A Simple Convex Program for Phase Retrieval Using Anchor Vectors

Sohail Bahmani and Justin Romberg
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

I. Background

Phase retrieval is often abstracted as the problem of finding an (approximate) solution $\widehat{\boldsymbol{x}}$ to the system of quadratic equations

$$y_m \approx \left|\boldsymbol{a}_m^* \boldsymbol{x}\right|^2 , \qquad m = 1, 2, \dots, M , \tag{1}$$

where $\boldsymbol{a}_m$s are given measurement vectors and approximate equations are due to noise. The existing convex programming methods for phase retrieval (e.g., *PhaseLift* [1]) utilize a lifting transform such as $\boldsymbol{x}\boldsymbol{x}^* \mapsto \boldsymbol{X}$ to convert the quadratic equations above to equations that are linear in $\boldsymbol{X}$. These methods then use semidefinite programming (SDP) to find a positive semidefinite matrix that is consistent with the linearized equations and has a small trace norm inducing the low-rank structure of the desired lifted solution. The main drawback of these SDP-based methods is that their computational cost does not scale well with the dimension of problem. Iterative non-convex methods such as the *AltMinPhase [2]* and the *Wirtinger Flow [3]* methods operate in the natural domain of the problem and thus avoid the computational drawback of the SDP-based methods. With proper initialization these methods produce accurate estimates of the signal. However, they are often not robust to changes in the model specifications (e.g., measurement distribution) which is the typical consequence of relying on "non-convex optimization". Furthermore, their analysis is generally tailored for a specific problem and is usually difficult to generalize.

In this paper, we propose a new convex relaxation for the phase retrieval problem that operates in the natural domain of the problem. The proposed method has a significantly lower computational cost than the SDP-based methods and competes with the non-convex iterative methods. It is also flexible in analysis and implementation as it relies on convex programming.

Shortly after we proposed and analyzed this new method in [4], the same algorithm is analyzed independently in [5] using different techniques. An alternative proof in the case of real-valued signals also appeared later in [6].

II. Convex Relaxation Using an Anchor Vector

Suppose that we observe quadratic measurements $y_1, y_2, \dots, y_m$ as in (1) for the ground truth $\boldsymbol{x}_\star \in \mathbb{C}^N$. As illustrated in Figure 1, with no measurement noise, $\boldsymbol{x}_\star$ can be viewed as an extreme point of the convex set $\mathcal{K} = \cap_{m=1}^M \mathcal{S}_m$ obtained by intersecting the slabs

$$\mathcal{S}_m = \left\{\boldsymbol{x} : \left|\boldsymbol{a}_m^*\boldsymbol{x}\right|^2 \le y_m\right\}.$$

To distinguish $\boldsymbol{x}_\star$ from the other extreme points we utilize an *anchor vector*. A vector $\boldsymbol{a}_0$ is an anchor vector with parameter $\delta \in (0,1)$ if

$$\left|\boldsymbol{a}_0^*\boldsymbol{x}_\star\right| \ge \delta \left\|\boldsymbol{a}_0\right\|_2 \left\|\boldsymbol{x}_\star\right\|_2 . \tag{2}$$

In words, the anchor vector has a "non-trivial" correlation with the ground truth. If the measurements are randomized, an anchor vector obeying (2) can be constructed from the given measurements. For example, the principal eigenvector of the matrix $\boldsymbol{S} = \frac{1}{M}\sum_{m=1}^M y_m \boldsymbol{a}_m\boldsymbol{a}_m^*$ or variations of it, with high probability, can be an anchor vector obeying (2). These *data-driven* anchor vectors are previously used as the initialization for some non-convex phase retrieval methods [2, 3, 7].

Given the anchor vector $\boldsymbol{a}_0$, the extreme point of $\mathcal{K}$ that is best aligned with $\boldsymbol{a}_0$ can be found by the convex program

$$\begin{aligned} &\underset{\boldsymbol{x}}{\operatorname{argmax}}\ \operatorname{Re}\left(\boldsymbol{a}_0^*\boldsymbol{x}\right) \\ &\text{subject to } \left|\boldsymbol{a}_m^*\boldsymbol{x}\right|^2 \le y_m , \qquad m = 1, 2, \dots, M . \end{aligned} \tag{3}$$

The convex program (3) is our proposed estimator for phase retrieval even with noisy measurements. We assume that the noise is bounded and non-negative; for a positive parameter η representing the SNR we have

$$0 \le y_m - \left|\boldsymbol{a}_m^*\boldsymbol{x}_\star\right|^2 \le \eta^{-1}\left\|\boldsymbol{x}_\star\right\|_2^2 , \qquad m = 1, 2, \dots, M .$$

The following lemma characterizes a sufficient condition to guarantee accuracy of the estimate produced by (3). Henceforth, without loss of generality, we assume that the ground truth $\boldsymbol{x}_\star$ is unit norm (i.e., $\left\|\boldsymbol{x}_\star\right\|_2 = 1$).

