On Measurement Number of Phase-only Signal Reconstruction

Junren Chen∗ Michael K. Ng†

Abstract

This work establishes a theoretical framework for signal reconstruction from phase of linear measurement or affine linear measurement, and obtains some results on measurement number. We consider \( d \)-dimensional complex signals. In the scenario of linear measurement, at least \( 2d \) but no more than \( 4d - 2 \) measurements are required to recover all signals up to a positive scalar. Moreover, we prove that \( 2d - 1 \) generic measurements can recover generic signals or a specific signal, and \( 2d - 1 \) is minimal in the sense that \( 2d - 2 \) measurements can hardly recover signals. In the scenario of affine linear measurement, at least \( 2d + 1 \) measurements but no more than \( 3d \) measurements are required to recover all signals exactly, while no less than \( 2d \) generic measurements can handle generic signals. Throughout the paper, we employ our theoretical results to understand existing results and algorithms of phase-only signal reconstruction.

1 Introduction

We consider a desired signal \( x \in \mathbb{C}^d \), a linear measurement matrix \( A \in \mathbb{C}^{m \times d} \). The resulting linear measurement \( Ax \) has a measurement number \( m \). We have the task of recovering \( x \) from partial information of \( Ax \). For example, in phase retrieval, only the magnitude of \( Ax \) is available. The problem of minimal number of measurements required to recover all signals in \( \mathbb{C}^d \) naturally arises, see for example [2,3,6,8,22]. In this paper, we are interested in recovering \( x \) from the phase of \( Ax \), namely phase-only signal reconstruction. We refer it to be the magnitude retrieval problem.

The magnitude retrieval problem arises in many real-world applications. For example, in blind deconvolution problem when the Fourier phase of the undesired signal vanishes, we can observe clean Fourier phase of the desired signal that can recover the underlying signal. Such case occurs in long-term exposure to atmospheric turbulence or when images are blurred by severely defocused lenses with circular aperture [1]. Besides, we may use phase-only reconstruction to reduce the bit-rate in some coding systems. In the literature, Oppehneim and Lim [19] demonstrated the importance of spectral phase that can preserve many important features of a signal. Hayes et al. [11] developed a new set of conditions under which a finite length real signal can be specified by its Fourier phase information. They also proposed two reconstruction algorithms, namely closed form solution method and iterative method. The

∗Department of Mathematics, The University of Hong Kong. E-mail: chenjr58@connect.hku.hk
†Department of Mathematics, The University of Hong Kong. E-mail: mng@maths.hku.hk
iterative method was accelerated by a modification in [18], while the recoverable conditions were extended to high dimensional real signals in [10]. Levi and Stark [15] established the POCS algorithm. Behar et al. [5] and Urieli [21] studied the algorithm for image reconstruction applications. Behar [5] also investigated the importance of localized phase, and showed that the algorithm surpasses global phase in representing images. By applying POCS algorithm, images of geometric form are most efficiently reconstructed while symmetric images have the poorest convergence characteristics, see [21]. Hua et al. [12], and Hua and Orchard [13] developed iterative image inpainting methods on geometrical modeling from the phase of Complex Wavelet Transform (CWT). Loveimi and Ahadi [16] compared phase-only and magnitude-only reconstructed speech and investigated the effects of the window shape and length. A theoretical framework for sparse signal phase-only reconstruction is built in [7]. In [4,17,20], it has been demonstrated that phase information is practicable in many applications such as image processing, speech recognition and three-dimensional object reconstruction. Recently, based on lifting $n$ variables to $n^2$ variables, Wu et al. [23] developed MagnitudeCut method for phase-only reconstruction, which surpasses iterative methods in [10,11,18] and the greedy algorithm in [9]. Later, Kishore et al. [14] presented another approach to formulate phase-only reconstruction problem as a quadratic programming problem. Compared to PhaseCut, their approach was shown to have comparable behaviour, and offer significant computational savings since no lifting is required.

Most of the existing works for the magnitude retrieval problem concentrate on the algorithms and applications. However, theoretical results are not well-studied. It is interesting to note that the results in [10,11] contain some conditions to characterize recoverable signals, but they are restricted in the sense that $x \in \mathbb{R}^d$ and $A = F$ (the discrete Fourier matrix only).

In this paper, we establish theoretical conditions for recovery of complex signals under general measurement matrices.

For $z \in \mathbb{C}$, we use $\text{sign}(z)$ to denote the phase of $z$, where $\text{sign} : \mathbb{C} \rightarrow \mathbb{C}$ is defined as:

$$
0 \neq z \mapsto \frac{z}{|z|} \quad 0 \mapsto 0.
$$

The real part, the imaginary part and the magnitude (absolute value) of $z$ are given by $\text{Re}(z)$, $\text{Im}(z)$ and $|z|$ respectively. We allow them to operate on a vector or a matrix in an entry-wise manner. More precisely, given $x \in \mathbb{C}^d$ and $A \in \mathbb{C}^{m \times d}$, we study the reconstruction of $x$ from $\text{sign}(Ax)$, a magnitude retrieval problem. Note that for any $\lambda \in \mathbb{R}_+ = \mathbb{R} \cap (0, +\infty)$, we have $\text{sign}(A(\lambda x)) = \text{sign}(\lambda(Ax)) = \text{sign}(Ax)$, which indicates that the phase-only measurements cannot distinguish the set of signals $\mathbb{R}_+x = \{\lambda x : \lambda > 0\}$. Generally speaking, we can never expect to recover $x$ exactly, so we regard the reconstruction up to a positive number as a successful case and the set of recoverable signals is written as

$$
\mathbb{W}_A = \{x \in \mathbb{C}^d : \text{sign}(Ay) = \text{sign}(Ax) \text{ implies } y \in \mathbb{R}_+x\}.
$$

The outline of this paper is given as follows. In Section 2 we study crucial discriminant matrices whose rank characterizes recoverable signals. In Section 3 we investigate the measurement number required for recovering all signals in $\mathbb{C}^d$, namely signal recovery. In Section 4, we prove that $2d-1$ generic measurements can recover generic signals, and $2d-1$ is minimal to some extent. In Section 5, we analyse the recovery of a specific signal. We extend all results to affine linear measurement phase in section 6. Finally, some concluding remarks and future research works are given in Section 7.
1.1 Preliminaries

Let us introduce the definition of generic set in this paper. We need to consider Zariski topology on \( \mathbb{R}^n \), which is defined by declaring affine varieties to be closed subsets, and recall that affine variety is defined by the vanishing of finitely many polynomials in \( \mathbb{R}[x_1, x_2, ..., x_n] \). Suppose \( V = \mathbb{C}^s \) or \( s \)-dimensional linear subspace/submanifold of \( \mathbb{C}^n \), \( V - v \) can be viewed as a \( 2s \)-dimensional real linear space for any \( v \in V \). We further declare \( G \subset V \) to be Zariski-open if there exists a specific point \( v_0 \in V \) and a linear isomorphism \( A \) from \( V - v_0 \) to \( \mathbb{R}^{2s} \) such that \( A(G - v_0) \) is Zariski-open. Since a non-empty Zariski open set is very large, we say a subset \( U \) is generic in \( V \) if \( U \) contains a non-empty Zariski open set of \( V \). For example, assume \( f_1, f_2, ..., f_t \) are nonzero polynomials of \( Re(x) \) and \( Im(x) \) with complex coefficients, then \{ \( x \in \mathbb{C}^d : f_j(x) \neq 0 \) for any \( 1 \leq j \leq t \} \) is a non-empty Zariski open set.

There is a useful mapping \( \varphi \) to transfrom a complex matrix to a real one:

\[
\varphi : \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{R}^{2n_1 \times 2n_2}
\]

\[
A \mapsto \begin{bmatrix} Re(A) & Im(A) \\ -Im(A) & Re(A) \end{bmatrix}.
\]

Also, we define

\[
\varphi_1(A) = \begin{bmatrix} Re(A) \\ -Im(A) \end{bmatrix}.
\]

It is evident that \( A = 0 \) iff \( \varphi(A) = 0 \) iff \( \varphi_1(A) = 0 \). And \( \varphi \) and \( \varphi_1 \) are useful mappings since they preserve addition and multiplication, more precisely, we have \( \varphi_1(A + B) = \varphi_1(A) + \varphi_1(B) \), \( \varphi_1(AB) = \varphi(A)\varphi_1(B) \), \( \varphi(AB) = \varphi(A)\varphi(B) \). In addition, we have \( rank(\varphi(A)) = 2 \times rank(A) \).

2 Discriminant Matrices

In this section, we would like to construct several matrices whose rank characterizes \( x \in \mathbb{W}_A \). We call them discriminant matrices. They are our main tools for analysis.

Let us begin with some simple facts. Suppose \( rank(A) < d \), then \( dim(Ker(A)) \geq 1 \), which leads to \( \mathbb{W}_A = \emptyset \) trivially. This is undesired hence we only consider \( A \) of full column rank. Note that exchanging rows would only change the sequence of our measurements, hence makes no difference to \( \mathbb{W}_A \). Moreover, for any invertible \( P \in \mathbb{C}^d \), we have \( x \in \mathbb{W}_A \) if and only if \( P^{-1}x \in \mathbb{W}_{AP} \). The ”if” part, assume \( sign(Ax) = sign(Ay) \), i.e., \( sign(APP^{-1}x) = sign(APP^{-1}y) \), by using \( P^{-1}x \in \mathbb{W}_{AP} \), we have \( P^{-1}y \in \mathbb{R}^+_P \), hence \( y \in \mathbb{R}^+_x \), then \( x \in \mathbb{W}_A \) follows. The ”only if” part comes similarly. Therefore, we have \( \mathbb{W}_A = P\mathbb{W}_{AP} \), which implies \( \mathbb{W}_A \) and \( \mathbb{W}_{AP} \) are of the same size. For \( A \) with full column rank, by exchanging rows, we can ensure \( A = \begin{bmatrix} A' \\ A'' \end{bmatrix} \) and \( A' \in \mathbb{C}^{d \times d} \) is invertible. Then by multiplying \((A')^{-1}\) from the right side, \( A \) becomes \( A = \begin{bmatrix} I_d \\ A''(A')^{-1} \end{bmatrix} \). To sum up, we know that for any \( A \), \( rank(A) = d \), there always exists a matrix with the special form \( \begin{bmatrix} I_d \\ A_1 \end{bmatrix} \) that can recover as many signals as \( A \).

In the following discussion, when \( v \in \mathbb{C}^{n_1} \), we denote the diagonal matrix in \( \mathbb{C}^{n_1 \times n_1} \) with the main diagonal \( v \) by \( diag(v) \), the set of nonzero positions of \( v \) by \( N(v) \), for example,
$v = [1, 0, 2]^T$, then $N(v) = \{1, 3\}$. For a finite set $S$, let $|S|$ be the size of $S$. When $A \in \mathbb{C}^{n_1 \times n_2}$, and $S \subset [n_1]$, $T \subset [n_2]$ with $[k] = \{1, 2, \ldots, k\}$, we use $A^S_T$ to denote the submatrix of $A$ constituted by rows in $S$ and columns in $T$. For ease of notations, we write $A^S_{[n_2]} = A^S$, $A^T_{[n_1]} = A^T$.

2.1 General Matrix $A$

This part is devoted to give characterizations of $x \in \mathbb{W}_A$, where $A \in \mathbb{C}^{m \times d}$ is of full column rank, $x \in \mathbb{C}^d \setminus \{0\}$. We define

$$\mathbb{D}_A(x) = [\varphi(A) \varphi_1(diag(Ax))]$$

where $\varphi(A) = \begin{bmatrix} Re(A) & Im(A) & Re(diag(Ax)) \\ -Im(A) & Re(A) & -Im(diag(Ax)) \end{bmatrix}$

Consider the equation

$$Ay = diag(sign(Ax))_{N(Ax)} \Lambda,$$

where $y \in \mathbb{C}^d$, $\Lambda \in \mathbb{R}^{[N(Ax)]}$, we first give the following lemma.

**Lemma 1.** Suppose $x \neq 0$, $\text{rank}(A) = d$. Then $x \in \mathbb{W}_A$ if and only if \{ $\lambda \in \mathbb{R}^{[N(Ax)]}$ : (1) has solutions \} is a 1-dimensional linear subspace.

**Proof:** For simplicity, we write $U_x = diag(sign(Ax))$, and denote \{ $\lambda \in \mathbb{R}^{[N(Ax)]}$ : (1) has solutions \} by $V_x$. Evidently, $V_x$ is a linear subspace. Note that $(U_x)_{N(Ax)}Ax^{|N(Ax)|} = U_xAx = Ax$, hence $|Ax|_{N(Ax)} \in V_x$, then $\text{dim}(V_x) \geq 1$.

For the "if" part, suppose $\text{sign}(Ax) = \text{sign}(Ax)$, then $Ax = diag(\text{sign}(Ax))|Ax| = (U_x)_{N(Ax)}Ax$. It follows that $|Ax|_{N(Ax)} \in V_x$, by using $\text{dim}(V_x) = 1$, we know $|Ax|_{N(Ax)} \in \mathbb{R}_+|Ax|_{N(Ax)}$, then $Ax \in \mathbb{R}_+|Ax|$, i.e., $Ax \in \mathbb{R}_+|Ax|$, since $\text{rank}(A) = d$, we have $Ax \in \mathbb{R}_+|Ax|$, hence $x \in \mathbb{W}_A$.

For the "only if" part, we only need to rule out the possibility that $\text{dim}(V_x) > 1$. Assume $\text{dim}(V_x) > 1$, since $|Ax|_{N(Ax)} \in \mathbb{R}_+|Ax|_{N(Ax)}$, we can select $\Lambda_1 \in V_x \cap \mathbb{R}_+|Ax|_{N(Ax)}$, s.t. $\Lambda_1 \notin \mathbb{R}_+|Ax|_{N(Ax)}$. Then there exists $y_1$, s.t. $Ay_1 = (U_x)_{N(Ax)}\Lambda_1$, which implies $\text{sign}(Ay_1) = \text{sign}(Ax)$, by using the assumption $x \in \mathbb{W}_A$, we know $y_1 \in \mathbb{R}_+Ax$, then $Ay_1 \in \mathbb{R}_+Ax$, i.e., $(U_x)_{N(Ax)}\Lambda_1 \in \mathbb{R}_+(U_x)_{N(Ax)}|Ax|_{N(Ax)}$. Since $(U_x)_{N(Ax)}$ is of full column rank, we have $\Lambda_1 \in \mathbb{R}_+|Ax|_{N(Ax)}$, which contradicts our choice of $\Lambda_1$. \hfill $\square$

**Theorem 1.** Suppose $x \neq 0$, $\text{rank}(A) = d$, then $x \in \mathbb{W}_A$ if and only if $\text{rank}(\mathbb{D}_A(x)) = 2d + |N(Ax)| - 1$.

