

# Information and Resolution

(Invited Paper)

Albert Fannjiang

Department of Mathematics

UC Davis, CA 95616-8633.

fannjiang@math.ucdavis.edu.

**Abstract**—The issue of resolution with complex-field measurement is first reviewed with emphasis on superresolution, discrete versus continuous measurement, and full versus compressive measurement.

The main focus of the paper is on the other extreme: one-bit intensity measurement which collects the least amount of information from each sensor by proper thresholding. Reconstruction from one-bit intensity measurement is related to reconstruction from zero-crossings and is much harder than reconstruction from complex-field measurement. The proposed method, dubbed the null vector, is a linearization of phase retrieval and takes the form of a quadratic optimization with a row submatrix subject to an ellipsoidal constraint.

The null vector differs from the spectral vector in two ways: (i) the former uses an ellipsoidal constraint while the latter uses the spheroidal constraint; (ii) the former uses binary, threshold-crossing information while the latter uses the intensity information to weight the quadratic optimization (i.e. row submatrix versus weighted matrix). Both contribute to the superior performance of the null vector.

Further, the question of optimal weighting is explored to improve the performance of the null vector.

**Index Terms**—Resolution length, one-bit intensity measurement, null vector, spectral vector, optimal weighting

## I. INTRODUCTION

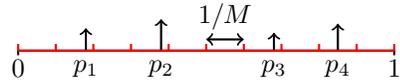
Consider the standard setup of imaging point objects located at  $p_j \in [0, 1], j = 1, \dots, s$  from signals received by sensors located at  $q_k, k = 1, \dots, N$

$$y[k] = \sum_{j=1}^s c[j] e^{-2\pi i q_k p_j} + \eta[j]$$

where  $\eta$  is the measurement noise. With wave field measured by a sufficient number of sensors uniformly deployed in  $[-\Omega, \Omega]$ , the resolution length (RL) is  $1/(2\Omega)$ , modulo some constant factor close to 1 [2].

Let us discretize the continuum system by replacing  $p_j$  by the closest grid point from the set  $\{m/M : m = 0, \dots, M - 1\}$ .

Invited paper to 2017 *Asilomar Conference on Signals, Systems, and Computers*. This work was supported by the National Science Foundation under Grant DMS-1413373.



and setting the unknown vector:  $x_0[m] = c[j]$  if  $m/M$  is the closest point to some  $p_j$  and zero otherwise. We then obtain the finite linear model:  $y = \Phi x_0 + e$  with the sensing matrix  $\Phi = [e^{-2\pi i q_k m/M}] \in \mathbb{C}^{N \times M}$  where the total error  $e$  now includes both the noise  $\eta$  and gridding (model) error. This object vector  $x_0$  can represent an extended object if the grid spacing is sufficiently small.

If  $e = 0$ , then the sampling theorem for unit-band limited, discrete  $M$ -periodic signals implies that finite Nyquist sampling contains the complete information about  $x_0$ . If  $e \neq 0$ , however, continuum sampling is not equivalent to finite Nyquist sampling. For example, if  $\eta = 0$  but  $e \neq 0$ , then the error-free signal  $y(t)$  is analytic in  $t \in \mathbb{R}$  and by analytic continuation can be recovered from sampling any infinite set with cluster points while this is generally not possible with finite sampling.

### A. Full-field measurement: Superresolution

The optimal recovery theory of Donoho [12] concerns the minimax error

$$E = \inf_x \sup_{x_0} \|x - x_0\|, \quad \text{s.t. } \|\Phi x - \Phi x_0\| < \epsilon.$$

for the linear inversion problem with a grid spacing  $= 1/F$  RL ( $F \geq 1$ ) and full continuous measurement data  $y(t), t \in [-\Omega, \Omega]$ .

