The Fisher Information Matrix and the CRLB in a Non-AWGN Model for the Phase Retrieval Problem

Radu Balan
Department of Mathematics and Center for Scientific Computation and Mathematical Modeling
University of Maryland, College Park, Md 20742
Email: rvbalan@math.umd.edu http://www.math.umd.edu/ rvbalan/

***Abstract*—In this paper we derive the Fisher information matrix and the Cramer-Rao lower bound for the non-additive white Gaussian noise model $y_k = |\langle x, f_k\rangle + \mu_k|^2$, $1 \leq k \leq m$, where $\{f_1, \cdots, f_m\}$ is a spanning set for $\mathbb{C}^n$ and $(\mu_1, \cdots, \mu_m)$ are i.i.d. realizations of the Gaussian complex process $\mathbb{C}\mathcal{N}(0, \rho^2)$. We obtain closed form expressions that include quadrature integration of elementary functions.**

I. Introduction

The *phaseless reconstruction* problem (also known as the *phase retrieval* problem) has gained a lot of attention recently. The problem is connected with several topics in mathematics and has applications in many areas of science and engineering. Consider a n-dimensional complex Hilbert space $H = \mathbb{C}^n$ endowed with a sesquilinear scalar product $\langle x, y\rangle$ (e.g. $\sum_{k=1}^n x_k\overline{y_k}$) that is linear in x and antilinear in y. Let $\mathcal{F} = \{f_1, \cdots, f_m\}$ be a spanning set for H. Since H is finite dimensional, $\mathcal{F}$ is also *frame* that is there are constants $0 < A \leq B < \infty$ called *frame bounds* such that for every $x \in H$,

$$A\|x\|^2 \leq \sum_{k=1}^m |\langle x, f_k\rangle|^2 \leq B\|x\|^2.$$

Consider the following nonlinear map

$$\alpha : H \to \mathbb{R}^m \quad , \quad \alpha(x) = (|\langle x, f_k\rangle|)_{1\leq k\leq m}\,.$$

Note $\alpha(e^{i\varphi}x) = \alpha(x)$ for any real φ. This suggests to replace H by the quotient space $\hat{H} = H/\sim$ where for $x, y \in H$, $x \sim y$ if and only if there is a unimodular scalar $z \in \mathbb{C}$, $|z| = 1$, so that $x = zy$. The elements of $\hat{H}$ are called *rays* in quantum mechanics. The nonlinear map α factors through the projection $H \searrow \hat{H}$ to a well-defined map also denoted by α that acts on $\hat{H}$ via

$$\alpha : \hat{H} \to \mathbb{R}^m \quad , \quad \alpha(\hat{x}) = (|\langle x, f_k\rangle|)_{1\leq k\leq m} \quad , \ \forall x \in \hat{x}.$$

The phaseless reconstruction problem refers to analysis of the nonlinear map α. By definition, we call $\mathcal{F}$ a *phase retrievable frame* if α is injective. There has been recent progress on the problems of injectivity, bi-Lipschitz continuity, and inversion algorithms ([3], [4], [10], [2], [6], [8], [9]). This paper refers to establishing information-theoretic performance bounds for the reconstruction problem. Consider the general measurement process:

$$y_k^{(p)} = |\langle x, f_k\rangle + \mu_k|^p + \nu_k \quad , \quad 1 \leq k \leq m \tag{1.1}$$

978-1-4673-7353-1/15/$31.00 ©2015 IEEE

where the noise variables $(\mu_1, \cdots, \mu_m, \nu_1, \cdots, \nu_m)$ are random variables with known statistics, and p is a known measurement parameter. Typically $p = 1$ or $p = 2$. In [5], [6], [8] the authors obtained the Fisher information matrix for the measurement process with additive white Gaussian noise (AWGN):

$$y_k^{(2)} = |\langle x, f_k\rangle|^2 + \nu_k \quad , \quad 1 \leq k \leq m \tag{1.2}$$

where $(\nu_1, \cdots, \nu_m)$ are i.i.d. $\mathcal{N}(0, \sigma^2)$. In the real case the Fisher information matrix has the form

$$\mathbb{I}^{AWGN,real}(x) = \frac{4}{\sigma^2}\sum_{k=1}^m |\langle x, f_k\rangle|^2 f_k f_k^T. \tag{1.3}$$

In the complex case, the Fisher information matrix takes the form:

$$\mathbb{I}^{AWGN,cplx}(x) = \frac{4}{\sigma^2}\sum_{k=1}^m \Phi_k\xi\xi^*\Phi_k \tag{1.4}$$

where $\xi = \mathbf{j}(x)$ and Φ_k are constructed from x and f_k from the realification process as described in the next section, see (2.9,2.11). In this paper we consider a non-additive white Gaussian noise model, namely the case with $\mu_k \neq 0$ and $\nu_k = 0$. We derive the Fisher information matrix and the Cramer-Rao Lower Bound (CRLB) for the case $p = 2$ but we show the bounds we obtain can easily be applied to other values of p. The noise model considered here is directly applicable to the case of noise reduction from measurements of the Short-Time Spectral Amplitude (STSA). For instance see [12] for a MMSE estimator of STSA that uses linear reconstruction of the signal x.

