The Iterative and Regularized Least Squares Algorithm for Phase Retrieval

Radu Balan

Naveed Haghani

Department of Mathematics, AMSC, CSCAMM and NWC
University of Maryland, College Park, MD

April 17, 2016
Special Session "Frames, Wavelets and Gabor Systems"
AMS Regional Meeting, Fargo ND

"This material is based upon work supported by the National Science Foundation under Grant No. DMS-1413249 and by ARO under Contract No. W911NF-16-1-0008. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation."

Table of Contents:

Phase Retrieval

The phase retrieval problem

Hilbert space $H = \mathbb{C}^n$, $\hat{H} = H/T^1$, frame $\mathcal{F} = \{f_1, \cdots, f_m\} \subset \mathbb{C}^n$ and measurements

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

The frame is said *phase retrievable* (or that it gives phase retrieval) if $\hat{x} \mapsto (|\langle x, f_k \rangle|)_{1 \leq k \leq m}$ is injective.

The general *phase retrieval problem* a.k.a. *phaseless reconstruction*: Decide when a given frame is phase retrievable, and, if so, find an algorithm to recover x from $y = (y_k)_k$ up to a global phase factor.

Our problem today: A reconstruction algorithm.

General Purpose Algorithms

Unstructured Frames. Unstructured Data

1. Iterative Algorithms:
   - Gerchberg-Saxton [Gerchberg&all]
   - Wirtinger flow - gradient descent [CLS14]
   - IRLS [B13]
2. Rank 1 Tensor Recovery:
   - PhaseLift; PhaseCut [Candes&all];[Waldspurger&all]
   - Higher-Order Tensor Recovery [B.]

Specialized Algorithms

Structured Frames and/or Structured Data

1. Structured Frames:
   - Fourier Frames: 4n-4 [BH13]; Masking DFT [CLS13]; STFT/Spectograms [B.][Eldar&all][Hayes&all]; Alternating Projections [GriffinLim][Fannjiang]
   - Polarization: 3-term [ABFM12], masking [BCM]
   - Shift-Invariant Spaces: Bandlimited [Thakur]; Filterbanks/Circulant Matrices [IVW2]; Other spaces [Chen&all]
   - X-Ray Crystallography – over 100 years old, lots of Nobel prizes ...
2. Special Signals:
   - Sparse general case: GESPAR[SBE14];
   - Specialized: sparse [IVW1]; speech [ARF03]

... and others – "phase retrieval" in title: 2680 papers

The IRLS Algorithm

First Motivation: Graduation Method. Homotopic Continuation

The IRLS algorithm belongs to the class of *Graduation Methods* , or *Homotopic Continuations*.

Idea:

The IRLS Algorithm

First Motivation: Graduation Method. Homotopic Continuation

The IRLS algorithm belongs to the class of *Graduation Methods* , or *Homotopic Continuations*.

Idea:

Our target is to optimize a complicated (possibly non-convex) optimization criterion $J(x)$, $argmin_{x \in D} J(x)$.

The IRLS Algorithm

First Motivation: Graduation Method. Homotopic Continuation

The IRLS algorithm belongs to the class of *Graduation Methods* , or *Homotopic Continuations*.

Idea:

Our target is to optimize a complicated (possibly non-convex) optimization criterion $J(x)$, $argmin_{x \in D} J(x)$.

However we know how to optimize a closely related criterion $J_0(x)$, $argmin_{x \in D_0} J_0(x)$.

The IRLS Algorithm

First Motivation: Graduation Method. Homotopic Continuation

The IRLS algorithm belongs to the class of *Graduation Methods* , or *Homotopic Continuations*.

Idea:

Our target is to optimize a complicated (possibly non-convex) optimization criterion $J(x)$, $argmin_{x \in D} J(x)$.

However we know how to optimize a closely related criterion $J_0(x)$, $argmin_{x \in D_0} J_0(x)$.

Then we introduce a monotonic sequence $0 \leq t_n \leq 1$ with $t_0 = 1$ and $t_n \to 0$ and solve iteratively

$$x^{n+1} = argmin_{x \in D_n} F(t_n, J(x), J_0(x))$$

using x^n as starting point. Here F is a continue function so that $F(1, J(x), J_0(x)) = J_0(x)$ and $F(0, J(x), J_0(x)) = J(x)$.