Donoho proves that unique inversion is more or less equivalent to the condition of object separation  $> 2$  RL and that if the minimum separation  $> 4$  RL, then  $E \leq c\epsilon F^{2s+1}$ . Demanet & Nguyen [11] extend the results and prove  $E \sim \epsilon F^{2s-1}$  for  $s$  point objects separated at least by  $1/F$  RL. Neither [12] nor [11] provides any practical reconstruction scheme.

### B. Full-field measurement with separated objects

Candès & Fernandez-Granda [3], [4] consider the L1-min principle

$$\min \|x\|_1, \quad \|y - \Phi x\|_1 < \epsilon \quad (1)$$

with the minimum separation  $\geq 4$  RL and full continuum measurement data  $y(t), t \in [-\Omega, \Omega]$  and prove  $\|\tilde{x} - x\|_1 \leq c\epsilon F^2$ , which is an improvement over the error bound of [12].

For comparison, the single-snapshot MUSIC algorithm [20] is a gridless method and can uniquely recover the object vector from the error-free data without any assumption of minimum separation if the number of equally spaced measurement data  $N \geq 2s - 1$ . Liao and Fannjiang [14], [17] prove noise stability for the single-snapshot MUSIC if the minimum separation  $> 2$  RL by deriving a discrete Ingham inequality. The discrete Ingham inequality implies that the condition number of the MUSIC sensing matrix is less than  $3x_{\max}/x_{\min}$  where  $x_{\max}$  and  $x_{\min}$  are respectively the largest and smallest (nonzero) components of the object vector whenever the minimum separation is greater than  $2(1 - 4\pi/N)^{-1/2}$  RL. Moreover, for minimum separation  $\delta < 1$  RL, empirical evidence shows an error amplification factor  $\delta^{-\alpha}$  with an exponent  $\alpha$  slightly greater than  $2s - 1$ , indicating the superresolution of MUSIC is not far from the one given in [11].

### C. Compressive measurement with separated objects

The above works assume either full continuous or discrete samples. We next discuss gridless compressive imaging results with sparse ( $\mathcal{O}(s)$  or  $\mathcal{O}(s^2)$ ) random samples.

Tang *et al.* [21] consider the error-free case ( $\epsilon = 0$  in (1)) and prove exact recovery by L1-minimization with  $\mathcal{O}(s)$  random samples if the minimum separation  $\geq 4$  RL.

On the other hand, multi-shot MUSIC with joint sparsity is developed in [13] which assumes  $\mathcal{O}(s^2)$  random samples for performance guarantee of object support recovery for  $e \neq 0$  if the minimum separation =  $\mathcal{O}(1)$  RL independent of grid spacing.

Fannjiang and Liao [16] develop another gridless approach to compressive imaging with  $\mathcal{O}(s^2)$  random samples based on the idea of coherence band. The band-exclusion and local optimization techniques therein can improve standard compressed sensing algorithms for accurate and efficient recovery if the minimum separation  $\geq 3$  RL.

A different approach to resolution analysis is to set the grid spacing = 1 RL (i.e.  $F = 1$ ) and consider measurement of physical quantities less informative than the complex wave field. For the rest of the paper, we discuss the extreme case of signal recovery from one-bit intensity measurement and its extension.

## II. ONE-BIT INTENSITY MEASUREMENT

For intensity-only as well as one-bit measurements, it is more convenient to begin with the two-dimensional case.

Let  $x[n_1, n_2], n_1, n_2 \in \mathbb{N}$  be a finite array. Let  $X(z_1, z_2), z_1, z_2 \in \mathbb{C}$ , denote the two-dimensional  $z$ -transform of  $x$  and  $X(\omega_1, \omega_2)$  the two-dimensional Fourier transform

of  $x$  where  $\omega_1, \omega_2 \in \mathbb{R}$ , are the Fourier variables. Recall Bezout's theorem, a generalization of the fundamental theorem of algebra:

If  $X(z_1, z_2)$  and  $Y(z_1, z_2)$  are complex polynomials of degrees  $c$  and  $d$  with no common factors, then  $X$  and  $Y$  have at most  $c \cdot d$  distinct common zeros over  $\mathbb{C}^2$ .

