Algorithms for phase retrieval problems
Part II

Irène Waldspurger

CNRS et CEREMADE (Université Paris Dauphine)
Équipe MOKAPLAN (INRIA)

11 décembre 2017

Journées de géométrie algorithmique
Aussois

Setting

Phase retrieval in finite dimension :

$$\boxed{\text{Reconstruct } x \in \mathbb{C}^n \text{ from } (|\langle x, f_k\rangle|)_{1\le k\le m}\,?}$$

(The $f_1, \dots, f_m$ are known elements of $\mathbb{C}^n$.)

Setting

Phase retrieval in finite dimension :

$$\boxed{\text{Reconstruct } x \in \mathbb{C}^n \text{ from } (|\langle x, f_k\rangle|)_{1\le k\le m}\,?}$$

(The $f_1, \dots, f_m$ are known elements of $\mathbb{C}^n$.)

We assume the f_k's to be random realizations of independent normal distributions :

$$f_k \overset{i.i.d.}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathrm{Id}_n), \quad k = 1, \dots, m.$$

When $m \geq \alpha n$, for some fixed constant α,
the phase retrieval problem is well-posed with high probability.

Which reconstruction algorithms can we use ?

We want algorithms that have the two following properties :

1. They are practical to use.
   - Reasonably fast.
   - Stable to noise.
   - (Ideally,) not too complex to implement.
2. It is possible to rigorously prove that they succeed (with high probability).

Reconstruct $x \in \mathbb{C}^n$ from $(|\langle x, f_k \rangle|)_{1 \leq k \leq m}$?

This is a non-convex problem.

To oversimplify, problems that involve only convex constraints can be efficiently solved.

But for $k \leq m$, $b_k \in \mathbb{R}^+$, the constraint

$$|\langle x, f_k \rangle| = b_k,$$

defines a non-convex set in $\mathbb{C}^n$.

Traditional algorithms

Several phase retrieval algorithms were proposed from the 70s.

They typically relied on simple heuristics.

Numerically, they were shown to perform well in some cases.

But because of non-convexity, they could also get stuck in "local minima" and fail to solve the problem.

No theoretical understanding of when they succeeded and when they did not.

Convexification methods

The picture changed with the introduction of convexification methods.

Principle

- Approximate the problem by a convex one ("easy" to solve).
- Prove that the original problem and the approximated one actually have the same solution.

We will present two convexification methods :

- PhaseLift ($\sim$ 2011)
- PhaseMax ($\sim$ 2015)

Lifting

Reconstruct x
from $(|\langle x, f_k \rangle|)_{k \leq m}$.

Lifting

$(x^* \stackrel{def}{=} \overline{x^T})$

$(n \times n \text{ matrix})$

Reconstruct x from $(|\langle x, f_k \rangle|)_{k \leq m}$. $\iff$ Reconstruct xx^* from $(|\langle x, f_k \rangle|^2)_{k \leq m}$.

Lifting

$= (f_k^* x)(x^* f_k)$ $\qquad (x^* \stackrel{def}{=} \overline{x^T})$

$(n \times n \text{ matrix})$

Reconstruct x from $(\lvert\langle x, f_k\rangle\rvert)_{k\le m}$.	$\Longleftrightarrow$	Reconstruct xx^* from $(\lvert\langle x, f_k\rangle\rvert^2)_{k\le m}$.

Lifting

$= (f_k^* x)(x^* f_k)$ $\quad (x^* \overset{def}{=} \overline{x^T})$

$(n \times n \text{ matrix})$

Reconstruct x from $(\lvert\langle x, f_k\rangle\rvert)_{k\le m}$.	$\Longleftrightarrow$	Reconstruct xx^* from $(\lvert\langle x, f_k\rangle\rvert^2)_{k\le m}$.
	$\Longleftrightarrow$	Reconstruct xx^* from $(\mathrm{Tr}(f_k^* xx^* f_k))_{k\le m}$.

Lifting

$= (f_k^* x)(x^* f_k)$

$(x^* \stackrel{def}{=} \overline{x^T})$

$(n \times n \text{ matrix})$

Reconstruct x from $(\lvert\langle x, f_k\rangle\rvert)_{k\le m}$.	$\Longleftrightarrow$	Reconstruct xx^* from $(\lvert\langle x, f_k\rangle\rvert^2)_{k\le m}$.
	$\Longleftrightarrow$	Reconstruct xx^* from $(\mathrm{Tr}(f_k^* xx^* f_k))_{k\le m}$.

