The Cramér-Rao Lower Bound in the Phase Retrieval Problem

Radu Balan, David Bekkerman

Department of Mathematics and CSCAMM
University of Maryland, College Park, MD 20742

July 10, 2019
SampTA 2019 Conference
Bordeaux, FRANCE
Notations and Assumptions
Phase Retrievability and Identifiability

- Hilbert space $H = \mathbb{C}^n$, $\hat{H} = H/T^1$, frame $\mathcal{F} = \{f_1, \cdots, f_m\} \subset \mathbb{C}^n$ and
  $$\alpha : \hat{H} \rightarrow \mathbb{R}^m, \quad \alpha(x) = (|\langle x, f_k \rangle|)_{1 \leq k \leq m}.$$  

- We assume the frame is phase retrievable, i.e., $\alpha$ is injective. Hence $(|\langle x, f_k \rangle|)_{1 \leq k \leq m}$ determine uniquely $x$ up to a global phase factor.
Notations and Assumptions

Phase Retriviality and Identifiability

- Hilbert space $H = \mathbb{C}^n$, $\hat{H} = H / T^1$, frame $\mathcal{F} = \{f_1, \cdots, f_m\} \subset \mathbb{C}^n$ and
  
  $$\alpha : \hat{H} \rightarrow \mathbb{R}^m \; , \; \alpha(x) = (|\langle x, f_k \rangle|)_{1 \leq k \leq m} .$$

- We assume the frame is phase retrievable, i.e., $\alpha$ is injective. Hence $(|\langle x, f_k \rangle|)_{1 \leq k \leq m}$ determine uniquely $x$ up to a global phase factor.

- Measurement process: $y = (y_k)_{1 \leq k \leq m}$. We assume the distribution of $y$, $p(y; x)$ depends on $\alpha(x)$ only. For instance:
  
  $$y_k = |\langle x, f_k \rangle|^a + \nu_k \; , \; \mu_k \sim \mathbb{C}N(0, \rho^2) \; , \; \nu_k \sim \mathbb{N}(0, \sigma^2)$$

  Specifically: $p(y; x) = F(s_1, \cdots, s_m, y)$, where $s_k = |\langle x, f_k \rangle|$.

- We assume identifiability and regularity: (1) If $\forall y \in \mathbb{R}^m$, $F(s^{[1]}, y) = F(s^{[2]}, y)$ then $s^{[1]} = s^{[2]}$; and, (2) The Fisher Infomatrix $\mathbb{E}[\frac{\partial \log(F)}{\partial s_k} \frac{\partial \log(F)}{\partial s_j}]$ is continuous and has constant rank on an open neighborhood of the operating point [Rthbrg71].
Assumptions:

\[ x \xrightarrow{(f_1, \ldots, f_m)} (s_1, \ldots, s_m) \xrightarrow{y} \]

- Phase Retrievable
- Identifiable
Problem Statement
FIM vs. CRLB

Assumptions:

In previous works we derived various Fisher Information Matrix expressions. We have also derived a Cramér-Rao Lower Bound (CRLB) for a specific estimation model. In this paper we analyze a second identification problem and compare the two CRLBs:

Problem

*The problem is not how to compute the Fisher Information Matrix (FIM). The problem is how to use FIM, to derive Cramér-Rao Lower Bounds.*
Fisher Info Matrix for the AWGN Model

- For the AWGN model:

\[ y_k = |\langle x, f_k \rangle|^2 + \nu_k , \quad 1 \leq k \leq m \]

with \( \nu_k \sim \mathbb{C}\mathcal{N}(0, \sigma^2) \) i.i.d. the Fisher Information Matrix:

\[ I = \mathbb{E} [(\nabla_x \log p(y; x))(\nabla_x \log p(y; x))^*] \]
Fisher Info Matrix for the AWGN Model

- For the AWGN model:

\[ y_k = |\langle x, f_k \rangle|^2 + \nu_k , \quad 1 \leq k \leq m \]

with \( \nu_k \sim \mathcal{CN}(0, \sigma^2) \) i.i.d. the Fisher Information Matrix:

\[ \mathbb{I} = \mathbb{E} [ (\nabla_x \log p(y; x))(\nabla_x \log p(y; x))^*] \]