As a corollary, if  $X(z_1, z_2)$  is irreducible over  $\mathbb{C}^2$ ,  $X(\omega_1, \omega_2)$  is real-valued and changes sign over  $\mathbb{R}^2$ , then  $\text{sgn}\{X(\omega_1, \omega_2)\}$  (i.e. the set of zero crossings) determines  $X$  uniquely, up to a constant factor [8], [9]. This is because in two (or higher) dimensions having zero crossings implies the existence of infinite zeros for  $X(\omega_1, \omega_2)$ . A similar result holds for general band-limited functions [10] and 1-D bandpass signals of bandwidth less than one octave [18]. Zero crossings are a type of one-bit measurement data.

Although Bezout's theorem implies unique determination of signal from zero crossings in the Fourier transform, actual reconstruction from zero crossings is difficult because in practice only a finite set of Fourier samples (on the order of the degree of  $X$ ) are available. Also the case of complex  $X(\omega_1, \omega_2)$  is left out.

For general signals we propose a practical reconstruction method based on threshold crossings of  $|X(\omega_1, \omega_2)|$ .

Consider the nonlinear signal model:  $b = |y|, y = A^*x_0 + e, A^* \in \mathbb{C}^{N \times M}$ , where  $|\cdot|$  denotes entrywise modulus. We propose a thresholding rule for the magnitude data  $b$  and a method for recovering  $x_0$  from the resulting binary data.

For our examples,  $A^*$  is either a random Gaussian matrix or 1-D coded Fourier matrix  $A^* = \Phi \text{diag}[\mu]$  where  $\Phi$  is the oversampled DFT and  $\mu = (\exp\{i\theta[j]\})_{j=1}^M$  with i.i.d. uniform random variables  $\theta[j]$  over  $[0, 2\pi)$ . To increase the number of measurement data in the latter case, more than one random masks may be used.

### A. Intensity threshold-crossing

We want to select a threshold to separate the ‘‘weak’’ signals from the ‘‘strong’’ signals. Let  $I \subset \{1, \dots, N\}$  be the support set of the weak signals and  $I_c$  its complement such that  $b[i] \leq b[j]$  for all  $i \in I, j \in I_c$ . In other words,  $\{b[i] : i \in I_c\}$  are the strong signals. Denote the sub-column matrices consisting of  $\{a_i\}_{i \in I}$  and  $\{a_j\}_{j \in I_c}$  by  $A_I$  and  $A_{I_c}$ , respectively. Let  $b_I = |A_I^*x_0|$  and  $b_{I_c} = |A_{I_c}^*x_0|$ .

The significance of the weak signal support  $I$  lies in the fact that  $I$  contains the best loci to ‘‘linearize’’ the problem since  $A_I^*x_0 \approx 0$ .

For a full rank  $A$ , let  $A^* = QR$  be the QR-decomposition of  $A^*$  where  $Q$  is isometric and  $R$  is an invertible upper-triangular square matrix. Let  $Q_I$  and  $Q_{I_c}$  be the sub-row matrices of  $Q$

corresponding to the index sets  $I$  and  $I_c$ , respectively. Clearly,  $A_I^* = Q_I R$  and  $A_{I_c}^* = Q_{I_c} R$ . Let  $|I|$  be the cardinality of the set  $I$ . We always assume  $|I| \geq M$  so that  $A_I^*$  and  $Q_I$  have a trivial null space and hence preserve the information of  $x_0$ .