The IRLS Algorithm

LARS Algorithm

Least Angle Regression (LARS) [EHJT04] designed to solve LASSO, or variants:

$$argmin_x \|y - Ax\|_2^2 + \lambda \|x\|_1$$

The IRLS Algorithm

LARS Algorithm

Least Angle Regression (LARS) [EHJT04] designed to solve LASSO, or variants:

$$argmin_x \|y - Ax\|_2^2 + \lambda \|x\|_1$$

It is proved the optimizer $x_{opt} = x(\lambda)$ is a continuous and piecewise differentiable function of λ (linear, in the case of LASSO).

The IRLS Algorithm

LARS Algorithm

Least Angle Regression (LARS) [EHJT04] designed to solve LASSO, or variants:

$$argmin_x \|y - Ax\|_2^2 + \lambda \|x\|_1$$

It is proved the optimizer $x_{opt} = x(\lambda)$ is a continuous and piecewise differentiable function of λ (linear, in the case of LASSO).
Method: Start with $\lambda = \lambda_0 = \frac{2}{\|A^T y\|_2}$ and the optimal solution is $x^0 = 0$.

The IRLS Algorithm

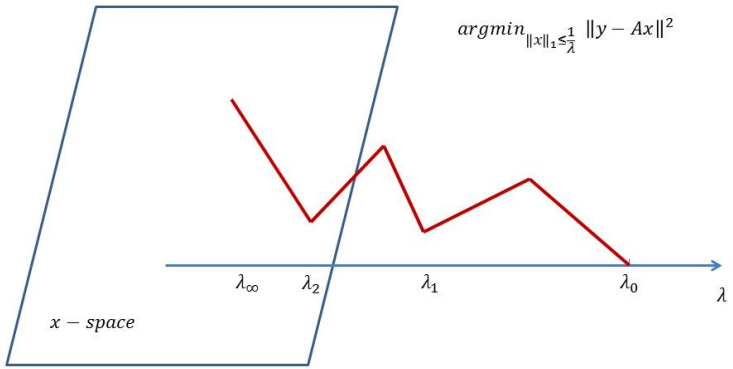
LARS Algorithm

Least Angle Regression (LARS) [EHJT04] designed to solve LASSO, or variants:

$$argmin_x \|y - Ax\|_2^2 + \lambda \|x\|_1$$

It is proved the optimizer $x_{opt} = x(\lambda)$ is a continuous and piecewise differentiable function of λ (linear, in the case of LASSO).
Method: Start with $\lambda = \lambda_0 = \frac{2}{\|A^T y\|_2}$ and the optimal solution is $x^0 = 0$.

Then LARS finds monotonically decreasing λ values where the slope (and support) of $x(\lambda)$ changes. The algorithm ends at the desired value of $\lambda = \lambda_\infty$ (see also Hierarchical Decompositions of Tadmor&all).



The IRLS Algorithm

IRLS Algorithm

The Iterative Regularized Least-Squares Algorithm attempts to fnd the global minimum of the non-convex problem

$$argmin_x \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda_\infty \|x\|_2^2$$

The IRLS Algorithm

IRLS Algorithm

The Iterative Regularized Least-Squares Algorithm attempts to fnd the global minimum of the non-convex problem

$$argmin_x \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda_\infty \|x\|_2^2$$

using a sequence of iterative least-squares problems:

$$x^{(t+1)} = argmin_x \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|^2$$

The IRLS Algorithm

IRLS Algorithm

The Iterative Regularized Least-Squares Algorithm attempts to fnd the global minimum of the non-convex problem

$$argmin_x \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda_\infty \|x\|_2^2$$

using a sequence of iterative least-squares problems:

$$x^{(t+1)} = argmin_x \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|^2$$

together with a polarization relaxation:

$$|\langle x, f_k \rangle|^2 \approx \frac{1}{2}(\langle x, f_k \rangle \langle f_k, x^{(t)} \rangle + \langle x^{(t)}, f_k \rangle \langle f_k, x \rangle)$$

The IRLS Algorithm

Main Optimization

The optimization problem:

$$
\begin{aligned}
x^{(t+1)} &= argmin_x \sum_{k=1}^{m} \left| y_k - \frac{1}{2}(\langle x, f_k\rangle\langle f_k, x^{(t)}\rangle + \langle x^{(t)}, f_k\rangle\langle f_k, x\rangle) \right|^2 + \\
&\quad + \lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|_2^2 + \lambda_t \|x^{(t)}\|_2^2 \\
&= argmin_x \;\; J(x, x^{(t)}; \lambda, \mu)
\end{aligned}
$$

