Non-Convex Phase Retrieval from STFT Measurements

Article · July 2016

CITATIONS
5

READS
18

2 authors, including:

Tamir Bendory
Technion - Israel Institute of Technology
14 PUBLICATIONS **79** CITATIONS

SEE PROFILE

Non-Convex Phase Retrieval from STFT Measurements

Tamir Bendory and Yonina C. Eldar, *Fellow IEEE*

***Abstract*—The problem of recovering a one-dimensional signal from its Fourier transform magnitude, called phase retrieval, is ill-posed in most cases. We consider the closely-related problem of recovering a signal from its phaseless short-time Fourier transform (STFT) measurements. This problem arises naturally in several applications, such as ultra-short pulse measurements and ptychography. The redundancy offered by the STFT enables unique recovery under mild conditions. We show that in some cases the unique solution can be obtained by the principle eigenvector of a matrix, constructed as the solution of a simple least-squares problem. When these conditions are not met, we suggest to use the principle eigenvector of this matrix as an initialization of a gradient algorithm, minimizing a non-convex loss function. We prove that under appropriate conditions, this initialization is close to the underlying signal. We analyze the geometry of the loss function and show empirically that the gradient algorithm converges to the underlying signal even with small redundancy in the measurements. Additionally, the algorithm is robust to noise. In contrast to previous works, our method is efficient and enjoys theoretical guarantees.**

***Index Terms*—phase retrieval, short-time Fourier transform, gradient descent, non-convex optimization, spectral initialization, least-squares, ptychography**

I. Introduction

The problem of recovering a signal from its Fourier transform magnitude arises in many areas in engineering and science, such as optics, X-ray crystallography, speech recognition, blind channel estimation and astronomy [20], [44], [29], [25], [2], [14]. This problem is called *phase retrieval* and can be viewed as a special case of a quadratic system of equations. The later area received considerable attention recently, partly due to its strong connections with the fields of compressed sensing and matrix completion, see for instance [5], [9], [7], [12], [6], [43]. We refer the reader to contemporary surveys of the phase retrieval problem in [38], [22].

Phase retrieval for one-dimensional (1D) signals is an ill-posed problem unless the signal has the minimum phase property [21], [36]. In this special case, the signal can be recovered by several tractable algorithms (see for instance Section 2.6 of [11]). Particularly, in [21] it was shown that a semi-definite program (SDP) relaxation achieves the optimal solution in a maximum likelihood sense. For general signals, two main approaches are followed. The first builds upon prior knowledge on the signal's support, such as sparsity or a portion of the underlying signal [15], [39], [33], [24], [37], [46]. The alternative strategy makes use of additional measurements. Such measurements can be obtained by structured illuminations and masks [7], [19] or by measuring the magnitude of the short-time Fourier transform (STFT). In [13], it was demonstrated that for the same number of measurements, the STFT magnitude leads to better performance than an oversampled discrete Fourier transform (DFT).

This paper deals with the problem of recovering a 1D signal from its STFT magnitude. The STFT of a 1D signal $\mathbf{x} \in \mathbb{C}^N$ can be interpreted as the Fourier transform of the signal multiplied by a real sliding window $\mathbf{g}$ of length W. Then, the STFT is defined as

$$\mathbf{X}[m,k] := \sum_{n=0}^{N-1} \mathbf{x}[n]\mathbf{g}[mL-n]e^{-2\pi jkn/N}, \qquad \text{(I.1)}$$

where $k = 0,\ldots,N-1$, $m = 0,\ldots,\left\lceil \frac{N}{L} \right\rceil - 1$ and L determines the separation in time between adjacent sections. In the sequel, all indices should be considered as modulo the signal's length N. We assume that $\mathbf{x}$ and $\mathbf{g}$ are periodically extended over the boundaries in (I.1) and that L divides N.

The problem of recovering a signal from its STFT magnitude $|\mathbf{X}[m,k]|^2$ arises in several applications in optics and speech processing [30], [18]. Particularly, it serves as the model for a popular variant of an ultra-short laser pulse measurement technique called Frequency-Resolved Optical Gating (referred to as X-FROG) [42]. Another application is ptychography in which a moving probe is used to sense multiple diffraction measurements [34], [28], [27].

Several algorithms were suggested to recover a signal from the magnitude of its STFT. The classic method, called Griffin-Lim algorithm (GLA) [18], is a modification of the alternating projection (or reduction error) algorithms of Gerchberg and Saxton [16] and Fineup [15]. The properties of this algorithm are not well-understood (for analysis of alternating projection algorithms in phase retrieval, see [28]). In [23], the authors prove that a non-vanishing signal can be recovered by an SDP with maximal overlapping between adjacent windows ($L = 1$). They also demonstrate empirically that the algorithm works well with less restrictive requirements on the window and is robust to noise (if the noise level in each section is known). Despite the appealing numerical performance, solving an SDP requires high computational resources. In contrast, we aim at developing a phase retrieval algorithm that reflects a practical setup, is computationally efficient and enjoys theoretical guarantees.

We begin by taking the 1D DFT of the acquired information with respect to the frequency variable (the second variable

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 646804-ERCCOG-BNYQ, and from the Israel Science Foundation under Grant no. 335/14. TB was partially funded by the Andrew and Erna Finci Viterbi Fellowship.

of the STFT). This transformation reveals the underlying structure of the data and greatly simplifies the analysis. As a direct consequence, we show that for $L = 1$ and sufficiently long windows $W \geq \left\lceil \frac{N+1}{2} \right\rceil$ (and some mild additional conditions), one can recover the signal by extracting the principle eigenvector of a designed matrix, constructed as the solution of a simple least-squares (LS) problem. We refer to this matrix as the *approximation matrix* since it approximates the correlation matrix $\mathbf{X} := \mathbf{x}\mathbf{x}^*$. When the conditions for a closed-form solution are not met, we propose using the principle eigenvector of the approximation matrix to initialize a gradient descent (GD) algorithm that minimizes a non-convex loss function. Our approach deviates in two important aspects from the recent line of work in non-convex phase retrieval [8], [10], [31], [47], [45]. First, all these papers focus their attention on the setup of phase retrieval with random sensing vectors and rely heavily on statistical considerations, while we consider a deterministic framework. Second, we construct our approximation matrix by the solution of a LS problem, whereas the aforementioned papers take a superposition of the measurements to approximate $\mathbf{X}$.

The properties of a GD algorithm depend heavily on the initialization method and the geometry of the loss function. For $L = 1$, we estimate the distance between the proposed initialization and the global minimum, which decays to zero as W tends to $\frac{N+1}{2}$. If the signal has unit module entries, then a slight modification of our initialization recovers the signal exactly for $W \geq 2$. In the later case, we also prove the existence of a *basin of attraction* around the global minimum of the loss function and estimate its size. In the basin of attraction, the GD algorithm is guaranteed to converge to a global minimum at a geometrical rate. We show empirically that a basin of attraction also exists for general (not necessarily unit module entries) signals. We note that while the theoretical guarantees of the algorithm are limited, the experimental performance is significantly better. Particularly, the algorithm performs well for general signals and is robust in the presence of noise.

The paper is organized as follows. We begin in Section II by formulating mathematically the problem of phase retrieval from STFT magnitude measurements. In Section III we discuss the uniqueness of the problem and present conditions under which it has a closed-form LS solution. Additionally, we present an algorithm that recovers signals with unit module entries under mild conditions. Section IV presents the GD algorithm with the proposed initialization. Section V shows numerical results and Section VI presents our theoretical findings. Related proofs are provided in Sections VII. Section VIII concludes the paper, discusses its main implications and draws potential future research directions.

Throughout the paper we use the following notation. Boldface small and capital letters denote vectors and matrices, respectively. We use $\mathbf{Z}^T$ and $\mathbf{Z}^*$ for the transpose and Hermitian of a matrix $\mathbf{Z}$; similar notation is used for vectors. We further use $\mathbf{Z}^\dagger$ and $\mathrm{tr}(\mathbf{Z})$ for the Moore–Penrose pseudo-inverse and the trace of the matrix $\mathbf{Z}$, respectively. The ℓth circular diagonal of a matrix $\mathbf{Z}$ is denoted by $\mathrm{diag}(\mathbf{Z}, \ell)$. Namely, $\mathrm{diag}(\mathbf{Z}, \ell)$ is a column vector with entries $\mathbf{Z}\left[i, (i+\ell) \bmod N\right]$ for $i = 0, \ldots, N-1$. We define the sign of a complex number a as $\mathrm{sign}\,(a) := \frac{a}{|a|}$ and use $'\circ'$ and $'*'$ for the Hadamard (point-wise) product and convolution, respectively. The set of all complex (real) signals of length N whose entries have module $a > 0$ are denoted by $\mathbb{C}_a^N$ ($\mathbb{R}_a^N$). Namely, $\mathbf{z} \in \mathbb{C}_a^N$ implies that $|\mathbf{z}[n]| = a$ for all n.

II. Problem Formulation

We aim at recovering the underlying signal $\mathbf{x}$ from the magnitude of its STFT, i.e. from measurements

$$\mathbf{Z}[m,k] = \left|\mathbf{X}[m,k]\right|^2. \tag{II.1}$$

Note that the signals $\mathbf{x}$ and $\mathbf{x}e^{j\phi}$ yield the same measurements for any *global phase* $\phi \in \mathbb{R}$ and therefore the phase ϕ cannot be recovered by any method. This *global phase ambiguity* leads naturally to the following definition:

Definition II.1. The distance between two vectors is defined as

$$d\left(\mathbf{z}, \mathbf{x}\right) = \min_{\phi \in [0, 2\pi)} \left\| \mathbf{z} - \mathbf{x}e^{j\phi} \right\|_2.$$

If $d\left(\mathbf{z}, \mathbf{x}\right) = 0$ then we say that $\mathbf{x}$ and $\mathbf{z}$ are equal *up to global phase*. The phase $\phi \in [0, 2\pi)$ attaining the minimum is denoted by $\phi(\mathbf{z})$, i.e.

$$\phi(\mathbf{z}) = \arg \min_{\phi \in [0, 2\pi)} \left\| \mathbf{z} - \mathbf{x}e^{j\phi} \right\|_2.$$

Instead of treating the measurements (II.1) directly, we consider the acquired data in a transformed domain by taking its 1D DFT with respect to the frequency variable (normalized by $\frac{1}{N}$). Then, our measurement model reads

$$\begin{aligned} \mathbf{Y}[m,\ell] &= \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{Z}[m,k] e^{-2\pi j k\ell/N} \\ &= \sum_{n=0}^{N-1} \mathbf{x}[n]\mathbf{x}^*[n+\ell]\mathbf{g}[mL-n]\mathbf{g}[mL-n-\ell]. \end{aligned} \tag{II.2}$$

When $W \leq \ell \leq (N - W)$, we have $\mathbf{Y}[m,\ell] = 0$ for all m. In this sense, $\mathbf{Y}[m,\ell]$ can be interpreted as a "W - bandlimited" function. The transformation (II.2) simplifies the quadratic system of equations and reveals its underlying structure. Observe that for fixed m, $\mathbf{Y}[m,\ell]$ is simply the auto-correlation of $\mathbf{x} \circ \mathbf{g}_{mL}$, where $\mathbf{g}_{mL} := \{\mathbf{g}[mL-n]\}_{n=0}^{N-1}$.

We will make repetitive use of two representations of (II.2). The first is based on a matrix formulation. Let $\mathbf{D}_{mL} \in \mathbb{R}^{N \times N}$ be a diagonal matrix composed of the entries of $\mathbf{g}_{mL}$. Let $\mathbf{P}_\ell$ be a matrix that shifts (circularly) the entries of a vector by ℓ locations, namely, $\left(\mathbf{P}_\ell \mathbf{x}\right)[n] = \mathbf{x}\left[n+\ell\right]$. Then, the correlation matrix $\mathbf{X} := \mathbf{x}\mathbf{x}^*$ is mapped linearly to $\mathbf{Y}[m,\ell]$ as follows:

$$\begin{aligned} \mathbf{Y}[m,\ell] &= \left(\mathbf{D}_{mL-\ell}\mathbf{D}_{mL}\mathbf{P}_\ell\mathbf{x}\right)^* \mathbf{x} \\ &= \mathbf{x}^*\mathbf{H}_{m,\ell}\mathbf{x} \\ &= \mathrm{tr}\left(\mathbf{X}\mathbf{H}_{m,\ell}\right), \end{aligned} \tag{II.3}$$

where

$$\mathbf{H}_{m,\ell} := \mathbf{P}_{-\ell}\mathbf{D}_{mL}\mathbf{D}_{mL-\ell}. \tag{II.4}$$

Observe that $\mathbf{P}_\ell^T = \mathbf{P}_{-\ell}$ and $\mathbf{H}_{m,\ell} = 0$ for $W \leq \ell \leq (N - W)$.

