

Phase Retrieval for Radar Waveform Design

Samuel Pinilla *Student Member, IEEE*, Kumar Vijay Mishra *Senior Member, IEEE*, Brian M Sadler *Fellow, IEEE*, Henry Arguello *Senior Member, IEEE*

Abstract—The ability of radar to discriminate in both range and velocity is completely characterized by the *ambiguity function* (AF) of its transmit waveform. Mathematically, it is obtained by correlating the waveform with its Doppler-shifted and delayed replicas. We consider the inverse problem of designing a radar transmit waveform that satisfies the specified ambiguity function magnitude. This process can be viewed as a signal reconstruction with some variation of phase retrieval methods. We provide a trust-region algorithm that minimizes a smoothed non-convex least-squares objective function to iteratively recover the underlying signal-of-interest for either time- or band-limited support. The method first approximates the signal using an iterative spectral algorithm and then refines the attained initialization based upon a sequence of gradient iterations. This holds uniquely as long as at least $(3S)3B$ measurements where (S) B is the (timewidth) bandwidth are available. Numerical experiments demonstrate that the proposed algorithm estimates time/band-limited signals from its radar ambiguity function for both complete and incomplete radar ambiguity function. The radar function is incomplete when only few shifts or Fourier frequencies are considered. Numerical results show that the proposed algorithm estimates the signal with mean-square-error of 1×10^{-6} , and 9×10^{-2} for complete-noiseless and incomplete-noisy cases, respectively.

Index Terms—Radar, phase retrieval, ambiguity function, Band-limited signals.

I. INTRODUCTION

RADAR is a detection system that uses radio waves to determine the position or velocity of objects [1]. This kind of signals arises in many fields in science and engineering such as synthetic aperture radar (SAR) and inverse SAR [2], hidden as in through-wall imaging and foliage penetration radar [3], complex as in ground penetration radar [4], and seemingly similar as in gesture radar [5]. As mentioned, the primary functions of a radar are to find targets and to estimate parameters that describe those targets such as *range* and *Doppler*. In fact, the radar's ability to discriminate in both range and velocity is completely characterized by the *ambiguity function* (AF) of its transmit waveform; it is obtained by correlating the waveform with its Doppler-shifted and delayed replicas.

Historically, it is known that the ambiguity function was first introduced by Ville [6], but the pioneer works [7], [8] of Woodward it is highly recognized due to the steps forward in characterizing the performance to identify the target parameters of range and Doppler based on the transmit waveform $x(t)$. This characterization was possible by analyzing that a desire waveform is able to distinguish between radar returns with different target parameters. In fact, Woodward performed this analysis by defining a mean-squared error metric between a known waveform $x(t)$ and a frequency-shifted and time-delayed version as

$$\Omega(\tau, f) = \int |x(t) - x(t - \tau)e^{-j2\pi ft}|^2 dt,$$

S. P. and H. A. are with Universidad Industrial de Santander, Bucaramanga, Santander 680002 Colombia, e-mail: samuel.pinilla@correo.uis.edu.co, henarfu@uis.edu.co.

K. V. M. and B. S. are with the United States Army Research Laboratory, Adelphi, MD 20783 USA, e-mail: kumarvijay-mishra@uiowa.edu, brian.m.sadler6.civ@mail.mil.

K. V. M. acknowledges support from the National Academies of Sciences, Engineering, and Medicine via Army Research Laboratory Harry Diamond Distinguished Postdoctoral Fellowship.

where Δt and Δf correspond to the sampling periods of time and Doppler domains. From this error metric Woodward concluded that the only term that depends on the parameters is the inner product between the original waveform and the time-delayed/frequency-shifted version where is magnitude squared is given by

$$A(\tau, f) = \left| \int x(t)^* x(t - \tau)e^{-j2\pi ft} dt \right|^2, \quad (1)$$

term which is called the ambiguity function. We recall that this definition of ambiguity function corresponds to the case of narrow radar signals [9, Chapter 3] which is the object of study in this work. In fact this definition has been changed to handle larger bandwidth signals, long duration signals, and targets with high velocity [9]–[11].

From the mathematical expression in (1) we have that the radar ambiguity function for narrow signals can be seen as the absolute value of the Fourier transform of the matched filter between the signal and its delayed replica. From this interpretation we have $A(\tau, f)$ concentrates the entire energy of the signal into an output peak at a predetermined additional delay. It is therefore optimal for causing the output to cross the threshold and identify a detected reflection at the corresponding delay in the presence of the receiver noise [1]. In fact, one requirement on a radar waveform is that it must be possible to search a large area of possible target locations (in both range and Doppler) with minimum losses, and a conflicting requirement is that it must be possible to resolve closely spaced targets and measure their positions with specified accuracy.

A. Waveform Design and Classical Radar Signals

The importance of the radar ambiguity function for narrow signals lies in the fact that it shows the distortion of a returned pulse because to the receiver matched filter due to the time-shift of the return from a moving target [7]. From the literature it is well known that an ambiguity function equal to zero except in one point is ideal for detection tasks [9, Chapter 3]. However, such a signal would not allow a radar to detect the target, because the probability of that target lying within the response region would be near zero [12]. Therefore, in the absence of an ideal ambiguity function, we need to choose signals which have ambiguity functions well suited to the detection task in radar. This is called *waveform design*, fact that motivates this work to study the ambiguity functions of various types of signals. For instance, classical basic types of waveforms are single-frequency pulse, linear frequency modulated signal, and stepped frequency pulse train [9] which are narrow enough pulses to detect a target in radar.

It is worth mentioning that the ambiguity function has the following important properties [9, Chapter 3]

- 1) The maximum value for the ambiguity function occurs at $(\Delta\tau, \Delta f) = (0, 0)$ and is equal to M

$$\max\{A(\tau, f)\} = A(0, 0) = M \\ A(\tau, f) \leq A(0, 0) \quad (2)$$

- 2) The total volume under the ambiguity function is the constant

$$\iint |A(\tau, f)|^2 dt df. \quad (3)$$

3) The ambiguity function is symmetric

$$A(\tau, f) = A(-\tau, -f) \quad (4)$$

4) If $x(t)$ has the ambiguity function $A(\tau, f)$, then

$$x(t) \exp(j\pi kt^2) \text{ has } A(\tau, f + k\tau) \quad (5)$$

The first property in (2) establishes that for normalized signals ($M = 1$), the maximum value of $A(\tau, f)$ is one at the origin. For the second property in (3) we have that the volume of $A(\tau, f)$ is a constant equal to one. The implication here is that if $A(\tau, f)$ is squeezed to a narrow peak near to the origin, then that peak cannot exceed the value of one. The third property in (4) indicates that the ambiguity function is symmetrical with respect to the origin. Finally, the fourth property in (5) says that multiplying the envelope of any signal by a quadratic phase (linear frequency) will shear the shape of the ambiguity function. Proofs for these rules are given in [1].

Some of these properties have been exploited for waveform design by proposing an inverse optimization problem to find a pulse that best fit any of these properties. For instance, boundedness, where its maximum value occurs at the origin, has been used in [13] to show that when the ambiguity function is bounded by a *Hermite function* (a polynomial times the standard normal distribution with variance $\sigma^2 = 1/(2\pi)$), then the unknown signal is also a Hermite function where the polynomial can be found from its ambiguity function by comparing coefficients. Another example is the volume property which has been successfully employed in [14] to solve a non-convex optimization problem to adaptively construct a Hermite function that best approximates the volume of a desired ambiguity function. More recent works such as [15], [16] have proposed to numerically design the signal to improve the delay resolution of the ambiguity function in order to better determine the position of objects.

Other research directions on radar for waveform design such as [17], [18], have mathematically studied the formulation of the ambiguity function in a different time-frequency representation such as the fractional Fourier transform. Roughly speaking, this analogous formulation is possible because the Fourier magnitude of the product of the unknown signal with a conjugate time-shifted version of itself, for several different shifts is equivalent to a frequency rotation using the fractional Fourier transform [17]. Mathematically this equivalent formulation is given as

$$A(-\zeta \sin(\alpha), \zeta \cos(\alpha)) = \left| \left| \int x(t) e^{-i\pi t^2 \cot(\alpha) - 2i\pi t \eta / \sin(\alpha)} dt \right|^2 \times e^{-j\zeta \eta} d\eta \right|^2 \quad (6)$$

Additionally, [17], [18] also have shown that a Hermite function or a rectangular pulse trains can be uniquely identified from its fractional radar ambiguity function. In fact any progress on extending the fractional radar formulation as in (6) to other functions than those two previously mentioned is an open problem of high interest.

B. Radar Phase Retrieval Problem

From the definition of the ambiguity function for narrow radar signals in (1) we have that $A(\tau, f)$ is a phaseless mapping. In fact, additional properties of $A(\tau, f)$ as in (2), (3), (4), and (5) have been identified in [13], [17], showing that the following transformed versions of $x(t)$ lead to the same $A(\tau, f)$

- 1) the rotated signal $e^{i\phi} x(t)$ for some $\phi \in \mathbb{R}$.
- 2) the translate signal $x(t - a)$ for some $a \in \mathbb{R}$.
- 3) the reflected signal $x(-t)$.
- 4) the scaled signal $e^{ibt} x(t)$ for some $b \in \mathbb{R}$.

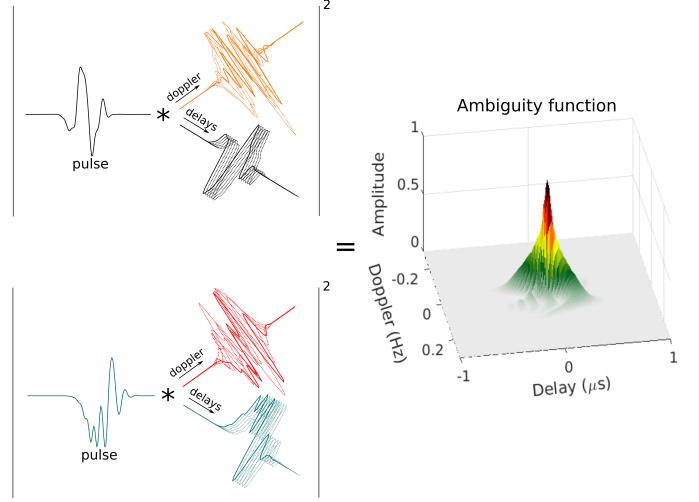


Fig. 1. Sketch how multiple signals maps to the same ambiguity function in radar. The left side correspond to the frequency and time processing performed over the pulse, where illustrations of shifted versions is presented. Hence, the goal is to recover a signal that fits the ambiguity function up to trivial transformations.

These transformations are called in the literature *trivial ambiguities*, and Fig. 1 illustrates how they appear. Considering this characterization over $A(\tau, f)$ the question we ask is thus what kind of waveforms have only trivial partners. In fact this strategy to design waveforms for radar is equivalent to a phase retrieval problem because $A(\tau, f)$ is a phaseless function [17].

The mathematical discrete form of (1) is provided in Table I. This formulation is also compared with different phaseless functions from other problems such as short-time Fourier transform (STFT) [22], frequency-resolve optical gating (FROG) [25], and fractional Fourier transform (FrFT) [18]. It can be concluded from Table I that the radar phase retrieval problem is the most challenging because it posses more trivial ambiguities compared with the rest of the formulations. Additionally, in Table I it is mentioned mathematical conditions in which uniqueness (up to trivial ambiguities) for the different formulations can be guaranteed. Also, reconstruction algorithms, if any exists for the different problems, are provided, where it can be concluded that any method for radar with a polynomial computational complexity has been developed in the literature.

We recall that phase retrieval constitutes an instance of nonconvex programming, that is generally known to be NP-hard [26]. For instance for real-valued signals, this problem can be understood as a combinatorial optimization since one seeks a series of signs over the entries of the target signal that obey the given ambiguity function. In the complex case, it becomes even more complicated, where instead of a set of signs, one must determine a collection of unimodular complex scalars that obeys the given ambiguity function. In particular, the case of radar, as established in Table I, belongs to the class of 1D phase retrieval problems [27]. This means that uniqueness cannot be ensure unless some conditions over the signal are assumed. For instance sparsity [28], [29], non-vanishing signals [22], [30], or band-limited pulses [24]. Additionally, in [27] was shown that, with the exception of a set of signals of measure zero, a real 2D-dimensional signal is uniquely specified by the magnitude of its continuous Fourier transform, up to the trivial ambiguities.

Nowdays, many advances in the phase retrieval problem has been done in terms of developing newer algorithms to recover a target

TABLE I

UNIQUENESS AND ALGORITHMS FOR 1D PHASE RETRIEVAL PROBLEMS RELATED TO RADAR. WE COMPARE THE MATHEMATICAL PHASELESS MODELS FOR DIFFERENT PHASE RETRIEVAL PROBLEMS WITH SOME ANALYTICAL RESULTS DESCRIBING UNIQUENESS UNDER MILD CONDITIONS. RECOVERY ALGORITHMS ARE ALSO MENTIONED IF ANY EXIST. ALSO, THE TRIVIAL AMBIGUITIES FOR THE DIFFERENT PROBLEMS ARE PROVIDED.