II. Fisher Information Matrix

Consider the measurement model:

$$y_k = |\langle x, f_k\rangle + \mu_k|^2 \quad , \quad 1 \leq k \leq m \tag{2.5}$$

where $\mathcal{F} = \{f_1, \cdots, f_m\}$ is a frame with bounds A, B for $\mathbb{C}^n$ and $(\mu_1, \cdots, \mu_m)$ are independent and identically distributed complex random variables with distribution $\mathbb{C}\mathcal{N}(0, \rho^2)$. Specifically the last statement means that the real parts and imaginary parts of the complex random variables $\mu_1, \cdots, \mu_m$ are i.i.d. with distribution $\mathcal{N}(0, \frac{\rho^2}{2})$.

We denote by $n(t;a,b)$ the probability density function of a Gaussian random variable T with mean a and variance b. Thus

$$n(t;a,b) = \frac{1}{\sqrt{2\pi b}} e^{-\frac{1}{2b}(t-a)^2}.$$

First we derive the likelihood function. Let δ_k be the phases from $\langle x, f_k\rangle = |\langle x, f_k\rangle| e^{-i\delta_k}$. Note that $e^{-i\delta_k}\mu_k$ has the same distribution $\mathbb{C}\mathcal{N}(0,\rho^2)$ as μ_k. Furthermore $(e^{-i\delta_1}\mu_1, \cdots, e^{-i\delta_m}\mu_m)$ remain independent and therefore identically distributed with distribution $\mathbb{C}\mathcal{N}(0,\rho^2)$. Let $u_k + iv_k = e^{-i\delta_k}\mu_k$. Thus the joint distribution of (u_k, v_k) has pdf $n(u_k; 0, \frac{\rho^2}{2}) n(v_k; 0, \frac{\rho^2}{2})$. Note

$$y_k = |\langle x, f_k\rangle|^2 + 2|\langle x, f_k\rangle| u_k + u_k^2 + v_k^2$$

Consider now the polar change of cordinates $(u_k, v_k) \mapsto (y_k, \theta_k)$ where

$$u_k = \sqrt{y_k} cos(\theta_k) - |\langle x, f_k\rangle| \quad , \quad v_k = \sqrt{y_k} sin(\theta_k).$$

The Jacobian of the inverse map $(y_k, \theta_k) \mapsto (u_k, v_k)$ is $\frac{1}{2}$. Thus the joint pdf of (y_k, θ_k) is given by

$$\begin{aligned} p(y_k, \theta_k; x) &= \frac{1}{2} n(\sqrt{y_k} cos(\theta_k) - |\langle x, f_k\rangle|; 0, \frac{\rho^2}{2}) \times \\ &\times \quad n(\sqrt{y_k} sin(\theta_k); 0, \frac{\rho^2}{2}) \end{aligned} \tag{2.6}$$

where x is the "clean" signal. By integrating over θ_k we obtain the marginal

$$p(y_k; x) = \frac{1}{\rho^2} exp\left\{ -\frac{y_k}{\rho^2} - \frac{|\langle x, f_k\rangle|^2}{\rho^2} \right\} I_0\left(\frac{2|\langle x, f_k\rangle| \sqrt{y_k}}{\rho^2} \right) \tag{2.7}$$

where I_0 is the modified Bessel function of the first kind and order 0 (see [1], (9.6.16)). Hence the likelihood function for $y = (y_k)_{1\leq k\leq m}$ is given by

$$\begin{aligned} p(y;x) &= \prod_{k=1}^m p(y_k; x) \\ &= \frac{1}{\rho^{2m}} exp\left\{ -\frac{1}{\rho^2} \left(\sum_{k=1}^m y_k + \sum_{k=1}^m |\langle x, f_k\rangle|^2 \right) \right\} \times \\ &\times \prod_{k=1}^m I_0\left(\frac{2|\langle x, f_k\rangle| \sqrt{y_k}}{\rho^2} \right). \end{aligned} \tag{2.8}$$

As we observed in an earlier paper ([6]) it is more advatageous to work with the realification of the problem. Let $\mathbf{j} : \mathbb{C}^n \to \mathbb{R}^{2n}$ dentote the $\mathbb{R}$-linear map

$$z \in \mathbb{C}^n \mapsto \zeta = \mathbf{j}(z) = \begin{bmatrix} real(z) \\ imag(z) \end{bmatrix}. \tag{2.9}$$

For $1 \leq k \leq m$ let

$$\xi = \mathbf{j}(x), \;\; \varphi_k = \mathbf{j}(f_k) \text{ and } J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \tag{2.10}$$

where I is the identity matrix of size n. Note $J^2 = -I$ (identity of order 2n), $J^T = -J$ and $\mathbf{j}(ix) = J\mathbf{j}(x)$.