An alternative useful representation of (II.2) is as multiple systems of linear equations. For fixed $\ell \in [-(W-1),\dots,(W-1)]$ we can write

$$\mathbf{y}_\ell = \mathbf{G}_\ell \mathbf{x}_\ell, \tag{II.5}$$

where $\mathbf{y}_\ell := \{\mathbf{Y}[m,\ell]\}_{m=0}^{\frac{N}{L}-1}$ and $\mathbf{x}_\ell := \operatorname{diag}(\mathbf{X},\ell)$. The (m,n)th entry of the matrix $\mathbf{G}_\ell \in \mathbb{R}^{\frac{N}{L}\times N}$ is given by $\mathbf{g}[mL-n]\mathbf{g}[mL-n-\ell]$. For $L=1$, $\mathbf{G}_\ell$ is a circulant matrix. We recall that a circulant matrix is diagonalized by the DFT matrix, namely, it can be factored as $\mathbf{G}_\ell = \mathbf{F}^*\mathbf{\Sigma}_\ell\mathbf{F}$, where $\mathbf{F}$ is the DFT matrix and $\mathbf{\Sigma}_\ell$ is a diagonal matrix, whose entries are given by the DFT of the first column of $\mathbf{G}_\ell$. In this case, the first column is given by $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$, where $\mathbf{g} := \{\mathbf{g}[n]\}_{n=0}^{N-1}$. Therefore the matrix $\mathbf{G}_\ell$ is invertible if and only if the DFT of $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$ is non-vanishing.

Our problem of recovering $\mathbf{x}$ from the measurements (II.1) can therefore be equivalently posed as a constrained LS problem, i.e.

$$\begin{aligned}&\min_{\tilde{\mathbf{x}}\in\mathbb{C}^N} \sum_{\ell=-(W-1)}^{W-1} \left\|\mathbf{y}_\ell - \mathbf{G}_\ell \operatorname{diag}\left(\tilde{\mathbf{X}},\ell\right)\right\|_2^2\\ &\text{subject to}\quad \tilde{\mathbf{X}} = \tilde{\mathbf{x}}\tilde{\mathbf{x}}^*.\end{aligned} \tag{II.6}$$

This problem is non-convex due to the quadratic constraint and is equivalent to

$$\begin{aligned}&\min_{\tilde{\mathbf{X}}\in\mathcal{H}^N} \sum_{\ell=-(W-1)}^{W-1} \left\|\mathbf{y}_\ell - \mathbf{G}_\ell \operatorname{diag}\left(\tilde{\mathbf{X}},\ell\right)\right\|_2^2\\ &\text{subject to}\quad \tilde{\mathbf{X}} \succeq 0,\quad \operatorname{rank}\left(\tilde{\mathbf{X}}\right) = 1,\end{aligned}$$

where $\mathcal{H}^N$ is the set of all Hermitian matrices of size $N\times N$. In the spirit of [17], [43], [5], [39], STFT phase retrieval can then be relaxed to a tractable SDP by dropping the rank constraint. While an SDP relaxation approach [23] (with a different formulation) has shown good results for the recovery from phaseless STFT measurements, it requires solving the problem in a lifted domain with N^2 variables. We take a different route to reduce the computational load. In the next section, we show that (II.6) admits a unique solution under moderate conditions. We further show that it has a closed-form LS solution when the window $\mathbf{g}$ is sufficiently long. If the conditions for the LS solution are not met, then we suggest applying a GD algorithm. To initialize the GD, we approximate (II.6) in two stages by first solving the LS objective function and then extracting its principal eigenvector. The GD algorithm is the main interest of this paper.

III. Uniqueness and Basic Algorithms

A fundamental question in phase retrieval problems is whether the quadratic measurement operator of (II.1), or equivalently the non-convex problem (II.6), determines the underlying signal $\mathbf{x}$ uniquely (up to global phase, see Definition II.1). In other words, under what conditions on the window $\mathbf{g}$ is the non-linear transformation that maps $\mathbf{x}$ to $\mathbf{Z}$ is injective. Before treating this question, we introduce some basic window definitions:

Definition III.1. A window $\mathbf{g}$ is called a *rectangular window of length* W if $\mathbf{g}[n] = 1$ for all $n = 0,\dots,W-1$ and zero elsewhere. It is a *non-vanishing window of length* W if $\mathbf{g}[n] \neq 0$ for all $n = 0,\dots,W-1$ and zero elsewhere.

According to (II.6), the injectivity of the measurement operator is related to the window's length W and the invertibility of the matrices $\mathbf{G}_\ell$ for $|\ell| < W$. For that reason, we give special attention to windows for which the associated matrices are invertible.

Definition III.2. A window $\mathbf{g}$ is called an *admissible window of length* W if for all $\ell = -(W-1),\dots,(W-1)$ the following two equivalent properties hold:

1) The DFT of the vector $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$ is non-vanishing.
2) The associated circulant matrices $\mathbf{G}_\ell$ as given in (II.5) are invertible.

The family of admissible windows is quite large. For instance, if α and N are coprime numbers for all $\alpha = 2,\dots,W$, then a rectangular window of length W is an admissible window as well:

Claim III.3. A rectangular window $\mathbf{g}$ of length $2 \le W \le N/2$ is an admissible window of length W if α and N are co-prime numbers for all $\alpha = 2,\dots,W$. This holds trivially if N is a prime number.

Proof: Observe that $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$ is a rectangular window of length $(W-|\ell|)$ for $\ell = -(W-1)\dots,(W-1)$. The DFT of a rectangular window of size $(W-|\ell|)$ is a Dirichlet kernel which is non-vanishing if $(W-|\ell|)$ and N are co-prime. ■

Equipped with the aforementioned definitions, we are ready to analyze the uniqueness of the measurement operator for the case[1] $L = 1$. Our results are constructive in the sense that their proofs provide an explicit scheme to recover the signal.

Our first uniqueness result concerns non-vanishing signals. We say that a signal $\mathbf{z}$ is non-vanishing if $\mathbf{z}[n] \neq 0$ for all n. In this case, the magnitude of the STFT determines the underlying signal uniquely under mild conditions. This result was already derived in [3] based on different considerations. Nevertheless, our result comes with an explicit recovery scheme (see Appendix A).

Proposition III.4. *Let $L = 1$. Suppose that $\mathbf{x}$ is non-vanishing and that the DFT of $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$ is non-vanishing for $\ell = 0, 1$. Then, $|\mathbf{X}[m,k]|^2$ determines $\mathbf{x}$ uniquely (up to global phase).*

Proof: See Appendix A. ■

A similar uniqueness result was derived in [13]. There, it is required that the DFT of $|\mathbf{g}[n]|^2$ is non-vanishing, $N \ge 2W-1$ and N and $W-1$ are co-prime numbers.

In the special case in which the signal is known to have unit module entries, the signal can be recovered as the principle eigenvector of a designed matrix as follows:

Proposition III.5. *Let $L = 1$. Suppose that $\mathbf{x} \in \mathbb{C}^N_{1/\sqrt{N}}$ and that $\mathbf{g}$ is an admissible window of length $W \ge 2$. Fix $M \in$*

[1] For uniqueness results for the cases of $L > 1$, see [30], [23].

$[1,\dots,W-1]$ *and let* $\mathbf{X}_0$ *be a matrix obeying*

$$\textit{diag}\left(\mathbf{X}_0,\ell\right)=\begin{cases}\mathbf{G}_\ell^{-1}\mathbf{y}_\ell, & \ell=0,M,\\ 0, & \textit{otherwise},\end{cases} \tag{III.1}$$

where $\mathbf{G}_\ell$ *and* $\mathbf{y}_\ell$ *are defined in (II.5). Then,* $\mathbf{x}$ *(up to global phase) is a principle eigenvector of* $\mathbf{X}_0$.

Proof: See Appendix B. ∎

For general signals (not necessarily non-vanishing) and $L=1$, we next derive a LS algorithm that stably recovers any complex signal if the window is sufficiently long. In the absence of noise, the recovery is exact (up to global phase). The method, summarized in Algorithm 1, is based on constructing an approximation matrix $\mathbf{X}_0$ that approximates the correlation matrix $\mathbf{X}:=\mathbf{x}\mathbf{x}^*$. The ℓth diagonal of $\mathbf{X}_0$ is chosen as the solution of the LS problem $\min_{\tilde{\mathbf{x}}\in\mathbb{C}^N}\|\mathbf{y}_\ell-\mathbf{G}_\ell\tilde{\mathbf{x}}\|_2$ (see (II.5)). If the matrix $\mathbf{G}_\ell$ is invertible, then

$$\operatorname{diag}\left(\mathbf{X}_0,\ell\right)=\mathbf{G}_\ell^{-1}\mathbf{y}_\ell=\operatorname{diag}\left(\mathbf{X},\ell\right).$$

Therefore, when all matrices $\mathbf{G}_\ell$ are invertible, $\mathbf{X}_0=\mathbf{X}$. In order to estimate $\mathbf{x}$, the (unit-norm) principle eigenvector of $\mathbf{X}_0$ is normalized by

$$\alpha=\sqrt{\sum_{n\in P}\left(\mathbf{G}_0^\dagger\mathbf{y}_0\right)[n]}, \tag{III.2}$$

where $P:=\{n:(\mathbf{G}_0^\dagger\mathbf{y}_0)[n]>0\}$. If $\mathbf{G}_0$ is invertible then

$$\sum_{n=0}^{N-1}\left(\mathbf{G}_0^{-1}\mathbf{y}_0\right)[n]=\sum_{n=0}^{N-1}\left(\operatorname{diag}\left(\mathbf{X},0\right)\right)[n]=\|\mathbf{x}\|_2^2=\lambda_0,$$

where λ_0 is the top eigenvector of $\mathbf{X}$. If $\mathbf{G}_0$ is not invertible or in the presence of noise, some terms of the vector $\mathbf{G}_0^\dagger\mathbf{y}_0$ might be negative. In this case, we estimate $\|\mathbf{x}\|_2$ by summing only the positive terms (the set P in (III.2)). Note that all matrix inversions can be performed efficiently using the FFT due to the circular structure of $\mathbf{G}_\ell$.

The following proposition shows that for $L=1$ Algorithm 1 recovers the underlying signal if the window is sufficiently long and satisfies some additional technical conditions. In [3], an equivalent uniqueness result was derived but without providing an algorithm. Algorithm 1 is equivalent to the discretized version of Wigner deconvolution that was suggested previously without theoretical analysis in [35], [48].

Proposition III.6. *Let* $L=1$ *and suppose that* $\mathbf{g}$ *is an admissible window of length* $W\geq\left\lceil\frac{N+1}{2}\right\rceil$ *(see Definition III.2). Then, Algorithm 1 recovers any complex signal uniquely (up to global phase) and efficiently.*

Proof: See Appendix C. ∎

In many cases, the window is shorter than $\left\lceil\frac{N+1}{2}\right\rceil$ so that (II.6) does not have a closed-form LS solution. In these cases, we suggest to recover the signal by minimizing a non-convex loss function using a gradient algorithm. In order to initialize the gradient algorithm, we use the same LS-based algorithm. However, for short windows we cannot estimate $\operatorname{diag}(\mathbf{X},\ell)$ for $\ell=W,\dots,(N-W)$ as the matrices $\mathbf{G}_\ell$ are simply zero. Nonetheless, we will show by both theoretical results and numerical experiments that under appropriate conditions, the principle eigenvector of the approximation matrix $\mathbf{X}_0$, with appropriate normalization, is a good estimate of $\mathbf{x}$.