Problem	Mathematical Model	Uniqueness/Algorithms
STFT	$\mathbf{A}[p, k] := \left \sum_{n=0}^{N-1} \mathbf{x}[n] \mathbf{g}[pL - n] e^{-\frac{2\pi i kn}{N}} \right ^2, L < N$	Uniqueness (up to a global) phase for almost all signals for some L 's and \mathbf{g}, \mathbf{x} non-vanishing [19]; Uniqueness if the first L samples of \mathbf{x} are known a priori for some L 's and \mathbf{g} is non-vanishing [20]; Uniqueness (up to a global phase) for some L 's, N 's and mild conditions on \mathbf{g} [21]. Non-convex algorithm in [22], employing an initialization followed by a gradient descent update rule.
FROG	$\mathbf{A}[p, k] := \left \sum_{n=0}^{N-1} \mathbf{x}[n] \mathbf{x}[pL + n] e^{-\frac{2\pi i kn}{N}} \right ^2, L < N$	Uniqueness (up to global phase, translated and reflected signal) for some L 's and band-limited pulses [23], [24]. Non-convex algorithm in [25], employing an initialization followed by a gradient descent update rule.
FrFT	$\mathbf{A}[\alpha, k] := \left \sum_{n=0}^{N-1} \left \sum_{t=0}^{N-1} \mathbf{x}[t] e^{-\frac{-2i\pi tn - i\pi t^2 \cos(\alpha)}{\sin(\alpha)}} \right ^2 e^{-\frac{-2i\pi nk}{N}} \right ^2$	Uniqueness (up to global phase, translated, reflected and scaled signal) for Hermite and compact support functions, rectangular pulse trains and linear combinations of Gaussians [18]. It has not been reported an algorithm that solve this problem in a polynomial time.
this paper	$\mathbf{A}[p, k] := \left \sum_{n=0}^{N-1} \mathbf{x}[n] \overline{\mathbf{x}[n-p]} e^{-\frac{-2i\pi nk}{N}} \right ^2$	Uniqueness (up to global phase, translated, reflected and scaled signal) for band/time-limited signals. Non-convex algorithm, employing an initialization followed by a gradient descent update rule.

signal from its phaseless data by using prior information such as sparsity [31] or signal from a constellation [32], using additional measurements such as short-time Fourier transform [22], [25], masked data, structured illumination and coded diffraction patterns [33]. These randomness and sparsity assumptions in the phase retrieval problem guarantee that it is no longer NP-hard and can be solved using traditional optimization strategies such as gradient descent [34] and semidefinite relaxations [33].

The above-mentioned algorithms are convex and non-convex formulations, where the later be described as a two-step procedure which starts with a carefully design initial guess that is then refined using a gradient descent strategy which takes polynomial time to converge. The initialization process is intended to reach the vicinity of the solution in order to ensure that a gradient descent update rule will converge zero i.e. a first optimal point. Therefore, the development of an algorithm that follows this two-step strategy to solve the radar phase retrieval problem is desired.

C. Motivation for Studying Phase Retrieval in Radar

The word “RADAR” is an acronym for Radio Detection And Ranging. As it was originally conceived, radio waves were used to detect the presence of a target and to determine its distance or range [9]. After its invention during World War II, a lot of advances from a technological point of view were done on espionage and intelligence gathering missions behind enemy lines [1], [9], [11]. Additionally, from a theoretical perspective efforts by many researchers have been done to improve the performance of radars by increasing the range resolution by changing the bandwidth of the pulse through the modulation of its amplitude, frequency or phase. In fact, there is a need to choose the correct type of waveform, since a correct decision in this respect is always a cost-effective decision. As discussed in Section I-A, choosing a well suited signal for detection tasks in radar is called waveform design.

The waveform design literature can be divided into two classes: First, exploiting the properties of the ambiguity function as modeled in (2), (3), (4), and (5); Second using the ambiguities exposed in Section (I-B). As discussed in Section I-A, for the first approach

there is a large body of literature in order to find pulses that allows to better detect a target. However, the second strategy, which is a phase retrieval problem, is a barely explored scenario in the state-of-the-art. Specifically, in 1970 Rudolf De Buda proves that Hermite functions can be recovered from its radar ambiguity function [13]. This is known as the first work in the radar phase retrieval literature. Some following works have extended this analysis to the fractional Fourier transform version of the radar ambiguity function as in (6), however, any theoretical progress on different kind of signals have been performed yet in the state-of-the-art.

The lack of analytical results in radar for a larger set of signals motivates this work to study the waveform design problem in radar from a different perspective than the literature. Specifically, despite the fact that in the radar literature different waveform designs strategies seem to successfully work, it is important mentioning that these designs are just restricted to Hermite or rectangular pulse trains functions without performance guarantees to estimate the signal. Another drawback of these works is that they do not provide a polynomial time computable procedure to estimate the signal from its ambiguity function. Thus, they also lack of performance guarantees to estimate the signal. Therefore, any forward steps on waveform design for different signals rather than Hermite functions or rectangular pulse trains is needed in the literature. For instance, time-limited and band-limited signals are highly employed in radar systems [35], [36], due to the increased electromagnetic congestion throughout the frequency bands at which radars operate, however any theoretical uniqueness result or recovering algorithm has been exposed in the literature.

D. Our contributions, their Significance and Novelty

In this paper, to estimate a (time) band-limited signal from the radar ambiguity function, a uniqueness theoretical result which states that the underlying signal can be recovered from at least $(3S)3B$ measurements where (S) B is the (pulse-width) bandwidth, respectively is presented. Additionally, a trust region algorithm that minimizes a smoothed non-convex least-squares objective function is proposed to iteratively estimate the band-limited signal of interest. This kind of methods define a region around the current iterate within

which they trust the model to be an adequate representation of the objective function, and then choose the step to be the approximate minimizer of the model in this region. In particular, the proposed method consists of two steps. First, we approximate the signal by an iterative spectral algorithm. Then, the attained initialization is refined based upon a sequence of gradient iterations. To the best of our knowledge this work is seminal in the sense of solving the radar phase retrieval problem for both time and band-limited signals. Simulations results suggest that the proposed algorithm is able to estimate a band-limited signal from the radar ambiguity function for both complete and incomplete radar ambiguity function. The radar function is incomplete when only few shifts or Fourier frequencies are considered. In Theorem 1 we provide theoretical justification for the success of the algorithm by showing that in the vicinity of the true solution, it converges to a first optimal point.

1) *Significance*: Considering that radars not only employ Hermite or pulse trains signals to detect objects it is imperative to extend the analytical results in radar for a larger set of pulses, such time/band-limited signals. Specifically, the provided analytical results will open the possibility to design waveform for radars via a phase retrieval, strategy that has not been explored in the literature.

2) *Novelty*: Contrary to prior works the contributions of this paper are the theoretical result staying that time/band-limited signals can be uniquely identified from its ambiguity function; the recovery algorithm consisting in a designed initialization that is then refined using a gradient update rule to recover the target signal from its ambiguity function; theoretical analysis on the convergence of the proposed reconstruction method for complete noiseless and noisy ambiguity function.

Throughout this paper, we denote by $\mathbb{R}_+ := \{w \in \mathbb{R} : w \geq 0\}$ and $\mathbb{R}_{++} := \{w \in \mathbb{R} : w > 0\}$ the sets of positive and strictly positive real numbers, respectively. The conjugate and the conjugate transpose of the vector $\mathbf{w} \in \mathbb{C}^N$ are denoted as $\bar{\mathbf{w}} \in \mathbb{C}^N$ and $\mathbf{w}^H \in \mathbb{C}^N$, respectively. The n th entry of a vector \mathbf{w} , which is assumed to be periodic, is written as $\mathbf{w}[n]$. We denote by $\tilde{\mathbf{w}}$ and $\hat{\mathbf{w}}$ the Fourier transform of a vector and its conjugate reflected version (that is, $\tilde{\mathbf{w}}[n] := \bar{\mathbf{w}}[-n]$). The notation $\text{diag}(\mathbf{W}, \ell)$ refers to a column vector with entries $\mathbf{W}[j, (j + \ell) \bmod N]$ for $j = 0, \dots, N - 1$. For vectors, $\|\mathbf{w}\|_p$ is the ℓ_p norm. Additionally, we use \odot , and $*$ for the Hadamard (point-wise) product, and convolution, respectively. Finally, $\mathbb{E}[\cdot]$ represents the expected value.

E. Organization

The paper is organized as follows. We begin in Section II by introducing necessary background on the radar phase retrieval problem. Sections IV and III presents the proposed initialization technique and introduces an iterative procedure to refine the solution by minimizing a smooth least-squares objective. Section VIII presents numerical results and compares our approach with competitive algorithms. Finally, Section IX concludes the paper.

II. PROBLEM FORMULATION

Mathematically, the radar *ambiguity function* of a discrete signal $\mathbf{x} \in \mathbb{C}^N$ is defined as

$$\mathbf{A}[p, k] := \left| \sum_{n=0}^{N-1} \mathbf{x}[n] \overline{\mathbf{x}[n-p]} e^{-2i\pi nk/N} \right|^2, \quad (7)$$

where $\bar{\mathbf{x}}$ is the conjugate of \mathbf{x} , and $i = \sqrt{-1}$. The ambiguity function defined in (7) can be considered as a map $\mathbb{C}^N \rightarrow \mathbb{R}_+^{N \times N}$ that has four types of symmetry, usually called *trivial ambiguities* in the radar PR literature. These ambiguities are summarized in Section I-B.

Our goal is to estimate the signal \mathbf{x} , up to trivial ambiguities, from the ambiguity function \mathbf{A} . In this work it is established that the signal \mathbf{x} can be uniquely identified (up to trivial ambiguities) from its ambiguity function under rather mild conditions as summarized in the Proposition 1 using the following definition of a band-limited signal.

A. Uniqueness for Band-limited Signals

In this section we provide uniqueness guarantees for band-limited signals. To that end, we introduce the following definition.

Definition 1. We say that $\mathbf{x} \in \mathbb{C}^N$ is a *B-band-limited* signal if its Fourier transform $\hat{\mathbf{x}} \in \mathbb{C}^N$ contains $N - B$ consecutive zeros. That is, there exists k such that $\hat{\mathbf{x}}[k] = \dots = \hat{\mathbf{x}}[N+k+B-1] = 0$.

Considering the above definition, we prove that the ambiguity function is able to uniquely identify a band-limited signal, which is proved in Proposition 1. In addition, before proving this, we have to recall two auxiliary lemmas as follows.

Lemma 1. ([30, Corollary IV.3]) If $m \geq 2|\mathcal{J} - \mathcal{J}| - 1 + 2|\mathcal{J}|$ and $N > |\mathcal{J}|$ (that is, at least one signal entry is known), then almost every $\mathbf{w} \in \mathbb{C}^N$ is determined uniquely by $\{|\tilde{\mathbf{w}}[k]|\}_{k=0}^{m-1}$. Here, \mathcal{J} is the set of indices of the unknowns, $|\mathcal{J}|$ represents its cardinality, and $\mathcal{J} - \mathcal{J} = \{n_1 - n_2 | n_1, n_2 \in \mathcal{J}\}$.

Lemma 2. ([37, Corollary 2]) Almost every complex-valued signal $\mathbf{w} \in \mathbb{C}^N$ can be uniquely recovered from $\{|\tilde{\mathbf{w}}[k]|\}_{k=0}^{N-1}$ and $\{|\mathbf{w}[n]|\}_{n=0}^{N-1}$ up to rotations.

Proposition 1. Let $\mathbf{x} \in \mathbb{C}^N$ be a *B*-band-limited signal as in Definition 1 for some $B \leq N/2$. Then almost all signals are uniquely determined from their ambiguity function $\mathbf{A}[p, k]$, up to trivial ambiguities, from $m \geq 3B$ measurements. If in addition we have access to the signal's power spectrum and $N \geq 3$, then $m \geq 2B$ measurements suffice.

Proof. We begin the proof by reformulating the measurement model to a more convenient structure. Applying the inverse Fourier transform we write $\mathbf{x}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{\mathbf{x}}[k] e^{2\pi i kn/N}$. Then, according to (7), we have

$$\mathbf{A}[p, k] = |\mathbf{S}[p, k]|^2, \quad (8)$$

where $\mathbf{S}[p, k]$ is defined as

$$\begin{aligned} \mathbf{S}[p, k] &= \sum_{n=0}^{N-1} \mathbf{x}[n] \overline{\mathbf{x}[n-p]} e^{-2i\pi nk/N} \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \left(\sum_{\ell_1=0}^{N-1} \hat{\mathbf{x}}[\ell_1] e^{2\pi i \ell_1 n/N} \right) \\ &\quad \times \left(\sum_{\ell_2=0}^{N-1} \bar{\hat{\mathbf{x}}}[\ell_2] e^{-2\pi i \ell_2 n/N} e^{2\pi i \ell_2 p/N} \right) e^{-2\pi i k n/N} \\ &= \frac{1}{N^2} \sum_{\ell_1, \ell_2=0}^{N-1} \hat{\mathbf{x}}[\ell_1] \bar{\hat{\mathbf{x}}}[\ell_2] e^{2\pi i \ell_2 p/N} \sum_{n=0}^{N-1} e^{2\pi i n(\ell_1 - \ell_2 - k)/N} \\ &= \frac{1}{N} \sum_{\ell=0}^{N-1} \hat{\mathbf{x}}[\ell + k] \bar{\hat{\mathbf{x}}}[\ell] e^{2\pi i \ell p/N}, \end{aligned} \quad (9)$$

since the later sum is equal to N if $\ell_1 = k + \ell_2$ and zero otherwise.

Assume that $B = N/2$, N is even, that $\hat{\mathbf{x}}[n] \neq 0$ for $k = 0, \dots, B - 1$, and that $\hat{\mathbf{x}}[n] = 0$ for $k = N/2, \dots, N - 1$. If the signal's nonzero Fourier coefficients are not in the interval $0, \dots, N/2 - 1$, then we can cyclically reindex the signal without

affecting the proof. If N is odd, then one should replace $N/2$ by $\lfloor N/2 \rfloor$ everywhere in the sequel. Clearly, the proof carries through for any $B \leq N/2$.

Considering (9), the bandlimit assumption on the signal forms a “inverted pyramid” structure. Here, each row represents fixed k and varying ℓ of $\tilde{\mathbf{x}}[\ell+k]\tilde{\mathbf{x}}[\ell]$ for $k = 0, \dots, N/2 - 1$

$$\begin{aligned} & |\tilde{\mathbf{x}}[0]|^2, |\tilde{\mathbf{x}}[1]|^2, \dots, |\tilde{\mathbf{x}}[B-1]|^2, 0, \dots, 0 \\ & \overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[1]}, \overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[2]}, \dots, \overline{\tilde{\mathbf{x}}[B-2]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0 \\ & \overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[2]}, \overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[3]}, \dots, \overline{\tilde{\mathbf{x}}[B-3]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0 \\ & \vdots \\ & \overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0, 0, \dots, 0 \\ & 0, 0, \dots, \overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0 \\ & 0, \dots, \overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[B-2]}, \overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0 \\ & \vdots \\ & 0, \mathbf{x}[0]\overline{\tilde{\mathbf{x}}[1]}, \overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[2]}, \dots, \overline{\tilde{\mathbf{x}}[B-2]\tilde{\mathbf{x}}[B-1]}, 0, \dots, 0. \end{aligned} \quad (10)$$

Then, $\mathbf{S}[p, k]$ as in (9) is a subsample of the Fourier transform of each one of the pyramid’s rows.