Denote further

$$\Phi_k = \varphi_k\varphi_k^T + J\varphi_k\varphi_k^T J^T \tag{2.11}$$

and

$$\mathbf{S} = \sum_{k=1}^m \Phi_k \tag{2.12}$$

which represents the frame operator acting on the realification space $H_{\mathbb{R}} = \mathbb{R}^{2n}$. A little algebra shows that for every $x, y \in H$ and $1 \leq k \leq m$:

$$\langle x, f_k\rangle = \langle \xi, \varphi_k\rangle + i\langle \xi, J\varphi_k\rangle \tag{2.13}$$

$$|\langle x, f_k\rangle|^2 = \langle \Phi_k\xi, \xi\rangle \tag{2.14}$$

$$|\langle x, f_k\rangle| = \sqrt{\langle \Phi_k\xi, \xi\rangle} \tag{2.15}$$

$$real(\langle x, f_k\rangle\langle f_k, y\rangle) = \langle \Phi_k\xi, \eta\rangle \tag{2.16}$$

where $\xi = \mathbf{j}(x)$ and $\eta = \mathbf{j}(y)$. Thus the log-likelihood becomes

$$\begin{aligned} log\, p(y; \xi = \mathbf{j}(x)) &= 2m \log\, \rho + \sum_{k=1}^m log\, I_0\left(\frac{2\sqrt{y_k\langle \Phi_k\xi, \xi\rangle}}{\rho^2} \right) \\ &\quad -\frac{1}{\rho^2}\sum_{k=1}^m y_k - \frac{1}{\rho^2}\langle \mathbf{S}\xi, \xi\rangle \end{aligned}$$

Next we compute the (column-vector) gradient

$$\begin{aligned} \nabla_\xi\, log\, p(y;\xi) &= \frac{2}{\rho^2}\sum_{k=1}^m \frac{I_1}{I_0}\left(\frac{2\sqrt{y_k\langle \Phi_k\xi, \xi\rangle}}{\rho^2} \right) \sqrt{\frac{y_k}{\langle \Phi_k\xi, \xi\rangle}} \Phi_k\xi \\ &- \frac{2}{\rho^2}\mathbf{S}\xi \end{aligned}$$

where $I_1 = I_0'$ is the modified Bessel function of the first kind and order 1 (see [1] (9.6.27)). While we shall not use explicitly the Hessian, a similar but slightly more tedious computation shows the Hessian matrix of the log-likelihood to be:

$$\begin{aligned} \nabla_\xi^2\, p(y;\xi) &= -\frac{2}{\rho^2}\mathbf{S} \\ &+ \frac{4}{\rho^4}\sum_{k=1}^m \frac{\frac{1}{2}I_2 I_0 + \frac{1}{2}I_0^2 - I_1^2}{I_0^2}\left(\frac{2\sqrt{y_k\langle \Phi_k\xi, \xi\rangle}}{\rho^2} \right) \\ &\quad \times \frac{y_k}{\langle \Phi_k\xi, \xi\rangle}\Phi_k\xi\xi^*\Phi_k \\ &+ \frac{2}{\rho^2}\sum_{k=1}^m \frac{I_1}{I_0}\left(\frac{2\sqrt{y_k\langle \Phi_k\xi, \xi\rangle}}{\rho^2} \right) \\ &\quad \times \sqrt{\frac{y_k}{\langle \Phi_k\xi, \xi\rangle}}\left(\Phi_k - \frac{1}{\langle \Phi_k\xi, \xi\rangle}\Phi_k\xi\xi^*\Phi_k \right) \end{aligned}$$

where $I_2 = 2I_1' - I_0$ is the modified Bessel function of the first kind and order 2 (see [1] (9.6.26-3)). For the Fisher information matrix we use the gradient:

$$\mathbb{I}(\xi = \mathbf{j}(x)) = \mathbb{E}\left[(\nabla_\xi\, log\, p(y;\xi)) \cdot (\nabla_\xi\, log\, p(y;\xi))^T \right]$$

which becomes:

$$\begin{aligned}\mathbb{I}(\xi) &= \frac{4}{\rho^4}\mathbf{S}\xi\xi^*\mathbf{S} \\ &- \frac{4}{\rho^4}\sum_{k=1}^m \mathbb{E}\left[\frac{I_1}{I_0}|_{\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}}\sqrt{\frac{y_k}{\langle\Phi_k\xi,\xi\rangle}}\right]\times \\ &\times(\Phi_k\xi\xi^*\mathbf{S}+\mathbf{S}\xi\xi^*\Phi_k) \\ +\frac{4}{\rho^4}\sum_{k,l=1}^m &\mathbb{E}\left[\frac{I_1}{I_0}|_{\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}}\frac{I_1}{I_0}|_{\frac{2\sqrt{y_l\langle\Phi_l\xi,\xi\rangle}}{\rho^2}}\sqrt{\frac{y_k y_l}{\langle\Phi_k\xi,\xi\rangle\langle\Phi_l\xi,\xi\rangle}}\right] \\ &\times\Phi_k\xi\xi^*\Phi_l\end{aligned}$$

Next we compute the expectations. Notice the double sum contains two types of terms: those with $k = l$ and those for $k \neq l$. If $k \neq l$ then the expectation factors as a product of the expectation involving the k-indexed term and the expectation of the l-indexed term. Let us denote

$$L_k = \mathbb{E}\left[\frac{I_1}{I_0}\left(\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\right)\sqrt{\frac{y_k}{\langle\Phi_k\xi,\xi\rangle}}\right] \quad (2.17)$$

$$Q_k = \mathbb{E}\left[\frac{I_1^2}{I_0^2}\left(\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\right)\frac{y_k}{\langle\Phi_k\xi,\xi\rangle}\right] \quad (2.18)$$

Then the Fisher information becomes

$$\begin{aligned}\mathbb{I}(\xi) = \frac{4}{\rho^4}\Bigg[&\mathbf{S}\xi\xi^*\mathbf{S} - \sum_{k=1}^m L_k(\Phi_k\xi\xi^*\mathbf{S}+\mathbf{S}\xi\xi^*\Phi_k)+ \quad (2.19) \\ &+(\sum_{k=1}^m L_k\Phi_k\xi)(\sum_{k=1}^m L_k\Phi_k\xi)^* + \sum_{k=1}^m (Q_k - L_k^2)\Phi_k\xi\xi^*\Phi_k\Bigg]\end{aligned}$$

Surprisingly this expression simplifies significantly once we establish the following lemma:

Lemma 2.1: For the non-additive white Gaussian noise model (2.5), $L_k = 1$. This means:

$$\mathbb{E}\left[\frac{I_1}{I_0}\left(\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\right)\sqrt{\frac{y_k}{\langle\Phi_k\xi,\xi\rangle}}\right] = 1$$

for all $1 \leq k \leq m$.

Proof. This result is obtained by direct computation. The expectation is taken with respect to (2.7):

$$e^{\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}}\mathbb{E}\left[\frac{I_1}{I_0}\left(\frac{2\sqrt{y_k\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\right)\sqrt{y_k}\right] =$$

$$= e^{\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}}\int_0^\infty I_1\left(\frac{2\sqrt{y\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\right)\sqrt{y}\frac{1}{\rho}^2 e^{-\frac{y}{\rho^2}-\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}}dy$$

Next use the series expansion (9.6.10) in [1] for $I_1(z) = \sum_{k=0}^\infty \frac{1}{k!(k+1)!}(\frac{z}{2})^{2k+1}$. Substitute in the formula above and integrate term by term:

$$\frac{1}{\rho^2}\sum_{k=0}^\infty \frac{1}{k!(k+1)!}\int_0^\infty\left(\frac{\sqrt{\langle\Phi_k\xi,\xi\rangle}}{\rho^2}\sqrt{y}\right)^{2k+1}\sqrt{y}e^{-\frac{y}{\rho^2}}dy =$$

$$\frac{\sqrt{\langle\Phi_k\xi,\xi\rangle}}{\rho^4}\sum_{k=0}^\infty\frac{1}{k!(k+1)!}\left(\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^4}\right)^k\int_0^\infty y^{k+1}e^{-y/\rho^2}dy =$$

$$\sqrt{\langle\Phi_k\xi,\xi\rangle}\sum_{k=1}^\infty\frac{1}{k!}\left(\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}\right)^k = \sqrt{\langle\Phi_k\xi,\xi\rangle}e^{\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}}$$

from where the lemma follows. □

This lemma allows us to simplify the Fisher information matrix to:

$$\mathbb{I}(\xi) = \frac{4}{\rho^4}\sum_{k=1}^m (Q_k - 1)\Phi_k\xi\xi^*\Phi_k \quad (2.20)$$