**Proof:** We continue to use the notation $U_x$ and $V_x$, then (1) becomes $Ay = (U_x)_{N(Ax)}\Lambda$. To compute $\text{dim}(V_x)$, we need to convert (1) to a real linear system.

By using $\varphi$ and $\varphi_1$, we have $Ay = (U_x)_{N(Ax)}\Lambda$ iff $\varphi_1(Ay) = \varphi_1((U_x)_{N(Ax)}\Lambda)$ iff $\varphi(A)\varphi_1(y) - \varphi((U_x)_{N(Ax)}\Lambda)\varphi_1(\Lambda) = 0$. Since $\Lambda \in \mathbb{R}^{[N(Ax)]}$, we have $\varphi_1(\Lambda) = \begin{bmatrix} \Lambda \\ 0 \end{bmatrix}$, hence the equation is equivalent to $\varphi(A)\varphi_1(y) - \varphi_1((U_x)_{N(Ax)}\Lambda) = 0$, i.e.,

$$\begin{bmatrix} \varphi(A) & -\varphi_1((U_x)_{N(Ax)}) \end{bmatrix} \begin{bmatrix} \varphi_1(y) \\ \Lambda \end{bmatrix} = 0.$$ 

(2)
Note that $\text{rank}(\varphi(A)) = 2\text{rank}(A) = 2d$, hence $\varphi_1(y)$ is uniquely determined by $\Lambda$. Therefore, the solution space of (2), i.e., $\text{Ker}([\varphi(A) - \varphi_1((U_x)_{N(Ax)}))]$, has the same dimension as $V_x$, which delivers that

$$\dim(V_x) = \dim(\text{Ker}([\varphi(A) - \varphi_1((U_x)_{N(Ax)})])) = 2d + |N(Ax)| - \text{rank}([\varphi(A) - \varphi_1((U_x)_{N(Ax)})]).$$

(3)

From simple observation, we have $\text{rank}(D_A(x)) = \text{rank}([\varphi(A) - \varphi_1((U_x)_{N(Ax)}))]$, combining with Lemma 1, the result follows.

By applying the Moore-Penrose inverse $A^\dagger = (A^* A)^{-1} A^*$, we propose two other forms of discriminant matrices. We have the following lemma that helps determine the dimension of the solution space with regard to a linear system with complex coefficients and real variables.

**Lemma 2.** Suppose $B \in \mathbb{C}^{m_1 \times n_2}$, $y \in \mathbb{R}^{n_2}$, denote the solution space of $By = 0$, $y \in \mathbb{R}^{n_2}$ by $(\text{Ker}(B))_\mathbb{R}$, then

$$\dim(\text{dim)((\text{Ker}(B))_\mathbb{R}) = n_2 - \text{rank}\left(\begin{bmatrix} \text{Re}(B) \\ \text{Im}(B) \end{bmatrix}\right) = n_2 - \text{rank}(\text{Re}(B^* B)).$$

**Proof:** The proof is based on two ways to convert $By = 0$ to real linear system.

The first way is to separate the real part and complex part, then $By = 0$ is equivalent to

$$\begin{bmatrix} \text{Re}(B) \\ \text{Im}(B) \end{bmatrix} y = 0,$$

hence $\dim((\text{Ker}(B))_\mathbb{R}) = n_2 - \text{rank}\left(\begin{bmatrix} \text{Re}(B) \\ \text{Im}(B) \end{bmatrix}\right)$ follows.

Then comes the other way. Note that $By = 0$ if and only if $y^* B^* By = 0$, for any $y \in \mathbb{R}^{n_2}$, $y^* B^* By \geq 0$, we have $y^* B^* By = \text{Re}(y^* B^* By) = y^T (\text{Re}(B^* B)) y$. Note that $y^T (\text{Re}(B^* B)) y \geq 0$, then $\text{Re}(B^* B)$ is positive semidefinite, from which we know there exists $B_1 \in \mathbb{R}^{m_1 \times n_2}$, such that $\text{Re}(B^* B) = B_1^T B_1$ and $\text{rank}(B_1) = \text{rank}(\text{Re}(B^* B))$. Using such decomposition, we have $y^T \text{Re}(B^* B) y = 0$ if and only if $y^T B_1^T B_1 y = 0$ if and only if $B_1 y = 0$. Therefore, $\dim(\text{dim}(\text{Ker}(B)_\mathbb{R})) = n_2 - \text{rank}(B_1) = n_2 - \text{rank}(\text{Re}(B^* B))$. 

**Theorem 2.** Suppose $x \neq 0$, $A \in \mathbb{C}^{m \times d}$ with $\text{rank}(A) = d$. Then $x \in \mathbb{W}_A$ iff

$$\text{rank}(\text{Re}(\text{diag}(Ax)^* (I - AA^\dagger) \text{diag}(Ax))) = |N(Ax)| - 1$$

iff

$$\text{rank}\left(\begin{bmatrix} \text{Re}((AA^\dagger - I) \text{diag}(Ax)) \\ \text{Im}((AA^\dagger - I) \text{diag}(Ax)) \end{bmatrix}\right) = |N(Ax)| - 1.$$

**Proof:** We continue to use the notation $U_x$ and $V_x$. Consider the equation

$$(AA^\dagger - I)(U_x)_{N(Ax)} \Lambda = 0, \ \Lambda \in \mathbb{R}^{|N(Ax)|}$$

(4)

and let us show that the solution space of (4) is $V_x$. Suppose $\Lambda_0$ is a solution of (4), i.e., $(AA^\dagger - I)(U_x)_{N(Ax)} \Lambda_0 = 0$, then $AA^\dagger (U_x)_{N(Ax)} \Lambda_0 = (U_x)_{N(Ax)} \Lambda_0$, hence $\Lambda_0 \in V_x$. Inversely, suppose $\Lambda_0 \in V_x$, then there exists $y_0$, such that $A y_0 = (U_x)_{N(Ax)} \Lambda_0$, since $\text{rank}(A) = d$, we must have $y_0 = A^\dagger (U_x)_{N(Ax)} \Lambda_0$, hence $(AA^\dagger - I)(U_x)_{N(Ax)} \Lambda_0 = 0$. Applying Lemma 2 to compute $\dim(V_x)$, and note that

$$\text{rank}(\text{Re}(\text{diag}(Ax)^* (I - AA^\dagger) \text{diag}(Ax))) = \text{rank}(\text{Re}((U_x)_{N(Ax)}^* (I - AA^\dagger))(U_x)_{N(Ax)})),$$

$$\text{rank}\left(\begin{bmatrix} \text{Re}((AA^\dagger - I) \text{diag}(Ax)) \\ \text{Im}((AA^\dagger - I) \text{diag}(Ax)) \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} \text{Re}((AA^\dagger - I)(U_x)_{N(Ax)}) \\ \text{Im}((AA^\dagger - I)(U_x)_{N(Ax)}) \end{bmatrix}\right),$$

the result follows.
2.2 Special Matrix

In this subsection, we study a special form of $A$ given as follows: $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$. Denote the $(j, k)$-th entry of $A_1$ by $r_{j,k}e^{i\theta_{j,k}}$, $d + 1 \leq j \leq m$, $1 \leq k \leq d$, $r_{j,k} \geq 0$. Let $\gamma_j^T$ be the $j$th row of $A$. We also write $x$ in a polar form $x = [\|x_k|e^{i\alpha_k}\|_{1 \leq k \leq d}]$. For $d + 1 \leq j \leq m$, if $\gamma_j^T x \neq 0$, assume $\text{sign}(\gamma_j^T x) = e^{i\delta_j}$, then write

$$\Psi_j(x) = [r_{j,1}\sin(\theta_{j,1} + \alpha_1 - \delta_j) \quad r_{j,2}\sin(\theta_{j,2} + \alpha_2 - \delta_j) \quad \ldots \quad r_{j,d}\sin(\theta_{j,d} + \alpha_d - \delta_j)];$$

Otherwise, $\gamma_j^T x = 0$, we write

$$\Psi_j(x) = \begin{bmatrix} r_{j,1}\sin(\theta_{j,1} + \alpha_1) & r_{j,2}\sin(\theta_{j,2} + \alpha_2) & \ldots & r_{j,d}\sin(\theta_{j,d} + \alpha_d) \\ r_{j,1}\cos(\theta_{j,1} + \alpha_1) & r_{j,2}\cos(\theta_{j,2} + \alpha_2) & \ldots & r_{j,d}\cos(\theta_{j,d} + \alpha_d) \end{bmatrix}.$$

We define

$$E_A^0(x) = \begin{bmatrix} \Psi_{d+1}(x) \\ \Psi_{d+2}(x) \\ \vdots \\ \Psi_m(x) \end{bmatrix} \quad \text{and} \quad E_A(x) = (E_A^0(x))_{N(x)},$$

and call $E_A(x)$ the discriminant matrix under the special form of $A$.

**Theorem 3.** Suppose $x \neq 0$ and $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix} \in \mathbb{C}^{m \times d}$, then $x \in W_A$ if and only if $\text{rank}(E_A(x)) = |N(x)| - 1$.

**Proof:** Assume $y \in \mathbb{C}^d$, and write $J_1 = \{d + 1 \leq j \leq m : \gamma_j^T x \neq 0\}$ and $J_2 = \{d + 1 \leq j \leq m : \gamma_j^T x = 0\}$, then $\text{sign}(Ay) = \text{sign}(Ax)$ iff

$$\begin{cases} 
\text{sign}(y) = \text{sign}(x) \\
\sum_{k=1}^d r_{j,k}e^{i\theta_{j,k}}y_k > 0, \quad \forall j \in J_1 \\
\sum_{k=1}^d r_{j,k}e^{i\theta_{j,k}}y_k = 0, \quad \forall j \in J_2.
\end{cases}$$

Since $\text{sign}(y) = \text{sign}(x)$ would imply $N(y) = N(x)$ and $\text{sign}(y_k) = \text{sign}(x_k) = e^{i\alpha_k}$ when $k \in N(x)$, it is equivalent to

$$\begin{cases} 
\text{sign}(y) = \text{sign}(x) \\
\sum_{k \in N(x)} r_{j,k}e^{i\theta_{j,k}+\alpha_k-\delta_j}|y_k| > 0, \quad \forall j \in J_1 \\
\sum_{k \in N(x)} r_{j,k}e^{i\theta_{j,k}+\alpha_k}|y_k| = 0, \quad \forall j \in J_2.
\end{cases}$$

By considering the real part and the imaginary part, we know it is also equivalent to

$$\begin{cases} 
\text{sign}(y) = \text{sign}(x) \\
E_A(x)|y_{N(x)}| = 0 \\
\sum_{k \in N(x)} r_{j,k}\cos(\theta_{j,k} + \alpha_k - \delta_j)|y_k| > 0, \quad \forall j \in J_1.
\end{cases}$$
Specifically, let $y = x$, we know
\[
\begin{cases}
E_A(x)|x^N(x)| = 0 \\
\sum_{k \in N(x)} r_{j,k} \cos(\theta_{j,k} + \alpha_k - \delta_j) |x_k| > 0, \quad \forall j \in J_1.
\end{cases}
\]

For the “if” part, assume $\text{sign}(Ay) = \text{sgin}(Ax)$ then we have $E_A(x)|y^{N(x)}| = 0$. By combining with $\text{rank}(E_A(x)) = |N(x)| - 1$ then we have $|y^{N(x)}|$ and $|x^{N(x)}|$ differ by a positive number. Note that $\text{sign}(y) = \text{sign}(x)$, then $y$ and $x$ differ by a positive number, hence $x \in \mathbb{W}_A$.

For the “only if” part, we first note that $E_A(x)|x^{N(x)}| = 0$ implies $\text{rank}(E_A(x)) \leq |N(x)|-1$. Assume that $\text{rank}(E_A(x)) < |N(x)| - 1$. Then $\text{dim}(\text{Ker}(E_A(x))) \geq 2$ and hence there exists $\Lambda_2 \in \mathbb{R}^{N(x)} \cap \text{Ker}(E_A(x))$ that is linearly independent with $|x^{N(x)}|$. We construct $y \in \mathbb{C}^d$ such that $\text{sign}(y) = \text{sign}(x)$ and $|y^{N(x)}| = \Lambda_2$. Now we can make $\Lambda_2$ sufficiently close to $|x^{N(x)}|$ to guarantee
\[
\sum_{k \in N(x)} r_{j,k} \cos(\theta_{j,k} + \alpha_k - \delta_j) |y_k| > 0, \quad \forall j \in J_1.
\]

Therefore, we have $\text{sign}(Ay) = \text{sign}(Ax)$, while obviously $y \notin \mathbb{R}_+x$. Hence $x \notin \mathbb{W}_A$, which is a contradiction. The result follows. \hfill \Box

**Remark 1.** Note that we have constructed four discriminant matrices to characterize recoverable signals, i.e., $\mathbb{D}_A(x)$, $\mathbb{E}_A(x)$ and two other matrices with the appearance of $A^1$. In general, $\mathbb{D}_A(x)$ is of a fixed size $2m \times (2d + m)$, and has relatively simple polynomials entries, hence it is suitable for a general analysis, see Theorem 7. Though $\mathbb{E}_A(x)$ only works for $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$ and has more complex entries, it can sometimes facilitate our analysis because of its smaller size, see Theorem 4 and 5. Because of $A^1$, the other two matrices seem unfavourable for theoretical analysis. We point out that $\text{Re}(\text{diag}(Ax)^*(I - AA^\dagger)\text{diag}(Ax))$ (or $\text{Re}(U_x^*(I - AA^\dagger)U_x)$) provides a formulation of the initial problem as a quadratic programming problem, see [14] for more details. Since this work focuses on theoretical part, we mainly use $\mathbb{D}_A(x)$ and $\mathbb{E}_A(x)$.