$= \mathrm{Tr}(xx^* f_k f_k^*)$

Lifting

$= (f_k^* x)(x^* f_k)$

$(x^* \stackrel{def}{=} \overline{x^T})$

$(n \times n \text{ matrix})$

Reconstruct x from $(|\langle x, f_k\rangle|)_{k\leq m}$.

$\iff$

Reconstruct xx^* from $(|\langle x, f_k\rangle|^2)_{k\leq m}$.

$\iff$

Reconstruct xx^* from $(\mathrm{Tr}(f_k^* xx^* f_k))_{k\leq m}$.

$= \mathrm{Tr}(xx^* f_k f_k^*)$

$\iff$

Reconstruct $X \in \mathbb{C}^{n\times n}$ from $(\mathrm{Tr}(X f_k f_k^*))_{k\leq m}$, with X positive, hermitian, rank 1.

Lifting

Reconstruct $X \in \mathbb{C}^{n\times n}$ from
$(\mathrm{Tr}(X f_k f_k^*))_{k\leq m}$,
with X positive, hermitian, rank 1.

$\iff$

$$\begin{aligned} &\text{Find } X \in \mathcal{H}_n \\ &\text{s.t. } \mathrm{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ &\qquad X \succeq 0 \\ &\qquad \mathrm{rank}(X) = 1. \end{aligned}$$

Set of hermitian matrices

Summary : a change of variable ($x \to xx^* = X$) has turned the modulus constraint into a linear constraint.

Not specific to phase retrieval (e.g. *MaxCut*).

Lifting

$$\begin{aligned} &\text{Find } X \in \mathcal{H}_n \\ &\text{s.t. } \operatorname{Tr}(X f_k f_k^*) = b_k, \quad \forall k \longrightarrow \text{convex} \\ &\qquad X \succeq 0 \longrightarrow \text{convex} \\ &\qquad \operatorname{rank}(X) = 1. \longrightarrow \text{non-convex} \end{aligned}$$

The problem is still non-convex :

$$\{X \in \mathcal{H}_n, \operatorname{rank}(X) = 1\} \quad \text{is not convex.}$$

But methods exist to deal with such constraints.

Digression : sparse recovery

Imagine you want to find $x \in \mathbb{R}^n$, solution of

$$\boxed{\text{Minimize } \operatorname{Card}\{i, x_i \neq 0\} \text{ under condition } \mathcal{L}(x) = a,}$$

with $\mathcal{L}$ linear and a known.

This problem is not convex.

Classical heuristic : replace "$\operatorname{Card}\{i, x_i \neq 0\}$" by the ℓ^1-norm.

$$\boxed{\text{Minimize } ||x||_1 \text{ under condition } \mathcal{L}(x) = a,}$$

The resulting problem is convex.

Convexification

Same idea : replace the rank condition by a convex surrogate.

A good convex surrogate is the nuclear norm :

$$||X||_1 = \sum_{k=1}^{n} |\lambda_k(X)|,$$

where $\lambda_1(X), \dots, \lambda_n(X)$ are the eigenvalues of X.

$$\boxed{\begin{aligned} &\text{Find } X \in \mathcal{H}_n \\ &\text{s.t. } \operatorname{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ &\qquad X \succeq 0 \\ &\qquad \operatorname{rank}(X) = 1. \end{aligned}}$$

Convexification

Same idea : replace the rank condition by a convex surrogate.

A good convex surrogate is the nuclear norm :

$$||X||_1 = \sum_{k=1}^{n} |\lambda_k(X)|,$$

where $\lambda_1(X), \ldots, \lambda_n(X)$ are the eigenvalues of X.

$$\begin{aligned} &\cancel{\text{Find } X \in \mathcal{H}_n} \quad \text{Minimize } ||X||_1 \\ &\text{s.t. } \operatorname{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ &\qquad X \succeq 0 \\ &\qquad \cancel{\operatorname{rank}(X) = 1.} \end{aligned}$$

Convexification

$$\begin{aligned} \text{Minimize } & ||X||_1 = \mathrm{Tr}(X) \\ \text{s.t. } & \mathrm{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ & X \succeq 0. \end{aligned}$$

This problem is convex. We can solve it in polynomial time.