Let us denote by G_1 and G_2 the following two scalar functions:

$$\begin{aligned}G_1(a) &= \frac{e^{-a}}{a}\int_0^\infty \frac{I_1^2(2\sqrt{at})}{I_0(2\sqrt{at})}te^{-t}dt \quad (2.21) \\ &= \frac{e^{-a}}{8a^3}\int_0^\infty\frac{I_1^2(t)}{I_0(t)}t^3e^{-\frac{t^2}{4a}}dt \\ G_2(a) &= a(G_1(a)-1) \quad (2.22)\end{aligned}$$

Then we obtain:

Lemma 2.2: For the non-additive white Gaussian noise model (2.5) we have

$$Q_k = G_1\left(\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}\right) \quad (2.23)$$

Proof. This follows by direct computation. □

In turn this lemma yields:

Theorem 2.3: The Fisher information matrix for the non-additive white Gaussian noise model (2.5) is given by

$$\begin{aligned}\mathbb{I}(\xi) &= \frac{4}{\rho^4}\sum_{k=1}^m\left(G_1\left(\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}\right)-1\right)\Phi_k\xi\xi^*\Phi_k \quad (2.24) \\ &= \frac{4}{\rho^2}\sum_{k=1}^m G_2\left(\frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}\right)\frac{1}{\langle\Phi_k\xi,\xi\rangle}\Phi_k\xi\xi^*\Phi_k \quad (2.25)\end{aligned}$$

For small Signal-To-Noise-Ratio $SNR = \frac{\langle\Phi_k\xi,\xi\rangle}{\rho^2}$, $G_1(SNR) \approx 2$ and $G_2(SNR) \approx SNR$. Thus:

$$\mathbb{I}(\xi) \approx \frac{4}{\rho^4}\sum_{k=1}^m \Phi_k\xi\xi^*\Phi_k \ , \text{ when } \frac{\max_k\langle\Phi_k\xi,\xi\rangle}{\rho^2} \ll 1 \quad (2.26)$$

For large SNR, $G_1(SNR) \approx 1 + \frac{1}{2SNR}$ and $\lim_{SNR\to\infty} G_2(SNR) = \frac{1}{2}$. Hence

$$\mathbb{I}(\xi) \approx \frac{2}{\rho^2}\sum_{k=1}^m\frac{1}{\langle\Phi_k\xi,\xi\rangle}\Phi_k\xi\xi^*\Phi_k \ , \text{ when } \frac{\min_k\langle\Phi_k\xi,\xi\rangle}{\rho^2} \gg 1 \quad (2.27)$$

Proof. Equation (2.25 follows from (2.20) and (2.23). The two asymptotical regimes follow from:

$$\lim_{a\to 0} G_1(a) = 2 \ \text{ and } \ \lim_{a\to\infty}\frac{G_1(a)}{1+\frac{1}{2a}} = 1$$

These limits are obtained as follows. For small SNR, we can approximate $I_0(t) \approx 1$ and $I_1(t) \approx \frac{t}{2}$ (see [1], (9.8.1) and (9.8.3)). Then substitute these expressions in (2.21) and obtain the first limit.

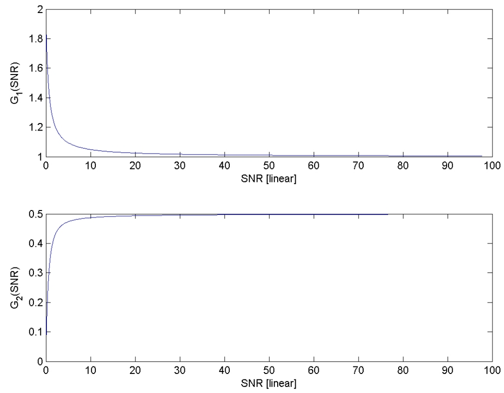

Fig. 1. Plots of G_1 (top) and G_2 (bottom) with SNR on a linear scale.

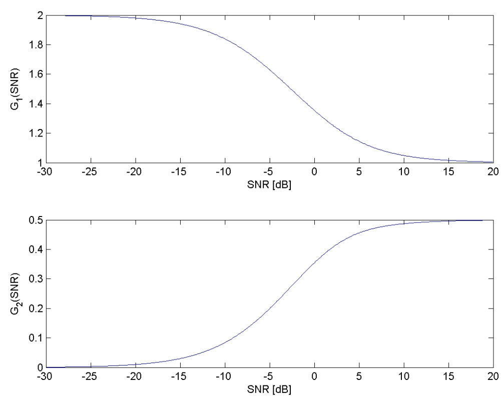

Fig. 2. Plots of G_1 (top) and G_2 (bottom) with SNR on a dB scale.