Next we give an example to illustrate the meanings and advantages of our general framework, that is, it helps us obtain an insight into some existing facts.

**Remark 2.** Many previous papers have reported the fact that the Fourier phase has difficulty in recovering symmetric signals. Most existing theorems exclude symmetric signals, e.g., [10, 15] and symmetric signals have the poorest convergence characteristics or even fail in applications, see [21]. In [22], their method can handle symmetric signals based on a preprocessing procedure. More precisely, a random mask matrix is used to destroy the symmetry. By using $\mathbb{E}_A(x)$, we would like to validate why symmetric signal cannot be recovered. Consider a symmetric signal $x = (x_1, \ldots, x_p, x_{p+1}, \ldots, x_{2p}) \in \mathbb{R}_{2p}^+$, where $x(k) = x(2p + 1 - k)$. We compute the Fourier transform of $x$ with frequency $e^{i\theta}$ as follows:

\[
\mathcal{F}_x(\theta) = \sum_{k=1}^{2p} e^{-ik\theta} x_k = \sum_{k=1}^p e^{-ik\theta} x_k + \sum_{k=p+1}^{2p} e^{-ik\theta} x_k
\]

\[
= \sum_{k=1}^p (e^{-ik\theta} + e^{-i(2p+1-k)\theta}) x_k = e^{-i(p+\frac{1}{2})\theta} \left[ \sum_{k=1}^p 2 \cos \left( p + \frac{1}{2} - k \right) \theta \right] x_k.
\]


Assume $A = \begin{bmatrix} I_{2p} & F_0 \end{bmatrix}$, where $F_0$ provides the Fourier phase sampling. For sampling frequency $e^{i\theta_j}$ where $0 < \theta_j < \frac{\pi}{2p-1}$, we obtain the phase information $e^{i\delta_j} = \text{sign}(F_x(\theta_j)) = e^{-i(p+\frac{1}{2})\theta_j}$, hence $\delta_j = -(p + \frac{1}{2})\theta_j$, then this measurement provides a row

$$
\begin{bmatrix}
\sin \left( \left( p + \frac{1}{2} - k \right) \theta_j \right) \\
\sin \left( \left( p - \frac{1}{2} \right) \theta_j \right) \\
\vdots \\
\sin \left( \frac{1}{2} \theta_j \right) \\
\sin \left( -\frac{1}{2} \theta_j \right) \\
\sin \left( \frac{1}{2} - p \right) \theta_j
\end{bmatrix}_{1 \leq k \leq 2p}
$$

for $E_A(x)$. Assume that we only sample in $0 < \theta < \frac{\pi}{2p-1}$, then for any $1 \leq k \leq 2d$, $[E_A(x)]_{\{k\}} = -[E_A(x)]_{\{2p+1-k\}}$, which implies $\text{rank}(E_A(x)) \leq p < 2p-1$, hence we cannot recover $x$. Actually we cannot recover $x$ no matter how many measurements are used.

### 3 Signal Recovery

In this section, we would like to search for $A \in \mathbb{C}^{m \times d}$ that can recover all signals, i.e., $W_A = \mathbb{C}^d$. As demonstrated above, there is no loss of generality if we only consider $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$. Let us verify the existence of such $A$ as the first step. We consider $A_1 \in \mathbb{C}^{d(d-1) \times d}$ that can provide the following measurements

$$\{\text{sign}(x_k + x_l), \text{sign}(x_k + ix_l) : 1 \leq k < l \leq d\}.$$

From the first $d$ measurements we know $\text{sign}(x)$, which indicates $N(x)$ and

$$\{\text{sign}(x_k) : k \in N(x)\}.$$

For any $k, l \in N(x), k \neq l$, given

$$\text{sign}(x_k), \text{sign}(x_l), \text{sign}(x_k + x_l), \text{sign}(x_k + ix_l),$$

we can easily determine $\frac{|x_k|}{|x_l|}$ as follows. When $\text{sign}(x_k) \neq \pm \text{sign}(x_l)$, then $x_k + x_l \neq 0$, we have

$$\frac{x_k + x_l}{\text{sign}(x_k + x_l)} > 0, \quad \text{i.e.,} \quad \frac{|x_k|\text{sign}(x_k) + |x_l|\text{sign}(x_l)}{\text{sign}(x_k + x_l)} > 0,$$

it implies

$$\text{Im} \left( \frac{\text{sign}(x_k)}{\text{sign}(x_k + x_l)} \right) \frac{|x_k|}{|x_l|} + \text{Im} \left( \frac{\text{sign}(x_l)}{\text{sign}(x_k + x_l)} \right) = 0. \quad (5)$$

Let us show $\text{Im} \left( \frac{\text{sign}(x_k)}{\text{sign}(x_k + x_l)} \right) \neq 0$, otherwise $\frac{\text{sign}(x_k)}{\text{sign}(x_k + x_l)} \in \mathbb{R}$, then $\frac{x_k}{x_k + x_l} \in \mathbb{R}$, hence $\frac{x_k}{x_l} \in \mathbb{R}$, which contradicts $\text{sign}(x_k) \neq \pm \text{sign}(x_l)$. Therefore, we can determine $\frac{|x_k|}{|x_l|}$ from (5). When $\text{sign}(x_k) = \pm \text{sign}(x_l), \text{sign}(x_k) \neq \pm \text{sign}(ix_l)$ must hold, by using $\text{sign}(x_k + ix_l), we
can similarly determine $|x_k|/|x_l|$, i.e., $|x_k|/|x_l|$. To sum up, from $d + d(d - 1) = d^2$ measurements, we can determine $x \in \mathbb{C}^d$ up to a positive scalar. Therefore, we know for any positive integer $d$, there exists $A \in \mathbb{C}^{d \times d}$, such that $\mathcal{W}_A = \mathbb{C}^d$.

Similar to the basic pursuit in minimal measurement number of phase retrieval, it is also of interest to explore

$$m(d) = \min\{m \in \mathbb{N}_+ : \exists A \in \mathbb{C}^{m \times d}, \ s.t. \mathcal{W}_A = \mathbb{C}^d\},$$

namely the minimal measurement number of magnitude retrieval. In the following analysis, we assume $A$ has no zero row, i.e., $\forall 1 \leq j \leq m, \gamma_j^T \neq 0$. It is trivial to remove such rows if they exist. We define the set

$$\mathbb{H}_A = \{x \in \mathbb{C}^d : \forall 1 \leq j \leq m, \gamma_j^Tx \neq 0\}$$

Note that $\text{Ker}(\gamma_j^T) = \{x \in \mathbb{C}^d : \gamma_j^T x = 0\}$ is $d - 1$ dimensional linear subspace in $\mathbb{C}^d$, by writing

$$\mathbb{H}_A = \bigcap_{j=1}^m [\text{Ker}(\gamma_j^T)]^c = \bigcup_{j=1}^m [\text{Ker}(\gamma_j^T)]^c,$$

where $Z^c$ refers to the complement of a set $Z$. It is clear $\mathbb{H}_A$ is Zariski open. By applying $\mathbb{E}_A(x)$, we can derive the following lower bound.

**Theorem 4.** When $d > 1$, $m(d) \geq 2d$.

**Proof:** Assume $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$ and $\mathcal{W}_A = \mathbb{C}^d$. For $x \in \mathbb{H}_A$, it is evident that $\mathbb{E}_A(x) = (r_{j,k} \sin(\theta_{j,k} + \alpha_k - \delta_j))_{d+1 \leq j \leq m, 1 \leq k \leq d} \in \mathbb{R}^{(m-d) \times d}$, and $\text{rank}(\mathbb{E}_A(x)) = |N(x)| - 1 = d - 1$. By setting the phase of $x$ by $\alpha_k = -\theta_{d+1,k}$, we have

$$x = [\lambda_1 e^{-i\delta_{d+1}}, \lambda_2 e^{-i\delta_{d+1}}, ..., \lambda_d e^{-i\delta_{d+1}}]^T = \Lambda E,$$

here $E = \text{diag}(e^{-i\delta_{d+1}}, ..., e^{-i\delta_{d+1}})$ is fixed, $\Lambda = [\lambda_1, ..., \lambda_d]^T \in \mathbb{R}^d$ remains to be designated. For $j \in \{d+1, ..., m\}$, if $\text{Re}(\gamma_j^T E) \neq 0$, we define $\xi_j^T = \text{Re}(\gamma_j^T E)$, otherwise, we let $\xi_j^T = 0$. Evidently, we have $\xi_j \in \mathbb{R}^d \setminus \{0\}$ for any $d + 1 \leq j \leq m$. Therefore, we can choose $\Lambda = \Lambda_0 = [\lambda_{1,0}, \lambda_{2,0}, ..., \lambda_{d,0}]^T \in \mathbb{R}^d_+$ s.t. $\xi_j^T \Lambda_0 \neq 0, \forall d + 1 \leq j \leq m$, and thus $x_0 = \Lambda_0 e^{i\delta_{d+1}}$ satisfies $\gamma_j^T x_0 \neq 0, \forall d + 1 \leq j \leq m$, then $x_0 \in \mathbb{H}_A$. Calculate that

$$e^{i\delta_{d+1}} = \text{sign} \left( \sum_{k=1}^d r_{d+1,k} \lambda_{k,0} e^{i(\theta_{d+1,k} + \alpha_k)} \right) = \text{sign} \left( \sum_{k=1}^d r_{d+1,k} \lambda_{k,0} \right) = 1.$$

hence $\delta_{d+1} = 0$, which delivers that $\sin(\theta_{d+1,k} + \alpha_k - \delta_{d+1}) = 0$, that is, the first row of $\mathbb{E}_A(x_0)$ equals 0. By applying $\text{rank}(\mathbb{E}_A(x_0)) = d - 1$, we have $m - d - 1 \geq d - 1$, then $m \geq 2d$. □

Consider $A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$, we can obtain an upper bound of $m(d)$. We need to decide all the $[I_d \\ A_1] \in \mathbb{C}^{m \times d}$ that fail to recover some signals. Let $\mathcal{X} = \{A = \begin{bmatrix} I_d \\ A_1 \end{bmatrix} \in \mathbb{C}^{m \times d} : \mathcal{W}_A \neq \mathbb{C}^d\}$, then we know $[I_d \\ A_1] \in \mathcal{X}$ if and only if there exist $x, y \in \mathbb{C}^d$, such that $\text{sign}(x) = \text{sign}(y)$,
\[ \text{sign}(A_1x) = \text{sign}(A_1y), \text{ and } y \notin \mathbb{R}_+x. \] Obviously, \( \text{sign}(x) = \text{sign}(y), y \notin \mathbb{R}_+x \) if and only if \( \text{sign}(x) = \text{sign}(y) \) and there exist \( 1 \leq p < q \leq d \), such that \( x_p \neq 0, x_q \neq 0, \) and \( \frac{x_q}{x_p} \neq \frac{y_q}{y_p}. \)

Now we want to normalize \( x_p, y_p \) to become 1, and then move the \( p \)-th, \( q \)-th entry of \( x, y \) to the first and the second entry respectively. Let \( E(j, k) \) be the matrix defined by exchanging the \( j \)-th row and \( k \)-th row of \( I_d \), for \( p < q \) we write

\[
\begin{cases}
A_1^{pq} = A_1E(1, p)E(2, q) \\
x = E(1, p)E(2, q)\frac{x}{x_p} \\
y = E(1, p)E(2, q)\frac{y}{y_p}
\end{cases}.
\]

Then we have \( \text{sign}(\hat{x}) = E(1, p)E(2, q)\text{sign}(x_p)^{-1}\text{sign}(x) = \text{sign}(E(1, p)E(2, q)\text{sign}(y_p)^{-1}\text{sign}(y) = \text{sign}(\hat{y}), \text{sign}(A_1^{pq}\hat{x}) = \text{sign}(x_p)^{-1}\text{sign}(A_1x) = \text{sign}(y_p)^{-1}\text{sign}(A_1y) = \text{sign}(A_1^{pq}\hat{y}), \) and \( \hat{x}_1 = \hat{y}_1 = 1, \hat{x}_2 \neq \hat{y}_2 \). Now \( \text{sign}(\hat{x}) = \text{sign}(\hat{y}) \) is equivalent to \( \hat{x}_2 \neq 0 \), and there exist \( \lambda_2 \in \mathbb{R}_+ \setminus \{1\}, \lambda_3, ..., \lambda_d \in \mathbb{R}_+ \), such that \( \hat{y}_k = \lambda_k \hat{x}_k \) for any \( 2 \leq k \leq d \). Let \( \lambda_1 = 1 \), then we also have \( \hat{y}_1 = \lambda_1 \hat{x}_1 \). For ease of notations we write \( A_1^{pq} = [b_{j,k}] \) where \( d+1 \leq j \leq m, 1 \leq k \leq d \). Then \( \text{sign}(A_1^{pq}\hat{x}) = \text{sign}(A_1^{pq}\hat{y}) \) is equivalent to there exist positive numbers \( \lambda_{d+1}, ..., \lambda_m \) such that for any \( d+1 \leq j \leq m \), it holds that \( \sum_{k=1}^{d} b_{j,k}\hat{x}_k = \lambda_j \sum_{k=1}^{d} b_{j,k}\lambda_k\hat{x}_k \).