Let  $z_0 = R x_0$ . Since  $b_I = |Q_I z_0|$  is small, the rows of  $Q_I$  are nearly orthogonal to  $z_0$ . A first approximation can be obtained from  $x_{\text{null}} = R^{-1} z_{\text{null}}$  where

$$z_{\text{null}} \in \arg \min \{ \|Q_I z\|^2 : z \in \mathbb{C}^M, \|z\| = \|b\| \}.$$

In view of the isometry property

$$\|z\|^2 = \|Q_I z\|^2 + \|Q_{I_c} z\|^2 = \|b\|^2$$

minimizing  $\|Q_I z\|^2$  is equivalent to maximizing  $\|Q_{I_c} z\|^2$  over  $\{z : \|z\| = \|b\|\}$ . This leads to the alternative variational principle

$$x_{\text{null}} \in \arg \max \{ \|A_{I_c}^* x\|^2 : x \in \mathbb{C}^M, \|R x\| = \|b\| \} \quad (2)$$

solvable by the power method (Algorithm 1).

---

### Algorithm 1: The power method for the null vector

---

- 1 **Input:**  $QR$ -decomposition of  $A^*$ ,  $I_c$ ,  $\|b\|$ .
  - 2 **Initialization:**  $z_1$
  - 3 **for**  $k = 1, 2, 3, \dots$  **do**
  - 4      $z'_k \leftarrow Q^*(\mathbf{1}_c \odot Q z_k)$ , where  $\mathbf{1}_c$  is the indicator function of  $I_c$
  - 5      $z_{k+1} \leftarrow z'_k / \|z'_k\|$
  - 6     **until**  $\|z_{k+1} - z_k\|$  is sufficiently small.
  - 7 **end**
  - 8 **Output:**  $x_{\text{null}} = \|b\| R^{-1} z_{\text{null}}$ ,  $z_{\text{null}} = z_{k+1}$ .
- 

### B. Performance guarantee

To the end of proving a performance guarantee, we consider the following simplified version of the null vector

$$\hat{x}_{\text{null}} \in \arg \min \{ \|A_I^* x\|^2 : x \in \mathbb{C}^M, \|x\| = \|x_0\| \} \quad (3)$$

which is close to  $x_{\text{null}}$  when the oversampling ratio  $L = N/M$  of the i.i.d. Gaussian matrix is large or when the measurement matrix is isometric ( $R = I$ ) as for the coded Fourier matrix. Like  $x_{\text{null}}$ ,  $\hat{x}_{\text{null}}$  can also be efficiently computed by the power method by iterating  $\lambda - A_I A_I^*$ , where the constant  $\lambda$  is chosen to be around the leading singular value of  $A_I$ . Numerical experiments show that  $\hat{x}_{\text{null}}$  as approximation of  $x_0$  is close to  $x_{\text{null}}$  for  $L \geq 8$ . But for  $L = 4$ ,  $x_{\text{null}}$  is a significantly better approximation than  $\hat{x}_{\text{null}}$ . Note that  $L = 4$  is near the threshold of having an injective intensity map:  $x_0 \rightarrow |A^* x_0|^2$  for a generic  $A^*$  [1]. We have the following performance guarantee for  $\hat{x}_{\text{null}}$ .

**Theorem 1.** [6], [7] *Let  $A$  be a  $M \times N$  i.i.d. complex Gaussian matrix and  $b = |A^* x_0|$ . Let  $\sigma := |I|/N < 1$ ,  $\nu = M/|I| < 1$ . Then for any  $x_0 \in \mathbb{C}^M$  the error bound*

$$\|x_0 x_0^* - \hat{x}_{\text{null}} \hat{x}_{\text{null}}^*\|_{\text{F}} / \|x_0\|^2 \leq c_0 \sqrt{\sigma} \quad (4)$$