For large SNR use $I_0(t) \approx \frac{e^t}{\sqrt{2\pi t}}(1 - \frac{3}{8t})$ and $I_1(t) \approx \frac{e^t}{\sqrt{2\pi t}}(1 + \frac{1}{8t})$ (see [1], (9.7.1)). Then substitute these expressions in (2.21) and obtain the second limit. □

It is useful to illustrate the two functions G_1 and G_2. Figures 1 and 2 contain the plots of these functions. In figure 1 we use a linear scale for SNR. In figure 2 we use a logarithmic scale (dB) for SNR. Specifically $SNR\,[dB] = 10\,log_{10}(SNR\,[linear])$.

III. The identifiability problem

It is interesting to note the relationship between the Fisher information matrix we derived in the previous section and conditions for phase retrievable frames. As we know the vector x is not identifiable from measurement y (in the absence of noise). At best its class $\hat{x}$ can be identified from y, in the absence of noise. This nonidentifiability is expressed in the fact that $I(\xi)$ is always rank deficient. In fact the vector $J\xi$ is always in the null space of $I(\xi)$. However the question is whether this is the only independent vector in the null space. The following result summarizes a necessary and sufficient condition for the frame $\mathcal{F}$ to be phase retrievable. For this let us introduce one more object:

$$\mathcal{R} : \mathbb{R}^{2n} \to Sym(\mathbb{R}^{2n}) \ , \ \mathcal{R}(\xi) = \sum_{k=1}^{m} \Phi_k \xi\xi^* \Phi_k \qquad (3.28)$$

where $Sym(\mathbb{R}^{2n})$ denotes the space of symmetric operators over $\mathbb{R}^{2n}$.

Theorem 3.1 ([6]): The following are equivalent:

1) The frame $\mathcal{F}$ is phase retrievable;
2) For every $0 \neq \xi \in \mathbb{R}^{2n}$, $rank(\mathcal{R}(\xi)) = 2n - 1$;
3) There is a constant $a_0 > 0$ so that for every $\xi \in \mathbb{R}^{2n}$ with $\|\xi\| = 1$,
$$\mathcal{R}(\xi) \geq a_0(I - J\xi\xi^* J^*) \qquad (3.29)$$
where the inequality is between quadratic forms;
4) There is a constant $a_0 > 0$ so that for every $\xi, \eta \in \mathbb{R}^{2n}$,
$$\sum_{k=1}^{m} |\langle \Phi_k \xi, \eta \rangle|^2 \geq a_0(\|\xi\|^2 \|\eta\|^2 - |\langle J\xi, \eta \rangle|^2). \qquad (3.30)$$

Furthermore, the constants a_0 at 3. and 4. can be chosen to be the same.

Now we show that $\mathcal{F}$ is phase retrievable if and only if $rank(\mathbb{I}(\xi)) = 2n - 1$ for all $\xi \neq 0$. Furthermore, we establish also a lower bound on $\mathbb{I}(\xi)$ in the sense of quadratic forms:

Theorem 3.2: Fix $\rho > 0$ and let B denote the upper frame bound. The following are equivalent:

1) The frame $\mathcal{F}$ is phase retrievable;
2) For every $0 \neq \xi \in \mathbb{R}^{2n}$, $rank(\mathbb{I}(\xi)) = 2n - 1$;
3) There is a constant $c_0 > 0$ that depends on ρ and frame $\mathcal{F}$ so that for every $\xi \in \mathbb{R}^{2n}$, $\|\xi\| \leq \rho\sqrt{\frac{10}{B}}$,
$$\mathbb{I}(\xi) \geq c_0(I - J\xi\xi^* J^*) \qquad (3.31)$$
where the inequality is between quadratic forms.

This result follows directly from Theorem 3.1 and the following lemma:

Lemma 3.3: Fix $\rho > 0$. Let A, B be the frame bounds. Set $D_0 = \frac{4}{\rho^4}$ and $d_0 = \frac{0.16}{\rho^4}$. Then:

1) For every $\xi \in \mathbb{R}^{2n}$, $\mathbb{I}(\xi) \leq D_0 \mathcal{R}(\xi)$.
2) For every $\xi \in \mathbb{R}^{2n}$ with $\|\xi\| \leq \rho\sqrt{\frac{10}{B}}$, $\mathbb{I}(x) \geq d_0 \mathcal{R}(\xi)$.
3) For every $\xi \in \mathbb{R}^{2n}$, $\mathbb{I}(\xi) \geq \frac{4}{\rho^4}\left(G_1(\frac{B\|\xi\|^2}{\rho^2}) - 1\right)\mathcal{R}(\xi)$.