Let us further distinguish whether \( \lambda_j \) equals 1 when \( d+1 \leq j \leq m \), we write \( Y = \{d+1 \leq j \leq m : \lambda_j \neq 1\} \) and \( Y_1 = [m] \setminus ([d] \cup Y) = \{d+1 \leq j \leq m : \lambda_j = 1\} \). Recall that \( \hat{x}_1 = \lambda_1 = 1 \), when \( j \in Y \), \( \sum_{k=1}^{d} b_{j,k}\hat{x}_k = \lambda_j \sum_{k=1}^{d} b_{j,k}\lambda_k\hat{x}_k \) is equivalent to

\[
b_{j,1} = \frac{\sum_{k=2}^{d} b_{j,k}\hat{x}_k(1-\lambda_j\lambda_k)}{\lambda_j-1}, \tag{6}
\]

and when \( j \in Y_1 \), since \( \hat{x}_2 \neq 0 \) and \( \lambda_2 \neq 1 \), \( \sum_{k=1}^{d} b_{j,k}\hat{x}_k = \lambda_j \sum_{k=1}^{d} b_{j,k}\lambda_k\hat{x}_k \) is equivalent to

\[
b_{j,2} = \frac{\sum_{k=3}^{d} b_{j,k}\hat{x}_k(1-\lambda_k)}{(\lambda_2-1)\hat{x}_2}. \tag{7}
\]

For \( Y \subset [m] \setminus [d] \), we define

\[
T_Y = \{(j, k) : d+1 \leq j \leq m, 1 \leq k \leq d, k \neq 1 \text{ when } j \in Y, k \neq 2 \text{ when } j \in Y_1\}.
\]

Therefore, \( \left[ I_d \over A_1 \right] \in \mathcal{X} \) if and only if we can find \( 1 \leq p < q \leq d \), \( Y \subset \{d+1, ..., m\} \), such that \( A_1^{pq} = [b_{j,k}] \) satisfies what follows:

There exist \( f_{j,k} \in \mathbb{C} \) where \( (j, k) \in T_Y, \hat{x}_2 \in \mathbb{C} \setminus \{0\} \), \( \hat{x}_k \in \mathbb{C} \) where \( 3 \leq k \leq d, \lambda_2 \in \mathbb{R}_+ \setminus \{1\}, \lambda_k \in \mathbb{R}_+ \) where \( 3 \leq k \leq d \), and \( \lambda_j \in \mathbb{R}_+ \setminus \{1\} \) where \( j \in Y \), such that \( b_{j,k} = f_{j,k} \) when \( (j, k) \in T_Y, \) \( \text{(6)} \) holds when \( j \in Y \), and \( \text{(7)} \) holds when \( j \in Y_1 \).
More precisely, we identify $A_1$ such that \[ \begin{bmatrix} I_d \\ A_1 \end{bmatrix} \in \mathcal{X} \] with the image of some mappings. We use $x_k$ to replace $\hat{x}_k$. Suppose $Y \subset \{d+1, ..., m\}$, note that $|T_Y| = (m-d)(d-1)$, we define $f_Y$ from $\Omega_Y = \mathbb{C}^{(m-d)(d-1)} \times (\mathbb{C} \setminus \{0\}) \times \mathbb{C}^{d-2} \times (\mathbb{R}_+ \setminus \{1\}) \times \mathbb{R}_+^{d-2} \times \mathbb{R}_+ \setminus \{1\})^{|Y|}$ to $\mathbb{C}^{(m-d)d}$ by letting the image of \( ([f_{j,k}]_{(j,k) \in T_Y}, x_2, [x_k]_{3 \leq k \leq d}; \lambda_2, [\lambda_k]_{3 \leq k \leq d}; [\lambda_j]_{j \in Y}) \in \Omega_Y \) be \[ b_{j,k} = f_{j,k} \] when $(j, k) \in T_Y$, $b_{j,1} = \sum_{k=2}^{d} f_{j,k} x_k (1-\lambda_j \lambda_k) \lambda_j - 1$ when $j \in Y$, and $b_{j,2} = \sum_{k=3}^{d} f_{j,k} x_k (1-\lambda_k) (\lambda_2 - 1)x_2$ when $j \in Y_1$. Let $P_{A_1}(\begin{bmatrix} I_d \\ A_1 \end{bmatrix}) = A_1$, we actually have proved that $P_{A_1}(\mathcal{X}) = \bigcup_{1 \leq p < q \leq d} \bigcup_{Y \subset \{m\} \setminus \{d\}} f_Y(\Omega_Y) \mathcal{E}(1, p) \mathcal{E}(2, q)$. (8)

**Theorem 5.** When $d \geq 2$, we have $m(d) \leq 4d - 2$.

**Proof:** By identifying $z \in \mathbb{C}$ with $(Re(z) \text{Im}(z))^T \in \mathbb{R}^2$, $f_Y$ can be viewed as a mapping from $\mathbb{R}^{2m-d(d-1)} \times \mathbb{R}^{2(d-2)} \times \mathbb{R}_+ \setminus \{1\} \times \mathbb{R}_+^{d-2} \times \mathbb{R}_+ \setminus \{1\})^{|Y|}$ to $\mathbb{R}^{2m-dd}$. Note that the definition domain is an open subset of $\mathbb{R}^{2m-2d+3(d-1)+|Y|}$. Let us show $f_Y$ is a smooth mapping. We use the notations used to define $f_Y$ above, then we know $Re(b_{j,k}) = Re(f_{j,k})$, $Im(b_{j,k}) = Im(f_{j,k})$ when $(j, k) \in T_Y$, and

\[
\begin{cases}
Re(b_{j,1}) = \frac{\sum_{k=2}^{d} (1-\lambda_j \lambda_k) (Re(f_{j,k}) Re(x_k) - Im(f_{j,k}) Im(x_k))}{\lambda_j - 1} \\
Im(b_{j,1}) = \frac{\sum_{k=2}^{d} (1-\lambda_j \lambda_k) (Re(f_{j,k}) Im(x_k) + Im(f_{j,k}) Re(x_k))}{\lambda_j - 1}
\end{cases}
\]

when $j \in Y$, and

\[
\begin{cases}
Re(b_{j,2}) = \frac{\sum_{k=3}^{d} (1-\lambda_k) (Re(f_{j,k}) Re(x_k) Re(x_2) + Re(f_{j,k}) Im(x_k) Im(x_2) - Im(f_{j,k}) Im(x_k) Re(x_2) + Im(f_{j,k}) Re(x_k) Im(x_2))}{\lambda_2 - 1} \\
Im(b_{j,2}) = \frac{\sum_{k=3}^{d} (1-\lambda_k) (Re(f_{j,k}) Im(x_k) Re(x_2) - Re(f_{j,k}) Re(x_k) Im(x_2) + Im(f_{j,k}) Re(x_k) Re(x_2) + Im(f_{j,k}) Im(x_k) Im(x_2))}{(\lambda_2 - 1) (Re(x_2)^2 + Im(x_2)^2)}
\end{cases}
\]

when $j \in Y_1$. Then the smoothness of $f_Y$ follows. When $m \geq 4d - 2$, we have

\[ 2(m-d)d > (2m-2d+3)(d-1) + m-d \geq (2m-2d+3)(d-1) + |Y|, \]

for any $Y \subset \{m\} \setminus \{d\}$, then from Sard Theorem, we know $f_Y(\Omega_Y)$ has zero Lebesgue measure, which implies that

\[ \bigcup_{1 \leq p < q \leq d} \bigcup_{Y \subset \{m\} \setminus \{d\}} f_Y(\Omega_Y) \mathcal{E}(1, p) \mathcal{E}(2, q) \]

has zero Lebesgue measure, hence it cannot be $\mathbb{R}^{2m-dd}$. From \cite{S}, we know $P_{A_1}(\mathcal{X}) \neq \mathbb{C}^{(m-d)d}$, hence there exists $A_1 \notin P_{A_1}(\mathcal{X})$, such that $A' = \begin{bmatrix} I_d \\ A_1 \end{bmatrix}$ satisfies $\mathbb{W}_{A'} = \mathbb{C}^d$. \( \square \)

**Remark 3.** Note that we have proved the Lebesgue measure of $P_{A_1}(\mathcal{X})$ is zero, hence almost all $A_1 \in \mathbb{C}^{m \times d}$ where $m \geq 3d - 2$ allow $\begin{bmatrix} I_d \\ A_1 \end{bmatrix}$ to recover all signals. Actually, we can further prove almost all $A \in \mathbb{C}^{m \times d}$ where $m \geq 4d - 2$ can recover all signals.
4 Generic Signal Recovery

From a practical point of view, a very large \( \mathcal{W}_A \) that contains most signals is satisfactory. We are going to demonstrate \( 2d - 1 \) generic measurements can handle recovering generic signals (see Section 4) or a specific signal (see Section 5), and hence \( 2d \) in most situations. Moreover, \( 2d \) are going to demonstrate nowhere dense with \( 2d \) are nowhere dense with \( 2d \) in our case.

To conduct more subtle analysis, we extend the definition of \( \mathbb{H}_A \). Suppose \( S \subset [m] \), write \( [m] \setminus S \) as \( S^c \), we define

\[
\mathbb{H}_A(S) = \{ x \in \mathbb{C}^d : \gamma_j^T x = 0, \forall j \in S; \gamma_j^T x \neq 0, \forall j \in S^c \},
\]

i.e.,

\[
\mathbb{H}_A(S) = \text{Ker}(A^S) \cap \mathbb{H}_{AS^c}.
\]

Note that we have the following partition

\[
\mathbb{C}^d = \bigcup_{S \subset [m]} \mathbb{H}_A(S),
\]

where \( \mathbb{H}_A(S) \cap \mathbb{H}_A(T) = \emptyset, \forall S \neq T \), and evidently

\[
\mathbb{H}_A(\emptyset) = \mathbb{H}_A\text{.}
\]

In addition, we use \([S]\) to denote \( |S| \), i.e.,

\[
[S] = \{1, 2, ..., |S|\}.
\]

**Theorem 6.** Suppose \( A \in \mathbb{C}^{m \times d} \) has no zero row with \( d > 1 \). When \( m \leq 2d - 2 \), \( \mathcal{W}_A \) is nowhere dense. When \( m \geq 2d - 1 \), there are generic \( A \) such that

\[
\forall S \subset [m], \mathcal{W}_A \cap \text{Ker}(A^S) \text{ is generic in } \text{Ker}(A^S).
\]

Specifically, let \( S = \emptyset \), we have \( \mathcal{W}_A \cap \text{generic in } \mathbb{C}^d \).

For ease of presentation, we begin with several lemmas. Lemma 3 aims at characterizing \( \mathcal{W}_A \cap \mathbb{H}_A \).

**Lemma 3.** \( x \in \mathcal{W}_A \cap \mathbb{H}_A \) if and only if \( \text{rank}(\mathbb{D}_A(x)) \geq 2d + m - 1 \).

**Proof:** The "only if" part comes directly from Theorem 1. For the "if" part, recall Lemma 1 we have \( \text{dim}(V_x) \geq 1 \), then from Lemma 3, we have \( \text{dim}(V_x) = 2d + |N(Ax)| - \text{rank}(\mathbb{D}_A(x)) \), hence \( \text{rank}(\mathbb{D}_A(x)) \leq 2d + |N(Ax)| - 1 \), combining with \( \text{rank}(\mathbb{D}_A(x)) \geq 2d + m - 1 \), we have \( 2d + |N(Ax)| - 1 \geq 2d + m - 1 \), i.e., \( |N(Ax)| \geq m \), then \( |N(Ax)| = m \), which implies \( x \in \mathbb{H}_A \) and \( \text{rank}(\mathbb{D}_A(x)) = 2d + |N(Ax)| - 1 \), from Theorem 1 again, we have \( x \in \mathcal{W}_A \), then \( x \in \mathcal{W}_A \cap \mathbb{H}_A \).

Next we propose a lemma of the rank of a fractional function matrix.

**Lemma 4.** Assume \( x \in \mathbb{C}^n, \Phi(x) = \left[ \begin{array}{c} f_{i,j}(x) \\ g_{i,j}(x) \end{array} \right] \in \mathbb{C}^{n_1 \times n_2} \), where

\[
\{ f_{i,j}(x), g_{i,j}(x) \} \subset \mathbb{C}[\text{Re}(x_1), \text{Re}(x_2), ..., \text{Re}(x_{n_1}), \text{Im}(x_1), \text{Im}(x_2), ..., \text{Im}(x_{n_2})], g_{i,j} \neq 0
\]

are polynomials of \( \text{Re}(x) \) and \( \text{Im}(x) \) with complex coefficients, and \( x \in \Omega = \{ x : g_{i,j}(x) \neq 0, \forall 1 \leq i \leq n_1, 1 \leq j \leq n_2 \} \). Given \( r \in \mathbb{N}_+ \), then \( \{ x : \text{rank}(\Phi(x)) \geq r \} \) is Zariski open set.
Proof: When \( r > n_1 \) or \( r > n_2 \), \( \{ x : \text{rank}(\Phi(x)) \geq r \} = \varnothing \) is Zariski open. When \( r \leq n_1 \) and \( r \leq n_2 \), we use \( \Phi_t(x), t \in T \) to denote all the \( r \times r \) submatrices of \( \Phi \), where \( T \) is a finite set. Then we have

\[
\{ x : \text{rank}(\Phi(x)) \geq r \} = \bigcup_{t \in T} \{ x \in \Omega : \det(\Phi_t(x)) \neq 0 \} = \bigcup_{t \in T} \{ x : g_{i,j}(x) \neq 0, \forall i, j; \det(\Phi_t(x)) \neq 0 \} = \bigcup_{t \in T} \{ x : g_{i,j}(x) \neq 0, \forall i, j; \prod_{i,j} g_{i,j}(x) \det(\Phi_t(x)) \neq 0 \},
\]

note that

\[
\prod_{i,j} g_{i,j}(x) \det(\Phi_t(x)) \in \mathbb{C}[\text{Re}(x_1), \text{Re}(x_2), ..., \text{Re}(x_n), \text{Im}(x_1), \text{Im}(x_2), ..., \text{Im}(x_n)].
\]

When \( \{ x : \text{rank}(\Phi(x)) \geq r \} = \varnothing \), the result is trivial. When \( \{ x : \text{rank}(\Phi(x)) \geq r \} \neq \varnothing \), there exists \( t_0 \in T \), such that \( \prod_{i,j} g_{i,j}(x) \det(\Phi_t(x)) \) is nonzero polynomial, hence \( \{ x : \text{rank}(\Phi(x)) \geq r \} \) is non-empty Zariski open set.

Next we need to establish a lemma to provide recoverable guarantee for the situation \( x \notin \mathbb{H}_A \), i.e., \( \exists S \subset [m], S \neq \varnothing \), such that \( x \in \mathbb{H}_A(S) \). This lemma allows us to convert the case \( x \notin \mathbb{H}_A \) to the preferable case \( x \in \mathbb{H}_A \), so Lemma 4 can be applied.

Lemma 5. Assume \( S \subset [m], 1 \leq |S| < d \), and \( \text{rank}(A^S_{|S|}) = |S|, x \in \text{Ker}(A^S) \). Define \( A(S) = A^S_{|S|} - A^S_{|S|} A^S_{[d]|S|}^{-1} A^S_{|d|\setminus S|} \in \mathbb{C}^{(m - |S| \times (d - |S|))} \). Then we have \( x \in \mathbb{W}_A \cap \mathbb{H}_A(S) \) if and only if \( x_{[d]|S|} \in \mathbb{W}_A(S) \cap \mathbb{H}_A(S) \).

