance distributions," *CoRR*, vol. abs/1804.02465, April 2018.

[6] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1711–1732, March 2014.

[7] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals: uniqueness conditions," *CoRR*, vol. abs/1308.3058, August 2013.

[8] T. Dakic, *On the turnpike problem*, Ph.D. thesis, Simon Fraser University, August 2000.

[9] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, pp. 1985–2007, April 2015.

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

[11] X. Jiang, S. Rajan, and X. Liu, "Wirtinger flow method with optimal stepsize for phase retrieval," *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1627–1631, Nov 2016.

[12] W. Zeng and H. So, "Coordinate descent algorithms for phase retrieval," *CoRR*, vol. abs/1706.03474, 2017.

[13] Y. Wang and Z. Xu, "Generalized phase retrieval: measurement number, matrix recovery and beyond," *Applied and Computational Harmonic Analysis*, September 2017.

[14] Z. Xu, "The minimal measurement number for low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 44, no. 2, pp. 497 – 508, March 2018.

[15] M. Huang and Z. Xu, "Solving Systems of Quadratic Equations via Exponential-type Gradient Descent Algorithm," *CoRR*, vol. abs/1806.00904, June 2018.

[16] A. Conca, D. Edidin, M. Hering, and C. Vinzant, "An algebraic characterization of injectivity in phase retrieval," *Applied and Computational Harmonic Analysis*, vol. 38, no. 2, pp. 346 – 356, March 2015.

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

[18] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, August 2010.

[19] Y. Eldar, D. Needell, and Y. Plan, "Uniqueness conditions for low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 33, no. 2, pp. 309 – 314, September 2012.

[20] M. Kabanava, R. Kueng, H. Rauhut, and U. Terstiege, "Stable low-rank matrix recovery via null space properties," *Information and Inference: A Journal of the IMA*, vol. 5, no. 4, pp. 405–441, August 2016.

[21] Y. Chen and M. J. Wainwright, "Fast low-rank estimation by projected gradient descent: general statistical and algorithmic guarantees," *CoRR*, vol. abs/1509.03025, September 2015.

[22] S. Tu, R. Boczar, M. Simchowitz, M. Soltanolkotabi, and B. Recht, "Low-rank solutions of linear matrix equations via procrustes flow," *Proceedings of the 33rd International Conference on International Conference on Machine Learning*, vol. 48, pp. 964–973, June 2016.

[23] S. Huang, S. Gupta, and I. Dokmanić, "Solving complex quadratic systems with full-rank random matrices," *CoRR*, vol. abs/1902.05612, February 2019.

[24] Y. M. Lu and G. Li, "Phase transitions of spectral initialization for high-dimensional nonconvex estimation," *CoRR*, vol. abs/1702.06435, August 2017.

[25] Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," *Advances in Neural Information Processing Systems*, pp. 739–747, December 2015.

[26] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, Feburary 2018.

[27] Y. Chen, Y. Chi, J. Fan, and C. Ma, "Gradient descent with random initialization: fast global convergence for nonconvex phase retrieval," *CoRR*, vol. 1803.07726, March 2018.