Frames and Phase Retrieval

Radu Balan
University of Maryland
Department of Mathematics and Center for Scientific Computation and Mathematical Modeling

AMS Short Course: JMM San Antonio, TX
Current Collaborators and Sponsors

• Mark Lai (UMD)
• Max Scharrenbroich (UMD)
• Dongmian Zou (UMD)

• Yang Wang (MSU)
• Heiko Claussen (Siemens)
• Justinian Rosca (Siemens)
Overview

• Problem Statement
• Analysis Results
  – Injectivity Results
  – Deterministic Bounds: Lipschitz bounds
  – Stochastic Bounds: CRLB
• The IRLS Algorithm
• Simulation Results
1. Problem Statement

• Consider the following nonlinear map:

\[ \alpha: H \rightarrow \mathbb{R}^m, (\alpha(x))_k = |\langle x, f_k \rangle| \]

Where \( f_1, \ldots, f_m \in H \) is a frame (spanning set) and \( H = \mathbb{R}^n \) or \( \mathbb{C}^n \).

• Replace \( H \) by \( \hat{H} = H/\sim \) where \( x \sim y \) iff there is a unit scalar \( z, |z| = 1 \), so that \( y = zx \).

• Nonlinear map is well defined:

\[ \alpha: \hat{H} \rightarrow \mathbb{R}^m, \alpha(x) = (|\langle x, f_k \rangle|)_{k=1}^m \]
Nonlinear map:
\[ \alpha: \hat{H} \rightarrow \mathbb{R}^m, \alpha(x) = (|\langle x, f_k \rangle|)_{k=1}^m \]

Problems:
1. **Invertibility**: When is \( \alpha \) injective?
2. **Algorithms**: How to efficiently invert \( \alpha \) ?

**Definition**: The frame \( \mathcal{F} = \{f_1, f_2, \ldots, f_m\} \) is said **phase retrievable** (or that it gives **phase retrieval**) if the nonlinear map \( \alpha \) is injective.
Relevance

- X-Ray Crystallography
- Audio and Signal processing
- Fiber optics data transmission
- THz Radar & imaging systems
- Beampattern design

\[ I(\vec{k}) = C \left| \int e^{i\vec{k} \cdot \vec{r}} R(\vec{r}) d\vec{r} \right|^2 \]

2. Invertibility: Real Case

Let
\[ \alpha : \mathbb{R}^n \to \mathbb{R}^m, \alpha(x) = (|\langle x, f_k \rangle|)_{k=1}^m = |F^* x| \]
where
\[ F = [f_1 \mid \ldots \mid f_m] \in \mathbb{R}^{n \times m}, \]
and let
\[ R(x) = \sum_{k=1}^m |\langle x, f_k \rangle|^2 f_k f_k^*. \]

**Theorem** [B.CasazzaEdidin’06, B.’12] The following are equivalent:

1. The frame $\mathcal{F}$ is phase retrievable
2. For any disjoint partition $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ either $\mathcal{F}_1$ or $\mathcal{F}_2$ spans $H$
3. There is a constant $a_0 > 0$ so that $R(x) \geq a_0 \|x\|^2 I_n$.

**Corollary** [BCE’06] If $\mathcal{F}$ is phase retrievable then $m \geq 2n - 1$. 
2. Invertibility: Complex case

- **Notations**: $\mathcal{F} = \{f_1, ..., f_m\} \subset \mathbb{C}^n$

\[
F_k = f_k f_k^* \leftrightarrow f_k \leftrightarrow \varphi_k = \begin{bmatrix}
\text{real}(f_k) \\
\text{imag}(f_k)
\end{bmatrix} \rightarrow \Phi_k = \varphi_k \varphi_k^T + J \varphi_k \varphi_k^T J^T
\]

where $J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$.

Note $\Phi_k$ is a rank-2 $2n \times 2n$ real symmetric matrix.