Proof: From \( x \in \text{Ker}(A^S) \) we have \( A^S x = 0 \), that is, \( [A^S_{|S|} A^S_{[d]|S|}] \begin{bmatrix} x_{|S|} \\ x_{[d]|S|} \end{bmatrix} = 0 \), which implies \( x_{|S|} = -(A^S_{|S|})^{-1} A^S_{[d]|S|} x_{[d]|S|} \), so we have \( x = \begin{bmatrix} -(A^S_{|S|})^{-1} A^S_{[d]|S|} y_{[d]|S|} \\ x_{[d]|S|} \end{bmatrix} \).

By simple calculation, we have \( A^{S^c} x = A(S)x_{[d]|S|} \). For the "if" part, \( x_{[d]|S|} \in \mathbb{H}_A(S) \) iff \( A(S)x_{[d]|S|} \) contains no zeroes, i.e., \( A^{S^c} x \) contains no zeroes. Note that \( A^{S^c} x = 0 \), we have \( x \in \mathbb{H}_A(S) \). Suppose \( y \in \mathbb{C}^d \), s.t. \( \text{sign}(Ay) = \text{sign}(Ax) \), i.e., \( \text{sign}(A^S y) = \text{sign}(A^S x) = 0 \) and \( \text{sign}(A^{S^c} y) = \text{sign}(A^{S^c} x) \). From \( \text{sign}(A^S y) = \text{sign}(A^S x) = 0 \), we know

\[
y = \begin{bmatrix} -(A^S_{|S|})^{-1} A^S_{[d]|S|} y_{[d]|S|} \\ y_{[d]|S|} \end{bmatrix},
\]

by combining with \( \text{sign}(A^{S^c} y) = \text{sign}(A^{S^c} x) \), we know \( \text{sign}(A(S)y_{[d]|S|}) = \text{sign}(A(S)x_{[d]|S|}) \), using \( x_{[d]|S|} \in \mathbb{W}_A(S) \), we have \( y_{[d]|S|} \in \mathbb{R}_+ x_{[d]|S|} \), and hence \( y \in \mathbb{R}_+ x \), this proves \( x \in \mathbb{W}_A \). For the "only if" part, we also have \( x_{[d]|S|} \in \mathbb{H}_A(S) \) since \( x \in \mathbb{H}_A(S) \). Suppose \( \hat{y} \in \mathbb{C}^{d-|S|} \), s.t. \( \text{sign}(A(S)\hat{y}) = \text{sign}(A(S)x_{[d]|S|}) \), consider \( x \) and \( y = \begin{bmatrix} -(A^S_{|S|})^{-1} A^S_{[d]|S|} y_{[d]|S|} \\ \hat{y} \end{bmatrix} \in \mathbb{C}^d \), then \( \text{sign}(A^{S^c} y) = \text{sign}(A(S)\hat{y}) = \text{sign}(A(S)x_{[d]|S|}) = \text{sign}(A^{S^c} x) \), it’s easy to verify that \( \text{sign}(A^S y) = 0 = \text{sign}(A^S x) \), hence \( \text{sign}(Ay) = \text{sign}(Ax) \), by using \( x \in \mathbb{W}_A \), we have \( y \in \mathbb{R}_+ x \), then \( \hat{y} \in \mathbb{R}_+ x_{[d]|S|} \) follows. \( \square \)
Now it is ready to show prove Theorem 6.

Proof of Theorem 6: When \( m \leq 2d - 2 \), since \( D_A(x) \in \mathbb{C}^{2m \times (2d + m)} \), we have \( \text{rank}(D_A(x)) \leq 2m < 2d + m - 1 \), from Lemma 5 we know \( \mathcal{W}_A \cap \mathbb{H}_A = \emptyset \), hence

\[
\mathcal{W}_A \subset (\mathbb{H}_A)^c = \bigcup_{j=1}^{m} \text{Ker}(\gamma_j^T),
\]

evidently \( \mathcal{W}_A \) is nowhere dense.

When \( m \geq 2d - 1 \), write \( S^c = [m] \setminus S \). In this proof we need to consider \( e_1 \) of different dimensions and hence let \( e_1(n_1) \) be \((1,0,...,0)^T \in \mathbb{C}^{n_1}\). Define property (a)(b)(c) as follows:

Property (a): \( \forall 0 < |S| \leq d \), \( \text{rank}(A^S_{[S]}) = |S| \).

When Property (a) holds, we define

\[
A(S) = A^{S^c}_{[d]\setminus[S]} - A^{S^c}_{[S]}(A^{S}_{[S]})^{-1}A^S_{[d]\setminus[S]} \in \mathbb{C}^{(m-|S|)\times (d-|S|)}
\]
as Lemma 5. We can further define Property (b).

Property (b): \( \forall 0 < |S| < d \), \( e_1(d - |S|) \in \mathcal{W}_A(S) \cap \mathbb{H}_A(S) \).

Property (c): \( e_1(d) \in \mathcal{W}_A \cap \mathbb{H}_A \).

We consider \( \Xi = \{ A \in \mathbb{C}^{m \times d} : A \text{ has property (a)(b)(c)} \} \). Let us verify \( \Xi \) is generic in \( \mathbb{C}^{m \times d} \) as a first step. Obviously, generic \( A \) have property (a). From Lemma 3 we know \( e_1(d) \in \mathcal{W}_A \cap \mathbb{H}_A \) if and only if \( \text{rank}(D_A(e_1(d))) \geq 2d + m - 1 \), since the entries of \( \text{rank}(D_A(e_1(d))) \) are rational functions (actually polynomials here) of \( \text{Re}(A) \) and \( \text{Im}(A) \), from Lemma 4, we know \( A \) having property (c) is Zariski open, let us further verify that there exists \( A \) such that property (c) holds. We consider

\[
A_0 = \begin{bmatrix}
1 \\
1 & 1 \\
1 & 1 & \ddots & \\
1 & & & & 1 \\
1 & i & & & \\
\vdots & & & & \\
1 & & & & \\
1 & & & & \\
\end{bmatrix} \in \mathbb{C}^{(2d-1)\times d} \quad \text{and} \quad A_0 = \begin{bmatrix} A_0' \\ A_0' \end{bmatrix} \in \mathbb{C}^{m \times d}, \quad (10)
\]

where \( A_0' = [1(m - (2d - 1)) 0] \), and \( e_1(n_1) = [1,1,...,1]^T \in \mathbb{C}^{n_1} \). \( A_0 e_1(d) = (A_0)_{[1]} = 1(m) \) has no zeroes, hence \( e_1(d) \in \mathbb{H}_A \). Suppose \( y \in \mathbb{C}^d \) and \( \text{sign}(A_0 y) = \text{sign}(A_0 e_1(d)) \), hence \( \text{sign}(A_0 y) = \text{sign}(A_0 e_1(d)) = 1(2d-1) \), which reads \( \forall 2 \leq k \leq d, y_1, y_1 + y_k, y_1 + iy_k \in \mathbb{R}^+ \), and hence we know \( y_k, iy_k \in \mathbb{R} \), then \( y_k = 0 \) when \( 2 \leq k \leq d \), then we have \( y \in \mathbb{R}^+ e_1(d) \). Therefore, \( e_1(d) \in \mathcal{W}_A \) follows. Note that we have verified that \( A_0 \) has property (c). Therefore, a non-empty Zariski open set has property (c), i.e., generic \( A \) have property (c). For property (b), we consider a specific \( S \) such that \( 0 < |S| < d \) first. From Lemma 3 we have \( e_1(d - |S|) \in \mathcal{W}_A(S) \cap \mathbb{H}_A(S) \) if and only if \( \text{rank}(D_A(S)(e_1(d - |S|))) \geq 2(d - |S|) + (m - |S|) - 1 \). Note that the entries of \( A(S) \) are rational functions of \( \text{Re}A \) and \( \text{Im}A \), so are those of \( D_A(S)(e_1(d - |S|)) \), hence by using Lemma 4 we know \( A \) with \( e_1(d - |S|) \in \mathcal{W}_A(S) \cap \mathbb{H}_A(S) \) are Zariski open. We can easily verify that there exists \( A \) such that \( e_1(d - |S|) \in \mathcal{W}_A(S) \cap \mathbb{H}_A(S) \), for example, let
$A_S^S[d\setminus|S|] = 0$, which leads to $A(S) = A_S^S[d\setminus|S|]$, due to $A(S) \in \mathbb{C}^{m - |S|} \times (d - |S|)$ and $m - |S| \geq 2d - 1 - |S| > 2(d - |S|) - 1$, we can let $A_S^S[d\setminus|S|]$ have the form of $A_0$ as in (10). Therefore, generic $A$ satisfy $e_1(d - |S|) \in \mathbb{W}_{A(S)} \cap \mathbb{H}_{A(S)}$ for a specific $S$, by finite intersection, we know generic $A$ have property (b). Therefore, $\Xi$ is generic in $\mathbb{C}^{m \times d}$.

It remains to prove that a specific $A \in \Xi$ satisfies (9). When $|S| \geq d$, $\text{ker}(A^S) = \{0\}$, and $0 \in \mathbb{W}_A$, (9) holds trivially. When $S = \emptyset$, from Lemma 3, $\mathbb{W}_A \cap \mathbb{H}_A = \{ x \in \mathbb{C}^d : \text{rank}(\mathbb{D}_A(x)) \geq 2d + m - 1 \}$, note that the entries of $\mathbb{D}_A(x)$ are polynomials of $\text{Re}x$ and $\text{Im}x$, using Lemma 4 again, we know $\mathbb{W}_A \cap \mathbb{H}_A$ are Zariski open in $\mathbb{C}^d$. Combining with property (e), $\mathbb{W}_A \cap \mathbb{H}_A$ is non-empty Zariski open, hence $\mathbb{W}_A \cap \mathbb{H}_A$ is generic in $\mathbb{C}^d$, which implies $\mathbb{W}_A$ is generic in $\mathbb{C}^d$. When $0 < |S| < d$, it’s analogous to prove $\mathbb{W}_{A(S)} \cap \mathbb{H}_{A(S)}$ is Zariski open in $\mathbb{C}^{d - |S|}$, combining with $e_1(d - |S|) \in \mathbb{W}_{A(S)} \cap \mathbb{H}_{A(S)}$, we know $\mathbb{W}_{A(S)} \cap \mathbb{H}_{A(S)}$ is generic in $\mathbb{C}^{d - |S|}$.

Assume $x \in \text{ker}(A^S)$, using Lemma 5, $x \in \mathbb{W}_A \cap \mathbb{H}_A(S)$ if and only if $x[d\setminus|S|] \in \mathbb{W}_{A(S)} \cap \mathbb{H}_{A(S)}$, hence $\mathbb{W}_A \cap \mathbb{H}_A(S)$ is generic in $\text{ker}(A^S)$, then $\mathbb{W}_A \cap \text{ker}(A^S)$ is generic in $\text{ker}(A^S)$. □

**Remark 4.** Let us review existing algorithms and use our Theorem 6 to understand the dependence of their performance on measurement number. In [11, 18], the theoretical results ensure that the Fourier phase of $d - 1$ different frequencies suffice to recover $d$-dimensional real signal. But $2d$ measurements are used in their algorithms. That is because they do not apply the priori that the signal is real in algorithms, while $2d - 1$ measurements are required to recover a $d$-dimensional complex signal. The MagnitudeCut algorithm [23] requires only $2d$ measurements, but it fails for $d$ measurements. These results can be explained by Theorem 6. In [11], the normalized inner product between the desired signal and the recovered signal is used as an indicator for the recovery performance. Both quadratic programming method [14] and MagnitudeCut algorithm [23] have shown in the plots that the inner product values are close to 1 when $2d$ measurements are used. Again it can be validated by using Theorem 6. However, the problem is ill-posed when $m < 2d - 1$. The results in [14] shows that a relatively good recovery can be obtained from quadratic programming with $3d/2$ measurements. Here we provide an explanation and conjecture that the quality of the recovered signal is further enhanced when the number $m$ of measurements increases from $d$ to $2d$. Recall from the proofs of Lemma 4 and Theorem 11, we know the solution space of the quadratic programming in [14] has dimension $\dim(V_x)$ (see Lemma 7), which equals $2d + |N(Ax)| - \text{rank}(\mathbb{D}_A(x))$ (see (3)). In the simulated settings of [14], it is about $2d + m - 2m = 2d - m$. Therefore, when $m$ increases, we are more likely to get a good recovery of $|Ax|$, and then $x$ from the quadratic programming.

## 5 Specific Signal Recovery

In this section, we study the set of recoverable signals $\mathbb{W}_A$ for a specific $A$. Now we would change the perspective and switch to the analysis for a specific signal $x$. We are interested in the set consisting of $A$ that can recover $x$:

$$\mathbb{W}_x(m) = \{ A \in \mathbb{C}^{m \times d} : x \in \mathbb{W}_A \}.$$ 

We also write

$$\mathbb{H}_x(m) = \{ A \in \mathbb{C}^{m \times d} : x \in \mathbb{H}_A \}.$$
Therefore,

\[ \mathbb{H}_x(m)^c = \bigcup_{j=1}^m \{ A \in \mathbb{C}^{m \times d} : \gamma_j^T x = 0 \} \]

The following result shows that for a specific signal, 2d - 2 measurements can hardly work, while 2d - 1 generic measurements can handle.

**Theorem 7.** Suppose \( x \in \mathbb{C}^d \setminus \{0\}, \ d > 1. \) When \( m \leq 2d - 2, \ \mathbb{W}_x(m) \) is nowhere dense; When \( m \geq 2d - 1, \) there are generic \( A \in \mathbb{W}_x(m). \)

**Proof:** Form Lemma 4 \( A \in \mathbb{W}_x(m) \cap \mathbb{H}_x(m) \) if and only if \( x \in \mathbb{W}_A \cap \mathbb{H}_A \) if and only if \( \text{rank}(\mathbb{D}_A(x)) = 2d + m - 1, \) recall that \( \mathbb{D}_A(x) \in \mathbb{C}^{2m \times (2d + m)}. \)

When \( m \leq 2d - 2, \ \text{rank}(\mathbb{D}_A(x)) \leq 2m < 2d + m - 1, \) so we have \( \mathbb{W}_x(m) \cap \mathbb{H}_x(m) = \emptyset, \) hence \( \mathbb{W}_x(m) \subset [\mathbb{H}_x(m)]^c \) is nowhere dense.

