Ann. Appl. Math.
doi: 10.4208/aam.OA-2021-0009

Vol. **37**, No. 4, pp. 437-512
November 2021

The Global Landscape of Phase Retrieval I: Perturbed Amplitude Models

Jian-Feng Cai[1], Meng Huang[1], Dong Li[2,*] and Yang Wang[1]

[1] *Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

[2] *SUSTech International Center for Mathematics and Department of Mathematics, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China*

Received 3 April 2021; Accepted (in revised version) 10 November 2021

Abstract. A fundamental task in phase retrieval is to recover an unknown signal $\boldsymbol{x}\in\mathbb{R}^n$ from a set of magnitude-only measurements $y_i=|\langle \boldsymbol{a}_i,\boldsymbol{x}\rangle|$, $i=1,\cdots,m$. In this paper, we propose two novel perturbed amplitude models (PAMs) which have a non-convex and quadratic-type loss function. When the measurements $\boldsymbol{a}_i\in\mathbb{R}^n$ are Gaussian random vectors and the number of measurements $m\geq Cn$, we rigorously prove that the PAMs admit no spurious local minimizers with high probability, i.e., the target solution $\boldsymbol{x}$ is the unique local minimizer (up to a global phase) and the loss function has a negative directional curvature around each saddle point. Thanks to the well-tamed benign geometric landscape, one can employ the vanilla gradient descent method to locate the global minimizer $\boldsymbol{x}$ (up to a global phase) without spectral initialization. We carry out extensive numerical experiments to show that the gradient descent algorithm with random initialization outperforms state-of-the-art algorithms with spectral initialization in empirical success rate and convergence speed.

AMS subject classifications: 94A12, 65K10, 49K45

Key words: Phase retrieval, landscape analysis, non-convex optimization.

---

1 Introduction

Emails: jfcai@ust.hk (J. Cai), menghuang@ust.hk (M. Huang), lid@sustech.edu.cn (D. Li), yangwang@ust.hk (Y. Wang)

1.1 Background

The basic amplitude model for phase retrieval can be written as

$$y_j = |\langle \boldsymbol{a}_j, \boldsymbol{x} \rangle|, \quad j = 1, \cdots, m,$$

where $\boldsymbol{a}_j \in \mathbb{R}^n$, $j=1,\cdots,m$ are given vectors and m is the number of measurements. The goal is to recover the unknown signal $\boldsymbol{x} \in \mathbb{R}^n$ based on the measurements $\{(\boldsymbol{a}_j, y_j)\}_{j=1}^m$. This problem arises in many fields of science and engineering such as X-ray crystallography [16, 24], microscopy [23], astronomy [7], coherent diffractive imaging [15,28] and optics [34] etc. In practical applications due to the physical limitations optical detectors can only record the magnitude of signals while losing the phase information. Despite its simple mathematical formulation, it has been shown that reconstructing a finite-dimensional discrete signal from the magnitude of its Fourier transform is generally an *NP-complete* problem [27].

Many algorithms have been designed to solve the phase retrieval problem, which can be categorized into convex algorithms and non-convex ones. The convex algorithms usually rely on a "matrix-lifting" technique, which lifts the phase retrieval problem into a low rank matrix recovery problem. By using convex relaxation one can recast the matrix recovery problem as a convex optimization problem. The corresponding algorithms include PhaseLift [2, 4], PhaseCut [33] etc. It has been shown [2] that PhaseLift can achieve the exact recovery under the optimal sampling complexity with Gaussian random measurements.

Although convex methods have good theoretical guarantees of convergence, they tend to be computationally inefficient for large scale problems. In contrast, many non-convex algorithms bypass the lifting step and operate directly on the lower-dimensional ambient space, making them much more computationally efficient. Early non-convex algorithms were mostly based on the technique of alternating projections, e.g., Gerchberg-Saxton [14] and Fineup [9]. The main drawback, however, is the lack of theoretical guarantee. Later Netrapalli et al. [25] proposed the AltMinPhase algorithm based on a technique known as *spectral initialization*. They proved that the algorithm linearly converges to the true solution with $\mathcal{O}(n\log^3 n)$ resampling Gaussian random measurements. This work led further to several other non-convex algorithms based on spectral initialization. A common thread is first choosing a good initial guess through spectral initialization, and then solving an optimization model through gradient descent. Two widely used optimization estimators are the intensity-based loss

$$\min_{\boldsymbol{z} \in \mathbb{R}^n} F(\boldsymbol{z}) = \sum_{j=1}^m \left(|\langle \boldsymbol{a}_j, \boldsymbol{z} \rangle|^2 - y_j^2 \right)^2; \tag{1.1}$$