Step 0: From the $(B-1)$ -th row of (10), we see that

$$|\mathbf{S}[p, B]| = |\tilde{\mathbf{x}}[0]||\tilde{\mathbf{x}}[B-1]|, \forall p = 0, \dots, N-1. \quad (11)$$

Considering that in the radar phase retrieval problem the translation ambiguity is continuous (second ambiguity in Section I-B), we can set $\tilde{\mathbf{x}}[0]$ to be real and, without loss of generality it can be assumed that $\tilde{\mathbf{x}}[0] = 1$ [23]. Note that in contrast to the FROG phase retrieval problem, this continuity property in the radar scenario is satisfied for general signals. Then, from (11) we obtain that

$$|\mathbf{S}[p, B-1]| = |\tilde{\mathbf{x}}[B-1]|, \forall p = 0, \dots, N-1. \quad (12)$$

Step 1: From the first row of (10), we conclude the following system of equations

$$|\mathbf{S}[p, 0]| = \frac{1}{N} \left| \sum_{\ell=0}^{B-1} |\tilde{\mathbf{x}}[\ell]|^2 e^{2\pi i \ell p/N} \right|, p = 0, \dots, N-1. \quad (13)$$

Given the fact that from the **Step 0** the entries $|\tilde{\mathbf{x}}[0]|, |\tilde{\mathbf{x}}[B-1]|$, and $\{|\mathbf{S}[p, 0]|\}_{p=0}^{N-1}$ are known, then appealing to Lemma 1 for almost all signals we have that $|\tilde{\mathbf{x}}[1]|, \dots, |\tilde{\mathbf{x}}[B-2]|$ are uniquely determined. It is worth mentioning that this previous argument does not imply that $\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[B-1]$ are uniquely determined. In fact there are up to 2^{B-1} vectors, modulo global phase, reflection and conjugation that satisfy the constraints in (11), and (13) [38, Section 3.1].

Step 2: Moving to analyze the second row of (10) we obtain the following system of equations

$$|\mathbf{S}[p, 1]| = \frac{1}{N} \left| \sum_{\ell=0}^{B-2} \tilde{\mathbf{x}}[\ell+1]\overline{\tilde{\mathbf{x}}[\ell]} e^{2\pi i \ell p/N} \right|, p = 0, \dots, N-1. \quad (14)$$

Fix one of the possible solutions for $\tilde{\mathbf{x}}[1]$ from **Step 1**. Then, since $\tilde{\mathbf{x}}[0]$ is known, Lemma 1 states that for almost all signals $\overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[2]}, \dots, \overline{\tilde{\mathbf{x}}[B-1]\tilde{\mathbf{x}}[B-2]}$ are uniquely determined.

Step 3: Considering the fact that $\tilde{\mathbf{x}}[0]$, and $\tilde{\mathbf{x}}[1]$ are known, from **Step 2** we can estimate $\tilde{\mathbf{x}}[2]$. Thus, since $\overline{\tilde{\mathbf{x}}[0]\tilde{\mathbf{x}}[2]}$ is known, appealing to Lemma 1 for almost all signals $\overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[3]}, \dots, \overline{\tilde{\mathbf{x}}[B-3]\tilde{\mathbf{x}}[B-1]}$ are uniquely determined. However, remark that at this stage the 2^{B-1} possible solutions from **Step 2** remains.

Despite the large amount of possible solutions, we can prove that at this step there is only one vector (up to trivial ambiguities) out of the 2^{B-1} possibilities of **Step 2**, that is consistent with the constraints in (11), (13), and (14). To see this, from **Step 1** we have that $|\tilde{\mathbf{x}}[0]|, \dots, |\tilde{\mathbf{x}}[B-1]|$ are uniquely determined. Therefore, from the knowledge of $\{|\tilde{\mathbf{x}}[\ell]|\}_{\ell=0}^{B-1}$, and $\{|\mathbf{S}[p, 0]|\}_{p=0}^{N-1}$, from Lemma 2 we have that $\overline{\tilde{\mathbf{x}}[1]\tilde{\mathbf{x}}[2]}, \dots, \overline{\tilde{\mathbf{x}}[B-1]\tilde{\mathbf{x}}[B-2]}$ are uniquely determined for almost all signals. This previous fact leads to a unique selection (up to trivial ambiguities) of $\tilde{\mathbf{x}}[1]$ in **Step 2**, and in consequence a unique selection of $\tilde{\mathbf{x}}[2]$ in this step.

Step $B-1$: Considering that from the $B-2$ previous steps the entries $\tilde{\mathbf{x}}[0], \dots, \tilde{\mathbf{x}}[B-2]$ were uniquely determined (up to trivial ambiguities), appealing again to Lemma 1 we have that $\tilde{\mathbf{x}}[B-1]$ can be also uniquely determined.

Finally, analyzing the construction process described above we have that at **Step 2**, the signal $\tilde{\mathbf{x}}$ can be uniquely determined, which means that $m \geq 3B$ measurements are needed to solve the radar phase retrieval problem for band-limited signals. If in addition, we have access to the spectrum signal $|\tilde{\mathbf{x}}|$, at **Step 1** we can uniquely determine $\tilde{\mathbf{x}}$ if $N \geq 3$, implying that under this scenario only $m \geq 2B$ measurements are needed. ■

By almost all signals Theorem 1 means that the set of signals which cannot be uniquely determined, up to trivial ambiguities, is contained in the vanishing locus of a nonzero polynomial. Observe that evidently, Proposition 1 states that not all the delay steps are needed to recover the signal, and therefore a method that works in this regime as well is desired. There are two aspects that it is important mentioning. First, the proof of Theorem 1 is a construction procedure that uses two classical results in phase retrieval, Corollary IV.3 in [30], and Corollary 2 in [37]. Second, the proof reveals that the first and the $(B-1)$ -th rows of the ambiguity function in (7) must be perfectly preserved in order to ensure uniqueness (up to trivial ambiguities). Then, since the radar phase retrieval problem is a design approach these two mentioned rows cannot be discarded or corrupted by any distortion noise in the design in order to guarantee uniqueness.

B. Uniqueness for Time-limited Signals

A direct consequence of Proposition 1 is the following corollary, under rather mild conditions, states that for almost all time-limited signals as in Definition 2 can be recovered.

Definition 2. We say that $\mathbf{x} \in \mathbb{C}^N$ is a S -timelimited signal if $\mathbf{x} \in \mathbb{C}^N$ contains $N-S$ consecutive zeros. That is, there exists k such that $\mathbf{x}[k] = \dots = \mathbf{x}[N+k+S-1] = 0$.

Corollary 1. Let $\mathbf{x} \in \mathbb{C}^N$ be a S -band-limited signal as in Definition 1 for some $S \leq N/2$. Then almost all signals are uniquely determined from their ambiguity function $\mathbf{A}[p, k]$, up to trivial ambiguities, from $m \geq 3S$ measurements. If in addition we have access to the signal’s power spectrum and $N \geq 3$, then $m \geq 2S$ measurements suffice.

Proof. Recall that according to (7), we have

$$\mathbf{A}[p, k] = \left| \sum_{n=0}^{N-1} \mathbf{x}[n] \overline{\mathbf{x}[n-p]} e^{-2\pi i \pi n k / N} \right|^2. \quad (15)$$

Assume that $S = N/2$, N is even, that $\mathbf{x}[n] \neq 0$ for $n = 0, \dots, S-1$, and that $\mathbf{x}[n] = 0$ for $n = N/2, \dots, N-1$. If the signal’s nonzero coefficients are not in the interval $0, \dots, N/2-1$, then we can cyclically reindex the signal without affecting the proof. If N is odd, then one should replace $N/2$ by $\lfloor N/2 \rfloor$ everywhere in the sequel. Clearly, the proof carries through for any $S \leq N/2$.

Considering (9), the bandlimit assumption on the signal forms a “inverted pyramid” structure. Here, each row represents fixed p and varying n of $\mathbf{x}[n-p]\mathbf{x}[n]$ for $p = 0, \dots, N/2 - 1$

$$\begin{aligned} & |\mathbf{x}[0]|^2, |\mathbf{x}[1]|^2, \dots, |\mathbf{x}[S-1]|^2, 0, \dots, 0 \\ & 0, \overline{\mathbf{x}[0]}\mathbf{x}[1], \overline{\mathbf{x}[1]}\mathbf{x}[2], \dots, \overline{\mathbf{x}[S-2]}\mathbf{x}[S-1], 0, \dots, 0 \\ & 0, 0, \overline{\mathbf{x}[0]}\mathbf{x}[2], \overline{\mathbf{x}[1]}\mathbf{x}[3], \dots, \overline{\mathbf{x}[S-3]}\mathbf{x}[S-1], 0, \dots, 0 \\ & \vdots \\ & 0, 0, \dots, 0, \overline{\mathbf{x}[0]}\mathbf{x}[S-1], 0, \dots, 0, 0, \dots, 0 \\ & \mathbf{x}[0]\overline{\mathbf{x}[S-1]}, 0, 0, \dots, 0 \\ & \mathbf{x}[0]\overline{\mathbf{x}[S-2]}, \mathbf{x}[1]\overline{\mathbf{x}[S-1]}, 0, 0, \dots, 0 \\ & \vdots \\ & \mathbf{x}[0]\overline{\mathbf{x}[1]}, \mathbf{x}[1]\overline{\mathbf{x}[2]}, \dots, \mathbf{x}[S-2]\overline{\mathbf{x}[S-1]}, 0, 0, \dots, 0. \end{aligned} \quad (16)$$

Then, $\mathbf{A}[p, k]$ as in (15) is a subsample of the Fourier transform of each one of the pyramid’s rows.

Therefore, performing an analogous construction procedure as in Proposition 1 over (16) we have that the signal $\tilde{\mathbf{x}}$ can be uniquely determined from $m \geq 3S$ measurements. If in addition, we have access to the spectrum signal $|\tilde{\mathbf{x}}|$, we can uniquely determine $\tilde{\mathbf{x}}$ if $N \geq 3$, implying that under this scenario only $m \geq 2S$ measurements are needed. ■

We remark here that the notion almost all signals is the same as in Theorem 1. Additionally, the proof of Corollary 1 is also a construction procedure, and that the first and the $(B-1)$ -th rows of the ambiguity function in (7) must be perfectly preserved in order to ensure uniqueness (up to trivial ambiguities).

C. Phase Retrieval Inverse Problem for radar

To take the ambiguities into account, we measure the relative error between the true signal \mathbf{x} and any $\mathbf{w} \in \mathbb{C}^N$ as

$$\text{dist}(\mathbf{x}, \mathbf{w}) := \frac{\|\sqrt{\mathbf{A}} - \sqrt{\mathbf{W}}\|_F}{\|\sqrt{\mathbf{A}}\|_F}, \quad (17)$$

where \mathbf{A} is the ambiguity function of \mathbf{x} according to (7), $\sqrt{\cdot}$ is the point-wise square root, \mathbf{W} is the ambiguity function of \mathbf{w} , and $\|\cdot\|_F$ denotes the Frobenius norm. Note that if $\text{dist}(\mathbf{x}, \mathbf{w}) = 0$, and the uniqueness conditions of Proposition 1 are met, then for almost all signals \mathbf{w} is equal to \mathbf{x} up to trivial ambiguities.

In recent years, many papers have examined the problem of recovering a signal from phaseless quadratic random measurements. A popular approach is to minimize the intensity least-squares objective; see for instance [34]. Recent works have shown that minimizing the amplitude least-squares objective leads to better reconstruction under noisy scenarios [39]–[41]. However, the latter cost function is non-smooth and thus may lead to a biased descent direction [32]. To overcome the non-smoothness of the objective function, we follow the smoothing strategy proposed in [32].

The smooth objective to recover the underlying signal considered in this work is

$$\min_{\mathbf{z} \in \mathbb{C}^N} h(\mathbf{z}, \mu) = \min_{\mathbf{z} \in \mathbb{C}^N} \frac{1}{N^2} \sum_{k,p=0}^{N-1} \ell_{k,p}(\mathbf{z}, \mu), \quad (18)$$

where

$$\ell_{k,p}(\mathbf{z}, \mu) := \left[\varphi_\mu \left(\left| \sum_{n=0}^{N-1} \mathbf{z}[n] \overline{\mathbf{z}[n-p]} e^{-2\pi i n k / N} \right| \right) - \sqrt{\mathbf{A}[p, k]} \right]^2. \quad (19)$$

The function $\varphi_\mu : \mathbb{R} \rightarrow \mathbb{R}_{++}$ in (19) is defined as

$$\varphi_\mu(w) := \sqrt{w^2 + \mu^2},$$

with $\mu \in \mathbb{R}_{++}$ (a tunable parameter). Notice that if $\mu = 0$, then (19) reduces to the non-smooth formulation. In [40], the authors addressed the non-smoothness by introducing truncation parameters into the gradient step in order to eliminate the errors in the estimated descent direction. However, this procedure can modify the search direction and increase the sample complexity of the phase retrieval problem [32].

In this work we propose a trust region algorithm based on the Cauchy point to solve (18), that is initialized by a spectral procedure which requires only a few iterations. Section III explains in detail the proposed algorithm.

III. RECONSTRUCTION ALGORITHM

In order to solve the optimization problem in (18), we develop a trust region algorithm, based on the Cauchy point, that is initialized by the outcome of a spectral method approximating the signal \mathbf{x} which will be explained in Section IV.