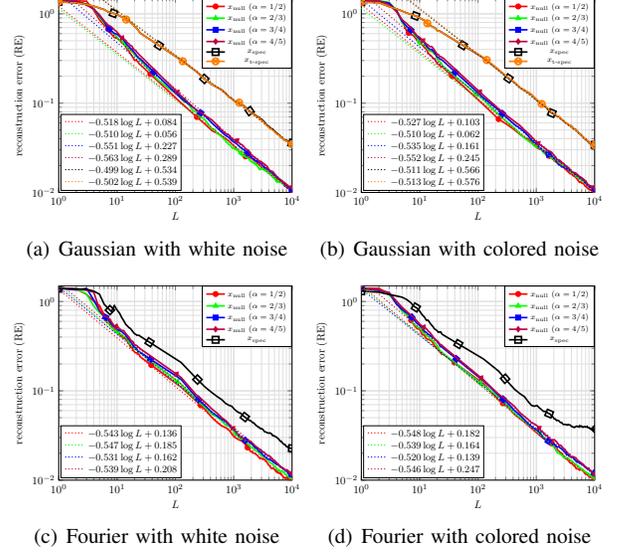


Figure 1. Log-log plot of RE for the coded Fourier case vs.  $L$  of the null vector method with various  $\alpha$ , the spectral and the truncated spectral method.

holds with probability at least  $1 - 5 \exp(-c_1 |I|^2/N) - 4 \exp(-c_2 M)$ . Here  $\|\cdot\|_{\text{F}}$  denotes the Frobenius norm.

By Theorem 1, we have that, for  $N = CM \ln M$  and  $|I| = CM, C > 1$ ,

$$\|x_0\|^{-2} \|x_0 x_0^* - \hat{x}_{\text{null}} \hat{x}_{\text{null}}^*\|_{\text{F}} \leq \frac{c}{\sqrt{\ln M}}$$

with probability exponentially (in  $M$ ) close to one, implying crude reconstruction from one-bit intensity measurement is easy.

Theorem 1 also gives a simple guideline

$$M < |I| \ll N \ll |I|^2$$

for the choice of  $|I|$  to achieve a small  $\sigma$  with high probability. In particular, the choice

$$|I| = \lceil ML^\alpha \rceil = \lceil M^{1-\alpha} N^\alpha \rceil, \quad \alpha \in [0.5, 1)$$

yields the (relative) error bound  $\mathcal{O}(L^{(\alpha-1)/2})$ , with probability exponentially (in  $n$ ) close to 1, achieving the asymptotic minimum at  $\alpha = 1/2$  (the geometric mean rule).

Given the wide range of effective thresholds, the null vector is robust as the noise tends to mess up the indices near the threshold and can be compensated by choosing a smaller  $I$ , unspoiled by noise and thus satisfying the error bound (4).

### C. Numerical experiments

For numerical experiments, we use the relative error,

$$\text{RE} := \|x_0\|^{-2} \|x_0 x_0^* - \hat{x} \hat{x}^*\|_{\text{F}},$$

as the figure of merit.

For comparison, we consider the spectral vector [19]

$$x_{\text{spec}} \in \arg \max \{ \|\text{diag}[b]A^*x\|^2 : x \in \mathbb{C}^M, \|x\| = \|x_0\| \}$$

computed by the power method (Algorithm 2).

---

**Algorithm 2: The power method for the spectral vector**

---

- 1 **Input:**  $A, b, \|x_0\|$ .
  - 2 **Initialization:**  $x_1$
  - 3 **for**  $k = 1, 2, 3, \dots$  **do**
  - 4      $x'_k \leftarrow A(|b|^2 \odot A^*x_k)$ ;
  - 5      $x_{k+1} \leftarrow x'_k / \|x'_k\|$ ;
  - 6     **Until**  $\|x_{k+1} - x_k\|$  is sufficiently small.
  - 7 **end**
  - 8 **Output:**  $\hat{x}_{\text{spec}} = x_k \|x_0\|$ .
- 

Two major differences between Algorithms 1 and 2 are (i) the QR step in Algorithm 1 and the lack of it in Algorithm 2; (ii) Algorithm 1 uses  $\mathbf{1}_c$  while Algorithm 2 uses  $|b|^2$  to weight the power method in step 4. Both factors contribute to the superior performance of Algorithm 1.