Proof Since G_1 is monotonically decreasing and $G_1(a) \leq 2$, then from (2.24) it follows that $\mathbb{I}(\xi) \leq \frac{4}{\rho^4}\mathcal{R}(\xi)$. For the second inequality, notice G_2 is monotonically increasing and concave. A lower bound is $G_2(a) \geq min(0.04a, 0.4)$, where the break point is for $SNR = 10$. Thus by (2.25)

$$\mathbb{I}(\xi) \geq \frac{4}{\rho^2} \sum_{k=1}^{m} \min(\frac{0.04}{\rho^2}, \frac{0.4}{\langle \Phi_k \xi, \xi \rangle}) \Phi_k \xi\xi^* \Phi_k$$

Since $\langle \Phi_k \xi, \xi \rangle = |\langle f_x, \mathbf{j}^{-1}(\xi) \rangle|^2 \leq B\|\xi\|^2 \leq 10\rho^2$ it follows

$$\min(\frac{0.04}{\rho^2}, \frac{0.4}{\langle \Phi_k \xi, \xi \rangle}) \geq \frac{0.04}{\rho^2}$$

Thus

$$\mathbb{I}(\xi) \geq \frac{0.16}{\rho^4} \sum_{k=1}^{m} \Phi_k \xi \xi^* \Phi_k$$

which proves the second statement. The third inequality follows from the fact that $max_k \langle \Phi_k \xi, \xi \rangle \leq B \|\xi\|^2$ and G_1 is monotonically decreasing. □

Proof of Theorem 3.2.

$1 \Leftrightarrow 2$. Note that $rank(\mathbb{I}(\xi)) = rank(\mathcal{R}(\xi))$. Thus the claim follows from Theorem 3.1(2).

$1 \Rightarrow 3$. If $\mathcal{F}$ is phase retrievable then by Theorem 3.1(3) and Lemma 3.3(3) it follows $\mathbb{I}(\xi) \geq d_0 \mathcal{R}(\xi) \geq d_0 a_0 (I - J\xi\xi^* J^*)$.

$3 \Rightarrow 1$. Equation (3.31) and Lemma 3.3(1) imply $\mathcal{R}(\xi) \geq \frac{c_0}{D_0}(I - J\xi\xi^* J^*)$ and thus the frame is phase retrievable by Theorem 3.1(3). □

Note the constant c_0 in Theorem 3.2 can be chosen as $c_0 = \frac{0.16 a_0}{\rho^4}$ with a_0 as in Theorem 3.1.

IV. The case of other exponents p

In the case the exponent p is different than 2, the Fisher information matrix can be easily obtained from (2.25). Indeed consider the model:

$$z_k = |\langle x, f_k \rangle + \mu_k|^p \quad , \quad 1 \leq k \leq m \tag{4.32}$$

where $p \neq 0$, $\mathcal{F} = \{f_1, \cdots, f_m\}$ is a phase retrievable frame and $(\mu_1, \cdots, \mu_m)$ are independent and identically distributed complex random variables with distribution $\mathbb{C}\mathcal{N}(0, \rho^2)$. The likelihood of $z = (z_k)_{1 \leq k \leq m}$ can be easily obtained from the distribution of y. Indeed the change of distribution is performed via $z_k = (y_k)^{p/2}$. Hence:

$$p_Z(z;\xi) = \frac{2}{p} z^{1-\frac{2}{p}} p_Y(z^{\frac{2}{p}};\xi).$$

Thus

$$\nabla_\xi \log p_Z(z;\xi) = \nabla_\xi \log p_Y(y;\xi) \quad ; \quad y_k = z_k^{2/p}$$

which implies that the Fisher information matrix for measurements model (4.32) is the same as for (2.5), hence also $\mathbb{I}(\xi)$.

V. The Cramer-Rao Lower Bound

Let us use now the Fisher information matrix derived in a previous section in order to derive performance bounds for statistical estimators. First we need to constraint the estimation problem so the signal to become identifiable.

Fix a unit-norm vector $z_0 \in H$, $\|z_0\| = 1$ and let $\zeta_0 = \mathbf{j}(z_0) \in H_\mathbb{R} = \mathbb{R}^{2n}$. Define the closed set $\Omega_{z_0} = \{\xi \in \mathbb{R}^{2n} \ , \ \langle \xi, \zeta_0 \rangle) \geq 0, \langle \xi, J\zeta_0 \rangle) = 0\}$ and its relative interior: $\mathring{\Omega}_{z_0} = \{\xi \in \mathbb{R}^{2n} \ , \ \langle \xi, \zeta_0 \rangle) > 0, \langle \xi, J\zeta_0 \rangle) = 0\}$. Let $E_{z_0} = span_\mathbb{R} \mathring{\Omega}_{z_0}$ be the real span of $\mathring{\Omega}_{z_0}$. Note E_{z_0} is the orthogonal complement of $J\zeta_0$, $E_{z_0} = \{J\zeta_0\}^\perp$. Let Π_{z_0} denote the orthogonal projection onto E_{z_0}, $\Pi_{z_0} = 1 - J\zeta_0\zeta_0^* J^*$.