- \( \mathbb{I}^{AWGN,real}(x) = \frac{4}{\sigma^2} \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 f_k f_k^T = \frac{4}{\sigma^2} \sum_{k=1}^{m} (f_k f_k^T)xx^T(f_k f_k^T) \)
For the AWGN model:

\[ y_k = |\langle x, f_k \rangle|^2 + \nu_k \quad , \quad 1 \leq k \leq m \]

with \( \nu_k \sim \mathbb{C}\mathcal{N}(0, \sigma^2) \) i.i.d. the Fisher Information Matrix:

\[ I = \mathbb{E}[(\nabla_x \log p(y; x))(\nabla_x \log p(y; x))^\ast] \]

- \( I^{AWGN, \text{real}}(x) = \frac{4}{\sigma^2} \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 f_k f_k^T = \frac{4}{\sigma^2} \sum_{k=1}^{m} (f_k f_k^T) x x^T (f_k f_k^T) \)

- \( I^{AWGN, \text{cplx}}(x) = \frac{4}{\sigma^2} \sum_{k=1}^{m} \Phi_k \xi^\ast \xi \Phi_k \) [Bal13, BCMN13] with \( \Phi_k \in \mathbb{R}^{2n \times 2n} \) and \( \xi \in \mathbb{R}^{2n} \).
Consider the Non-AWGN model:

\[ y_k = |\langle x, f_k \rangle + \mu_k|^2, \quad 1 \leq k \leq m \]

with \( \mu_k \sim \mathcal{CN}(0, \rho^2) \) i.i.d.
Consider the Non-AWGN model:

\[ y_k = |\langle x, f_k \rangle + \mu_k|^2, \quad 1 \leq k \leq m \]

with \( \mu_k \sim \mathcal{CN}(0, \rho^2) \) i.i.d.

The likelihood function:

\[
p(y; x) = \frac{1}{\rho^{2m}} \exp \left\{ -\frac{1}{\rho^2} \left( \sum_{k=1}^{m} y_k + \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 \right) \right\} \prod_{k=1}^{m} l_0 \left( \frac{2|\langle x, f_k \rangle|\sqrt{y_k}}{\rho^2} \right)
\]
Problem Statement

Consider the Non-AWGN model:

\[ y_k = |\langle x, f_k \rangle + \mu_k|^2, \quad 1 \leq k \leq m \]

with \( \mu_k \sim \mathcal{CN}(0, \rho^2) \) i.i.d.

The likelihood function:

\[
p(y; x) = \frac{1}{\rho^{2m}} \exp \left\{ - \frac{1}{\rho^2} \left( \sum_{k=1}^{m} y_k + \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 \right) \right\} \prod_{k=1}^{m} I_0 \left( \frac{2|\langle x, f_k \rangle| \sqrt{y_k}}{\rho^2} \right)
\]

Realification: \( x \mapsto \xi = [\text{real}(x) \ \text{imag}(x)]^T \) and \( |\langle x, f_k \rangle| = \sqrt{\langle \Phi_k \xi, \xi \rangle} \)

where \( \Phi_k \) is a rank-2 replacing \( f_k f_k^* \).
FIM for Non-AWGN

Consider the Non-AWGN model:

\[ y_k = |\langle x, f_k \rangle + \mu_k|^2, \quad 1 \leq k \leq m \]

with \( \mu_k \sim \mathbb{C}\mathcal{N}(0, \rho^2) \) i.i.d.

The likelihood function:

\[
p(y; x) = \frac{1}{\rho^{2m}} \exp \left\{ - \frac{1}{\rho^2} \left( \sum_{k=1}^{m} y_k + \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 \right) \right\} \prod_{k=1}^{m} I_0 \left( \frac{2|\langle x, f_k \rangle|\sqrt{y_k}}{\rho^2} \right)
\]

Realification: \( x \mapsto \xi = [\text{real}(x) \; \text{imag}(x)]^T \) and \( |\langle x, f_k \rangle| = \sqrt{\langle \Phi_k \xi, \xi \rangle} \)

where \( \Phi_k \) is a rank-2 replacing \( f_k f_k^* \).
FIM for Non-AWGN

Theorem (Bal15)

The Fisher Information Matrix for the Non-AWGN model is given by

\[
\mathbb{I}(\xi) = \frac{4}{\rho^4} \sum_{k=1}^{m} \left( G_1 \left( \frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) - 1 \right) \Phi_k \xi \xi^* \Phi_k
\]

\[
= \frac{4}{\rho^2} \sum_{k=1}^{m} G_2 \left( \frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^* \Phi_k
\]

where

\[
G_1(a) = \frac{e^{-a}}{8a^3} \int_{0}^{\infty} \frac{l_1^2(t)}{l_0(t)} t^3 e^{-\frac{t^2}{4a}} dt , \quad G_2(a) = a(G_1(a) - 1)
\]
FIM for Non-AWGN
Asymptotic Regimes