The truncated spectral vector [5] uses a still different weighting

$$x_{\text{t-spec}} \in \arg \max \{ \|A(\mathbf{1}_{\text{trim}} \odot |b|^2 \odot A^*x)\| : \|x\| = \|x_0\| \}$$

where  $\mathbf{1}_{\text{trim}}$  is the characteristic function of the set  $\{i : b(i) \leq \tau \frac{\|b\|}{\sqrt{N}}\}$  for some threshold  $\tau$ , meant to remove large spurious data.

We test the null vector and the spectral methods [5], [19] for two different signals: the white noise and a colored noise with  $M = 160$ . Fig. 1 is the log-log plots of RE versus  $L$  for (a)(b) the Gaussian matrix and (c)(d) the 1-D coded Fourier matrix. We see that the performance of the null vector is robust with respect to the choice of the measurement matrix and the signal. Moreover, the error curve for the null vector has a slope around 0.5 roughly independent of  $\alpha$  (legend), indicating a universal behavior for large  $L$ .

### III. TOWARD OPTIMALLY WEIGHTED POWER METHOD

To further improve the performance of the null vector method, let us consider a more general weight function  $w \in [0, 1]$  in Algorithm 1:

$$x_w \in \arg \max \{ \|\sqrt{w} \odot A^*x\|^2 : x \in \mathbb{C}^M, \|Rx\| = \|b\| \} \quad (5)$$

or equivalently

$$x_w \in \arg \min \{ \|\sqrt{1-w} \odot A^*x\|^2 : x \in \mathbb{C}^M, \|Rx\| = \|b\| \}$$

To demonstrate the importance of the QR step, we also test the version without QR:

$$\hat{x}_w \in \arg \max \{ \|\sqrt{w} \odot A^*x\|^2 : x \in \mathbb{C}^M, \|x\| = \|x_0\| \}. \quad (6)$$

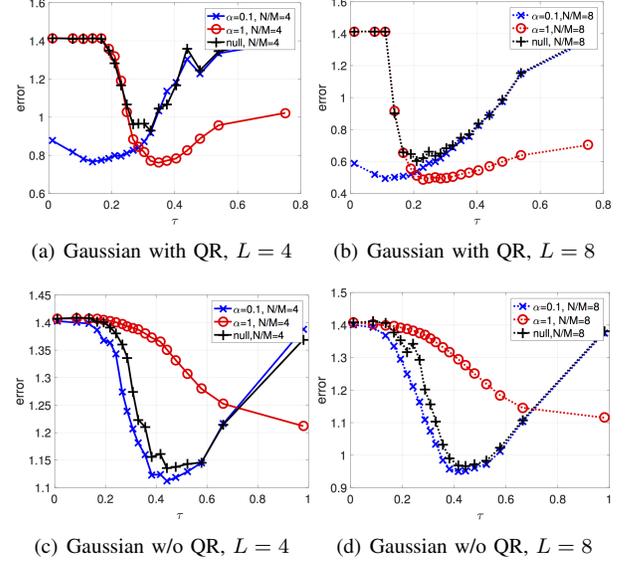


Figure 2. RE vs.  $\tau$  for the Gaussian case (a)(c) with or (b)(d) without QR at  $L = 4$  (top) and  $L = 8$  (bottom).

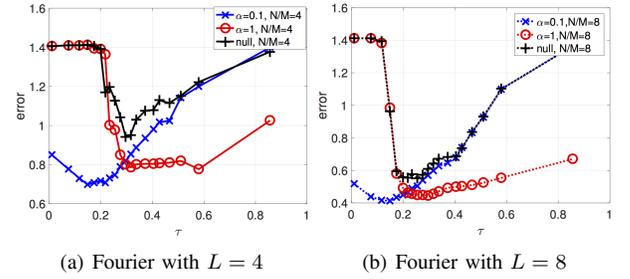


Figure 3. RE vs.  $\tau$  for the coded Fourier case at (a)  $L = 4$  or (b)  $L = 8$