This problem is called *PhaseLift*.
[Chai, Moscoso, and Papanicolaou, 2011]
[Candès, Eldar, Strohmer, and Voroninski, 2011]

Convexification

$$\begin{aligned} \text{Minimize } & ||X||_1 = \mathrm{Tr}(X) \\ \text{s.t. } & \mathrm{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ & X \succeq 0. \end{aligned}$$

This problem is convex. We can solve it in polynomial time.

This problem is called *PhaseLift*.
[Chai, Moscoso, and Papanicolaou, 2011]
[Candès, Eldar, Strohmer, and Voroninski, 2011]

This problem is only an approximation of the original problem.
Will its solution be the same as the original solution ?

Is the solution the same as the original one ?

Theorem (*PhaseLift* works)

There exist constants $\alpha, \gamma > 0$ such that, when

$$m \geq \alpha n,$$

the solution of the approximated problem is the same as the solution of the initial one,
with probability at least $1 - O(e^{-\gamma m})$.

[Candès, Strohmer, and Voroninski, 2013]
[Candès and Li, 2014]

Idea of proof

(Original problem)

$$\begin{aligned} &\text{Find } X \\ &\quad \text{s.t. } \operatorname{Tr}(X f_k f_k^*) = b_k, \forall k \\ &\quad\quad X \succeq 0 \\ &\quad\quad \operatorname{rank}(X) = 1. \end{aligned}$$

(Convex approximation)

$$\begin{aligned} &\text{Minimize } \operatorname{Tr}(X) \\ &\quad \text{s.t. } \operatorname{Tr}(X f_k f_k^*) = b_k, \forall k \\ &\quad\quad X \succeq 0. \end{aligned}$$

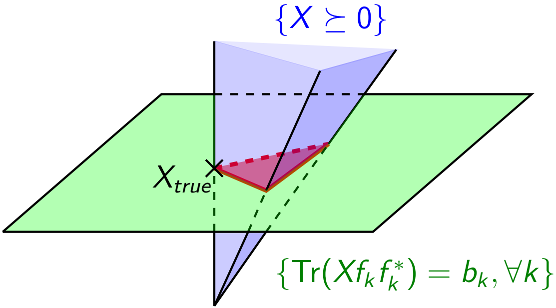
Let $X_{true} = x_{true} x_{true}^*$ be the solution of the original problem.

Let us show that it is a solution of the approximated problem.

To simplify, we assume $X_{true} = e_1 e_1^* = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 0 \end{pmatrix}$.

$X_{true} = e_1 e_1^*$ solution of

$$\begin{aligned} \text{Minimize } & \mathrm{Tr}(X) = \langle X, \mathrm{Id}_n \rangle \\ \text{s.t. } & \mathrm{Tr}(X f_k f_k^*) = b_k, \quad \forall k \\ & X \succeq 0 \quad ? \end{aligned}$$



X_{true} solution

$\Longleftrightarrow$

$-\mathrm{Id}_n \in$ Normal cone to the constraint set at X_{true}.

Dual certificate

We want to show :

$-\mathrm{Id}_n \in$ Normal cone to the constraint set at $X_{true} = e_1 e_1^*$.

with Constraint set $= \{X \succeq 0\} \cap \{\mathrm{Tr}(X f_k f_k^*) = b_k, \forall k\}$.

Dual certificate

We want to show :

$$-\mathrm{Id}_n \in \text{Normal cone to the constraint set at } X_{true} = e_1 e_1^*.$$

with $\quad$ Constraint set $= \{X \succeq 0\} \cap \{\mathrm{Tr}(X f_k f_k^*) = b_k, \forall k\}$.

The normal cone of the intersection contains the sum of the normal cones of each set :

$$\left\{ \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & \mathrm{M} & \\ 0 & & \end{pmatrix}, M \preceq 0 \right\} \quad + \quad \left\{ \sum_k c_k f_k f_k^*, c_1, \ldots, c_m \in \mathbb{R} \right\}.$$

Dual certificate

We want to show :

$$-\mathrm{Id}_n \in \text{Normal cone to the constraint set at } X_{true} = e_1 e_1^*.$$

with $\quad$ Constraint set $= \{X \succeq 0\} \cap \{\mathrm{Tr}(X f_k f_k^*) = b_k, \forall k\}$.