When \( m \geq 2d - 1, \) we first show that we only need to consider \( x = e_1. \) For nonzero \( x, \) there exists an invertible \( P \in \mathbb{C}^{d \times d} \) such that \( P e_1 = x, \) i.e., \( P^{-1} x = e_1. \) Since \( x \in \mathbb{W}_{A_{AP^{-1}}} \) if and only if \( P^{-1} x \in \mathbb{W}_A, \) we have \( AP^{-1} \in \mathbb{W}_x(m) \) if and only if \( A \in \mathbb{W}_{e_1}(m). \) Therefore, \( \mathbb{W}_x(m)P = \mathbb{W}_{e_1}(m). \) Then we come to analyse \( \mathbb{W}_{e_1}(m) \) where \( m \geq 2d - 1. \) Actually, we have proved there exist generic \( A \) such that \( e_1 \in \mathbb{W}_A \cap \mathbb{H}_A, \) while implies \( A \in \mathbb{W}_{e_1}(m), \) see the proof of Theorem 4 hence \( \mathbb{W}_{e_1}(m) \) is generic in \( \mathbb{C}^{m \times d}. \) From \( \mathbb{W}_x(m)P = \mathbb{W}_{e_1}(m), \) the result follows. \( \square \)

By applying \( \mathbb{E}_A(x), \) we know that if more than \( 2d - 1 \) measurements recover \( x, \) we can always select \( 2d - 1 \) measurements to recover \( x, \) while there is no such property for \( 2d - 2 \) measurements.

**Theorem 8.** If \( m > 2d - 1, \ x \in \mathbb{W}_A, \) then there exists \( S \subset [m], \ |S| = 2d - 1, \) such that \( x \in \mathbb{W}_{AS}. \) If \( m > 2d - 2, \ x \in \mathbb{W}_A, \) it is likely that for all \( IS \subset [m], \ |S| = 2d - 2, \ x \notin \mathbb{W}_{AS}. \)

**Proof:** Since for invertible \( P \in \mathbb{C}^{d \times d}, \ x \in \mathbb{W}_{AS} \) iff \( P^{-1} x \in \mathbb{W}_{AS}, \) there is no loss of generality if we only consider \( A = \begin{bmatrix} I_d & A_1 \end{bmatrix}. \)

Recall that

\[
\mathbb{E}_A(x) = \begin{bmatrix} [\Psi_{d+1}(x)]_{N(x)} \\ [\Psi_{d+2}(x)]_{N(x)} \\ \vdots \\ [\Psi_{m}(x)]_{N(x)} \end{bmatrix},
\]

from \( x \in \mathbb{W}_A \) we know \( \text{rank}(\mathbb{E}_A(x)) = |N(x)| - 1, \) hence we can find \( S_0 \subset [m - d], |S_0| = |N(x)| - 1, \) s.t. \( \text{rank}([\mathbb{E}_A(x)]^{S_0}) = |N(x)| - 1. \) Define

\[
J = \{ d + 1 \leq j \leq m : \text{at least 1 row of } [\Psi_j(x)]_{N(x)} \text{ appears in } [\mathbb{E}_A(x)]^{S_0} \},
\]

then \( |J| \leq |S_0| = |N(x)| - 1. \)

Consider \( A^{[d \cup J]}, \) we have

\[
|N(x)| - 1 = \text{rank}(\mathbb{E}_A(x)) \geq \text{rank}(\mathbb{E}_{A^{[d \cup J]}}(x)) \geq \text{rank}([\mathbb{E}_A(x)]^{S_0}) = |N(x)| - 1,
\]

thus \( x \in \mathbb{W}_{A^{[d \cup J]}}, \) since \( |d \cup J| = d + |J| \leq d + |N(x)| - 1 \leq 2d - 1, \) we can find \( S \subset [m], \)

such that \( |d \cup J| \subset S \) and \( |S| = 2d - 1, \) then \( x \in \mathbb{W}_{AS}. \)
However, from Theorem 6 we know there exists \( A \in \mathbb{C}^{(2d-1) \times d} \), \( \mathcal{W}_A \) has dense interior while \( \forall S \subset [2d-1], |S| = 2d - 2, \mathcal{W}_{As} \) nowhere dense, which delivers that

\[
\mathcal{W}_A \setminus \bigcup_{|S|=2d-2} \mathcal{W}_{As} \neq \emptyset,
\]

then the result follows immediately. \( \square \)

**Remark 5.** This remark is devoted to present a new perspective of the motivations and results of this paper. That is the study of phase-only system \( \text{sign}(Ax) = b \). For example, assume \( Ax = b \) has solutions, then the solution cannot be unique when \( m = d - 1 \), while when \( m \geq d \) generic \( A \) (which is of full column rank) would guarantee the uniqueness. Note that Theorem 6 says something similar for \( \text{sign}(Ax) = b \), for which such a leap appears from \( m = 2d - 2 \) to \( m = 2d - 1 \). Also, for linear system having unique solution, we can find \( S \subset [m], |S| = d, \) such that \( Ax = b \) has unique solution. It is interesting that Theorem 8 can be viewed as the counterpart for phase-only system.

Theorem 8 gives us an insight of the structure of \( \mathcal{W}_A \) or \( \mathcal{W}_x(m) \), more precisely, when \( m \geq 2d - 1 \), we have

\[
\mathcal{W}_A = \bigcup_{|S|=2d-1} \mathcal{W}_{As}, \\
\mathcal{W}_x(m) = \bigcup_{|S|=2d-1} \{ A : A^S \in \mathcal{W}_x(2d - 1) \}
\]

In addition, the notations of this section provide us a novel way to rephrase \( A \in \mathbb{C}^{m \times d} \) that can recover all signals, which may be of usage for further study of the minimal measurement number \( m(d) \). Obviously, \( \{ A \in \mathbb{C}^{m \times d} : \mathcal{W}_A = \mathbb{C}^d \} = \bigcap_{x \in \mathbb{C}^d} \mathcal{W}_x(m) \). Since for any nonzero \( x \),

\[
\mathcal{W}_0(m) = \{ A \in \mathbb{C}^{m \times d} : \text{rank}(A) = m \} \subset \mathcal{W}_x(m),
\]

and

\[
\forall x \neq 0, \exists P \in \text{GL}_d(\mathbb{C}), \ s.t. x = Pe_1
\]

we can write it as

\[
\bigcap_{P \in \text{GL}_d(\mathbb{C})} \mathcal{W}_{Pe_1}(m),
\]

combining with \( \mathcal{W}_{Pe_1}(m) = \mathcal{W}_{e_1}(m)P^{-1} \), it becomes

\[
\bigcap_{P \in \text{GL}_d(\mathbb{C})} \mathcal{W}_{e_1}(m)P
\]

We can also rephrase Theorem 4, 5 as

\[
\bigcap_{P \in \text{GL}_d(\mathbb{C})} \mathcal{W}_{e_1}(2d - 1)P = \emptyset; \quad \bigcap_{P \in \text{GL}_d(\mathbb{C})} \mathcal{W}_{e_1}(4d - 2)P \neq \emptyset.
\]
6 Affine Linear Measurement Phase

Suppose \( x \in \mathbb{C}^d, [A \ b] \in \mathbb{C}^{m \times (d+1)} \). This section is intended to study the reconstruction of \( x \) from \( \text{sign}(Ax + b) \). This is also a magnitude retrieval problem and we call \( \text{sign}(Ax + b) \) the affine linear measurement phase. This scenario occurs in real-world problems, for example, in the previous situation where some components of signal \( x \) are known as a priori. Though this problem can be viewed as a natural extension of the previous one, there is an essential difference because \( \text{sign}(Ax + b) \) can distinguish \( x \) from any other signals, hence the set of recoverable signals should be written as

\[
\mathcal{W}_{A, b} = \{ x \in \mathbb{C}^d : \text{sign}(Ay + b) = \text{sign}(Ax + b) \implies y = x \}.
\]

Denote the \( j-th \) row of \( A \) by \( \gamma_j^T \), We define

\[
\mathbb{H}_{A, b} = \{ x \in \mathbb{C}^d : \gamma_j^T x + b_j \neq 0, \forall 1 \leq j \leq m \}.
\]

With the setting \( \text{sign}(0) = 0 \), some measurements are actually linear measurements if \( x \notin \mathbb{H}_{A, b} \), hence our framework also embraces the reconstruction from a combination of phase-only measurements and linear measurements. In what follows, we will extend all results to this affine case. Most proofs are parallel to the traditional case, and only the sketch of the proof will be given if the proof is extremely similar.

Let us start with some simple facts. If \( \text{rank}(A) < d \), there exists nonzero \( y_0 \in \text{Ker}(A) \), then for any \( x \) we have \( \text{sign}(Ax + b) = \text{sign}(A(x + y_0) + b) \), which implies \( \mathcal{W}_{A, b} = \emptyset \). If \( \text{rank}(A) = d \) but \( b = Ax_0 \in \mathbb{AC}^d \), we have \( -x_0 \in \mathcal{W}_A \), however, for any \( x \neq -x_0 \), we have \( \text{sign}(A(2x + x_0) + b) = \text{sign}(A(2x + 2x_0)) = \text{sign}(Ax + Ax_0) = \text{sign}(Ax + b) \), note that \( 2x + x_0 \neq x \), hence \( x \notin \mathbb{W}_A \), so we know \( \mathcal{W}_{A, b} = \{ -x_0 \} \) is a singleton. Therefore, we will assume \( \text{rank}(A) = d \) and \( b \notin \mathbb{AC}^d \) beforehand.

For any invertible \( P \in \mathbb{C}^{d \times d}, \tilde{x} \in \mathbb{C}^d \), it’s easy to verify that \( x \in \mathcal{W}_{A, b} \) if and only if \( P^{-1}x + \tilde{x} \in \mathcal{W}_{A P,b - AP\tilde{x}} \), hence \( \mathcal{W}_{A, b} = P[\mathcal{W}_{A P,b - AP\tilde{x}}] \), which implies \( [A \ b] \) and \( [AP \ b - AP\tilde{x}] \) can recover a same amount of signals. Evidently, exchanging rows of \( [A \ b] \) cannot change \( \mathcal{W}_{A, b} \).

Given \( [A \ b] = \begin{bmatrix} A' & b' \\ A'' & b'' \end{bmatrix} \) where \( [A' \ b'] \in \mathbb{C}^{d \times (d+1)} \), by exchanging rows, we can assume \( \text{rank}(A') = d \). Note that \( [A(A')^{-1} \ b - A(A')^{-1}b'] = \begin{bmatrix} I_d \\ A''(A')^{-1} \\ b'' - A''(A')^{-1}b'' \end{bmatrix} \), hence we can always find a measurement matrix of the form \( \begin{bmatrix} I_d \\ 0 \\ A_1 b_1 \end{bmatrix} \) that can recover as many signals as \( [A \ b] \).

6.1 Discriminant Matrices

Given \( [A \ b] \in \mathbb{C}^{m \times (d+1)}, \text{rank}(A) = d, b \notin \mathbb{AC}^d, x \in \mathbb{C}^d \), we define

\[
\mathbb{D}_{A, b}(x) = [\varphi(A) \ \varphi_1(\text{diag}(Ax + b))]
\]

\[
= \begin{bmatrix}
\text{Re}(A) & \text{Im}(A) & \text{Re}((\text{diag}(Ax + b)) \\
-\text{Im}(A) & \text{Re}(A) & -\text{Im}(\text{diag}(Ax + b))
\end{bmatrix}.
\]

We consider the equation

\[
Ay + b = \text{diag}(\text{sign}(Ax + b))N(Ax+b)\Lambda,
\] (11)
where \( y \in \mathbb{C}^d, \Lambda \in \mathbb{R}^{[N(Ax+b)]} \). Note that in what follows we won’t repeat our assumption that \( \text{rank}(A) = d, b \notin AC^d \).

**Lemma 6.** \( x \in \mathbb{W}_{A,b} \) if and only if \( \{ \Lambda \in \mathbb{R}^{[N(Ax+b)]} : (\Pi) \text{ has solutions} \} \) is a 0-dimensional linear submanifold, i.e., a single nonzero point.

**Proof:** For ease of notation, let \( V_x = \{ \Lambda \in \mathbb{R}^{[N(Ax+b)]} : (\Pi) \text{ has solutions} \} \), and \( U_x = \text{diag}(\text{sign}(Ax+b)) \), hence \( (\Pi) \) becomes \( Ay + b = (U_x)\Lambda \). Since \( (U_x)\Lambda \) is a linear submanifold of \( \mathbb{R}^{[N(Ax+b)]} \). Given \( \Lambda_1, \Lambda_2 \in V_x \), there exist \( y_1 \) and \( y_2 \) such that \( Ay_1 + b = (U_x)\Lambda_1, Ay_2 + b = (U_x)\Lambda_2 \). Then we have \( t(Ay_1 + b) + (1-t)(Ay_2 + b) = t(U_x)\Lambda_1 + (1-t)(U_x)\Lambda_2 \), which implies \( t\Lambda_1 + (1-t)\Lambda_2 \in V_x \). Therefore, \( V_x \) is a linear submanifold.

For the "if" part, assume \( \text{sign}(Ax' + b) = \text{sign}(Ax + b) \), then \( |Ax' + b| = |Ax + b| \), then \( Ax' + b = Ax + b \), by using \( \text{rank}(A) = d \), we have \( x' = x \), hence \( x \in \mathbb{W}_{A,b} \). For the "only if" part, let’s assume \( x \notin \mathbb{W}_{A,b} \) and then rule out the probability that \( \text{dim}(V_x) \geq 1 \). Note that \( |Ax + b| = |Ax + b| \), \( |Ax' + b| = |Ax + b| \), \( Ax' + b = Ax + b \), \( \text{rank}(A) = d \), \( x' = x \), hence the result follows.