Let us consider two types of continuous weight functions with a threshold  $\tau$  in terms of  $\xi = \tilde{b}[j] \equiv b[j]/\|b\|_\infty$ :

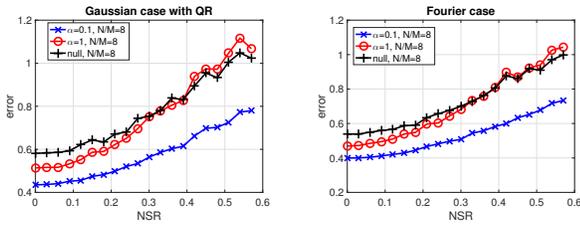
$$w_1(\xi) = \begin{cases} \left| \frac{\xi}{\tau} \right|^\alpha & \xi \leq \tau, \\ 1 & \xi > \tau, \end{cases} \quad \alpha = 1 \quad (7)$$

$$w_2(\xi) = \begin{cases} 0 & \xi < \tau, \\ \left| \frac{\xi - \tau}{1 - \tau} \right|^\alpha & \xi \geq \tau, \end{cases} \quad \alpha = 0.1. \quad (8)$$

The Heaviside function  $H_\tau = I_c$  in Algorithm 1 can be viewed as the limiting case  $\alpha \rightarrow \infty$  in (7) or  $\alpha \rightarrow 0$  in (8) while the spectral vector corresponds to  $\tau = 1, \alpha = 2$  in (7) or  $\tau = 0, \alpha = 2$  in (8).

We consider  $L = 4, 8$  for numerical demonstration. We set  $M = 200$  for the subsequent figures.

Fig. 2 (a)(b) shows RE for (5), while (c)(d) for (6), with  $w = w_1, w_2, H_\tau$  versus  $\tau$  in the Gaussian case. Clearly the QR step has a significant effect on the performance of all three cases.



(a) Gaussian with  $L = 8$

(b) Fourier with  $L = 8$

Figure 4. Average RE of 10 trials versus NSR with  $\tau = 0.3, 0.15, 0.3$  for  $w_1, w_2, H_\tau$ , respectively.

Fig. 3 shows the corresponding results for the 1-D coded Fourier case for  $L = 4, 8$ . Since the 1-D phase-coded oversampled Fourier matrix is isometric, the QR step is not needed.

Clearly from Fig. 2 and Fig. 3 the power method with either  $w_1$  or  $w_2$  performs better than with  $H_\tau$  for all  $\tau$ . Interestingly, the error curve for  $H_\tau$  is close to the larger of the two error curves for  $w_1$  and  $w_2$ . The behaviors for  $w_1$  and  $w_2$  appear to be complementary as  $\tau$  is a “lower” threshold of  $w_2$  and a “upper” threshold of  $w_1$ .

Fig. 2 (a)(b) and Fig. 3 also demonstrate the existence of an optimal threshold, not far from  $\tau = 0.15$  for  $w_2$  and  $\tau = 0.3$  for  $w_1$  and  $H_\tau$ . For noisy data, Fig. 4 shows the average RE versus the noise-to-signal (NSR) ratio. For these parameters  $w_2$  is the best performer.

#### IV. CONCLUSION

We have demonstrated the feasibility of complex signal reconstruction from one-bit intensity measurement by proper thresholding in Theorem 1 and numerical experiments. The proposed method, dubbed the null vector, can be viewed as a linearization of phase retrieval and takes the form of a quadratic optimization with a row submatrix subject to an ellipsoidal constraint, determined by the QR factorization of the measurement matrix.

The simplicity of the null vector method makes it an effective initialization scheme for other more sophisticated approaches to phase retrieval [6], [7].

From the perspective of the weighted power method, the null vector method corresponds to weighting which has a lower threshold and is thus robust w.r.t. noise. By keeping the same feature and adjusting the weighting profile as in (8) we can significantly improve the performance of the weighted power method.

#### ACKNOWLEDGMENT

I thank Pengwen Chen (Fig. 2-4) and Gi-Ren Liu (Fig. 1) for preparing the figures.

