Appl. Comput. Harmon. Anal. 36 (2014) 473–494

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Phase retrieval: Stability and recovery guarantees ☆

Yonina C. Eldar [a,*], Shahar Mendelson [b]

[a] Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel
[b] Department of Mathematics, Technion—Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 16 November 2012
Received in revised form 31 July 2013
Accepted 17 August 2013
Available online 4 September 2013
Communicated by Bruno Torresani

Keywords:
Phase retrieval
Compressed sensing

ABSTRACT

We consider stability and uniqueness in real phase retrieval problems over general input sets, when the data consists of random and noisy quadratic measurements of an unknown input $x_0 \in \mathbb{R}^n$ that lies in a general set T. We study conditions under which x_0 can be stably recovered from the measurements. In the noise-free setting we show that the number of measurements needed to ensure a unique and stable solution depends on the set T through its Gaussian mean-width, which can be computed explicitly for many sets of interest. In particular, for k-sparse inputs, $O(k\log(n/k))$ measurements suffice, while if x_0 is an arbitrary vector in $\mathbb{R}^n$, $O(n)$ measurements are sufficient.

In the noisy case, we show that if the empirical risk is bounded by a given, computable constant that depends only on statistical properties of the noise, the error with respect to the true input is bounded by the same Gaussian parameter (up to logarithmic factors). Therefore, the number of measurements required for stable recovery is the same as in the noise-free setting up to log factors.

It turns out that the complexity parameter for the quadratic problem is the same as the one used for analyzing stability in linear measurements under very general conditions. Thus, no substantial price has to be paid in terms of stability when there is no knowledge of the phase of the measurements.

1. Introduction

Recently, there has been growing interest in recovering an input vector $x_0 \in \mathbb{R}^n$ from quadratic measurements

$$y_i = \left|\langle a_i, x_0\rangle\right|^2 + w_i, \quad i = 1, \ldots, N. \tag{1.1}$$

Here, we focus on the case in which $(a_i)_{i=1}^N$ are selected independently according to a random vector a on $\mathbb{R}^n$, and $(w_i)_{i=1}^N$ are selected independently according to the noise w, and are assumed to be independent of $(a_i)_{i=1}^N$.

Since only the magnitude of $\langle a_i, x_0\rangle$ is measured, and not the phase (or the sign, in the real case), this family of problems is referred to as *phase retrieval*. These problems arise in many areas of optics, where the detector can only measure the magnitude of the received optical wave. Several applications of phase retrieval include X-ray crystallography, transmission electron microscopy and coherent diffractive imaging [39,20,19,46].

☆ The work of Y. Eldar is supported in part by the Israel Science Foundation under Grant no. 170/10, in part by the Ollendorf Foundation, in part by a Magnet grant Metro450 from the Israel Ministry of Industry and Trade, and in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI). The work of S. Mendelson is supported in part by the Centre for Mathematics and Its Applications, The Australian National University, Canberra, ACT 0200, Australia. Additional support was given by the European Community's Seventh Framework Programme (FP7/2007–2013) under ERC grant agreement 203134, and by the Israel Science Foundation grant 900/10.

* Corresponding author.
E-mail addresses: yonina@ee.technion.ac.il (Y.C. Eldar), shahar@tx.technion.ac.il (S. Mendelson).

1063-5203/$ – see front matter
http://dx.doi.org/10.1016/j.acha.2013.08.003

Many algorithmic methods have been developed for phase recovery (see, e.g., [20]) which often rely on prior information on the signal, such as positivity or support constraints. One of the most popular techniques is based on alternating projections, such as the Gerchberg and Saxton [16] and Fienup [15] iterations. To circumvent the difficulties associated with convergence of alternating projections, more recently, phase retrieval problems have been treated using semidefinite relaxation, and low-rank matrix recovery ideas [7,41]. Another possible approach that potentially leads to robust solutions is to assume that the input signal x_0 is sparse, namely, that it contains only a few non-zeros values in an appropriate basis expansion. Both the semidefinite relaxation [41,21,37] and greedy recovery methods [36,5,40] can be extended to phase retrieval of sparse inputs.