Algorithm 1 Least-squares algorithm for $L=1$

Input: The measurements $\mathbf{Z}[m,k]$ as given in (II.1).
Output: $\mathbf{x}_0$: estimation of $\mathbf{x}$.

1) Compute $\mathbf{Y}[m,\ell]$, the 1D DFT with respect to the second variable of $\mathbf{Z}[m,k]$ as given in (II.2).
2) Construct a matrix $\mathbf{X}_0$ such that

$$\operatorname{diag}\left(\mathbf{X}_0,\ell\right)=\begin{cases}\mathbf{G}_\ell^\dagger\mathbf{y}_\ell & \ell=-\left(W-1\right),\cdots,\left(W-1\right),\\ 0 & \text{otherwise},\end{cases}$$

where $\mathbf{G}_\ell\in\mathbb{R}^{N\times N}$ are defined in (II.5).
3) Let $\mathbf{x}_p$ be the principle (unit-norm) eigenvector of $\mathbf{X}_0$. Then,

$$\mathbf{x}_0=\sqrt{\sum_{n\in P}\left(\mathbf{G}_0^\dagger\mathbf{y}_0\right)[n]}\mathbf{x}_p,$$

where $P:=\left\{n:\left(\mathbf{G}_0^\dagger\mathbf{y}_0\right)[n]>0\right\}$.

IV. Gradient Descent Algorithm

In this section we present our main algorithmic approach to recover a signal from its STFT magnitude (II.1). Recall that by taking the DFT with respect to the frequency variable, the measurement model reads $\mathbf{Y}[m,\ell]=\mathbf{x}^*\mathbf{H}_{m,\ell}\mathbf{x}$, where $\mathbf{H}_{m,\ell}$ is defined in (II.4). It is therefore natural to minimize the following non-convex loss function (frequently called non-convex least-squares or empirical risk):

$$f(\mathbf{z})=\frac{1}{2}\sum_{m=0}^{N/L-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^*\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{Y}[m,\ell]\right)^2. \tag{IV.1}$$

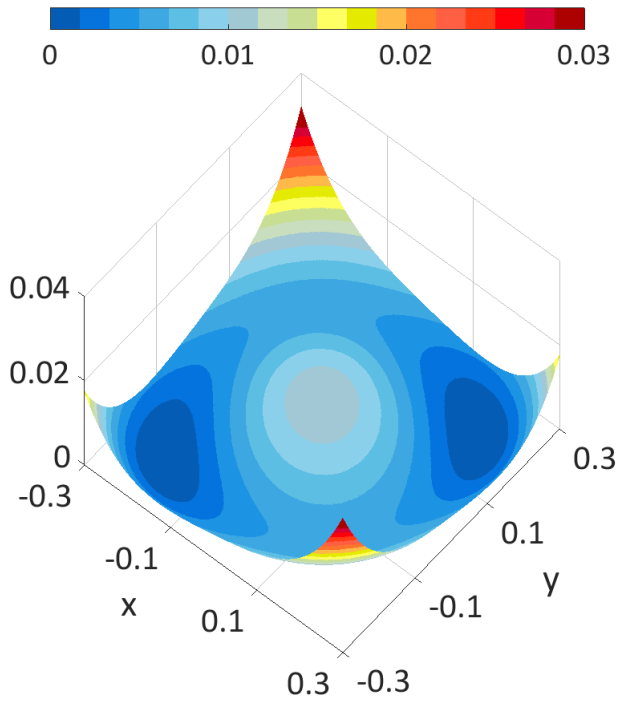
Figure IV.1 presents the two-dimensional (first two variables) plane of the loss function (IV.1) for the signal $\mathbf{x}=[0.2,0.2,0,0,0]$ (i.e. $N=5$) with $L=1$ and a rectangular window of length $W=2$. The function has no sharp transitions and contains two saddle points and two global minima (as a result of the global phase ambiguity). Accordingly, in this specific case it seems that a gradient descent algorithm will converge to a global minimum from almost any arbitrary initialization (see also [26]). While this phenomenon does not occur for any arbitrary parameter selection, this example motivates applying a gradient algorithm directly on the non-convex loss function (for a similar demonstration of the loss function with random sensing vectors, see [41]). To keep the framework simple, we focus here on real signals. If the signal is complex, we replace the inner term of (IV.1) by $|\mathbf{z}^*\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{Y}[m,\ell]|^2$.

We suggest minimizing (IV.1) by employing a gradient algorithm, where the kth iteration takes on the form

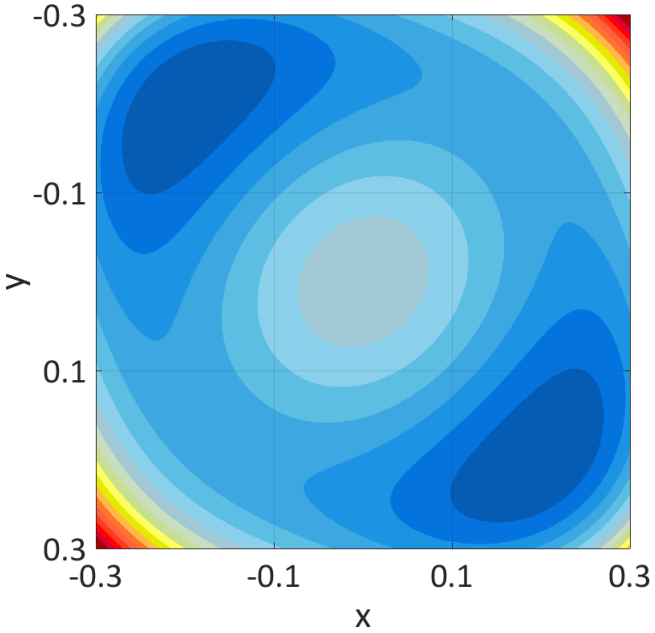
$$\mathbf{x}_k=\mathbf{x}_{k-1}-\mu\nabla f\left(\mathbf{x}_{k-1}\right),$$

for step size μ. Direct computation of the gradient gives

$$\nabla f(\mathbf{z})=\sum_{m=0}^{N/L-1}\sum_{\ell=-(W-1)}^{W-1}\left(h(\mathbf{z})-\mathbf{Y}[m,\ell]\right)\nabla h(\mathbf{z}), \tag{IV.2}$$



(a) Side view



(b) View from above

Figure IV.1: The two-dimensional (first two variables) plane of the loss function (IV.1) of the signal $\mathbf{x} = [0.2, 0.2, 0, 0, 0]$ (i.e. $N = 5$) with $L = 1$ and a rectangular window of length $W = 2$.

Algorithm 2 Gradient descent algorithm

Input: The measurements $\mathbf{Z}[m,k]$ as given in (II.1) and (optional) thresholding parameter $B > 0$.

Output: Estimation of $\mathbf{x}$.

1) Initialization by Algorithm 1 (for $L = 1$) or Algorithm 3 (for $L > 1$).
2) Apply the update rule until convergence:
   a) (gradient step)
   $$\tilde{\mathbf{x}}_k = \mathbf{x}_{k-1} - \mu \nabla f\left(\mathbf{x}_{k-1}\right),$$
   for step size μ and ∇f given in (IV.2).
   b) (optional thresholding)
   $$\mathbf{x}_k[n] = \begin{cases} \tilde{\mathbf{x}}_k[n] & \text{if } |\tilde{\mathbf{x}}_k[n]| \le B, \\ B \cdot \text{sign}\left(\tilde{\mathbf{x}}_k[n]\right) & \text{if } |\tilde{\mathbf{x}}_k[n]| > B. \end{cases}$$

where

$$h(\mathbf{z}) := \mathbf{z}^T \mathbf{H}_{m,\ell} \mathbf{z},$$
$$\nabla h(\mathbf{z}) = \left(\mathbf{H}_{m,\ell} + \mathbf{H}_{m,\ell}^T\right) \mathbf{z}.$$

The gradient step is followed by an optional thresholding that can be used when the signal is known to be bounded. The algorithm is summarized in Algorithm 2. The code is publicly available on

http://webee.technion.ac.il/Sites/People/YoninaEldar

A. *Initialization for* $L = 1$

While the GD method suggests a simple minimization procedure, the function (IV.1) is non-convex and therefore it is not clear whether an arbitrary initialization will converge to a global minimum. When $L = 1$, we propose initializing the GD algorithm by using Algorithm 1. As explained in Section III, for $W \ge \left\lceil \frac{N+1}{2} \right\rceil$ the algorithm returns $\mathbf{x}$ exactly. However, when $W < \left\lceil \frac{N+1}{2} \right\rceil$, $\mathbf{G}_\ell = 0$ for $\ell = W, \ldots, (N - W)$ so that the output is not necessarily $\mathbf{x}$. Nevertheless, in Section VI we provide theoretical guarantees establishing that under appropriate conditions, this initialization results in a good approximation of the underlying signal $\mathbf{x}$.

B. *Initialization for* $L > 1$

Until now we focused on maximal overlapping between adjacent windows $L = 1$. When $L > 1$, (II.5) results in an underdetermined system of equations as $\mathbf{y}_\ell \in \mathbb{R}^{\frac{N}{L}}$, $\mathbf{G}_\ell \in \mathbb{R}^{\frac{N}{L} \times N}$ and $\mathbf{x}_\ell \in \mathbb{R}^N$. The LS solution $\mathbf{G}_\ell^\dagger \mathbf{y}_\ell$ is the vector with minimal ℓ_2 norm among the set of feasible solutions. This approximation is quite poor in general.

We notice that the measurements $\mathbf{y}_\ell$ are a downsampled version by a factor L of the case of maximal overlapping ($L = 1$). Therefore, we suggest upsampling $\mathbf{y}_\ell$ to approximate the case of maximal overlap based on the averaging nature of the window $\mathbf{g}$. In order to motivate our approach, we start by considering an ideal situation. Suppose that for some ℓ, the

DFT of the first column of $\mathbf{G}_\ell$, denoted by $\hat{\mathbf{g}}_\ell$, is an ideal low-pass with bandwidth N/L_{BW}. Namely,

$$\hat{\mathbf{g}}_\ell[k] = \begin{cases} 1, & k = 0, \dots N/L_{BW} - 1, \\ 0, & \text{otherwise.} \end{cases}$$

The following lemma states that in this case, no information is lost by taking $L = L_{BW}$ compared to taking maximal overlap $L = 1$. Moreover, it suggests to upsample the measurement vector by expansion and low-pass interpolation. Our technique resembles standard upsampling arguments in digital signal processing (DSP) (see for instance Section 4.6 of [32]).

Lemma IV.1. *Let $\tilde{\mathbf{g}} := \{\mathbf{g}[(-n) mod N]\}_{n=0}^{N-1}$. Suppose that $\tilde{\mathbf{g}} \in \mathbb{R}^N$ is an ideal low-pass with bandwidth N/L and $\mathbf{y} = \mathbf{g} * \mathbf{x}$ for some $\mathbf{x} \in \mathbb{C}^N$ (or equivalently, $\mathbf{y} = \mathbf{G}\mathbf{x}$, where $\mathbf{G}$ is a circualnt matrix whose first column is $\tilde{\mathbf{g}}$). Let $\mathbf{y}_L \in \mathbb{C}^{N/L}$ be its L-downsampled version, i.e.*

$$\mathbf{y}_L[n] = \mathbf{y}[nL], \quad n = 0, \dots, N/L - 1.$$

Then, $\mathbf{y} = \left(\mathbf{F}_p^* \mathbf{F}_p\right) \tilde{\mathbf{y}}_L$, where

$$\tilde{\mathbf{y}}_L[n] = \begin{cases} \mathbf{y}_L[m], & n = mL, \\ 0, & \textit{otherwise,} \end{cases} \tag{IV.3}$$

and $\mathbf{F}_p$ is a partial Fourier matrix consisting of the first N/L rows of the DFT matrix $\mathbf{F}$.