#### REFERENCES

- [1] R. Balan, P. Casazza and D. Edidin, “On signal reconstruction without phase,” *Appl. Comput. Harm. Anal.* vol. 20, pp. 345-356, 2006.
- [2] M. Born and E. Wolf, *Principles of Optics*, 7-th edition, Cambridge University Press, 1999.
- [3] E. J. Candès & C. Fernandez-Granda, “Super-resolution from noisy data”, *J. Fourier Anal. Appl.* vol. 19(6), pp. 1229-1254, 2013.
- [4] E.J. Candès & C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Commun. Pure Appl. Math.* vol 67(6), pp. 906-956, 2014.
- [5] Y. Chen and E.J. Candès, “Solving random quadratic systems of equations is nearly as easy as solving linear systems,” arXiv:1505.05114.
- [6] P. Chen, A. Fannjiang, G. Liu, “Phase retrieval with one or two diffraction patterns by alternating projection with null initialization,” arXiv:1510.07379, 2015.
- [7] P. Chen, A. Fannjiang, G. Liu, “Phase retrieval by linear algebra,” *SIAM J. Matrix Anal. Appl.* vol. 38, pp. 854 - 868, 2017.
- [8] S. Curtis, A. Oppenheim. and J. Lim, “Signal reconstruction from Fourier transform sign information,” *IEEE Trans. Acoust., Speech Signal Proc.*, vol. 33, pp. 643-657, 1985.
- [9] S.R. Curtis and A.V. Oppenheim, “Reconstruction of multidimensional signals from zero crossings,” *J. Opt. Soc. Am. A* Vol 4, pp. 221-231, 1987.
- [10] S. Curtis, S. Shitz and A.V. Oppenheim, “Reconstruction of non-periodic two-dimensional signals from zero crossings,” *IEEE Trans. Acoust., Speech, Signal Proc.* vol. ASSP-35, 890-893.
- [11] L. Demanet & N. Nguyen, “The recoverability limit for superresolution via sparsity,” arXiv:1502.01385.
- [12] D. L. Donoho, “Superresolution via sparsity constraints”, *SIAM J. Math. Anal.* vol. 23, pp.1309-1331, 1992.
- [13] A. Fannjiang, “The MUSIC algorithm for sparse objects: a compressed sensing analysis”, *Inverse Problems* vol. 27, 035013, 2011.
- [14] A. Fannjiang, “Compressive spectral estimation with single-snapshot ESPRIT: stability and resolution,” arXiv:1607.01827
- [15] A. Fannjiang, “Compressive sensing theory for optical systems described by a continuous model,” in *Optical Compressive Imaging*, in *Optical Compressive Imaging*, Taylor & Francis 2017.
- [16] A. Fannjiang & W. Liao, “Coherence pattern-guided compressive sensing with unresolved grids,” *SIAM J. Imaging Sci.* vol. 5, pp. 179-202, 2012.
- [17] W. Liao & A. Fannjiang, “MUSIC for single-snapshot spectral estimation: Stability and super-resolution,” *Appl. Comput. Harm. Anal.* vol. 40, pp 33-67, 2016.
- [18] B. Logan, “Information in the zero crossings of bandpass signals,” *Bell Syst. Tech. J.* vol 56 (4), pp. 487-510, 1977.
- [19] P. Netrapalli, P. Jain, S. Sanghavi, “Phase retrieval using alternating minimization,” *IEEE Trans. Signal Proc.* vol. 63, pp. 4814-4826, 2015.
- [20] R. O. Schmidt, “Multiple emitter location and signal parameter estimation”, *IEEE Trans. Antennas Propagat.* vol. 34(3), pp. 276-280, 1986.
- [21] G. Tang, B. Bhaskar, P. Shah & B. Recht, “Compressed sensing off the grid”, *IEEE Trans. Inform. Th.* vol. 59, pp. 7465-7490, 2013.