For any two vectors $u, v$ the symmetric outer product:

\[
[u, v] = \frac{1}{2} (uv^* + vu^*)
\]

For any $\mathbb{C}^n \ni x \leftrightarrow \xi = \begin{bmatrix} \text{real}(x) \\ \text{imag}(x) \end{bmatrix} \in \mathbb{R}^{2n}$ denote

\[
R(x) = R(\xi) = \sum_{k=1}^{m} \Phi_k \xi \xi^T \Phi_k = \sum_{k=1}^{m} [\Phi_k \xi, \Phi_k \xi]
\]
2: Invertibility: Complex Case (2)

**Theorem** [B’13,BandeiraCahillMixonNelson’13] TFAE:
1. The frame $\mathcal{F}$ is phase retrievable
2. There is a constant $a_0 > 0$ so that for all $u, v \in \mathbb{C}^n$
   \[
   \sum_{k=1}^{m} \left( \text{real}(\langle u, f_k \rangle \langle f_k, v \rangle) \right)^2 = \sum_{k=1}^{m} |\langle F_k, [u, v] \rangle|^2 \geq a_0 \| [u, v] \|_1^2 = a_0 \left[ \| u \|^2 \| v \|^2 - (\text{imag}(\langle u, v \rangle))^2 \right].
   \]
3. For any $\xi \neq 0$, $\text{rank}(V(\xi))=2n-1$, where $V(\xi) = [\Phi_1 \xi | \ldots | \Phi_m \xi]$;
4. There is a constant $a_0 > 0$ so that for all $\xi \in \mathbb{R}^{2n}$
   \[
   R(\xi) = \sum_{k=1}^{m} \Phi_k \xi \xi^T \Phi_k \geq a_0 (\| \xi \|^2 I_{2n} - J \xi \xi^T J^T) = a_0 \| \xi \|^2 P_{J^T \xi}^+.\]

**Remark.** Critical redundancy is $4n - O(\log(n)) \leq m \leq 4n - 2(3)$
3. Robustness Measures: CRLB

Consider the following additive white Gaussian noise (AWGN) model:
\[ y_k = \alpha^2(x)_k + \nu_k = |\langle x, f_k \rangle|^2 + \nu_k, \quad 1 \leq k \leq m \]
where \( \nu_k \sim \mathcal{N}(0, \sigma^2) \) are i.i.d.

Theorem [B’12, B.’13, BCMN’13] The Fisher Information matrix for \( x \) is given by

\[
I(x) = \frac{4}{\sigma^2} R(x) = \begin{cases} 
\frac{4}{\sigma^2} \sum_{k=1}^{m} [F_k x, F_k x] & H = \mathbb{R}^n \\
\frac{4}{\sigma^2} \sum_{k=1}^{m} [\Phi_k \xi, \Phi_k \xi] & H = \mathbb{C}^n 
\end{cases}
\]
3. Robustness Measures: CRLB (2)

The No-Oracle Case:

**Theorem [B’13, BCMN’13]** Fix $z_0 \in H$ and denote $\Omega_{z_0} = \{ x \in H : \langle x, z_0 \rangle > 0 \}$. Then the covariance matrix of any unbiased estimator $\omega : \mathbb{R}^m \rightarrow \Omega_{z_0} \subset H$ of an unknown vector $x$ known to be in $\Omega_{z_0}, x \in \Omega_{z_0}$, is bounded below by:

- (Real Case): $\text{Cov}[\omega] \geq \frac{\sigma^2}{4} \left( R(x) \right)^{-1}$
- (Complex Case): $\text{Cov}[\omega] \geq \frac{\sigma^2}{4} \left( R(x) \right)^{+}$

[CRLB Bounds]
3. Robustness Measures: CRLB (3)

The Oracle Case:

\[ y = \alpha^2(x) + \nu \]
\[ \omega_1(y) \]
\[ \omega(y; x) = \omega_1(y)\epsilon(x, \omega_1(y)) \]
\[ \epsilon(x, \omega_1(y)) \]
\[ \text{Oracle} \]
\[ x \]

where: \( \epsilon(x, z) = \arg \min_{|a|=1} \|x - az\| = \frac{\langle x, z \rangle}{|\langle x, z \rangle|} \)

Assume \( \mathbb{E}[\omega(y; x)|x] = x \) (unbiased estimator).

Problem: Find a lower bound for MSE:
\[
MSE = \mathbb{E}[\|x - \omega(y; x)\|^2|x] 
\]

Still an open problem!
3. Robustness Measures: Lipschitz bounds (1)

We insist on exact inversion on the range: \( \omega(\alpha(\hat{x})) = \hat{x} \)
for all \( x \). Then we need to characterize bi-Lipschitz maps.

- We shall consider two distances on \( \widehat{H} \):

1. \( d(x, y) = \min_{|\phi|=1} \| x - \phi y \| \)
2. \( d_1(x, y) = \| [x, x] - [y, y] \|_1 \)

- We want to compute:

\[
\begin{align*}
\inf_{x \neq y} \frac{\|\alpha(x) - \alpha(y)\|}{d(x, y)}, & \sup_{x \neq y} \frac{\|\alpha(x) - \alpha(y)\|}{d(x, y)} \\
\inf_{x \neq y} \frac{\|\alpha^2(x) - \alpha^2(y)\|}{d_1(x, y)}, & \sup_{x \neq y} \frac{\|\alpha^2(x) - \alpha^2(y)\|}{d_1(x, y)}
\end{align*}
\]
3. Robustness Measures: Lipschitz bounds (2)

\[ d_1(x, y) = \| [x, x] - [y, y] \|_1 \]

This distance makes the embedding

\[ \tilde{H} \mapsto Sym(H), \quad x \mapsto [x, x] = xx^* \]

isometric, with nuclear norm on \( Sym \).

Theorem [R.B., Wang '13]. When \( \alpha \) is injective, then \( \alpha^2 \) is bi-Lipschitz: there are positive constants \( 0 < a_0 \leq b_0 \) so that for every \( x, y \in \tilde{H} \)

\[ \sqrt{a_0} d_1(x, y) \leq \| \alpha^2(x) - \alpha^2(y) \| \leq \sqrt{b_0} d_1(x, y) \]

Explicitly:

\[ a_0 \leq \sum_{k=1}^{m} \frac{\| \langle x, f_k \rangle^2 \rangle - \langle y, f_k \rangle^2 \|^2}{\| [x, x] - [y, y] \|_1^2} \leq b_0 \]

Same \( a_0 \) as in CRLB

Mixed \( l^{2,4} \) norm of the analysis operator
3. Robustness Measures: Lipschitz bounds (3)

In the real case:

$$||[x, x] - [y, y]||_1 = ||[x - y, x + y]||_1 = ||x - y|| \cdot ||x + y||$$

Hence:

$$a_0 \leq \frac{\sum_{k=1}^{m} |\langle x, f_k \rangle|^2 - |\langle y, f_k \rangle|^2|^2}{||x - y||^2 \cdot ||x + y||^2} \leq b_0$$

With:  
$$a_0 = \min_{||x||=1} \lambda_{min}(R(x)), \quad b_0 = \max_{||x||=1} \lambda_{max}(R(x))$$

Lower bound similar to [EldarMendelson'12]

In the complex case:

$$||[x, x] - [y, y]||_1 = ||[x - y, x + y]||_1 = \sqrt{||x - y||^2 ||x + y||^2 + 4(imag(\langle x, y \rangle))^2}$$

Hence:

$$a_0 \leq \frac{\sum_{k=1}^{m} |\langle x, f_k \rangle|^2 - |\langle y, f_k \rangle|^2|^2}{||x - y||^2 ||x + y||^2 + 4(imag(\langle x, y \rangle))^2} \leq b_0$$

With:  
$$a_0 = \min_{||x||=1} \lambda_{2n-1}(R(x)), \quad b_0 = \max_{||x||=1} \lambda_{max}(R(x))$$
3. Robustness Measures: Lipschitz bounds (4)

\[ d(x, y) = \min_{|\varphi|=1} \|x - \varphi y\| = \begin{cases} \min(\|x - y\|, \|x + y\|) & , \text{ real case} \\ \sqrt{\|x\|^2 + \|y\|^2 - 2|\langle x, y \rangle|^2} & , \text{ complex case} \end{cases} \]

Theorem [R.B., Wang ‘13]. In the real case \((H = \mathbb{R}^n)\) when \(\alpha\) is injective, then \(\alpha\) is also bi-Lipschitz: there are positive constants \(0 < \rho_\infty \leq \nu_0\) so that for every \(x, y \in \widehat{H}\)

\[
\sqrt{\rho_\infty} d(x, y) \leq \|\alpha(x) - \alpha(y)\| \leq \sqrt{\nu_0} d(x, y)
\]

Explicitly:

\[
\rho_\infty \leq \frac{\Sigma_{k=1}^{m} ||\langle x, f_k \rangle| - |\langle y, f_k \rangle||^2}{\min(\|x - y\|^2, \|x + y\|^2)} \leq \nu_0
\]

with \(\rho_\infty = \min_{S} \left( \sqrt{A[S] + A[S^c]} \right)\) and \(\nu_0 = B\)

where \(A[S]\) denotes the lower frame bound of subset indexed by \(S\), and \(B\) denotes the upper frame bound.