Assume now the following scenario. We assume the vector to-be-estimated x satisfies $real(\langle x, z_0 \rangle) > 0$ and $imag(\langle x, z_0 \rangle) = 0$. For $\xi = \mathbf{j}(x)$ this means $\xi \in \mathring{\Omega}_{z_0}$. Then following the discussion in [6] we obtain the Fisher information matrix for this scenario is

$$\mathbb{I}_{z_0}(\xi) = \Pi_{z_0} \mathbb{I}(\xi) \Pi_{z_0} \tag{5.33}$$

Next we restrict to the class of *unbiased estimators*, that are functions $\omega : \mathbb{R}^m \to \Omega_{z_0}$ so that $\mathbb{E}[\omega(y);\xi] = \xi$ for all $\xi \in \mathring{\Omega}_{z_0}$. Again following Theorem 4.3 in [6] we obatin:

Theorem 5.1: Assume the model (2.5) with $\xi = \mathbf{j}(x) \in \mathring{\Omega}_{z_0}$. Then the covariance of any unbiased estimator is bounded below by:

$$Cov[\omega(y);\xi] \geq (\Pi_{z_0} \mathbb{I}_{z_0}(\xi) \Pi_{z_0})^\dagger \tag{5.34}$$

for every $\xi \in \mathring{\Omega}_{z_0}$, where $\dagger$ denotes the pseudo-inverse operation. In particular the Mean-Square Error (MSE) of such an estimator is bounded below by:

$$\mathbb{E}[\|\omega(y) - \xi\|^2;\xi] \geq trace\left\{(\Pi_{z_0} \mathbb{I}_{z_0}(\xi) \Pi_{z_0})^\dagger\right\} \tag{5.35}$$

for every $\xi \in \mathring{\Omega}_{z_0}$.

VI. Conclusion

In this paper we analyzed the Fisher information matrix and the Cramer-Rao lower bound for a non-additive white Gaussian noise model in the phase retrieval problem. Specifically we obtained a closed-form expression for these objects that involves parametric integrals of modified Bessel functions. The rank condition is similar to the case of an AWGN model.

Acknowledgment

The author was partially supported by NSF under DMS-1109498 and DMS-1413249 grants.

References

[1] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, US Dept.Comm., Nat.Bur.Stand. Applied Math. Ser. 55, 10th Printing, 1972.

[2] B. Alexeev, A. S. Bandeira, M. Fickus, D. G. Mixon, Phase retrieval with polarization, SIAM J. Imaging Sci., **7** (1) (2014), 35–66.

[3] R. Balan, P. Casazza, D. Edidin, On signal reconstruction without phase, Appl.Comput.Harmon.Anal. **20** (2006), 345–356.

[4] R. Balan, B. Bodmann, P. Casazza, D. Edidin, Painless reconstruction from Magnitudes of Frame Coefficients, J.Fourier Anal.Applic., **15** (4) (2009), 488–501.

[5] R. Balan, Reconstruction of Signals from Magnitudes of Frame Representations, arXiv submission arXiv:1207.1134

[6] R. Balan, Reconstruction of Signals from Magnitudes of Redundant Representations: The Complex Case, available online arXiv:1304.1839v1, Found.Comput.Math. 2015, http://dx.doi.org/10.1007/s10208-015-9261-0

[7] R. Balan and Y. Wang, Invertibility and Robustness of Phaseless Reconstruction, available online arXiv:1308.4718v1, Appl. Comp. Harm. Anal., 38 (2015), 469–488.

[8] A. S. Bandeira, J. Cahill, D. Mixon, A. A. Nelson, Saving phase: Injectivity and Stability for phase retrieval, arXiv submission , arXiv: 1302.4618, Appl. Comp. Harm. Anal. 37 (1) (2014), 106–125.

[9] B. G. Bodmann, N. Hammen, Stable Phase Retrieval with Low-Redundancy Frames, arXiv submission:1302.5487v1, Adv. Comput. Math., accepted 10 April 2014.

[10] E. Candés, T. Strohmer, V. Voroninski, PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming, Communications in Pure and Applied Mathematics 66 (2013), 1241–1274.

[11] P. Casazza, The art of frame theory, Taiwanese J. Math., 4(2) (2000), 129–202.

[12] Y. Ephraim, D. Malah, Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator, IEEE Trans. ASSP, 32(6) (1984), 1109-1121.