The standard update rule in this kind of methods takes the form of

$$\mathbf{x}^{(t+1)} := \mathbf{x}^{(t)} + \alpha^{(t)} \mathbf{b}^{(t)}, \quad (20)$$

where $\alpha^{(t)}$ is the step size at iteration t and the vector $\mathbf{b}^{(t)}$ is chosen in this work as

$$\begin{aligned} \mathbf{b}^{(t)} &:= \arg \min_{\mathbf{b} \in \mathbb{C}^n} h(\mathbf{x}^{(t)}, \mu^{(t)}) + 2\mathcal{R}(\mathbf{b}^H \mathbf{d}^{(t)}), \\ s.t. \quad & \|\mathbf{b}\|_2 \leq \mu^{(t)} \end{aligned} \quad (21)$$

with $\mathcal{R}(\cdot)$ as the real part function, and $\mathbf{d}^{(t)}$ as the gradient of $h(\mathbf{z}, \alpha)$ with respect to $\bar{\mathbf{z}}$ at iteration t . The solution to (21) is given by [42, Chapter 4]

$$\mathbf{b}^{(t)} = -\frac{\mu^{(t)}}{\|\mathbf{d}^{(t)}\|_2} \mathbf{d}^{(t)}. \quad (22)$$

To mathematically compute $\mathbf{d}^{(t)}$, the Wirtinger derivatives as introduced in [43] are employed. Let us define the vector \mathbf{f}_k^H as

$$\mathbf{f}_k^H := [\omega^{-0(k-1)}, \omega^{-1(k-1)}, \dots, \omega^{-(N-1)(k-1)}], \quad (23)$$

with $\omega = e^{\frac{2\pi i}{n}}$ the n th root of unity. Then, the Wirtinger derivative of $h(\mathbf{z}, \mu)$ in (18) with respect to $\mathbf{z}[\ell]$ is given by

$$\begin{aligned} \frac{\partial h(\mathbf{z}, \mu)}{\partial \mathbf{z}[\ell]} &:= \frac{1}{N^2} \sum_{k,p=0}^{N-1} \left(\mathbf{f}_k^H \mathbf{g}_p - v_{k,p} \right) \mathbf{z}[\ell - p] e^{2\pi i \ell k / N} \\ &+ \frac{1}{N^2} \sum_{k,p=0}^{N-1} \left(\mathbf{f}_k^T \overline{\mathbf{g}_p} - v_{k,p} \right) \mathbf{z}[\ell + p] e^{-2\pi i (\ell + p) k / N}, \end{aligned} \quad (24)$$

where $v_{k,p} := \sqrt{\mathbf{A}[p, k]} \frac{\mathbf{f}_k^H \mathbf{g}_p}{\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p|)}$, and

$$\mathbf{g}_p := [\mathbf{z}[0]\overline{\mathbf{z}[p]}, \dots, \mathbf{z}[N-1]\overline{\mathbf{z}[N-1+p]}]^T. \quad (25)$$

The gradient $\mathbf{d}^{(t)}$ is then given by

$$\mathbf{d}^{(t)} := \left[\frac{\partial h(\mathbf{z}^{(t)}, \mu)}{\partial \mathbf{z}[0]}, \dots, \frac{\partial h(\mathbf{z}^{(t)}, \mu)}{\partial \mathbf{z}[N-1]} \right]^T. \quad (26)$$

To alleviate the memory requirements and computational complexity required for large N , we suggest a block stochastic gradient

descent strategy. Instead of calculating (24), we choose only a random subset of the sum for each iteration t , that is,

$$\begin{aligned} \mathbf{d}_{\Gamma(t)}[\ell] &= \sum_{p,k \in \Gamma(t)} \left(\mathbf{f}_k^H \mathbf{g}_p^{(t)} - v_{k,p,t} \right) \mathbf{x}^{(t)}[\ell-p] e^{2\pi i \ell k / N} \\ &+ \sum_{p,k \in \Gamma(t)} \left(\mathbf{f}_k^T \bar{\mathbf{g}}_p^{(t)} - v_{k,p} \right) \mathbf{x}^{(t)}[\ell+p] e^{-2\pi i (\ell+p) k / N}, \end{aligned} \quad (27)$$

where the set $\Gamma(t)$ is chosen uniformly and independently at random at each iteration t from subsets of $\{1, \dots, N\}^2$ with cardinality Q . Specifically, the gradient in (26) is uniformly sampled using a minibatch of data, in this case of size Q for each step update, such that in expectation is (24) [44, page 130].

As mentioned in Section III, choosing $\mu > 0$ prevents bias in the update direction. Since the function h is smooth, we are able to construct a descent rule for μ (Line 13 of Algorithm 1) in order to guarantee convergence to a first-order optimal point, that is, a point with zero gradient, in the vicinity of the solution.

Algorithm 1 Recovery from the ambiguity function

- 1: **Input:** Data $\{\mathbf{A}[p,k] : k, p = 0, \dots, N-1\}$. Choose constants $\gamma_1, \gamma, \alpha \in (0, 1)$, $\mu^{(0)} \geq 0$, cardinality $Q \in \{1, \dots, N^2\}$, and tolerance $\epsilon > 0$.
 - 2: Initial point $\mathbf{x}^{(0)} \leftarrow$ Algorithm 2($\mathbf{A}[p,k], T$).
 - 3: **while** $\|\mathbf{b}_{\Gamma(t)}\|_2 \geq \epsilon$ **do**
Choose $\Gamma(t)$ uniformly at random from the subsets of $\{1, \dots, N\}^2$ with cardinality Q per iteration $t \geq 0$.
 - 4: $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \alpha^{(t)} \mathbf{b}_{\Gamma(t)} = \mathbf{x}^{(t)} - \alpha^{(t)} \frac{\mu^{(t)}}{\|\mathbf{d}_{\Gamma(t)}\|_2} \mathbf{d}_{\Gamma(t)},$
where
 - 5:
 - 6: $v_{k,p,t} = \sqrt{\mathbf{A}[p,k]} \frac{\mathbf{f}_k^H \mathbf{g}_p^{(t)}}{\varphi_{\mu^{(t)}}(|\mathbf{f}_k^H \mathbf{g}_p^{(t)}|)}.$
 - 7: $\mathbf{g}_p^{(t)} = [\mathbf{x}^{(t)}[0] \bar{\mathbf{x}}[p]^{(t)}, \dots, \mathbf{x}^{(t)}[N-1] \bar{\mathbf{x}}[N-1+p]^{(t)}]^T.$
 - 8: **if** $\|\mathbf{d}_{\Gamma(t)}\|_2 \geq \gamma \mu^{(t)}$ **then**
 - 9: $\mu^{(t+1)} = \mu^{(t)}.$
 - 10: **else**
 - 11: $\mu^{(t+1)} = \gamma_1 \mu^{(t)}.$
 - 12: **end if**
 - 13: **end while**
 - 14: **return:** $\mathbf{x}^{(T)}.$
-

Theorem 1. Let \mathbf{x} be S -time-limited or B -band-limited for some $S \leq N/2$ or $B \leq N/2$, respectively, satisfying $\text{dist}(\mathbf{x}, \mathbf{x}^{(t)}) \leq \rho$ for some sufficiently small constant $\rho > 0$. Suppose that $\Gamma(t)$ is sampled uniformly at random from all subsets of $\{1, \dots, N\}^2$ with cardinality Q , independently for each iteration. Then for almost all signals, Algorithm 1 with step size $\alpha \in (0, \frac{2}{U}]$ satisfies

$$\lim_{t \rightarrow \infty} \mu^{(t)} = 0, \text{ and } \lim_{t \rightarrow \infty} \|\mathbf{d}^{(t)}\|_2 = 0, \quad (28)$$

for some constant $U > 0$ depending on ρ .

Proof. See Section V. ■

IV. INITIALIZATION ALGORITHM

In this section we devise a method to initialize the gradient iterations. This strategy approximates the signal \mathbf{x} from the ambiguity function as the leading eigenvector of a carefully designed matrix.

Instead of directly dealing with the ambiguity function in (7), we consider the acquired data in a transformed domain by taking its 1D DFT with respect to the frequency variable (normalized by $1/N$). Our measurement model is then

$$\begin{aligned} \mathbf{Y}[p, \ell] &= \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{A}[p, k] e^{-2\pi i k \ell / N} \\ &= \frac{1}{N} \sum_{k,n,m=0}^{N-1} \mathbf{x}[n] \bar{\mathbf{x}}[n-p] \mathbf{x}[m-p] \bar{\mathbf{x}}[m] e^{-2\pi i k \frac{(m-n-\ell)}{N}} \\ &= \sum_{n=0}^{N-1} \mathbf{x}[n] \bar{\mathbf{x}}[n-p] \mathbf{x}[n+\ell-p] \bar{\mathbf{x}}[n+\ell], \end{aligned} \quad (29)$$

where $p, \ell = 0, \dots, N-1$. Observe that for fixed p , $\mathbf{Y}[p, \ell]$ is the autocorrelation of $\mathbf{x} \odot \bar{\mathbf{x}}_p$, where $\mathbf{x}_p[n] = \mathbf{x}[n-p]$.

Let $\mathbf{D}_p \in \mathbb{C}^{N \times N}$ be a diagonal matrix composed of the entries of \mathbf{x}_p , and let \mathbf{C}_ℓ be a circulant matrix that shifts the entries of a vector by ℓ locations, namely, $(\mathbf{C}_\ell \mathbf{x})[n] = \mathbf{x}[n+\ell]$. Then, the matrix $\mathbf{X} := \mathbf{x} \mathbf{x}^H$ is linearly mapped to $\mathbf{Y}[p, \ell]$ as follows:

$$\begin{aligned} \mathbf{Y}[p, \ell] &= (\bar{\mathbf{D}}_{p+\ell} \mathbf{D}_p \mathbf{C}_\ell \mathbf{x})^H \mathbf{x} = \mathbf{x}^H \mathbf{A}_{p,\ell} \mathbf{x} \\ &= \text{tr}(\mathbf{X} \mathbf{A}_{p,\ell}), \end{aligned} \quad (30)$$

where $\mathbf{A}_{p,\ell} = \mathbf{C}_{-\ell} \bar{\mathbf{D}}_p \mathbf{D}_{p+\ell}$, and $\text{tr}(\cdot)$ denotes the trace function. Observe that $\mathbf{C}_\ell^T = \mathbf{C}_{-\ell}$. Thus, we have that

$$\mathbf{y}_\ell = \mathbf{G}_\ell \mathbf{x}_\ell, \quad (31)$$

for a fixed $\ell \in \{0, \dots, N-1\}$, where $\mathbf{y}_\ell[n] = \mathbf{Y}[n, \ell]$ and $\mathbf{x}_\ell = \text{diag}(\mathbf{X}, \ell)$. The (p, n) th entry of the matrix $\mathbf{G}_\ell \in \mathbb{C}^{\lceil \frac{N}{L} \rceil \times N}$ is given by

$$\mathbf{G}_\ell[p, n] := \bar{\mathbf{x}}_p[n] \mathbf{x}_p[n+\ell]. \quad (32)$$

From (32) it follows that \mathbf{G}_ℓ is a circulant matrix. Therefore, \mathbf{G}_ℓ is invertible if and only if the DFT of its first column, in this case $\bar{\mathbf{x}} \odot (\mathbf{C}_\ell \mathbf{x})$, is non-vanishing.

Using (31), we propose a method to estimate the signal \mathbf{x} from measurements (7) using an alternating scheme: fixing \mathbf{G}_ℓ , solving for \mathbf{x}_ℓ , updating \mathbf{G}_ℓ and so forth.

We start the alternating scheme with the initial point

$$\mathbf{x}_{\text{init}}[p] := \mathbf{v}[p] \exp(i\theta[p]), \quad (33)$$

where $\theta[r] \in [0, 2\pi)$ is chosen uniformly at random for all $r \in \{0, \dots, N-1\}$. The r th entry of \mathbf{v} corresponds to the summation of the measured ambiguity function over the frequency axis:

$$\begin{aligned} \mathbf{v}[p] &:= \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{A}[p, k] = \sum_{k=0}^{N-1} \left| \sum_{n=0}^{N-1} \mathbf{x}[n] \bar{\mathbf{x}}[n-p] e^{-2\pi i n k / N} \right|^2 \\ &:= \sum_{n=0}^{N-1} |\mathbf{x}[n]|^2 |\mathbf{x}[n-p]|^2. \end{aligned} \quad (34)$$

Once the vector \mathbf{x}_{init} is constructed, the vectors $\mathbf{x}_\ell^{(t)}$ at $t = 0$ can be built as

$$\mathbf{x}_\ell^{(0)} = \text{diag}(\mathbf{X}_0^{(0)}, \ell), \quad (35)$$

where

$$\mathbf{X}_0^{(0)} = \mathbf{x}_{\text{init}} \mathbf{x}_{\text{init}}^H. \quad (36)$$

Then, from (35) we proceed with an alternating procedure between estimating the matrix \mathbf{G}_ℓ , and updating the vector \mathbf{x}_ℓ as follows.

- *Update rule for \mathbf{G}_ℓ :* In order to update \mathbf{G}_ℓ , we update the matrix $\mathbf{X}_0^{(t)}$ as

$$\text{diag}(\mathbf{X}_0^{(t)}, \ell) = \mathbf{x}_\ell^{(t)}. \quad (37)$$

Observe that if $\mathbf{x}_\ell^{(t)}$ is close to \mathbf{x}_ℓ for all ℓ , then $\mathbf{X}_0^{(t)}$ is close to $\mathbf{x}\mathbf{x}^H$. Letting $\mathbf{w}^{(t)}$ be the leading (unit-norm) eigenvector of the matrix $\mathbf{X}_0^{(t)}$ constructed in (37), from (32) each matrix $\mathbf{G}_\ell^{(t)}$ at iteration t is given by

$$\mathbf{G}_\ell^{(t)}[p, n] = \overline{\mathbf{x}_p^{(t)}}[n] \mathbf{x}_p^{(t)}[n + \ell], \quad (38)$$

where $\mathbf{x}_p^{(t)}[n] = \mathbf{w}^{(t)}[n - p]$.