Despite the vast interest in phase retrieval, there has been little theoretical work on the fundamental limits of this problem, which is the focus of this article. One question in this context is to estimate the number of measurements that are needed in order to ensure robust recovery of the input x_0 – and regardless of the specific recovery method used. Several recent works treat this problem; most, study the case in which x_0 is a general input, namely, there is no sparsity (or other) constraints on x_0. For example, in this case, with probability one, $N = 4n - 2$ randomized equations are sufficient for recovery using a brute force (intractable) method, when there is no noise [2]. However, it is not clear, even in that restricted scenario, whether a stable recovery method exists with this number of measurements. In [8,9] the authors consider the case in which a_i are real or complex vectors that are either uniform on the sphere of radius $\sqrt{n}$, or iid zero-mean Gaussian vectors with unit variance. Under these assumptions they show that on the order of n measurements are needed in order to recover a generic x_0 (and while using a semidefinite relaxation approach). In the presence of noise, it is shown in [8] that one can find an estimate $\hat{x}$ satisfying

$$\left\|\hat{x} - e^{i\phi}x\right\|_2 \leqslant C_0 \min\left(\|x\|_2, \frac{\|w\|_1}{N\|x\|_2}\right), \tag{1.2}$$

for some ϕ, where C_0 is a constant and w is the noise vector that is assumed to be bounded so that $\|w\|_1$ is finite.

The article [27] treats the case in which the input x_0 is of norm one and k-sparse, and a_i are independent, zero-mean normal vectors. It shows that if N is on the order of $k^2 \log n$, then recovery is possible (by using a sparse semidefinite relaxation approach).

Here, we treat the real case and random measurements, using reasonable ensembles. The first question addressed is that of stable uniqueness, namely, identifying conditions under which a unique solution can be found in a stable way. Though the results presented here apply to arbitrary sets $T \subset \mathbb{R}^n$, the examples we consider are $T = \mathbb{R}^n$, and the class of k-sparse vectors. For the latter, $O(k\log(n/k))$ measurements suffice for stability. This result is better by a factor of k than the estimate from [27]. Also, when x_0 can be any vector in $\mathbb{R}^n$, $O(n)$ measurements suffice, which is also the bound derived in [8].

It turns out that the same complexity parameter, the Gaussian mean-width, captures both linear and quadratic problems. This observation will be discussed in Section 5. It implies that in a rather general sense, the number of measurements required for stable recovery in the quadratic setting, is of the same order of magnitude as the one needed to ensure stability under linear sampling.

The second main result of this article deals with the noisy phase retrieval problem; more specifically, recovery from noisy measurements of the form (1.1), generated by $x_0 \in T$. A straightforward approach is to select $\hat{x}$ that minimizes the empirical risk, but since this leads to a nonconvex problem, finding its global solution is in general not possible. Nonetheless, one can show that if the empirical risk of $\hat{x}$ is bounded by a given, computable constant (and that depends only on statistical properties of the noise), then $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2$ may be controlled using the Gaussian mean-width of the set. In particular, for reasonable noise levels, in the case of k-sparse vectors on the Euclidean sphere S^{n-1}, one can guarantee stable recovery from $O(k\log(n/k)(\log^2 k + \log^2 \log(n/k)))$ noisy measurements, and when x_0 can be any vector in S^{n-1}, $O(n\log^2 n)$ noisy measurements are sufficient. An exact formulation of both main results is presented in the next section.

A conclusion that could have practical importance is that although the squared error for nonlinear measurements as in (1.1) cannot be minimized directly, it is sufficient to find a point for which the empirical error is bounded by a known constant. Thus, one may use any desired recovery algorithm and check whether the solution $\hat{x}$ satisfies the bound. For this purpose, methods such as those developed in [40] are advantageous since they allow for arbitrary initial points. As different initializations lead to different choices of $\hat{x}$, the algorithm can be used several times until an appropriate value of $\hat{x}$ is found. The theoretical analysis ensures that such an $\hat{x}$ is sufficiently close to x_0 or to $-x_0$ if enough measurements are used.

The reminder of the article is organized as follows. The problem and main results are formulated in Section 2. Stability results in the noise-free setting are developed in Section 3, while the noisy setting is treated in Section 4. In Section 5 the relation between the results in the quadratic case and those in the linear setting is discussed.

Throughout the article we use the following notation. All absolute constants (that is, fixed positive numbers) are denoted by c_1, c_2, etc. Their values may change from line to line. The expectation is denoted by $\mathbb{E}$, and if the probability space is a product space $(\Omega \times \Omega', \mu \otimes \mu')$, then $\mathbb{E}_\mu$ and $\mathbb{E}_{\mu'}$ are the conditional expectations. In the context of an empirical process, $P_N f$ denotes the empirical mean of f while Pf is its expectation. If X is a random variable, then $\|X\|_{L_p} = (\mathbb{E}|X|^p)^{1/p}$. If $x \in \mathbb{R}^n$, $\| \ \|_p$ denotes its ℓ_p norm. ℓ_p^n is the normed space $(\mathbb{R}^n, \| \ \|_p)$, the corresponding unit ball is B_p^n and S^{n-1} is the Euclidean sphere in $\mathbb{R}^n$.