The IRLS Algorithm

Main Optimization

The optimization problem:

$$
\begin{aligned}
x^{(t+1)} &= argmin_x \sum_{k=1}^{m} \left| y_k - \frac{1}{2}(\langle x, f_k\rangle\langle f_k, x^{(t)}\rangle + \langle x^{(t)}, f_k\rangle\langle f_k, x\rangle) \right|^2 + \\
&\quad + \lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|_2^2 + \lambda_t \|x^{(t)}\|_2^2 \\
&= argmin_x \;\; J(x, x^{(t)}; \lambda, \mu)
\end{aligned}
$$

Note:

The IRLS Algorithm

Main Optimization

The optimization problem:

$$
\begin{aligned}
x^{(t+1)} &= argmin_x \sum_{k=1}^{m} \left| y_k - \frac{1}{2}(\langle x, f_k\rangle\langle f_k, x^{(t)}\rangle + \langle x^{(t)}, f_k\rangle\langle f_k, x\rangle) \right|^2 + \\
&\quad + \lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|_2^2 + \lambda_t \|x^{(t)}\|_2^2 \\
&= argmin_x \;\; J(x, x^{(t)}; \lambda, \mu)
\end{aligned}
$$

Note:

- $J(x, .; ., .)$ is quadratic in $x \Rightarrow$ hence a least-squares problem!

The IRLS Algorithm

Main Optimization

The optimization problem:

$$
\begin{aligned}
x^{(t+1)} &= argmin_x \sum_{k=1}^{m} \left| y_k - \frac{1}{2}(\langle x, f_k\rangle\langle f_k, x^{(t)}\rangle + \langle x^{(t)}, f_k\rangle\langle f_k, x\rangle) \right|^2 + \\
&\quad + \lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|_2^2 + \lambda_t \|x^{(t)}\|_2^2 \\
&= argmin_x \;\; J(x, x^{(t)}; \lambda, \mu)
\end{aligned}
$$

Note:

- $J(x, .; ., .)$ is quadratic in $x \Rightarrow$ hence a least-squares problem!
- $J(x, x; \lambda, \mu) = \sum_{k=1}^{m} |y_k - |\langle x, f_k\rangle|^2|^2 + 2\lambda\|x\|_2^2 \Rightarrow$ Fixed points of IRLS are local minima of the original problem.

The IRLS Algorithm

Second Motivation: Relaxation of Constraints

Another motivation: seek $X = xx^*$ that solves

$$\min_{X \geq 0, rank(X)=1} \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda trace(X).$$

The IRLS Algorithm

Second Motivation: Relaxation of Constraints

Another motivation: seek $X = xx^*$ that solves

$$\min_{X \geq 0, rank(X)=1} \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda trace(X).$$

PhaseLift algorithm removes the condition $rank(X) = 1$ and shows (for large λ) this produces the desired result with high probability.

The IRLS Algorithm

Second Motivation: Relaxation of Constraints

Another motivation: seek $X = xx^*$ that solves

$$\min_{X \geq 0, rank(X)=1} \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda trace(X).$$

PhaseLift algorithm removes the condition $rank(X) = 1$ and shows (for large λ) this produces the desired result with high probability.
Another way to relax the problem is to search for X in a larger space. The IRLS is essentially equivalent to optimize a convex functional of X on the larger space

$$\mathcal{S}^{1,1} = \{T = T^* \in \mathbb{C}^{n \times n}, \ T \text{ has at most one positive eigenvalue and at most one negative eigenvalue}\}.$$

The IRLS Algorithm

Second Formulation

Consider the following three convex criteria:

$$
\begin{aligned}
J_1(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2(\lambda+\mu)\|X\|_1 - 2\mu trace(X) \\
J_2(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda eig_{max}(X) - (2\lambda + 4\mu) eig_{min}(X) \\
J_3(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda\|X\|_1 - 4\mu eig_{min}(X)
\end{aligned}
$$

which coincide on $\mathcal{S}^{1,1}$.

The IRLS Algorithm

Second Formulation

Consider the following three convex criteria:

$$
\begin{aligned}
J_1(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2(\lambda+\mu)\|X\|_1 - 2\mu trace(X) \\
J_2(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda eig_{max}(X) - (2\lambda + 4\mu) eig_{min}(X) \\
J_3(X;\lambda,\mu) &= \sum_{k=1}^{m} |y_k - \langle X, f_k f_k^* \rangle_{HS}|^2 + 2\lambda\|X\|_1 - 4\mu eig_{min}(X)
\end{aligned}
$$

which coincide on $\mathcal{S}^{1,1}$. Consider the optimization problem

$$
(Jopt, X) = \min_{X \in \mathcal{S}^{1,1}} J_k(X;\lambda,\mu) \quad , \quad 1 \le k \le 3
$$