Form 1: Low SNR
\[ I(\xi) = 4\rho^4 \sum_{m} G_1(\langle \Phi_k \xi, \xi \rangle \rho^2) - 1 \approx 4\rho^4 \sum_{m} \Phi_k \xi \xi^* \Phi_k \xi \]

Form 2: High SNR
\[ I(\xi) = 4\rho^2 \sum_{m} G_2(\langle \Phi_k \xi, \xi \rangle \rho^2 \langle \Phi_k \xi, \xi \rangle) \approx 2\rho^2 \sum_{m} \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^* \Phi_k \xi \]
FIM for Non-AWGN
Asymptotic Regimes

Form 1: Low SNR

$$\mathcal{I}(\xi) = \frac{4}{\rho^4} \sum_{k=1}^{m} \left( G_1 \left( \frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) - 1 \right) \Phi_k \xi \xi^* \Phi_k$$

$$\approx \frac{4}{\rho^4} \sum_{k=1}^{m} \Phi_k \xi \xi^* \Phi_k$$
FIM for Non-AWGN
Asymptotic Regimes

**Form 1: Low SNR**

\[ \mathbb{I}(\xi) = \frac{4}{\rho^4} \sum_{k=1}^{m} \left( G_1 \left( \frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) - 1 \right) \Phi_k \xi \xi^* \Phi_k \]

\[ \approx \frac{4}{\rho^4} \sum_{k=1}^{m} \Phi_k \xi \xi^* \Phi_k \]

**Form 2: High SNR**

\[ \mathbb{I}(\xi) = \frac{4}{\rho^2} \sum_{k=1}^{m} G_2 \left( \frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^* \Phi_k \]

\[ \approx \frac{2}{\rho^2} \sum_{k=1}^{m} \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^* \Phi_k \]
Setup 1: Reference signal based estimation

In the first setup we fix a reference unit-norm signal $z_0 \in \mathbb{C}^n$. The unknown (to-be-estimated) signal $x$ is assumed to come from set:

$$V_{z_0} = \{ x \in \mathbb{C}^n : \text{imag}(\langle x, z_0 \rangle) = 0 , \text{real}(\langle x, z_0 \rangle) > 0 \}.$$ 

The estimator has access to the reference signal $z_0$: 

![Diagram of measurement device and estimator]
Setup 1: Reference signal based estimation

In the first setup we fix a reference unit-norm signal $z_0 \in \mathbb{C}^n$. The unknown (to-be-estimated) signal $x$ is assumed to come from set:

$$V_{z_0} = \{ x \in \mathbb{C}^n : \text{imag}(\langle x, z_0 \rangle) = 0 , \text{real}(\langle x, z_0 \rangle) > 0 \}.$$  

The estimator has access to the reference signal $z_0$:

Let $V_{\zeta_0} = \{ \xi \in \mathbb{R}^{2n} , \langle \xi, \zeta_0 \rangle \geq 0 , \langle \xi, J\zeta_0 \rangle ) = 0 \}$. , $\mathcal{E}_{\zeta_0} = \text{span}_\mathbb{R}(V_{\zeta_0})$ with $\zeta_0 = [\text{real}(z_0) \ \text{imag}(z_0)]^T$. The estimator $o : \mathbb{R}^m \rightarrow \mathcal{E}_{\zeta_0}$ is unbiased if $\mathbb{E}[o(y) ; \xi] = \xi$ for every $x \in V_{z_0}$, with $\xi = [\text{real}(x) ; \text{imag}(x)]$. 

Radu Balan, David Bekkerman (UMD) CRLB SampTA Conference July 10, 2019
Setup 1: Positive correlation with a reference signal

The CRL Bound

Let $\Pi_\eta = 1 - \frac{1}{\|\eta\|^2} J\eta\eta^T J^T$ and $L = I - \frac{1}{\langle \xi, \zeta_0 \rangle} J\zeta_0\zeta^T J^T$, with $J$ the symplectic form matrix $[0, -I; I, 0]$.