- *Optimization with respect to \mathbf{x}_ℓ :* Fixing $\mathbf{G}_\ell^{(t-1)}$, one can estimate $\mathbf{x}_\ell^{(t)}$ at iteration t by solving the linear least-squares (LS) problem

$$\min_{\mathbf{p}_\ell \in \mathbb{C}^N} \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{p}_\ell\|_2^2. \quad (39)$$

The relationship between the vectors $\mathbf{x}_\ell^{(t)}$ is ignored at this stage. If $\mathbf{G}_\ell^{(t-1)}$ is invertible, then the solution to this problem is given by $(\mathbf{G}_\ell^{(t-1)})^{-1} \mathbf{y}_\ell$. Since $\mathbf{G}_\ell^{(t-1)}$ is a circulant matrix, it is invertible if and only if the DFT of $\overline{\mathbf{x}}^{(t-1)} \odot (\mathbf{C}_\ell \mathbf{x}^{(t-1)})$ is non-vanishing. This condition cannot be ensured in general. Thus, we propose a surrogate proximal optimization problem to estimate $\mathbf{x}_\ell^{(t)}$ by

$$\min_{\mathbf{p}_\ell \in \mathbb{C}^N} \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{p}_\ell\|_2^2 + \frac{1}{2\lambda_{(t)}} \|\mathbf{p}_\ell - \mathbf{x}_\ell^{(t-1)}\|_2^2, \quad (40)$$

where $\lambda_{(t)} > 0$ is a regularization parameter. In practice $\lambda_{(t)}$ is a tunable parameter [45]. In particular, for this work the value of $\lambda_{(t)}$ was determined using a cross-validation strategy such that each simulation uses the value that results in the smallest relative error according to (17). The surrogate optimization problem in (40) is strongly convex [45], and admits the following closed form solution

$$\mathbf{x}_\ell^{(t)} = \mathbf{B}_{\ell,t}^{-1} \mathbf{e}_{\ell,t}, \quad (41)$$

where

$$\begin{aligned} \mathbf{B}_{\ell,t} &= \left(\mathbf{G}_\ell^{(t-1)} \right)^H \left(\mathbf{G}_\ell^{(t-1)} \right) + \frac{1}{2\lambda} \mathbf{I}, \\ \mathbf{e}_{\ell,t} &= \left(\mathbf{G}_\ell^{(t)} \right)^H \mathbf{y}_\ell + \frac{1}{2\lambda} \mathbf{x}_\ell^{(t-1)}, \end{aligned} \quad (42)$$

with $\mathbf{I} \in \mathbb{R}^{N \times N}$ the identity matrix. Clearly $\mathbf{B}_{\ell,t}$ in (42) is always invertible. The update step for each $\mathbf{x}_\ell^{(t)}$ is computed in Line 9 of Algorithm 2.

Finally, in order to estimate \mathbf{x} , the (unit-norm) principal eigenvector of $\mathbf{X}_0^{(T)}$ is normalized by

$$\beta = \sqrt[4]{\sum_{n \in \mathcal{S}} (\mathbf{B}_{0,T}^{-1} \mathbf{e}_{0,T})[n]}, \quad (43)$$

where $\mathcal{S} := \{n : (\mathbf{B}_{0,T}^{-1} \mathbf{e}_{0,T})[n] > 0\}$. Observe that (43) results from the fact that $\sum_{n=0}^{N-1} \text{diag}(\mathbf{X}, 0)[n] = \|\mathbf{x}\|_2^4$.

After a few iterations of this two-step procedure, the output is used to initialize the gradient algorithm described in Section III. This alternating scheme is summarized in Algorithm 2. Next, we theoretically established that the attained $\mathbf{x}^{(0)}$ following the update rules in (38) and (41) leads to a close estimation of the real unknown signal \mathbf{x} for complete radar ambiguity function.

Theorem 2. Suppose that \mathbf{A} in (7) is complete. Assuming that the regularization parameter $\lambda_{(t)}$ satisfies $\lambda_{(t)} \sigma_{\min}^2(\mathbf{G}^{(t-1)}) > 1/2$ for

Algorithm 2 Initialization Procedure

- 1: **Input:** The measurements $\mathbf{A}[p, k]$, T the number of iterations, and $\lambda > 0$.
 - 2: **Output:** $\mathbf{x}^{(0)}$ (estimation of \mathbf{x}).
 - 3: **Initialize:** $\mathbf{x}_{init}[p] = \mathbf{v}[p] \exp(i\theta[p])$, and $\mathbf{v}[p] = \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{A}[p, k]$, $\theta[p] \in [0, 2\pi)$ is chosen uniformly and independently at random.
 - 4: Compute $\mathbf{Y}[p, \ell]$ the 1D inverse DFT with respect to k of $\mathbf{A}[p, k]$.
 - 5: **for** $t = 1$ to T **do**
 - 6: Construct $\mathbf{G}_\ell^{(t)}$ according to (38).
 - 7: Compute $\mathbf{B}_{\ell,t} = (\mathbf{G}_\ell^{(t)})^H (\mathbf{G}_\ell^{(t)}) + \frac{1}{2\lambda} \mathbf{I}$.
 - 8: Compute $\mathbf{e}_{\ell,t} = (\mathbf{G}_\ell^{(t)})^H \mathbf{y}_\ell + \frac{1}{2\lambda} \mathbf{x}_\ell^{(t-1)}$.
 - 9: Construct the matrix $\mathbf{X}_0^{(t)}$ such that
 - $$\text{diag}(\mathbf{X}_0^{(t)}, \ell) = \mathbf{B}_{\ell,t}^{-1} \mathbf{e}_{\ell,t}, \quad \ell = 0, \dots, N-1.$$
 - 10: Let $\mathbf{w}^{(t)}$ be the leading (unit-norm) eigenvector of $\mathbf{X}_0^{(t)}$.
 - 11: Take $\mathbf{x}_p^{(t)}[n] = \mathbf{w}^{(t)}[n - p]$.
 - 12: **end for**
 - 13: Compute vector $\mathbf{x}^{(0)}$ as
 - $$\mathbf{x}^{(0)} := \sqrt[4]{\sum_{n \in \mathcal{S}} (\mathbf{B}_{0,T}^{-1} \mathbf{e}_{0,T})[n]} \mathbf{w}^{(T)},$$
 - where $\mathcal{S} := \{n : (\mathbf{B}_{0,T}^{-1} \mathbf{e}_{0,T})[n] > 0\}$.
 - 14: **return:** $\mathbf{x}^{(0)}$.
-

all $t > 0$, and starting from \mathbf{x}_{init} in (33), then the returned vector $\mathbf{x}^{(0)}$ by Algorithm 2 satisfies

$$\|\mathbf{x}^{(0)} (\mathbf{x}^{(0)})^H - \mathbf{x}\mathbf{x}^H\|_F < \tau \|\mathbf{x}_{init} \mathbf{x}_{init}^H - \mathbf{x}\mathbf{x}^H\|_F, \quad (44)$$

for some $\tau \in (0, 1)$.

Proof. See Section VII. ■

V. PROOF OF THEOREM 1

Proof. Let us define the search set as

$$\mathcal{J} := \{\mathbf{z} \in \mathbb{C}^N, B\text{-bandlimited} : \text{dist}(\mathbf{x}, \mathbf{z}) \leq \rho, B \leq N/2\}, \quad (45)$$

for some small constant $\rho > 0$. Recall that \mathbf{z} is a B -bandlimited signal if there exists k such that $\tilde{\mathbf{z}}[k] = \dots = \tilde{\mathbf{z}}[N+k+B-1] = 0$, where $\tilde{\mathbf{z}}$ is the Fourier transform of \mathbf{z} . The bandlimit condition guarantees that we have unique solution, according to Proposition 1.

In order to prove Theorem 1, the function $h(\mathbf{z}, \mu)$ in (18) needs to satisfy the four requirements stated in the following lemma, which are used in the analysis of convergence for stochastic gradient methods [46].

Lemma 3. The function $h(\mathbf{z}, \mu)$ in (18) and its Wirtinger derivative in (26) satisfy the following properties.

- 1) The cost function $h(\mathbf{z}, \mu)$ in (18) is bounded below.
- 2) The set \mathcal{J} as defined in (45) is closed and bounded.
- 3) There exists a constant $U > 0$, such that

$$\left\| \frac{\partial h(\mathbf{z}_1, \mu)}{\partial \bar{\mathbf{z}}} - \frac{\partial h(\mathbf{z}_2, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2 \leq U \|\mathbf{z}_1 - \mathbf{z}_2\|_2, \quad (46)$$

holds for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{J}$.

- 4) For all $\mathbf{z} \in \mathcal{J}$

$$\mathbb{E}_{\Gamma(t)} \left[\left\| \mathbf{d}_{\Gamma(t)} - \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2^2 \right] \leq \zeta^2, \quad (47)$$

for some $\zeta > 0$, where $\mathbf{d}_{\Gamma(t)}$ is as in Line 9 of Algorithm 1.

Proof. See Appendix VI. ■

To prove Theorem 1, denote the set $\mathcal{K}_1 := \{t|\mu^{(t+1)} = \gamma_1\mu^{(t)}\}$ with $\gamma_1 \in (0, 1)$, which is a tunable parameter [47]. If the set \mathcal{K}_1 is finite, then according to Lines 13-16 in Algorithm 1 there exists an integer \hat{t} , such that, for all $t > \hat{t}$

$$\left\| \mathbf{d}_{\Gamma(t)} \right\|_2 \geq \gamma \mu^{(\hat{t})}, \quad (48)$$

with $\gamma \in (0, 1)$. Taking $\dot{\mu} = \mu^{(\hat{t})}$, the optimization problem (18) reduces to

$$\min_{\mathbf{z} \in \mathbb{C}^N} h(\mathbf{z}, \dot{\mu}). \quad (49)$$

Now, considering the properties stated in Lemma 3, from [46, Theorem 2.1] we get

$$\lim_{t \rightarrow \infty} \left\| \frac{\partial h(\mathbf{x}^{(t)}, \mu^{(t)})}{\partial \bar{\mathbf{z}}} \right\|_2 = \lim_{t \rightarrow \infty} \left\| \mathbb{E}_{\Gamma(t)} [\mathbf{d}_{\Gamma(t)}] \right\|_2 = 0. \quad (50)$$

It can be readily seen that (50) contradicts the assumption $\left\| \mathbf{d}_{\Gamma(t)} \right\|_2 \geq \gamma \mu^{(\hat{t})}$, for all $t > \hat{t}$. This shows that \mathcal{K}_1 must be infinite and $\lim_{t \rightarrow \infty} \mu^{(t)} = 0$.

Given that \mathcal{K}_1 is infinite, we deduce that

$$\begin{aligned} \lim_{t \rightarrow \infty} \left\| \frac{\partial h(\mathbf{x}^{(t)}, \mu^{(t)})}{\partial \bar{\mathbf{z}}} \right\|_2 &= \lim_{t \rightarrow \infty} \left\| \mathbb{E}_{\Gamma(t)} [\mathbf{d}_{\Gamma(t)}] \right\|_2 \\ &\leq \lim_{t \rightarrow \infty} \mathbb{E}_{\Gamma(t)} \left[\left\| \mathbf{d}_{\Gamma(t)} \right\|_2 \right] \leq \gamma \lim_{t \rightarrow \infty} \mu^{(t)} = 0, \end{aligned} \quad (51)$$

where the second line follows from the Jensen inequality. Therefore, from (51) the result of Theorem 1 holds.

VI. PROOF OF LEMMA 3

The proof of Lemma 3 is obtained by individually proving the following four requirements.

1) Following from the definition of $h(\mathbf{z}, \mu)$ in (18) it is clear that $h(\mathbf{z}, \mu) \geq 0$ and thus bounded below.

2) This holds by definition.

3) From (24) it follows that the ℓ -th entry of $\frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}}$ is given by

$$\begin{aligned} \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}}[\ell] &= \frac{1}{N^2} \sum_{k,p=0}^{N-1} \left(\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}) - v_{k,p} \right) \mathbf{z}[\ell - p] e^{2\pi i \ell k / N} \\ &\quad + \frac{1}{N^2} \sum_{k,p=0}^{N-1} \left(\mathbf{f}_k^T \overline{\mathbf{g}_p} - v_{k,p} \right) \mathbf{z}[\ell + p] e^{-2\pi i (\ell + p) k / N}, \end{aligned} \quad (52)$$

where $v_{k,p} = \sqrt{\mathbf{A}[p, k]} \frac{\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})}{\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})|)}$, and

$$\mathbf{g}_p := \left[\mathbf{z}[0] \overline{\mathbf{z}[p]}, \dots, \mathbf{z}[N-1] \overline{\mathbf{z}[N-1+p]} \right]^T. \quad (53)$$

Let $\mathbf{D}_p(\mathbf{z})$ be a diagonal matrix composed of the entries of $\mathbf{z}_p[n] = \mathbf{z}[n-p]$. Thus,

$$\frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} = \frac{1}{N^2} \sum_{p,k=0}^{N-1} f_{k,p}(\mathbf{z}) + g_{k,p}(\mathbf{z}), \quad (54)$$

where

$$\begin{aligned} f_{k,p}(\mathbf{z}) &= \rho_{k,p}(\mathbf{z}) \mathbf{D}_p(\mathbf{z}) \mathbf{f}_k, \\ g_{k,p}(\mathbf{z}) &= \omega^{-kp} \overline{\rho_{k,p}(\mathbf{z})} \mathbf{D}_{-p}(\mathbf{z}) \mathbf{f}_k, \end{aligned} \quad (55)$$

and

$$\rho_{k,p}(\mathbf{z}) = \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}) - \sqrt{\mathbf{A}[p, k]} \frac{\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})}{\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})|)}. \quad (56)$$

To prove 3) we establish that any $f_{k,p}(\mathbf{z})$ and $g_{k,p}(\mathbf{z})$ satisfy

$$\|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \leq r_{k,p} \|\mathbf{z}_1 - \mathbf{z}_2\|_2, \quad (57)$$

and

$$\|g_{k,p}(\mathbf{z}_1) - g_{k,p}(\mathbf{z}_2)\|_2 \leq s_{k,p} \|\mathbf{z}_1 - \mathbf{z}_2\|_2, \quad (58)$$

for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{J}$ and some constants $r_{k,p}, s_{k,p} > 0$. In fact, once we prove (57), it can be performed a similar analysis for $g_{k,p}(\mathbf{z})$, and thus the result of this third part holds.

From the definition of $f_{k,p}(\mathbf{z})$, for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{J}$ we have that

$$\frac{1}{\sqrt{N}} \|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \leq \|\rho_{k,p}(\mathbf{z}_1) \bar{\mathbf{z}}_1 - \rho_{k,p}(\mathbf{z}_2) \bar{\mathbf{z}}_2\|_2, \quad (59)$$

considering that $\mathbf{D}_p(\mathbf{z}_1)$ and $\mathbf{D}_p(\mathbf{z}_2)$ are diagonal matrices, and $\|\mathbf{f}_k\|_2 = \sqrt{N}$. Observe that from (56) and (59) it can be obtained that

$$\begin{aligned} &\frac{1}{\sqrt{N}} \|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \\ &\leq \frac{|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)|}{\mu} \left(\varphi_\mu \left(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)| \right) + \sqrt{\mathbf{A}[p, k]} \right) \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\quad + \|\mathbf{z}_2\|_2 \underbrace{|\rho_{k,p}(\mathbf{z}_1) - \rho_{k,p}(\mathbf{z}_2)|}_{p_1}, \end{aligned} \quad (60)$$

where the second inequality comes from the fact that $\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)|) \geq \mu$. The term p_1 in (60) can be upper bounded as

$$\begin{aligned} p_1 &\leq \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \sqrt{\mathbf{A}[p, k]} \left| \frac{\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)}{\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)|)} - \frac{\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)}{\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)|)} \right| \\ &\leq \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \frac{\sqrt{\mathbf{A}[p, k]}}{\mu^2} \varphi_\mu \left(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| \right) \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \frac{\sqrt{\mathbf{A}[p, k]}}{\mu^2} \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \left| \varphi_\mu \left(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)| \right) - \varphi_\mu \left(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| \right) \right|. \end{aligned} \quad (61)$$

Recall that, by Heine-Borel Theorem, \mathcal{J} is a closed bounded set, and thus compact. Since $\varphi_\mu(\cdot)$ is a continuous function, there exists a constant M_{φ_μ} such that $\varphi_\mu(|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})|) \leq M_{\varphi_\mu}$ for all $\mathbf{z} \in \mathcal{J}$. Also, from Lemma 2 in [32] we have that $\varphi_\mu(\cdot)$ is a 1-Lipschitz function. Combining this with (61) we get

$$\begin{aligned} p_1 &\leq \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \frac{\sqrt{\mathbf{A}[p, k] M_{\varphi_\mu}}}{\mu^2} \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \frac{\sqrt{\mathbf{A}[p, k]}}{\mu^2} \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right|, \end{aligned} \quad (62)$$

and thus

$$\begin{aligned} p_1 &\leq \left(\frac{\sqrt{\mathbf{A}[p, k] M_{\varphi_\mu}}}{\mu^2} + 1 \right) \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \\ &\quad + \frac{\sqrt{\mathbf{A}[p, k]}}{\mu^2} \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right| \left| \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2) \right|, \end{aligned} \quad (63)$$

where (63) results from applying the triangular inequality. Putting together (60) and (63) we obtain that

$$\begin{aligned} & \frac{1}{\sqrt{N}} \|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \\ & \leq \frac{|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1)|}{\mu} \left(M_{\varphi_\mu} + \sqrt{\mathbf{A}[p, k]} \right) \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ & + \|\mathbf{z}_2\|_2 \left(\frac{\sqrt{\mathbf{A}[p, k]} M_{\varphi_\mu}}{\mu^2} + 1 \right) |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| \\ & + \frac{\|\mathbf{z}_2\|_2 \sqrt{\mathbf{A}[p, k]}}{\mu^2} |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)|. \end{aligned} \quad (64)$$

Observe that the upper bound in (64) directly depends on a term of the form $\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})$ for some $\mathbf{z} \in \mathcal{J}$, which might be zero. However, Lemma 4 proves that $|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})| > 0$ or equivalently $\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}) \neq 0$, for almost all $\mathbf{z} \in \mathcal{J}$.

Lemma 4. Let $\mathbf{z} \in \mathcal{J}$ where \mathcal{J} as defined in (45). Then, for almost all $\mathbf{z} \in \mathcal{J}$ the following holds

$$|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})| > 0, \quad (65)$$

for all $k, p \in \{0, \dots, N-1\}$, with $\mathbf{g}_p(\mathbf{z})$ as in (53).

Proof. We prove this lemma by contradiction. Suppose that $|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})| = 0$. Then, from (7) we have that

$$\begin{aligned} |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})|^2 &= \left| \sum_{n=0}^{N-1} \mathbf{z}[n] \overline{\mathbf{z}[n-p]} e^{-2\pi i n k / N} \right|^2 \\ &= \sum_{n,m=0}^{N-1} \left(\mathbf{z}[n] \overline{\mathbf{z}[n-p]} \mathbf{z}[m-p] \overline{\mathbf{z}[m]} \right) e^{\frac{2\pi i (m-n)k}{N}} = 0. \end{aligned} \quad (66)$$

Observe that (66) is a quartic polynomial equation with respect to the entries of \mathbf{z} . However, for almost all signals $\mathbf{z} \in \mathcal{J}$ the left hand side of (66) will not be equal to zero which leads to a contradiction [23]. \blacksquare

Then, proceeding to bound the term $|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})|$, notice that from (7) we have that

$$\begin{aligned} |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z})| &= \left| \sum_{n=0}^{N-1} \mathbf{z}[n] \overline{\mathbf{z}[n-p]} e^{-2\pi i n k / N} \right| \\ &\leq \sum_{n=0}^{N-1} |\mathbf{z}[n] \overline{\mathbf{z}[n-p]}| \leq N \|\mathbf{z}\|_2, \end{aligned} \quad (67)$$

in which the second inequality arises from $\|\mathbf{z}\|_2 \leq \sqrt{N} \|\mathbf{z}\|_\infty$ and $\|\mathbf{z}\|_1 \leq \sqrt{N} \|\mathbf{z}\|_2$. Combining (64) and (67) we get

$$\begin{aligned} & \frac{1}{\sqrt{N}} \|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \\ & \leq \frac{N \|\mathbf{z}_1\|_2}{\mu} \left(M_{\varphi_\mu} + \sqrt{\mathbf{A}[p, k]} \right) \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ & + \|\mathbf{z}_2\|_2 \left(\frac{\sqrt{\mathbf{A}[p, k]} M_{\varphi_\mu}}{\mu^2} + 1 \right) |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| \\ & + \frac{N \|\mathbf{z}_2\|_2^2 \sqrt{\mathbf{A}[p, k]}}{\mu^2} |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)|. \end{aligned} \quad (68)$$

Now, we have to analyze the term $|\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)|$ in (68). Specifically, from (7) it can be obtained that

$$\begin{aligned} & |\mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_1) - \mathbf{f}_k^H \mathbf{g}_p(\mathbf{z}_2)| \\ & \leq \sum_{n=0}^{N-1} |\mathbf{z}_1[n] \overline{\mathbf{z}_1[n-p]} - \mathbf{z}_2[n] \overline{\mathbf{z}_2[n-p]}| \\ & \leq N (\|\mathbf{z}_1\|_2 + \|\mathbf{z}_2\|_2) \|\mathbf{z}_1 - \mathbf{z}_2\|_2, \end{aligned} \quad (69)$$

where the second inequality results from $\|\mathbf{z}\|_2 \leq \sqrt{N} \|\mathbf{z}\|_\infty$ and $\|\mathbf{z}\|_1 \leq \sqrt{N} \|\mathbf{z}\|_2$. Combining (68) and (69) we obtain that

$$\|f_{k,p}(\mathbf{z}_1) - f_{k,p}(\mathbf{z}_2)\|_2 \leq r_{k,p} \|\mathbf{z}_1 - \mathbf{z}_2\|_2, \quad (70)$$

where $r_{k,p}$ is given by

$$\begin{aligned} r_{k,p} &= \frac{N \sqrt{N} \|\mathbf{z}_1\|_2}{\mu} \left(M_{\varphi_\mu} + \sqrt{\mathbf{A}[p, k]} \right) \\ &+ N \sqrt{N} (\|\mathbf{z}_1\|_2 + \|\mathbf{z}_2\|_2) \|\mathbf{z}_2\|_2 \left(\frac{\sqrt{\mathbf{A}[p, k]} M_{\varphi_\mu}}{\mu^2} + 1 \right) \\ &+ N^2 \sqrt{N} (\|\mathbf{z}_1\|_2 + \|\mathbf{z}_2\|_2) \frac{\|\mathbf{z}_2\|_2^2 \sqrt{\mathbf{A}[p, k]}}{\mu^2}. \end{aligned} \quad (71)$$

Since the set \mathcal{J} is bounded, then $\|\mathbf{z}\|_2 < \infty$ for all $\mathbf{z} \in \mathcal{J}$. Therefore, $0 < r_{k,p} < \infty$, and from (70) the result holds.

4) We proceed to prove (47). Observe that

$$\begin{aligned} & \mathbb{E}_{\Gamma(t)} \left[\left\| \mathbf{d}_{\Gamma(t)} - \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2^2 \right] \\ & \leq \mathbb{E}_{\Gamma(t)} \left[2 \left\| \mathbf{d}_{\Gamma(t)} \right\|_2^2 \right] + 2 \left\| \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2^2, \end{aligned} \quad (72)$$

in which the inequality comes from the fact that $\|\mathbf{w}_1 + \mathbf{w}_2\|_2^2 \leq 2 (\|\mathbf{w}_1\|_2^2 + \|\mathbf{w}_2\|_2^2)$ for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{C}^N$. Combining (46) and (72) we have that

$$\mathbb{E}_{\Gamma(t)} \left[\left\| \mathbf{d}_{\Gamma(t)} - \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2^2 \right] \leq \mathbb{E}_{\Gamma(t)} \left[2 \left\| \mathbf{d}_{\Gamma(t)} \right\|_2^2 \right] + 2U \|\mathbf{z}\|_2^2, \quad (73)$$

for some $U > 0$. Recall that $\Gamma(t)$ is sampled uniformly at random from all subsets of $\{1, \dots, N\} \times \{1, \dots, R\}$ with cardinality Q . From the definition of $\mathbf{d}_{\Gamma(t)}$ in Line 9 of Algorithm 1, it can be concluded that

$$\begin{aligned} \mathbb{E}_{\Gamma(t)} \left[2 \left\| \mathbf{d}_{\Gamma(t)} \right\|_2^2 \right] &\leq \frac{4Q}{N^2} \sum_{p,k=0}^{N-1} \|f_{k,p}(\mathbf{z}) + g_{k,p}(\mathbf{z})\|_2^2 \\ &\leq \frac{8Q}{N^2} \sum_{p,k=0}^{N-1} \|f_{k,p}(\mathbf{z})\|_2^2 + \|g_{k,p}(\mathbf{z})\|_2^2, \end{aligned} \quad (74)$$

using the fact that $\|\mathbf{w}_1 + \mathbf{w}_2\|_2^2 \leq 2 (\|\mathbf{w}_1\|_2^2 + \|\mathbf{w}_2\|_2^2)$ for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{C}^N$. Furthermore, since $f_{k,p}(\mathbf{z})$ and $g_{k,p}(\mathbf{z})$ satisfy (57) and (58), respectively, we conclude that

$$\mathbb{E}_{\Gamma(t)} \left[2 \left\| \mathbf{d}_{\Gamma(t)} \right\|_2^2 \right] \leq \frac{8Q \|\mathbf{z}\|_2^2}{N^2} \sum_{p,k=0}^{N-1} r_{k,p}^2 + s_{k,p}^2, \quad (75)$$

for some constants $r_{k,p}, s_{k,p} > 0$. Thus, combining (73) and (75) we obtain that

$$\mathbb{E}_{\Gamma(t)} \left[\left\| \mathbf{d}_{\Gamma(t)} - \frac{\partial h(\mathbf{z}, \mu)}{\partial \bar{\mathbf{z}}} \right\|_2^2 \right] \leq \zeta^2, \quad (76)$$

where ζ is defined as

$$\zeta = \|\mathbf{z}\|_2 \sqrt{\frac{8Q}{N^2} \sum_{p,k=0}^{N-1} r_{k,p}^2 + s_{k,p}^2 + 2U}. \quad (77)$$

Notice $\zeta < \infty$ because the set \mathcal{J} is bounded. Thus, from (76) the result holds. \blacksquare

VII. PROOF OF THEOREM 2

Proof. Notice that the function as defined in (40) given by

$$\mathbf{p}_\ell \rightarrow \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{p}_\ell\|_2^2 + \frac{1}{2\lambda_{(t)}} \|\mathbf{p}_\ell - \mathbf{x}_\ell^{(t-1)}\|_2^2$$

is $\frac{1}{\lambda_{(t)}}$ -strongly convex in \mathbf{p}_ℓ . Here we allow inaccuracy in the solutions. More precisely, if we define $\mathbf{x}_{\ell,*}^{(t)}$ to be the exact minimizer of (40) the standard optimality conditions for strongly convex minimization imply that there exists a sequence of additive accuracy parameters $\epsilon_{(t)} \geq 0$ such that the numerically attained solution $\mathbf{x}_\ell^{(t)}$ satisfies

$$\begin{aligned} & \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell^{(t)}\|_2^2 + \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell^{(t-1)}\|_2^2 \\ & \leq \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_{\ell,*}^{(t)}\|_2^2 + \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_{\ell,*}^{(t)} - \mathbf{x}_\ell^{(t-1)}\|_2^2 + \epsilon_{(t)} \\ & \leq \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 + \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell - \mathbf{x}_\ell^{(t-1)}\|_2^2 + \epsilon_{(t)}, \end{aligned} \quad (78)$$

where \mathbf{x}_ℓ is the unknown desired signal and $\frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 \leq \epsilon_{(t)}$ [48]. Then from (78) it can be obtained that

$$\begin{aligned} \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell^{(t)}\|_2^2 & \leq \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 \\ & + \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell - \mathbf{x}_\ell^{(t-1)}\|_2^2 + \epsilon_{(t)}. \end{aligned} \quad (79)$$

Recall that the radar phase retrieval problem has several ambiguities. In fact, the feasible set for each ℓ , denoted by \mathcal{A}_ℓ , is given by

$$\mathcal{A}_\ell = \{e^{j\beta} \mathbf{x}_\ell\} \cup \{\mathbf{x}_\ell[n-r]\} \cup \{\mathbf{x}_\ell[-n]\} \cup \{e^{ibn} \mathbf{x}_\ell[n]\}, \quad (80)$$

for all $r = 0, \dots, N-1$, and $b, \beta \in \mathbb{R}$. Thus, performing a parallel derivation with $\mathbf{q}_\ell^{(t-1)} \in \mathcal{A}_\ell$ replacing $\mathbf{x}_\ell^{(t-1)}$ and simplifying, we have from (78) and (79) that

$$\begin{aligned} \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell^{(t)}\|_2^2 & \leq \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 \\ & + \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell + \mathbf{q}_\ell^{(t-1)}\|_2^2 + \epsilon_{(t)}. \end{aligned} \quad (81)$$

Notice that the term at the left side of the inequality in (81) can be rewritten as

$$\begin{aligned} & \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell^{(t)}\|_2^2 \\ & = \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell^{(t)} - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell + \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 \\ & = \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 + \|\mathbf{G}_\ell^{(t-1)} \mathbf{z}_\ell^{(t)} - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 \\ & - 2\Re \left(\left(\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell \right)^H \left(\mathbf{G}_\ell^{(t-1)} \mathbf{z}_\ell^{(t)} - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell \right) \right) \\ & \geq \|\mathbf{G}_\ell^{(t-1)} \mathbf{z}_\ell^{(t)} - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 \\ & - 2 \left| \left(\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell \right)^H \left(\mathbf{G}_\ell^{(t-1)} \mathbf{z}_\ell^{(t)} - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell \right) \right| \\ & \geq \rho_1^{(t)} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 - 2\rho_2^{(t)} \|\mathbf{x}_\ell\|_2 \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2, \end{aligned} \quad (82)$$

where $\rho_1^{(t)}$ is the squared smallest singular value of $\mathbf{G}_\ell^{(t)}$ greater than zero, $\rho_2^{(t)} = \sigma_{\max}(\mathbf{G}_\ell - \mathbf{G}_\ell^{(t-1)}) \sigma_{\max}(\mathbf{G}_\ell^{(t-1)})$ with $\sigma_{\max}(\cdot)$ representing the largest singular value, and $\Re(\cdot)$ is the real part function. Then, combining (81) and (82) we have that

$$\begin{aligned} \rho_1^{(t)} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 & \leq \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell + \mathbf{q}_\ell^{(t-1)}\|_2^2 \\ & + 2\rho_2^{(t)} \|\mathbf{x}_\ell\|_2 \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2 \\ & + \|\mathbf{y}_\ell - \mathbf{G}_\ell^{(t-1)} \mathbf{x}_\ell\|_2^2 + \epsilon_{(t)} \end{aligned} \quad (83)$$

Considering the construction of matrix $\mathbf{G}_\ell^{(t-1)}$ in (16), the term $\|\mathbf{G}_\ell - \mathbf{G}_\ell^{(t-1)}\|_F^2$ in (83) can be bounded as

$$\|\mathbf{G}_\ell - \mathbf{G}_\ell^{(t-1)}\|_F^2 \leq 2N\lambda_{(t)}\epsilon_{(t-1)}, \quad (84)$$

because $\frac{1}{2\lambda_{(t-1)}} \|\mathbf{x}_\ell^{(t-1)} - \mathbf{x}_\ell\|_2 \leq \epsilon_{(t-1)}$. Thus, combining (83) and (84) we have that

$$\begin{aligned} \rho_1^{(t)} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 & \leq \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell + \mathbf{q}_\ell^{(t-1)}\|_2^2 \\ & + 2\rho_2^{(t)} \|\mathbf{x}_\ell\|_2 \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2 \\ & + 2N\lambda_{(t)}\epsilon_{(t-1)} \|\mathbf{x}_\ell\|_2^2 + \epsilon_{(t)}. \end{aligned} \quad (85)$$

It is worth to mention that inequality (85) is valid for all the trivial ambiguities of \mathbf{x}_ℓ for the radar problem. In particular, if we take the minimum over the ambiguities of \mathbf{x}_ℓ in both sides of (85), we find that

$$\begin{aligned} \rho_1^{(t)} \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 & \leq \frac{1}{2\lambda_{(t)}} \|\mathbf{x}_\ell^{(t-1)} - \mathbf{x}_\ell\|_2^2 \\ & + 2N\lambda_{(t)}\epsilon_{(t-1)} \|\mathbf{x}_\ell\|_2^2 + \epsilon_{(t)} \\ \|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2^2 & \leq \frac{1}{2\lambda_{(t)}\rho_1^{(t)}} \|\mathbf{x}_\ell^{(t-1)} - \mathbf{x}_\ell\|_2^2 \\ & + 2N\lambda_{(t)}\epsilon_{(t-1)} \|\mathbf{x}_\ell\|_2^2 + \epsilon_{(t)}. \end{aligned} \quad (86)$$

Notice that if we inductively apply inequality (86) when $\epsilon_{(t)} = 0$ for all t yields to

$$\|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2 \leq \frac{1}{\sqrt{2\lambda_{(t)}\rho_1^{(t)}}} \|\mathbf{x}_\ell^{(0)} - \mathbf{x}_\ell\|_2, \quad (87)$$

in which if we choose $\lambda_{(t)}$ such that $\lambda_{(t)}\rho_1^{(t)} > 1/2$ a reduction of the error in the estimation of \mathbf{x}_ℓ is guaranteed. Then, from (87) it can be concluded that

$$\|\mathbf{x}_\ell^{(t)} - \mathbf{x}_\ell\|_2 < \tau \|\mathbf{x}_\ell^{(0)} - \mathbf{x}_\ell\|_2, \quad (88)$$

where τ is given by $\tau = \prod_t \frac{1}{\sqrt{2\lambda_{(t)}\rho_1^{(t)}}} < 1$. Finally, since $\text{diag}(\mathbf{x}\mathbf{x}^H, \ell) = \mathbf{x}_\ell$, then we have from (88) that

$$\|\mathbf{x}^{(0)}(\mathbf{x}^{(0)})^H - \mathbf{x}\mathbf{x}^H\|_F < \tau \|\mathbf{x}_{\text{init}}\mathbf{x}_{\text{init}}^H - \mathbf{x}\mathbf{x}^H\|_F, \quad (89)$$

where $\mathbf{x}^{(0)}$ is the returned vector of Algorithm 2, and \mathbf{x}_{init} as in (33). Thus, the result holds. \blacksquare

VIII. NUMERICAL RESULTS

This section evaluates the numerical performance of the proposed method. We used the following parameters for Algorithm 1: $\gamma_1 = 0.1$, $\gamma = 0.1$, $\alpha = 0.6$, $\mu_0 = 65$, and $\epsilon = 1 \times 10^{-10}$. The number of indices that are chosen uniformly at random is fixed as $Q = N$.

The signals used in the simulations were constructed as follows. For all tests, we built a set of $\lceil \frac{N-1}{2} \rceil$ -band-limited and time-limited signals that conform to a Gaussian power spectrum centered at 800 nm. Specifically, each signal ($N = 128$ grid points) is produced via the Fourier transform of a complex vector with a Gaussian-shaped amplitude with a cutoff frequency of 150 microseconds⁻¹ (usec⁻¹). Next, we multiply the obtained power spectrum by a uniformly distributed random phase. In the experiments we used the inverse Fourier of this signal as the underlying signal.¹

The tests are divided in three sections to study the performance of Algorithm 1 for complete and incomplete AF, and additional type of time, band limited signals under noisy and noiseless scenarios at different values of signal-to-noise-ratio (SNR), defined as $\text{SNR} = 10 \log_{10}(\|\mathbf{A}\|_F^2 / \|\boldsymbol{\sigma}\|_2^2)$, where $\boldsymbol{\sigma}$ is the variance of the noise. The radar function is incomplete when only few shifts or Fourier frequencies are considered. In the first section we examine the ability

¹All simulations were implemented in Matlab R2019a on an Intel Core i7 3.4GHz CPU with 32 GB RAM.

of Algorithm 1 to recover the signal from complete data. The second section assesses the performance of Algorithm 1 to recover the underlying signal when the AF is incomplete. The last test studies the ability of the proposed method to estimate different type of radar signals than the described above.

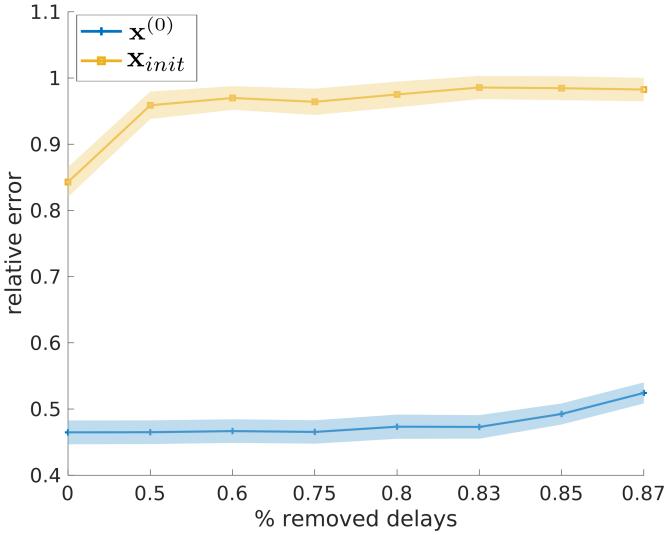


Fig. 2. Relative error comparison between the initial vector x_{init} as defined in (33), and the returned initial guess $x^{(0)}$ for different percentages of removed delays (uniformly) in the absence of noise. An average of the relative error was computed among 100 trials.

A. Relative Error of the Initialization Procedure

This section examines the impact of the designed initialization described in Algorithm 2, under noisy and noiseless scenarios. We compare the relative error between the starting vector in (33), and the returned solution $x^{(0)}$ of the proposed initialization procedure. The number of iterations to attain the vector $x^{(0)}$ using the designed initialization was fixed as $T = 2$, and we numerically determine the relative error averaged over 100 trials. These numerical results are summarized in Figs. 2, and 3 indicate that the proposed initialization algorithm outperforms x_{init} .

B. Signal Reconstruction from Complete Data

The performance of Algorithm 1 is presented to recover time and band-limited signals under noiseless and noisy scenarios using the complete radar ambiguity function. The results are presented in Figs. 6, 7, where the attained relative errors by the proposed algorithm are included. For the second scenario, the radar ambiguity function trace is corrupted by Gaussian noise with SNR = 20dB. Specifically, in the noisy case we are assuming that ambiguity function is not perfectly designed which allows to evaluate the robustness of Algorithm 1. The results in Figs. 6 and 7 suggest that the proposed method is able to estimate the signals.

C. Signal Reconstruction from Incomplete Data

The success rate of Algorithm 1 is evaluated when the ambiguity function is incomplete. To this end, Algorithm 1 is initialized at $\mathbf{x}^{(0)} = \mathbf{x} + \delta\zeta$, where δ is a fixed constant and ζ takes values on $\{-1, 1\}$ with equal probability, while a percentage of the delays are set to zero. A trial is declared successful when the returned estimate attains a relative error as in (17) that is smaller than 10^{-6} . We

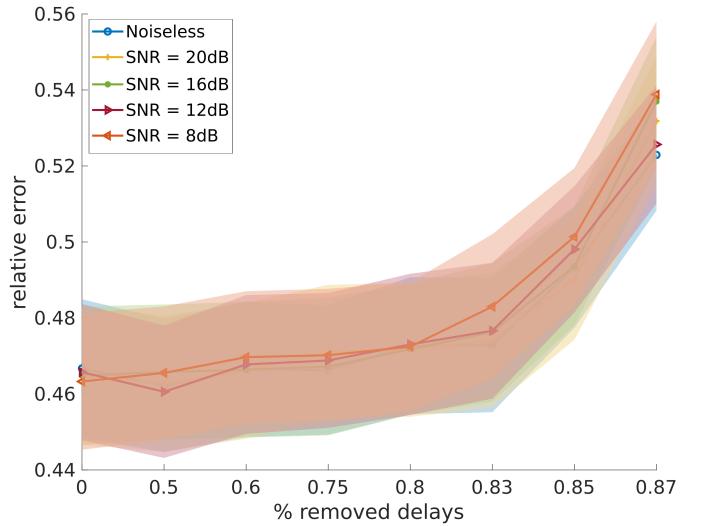


Fig. 3. Performance of the proposed initialization described in Algorithm 2 at different SNR levels, for different percentages of removed delays (uniformly). The relative error was averaged over 100 trials.

numerically determine the empirical success rate among 100 trials. Fig. 4 summarizes these results, and shows that Algorithm 1 is able to estimate the pulse when the AF is incomplete.

In Fig. 4 the % of removed delays are performed uniformly, which means for instance in the case of 50% every two delays, starting from the first one, are preserved. Additionally, in the case of 75% means every three delays, starting from the first one, are preserved. From these results it is numerically validate Proposition 1 (in consequence Corollary 1) that not all the delays are need to estimate the underlying signal. To illustrate this, Figs. 6 and 8 show the estimated time and band-limited signals from noisy incomplete AF (50% and 75% of the delays are removed respectively). Observe that Algorithm 1 is able to return a close estimation of the signal even when the incomplete AF is assumed imperfectly designed, suggesting the effectiveness of Algorithm 1.