**Theorem 9.** \( x \in \mathbb{W}_{A,b} \) if and only if \( \text{rank}(\mathbb{D}_{A,b}(x)) = 2d + |N(Ax+b)| \).

**A sketch of proof:** Continue to use notations introduced in proof of Lemma 6. Use \( \varphi_1 \) to convert \( (\Pi) \) to a real linear system, which reads \( \varphi(A)\varphi_1(y) = \varphi((U_x)N(Ax+b))\Lambda = -\varphi_1(b) \). Then we have \( \text{dim}(V_x) = 2d + |N(Ax+b)| - \text{rank}([\varphi(A)\varphi_1((U_x)N(Ax+b))]) = 2d + |N(Ax+b)| - \text{rank}(\mathbb{D}_{A,b}(x)) \), hence the result follows.

**Theorem 10.** \( x \in \mathbb{W}_{A,b} \) if and only if \( \text{rank}(\text{Re}(\text{diag}(Ax+b)^*(I-AA^\dagger)\text{diag}(Ax+b))) \) and \( \text{Im}(\text{diag}(Ax+b))) = |N(Ax+b)| \).

**A sketch of proof:** Continue to use notations used above. By using Moore-Penrose inverse \( A^\dagger \), we can easily verify \( \text{dim}(V_x) = 0 \) if and only if \( (AA^\dagger - I)(U_x)N(Ax+b)\Lambda = (AA^\dagger - I)b \) has unique solution in \( \mathbb{R}^{[N(Ax+b))] \}. By separating real part and imaginary part, it’s equivalent to \( \text{rank}\left(\begin{bmatrix} \text{Re}(AA^\dagger - I)\Lambda, & \text{Im}(AA^\dagger - I)\Lambda \end{bmatrix}ight) = |N(Ax+b)| \), by combining with the observation \( \text{rank}\left(\begin{bmatrix} \text{Re}\left(\begin{bmatrix} \text{Re}(AA^\dagger - I)\Lambda \text{Im}(AA^\dagger - I)\Lambda \end{bmatrix}ight) \right) = \text{rank}\left(\begin{bmatrix} \text{Re}(AA^\dagger - I)\Lambda \text{Im}(AA^\dagger - I)\Lambda \end{bmatrix} \right) \right) \) and Lemma 2 we complete the proof.

Consider \( [A b] = \begin{bmatrix} I_d & 0 \\ A_1 & b_1 \end{bmatrix} \) and denote the entry of \( [A_1 b_1] \) by \( r_j,k e^{i\theta_j,k} \), \( d + 1 \leq j \leq m, 1 \leq k \leq d + 1, r_j,k \geq 0 \). We also write the signal in a polar form \( x = [|x_k|e^{i\alpha_k}]^T_{1 \leq k \leq d} \). For \( d + 1 \leq j \leq m \), if \( \gamma_j^T x + b_j = 0 \), assume \( \text{sign}(\gamma_j^T x + b_j) = e^{i\delta_j} \), then write \( \Psi_j(x) = [r_j,1\sin(\theta_j,1 + \alpha_1 - \delta_j), r_j,2\sin(\theta_j,2 + \alpha_2 - \delta_j), \ldots, r_j,d\sin(\theta_j,d + \alpha_d - \delta_j)] \);

Otherwise, \( \gamma_j^T x + b_j = 0 \), we write \( \Psi_j(x) = [r_j,1\sin(\theta_j,1 + \alpha_1), r_j,2\sin(\theta_j,2 + \alpha_2), \ldots, r_j,d\sin(\theta_j,d + \alpha_d)] \).

19
We define

$$\mathbb{E}^0_{A,b}(x) = \begin{bmatrix} \Psi_{d+1}(x) \\ \Psi_{d+2}(x) \\ \vdots \\ \Psi_m(x) \end{bmatrix}; \mathbb{E}_{A,b}(x) = (\mathbb{E}^0_{A}(x))_{N(x)}.$$  

**Theorem 11.** Suppose \( x \neq 0 \), \([A b] = \begin{bmatrix} I_d & 0 \\ A_1 & b_1 \end{bmatrix} \in \mathbb{C}^{m \times (d+1)}\), then \( x \in \mathbb{W}_{A,b} = |N(Ax + b)|\) if and only if \( \text{rank}(\mathbb{E}_{A,b}(x)) = |N(x)|\).

A sketch of proof: Let \( J_1 = \{d + 1 \leq j \leq m : \gamma_j^T x + b_j \neq 0\}\), \( J_2 = \{d + 1 \leq j \leq m : \gamma_j^T x + b_j = 0\}\). For \( j \in J_1\), define \( b_j(x) = [r_{j,d+1} \sin(\theta_{j,d+1} - \delta_j)]\), for \( j \in J_1\), define \( b_j(x) = [r_{j,d+1} \cos(\theta_{j,d+1})]\), then we let \( b(x) = [b_{d+1}(x)^T, b_{d+2}(x)^T, \ldots, b_m(x)^T]^T\). By simple computation, we have \( \text{sign}(Ay + b) = \text{sign}(Ax + b)\) if and only if

$$ \begin{cases} \text{sign}(y) = \text{sign}(x) \\ \mathbb{E}_{A,b}(x)|y|^{N(x)} + b(x) = 0 \\ \sum_{k \in N(x)} r_{j,k} \cos(\theta_{j,k} + \alpha_k - \delta_j)|y_k| + r_{j,d+1} \cos(\theta_{j,d+1} - \delta_j) > 0, \forall j \in J_1 \end{cases}$$

Let \( y = x\), we have \( \mathbb{E}_{A,b}(x)|x|^{N(x)} + b(x) = 0\) and \( \sum_{k \in N(x)} r_{j,k} \cos(\theta_{j,k} + \alpha_k - \delta_j)|x_k| + r_{j,d+1} \cos(\theta_{j,d+1} - \delta_j) > 0, \forall j \in J_1\).

The ”if” part holds directly. For the ”only if” part, we assume \( \text{rank}(\mathbb{E}_{A,b}(x)) < |N(x)|\), by which we can find \( Y \in \mathbb{R}^{[N(x)]}, Y \neq |x|^{N(x)}\), such that \( \mathbb{E}_{A,b}(x)Y + b(x) = 0\), and we construct \( y^0\) such that \( \text{sign}(y^0) = \text{sign}(x)\), and \( |y^0|^{N(x)} = Y\). We can further make \( Y\) sufficiently close to \( |x|^{N(x)}\) to ensure \( \sum_{k \in N(x)} r_{j,k} \cos(\theta_{j,k} + \alpha_k - \delta_j)|y_k^0| + r_{j,d+1} \cos(\theta_{j,d+1} - \delta_j) > 0, \forall j \in J_1\). Therefore, \( \text{sign}(Ay^0 + b) = \text{sign}(Ax + b)\), while evidently \( y^0 \neq x\), this contradicts \( x \in \mathbb{W}_{A,b}\). \( \square \)

**Remark 6.** For general comparisons of several discriminant matrices, see Remark \( \square \). For affine linear measurement phase, however, we cannot formulate the problem as a quadratic programming by using the positive semi-definite matrix \( \text{Re}(U_x^*(I - AA^*)U_x)\), since the desired \( |Ax + b|^{N(Ax+b)}\) cannot set the corresponding quadratic form to be vanished, see Theorem \( \square \). Here we suggest we convert this problem to a traditional one. Note that \( \text{sign}(Ax + b) = \text{sign}([A b] \begin{bmatrix} x \\ 1 \end{bmatrix})\), we use \([A b]\) and the information \( \text{sign}(Ax + b)\) to get \( \hat{x} \in \mathbb{C}^{d+1}\) from existing algorithms. Finally we rescale \( \hat{x}\) to make its last entry become 1, then \( \hat{x}^{[d]}\) can serve as an estimation of origin signal.

### 6.2 Signal Recovery

We are interested in the minimal measurement number

$$m'(d) = \min\{m \in \mathbb{N}_+ : \exists [A b] \in \mathbb{C}^{m \times (d+1)}, s.t. \mathbb{W}_{A,b} = \mathbb{C}^d\}.$$ 

By using \( \mathbb{E}_{A,b}(x)\), we derive \( 2d + 1 \) as a lower bound.
Theorem 12. For $d \in \mathbb{N}_+$, $m'(d) \geq 2d + 1$.

A sketch of proof: Let $\alpha_k = \theta_{d+1,d+1} - \theta_{d+1,k}$ when $1 \leq k \leq d$, and choose $\lambda_1, \lambda_2, \ldots, \lambda_d > 0$ such that $x = [\lambda_1 e^{(\theta_{d+1,d+1}-\theta_{d+1,1})}, \ldots, \lambda_d e^{(\theta_{d+1,d+1}-\theta_{d+1,d})}]^T \in \mathbb{H}_{A,b}$. We can verify that $\delta_{d+1} = \theta_{d+1,d+1}$ and hence the first row of $D_{A,b}(x)$ equals zero, by using Theorem 11, we have $m - d - 1 \geq d$, hence $m \geq 2d + 1$. \hfill \Box

We can follow the method in Theorem 5, which brings the the upper bound $4d + 1$. However, thanks to the appearance of $b$, we can easily construct an $[A \ b]$ that recover all signals from $3d$ measurements.

Theorem 13. For $d \in \mathbb{N}_+$, $m'(d) \leq 3d$.

Proof: First prove the mapping $\sigma :$

$$
\mathbb{C} \longrightarrow \mathbb{C}^3
$$

$$
x \longmapsto \begin{bmatrix}
\text{sign}(x) \\
\text{sign}(x + 1) \\
\text{sign}(x + i)
\end{bmatrix}
$$

is injective. Assume $y \in \mathbb{C}$ such that $[\text{sign}(y) \text{sign}(y + 1) \text{sign}(y + i)]^T = [\text{sign}(x) \text{sign}(x + 1) \text{sign}(x + i)]^T$, if $[\text{sign}(x) \text{sign}(x + 1) \text{sign}(x + i)]^T$ contains zeroes, it is trivial that $y = x$, so we assume $x, x + 1, x + i \neq 0$. When $\text{sign}(y) = \text{sign}(x) \neq \pm 1$, we have

$$
\frac{x + 1}{\text{sign}(x + 1)} > 0 \text{ i.e., } \frac{|x|\text{sign}(x) + 1}{\text{sign}(x + 1)} > 0,
$$

it implies

$$
\text{Im} \left( \frac{\text{sign}(x)}{\text{sign}(x + 1)} \right) |x| + \text{Im} \left( \frac{1}{\text{sign}(x + 1)} \right) = 0.
$$

It can be easily verify that $\text{Im} \left( \frac{\text{sign}(x)}{\text{sign}(x + 1)} \right) \neq 0$, hence $|x|$ is uniquely determined, since $\text{sign}(y + 1) = \text{sign}(x + 1)$, $|y| = |x|$, so $y = x$. When $\text{sign}(y) = \text{sign}(x) = \pm 1$, by using $\text{sign}(x + i)$, we can prove $y = x$ similarly.

Use $1$ to denote $[1, 1, \ldots, 1]^T \in \mathbb{C}^d$, we consider

$$
[A \ b] = \begin{bmatrix}
I_d & 0 \\
I_d & 1 \\
I_d & i1
\end{bmatrix} \in \mathbb{C}^{3d \times (d+1)}.
$$

From the fact that $\sigma$ is injective, we know

$$
\mathbb{C}^{3d} \longrightarrow \mathbb{C}^{3d}
$$

$$
x \longmapsto \text{sign}(Ax + b)
$$

is injective, thus $\mathbb{W}_{A,b} = \mathbb{C}^d$. \hfill \Box
6.3 Generic Signal Recovery

This part is devoted to demonstrate $2d$ generic measurements can recover a generic set of signals, and $2d$ is minimal in the sense that $2d - 1$ can hardly handle any signal (nowhere dense). We use $\text{Ker}(A, b^1)$ to denote $\{x : Ax + b^1 = 0\}$. Suppose $S \subset [m]$, use $S^c$ to denote $[m] \setminus S$, we define

$$H_{A,b}(S) = \{x \in \mathbb{C}^d : \gamma_j^T x + b_j = 0, \forall j \in S; \gamma_j^T x + b_j \neq 0, \forall j \in S^c\},$$

i.e.,

$$H_{A,b}(S) = \text{Ker}(A^S, b^S) \cap H_{A^S, b^S}.$$

Note that we have the following partition

$$\mathbb{C}^d = \bigcup_{S \subset [m]} H_{A,b}(S), \text{ where } H_{A,b}(S) \cap H_{A,b}(T) = \emptyset, \forall S \neq T.$$

For ease to present, let us introduce two Lemma\textsuperscript{7,8}. The former characterizes $\mathbb{W}_{A,b} \cap H_{A,b}$, and the latter allows us to convert $x \notin H_{A,b}$ to the situation where no measurements equal zero.

**Lemma 7.** $x \in \mathbb{W}_{A,b} \cap H_{A,b}$ if and only if $\text{rank}(D_{A,b}(x)) \geq 2d + m$.

*Proof:* The "only if" part comes from Theorem\textsuperscript{9} directly. For the "if" part, recall that we have $\text{rank}(D_{A,b}(x)) \leq \text{rank}(\varphi(A)) + \text{rank}(\varphi_1((\text{diag}(Ax + b))) \leq 2d + |N(Ax + b)|$, combining with $\text{rank}(D_{A,b}(x)) \geq 2d + m$, we have $|N(Ax + b)| \geq m$, which implies $x \in H_{A,b}$, hence $|N(Ax + b)| = m$. Using Theorem\textsuperscript{9} again, the result follows. \hfill $\Box$

**Lemma 8.** Assume $S \subset [m]$, $1 \leq |S| < d$, and $\text{rank}(A_{[S]}^S) = |S|$, $x \in \text{Ker}(A^S, b^S)$. Define

$$A'(S) = A_{[d]\setminus[S]}^S - A_{[d]\setminus[S]}^S A_{[S]}^S (A_{[d]\setminus[S]}^S)^{-1} A_{[d]\setminus[S]}^S \in \mathbb{C}^{(m-|S|) \times (d-|S|)};$$

$$b'(S) = b^S - A_{[S]}^S (A_{[d]\setminus[S]}^S)^{-1} b^S \in \mathbb{C}^{m-|S|}.$$

Then we have $x \in \mathbb{W}_{A,b} \cap H_{A,b}(S)$ if and only if $x_{[d]\setminus[S]} \in \mathbb{W}_{A'(S), b'(S)} \cap H_{A'(S), b'(S)}$.