Lemma 1. *Define* $\mathcal{R}_\delta \stackrel{\text{def}}{=} \left\{\boldsymbol{h} : \left\|\boldsymbol{h} - (\boldsymbol{x}_\star^*\boldsymbol{h})\,\boldsymbol{x}_\star\right\|_2 \ge \delta \left|\operatorname{Im}\left(\boldsymbol{x}_\star^*\boldsymbol{h}\right)\right|\right\}$. *The estimate produced by (3) obeys* $\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}_\star\right\|_2 \le \eta^{-1}$, *if there is no* $\boldsymbol{h} \in \mathcal{R}_\delta$ *with* $\left\|\boldsymbol{h}\right\|_2 > \eta^{-1}$ *such that*

$$\begin{aligned} \operatorname{Re}\left(\boldsymbol{a}_0^*\boldsymbol{h}\right) &\ge 0 \\ \left|\operatorname{Re}\left(\boldsymbol{a}_m\boldsymbol{a}_m^*\boldsymbol{x}_\star, \boldsymbol{h}\right)\right| &\le \frac{1}{2}\eta^{-1} , \qquad m = 1, 2, \dots, M , \end{aligned}$$

hold simultaneously.

For measurements that are drawn i.i.d. from a distribution with mild regularity conditions, we have shown, using classic results from *statistical learning theory* [8] on linear classification, that the sufficient condition stated in Lemma 1 holds with high probability. The requirement is that the number of measurements (i.e., M) should grow with the dimension of the problem (i.e., N) as $M \overset{\delta}{\gtrsim} N$ with the hidden constant factor depending on the parameter δ of the anchor vector. The following theorem summarizes our result in the special case where the measurements are drawn from the standard (complex) Normal distribution.

Theorem 2. *Let the measurement vectors* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_M$ *be i.i.d. copies of the random vector* $\boldsymbol{a} \sim \operatorname{Normal}\left(0, \frac{1}{2}\boldsymbol{I}\right) + \imath\operatorname{Normal}\left(0, \frac{1}{2}\boldsymbol{I}\right)$. *If for* $\varepsilon \in (0,1)$ *the number of measurements obeys*

$$M \ge C_\delta\left(N + \log\frac{1}{\varepsilon}\right) ,$$

where C_δ *is a constant depending on* δ *in (2), then, with probability* $\ge 1 - \varepsilon$, *the estimate* $\widehat{\boldsymbol{x}}$ *produced by (3) obeys*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}_\star\right\|_2 \le \eta^{-1} .$$

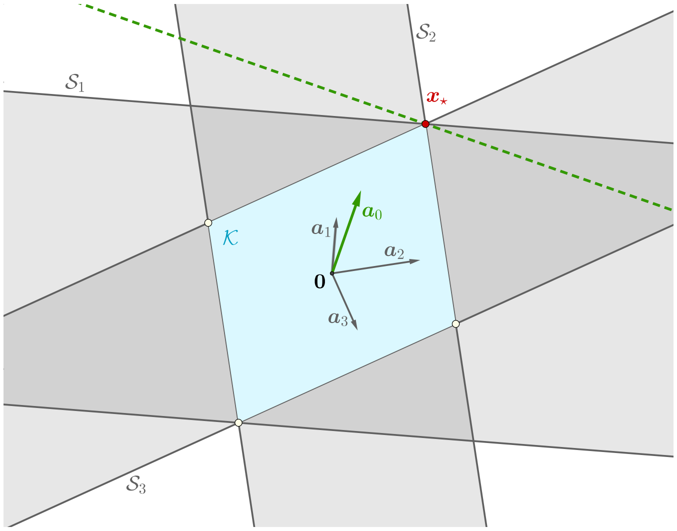

Figure 1. The slabs $\mathcal{S}_1, \mathcal{S}_2, \ldots$ corresponding to each measurement intersect at $\mathcal{K}$. With no measurement noise, $\boldsymbol{x}_\star$ (and its equivalent points) are extreme points of $\mathcal{K}$ which can be identified using the anchor vector $\boldsymbol{a}_0$.

References

[1] E. J. Candès, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.

[2] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, 2013, pp. 2796–2804.

[3] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *Information Theory, IEEE Transactions on*, vol. 61, no. 4, pp. 1985–2007, Apr. 2015.

[4] S. Bahmani and J. Romberg, "Phase retrieval meets statistical learning theory: A flexible convex relaxation," preprint arXiv:1610.04210 [cs.IT].

[5] T. Goldstein and C. Studer, "Phasemax: Convex phase retrieval via basis pursuit," preprint arXiv:1610.07531 [cs.IT].

[6] P. Hand and V. Voroninski, "An elementary proof of convex phase retrieval in the natural parameter space via the linear program phasemax," preprint arXiv:1611.03935 [cs.IT].

[7] Y. Chen and E. Candès, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Advances in Neural Information Processing Systems 28 (NIPS'15)*. Curran Associates, Inc., Dec. 2015, pp. 739–747.

[8] V. N. Vapnik, *Statistical Learning Theory*. Wiley, 1998.