Complex Case: Still Open!
4. The Iterative Regularized Least Squares (IRLS) Algorithm – The complex case (1)

We want to minimize:

\[ J(x) = \sum_{k=1}^{m} |y_k - \langle F_k x, x \rangle|^2 = \sum_{k=1}^{m} |y_k - \langle \Phi_k \xi, \xi \rangle|^2 = J(\xi) \]

Instead we introduce an homotopic optimization procedure that uses regularization and splitting:

\[
J(\xi, \eta; \lambda, \mu) = \sum_{k=1}^{m} |y_k - \langle \Phi_k \xi, \eta \rangle|^2 + \lambda \| \xi \|^2 + \mu \| \xi - \eta \|^2 + \lambda \| \eta \|^2
\]

\[
= \sum_{k=1}^{m} |y_k - \langle F_k u, v \rangle|^2 + \lambda \| u \|^2 + \mu \| u - v \|^2 + \lambda \| v \|^2 = J(u, v; \lambda, \mu)
\]

Main iteration: \( \xi^{t+1} = \text{argmin}_{\eta} J(\eta, \xi^t; \lambda_t, \mu_t) \)
4. The IRLS Algorithm (2)

1. Initialization.

For large $\lambda$, $J(u, u; \lambda, \mu)$ has global optimizer $u=0$. The largest $\lambda$ for a nontrivial optimizer:

$$
\lambda_{\text{max}}(R), R = \sum_{k=1}^{m} y_k F_k
$$

Choose $\rho < 1$. Then initialize:

$$
\mu^0 = \lambda^0 = \rho \lambda_{\text{max}}(R)
$$

$$
u^0 = \sqrt{\frac{(1 - \rho) \lambda_{\text{max}}(R)}{\sum_{k=1}^{m} |\langle e_{\text{max}}, f_k \rangle|^4}} e_{\text{max}} \rightarrow \xi^0 = j(u^0)$$
4. The IRLS Algorithm (3)

2. **Iteration.**

Compute \( \xi^{t+1} = \arg\min_{\eta} J(\eta, \xi^t; \lambda_t, \mu_t) \)

Adapt: \( \lambda^{t+1} = \rho \lambda^t, \mu^{t+1} = \max(\rho \mu^t, \mu_{\text{min}}) \)

Solution to the optimization problem: Solve

\[
\left[ \sum_{k=1}^{m} \Phi_k \xi^T \xi \Phi_k + (\lambda + \mu)I_{2n} \right] \xi^{t+1} = \left( \sum_{k=1}^{m} y_k \Phi_k + \mu I_{2n} \right) \xi^t
\]
4. The IRLS Algorithm (4)

3. Stopping criterion

Stop when regularization parameter $\lambda^t$ reaches a preset minimal value:

$$\lambda^T \leq \lambda^{min}$$

Alternatively: Choose a signal-to-noise level, say SNR, and stop when

$$\frac{||\xi^T||^2}{||y - \alpha(\xi^T)||^2} \leq SNR$$
5. Simulation Results

Results for \( n=100 \), \( m=800 \) (redundancy 8) complex case

\[
SNR = 10 \log_{10} \left( \frac{\sum_{k=1}^{m} |\langle x, f_k \rangle|^2}{\sigma^2} \right)
\]
Redundancy 6: n=100, m=600
Redundancy 4: n=100, m=400
Analysis for Redundancy=8, n=100, m=800

Number of iterations: left plot; Components of the Mean-Square error: Bias and Variance (right plot)
Redundancy 8: Trace of a realization; SNR = 0dB
Redundancy 8: Trace of a realization; SNR = 40dB
6. Conclusions

We analyzed the phase retrieval problem (phaseless reconstruction):

1. We presented necessary and sufficient conditions for injectivity of the nonlinear map;
2. Explicit expressions of the Fisher Information Matrix and Cramer-Rao Lower Bound;
3. Explicit expressions for bi-Lipschitz constants
Thank you!