The IRLS Algorithm

Second Formulation -2

The following are true:

1. Optimization in $\mathcal{S}^{1,1}$:

$$\min_{X \in \mathcal{S}^{1,1}} J_k(X; \lambda, \mu) = \min_{u,v \in \mathbb{C}^n} J(u, v; \lambda, \mu)$$

If $\hat{X}$ and $(\hat{u}, \hat{v})$ denote optimizers so that $imag(\langle \hat{u}, \hat{v} \rangle) = 0$, then $\hat{X} = \frac{1}{2}(\hat{u}\hat{v}^* + \hat{v}\hat{u}^*)$.

The IRLS Algorithm

Second Formulation -2

The following are true:

1. Optimization in $\mathcal{S}^{1,1}$:

$$\min_{X \in \mathcal{S}^{1,1}} J_k(X; \lambda, \mu) = \min_{u,v \in \mathbb{C}^n} J(u, v; \lambda, \mu)$$

   If $\hat{X}$ and $(\hat{u}, \hat{v})$ denote optimizers so that $imag(\langle \hat{u}, \hat{v} \rangle) = 0$, then $\hat{X} = \frac{1}{2}(\hat{u}\hat{v}^* + \hat{v}\hat{u}^*)$.

2. Optimization in $\mathcal{S}^{1,0}$:

$$\min_{X \in \mathcal{S}^{1,0}} J_k(X; \lambda, \mu) = \min_{x \in \mathbb{C}^n} J(x, x; \lambda, \mu)$$

   If $\hat{X}$ and $\hat{x}$ denote optimizers, then $\hat{X} = \hat{x}\hat{x}^*$. $\mathcal{S}^{1,0} = \{xx^*\}$.

The IRLS Algorithm

Initialization

For $\lambda \geq eig_{max}(R(y))$, where $R(y) = \sum_{k=1}^{m} y_k f_k f_k^*$, $J(x;\lambda) = \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2 + 2\lambda \|x\|_2^2$ is convex. The unique global minimum is $x^0 = 0$.

The IRLS Algorithm

Initialization

For $\lambda \geq eig_{max}(R(y))$, where $R(y) = \sum_{k=1}^{m} y_k f_k f_k^*$, $J(x;\lambda) = \sum_{k=1}^{m} |y_k - |\langle x, f_k\rangle|^2|^2 + 2\lambda\|x\|_2^2$ is convex. The unique global minimum is $x^0 = 0$.

Initialization Procedure:

- Solve the principal eigenpair (e, eig_{max}) of matrix $R(y)$ using e.g. the power method;
- Set

$$\lambda_0 = (1-\varepsilon)eig_{max} \ , \ x^0 = \sqrt{\frac{(1-\varepsilon)eig_{max}}{\sum_{k=1}^{m} |\langle e, f_k\rangle|^4}}\, e.$$

  Here $\varepsilon > 0$ is a parameter that depends on the frame set as well as the spectral gap of $R(y)$.
- Set $\mu_0 = \lambda_0$ and $t = 0$.

The IRLS Algorithm

Iterations

Repeat the following steps until stopping:

- Optimization: Solve the least-square problem:

$$
\begin{aligned}
x^{(t+1)} &= argmin_x \sum_{k=1}^{m} \left| y_k - \frac{1}{2}(\langle x, f_k\rangle\langle f_k, x^{(t)}\rangle + \langle x^{(t)}, f_k\rangle\langle f_k, x\rangle)\right|^2 + \\
&\quad +\lambda_t \|x\|_2^2 + \mu_t \|x - x^{(t)}\|_2^2 + \lambda_t \|x^{(t)}\|_2^2 \\
&= argmin_x \;\; J(x, x^{(t)}; \lambda, \mu)
\end{aligned}
$$

- Update: $\lambda_{t+1} = \gamma\lambda_t$, $\mu_{t+1} = \max(\gamma\mu_t, \mu^{\min})$, $t = t + 1$. Here γ is the learning rate, and μ^{min} is related to performance.

The IRLS Algorithm

Performance

Let $y_k = |\langle x, f_k \rangle|^2 + \nu_k$. Assume the algorithm is stopped at some T so that

$$J(x^{(T)}, x^{(T-1)}; \lambda, \mu) \leq J(x, x; \lambda, \mu).$$

Denote $\hat{X} = \frac{1}{2}(x^{(T)} x^{(T-1)*} + x^{(T-1)} x^{(T)*})$ and $\hat{x}\hat{x}^* = P_+(\hat{X})$.