The normal cone of the intersection contains the sum of the normal cones of each set :

$$\left\{ \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \mathrm{M} & \\ 0 & & \end{pmatrix}, M \preceq 0 \right\} \quad + \quad \left\{ \sum_k c_k f_k f_k^*, c_1, \dots, c_m \in \mathbb{R} \right\}.$$

Let us find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

$$\boxed{-\mathrm{Id}_n = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \mathrm{M} & \\ 0 & & \end{pmatrix} + \sum_{k \le m} c_k f_k f_k^*.}$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

$$-\mathrm{Id}_n = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \mathsf{M} & \\ 0 & & \end{pmatrix} + \sum_{k \leq m} c_k f_k f_k^*.$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

(Approximate equality is actually enough)

$$-\mathrm{Id}_n \underset{\approx}{\cancel{=}} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & M & \\ 0 & & \end{pmatrix} + \sum_{k \le m} c_k f_k f_k^*.$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

(Approximate equality is actually enough)

$$-\mathrm{Id}_n \cancel{=} \approx \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \text{M} & \\ 0 & & \end{pmatrix} + \sum_{k \leq m} c_k f_k f_k^*.$$

(Idea)

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} - \frac{1}{2m} \sum_{k \leq m} |\langle e_1, f_k \rangle|^2 f_k f_k^*.$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

(Approximate equality is actually enough)

$$-\mathrm{Id}_n \;\cancel{=}\; \approx \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \mathrm{M} & \\ 0 & & \end{pmatrix} + \sum_{k \le m} c_k f_k f_k^* .$$

(Idea)

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} - \frac{1}{2m} \sum_{k \le m} \underbrace{|\langle e_1, f_k \rangle|^2 f_k f_k^*}_{= \begin{pmatrix} 2 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \text{ in expectation}} .$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

(Approximate equality is actually enough)

$$-\mathrm{Id}_n \underset{\approx}{\cancel{=}} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & M & \\ 0 & & \end{pmatrix} + \sum_{k \le m} c_k f_k f_k^*.$$

(Idea)

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} - \frac{1}{2m} \sum_{k \le m} \underbrace{|\langle e_1, f_k\rangle|^2 f_k f_k^*}_{= \begin{pmatrix} 2 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \text{ in expectation}}.$$

$$\underbrace{}_{\approx \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} \text{ by concentration inequalities}}$$

We want to find $M \preceq 0$ and $c_1, \dots, c_m \in \mathbb{R}$ such that

(Approximate equality is actually enough)
$$-\mathrm{Id}_n \underset{\approx}{\cancel{=}} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & M & \\ 0 & & \end{pmatrix} + \sum_{k \le m} c_k f_k f_k^*.$$

(Idea)
$$\boxed{\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} - \frac{1}{2m} \sum_{k \le m} \underbrace{|\langle e_1, f_k \rangle|^2 f_k f_k^*}_{= \begin{pmatrix} 2 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \text{ in expectation}}.}$$

$$\approx -\mathrm{Id}_n$$

$$\approx \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -\frac{1}{2} \end{pmatrix} \text{ by concentration inequalities}$$

End of the proof

We have constructed an (approximate) dual certificate.

This shows that the solution $X_{true} = e_1 e_1^*$ of the original problem is also a solution of the approximated problem.

With the dual certificate, we can also prove that the solution of the approximated problem is unique.

So with high probability, *PhaseLift* works.

PhaseLift is stable to noise

$$\begin{aligned} \text{Minimize} \quad & \mathrm{Tr}(X) \\ \text{s.t.} \quad & \mathrm{Tr}(X f_k f_k^*) = b_k + \epsilon_k, \quad \forall k \\ & X \succeq 0 \quad ? \end{aligned}$$

For simplicity, assume $||x||_2 = 1$.

Theorem (Candès and Li [2014])

Under the same hypotheses as previously, *PhaseLift* allows to recover a vector x_{noise} such that, with high probability,

$$||x_{true} - x_{noise}||_2 \leq \text{constant} \times \frac{||\epsilon||_1}{m}.$$

Possible extensions

- The probability distribution of the f_k's is something else than a normal law.
  $\rightarrow$ "Coded diffraction patterns"

  [Candès, Li, and Soltanolkotabi, 2015]
  [Gross, Krahmer, and Kueng, 2015]

- The noise contains some very large entries.