*Proof:* We note that for $x \in \text{Ker}(A^S, b^S)$, we have $A^S x + b^S = A'(S) x_{[d]\setminus[S]} + b'(S)$. Then the remaining part is direct verification following the basic routine, which is analogous to Lemma\textsuperscript{5} \hfill $\Box$

**Theorem 14.** Suppose $[A b] \in \mathbb{C}^{m \times (d+1)}$. When $m \leq 2d - 1$, $\mathbb{W}_{A,b}$ is nowhere dense; When $m \geq 2d$, there are generic $[A b]$ such that

$$\forall S \subset [m], \text{ with non-empty } \text{Ker}(A^S, b^S), \mathbb{W}_{A,b} \cap \text{Ker}(A^S, b^S) \text{ is generic in } \text{Ker}(A^S, b^S).$$

Specifically, let $S = \emptyset$, we have $\mathbb{W}_{A,b}$ is generic in $\mathbb{C}^d$.

*Proof:* Here we give a sketch of proof. The idea is similar to the proof of Theorem 6. When $m \leq 2d - 1$, from Lemma\textsuperscript{7} we have $\mathbb{W}_{A,b} \cap H_{A,b} = \emptyset$, hence $\mathbb{W}_{A,b} \subset (H_{A,b})^c$ is nowhere dense. When $m \geq 2d$, we use $0(n_1)$ to denote $[0, 0, ..., 0]^T \in \mathbb{C}^{n_1}$, $1(n_1)$ to denote $[1, 1, ..., 1]^T \in \mathbb{C}^{n_1}$, and we propose Property (a)(b)(c) as follows:

- **Property (a):** $\forall 0 < |S| \leq d$, $\text{rank}(A_{[S]}^S) = |S|$. 

22
Based on having Property (a), we further define

Property (b): ∀0 < |S| < d, \(0(D - |S|) \in W_{A^t(S),b^t(S)} \cap H_{A^t(S),b^t(S)} \), where \(A^t(S), b^t(S) \) follow the definitions in Lemma 5.

Property (c): \(0(d) \in W_{A,b} \cap H_{A,b} \).

We consider \(\Xi = \{[A b] \in C^{m \times (d+1)} : [A b] \) has property (a)(b)(c)\}. We first come to prove \(\Xi \) is generic in \(C^{m \times (d+1)} \). Evidently, generic \([A b] \) have property (a). Combining Lemma 7 and 4 we know the set of \([A b] \) having property (c) is Zariski open. Consider \([A_0 b_0] = \begin{bmatrix} I_d & 1(d) \\
 iI_d & 1(d) \\
 0 & 1(m - 2d) \end{bmatrix} \), it’s easy to confirm \([A_0 b_0] \) have property (c). Therefore, generic \([A b] \) have property (c). For a specific \(S \subset [m] \) and \(0 < |S| < d \), using Lemma 7 and 4 it holds similarly that \(\{[A b] : 0(D - |S|) \in W_{A^t(S),b^t(S)} \cap H_{A^t(S),b^t(S)} \} \) is Zariski open. And it’s non-empty, for example, we let \(A^S \neq 0 \) and \(b^S = 0 \), then \(A^t(S) = A^S \), \(b^t(S) = b^S \), note that \(m - |S| \geq 2d - |S| > 2(d - |S|) \), therefore, we can construct \(A^S \) and \(b^S \) just like \([A_0 b_0] \) mentioned above.

Then we prove a specific \(A \in \Xi \) would satisfy (12). Note that \(|S| \geq d \) would be trivial case. And for \(|S| < d \), using Lemma 7 and 4 we know \(W_{A^t(S),b^t(S)} \cap H_{A^t(S),b^t(S)} \) is generic in \(C^{d - |S|} \). Then from Lemma 5 we have \(W_{A,b} \cap H_{A,b}(S) \) is generic in \(Ker(\mathcal{A}^2, b^S) \), the result follows. □

6.4 Specific Signal Recovery

For a specific \(x \in C^d \), we define the set of \([A b] \) that can recover \(x \) by \(W_x(m) = \{[A b] \in C^{m \times (d+1)} : x \in W_{A,b} \} \), and the set of \([A b] \) such that no measurement equals zero by \(\mathbb{H}_x(m) = \{[A b] \in C^{m \times (d+1)} : x \in W_{A,b} \} \). The following result shows that \(2d \) measurements are required for recovering a specific \(x \).

Theorem 15. When \(m \leq 2d - 1 \), \(W_x(m) \) is nowhere dense; When \(m \geq 2d \), there are generic \([A b] \in W_x(m) \).

Proof: Note that \([A b] \in W_x(m) \cap \mathbb{H}_x(m) \) if and only if \(x \in W_{A,b} \cap H_{A,b} \), while for \(m < 2d \) we have \(\text{rank}(\mathcal{D}_{A,b}(x)) \leq 2m < m + 2d \), hence from Lemma 7 we know \(W_{A,b} \cap H_{A,b} = \emptyset \). Therefore, \(W_x(m) \cap \mathbb{H}_x(m) = \emptyset \), hence \(W_x(m) \subset \mathbb{H}_x(m) \) is nowhere dense. Note that \(x \in W_{A,b} \) if and only if \(0 \in W_{A,b} \), hence \([A b] \in W_x(m) \) if and only if \([A b] \begin{bmatrix} I_d & x \\
 0 & 1 \end{bmatrix} \in W_x(m) \).

So we have

\[ W_x(m) = W_0(m) \begin{bmatrix} I_d & -x \\
 0 & 1 \end{bmatrix} \] (13)

while it has been shown in Theorem 14 that \(W_0(m) \) generic when \(m \geq 2d \). □

Theorem 16. If \(m > 2d \), \(x \in W_{A,b} \), then there exists \(S \subset [m], |S| = 2d \), s.t. \(x \in W_{A,b^S} \). If \(m > 2d - 1 \), \(x \in W_{A,b} \), it is likely that \(\forall S \subset [m], |S| = 2d - 1 \), \(x \notin W_{A,b^S} \).

Proof: Without loss of generality, we consider \([A b] = \begin{bmatrix} I_d & 0 \\
 A_1 & b_1 \end{bmatrix} \). Recall that

\[ E_{A,b}(x) = \begin{bmatrix} \Psi_{d+1}(x)_{N(x)} \\
 \Psi_{d+2}(x)_{N(x)} \\
 \vdots \\
 \Psi_m(x)_{N(x)} \end{bmatrix} \]
since \( x \in \mathbb{W}_{A,b} \) we know \( \text{rank}(\mathbb{E}_{A,b}(x)) = |N(x)| \), hence we can find \( S_0 \subset [m-d] \), such that \( |S_0| = |N(x)| \) and \( \text{rank}(\mathbb{E}_{A,b}(x)|^{S_0}) = |N(x)| \). Define \( J = \{d + 1 \leq j \leq m : \) at least 1 row appears in \((\mathbb{E}_{A,b}(x))_{S_0}\)\}. Then we can prove \( \text{rank}(\mathbb{E}_{A[d \cup J],b[d \cup J]}(x)) = |N(x)| \), which implies \( x \in \mathbb{W}_{A[d \cup J],b[d \cup J]} \). Since \( |d \cup S_0| \leq d + |S_0| \leq 2d \), we can find \( S \subset [m] \), \( |S| = 2d \) such that \( x \in \mathbb{W}_{A,S,b} \). For the latter part, from Theorem \([14]\) there exists \([A\ b] \in \mathbb{C}^{2d \times d} \), such that \( \mathbb{W}_{A,b} \) generic while \( \mathbb{W}_{A,S,b} \) where \( |S| = 2d - 1 \) is nowhere dense. \( \square \)

From Theorem \([16]\) we know that when \( m \geq 2d \)

\[
\mathbb{W}_{A,b} = \bigcup_{|S| = 2d} \mathbb{W}_{A,S,b}^s,
\]

\[
\mathbb{W}'_x(m) = \bigcup_{|S| = 2d} \{[A\ b] : [A^S\ b^S] \in \mathbb{W}'_x(2d)\}.
\]

Using the notation \( \mathbb{W}'_x(m) \), we can determine \([A\ b] \in \mathbb{C}^{m \times (d+1)} \) s.t. \( \mathbb{W}_{A,b} = \mathbb{C}^d \) by

\[
\bigcap_{x \in \mathbb{C}^d} \mathbb{W}'_x(m).
\]

Recall \([13]\), we can rephrase \([A\ b] \in \mathbb{C}^{m \times (d+1)} \) that can recover \( \mathbb{C}^d \) by

\[
\bigcap_{x \in \mathbb{C}^d} \mathbb{W}'_0(m) \left[ \begin{array}{cc} I_d & x \\ 0 & 1 \end{array} \right].
\]

And we can restate Theorem \([11]\) and \([10]\) as follows:

\[
\bigcap_{x \in \mathbb{C}^d} \mathbb{W}'_0(2d) \left[ \begin{array}{cc} I_d & x \\ 0 & 1 \end{array} \right] = \emptyset; \bigcap_{x \in \mathbb{C}^d} \mathbb{W}'_0(3d) \left[ \begin{array}{cc} I_d & x \\ 0 & 1 \end{array} \right] \neq \emptyset.
\]

### 7 Conclusion

In this paper, we build a theoretical framework for magnitude retrieval from \( \text{sign}(Ax) \) or \( \text{sign}(Ax + b) \). We remove several restrictions in traditional phase-only signal reconstruction (e.g. \([10,11,15]\)) by considering general matrices \( A \) and complex vectors \( x \). We also explore the minimal measurement number in magnitude retrieval and obtain some results. We demonstrate that in most situations, \( \text{sign}(Ax) \) where \( m = 2d - 1 \) suffices to recover \( x \) up to a positive number, and \( \text{sign}(Ax + b) \) where \( m = 2d \) suffices to recover \( x \) exactly.

Most existing papers of phase-only reconstruction focus on algorithms or its applications. Since this paper encompass the situations discussed in their settings, it manages to fill the theoretical vacancy. Our general framework can understand previous works. For example, we use \( \mathbb{E}_{A}(x) \) to demonstrate why symmetric signals may fail the reconstruction from Fourier phase, see Remark \([2]\). Besides, almost all experimental results in \([14,23]\) seem natural from the viewpoint of our theorems, see Remark \([4]\). We mention that thanks to our setting of \( \text{sign}(0) = 0 \), our framework also embraces the situation where both phase-only measurements and linear measurements are obtainable. Beyond that, our results are actually some properties of phase-only system, which is somewhat analogous to linear system, see Remark \([5]\).
References

[1] H. C. Andrews and B. R. Hunt. Digital Image Restoration. 1977.

[2] R. Balan, P. Casazza, and D. Edidin. On signal reconstruction without phase. Applied and Computational Harmonic Analysis, 20(3):345–356, 2006.

[3] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson. Saving phase: Injectivity and stability for phase retrieval. Applied and Computational Harmonic Analysis, 37(1):106–125, 2014.

[4] I. Bartolini, P. Ciaccia, and M. Patella. Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(1):142–147, 2005.

[5] J. Behar, M. Porat, and Y. Y. Zeevi. Image reconstruction from localized phase. IEEE Transactions on Signal Processing, 40(4):736–743, 1992.

[6] B. G. Bodmann and N. Hammen. Stable phase retrieval with low-redundancy frames. Advances in Computational Mathematics, 41(2):317–331, 2015.

[7] P. T. Boufounos. Sparse signal reconstruction from phase-only measurements. In Proc. Int. Conf. Sampling Theory and Applications, 2013.

[8] A. Conca, D. Edidin, M. Hering, and C. Vinzant. An algebraic characterization of injectivity in phase retrieval. Applied and Computational Harmonic Analysis, 38(2):346–356, 2015.

[9] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. Journal of the ACM, 42(6):1115–1145, 1995.

[10] M. Hayes. The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform. IEEE Transactions on Acoustics, Speech, and Signal Processing, 30(2):140–154, 1982.

[11] M. Hayes, J. Lim, and A. Oppenheim. Signal reconstruction from phase or magnitude. IEEE Transactions on Acoustics, Speech, and Signal Processing, 28(6):672–680, 1980.

[12] G. Hua and M. T. Orchard. Image inpainting based on geometrical modeling of complexwavelet coefficients. In 2007 IEEE International Conference on Image Processing, volume 1, pages I–553. IEEE, 2007.

[13] G. Hua and M. T. Orchard. Image reconstruction from the phase or magnitude of its complex wavelet transform. In 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, pages 3261–3264. IEEE, 2008.

[14] V. Kishore, S. Mukherjee, and C. S. Seelamantula. Phasesense—signal reconstruction from phase-only measurements via quadratic programming. In 2020 International Conference on Signal Processing and Communications (SPCOM), pages 1–5. IEEE, 2020.
[15] A. Levi and H. Stark. Signal restoration from phase by projections onto convex sets. *JOSA*, 73(6):810–822, 1983.

[16] E. Loveimi and S. M. Ahadi. Objective evaluation of magnitude and phase only spectrum-based reconstruction of the speech signal. In *2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 1–4. IEEE, 2010.

[17] O. Matoba, T. J. Naughton, Y. Frauel, N. Bertaux, and B. Javidi. Real-time three-dimensional object reconstruction by use of a phase-encoded digital hologram. *Applied optics*, 41(29):6187–6192, 2002.

[18] A. V. Oppenheim, M. H. Hayes, and J. S. Lim. Iterative procedures for signal reconstruction from phase. In *1980 Intl Optical Computing Conf I*, volume 231, pages 121–129. International Society for Optics and Photonics, 1980.

[19] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, 1981.

[20] Z. Shao, H. Shu, J. Wu, Z. Dong, G. Coatrieux, and J. L. Coatrieux. Double color image encryption using iterative phase retrieval algorithm in quaternion gyrator domain. *Optics Express*, 22(5):4932–4943, 2014.

[21] S. Urieli, M. Porat, and N. Cohen. Optimal reconstruction of images from localized phase. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 7(6):838–853, 1998.

[22] C. Vinzant. A small frame and a certificate of its injectivity. In *2015 International Conference on Sampling Theory and Applications*, pages 197–200. IEEE, 2015.

[23] J. Wu, J. Liu, Y. Kong, X. Han, L. Senhadji, and H. Shu. Phase-only signal reconstruction by magnitudecut. *arXiv preprint arXiv:1603.00210*, 2016.