The IRLS Algorithm

Performance

Let $y_k = |\langle x, f_k \rangle|^2 + \nu_k$. Assume the algorithm is stopped at some T so that

$$J(x^{(T)}, x^{(T-1)}; \lambda, \mu) \leq J(x, x; \lambda, \mu).$$

Denote $\hat{X} = \frac{1}{2}(x^{(T)}x^{(T-1)*} + x^{(T-1)}x^{(T)*})$ and $\hat{x}\hat{x}^* = P_+(\hat{X})$.
Then the following hold true:

1. Matrix norm error:

$$\|\hat{X} - xx^*\|_1 \leq \frac{\lambda}{C_0} + \sqrt{C_0}\|\nu\|$$

2. Natural distance:

$$D(\hat{x}, x)^2 = \|\hat{X} - xx^*\|_1 + |eig_{min}(\hat{X})| \leq \frac{\lambda}{C_0} + \sqrt{C_0}\|\nu\| + \frac{\|\nu\|^2}{4\mu} + \frac{\lambda\|x\|^2}{2\mu}$$

where C_0 is a frame dependent constant (lower Lipschitz constant in $\mathcal{S}^{1,1}$).

Numerical Simulations

Setup

The algorithm requires $O(m)$ memory. Simulations with $m = Rn$ (complex case) with $n = 1000$ and $R \in \{4, 6, 8, 12\}$. Frame vectors corresponding to masked (windowed) DFT:

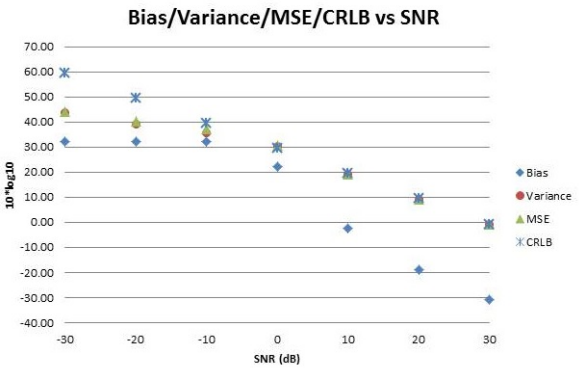
$$f_{jn+k} = \frac{1}{\sqrt{8n}} \left(w_l^j e^{2\pi i k(l-1)/n} \right)_{0 \le l \le n-1}, \quad 1 \le j \le R, 1 \le k \le n$$

$$\begin{bmatrix} f_1 & f_2 & \cdots & f_m \end{bmatrix} = \begin{bmatrix} Diag(w^1) & \cdots & Diag(w^R) \end{bmatrix} \begin{bmatrix} DFT_n & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & DFT_n \end{bmatrix}$$

Parameters: $\varepsilon = 0.1$, $\gamma = 0.95$, $\mu^{min} = \frac{\mu^0}{10}$. Power method tolerance: 10^{-8} Conjugate gradient tolerance: 10^{-14}.