Proof: See Appendix D. ∎

While Lemma IV.1 shows that no information is lost for an ideal low-pass window with bandwidth N/L, in practice we do not use these windows. Figure IV.2 presents the DFT magnitude of two typical choices of windows. In the spirit of standard DSP practice (see Section 4.6.1 of [32]) we approximate the low-pass interpolation of $\left(\mathbf{F}_p^* \mathbf{F}_p\right)$ as suggested in Lemma IV.1 by a simple smooth interpolation. This gives us better numerical results and reduces the computational complexity. In Section V we show some results with both linear and cubic interpolations. The figure also indicates the well-known fact that the wider the window is in the time domain, the smaller the bandwidth in the frequency domain. Hence, Lemma IV.1 implies that larger L requires longer window (larger W). This conclusion is supported by the numerical experiments of Section V and by previous results in the literature (see for instance [23]).

Following the upsampling stage, the algorithm proceeds as for $L = 1$ by extracting the principle eigenvector (with the appropriate normalization) of an approximation matrix. This initialization is summarized in Algorithm 3.

Algorithm 3 Least-squares initialization for $L > 1$

Input: The measurements $\mathbf{Z}[m, k]$ as given in (II.1) and a smooth interpolation filter $\mathbf{h}_L \in \mathbb{R}^N$ that approximates a low-pass filter with bandwidth N/L.

Output: $\mathbf{x}_0$: estimation of $\mathbf{x}$.

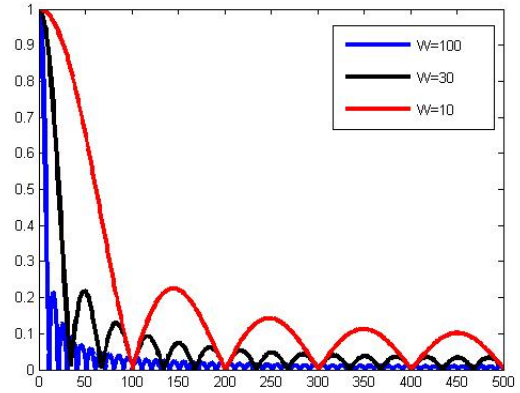
1) Compute $\mathbf{Y}[m, \ell]$, the 1D DFT with respect to the second variable of $\mathbf{Z}[m, k]$ as given in (II.2).
2) (upsampling) For each $\ell \in [-(W-1), \dots, (W-1)]$:
   a) Let $\mathbf{y}_\ell[m] := \{\mathbf{Y}[m, \ell]\}_{m=0}^{\frac{N}{L}-1}$ for fixed ℓ.
   b) (expansion)
   $$\tilde{\mathbf{y}}_\ell[n] := \begin{cases} \mathbf{y}_\ell[m], & n = mL, \\ 0, & \text{otherwise.} \end{cases}$$
   c) (interpolation)
   $$\bar{\mathbf{y}}_\ell = \tilde{\mathbf{y}}_\ell * \mathbf{h}_L.$$
3) Construct a matrix $\mathbf{X}_0$ such that
   $$\text{diag}\left(\mathbf{X}_0, \ell\right) = \begin{cases} \mathbf{G}_\ell^\dagger \bar{\mathbf{y}}_\ell & \ell = -(W-1), \cdots, (W-1), \\ 0 & \text{otherwise,} \end{cases}$$
   where $\mathbf{G}_\ell \in \mathbb{R}^{N \times N}$ are defined as in (II.5) for $L = 1$.
4) Let $\mathbf{x}_p$ be the principle (unit-norm) eigenvector of $\mathbf{X}_0$. Then,
   $$\mathbf{x}_0 = \sqrt{\sum_{n \in P} \left(\mathbf{G}_0^\dagger \mathbf{y}_0\right)[n]}\, \mathbf{x}_p,$$
   where $P := \left\{n : \left(\mathbf{G}_0^\dagger \mathbf{y}_0\right)[n] > 0\right\}$.

V. Numerical Results

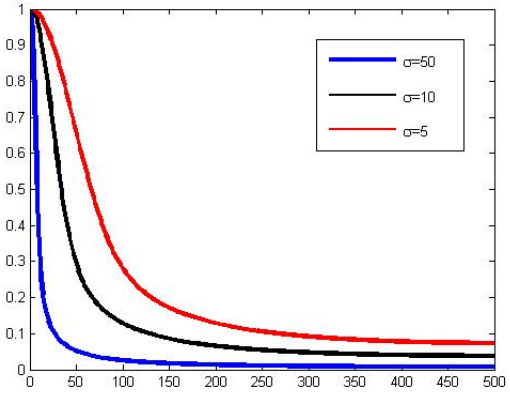
This section is devoted to numerical experiments examining the GD algorithm. In all experiments, the underlying signal was drawn from $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix. The measurements were contaminated with additive noise that was drawn from the same distribution with the appropriate standard deviation according to the desired signal to noise (SNR) ratio. The recovery error is computed by $\frac{d(\mathbf{x}, \hat{\mathbf{x}})}{\|\mathbf{x}\|_2}$, where $\hat{\mathbf{x}}$ is the estimated signal and the distance function $d(\cdot, \cdot)$ is defined in Definition II.1.

The first experiment examines the estimation quality of the initialization described in Algorithm 3. Figure V.1 presents the initialization error as a function of the window's length. We considered a Gaussian window defined by $\mathbf{g}[n] = e^{\frac{-n^2}{2\sigma^2}}$ and cubic and linear interpolations. For $n > 3\sigma$, we set the entries of the window to be zero so that $W = 3\sigma$. The results demonstrate the effectiveness of the smooth interpolation technique. For low values of L, it seems that the two interpolations achieve similar performance. For larger L, namely, less measurements, the cubic interpolation outperforms the linear interpolation.

The next experiment aims to estimate the *basin of attraction* of the GD algorithm in the absence of noise. That is to say, the area in which the GD algorithm will converge to a global minimum. To do that, we set the initialization point of Algorithm 2 to be $\mathbf{x}_0 = \mathbf{x} + \mathbf{z}$, where $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I})$ is the underlying signal. The perturbation vector $\mathbf{z}$ takes the values of $\pm\sigma$ (with random signs) for some $\sigma > 0$ so that $d(\mathbf{x}_0, \mathbf{x}) = \sqrt{N}\sigma$. Then, we applied the GD algorithm and checked whether the algorithm converges to $\mathbf{x}$. As can be seen in Figure V.2, the GD converges to the global minimum as long as $\sigma \leq 0.3$ for $L = 1, 2$ (the case of $L = 1$ is not presented in the figure) and $\sigma \leq 0.25$ for $L = 4$. A similar experiment showed that the size of the basin of attraction is independent of the value of W. These experimental results indicate that the actual basin of attraction is larger than our

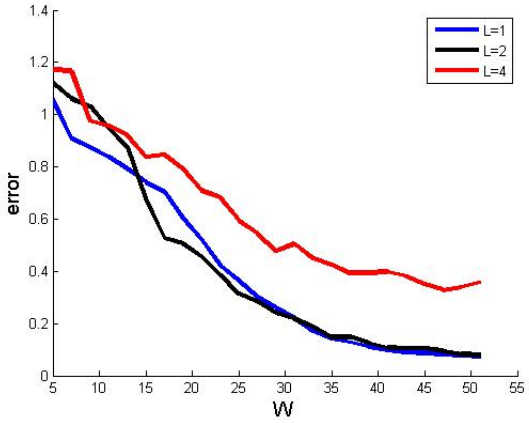


(a) DFT of rectangular windows with different lengths W.

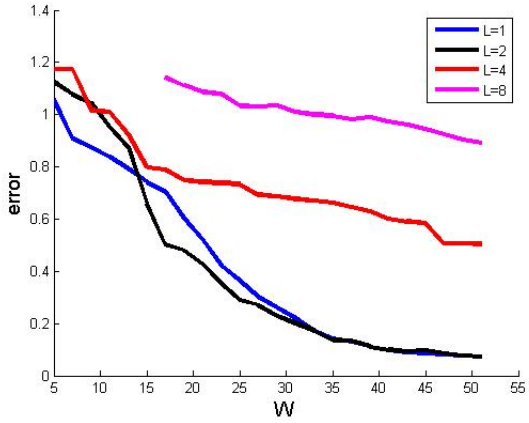


(b) DFT of Gaussian windows with different standard deviations σ.

Figure IV.2: The absolute value of the DFT of rectangular and Gaussian windows with $N = 1000$. We illustrate only $N/2$ entries due to symmetry considerations.



(a) Initialization with linear interpolation



(b) Initialization with cubic interpolation

Figure V.1: The average error (over 50 experiments) of the initialization of Algorithm 3 as a function of W and L. The experiments were conducted on signal of length $N = 101$ with a Gaussian window $e^{-\frac{n^2}{2\sigma^2}}$ and linear or cubic interpolation. The window length was set to be $W = 3\sigma$.

theoretical estimation in Section VI and Theorem VI.4.

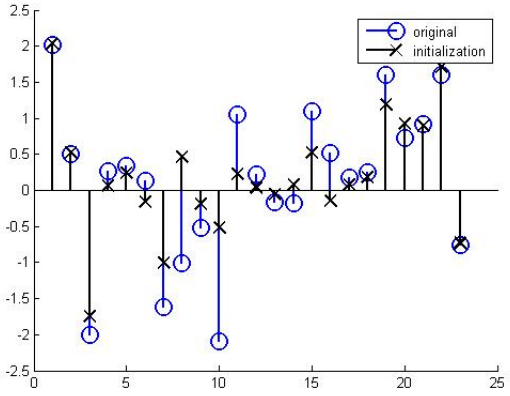
Figure V.3 shows a representative example of the performance of Algorithm 2. The experiment was conducted on a signal of length $N = 23$ with a rectangular window in a noisy environment of SNR$=$ 20 db. The step size was set to be $\mu = 5 \times 10^{-3}$. Figure V.4 presents the average recovery error as a function of the SNR level. We worked with signals of length $N = 53$, rectangular window of length $W = 19$ and $L = 2, 4$. We compared the algorithm's performance with the classical iterative Griffin-Lim Algorithm (GLA) [18]. As can be seen, the GD algorithm outperforms the GLA especially in high noise levels.

VI. Theory

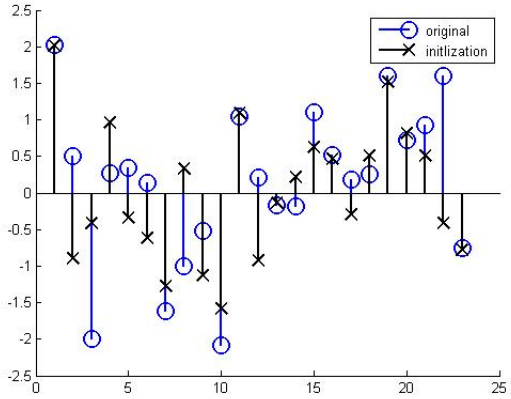
This section presents the theoretical contribution of this work, focusing on the case of maximum overlap between adjacent windows ($L = 1$). In Theorem VI.1 we first analyze the initialization algorithm presented in Algorithm 1 and estimate the distance between the initialization point and the global minimum. Next, we study the geometry of the loss function (IV.1), which controls the behavior of the GD algorithm. Particularly, we show in Theorem VI.4 the existence of a *basin of attraction* of size $\frac{1}{8\sqrt{N}W^2}$ around the global minimum for signals with unit module entries. In the basin of attraction, a GD algorithm is guaranteed to converge to a global minimum at a geometric rate.

A crucial condition for the success of gradient algorithms is that its initialization will be sufficiently close to the global minimum. The following result quantifies the estimation error of the proposed initialization presented in Algorithm 1 for bounded signals and $L = 1$. The error reduces to zero as W approaches $\frac{N+1}{2}$. The case of $L > 1$ is discussed briefly in Section IV.