Theorem

Assume the measurement model $y = (y_k)_{1 \leq k \leq m}$ where the likelihood function $p(y; x) = F(|\langle x, f_1 \rangle|, \cdots, |\langle x, f_m \rangle|, y)$ is identifiable and regular. Then the covariance of any unbiased estimator $\omega : \mathbb{R}^m \rightarrow \mathcal{E}_{\zeta_0}$ is bounded below by

$$\text{Cov}[\omega(y); \xi] \geq (\Pi_{z_0} \Pi(\xi) \Pi_{z_0})^\dagger = L^T (\Pi(\xi))^\dagger L.$$ 

In particular:

$$\mathbb{E}[\|\omega(y) - \xi\|^2; \xi] \geq \text{trace} \left\{ (\Pi_{z_0} \Pi(\xi) \Pi_{z_0})^\dagger \right\} =$$

$$\text{trace}(\Pi(\xi))^\dagger + \frac{\|\xi\|^2}{|\langle \xi, \zeta_0 \rangle|^2} \langle (\Pi(\xi))^\dagger J\zeta_0, J\zeta_0 \rangle.$$ 

Remark: First inequality was derived in 2015 paper; the second equality is new.
Consider now a different setup, where \( x \in \mathbb{C}^n \) is unconstrained and the estimation is performed in two stages: (i) the first stage returns a ”class” estimate through \( o : \mathbb{R}^m \rightarrow \mathbb{C}^n \); (ii) in the second stage, an oracle provides the optimal global phase \( \langle x, o(y) \rangle \). Thus, the overall estimator:

\[
\tilde{o} : \mathbb{R}^m \rightarrow \mathbb{C}^n, \quad \tilde{o}(y) = o(y) \frac{\langle x, o(y) \rangle}{|\langle x, o(y) \rangle|}.
\]

The estimator is **unbiased** if \( \mathbb{E}[\tilde{o}(y); x] = x \) for every \( x \in \mathbb{C}^n \).
Setup 2: Oracle-based signal estimation

The CRL Bound

**Theorem**

Assume the measurement model $y = (y_k)_{1 \leq k \leq m}$ where the likelihood function $p(y; x) = F(|\langle x, f_1 \rangle|, \cdots, |\langle x, f_m \rangle|, y)$ is identifiable and regular.

Let $\tilde{o} : \mathbb{R}^m \to \mathbb{C}^n$ be an unbiased estimator in Setup 2 (Oracle-based estimator). Denote by $\omega(y) = [\text{real}(o(y)); \text{imag}(o(y))]$ and $\omega(y) = [\text{real}(\tilde{o}(y)); \text{imag}(\tilde{o}(y))]$. Then for any $\xi = [\text{real}(x); \text{imag}(x)] \neq 0$,

$$\text{Cov}[\tilde{\omega}(y); \xi] \geq (I - \Delta)(\Pi(\xi))^\dagger(I - \Delta)$$

where $\Delta =$

$$\begin{bmatrix}
\frac{(\langle \omega, J \xi \rangle)^2}{((\langle \omega, \xi \rangle)^2 + (\langle \omega, J \xi \rangle)^2)^{3/2}} \omega \omega^T + \frac{\langle \omega, \xi \rangle \langle \omega, J \xi \rangle}{((\langle \omega, \xi \rangle)^2 + (\langle \omega, J \xi \rangle)^2)^{3/2}} (J\omega \omega^T + \omega \omega^T J^T) + \frac{(\langle \omega, \xi \rangle)^2}{((\langle \omega, \xi \rangle)^2 + (\langle \omega, J \xi \rangle)^2)^{3/2}} J \omega \omega^T J^T
\end{bmatrix}$$

and satisfies $\Delta = \Delta^T \geq I - \Pi^T \xi \geq 0$, $\Delta J \xi = J \xi$ and $\Delta \xi = 0$. 

Radu Balan, David Bekkerman (UMD)  
CRLB  
SampTA Conference July 10, 2019
Conclusions and Open Questions

We obtained Cramér-Rao (type) Lower Bounds for two setups:

1. Positive Correlation with a reference signal: CRLB has a simple form.
2. Oracle-based global phase: CRLB seems very complicated, and estimator dependent. (Remark: Estimator dependency is known for other classes of estimators)
Conclusions and Open Questions

We obtained Cramér-Rao (type) Lower Bounds for two setups:

1. Positive Correlation with a reference signal: CRLB has a simple form.
2. Oracle-based global phase: CRLB seems very complicated, and estimator dependent. (Remark: Estimator dependency is known for other classes of estimators)

Open Question: Which of the two CRL bounds is smaller?

Intuitively, Oracle-based estimator seems to have more information than the reference signal based estimator. But is this true/quantifiable?

Easy case: $CRLB_{Setup\ 1} \to \infty$ as $\xi \perp \zeta_0$. 
Thank you! Merci!

Questions?
Problem Statement

Existing results: FIM

References


E. Candés, T. Strohmer, V. Voroninski, PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex