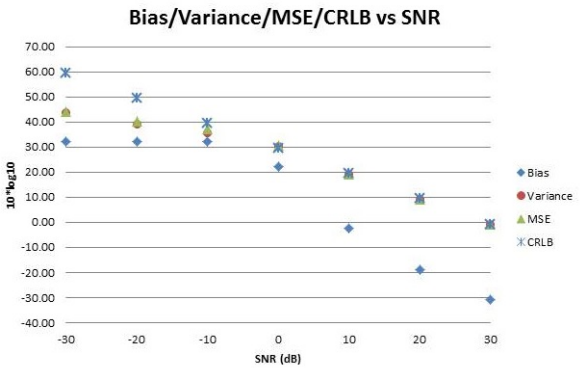
Numerical Simulations

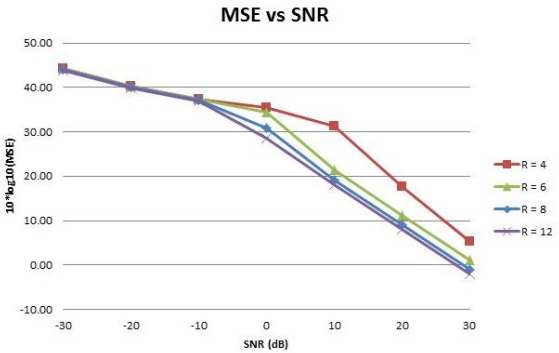
MSE Plots

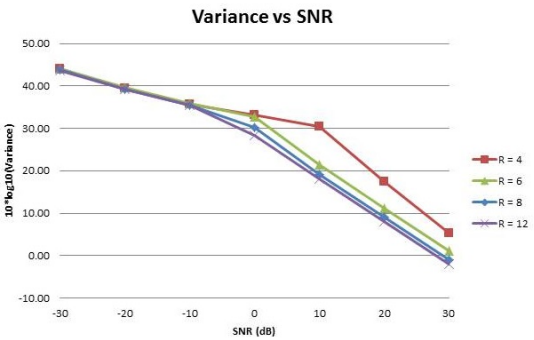


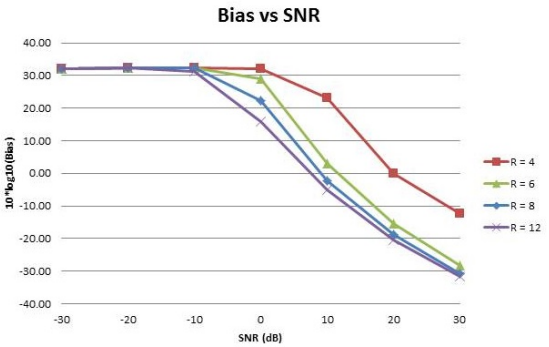
Numerical Simulations

MSE Plots



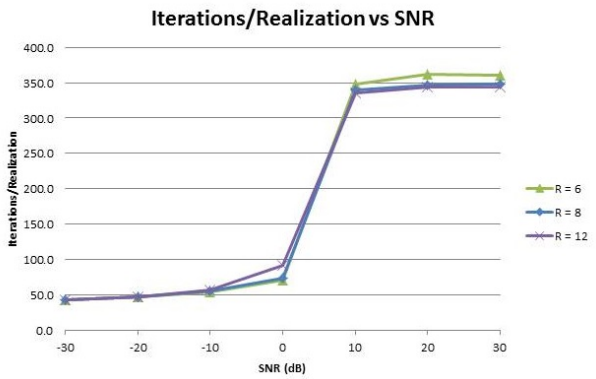


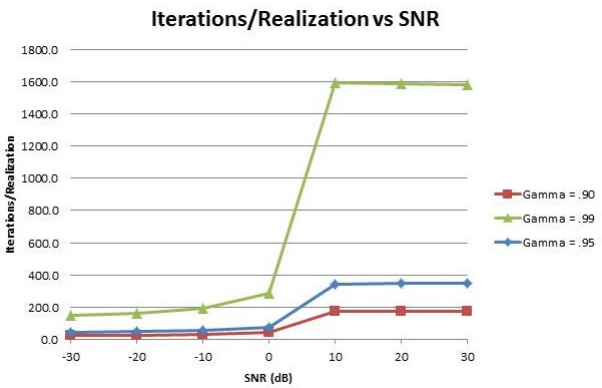




Numerical Simulations

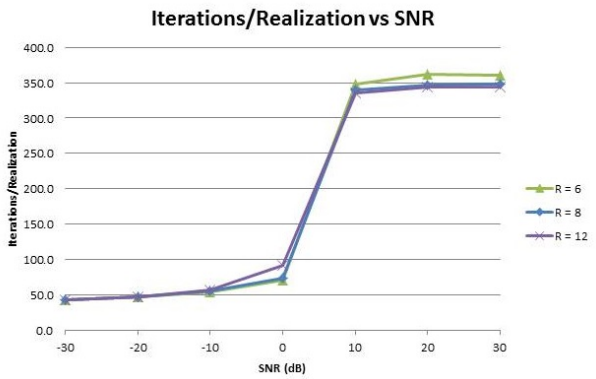
Performance

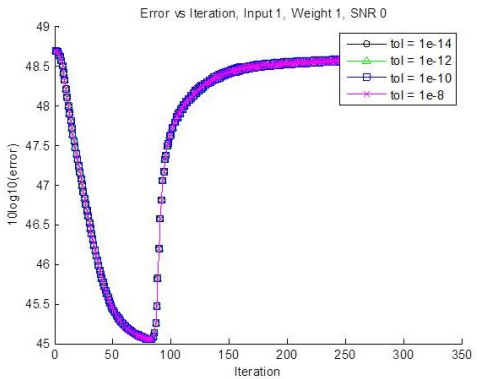


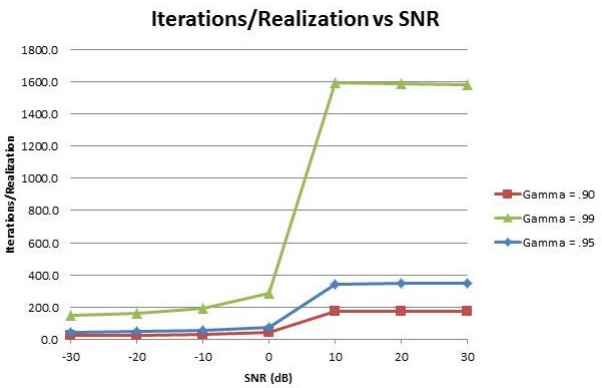


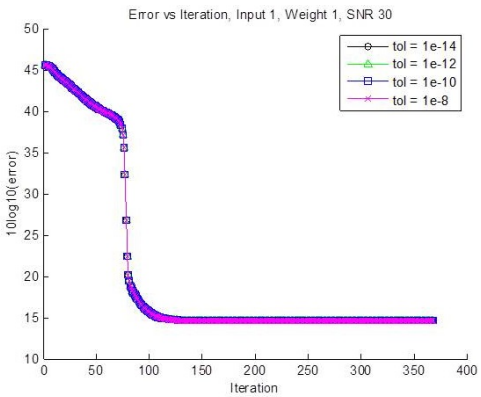
Numerical Simulations

Performance



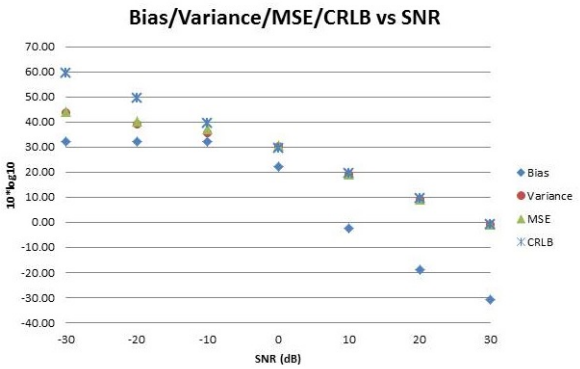






Numerical Simulations

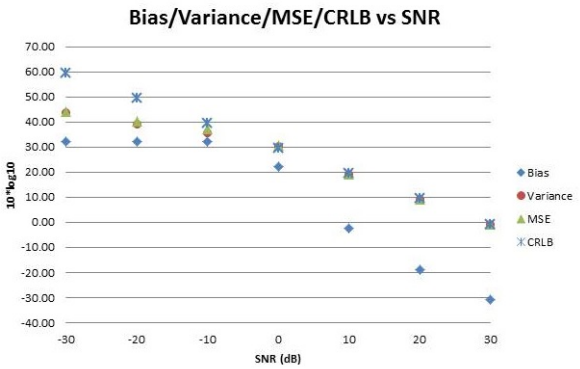
Performance - 2



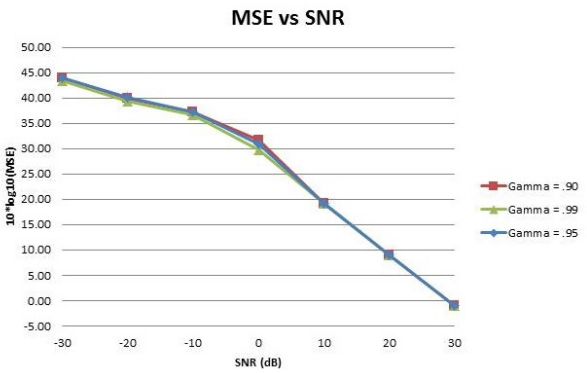
SNR	$10 \cdot log_{10}(Bias)$	$10 \cdot log_{10}(Variance)$	$10 \cdot log_{10}(MSE)$	CRLB
-30	32.13	43.78	44.07	59.39
-20	32.39	39.29	40.09	49.39
-10	32.27	35.56	37.23	39.39
0	22.17	30.24	30.87	29.39
10	-2.21	19.16	19.19	19.39
20	-18.88	9.05	9.05	9.39
30	-30.63	-0.96	-0.96	-0.61

Numerical Simulations

Performance - 2



SNR	$10 \cdot log_{10}(Bias)$	$10 \cdot log_{10}(Variance)$	$10 \cdot log_{10}(MSE)$	CRLB
-30	32.13	43.78	44.07	59.39
-20	32.39	39.29	40.09	49.39
-10	32.27	35.56	37.23	39.39
0	22.17	30.24	30.87	29.39
10	-2.21	19.16	19.19	19.39
20	-18.88	9.05	9.05	9.39
30	-30.63	-0.96	-0.96	-0.61