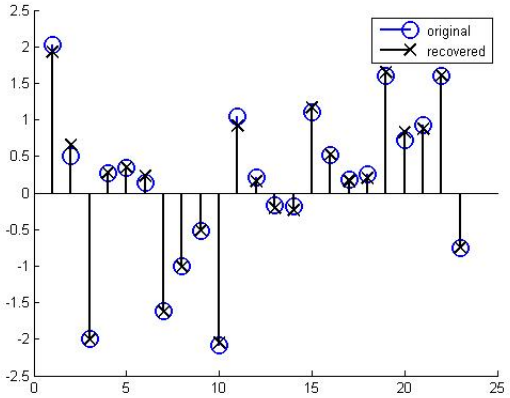
Theorem VI.1. *Suppose that $L = 1$, $\mathbf{g}$ is an admissible window of length $W \geq 2$ and that $\|\mathbf{x}\|_\infty \leq \sqrt{\frac{B}{N}}$ for some*



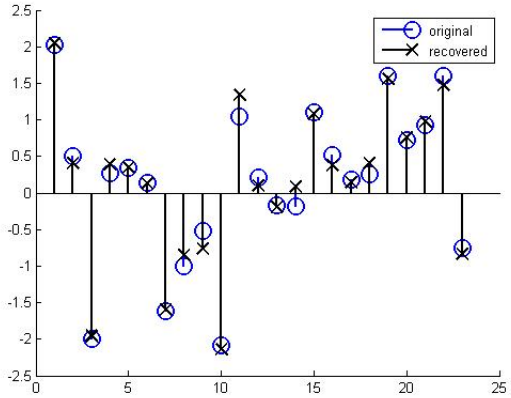
(a) Initialization with $W = 7$ and $L = 1$



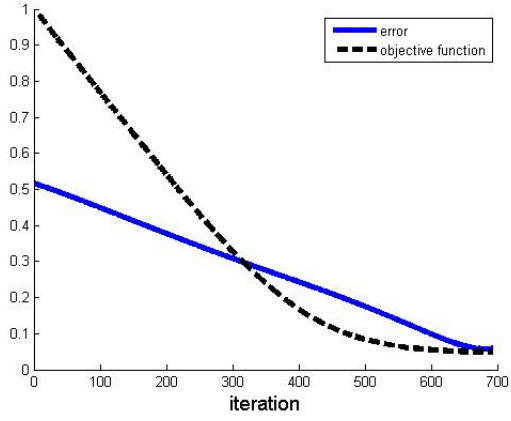
(b) Initialization with $W = 11$ and $L = 3$



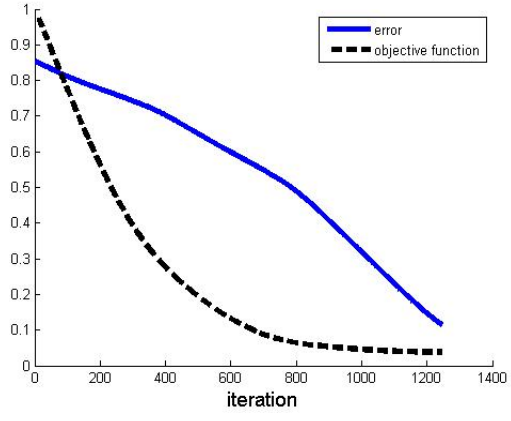
(c) Recovery by Algorithm 2 with $W = 7$ and $L = 1$



(d) Recovery by Algorithm 2 with $W = 11$ and $L = 3$

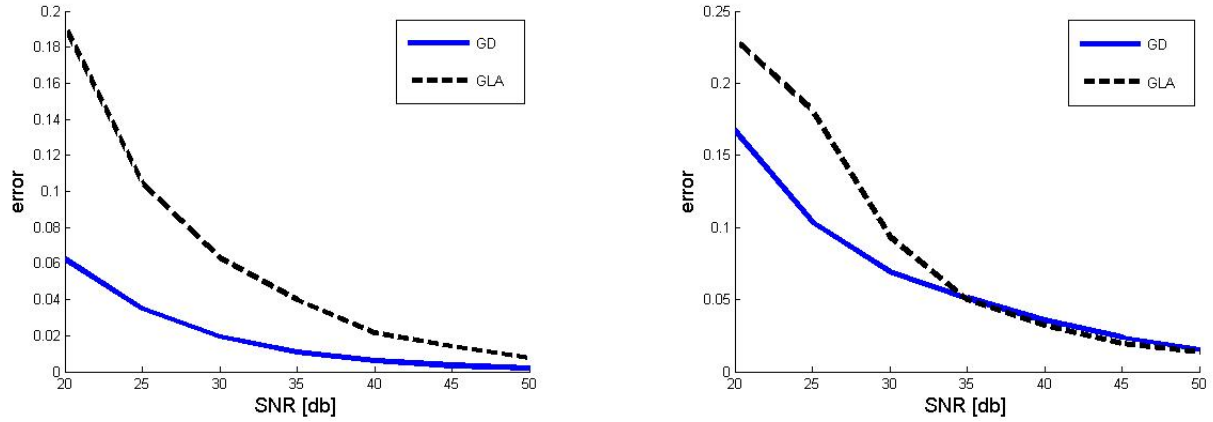


(e) The normalized objective function value and the error curves as a function of iterations for $W = 7$ and $L = 1$



(f) The normalized objective function value and the error curves as a function of iterations for $W = 11$ and $L = 3$

Figure V.3: Recovery of a signal of length $N = 23$ with a rectangular window in a noisy environment of SNR$= 20$ db. The step size was set to be $\mu = 5 \times 10^{-3}$. The experiments were conducted with $W = 7$ and $L = 1$ and $W = 11$ and $L = 3$ in the left and the right columns, respectively.



(a) Recovery error with $L = 2$

(b) Recovery error with $L = 4$

Figure V.4: The average recovery error (over 20 experiments) of the GD and GLA algorithms in the presence of noise. The experiments were conducted on signal of length $N = 53$ with a rectangular window of length $W = 19$, step size $\mu = 5 \times 10^{-3}$ and $L = 2, 4$.

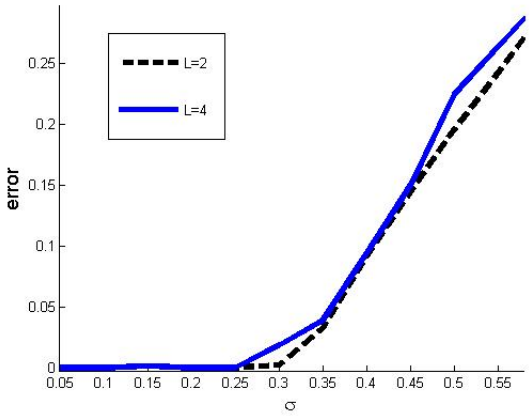


Figure V.2: The experiment was conducted on signal of length $N = 43$, $L = 1$ and step size of $\mu = 0.1$. The initialization point of the GD algorithm was set to be $\mathbf{x}_0 = \mathbf{x} + \mathbf{z}$, where $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I})$ is the underlying signal and the perturbation vector $\mathbf{z}$ takes the values of $\pm\sigma$ for some $\sigma > 0$ where the sign is randomly drawn. The figure presents the average recovery error over 100 experiments for each value of σ and L.

$0 < B \leq \frac{N}{2(N-2W+1)}$. *Then under the measurement model of (II.1), the initialization point as given in Algorithm 1 satisfies*

$$d^2(\mathbf{x}_0, \mathbf{x}) \leq \|\mathbf{x}\|_2^2 \left(1 - \sqrt{1 - 2B\frac{N-2W+1}{N}}\right).$$

Proof: See Section VII-A. ∎

The properties of the gradient algorithm rely on the geometry of the loss function (IV.1) near the global minimum. We begin the geometry analysis by two lemmas for signals with unit module entries that pave the way for the main result in Theorem VI.4. This result estimates the size of the basin of attraction of the loss function (IV.1).

The first result shows that the gradient of the loss function (IV.1), given explicitly in (IV.2), is bounded near its global minimum. This implies that the loss function is smooth. We consider here only the case of a rectangular window $\mathbf{g}$ of length W. The extension to non-vanishing windows of length W is straightforward (see Remark VII.1):

Lemma VI.2. *Suppose that* $\mathbf{x} \in \mathbb{R}^N_{1/\sqrt{N}}$, $\|\mathbf{z}\|_\infty \leq \frac{1}{\sqrt{N}}$ *and* $d(\mathbf{x}, \mathbf{z}) \leq \frac{1}{\sqrt{N}}$. *Let* $\mathbf{g}$ *be a rectangular window of length* W. *Then,* $\nabla f(\mathbf{z})$ *as given in (IV.2) satisfies*

$$\|\nabla f(\mathbf{z})\|_2 \leq \frac{8}{L} W^2 \sqrt{N} d(\mathbf{x}, \mathbf{z}).$$

Proof: See Section VII-B. ∎

The second lemma shows that the inner product between the gradient and the vector $\mathbf{z} - \mathbf{x}e^{j\phi(\mathbf{z})}$ is positive if $d(\mathbf{x}, \mathbf{z}) \leq \frac{1}{8\sqrt{N}W^2}$. This result implies that $-\nabla f(\mathbf{z})$ points approximately towards $\mathbf{x}$. As in Lemma VI.2, we consider for simplicity rectangular windows of length W. Yet, the analysis can be extended to non-vanishing windows of length W. In this case, the bounds are dependent on the dynamic range of $\mathbf{g}$ (for details, see Remark VII.2).

Lemma VI.3. *Suppose that* $L = 1$ *and* $\mathbf{g}$ *is a rectangular window of length* W. *For any* $\mathbf{x} \in \mathbb{R}^N_{1/\sqrt{N}}$ *and* $\|\mathbf{z}\|_\infty \leq \frac{1}{\sqrt{N}}$, *if* $d(\mathbf{x}, \mathbf{z}) \leq \frac{1}{8\sqrt{N}W^2}$, *then*

$$\left\langle \nabla f(\mathbf{z}), \mathbf{z} - \mathbf{x}e^{j\phi(\mathbf{z})} \right\rangle \geq \frac{W d^2(\mathbf{x}, \mathbf{z})}{2N},$$

where $\nabla f(\mathbf{z})$ *is given in (IV.2).*

Proof: See section VII-C. ∎

The last result quantifies the basin of attraction of the loss function (IV.1), namely, the area in which a gradient algorithm is guaranteed to converge to a global minimum at a geometric rate. As demonstrated in Figure V.2, in practice the basin of attraction is quite large and exists for a broad family of signals.

Theorem VI.4. *Let* $L = 1$ *and suppose that* $\mathbf{x} \in \mathbb{R}^N_{1/\sqrt{N}}$ *and* $\mathbf{g}$ *is a rectangular window of length* W. *Additionally, suppose that* $d(\mathbf{x}_0, \mathbf{x}) \leq \frac{1}{8\sqrt{N}W^2}$, *where* $\mathbf{x}_0$ *obeys* $\|\mathbf{x}_0\|_\infty \leq \frac{1}{\sqrt{N}}$.

Then, under the measurement model (II.1), Algorithm 2 with thresholding parameter $B = \frac{1}{\sqrt{N}}$ and step size $0 < \mu \leq 2/\beta$ achieves the following geometrical convergence

$$d^2\left(\mathbf{x}_k, \mathbf{x}\right) \leq \left(1 - \frac{2\mu}{\alpha}\right)^k d^2\left(\mathbf{x}_{k-1}, \mathbf{x}\right),$$

where $\alpha \geq \frac{4N}{W}$ and $\beta \geq 256N^2W^3$.

Proof: See Section VII-D. ∎

Combining Theorems VI.1 and VI.4 leads to the following corollary:

Corollary VI.5. *Suppose that $L = 1$, $\mathbf{x} \in \mathbb{R}^N_{1/\sqrt{N}}$, N is a prime number and $\mathbf{g}$ is a rectangular window of length*

$$2W - 1 + \frac{1}{128W^4} \geq N.$$

Then, under the measurement model of (II.1), Algorithm 2, initialized by Algorithm 1, with thresholding parameter $B = \frac{1}{\sqrt{N}}$ and step size $0 < \mu \leq 2/\beta$ achieves the following geometrical convergence

$$d^2\left(\mathbf{x}_k, \mathbf{x}\right) \leq \left(1 - \frac{2\mu}{\alpha}\right)^k d^2\left(\mathbf{x}_{k-1}, \mathbf{x}\right),$$

where $\alpha \geq \frac{4N}{W}$ and $\beta \geq 256N^2W^3$.

Proof: See Section VII-E. ∎

We mention that the result of Corollary VI.5 is good merely for long windows. However, in practice we observe that the algorithm works well also for short windows. As we discuss in Section VIII, bridging this theoretical gap is an important direction for future research.

VII. Proofs

A. Proof of Theorem VI.1

The initialization is based on extracting the principle eigenvector of the matrix $\mathbf{X}_0$ defined in Algorithm 1. By assumption, $\mathbf{G}_\ell$ are invertible matrices for $\ell = -(W-1), \ldots, W-1$ for some $W \geq 2$ and hence we can compute (see (II.5))

$$\operatorname{diag}\left(\mathbf{X}_0, \ell\right) = \mathbf{G}_\ell^{-1}\mathbf{y}_\ell = \operatorname{diag}\left(\mathbf{X}, \ell\right).$$

For $\ell = W, \ldots, (N-W)$ we have $\operatorname{diag}\left(\mathbf{X}_0, \ell\right) = 0$. Let us take a look at the matrix $\mathbf{E} := \mathbf{X} - \mathbf{X}_0$. Clearly, $\mathbf{E}$ is not zero at most on $N - 2W + 1$ diagonals. In other words, in each row and column, there are at most $N - 2W + 1$ non-zero values. Let Ω_i be the set of non-zero values of the ith row of $\mathbf{E}$ with cardinality $|\Omega_i| \leq N - 2W + 1$. Using the fact that $\|\mathbf{x}\|_\infty = \sqrt{\frac{B}{N}}$ we can estimate

$$\begin{aligned}
\|\mathbf{E}\|_\infty &:= \max_i \sum_j \left|\mathbf{X}\left[i,j\right] - \mathbf{X}_0\left[i,j\right]\right| \\
&= \max_i \sum_{j \in \Omega_i} \left|\mathbf{X}\left[i,j\right]\right| \\
&= \max_i \sum_{j \in \Omega_i} \left|\mathbf{x}\left[i\right]\mathbf{x}\left[j\right]\right| \\
&\leq \frac{B\left(N - 2W + 1\right)}{N}.
\end{aligned}$$

The same bound holds for $\|\mathbf{E}\|_1 := \max_j \sum_i |\mathbf{E}[i,j]|$ and therefore by Hölder inequality we get

$$\|\mathbf{E}\|_2 \leq \sqrt{\|\mathbf{E}\|_\infty \|\mathbf{E}\|_1} = \frac{B\left(N - 2W + 1\right)}{N}.$$

In order to complete the proof, we still need to show that if $\|\mathbf{X} - \mathbf{X}_0\|_2$ is small, then $d\left(\mathbf{x}, \mathbf{x}_0\right)$ is small as well, where $\mathbf{x}_0$ be the principle eigenvector $\mathbf{X}_0$ with appropriate normalization. To show that, we follow the outline of Section 7.8 in [8]. Observe that as $\mathbf{G}_0$ is invertible by assumption, the norm of $\mathbf{x}$ is known by

$$\|\mathbf{x}\|_2^2 = \sum_{n=0}^{N-1} \left(\operatorname{diag}\left(\mathbf{X}, 0\right)[n] = \sum_{n=0}^{N-1} \left(\mathbf{G}_0^{-1}\mathbf{y}_0\right)[n].\right.$$

Accordingly, we assume hereinafter without loss of generality that $\mathbf{x}$ and $\mathbf{x}_0$ have unit norm. Let λ_0 be the top eigenvalue of $\mathbf{X}_0$, associated with $\mathbf{x}_0$. We observe that

$$\begin{aligned}
\left|\lambda_0 - \left|\mathbf{x}_0^*\mathbf{x}\right|^2\right| &= \left|\mathbf{x}_0^*\mathbf{X}_0\mathbf{x}_0 - \mathbf{x}_0^*\mathbf{x}\mathbf{x}^*\mathbf{x}_0\right| \\
&\leq \left\|\mathbf{X}_0 - \mathbf{x}\mathbf{x}^*\right\|_2.
\end{aligned}$$

Furthermore, as $\|\mathbf{x}\|_2 = 1$ we also have

$$\begin{aligned}
\lambda_0 &\geq \mathbf{x}^*\mathbf{X}_0\mathbf{x} = \mathbf{x}^*\left(\mathbf{X}_0 - \mathbf{x}\mathbf{x}^*\right)\mathbf{x} + 1 \\
&\geq 1 - \left\|\mathbf{X}_0 - \mathbf{x}\mathbf{x}^*\right\|_2.
\end{aligned}$$

Combining the last two inequalities we get

$$\begin{aligned}
\left|\mathbf{x}_0^*\mathbf{x}\right|^2 &\geq 1 - 2\left\|\mathbf{X}_0 - \mathbf{x}\mathbf{x}^*\right\|_2 \\
&\geq 1 - 2B\frac{N - 2W + 1}{N}.
\end{aligned}$$

It then follows

$$d^2\left(\mathbf{x}_0, \mathbf{x}\right) \leq 2\left(1 - \sqrt{1 - 2B\frac{N - 2W + 1}{N}}\right),$$

where the term in the square root is positive by assumption. This completes the proof.

B. Proof of Lemma VI.2

Recall that

$$\begin{aligned}
\nabla f(\mathbf{z}) = \sum_{m=0}^{N/L-1} \sum_{\ell=-(W-1)}^{W-1} &\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z} - \mathbf{Y}\left[m,\ell\right]\right) \\
&\cdot \left(\mathbf{H}_{m,\ell} + \mathbf{H}_{m,\ell}^T\right)\mathbf{z},
\end{aligned}$$

where

$$\mathbf{H}_{m,\ell} := \mathbf{P}_{-\ell}\mathbf{D}_{mL}\mathbf{D}_{mL-\ell},$$

$\mathbf{D}_{mL}$ is a diagonal matrix whose entries are $\{\mathbf{g}\left[mL - n\right]\}_{n=0}^{N-1}$ for fixed m and $\mathbf{P}_\ell$ is a matrix that shifts (circularly) the entries of an arbitrary vector by ℓ entries. We observe that for a rectangular window of length W and $\|\mathbf{z}\|_\infty \leq \frac{1}{\sqrt{N}}$, we have $\|\mathbf{z}\|_2 \leq 1$ and

$$\left\|\mathbf{H}_{m,\ell}\mathbf{z}\right\|_2 \leq \left\|\mathbf{H}_{m,\ell}\right\|_2 \|\mathbf{z}\|_2 \leq 1,$$

and hence

$$\left\|\nabla f(\mathbf{z})\right\|_2 \leq 2\sum_{m=0}^{N/L-1} \sum_{\ell=-(W-1)}^{W-1} \left|\mathbf{Y}\left[m,\ell\right] - \mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}\right|. \tag{VII.1}$$

For convenience, let us denote $d(\mathbf{x},\mathbf{z}) = \frac{\varepsilon}{\sqrt{N}}$ for some $\varepsilon \leq 1$ and therefore $|\mathbf{z}[n]| \geq \frac{1-\varepsilon}{\sqrt{N}}$ for all n. Accordingly, for any (n,k)

$$\frac{(1-\varepsilon)^2}{N} \leq |\mathbf{z}[n]\mathbf{z}[n+k]| \leq \frac{1}{N}.$$

Since $|\mathbf{x}[n]-\mathbf{z}[n]| \leq \frac{1}{\sqrt{N}}$, we have $\text{sign}\left(\mathbf{z}\left[n\right]\mathbf{z}\left[n+k\right]\right) = \text{sign}\left(\mathbf{x}\left[n\right]\mathbf{x}\left[n+k\right]\right)$. Accordingly,

$$\begin{aligned}|\mathbf{x}[n]\mathbf{x}[n+k]-\mathbf{z}[n]\mathbf{z}[n+k]| &\leq \frac{1}{N}\left|1-(1-\varepsilon)^2\right| \\ &\leq \frac{2\varepsilon}{N},\end{aligned}$$

and for all $m,\ell \geq 0$,

$$\begin{aligned}&\left|\mathbf{Y}\left[m,\ell\right]-\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}\right| \\ &\leq \sum_{k=m-(W-1)}^{m-\ell} |\mathbf{x}[n]\mathbf{x}[n+k]-\mathbf{z}[n]\mathbf{z}[n+k]| \\ &\leq \sum_{k=m-(W-1)}^{m-\ell} \frac{2\varepsilon}{N} \leq \frac{2Wd(\mathbf{x},\mathbf{z})}{\sqrt{N}}.\end{aligned} \tag{VII.2}$$

The same bound holds for $\ell < 0$. Combining (VII.1) and (VII.2) we conclude that

$$\begin{aligned}\left\|\nabla f(\mathbf{z})\right\|_2 &\leq 2\sum_{m=0}^{N/L-1}\sum_{\ell=-(W-1)}^{W-1}\frac{2Wd(\mathbf{x},\mathbf{z})}{\sqrt{N}} \\ &= \frac{8}{L}W^2\sqrt{N}d(\mathbf{x},\mathbf{z}).\end{aligned}$$

Remark VII.1. In case of non-vanishing window of length W, one can easily bound the gradient using the same technique, while taking into account $\max_n |\mathbf{g}[n]|$ in the inequalities.

C. Proof of Lemma VI.3

Recall that (see (IV.2))

$$\begin{aligned}&\left\langle \nabla f(\mathbf{z}), \mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right\rangle \\ &= \sum_{m=0}^{N/L-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)\cdot \\ &\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\left(\mathbf{H}_{m,\ell}+\mathbf{H}_{m,\ell}^T\right)\mathbf{z}.\end{aligned}$$

Since $\mathbf{x}^T\mathbf{H}_{m,\ell}^T\mathbf{z} = \mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{x}$ we get for fixed (m,ℓ) and $\phi(\mathbf{z}) = \{0,\pi\}$:

$$\begin{aligned}&\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\left(\mathbf{H}_{m,\ell}+\mathbf{H}_{m,\ell}^T\right)\mathbf{z} \\ &= \left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\mathbf{H}_{m,\ell}\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right) \\ &+ \left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right).\end{aligned}$$

Therefore we get

$$\begin{aligned}&\left\langle \nabla f(\mathbf{z}), \mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right\rangle \\ &= \sum_{m=0}^{N-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)^2 \\ &+ \sum_{m=0}^{N-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)\cdot \\ &\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\mathbf{H}_{m,\ell}\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right).\end{aligned} \tag{VII.3}$$

Clearly, if $\mathbf{z} = \mathbf{x}e^{j\phi(\mathbf{z})}$ then $\left\langle \nabla f(\mathbf{z}), \mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right\rangle = 0$. Otherwise, the first term of (VII.3) is strictly positive. Hence, in order to achieve a lower bound on (VII.3), we first derive an upper bound on the second term and then bound the first term from below.

By assumption $d(\mathbf{x},\mathbf{z}) \leq \frac{1}{\sqrt{N}}$ and denote $\left|\mathbf{x}[n]e^{j\phi(\mathbf{z})}-\mathbf{z}[n]\right| := \frac{\varepsilon_n}{\sqrt{N}}$ for some $\varepsilon_n \leq 1$. We observe that $\sum_n(\frac{\varepsilon_n}{\sqrt{N}})^2 = d^2(\mathbf{x},\mathbf{z})$. For fixed $\ell \geq 0$, we can use the Cauchy-Schwartz inequality to obtain:

$$\begin{aligned}&\sum_{m=0}^{N-1}\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\mathbf{H}_{m,\ell}\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right) \\ &= \sum_{m=0}^{N-1}\sum_{n=m-(W-1)}^{m-\ell}\left(\mathbf{z}\left[n\right]-\mathbf{x}\left[n\right]e^{j\phi(\mathbf{z})}\right)\left(\mathbf{z}\left[n+\ell\right]-\mathbf{x}\left[n+\ell\right]e^{j\phi(\mathbf{z})}\right) \\ &\leq W\left|\sum_{m=0}^{N-1}\left(\mathbf{z}\left[m\right]-\mathbf{x}\left[m\right]e^{j\phi(\mathbf{z})}\right)\left(\mathbf{z}\left[m+\ell\right]-\mathbf{x}\left[m+\ell\right]e^{j\phi(\mathbf{z})}\right)\right| \\ &\leq W\sqrt{\sum_{m=0}^{N-1}\frac{\varepsilon_m^2}{N}}\sqrt{\sum_{m=0}^{N-1}\frac{\varepsilon_{m+\ell}^2}{N}} = Wd^2(\mathbf{x},\mathbf{z}).\end{aligned}$$

The same bound holds for $\ell < 0$. Combining the last result with (VII.2) we get for the second term in (VII.3) that

$$\begin{aligned}&\sum_{m=0}^{N-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)\cdot \\ &\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right)^T\mathbf{H}_{m,\ell}\left(\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right) \leq \frac{4}{\sqrt{N}}W^3d^3(\mathbf{x},\mathbf{z}).\end{aligned} \tag{VII.4}$$

Next, we aim to bound the first term of (VII.3) from below as follows:

$$\begin{aligned}&\sum_{m=0}^{N-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)^2 \\ &\geq \sum_{m=0}^{N-1}\left(\mathbf{z}^T\mathbf{H}_{m,0}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,0}\mathbf{x}\right)^2 \\ &= \sum_{m=0}^{N-1}\left(\sum_{n=m-(W-1)}^{m}\mathbf{z}^2\left[n\right]-\mathbf{x}^2\left[n\right]\right)^2 \\ &\geq \sum_{m=0}^{N-1}\sum_{n=m-(W-1)}^{m}\left(\mathbf{z}^2\left[n\right]-\mathbf{x}^2\left[n\right]\right)^2,\end{aligned} \tag{VII.5}$$

where the last inequality is true since $\mathbf{x}^2[n] \geq \mathbf{z}^2[n]$ and for any positive (or negative) sequence $\{a_i\}$ we have $\left(\sum_i a_i\right)^2 \geq \sum_i a_i^2$. Furthermore, since $|\mathbf{z}[n]| = \frac{1-\varepsilon_n}{\sqrt{N}}$ we have

$$\begin{aligned}
&\sum_{m=0}^{N-1}\sum_{n=m-(W-1)}^{m}\left(\mathbf{z}^2[n]-\mathbf{x}^2[n]\right)^2 \\
&= \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=m-(W-1)}^{m}\left(1-(1-\varepsilon_n)^2\right)^2 \\
&= \frac{W}{N^2}\sum_{n=0}^{N-1}\left(2\varepsilon_n-\varepsilon_n^2\right)^2.
\end{aligned}$$

Therefore, since $\varepsilon_n \leq 1$ for all n we conclude that

$$\begin{aligned}
&\sum_{m=0}^{N-1}\sum_{\ell=-(W-1)}^{W-1}\left(\mathbf{z}^T\mathbf{H}_{m,\ell}\mathbf{z}-\mathbf{x}^T\mathbf{H}_{m,\ell}\mathbf{x}\right)^2 \\
&\geq \frac{W}{N^2}\sum_{n=0}^{N-1}\varepsilon_n^2 = \frac{Wd^2(\mathbf{x},\mathbf{z})}{N}.
\end{aligned} \tag{VII.6}$$

Remark VII.2. Observe that the analysis for non-vanishing windows of length W requires only a small modification. In this case, one should use the maximal and the minimal values of the window in the above inequalities. For instance, one would need to take $\mathrm{g}_{\min} := \min_{n=0,\ldots,W-1}|\mathbf{g}[n]|$ into account in (VII.5).

Plugging (VII.4) and (VII.6) into (VII.3) yields

$$\begin{aligned}
\langle\nabla f(\mathbf{z}),\mathbf{z}-\mathbf{x}\rangle &\geq \frac{Wd^2(\mathbf{x},\mathbf{z})}{N}\left(1-4\sqrt{N}W^2d(\mathbf{x},\mathbf{z})\right) \\
&\geq \frac{Wd^2(\mathbf{x},\mathbf{z})}{2N},
\end{aligned}$$

where the last inequality holds for $d(\mathbf{x},\mathbf{z}) \leq \frac{1}{8\sqrt{N}W^2}$. This concludes the proof.

D. Proof of Theorem VI.4

For fixed $\mathbf{x}$, let $\mathcal{E}$ be the set of vectors in $\mathbb{R}^N$ satisfying $\|\mathbf{z}\|_\infty \leq \frac{1}{\sqrt{N}}$ and $d(\mathbf{x},\mathbf{z}) \leq \frac{1}{8\sqrt{N}W^2}$. We first need the following definition:

Definition VII.3. We say that a function f satisfies the *regularity condition* in $\mathcal{E}$ if for all vectors $\mathbf{z} \in \mathcal{E}$ we have

$$\left\langle\nabla f(\mathbf{z}),\mathbf{z}-\mathbf{x}e^{j\phi(\mathbf{z})}\right\rangle \geq \frac{1}{\alpha}d^2(\mathbf{z},\mathbf{x})+\frac{1}{\beta}\|\nabla f(\mathbf{z})\|_2^2,$$

for some positive constants α,β.

The following lemma states that if the regularity condition is met, then the gradient step converges to a global minimum at a geometric rate.

Lemma VII.4. *Assume that f satisfies the regularity condition for all $\mathbf{z} \in \mathcal{E}$. Consider the following update rule*

$$\mathbf{z}_k = \mathbf{z}_{k-1} - \mu\nabla f(\mathbf{z}_{k-1}),$$

for $0 < \mu \leq 2/\beta$. Then,

$$d^2(\mathbf{z}_k,\mathbf{x}) \leq \left(1-\frac{2\mu}{\alpha}\right)^k d^2(\mathbf{z}_{k-1},\mathbf{x}).$$

Proof: See Section 7.4 in [8]. ∎

We notice that the thresholding stage of Algorithm 2 cannot increase the error as the signal is assumed to be bounded. Hence, we can directly leverage lemmas VI.2 and VI.3 and see that Definition VII.3 holds in our case with the constants $\alpha \geq \frac{4N}{W}$ and $\beta \geq 256N^2W^3$.

E. Proof of Corollary VI.5

As N is a prime number, $\mathbf{g}$ is an admissible window of length W (see Claim III.3). According to Theorem VI.4, we merely need to show that the initialization point is within the basin of attraction, namely, $d(\mathbf{x},\mathbf{x}_0) \leq \frac{1}{8\sqrt{N}W^2}$. From Lemma VI.1, we know that the initialization point obeys

$$d^2(\mathbf{x}_0,\mathbf{x}) \leq 2\left(1-\sqrt{1-2\frac{N-2W+1}{N}}\right).$$

Using the fact that $a \leq \sqrt{a}$ for all $0 \leq a \leq 1$ and some standard algebraic calculations, we conclude that the initialization of Algorithm 1 is within the basin of attraction as long as

$$2W-1+\frac{1}{128W^4} \geq N.$$

This completes the proof.

VIII. Discussion

This paper aims at suggesting a practical, efficient phase retrieval algorithm with theoretical guarantees. The proposed algorithm minimizes a non-convex loss function. Similar approaches were taken recently for phase retrieval problems, however, they are mainly focused on random setups and their analysis is greatly based on statistical considerations. Here, we pursue a different approach. The algorithm begins by taking the DFT of the measurements (II.1). This step simplifies the structure of the data and implies almost directly uniqueness results. For sufficiently long windows, we show that a LS algorithm recovers the signal efficiently. For general settings, we apply a gradient descent algorithm. The algorithm is initialized by the principle eigenvector of a designed matrix, constructed as the solution of a LS problem. For $L = 1$, we estimate the distance between the initialization and the global minimum (see Theorem VI.1). The case of $L > 1$ raises some interesting questions. As a heuristic, we have suggested to smoothly interpolate the missing entries. This practice works quite well since the window acts as an averaging operator. Clearly, the interpolation method depends on the window shape. A main challenge for future research is analyzing the setting of $L > 1$. Additionally, a slight modification of our initialization algorithm recovers exactly signals with unit module entries (see Proposition III.5). This result implies a potential applicability to angular synchronization [40], [1], [4].

The analysis of the non-convex algorithm relies on the geometry of the loss function (IV.1). We prove in Theorem VI.4 that for signals with unit module entries, there exists a basin of attraction. We show numerically that the actual basin of attraction is larger than the theoretical bound and exists for a broader family of signals. The gap between the actual

basin of attraction and the theoretical result is the bottleneck that prevents a full theoretical understanding of the algorithm. Specifically, improving Lemma VI.3 will lead directly to tighter estimation of the basin of attraction. Bridging that gap is an additional major goal of a future work.

ACKNOWLEDGMENT

We would like to thank Mahdi Soltanolkotabi for his remarks on an initial draft of this paper. TB would also like to thank Kumar Vijay Mishra for his kind help with Figure IV.1.

REFERENCES

[1] A.S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *arXiv preprint arXiv:1411.3272*, 2014.

[2] B. Baykal. Blind channel estimation via combining autocorrelation and blind phase estimation. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 51(6):1125–1131, 2004.

[3] I. Bojarovska and A. Flinth. Phase retrieval from gabor measurements. *Journal of Fourier Analysis and Applications*, pages 1–26, 2015.

[4] N. Boumal. Nonconvex phase synchronization. *arXiv preprint arXiv:1601.06114*, 2016.

[5] E.J. Candes, Y.C. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Review*, 57(2):225–251, 2015.

[6] E.J. Candès and X. Li. Solving quadratic equations via phaselift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, 14(5):1017–1026, 2014.

[7] E.J. Candes, X. Li, and M. Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, 2015.

[8] E.J. Candes, X. Li, and M. Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. *Information Theory, IEEE Transactions on*, 61(4):1985–2007, 2015.

[9] E.J. Candes, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.

[10] Y. Chen and E. Candes. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *To appear in Communications on Pure and Applied Mathematics*, 2015.

[11] B.A. Dumitrescu. *Positive trigonometric polynomials and signal processing applications*. Springer Science & Business Media, 2007.

[12] Y.C. Eldar and S. Mendelson. Phase retrieval: Stability and recovery guarantees. *Applied and Computational Harmonic Analysis*, 36(3):473–494, 2014.

[13] Y.C. Eldar, P. Sidorenko, D.G. Mixon, S. Barel, and O. Cohen. Sparse phase retrieval from short-time fourier measurements. *Signal Processing Letters, IEEE*, 22(5):638–642, 2015.

[14] C. Fienup and J. Dainty. Phase retrieval and image reconstruction for astronomy. *Image Recovery: Theory and Application*, pages 231–275, 1987.

[15] J.R. Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.

[16] R.W. Gerchberg. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237, 1972.

[17] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[18] D.W. Griffin and J.S Lim. Signal estimation from modified short-time fourier transform. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 32(2):236–243, 1984.

[19] D. Gross, F. Krahmer, and R. Kueng. Improved recovery guarantees for phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 2015.

[20] R.W. Harrison. Phase problem in crystallography. *JOSA A*, 10(5):1046–1055, 1993.

[21] K. Huang, Y.C. Eldar, and N.D. Sidiropoulos. Phase retrieval from 1D Fourier measurements: Convexity, uniqueness, and algorithms. *arXiv preprint arXiv:1603.05215*, 2016.

[22] K. Jaganathan, Y.C. Eldar, and B. Hassibi. Phase retrieval: An overview of recent developments. *arXiv preprint arXiv:1510.07713*, 2015.

[23] K. Jaganathan, Y.C. Eldar, and B. Hassibi. STFT phase retrieval: Uniqueness guarantees and recovery algorithms. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):770–781, 2016.

[24] K. Jaganathan, S. Oymak, and B. Hassibi. Sparse phase retrieval: Uniqueness guarantees and recovery algorithms. *arXiv preprint arXiv:1311.2745*, 2013.

[25] B. Juang and L. Rabiner. Fundamentals of speech recognition. *Signal Processing Series. Prentice Hall, Englewood Cliffs, NJ*, 1993.

[26] J.D Lee, M. Simchowitz, M.I. Jordan, and B. Recht. Gradient descent converges to minimizers. *arXiv preprint arXiv:1602.04915*, 2016.

[27] A.M. Maiden, M.J. Humphry, F. Zhang, and J.M. Rodenburg. Super-resolution imaging via ptychography. *JOSA A*, 28(4):604–612, 2011.

[28] S. Marchesini, Y. Tu, and H. Wu. Alternating projection, ptychographic imaging and phase synchronization. *Applied and Computational Harmonic Analysis*, 2015.

[29] R.P. Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.

[30] S.H. Nawab, T.F. Quatieri, and J.S. Lim. Signal reconstruction from short-time fourier transform magnitude. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 31(4):986–998, 1983.

[31] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. *Signal Processing, IEEE Transactions on*, 63(18):4814–4826, 2015.

[32] A.V. Oppenheim and R.W. Schafer. *Discrete-time signal processing*. Pearson Higher Education, 2010.

[33] J. Ranieri, A. Chebira, Y.M. Lu, and M. Vetterli. Phase retrieval for sparse signals: Uniqueness conditions. *arXiv preprint arXiv:1308.3058*, 2013.

[34] J.M. Rodenburg. Ptychography and related diffractive imaging methods. *Advances in Imaging and Electron Physics*, 150(07):87–184, 2008.

[35] J.M. Rodenburg and R. Bates. The theory of super-resolution electron microscopy via wigner-distribution deconvolution. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 339(1655):521–553, 1992.

[36] C. Rusu and J. Astola. Extending a sequence into a minimum-phase sequence. In *In: Bregovic, R. & Gotchev, A.(eds.). Proceedings of the 2007 International TICSP Workshop on Spectral Methods and Multirate Signal Processing, SMMSP 2007, Moscow, Russia, 1-2 September 2007*, 2007.

[37] Y. Shechtman, A. Beck, and Y.C. Eldar. Gespar: Efficient phase retrieval of sparse signals. *Signal Processing, IEEE Transactions on*, 62(4):928–938, 2014.

[38] Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, and M. Segev. Phase retrieval with application to optical imaging: a contemporary overview. *Signal Processing Magazine, IEEE*, 32(3):87–109, 2015.

[39] Y. Shechtman, Y.C. Eldar, A. Szameit, and M. Segev. Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing. *Optics express*, 19(16):14807–14822, 2011.

[40] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1):20–36, 2011.

[41] J. Sun, Q. Qu, and J. Wright. A geometric analysis of phase retrieval. *arXiv preprint arXiv:1602.06664*, 2016.

[42] R. Trebino. Frequency resolved optical gating: The measurement of ultrashort optical pulses. *Kluwer Academic Publishers, Boston*, 5:70–73, 2002.

[43] I. Waldspurger, A. dAspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *Mathematical Programming*, 149(1-2):47–81, 2015.

[44] A. Walther. The question of phase retrieval in optics. *Journal of Modern Optics*, 10(1):41–49, 1963.

[45] G. Wang, G.B. Giannakis, and Y.C. Eldar. Solving random systems of quadratic equations via truncated generalized gradient flow. *arXiv preprint arXiv:1605.08285*, 2016.

[46] Y. Wang and Z. Xu. Phase retrieval for sparse signals. *Applied and Computational Harmonic Analysis*, 37(3):531–544, 2014.

[47] C.D. White, S. Sanghavi, and R. Ward. The local convexity of solving systems of quadratic equations. *arXiv preprint arXiv:1506.07868*, 2015.

[48] C. Yang, J. Qian, A. Schirotzek, F. Maia, and S. Marchesini. Iterative algorithms for ptychographic phase retrieval. *arXiv preprint arXiv:1105.5628*, 2011.

APPENDIX

A. Proof of Proposition III.4

By assumption, the DFT of $\mathbf{g}\circ(\mathbf{P}_{-\ell}\mathbf{g})$ is non-vanishing for $\ell=0,1$. Consequentially, the matrices $\mathbf{G}_\ell$, $\ell=0,1$ as given in (II.5) are invertible and we can compute

$$\mathbf{x}_\ell=\mathbf{G}_\ell^{-1}\mathbf{y}_\ell,\quad \ell=0,1,$$

where $\mathbf{X}=\mathbf{x}\mathbf{x}^*$, $\mathbf{x}_\ell=\mathrm{diag}\left(\mathbf{X},\ell\right)$ and $\mathbf{y}_\ell:=\{\mathbf{Y}\left[m,\ell\right]\}_{m=0}^{N-1}$. Because of the fundamental ambiguity of phase retrieval, the first entry can be set arbitrarily to be $\sqrt{\mathbf{x}_0\left[0\right]}=\left|\mathbf{x}\left[0\right]\right|$. Then, as we assume non-vanishing signals, the rest of the entries can be determined recursively for $n=1\ldots,N-1$ by

$$\frac{\mathbf{x}_1\left[n-1\right]}{\mathbf{x}\left[n-1\right]}=\frac{\mathbf{x}\left[n-1\right]\mathbf{x}\left[n\right]}{\mathbf{x}\left[n-1\right]}=\mathbf{x}\left[n\right].$$

This completes the proof.

B. Proof of Proposition III.5

By assumption, $\mathbf{G}_\ell$ is an invertible matrix for $|\ell|\leq W-1$ for some $W\geq 2$ (see (II.5)). Hence, we can compute $\mathrm{diag}(\mathbf{X},\ell)=\mathbf{G}_\ell^{-1}\mathbf{y}_\ell$ for $\ell=0,M$ for any $1\leq M\leq W-1$. The proof is a direct corollary of the following lemma:

Lemma. *Let $L=1$. Suppose that $\mathbf{x}\in\mathbb{C}^N_{1/\sqrt{N}}$ and let $\mathbf{X}=\mathbf{x}\mathbf{x}^*$. Fix $M\in[1,\ldots,N-1]$ and let $\mathbf{X}_0$ be a matrix obeying*

$$\textit{diag}\left(\mathbf{X}_0,\ell\right)=\begin{cases}\textit{diag}\left(\mathbf{X},\ell\right), & \ell=0,M,\\ 0, & \textit{otherwise}.\end{cases}$$

Then, $\mathbf{x}$ is a principle eigenvectors of $\mathbf{X}_0$ (up to global phase).

Proof: Based on the special structure of $\mathbf{X_0}$, the following calculation shows that $\mathbf{x}$ is an eigenvector of $\mathbf{X}_0$ with $\frac{2}{N}$ as the associated eigenvalue:

$$\begin{aligned}\left(\mathbf{X}_0\mathbf{x}\right)[i] &= \sum_{j=1}^{N}\mathbf{X}_0[i,j]\mathbf{x}[j]\\ &= \mathbf{X}_0[i,i]\mathbf{x}[i]+\mathbf{X}_0[i,i+M]\mathbf{x}[i+M]\\ &= \mathbf{x}[i]\left|\mathbf{x}[i]\right|^2+\mathbf{x}[i]\left|\mathbf{x}[i+M]\right|^2\\ &= \frac{2}{N}\mathbf{x}[i].\end{aligned}$$

We still need to show that $\mathbf{x}$ is a principle eigenvector of $\mathbf{X}_0$. Since each column and row of $\mathbf{X}_0$ is composed of two non-zero values, it is evident that

$$\left\|\mathbf{X}_0\right\|_\infty:=\max_i\sum_j\left|\mathbf{X}_0\left[i,j\right]\right|=\frac{2}{N}.$$

In the same manner

$$\left\|\mathbf{X}_0\right\|_1:=\max_j\sum_i\left|\mathbf{X}_0\left[i,j\right]\right|=\frac{2}{N}.$$

Hence by Hölder inequality we get

$$\left\|\mathbf{X}_0\right\|_2\leq\sqrt{\left\|\mathbf{X}_0\right\|_1\left\|\mathbf{X}_0\right\|_\infty}=\frac{2}{N}.$$

This concludes the proof. ∎

C. Proof of Proposition III.6

As the matrices $\mathbf{G}_\ell$ are invertible by assumption for all $\ell=-(W-1),\ldots,(W-1)$, we can compute

$$\mathrm{diag}\left(\mathbf{X}_0,\ell\right)=\mathbf{G}_\ell^{-1}\mathbf{y}_\ell=\mathrm{diag}\left(\mathbf{X},\ell\right).$$

The assumption $W\geq\left\lceil\frac{N+1}{2}\right\rceil$ implies that $\mathbf{X}_0=\mathbf{X}$. Specifically, observe that it is sufficient to consider only $W=\left\lceil\frac{N+1}{2}\right\rceil$ since for any $|\ell_1|>\left\lceil\frac{N+1}{2}\right\rceil$, the window $\mathbf{g}\circ(\mathbf{P}_{-\ell_1})$ is equal to another window $\mathbf{g}\circ(\mathbf{P}_{-\ell_2})$ for some $|\ell_2|\leq\left\lceil\frac{N+1}{2}\right\rceil$.

Let $\tilde{\mathbf{x}}:=\mathbf{x}/\|\mathbf{x}\|_2$. Then, $\tilde{\mathbf{x}}$ is the principle eigenvector of $\mathbf{X}$ and the normalization stage of Algorithm 1 gives

$$\sqrt{\sum_{n=0}^{N-1}\left(\mathbf{G}_0^{-1}\mathbf{y}_0\right)[n]}=\|\mathbf{x}\|_2.$$

This completes the proof.

D. Proof of Lemma IV.1

We identify the convolution $\mathbf{g}*\mathbf{x}$ by the matrix-vector product $\mathbf{G}\mathbf{x}$, where $\mathbf{G}\in\mathbb{R}^{N\times N}$ is a circulant matrix whose first column is given by $\tilde{\mathbf{g}}:=\{\mathbf{g}[(-n)\mathrm{mod}N]\}_{n=0}^{N-1}$. For $L=1$, we can then write

$$\mathbf{y}=\mathbf{G}\mathbf{x}=\mathbf{F}^*\mathbf{\Sigma}\mathbf{F}\mathbf{x},$$

where $\mathbf{F}$ is a DFT matrix and $\mathbf{\Sigma}$ is a diagonal matrix whose entries are the DFT of $\tilde{\mathbf{g}}$. By assumption, the first N/L entries of $\mathbf{\Sigma}$ are ones and the rest are zeros. Hence, we may write

$$\mathbf{y}=\mathbf{F}_p^*\mathbf{F}_p\mathbf{x},\tag{D.1}$$

where $\mathbf{F}_p\in\mathbb{C}^{N/L\times N}$ consists of the first N/L rows of $\mathbf{F}$.

Let $\mathbf{G}_L\in\mathbb{R}^{\frac{N}{L}\times N}$ be a matrix consists of the $\{jL:j=0,\ldots,N/L-1\}$ rows of $\mathbf{G}$. For $L>1$, we get the downsampled system of equations

$$\mathbf{y}_L=\mathbf{G}_L\mathbf{x}=\mathbf{F}_L^*\mathbf{\Sigma}\mathbf{F}\mathbf{x},$$

where $\mathbf{F}_L$ consists of the $\{jL:j=0,\ldots,N/L-1\}$ columns of $\mathbf{F}$ (notice the difference between $\mathbf{F}_L$ and $\mathbf{F}_p$). We aim at showing that expanding and interpolating $\mathbf{y}_L$ as explained in Lemma IV.1 results in $\mathbf{y}$. Direct computation shows that the expansion stage as described in (IV.3) is equivalent to multiplying both sides by $\mathbf{F}^*\mathbf{F}_L$:

$$\tilde{\mathbf{y}}_L=\mathbf{F}^*\mathbf{F}_L\mathbf{y}_L=\mathbf{F}^*\left(\mathbf{F}_L\mathbf{F}_L^*\right)\mathbf{\Sigma}\mathbf{F}\mathbf{x}.$$

Let us denote $\mathbf{T}:=\mathbf{F}_L\mathbf{F}_L^*$, which is a Toeplitz matrix with L on the $\frac{jN}{L}$ diagonals for $j=0,\ldots,N/L-1$ and zero otherwise. Because of the structure of $\mathbf{\Sigma}$ we can then write

$$\tilde{\mathbf{y}}_L=\mathbf{F}^*\mathbf{T}_p\mathbf{F}_p\mathbf{x},$$

where $\mathbf{T}_p\in\mathbb{R}^{N\times N/L}$ consists of the first N/L columns of $\mathbf{T}$. Direct calculation shows that $\mathbf{F}_p\mathbf{F}^*\mathbf{T}_p=\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Therefore we conclude that

$$\left(\mathbf{F}_p^*\mathbf{F}_p\right)\tilde{\mathbf{y}}_L=\mathbf{F}_p^*\mathbf{F}_p\mathbf{x}.\tag{D.2}$$

Comparing (D.2) with (D.1) completes the proof.