  [Hand, 2017]

Computational complexity

It depends on which algorithm we use to solve *PhaseLift*.

We assume $m = O(n)$; let ϵ be the precision.

- Interior-point solvers :

$$O\left(n^{4.5}\log\left(\frac{1}{\epsilon}\right)\right).$$

- First-order methods :

$$O\left(\frac{n^3}{\epsilon}\right).$$

The problem is that we have lifted.
The matrix X has n^2 entries, while x had only n.

How to make it faster ?

- Perform lifting in a different way, so that the lifted problem has a more favorable structure.
  [Waldspurger, d'Aspremont, and Mallat, 2015]
- Use the fact that the solution will be low-rank.
  $\rightarrow$ Represent X by its eigenvectors.
  [Yurtsever, Udell, Tropp, and Cevher, 2017]
- Find a convexification method with no lifting ?
  $\rightarrow$ *PhaseMax*
  [Bahmani and Romberg, 2017]
  [Goldstein and Studer, 2016]

Convexification in the natural parameter space

$$\boxed{\begin{aligned} &\text{Find } x \in \mathbb{C}^n \\ &\text{such that } |\langle x, f_k\rangle| = b_k, \quad \forall k \leq m. \end{aligned}}$$

Recall that the problem is that the set

$$\{x \in \mathbb{C}^n, |\langle x, f_k\rangle| = b_k\}$$

is non-convex.

We replace the equality by an inequality. The set

$$\{x \in \mathbb{C}^n, |\langle x, f_k\rangle| \leq b_k\}$$

is convex.

Convexification in the natural parameter space

$$\begin{aligned} &\text{Find } x \in \mathbb{C}^n \\ &\text{such that } |\langle x, f_k \rangle| \overset{\leq}{\cancel{=}} b_k, \quad \forall k \leq m. \end{aligned}$$

Recall that the problem is that the set

$$\{x \in \mathbb{C}^n, |\langle x, f_k \rangle| = b_k\}$$

is non-convex.

We replace the equality by an inequality. The set

$$\{x \in \mathbb{C}^n, |\langle x, f_k \rangle| \leq b_k\}$$

is convex.

Convexification in the natural parameter space

We get a convex approximated problem.

$$\begin{aligned} &\text{Find } x \in \mathbb{C}^n \\ &\text{such that } |\langle x, f_k \rangle| \leq b_k, \quad \forall k \leq m. \end{aligned}$$

Convexification in the natural parameter space

We get a convex approximated problem.

$$\boxed{\begin{aligned} &\text{Find } x \in \mathbb{C}^n \\ &\text{such that } |\langle x, f_k \rangle| \leq b_k, \quad \forall k \leq m. \end{aligned}}$$

But this problem has many solutions that are not the correct one (0, for instance).

Convexification in the natural parameter space

We get a convex approximated problem.

$$\boxed{\begin{aligned} &\text{Find } x \in \mathbb{C}^n \\ &\text{such that } |\langle x, f_k\rangle| \leq b_k, \quad \forall k \leq m. \end{aligned}}$$

But this problem has many solutions that are not the correct one (0, for instance).

Let us assume that we can compute an approximation of x :

$$x_{anchor} \approx x_{true}.$$

Let us pick the solution that "looks most" like x_{anchor}.

Convexification in the natural parameter space

We get a convex approximated problem.

$$\boxed{\begin{array}{c} \cancel{\text{Find } x \in \mathbb{C}^n} \quad \text{Max } \langle x, x_{anchor} \rangle \\ \text{such that } |\langle x, f_k \rangle| \leq b_k, \quad \forall k \leq m. \end{array}}$$

But this problem has many solutions that are not the correct one (0, for instance).

Let us assume that we can compute an approximation of x :

$$x_{anchor} \approx x_{true}.$$

Let us pick the solution that "looks most" like x_{anchor}.

PhaseMax works

Theorem

Let $\theta_0 \in]0; \frac{\pi}{2}[$ be fixed.
There exists $\alpha, \gamma > 0$ such that, if

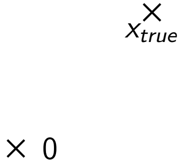
$$m \geq \alpha n \quad \text{and} \quad \mathrm{angle}(x_{anchor}, x_{true}) < \theta_0,$$

then *PhaseMax* recovers the correct solution,
with probability $1 - O(e^{-\gamma m})$.

[Bahmani and Romberg, 2017]
[Goldstein and Studer, 2016]

Intuition

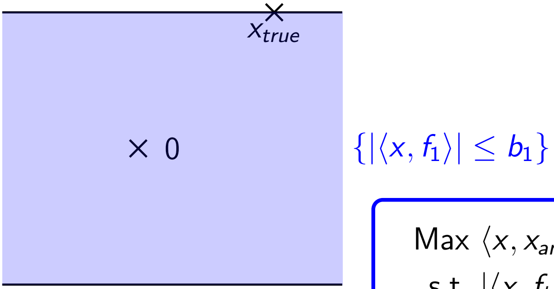
To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\boxed{\begin{array}{l}\text{Max } \langle x, x_{anchor} \rangle \\ \quad \text{s.t. } |\langle x, f_k \rangle| \leq b_k, \forall k\end{array}}$$

Intuition

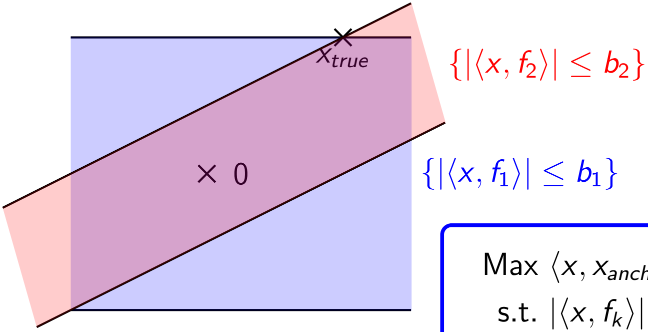
To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\boxed{\begin{aligned} &\text{Max } \langle x, x_{anchor}\rangle \\ &\text{ s.t. } |\langle x, f_k\rangle| \leq b_k, \forall k \end{aligned}}$$

Intuition

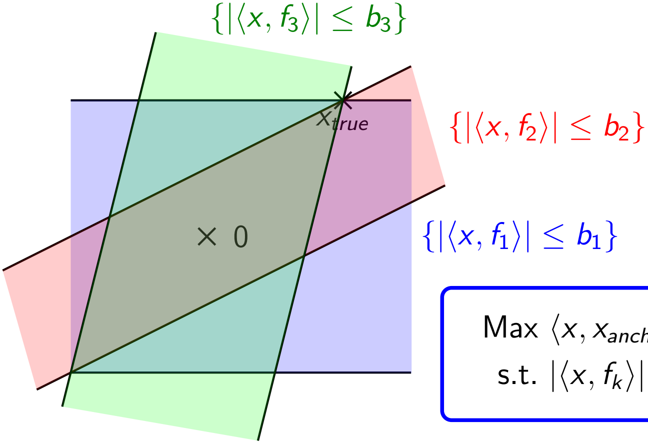
To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\boxed{\begin{aligned} &\text{Max } \langle x, x_{anchor}\rangle \\ &\text{ s.t. } |\langle x, f_k\rangle| \leq b_k, \forall k \end{aligned}}$$

Intuition

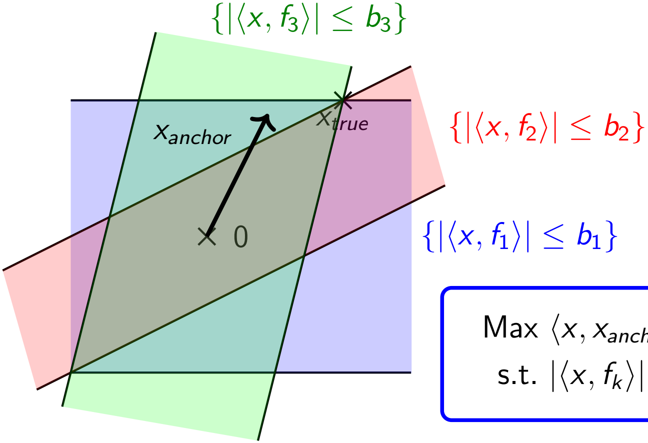
To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\begin{aligned} &\text{Max } \langle x, x_{anchor}\rangle \\ &\text{s.t. } |\langle x, f_k\rangle| \leq b_k, \forall k \end{aligned}$$

Intuition

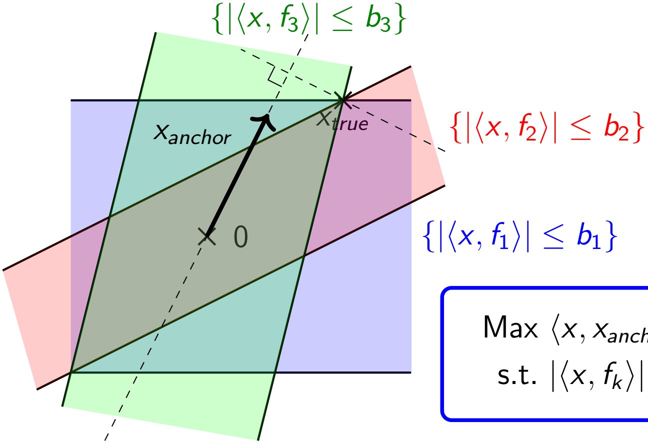
To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\begin{aligned} &\text{Max } \langle x, x_{anchor} \rangle \\ &\text{s.t. } |\langle x, f_k \rangle| \leq b_k, \forall k \end{aligned}$$

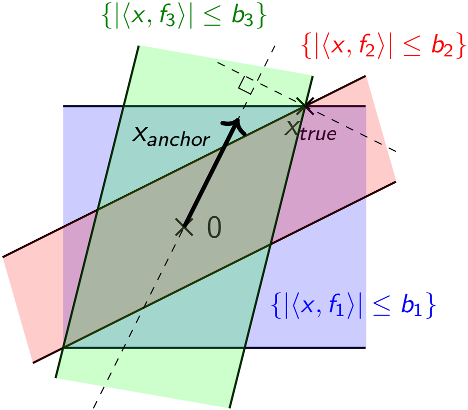
Intuition

To simplify, assume $x, f_1, \ldots, f_m$ have real coordinates.



$$\boxed{\begin{array}{l} \text{Max } \langle x, x_{anchor}\rangle \\ \text{ s.t. } |\langle x, f_k\rangle| \leq b_k, \forall k \end{array}}$$

Vague idea of proof

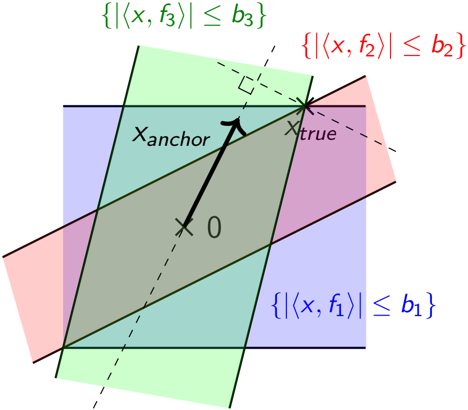


We show : for all δ, if

$$\langle x_{true} + \delta, x_{anchor}\rangle \geq \langle x_{true}, x_{anchor}\rangle,$$

then $x_{true} + \delta$ is not in the intersection of the slabs.

Vague idea of proof

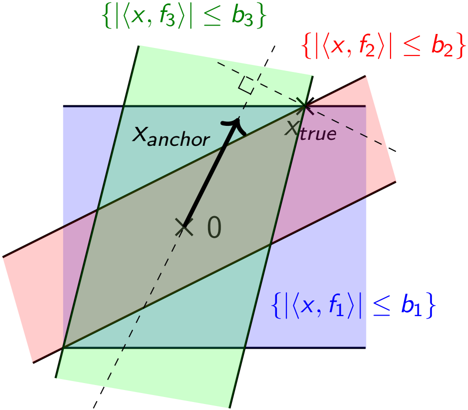


We show : for all δ, if

$$\langle \delta, x_{anchor}\rangle \geq 0$$

then $x_{true} + \delta$ is not in the intersection of the slabs.

Vague idea of proof



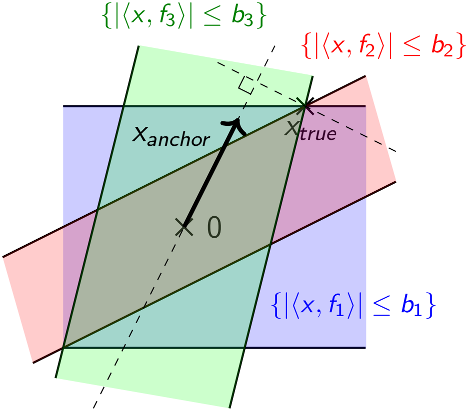
We show : for all δ, if

$$\langle \delta, x_{anchor}\rangle \geq 0$$

then, for some k,

$$|\langle x_{true} + \delta, f_k\rangle| > b_k$$
$$= |\langle x_{true}, f_k\rangle|.$$

Vague idea of proof



We show : for all δ, if

$$\langle \delta, x_{anchor}\rangle \geq 0$$

then, for some k,

$$|\langle x_{true} + \delta, f_k\rangle| > b_k$$
$$= |\langle x_{true}, f_k\rangle|.$$

Enough to show that, $\forall\delta$, if

$$\langle \delta, x_{anchor}\rangle \geq 0,$$

then, for some k,

$$\mathrm{sign}(\langle x_{true}, f_k\rangle)\langle \delta, f_k\rangle > 0.$$

Vague idea of proof

In other words, it is enough to show that, with high probability,

$$\text{A specific half-space} \subset \text{Some union of random half-spaces.}$$

The probability that the inclusion holds can be precisely lower-bounded.

Stability to noise ?

- *PhaseMax* is stable to bounded non-negative noise.
  [Bahmani and Romberg, 2017]
  [Goldstein and Studer, 2016]
- A modified version is stable to sparse arbitrary noise.
  [Hand and Voroninski, 2016]
- More realistic noise models ?

Computational complexity

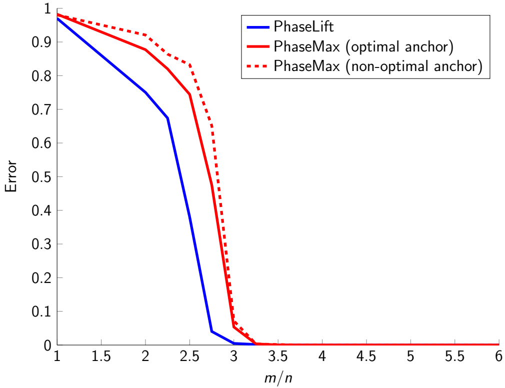
- In *PhaseMax*, the unknown is a vector, not a matrix as in *PhaseLift*.
- The most costly operations in solving *PhaseMax* are matrix-vector multiplications : $O(n^2)$ operations.
- Solving a penalized version of *PhaseMax* with a first order method :

$$O\left(\frac{n^2}{\epsilon^{3/2}}\right),$$

where $\epsilon > 0$ is the desired precision, and $m = O(n)$.

(For *PhaseLift*, the term in n was n^3 or $n^{4.5}$.)

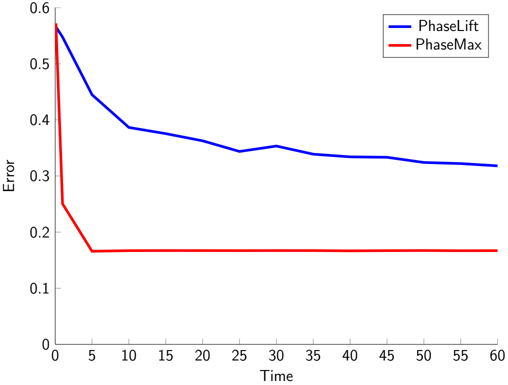
Numerical results



Median error as a function of m/n for $n = 64$.

[Chandra, Zhong, Hontz, McCulloch, Studer, and Goldstein, 2017]

Numerical results



Error as a function of the time, for $n = 64$ and $m = 256$.

Summary

- *PhaseLift*
  - Lifting the problem $\rightarrow$ convenient convex approximation.
  - *PhaseLift* has very good theoretical guarantees.
  - Because of lifting, solving *PhaseLift* is slow.
- *PhaseMax*
  - Convexification without lifting.
  - *PhaseMax* has good theoretical guarantees.
  - Maybe a bit less precise than *PhaseLift*, but much faster.

Tomorrow

Convexification is not the only way to handle non-convexity.

We can also ignore the non-convexity, and try to solve the problem "as if it was convex".

$\rightarrow$ Non-convex methods.