References

- K. Achan, S.T. Roweis, B.J. Frey, *Probabilistic Inference of Speech Signals from Phaseless Spectrograms*, NIPS 2003.
- B. Alexeev, A. S. Bandeira, M. Fickus, D. G. Mixon, *Phase Retrieval with Polarization*, SIAM J. Imaging Sci., **7** (1) (2014), 35–66.
- R. Balan, P. Casazza, D. Edidin, On signal reconstruction without phase, Appl.Comput.Harmon.Anal. **20** (2006), 345–356.
- R. Balan, B. Bodmann, P. Casazza, D. Edidin, Painless reconstruction from Magnitudes of Frame Coefficients, J.Fourier Anal.Applic., **15** (4) (2009), 488–501.
- R. Balan, Reconstruction of Signals from Magnitudes of Frame Representations, arXiv submission arXiv:1207.1134 (2012).
- 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

R. Balan, The Fisher Information Matrix and the Cramer-Rao Lower Bound in a Non-Additive White Gaussian Noise Model for the Phase Retrieval Problem, proceedings of SampTA 2015.

A.S. Bandeira, Y. Chen, D.G. Mixon, *Phase Retrieval from Power Spectra of Masked Signals*, arXiv:1303.4458v1 (2013).

B. G. Bodmann and N. Hammen, *Stable Phase Retrieval with Low-Redundancy Frames*, available online arXiv:1302.5487v1. Adv. Comput. Math., accepted 10 April 2014.

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

- E. Candés, Y. Eldar, T. Strohmer, V. Voroninski, *Phase Retrieval via Matrix Completion Problem*, SIAM J. Imaging Sci., **6**(1) (2013), 199–225.
- E. Candès, X. Li, *Solving Quadratic Equations Via PhaseLift When There Are As Many Equations As Unknowns*, available online arXiv:1208.6247
- E. Candès, X. Li, M. Soltanolkotabi, *Phase Retrieval from Coded Diffraction Patterns*,
- E. Candès, X. Li and M. Soltanolkotabi, *Phase retrieval via Wirtinger flow: theory and algorithms*, IEEE Transactions on Information Theory 61(4), (2014) 1985–2007.
- Yang Chen, Cheng Cheng, Qiyu Sun and Haichao Wang, Phase Retrieval of Real-Valued Signals in a Shift-Invariant Space, arXiv:1603.01592 (2016).

- Y. C. Eldar, P. Sidorenko, D. G. Mixon, S. Barel and O. Cohen, Sparse phase retrieval from short-time Fourier measurements, IEEE Signal Processing Letters 22, no. 5 (2015): 638-642.
- B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, *Least Angle Regression*, The Annals of Statistics, vol. **32**(2), 407–499 (2004).
- A. Fannjiang, W. Liao, *Compressed Sensing Phase Retrieval*, Asilomar 2011.
- M. Fickus, D.G. Mixon, A.A. Nelson, Y. Wang, *Phase retrieval from very few measurements*, available online arXiv:1307.7176v1. Linear Algebra and its Applications **449** (2014), 475–499
- J.R. Fienup. *Phase retrieval algorithms: A comparison*, Applied Optics, 21(15):2758–2768, 1982.

- R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik 35, 237 (1972).
- D. Griffin and J.S. Lim, *Signal Estimation from Modified Short-Time Fourier Transform*, ICASSP 83, Boston, April 1983.
- M. H. Hayes, J. S. Lim, and A. V. Oppenheim, *Signal Reconstruction from Phase and Magnitude*, IEEE Trans. ASSP **28**, no.6 (1980), 672–680.
- M. Iwen, A. Viswanathan, Y. Wang, *Robust Sparse Phase Retrieval Made Easy*, preprint.
- M. Iwen, A. Viswanathan, Y. Wang, *Fast Phase Retrieval for High-Dimensions*, preprint.
- [Overview2] K. Jaganathany Y.Eldar B.Hassibiy, Phase Retrieval: An Overview of Recent Developments, arXiv:1510.07713 (2015).

- Y. Shechtman, A. Beck and Y. C. Eldar, GESPAR: Efficient phase retrieval of sparse signals, IEEE Transactions on Signal Processing 62, no. 4 (2014): 928-938.
- J. Sun, Q. Qu, J. Wright, *A Geometric Analysis of Phase Retrieval*, preprint 2016.
- G. Thakur, Reconstruction of bandlimited functions from unsigned samples, J. Fourier Anal. Appl., 17(2011), 720–732.