To complement the results in Figs. 4, 7 and 8, here is also presented the performance of Algorithm 1 when a percentage of the delays

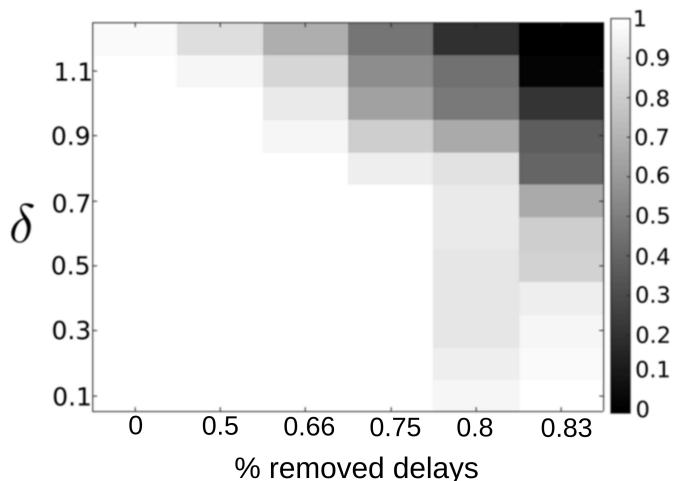


Fig. 4. Empirical success rate of Algorithm 1 as a function of % removed delays (uniformly) and δ in the absence of noise.

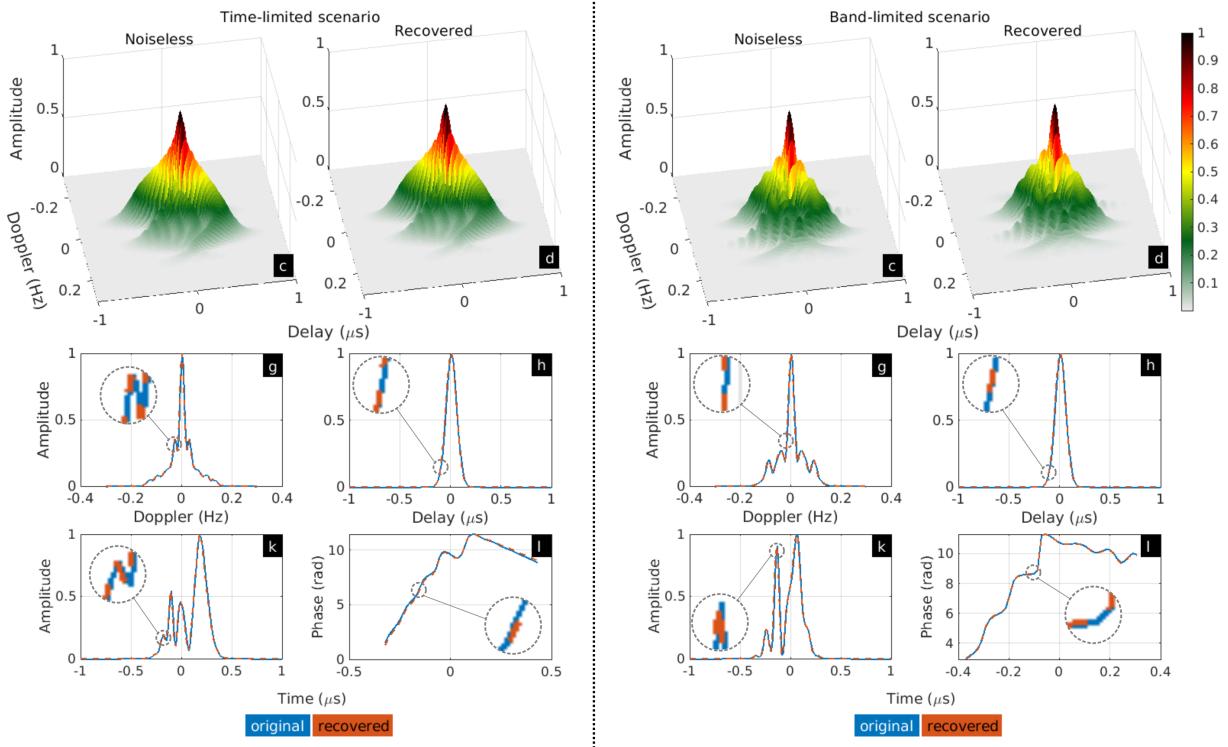


Fig. 5. Reconstructed time and band-limited signals with their ambiguity functions in the absent of noise. The attained relative error as in (17) was 1×10^{-6} for both signals. (a), (c) and (b), (d) are the original and recovered ambiguity functions, respectively. (e), (g), and (f), (h) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (i), (k) and (j), (l) correspond to the recovered magnitude and phase of the estimated signals, respectively.

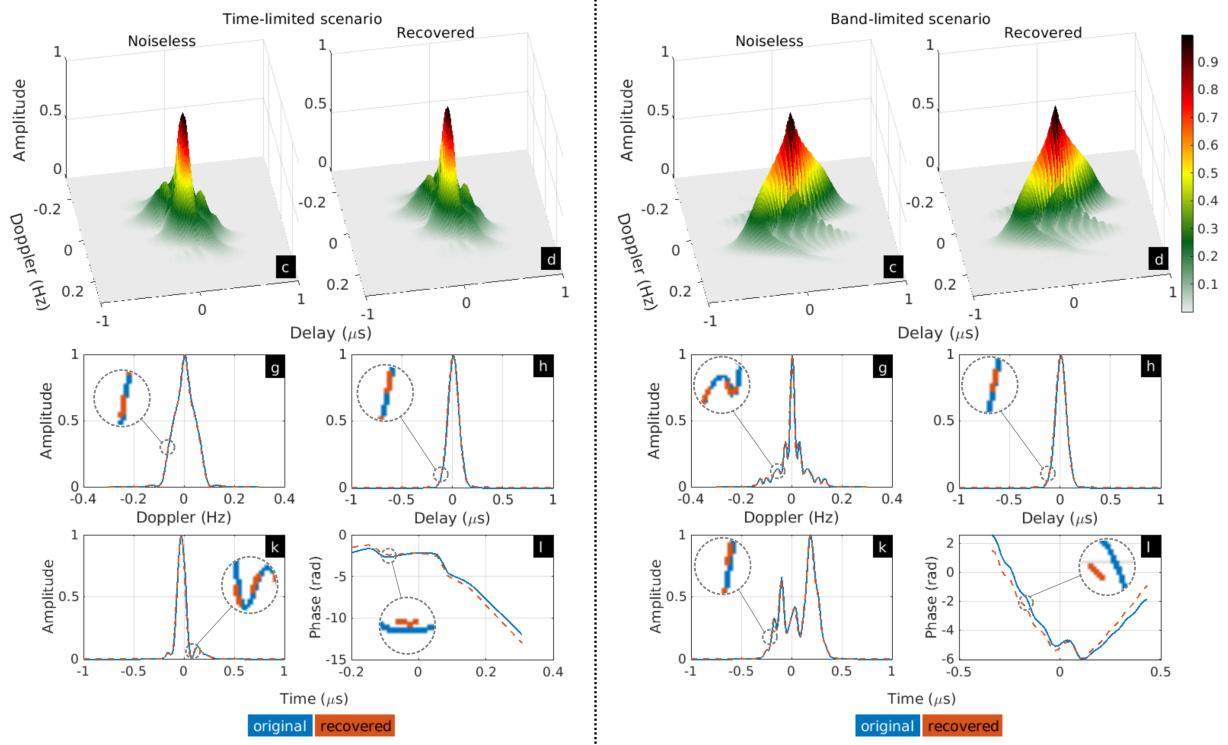


Fig. 6. Reconstructed time and band-limited signals with their ambiguity functions in the present of noise with $\text{SNR} = 20\text{dB}$. The attained relative error as in (17) was 5×10^{-2} for both signals. (a), (c) and (b), (d) are the noiseless (ideal) and recovered ambiguity functions, respectively. (e), (g), and (f), (h) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (i), (k) and (j), (l) correspond to the recovered magnitude and phase of the estimated signals, respectively.

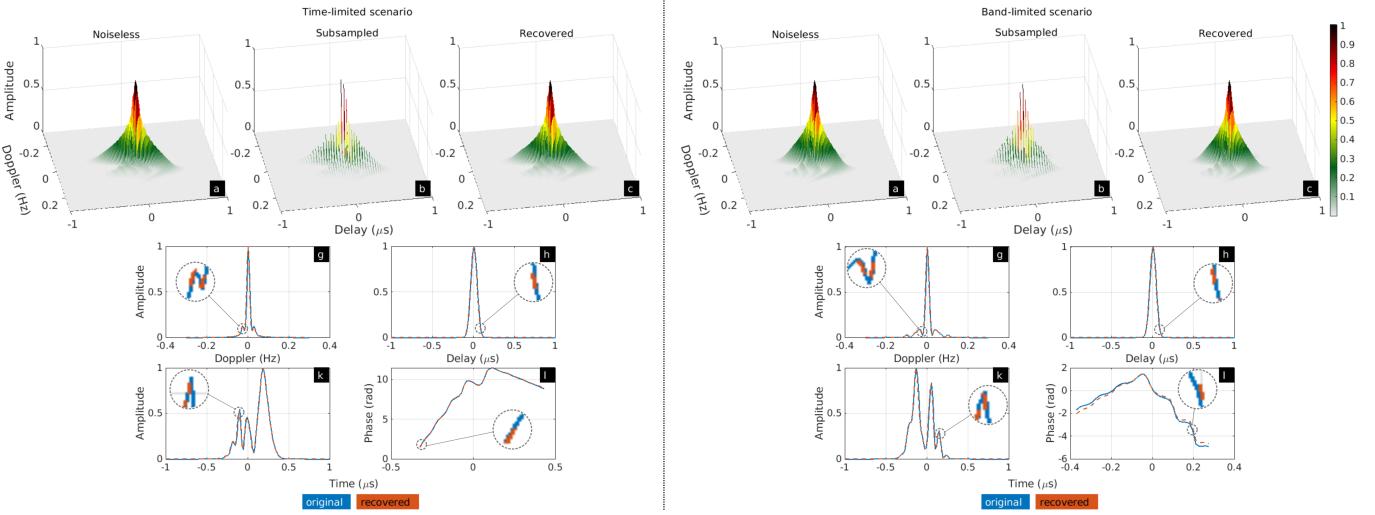


Fig. 7. Reconstructed time and band-limited signals when a 50% of the delays of their ambiguity functions are uniformly removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 5×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (l), (n) correspond to the recovered magnitude and phase of the estimated signals.

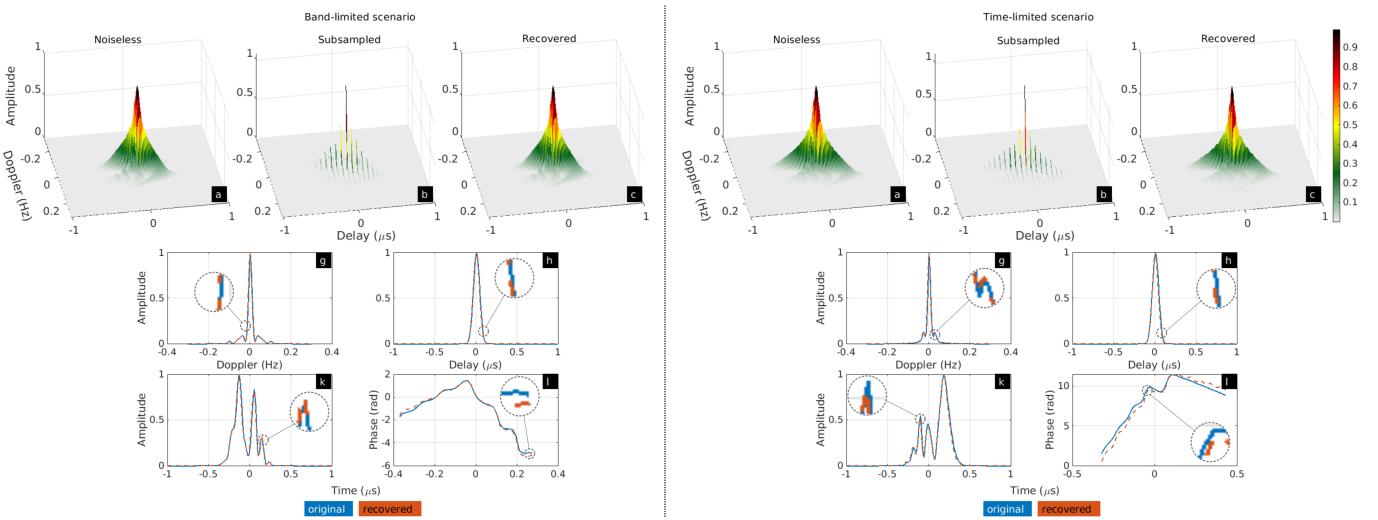


Fig. 8. Reconstructed time and band-limited signals when a 75% of the delays of their ambiguity functions are uniformly removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 5×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (l), (n) correspond to the recovered magnitude and phase of the estimated signals.

and the Fourier frequencies of the AF are non-uniformly removed, illustrated in Figs. 9 and 10. Specifically, 28% of the first and last delays/ frequencies of the AF were set to zero in Fig. 9/10, respectively. These results suggest that a non-uniform selection of the delays to be removed reduces the ability of Algorithm 1 to estimate the analyzed pulse compared with a uniform strategy. In contrast, in the case of a non-uniform modality to remove frequencies it can be concluded that the performance of Algorithm 1 is close to the uniform selection of the delays to be removed.

D. Additional Type of Signals

In this section we investigate the performance of Algorithm 1 to estimate Linear/Non-linear Frequency Modulated (LFM/NLFM)

pulses from its incomplete noisy ambiguity function. These kind of signals are modeled as

$$\mathbf{x}[n] = \mathbf{a}[n] e^{j\pi\varphi[n]}, \quad (90)$$

where $\varphi[n]$ is given by

$$\varphi[n] = \pi k(\Delta tn)^2, \quad (\text{LFM})$$

$$\varphi[n] = \pi kt^2 + \sum_{l=1}^L \alpha_l \cos(2\pi l \Delta tn / T) \quad (\text{NLFM}), \quad (91)$$

with T as the duration of the pulse, Δt as the sampling size in time, $k = \frac{\Delta f}{T}$ such that Δf is the swept bandwidth, and $L > 0$ is an integer. The values for α_l are given by $\alpha_l = \frac{0.4T}{l}$. For this experiment $\Delta f = 128 \times 10^3$, and $\Delta t = 0.4 \times 10^{-6}$. The values of

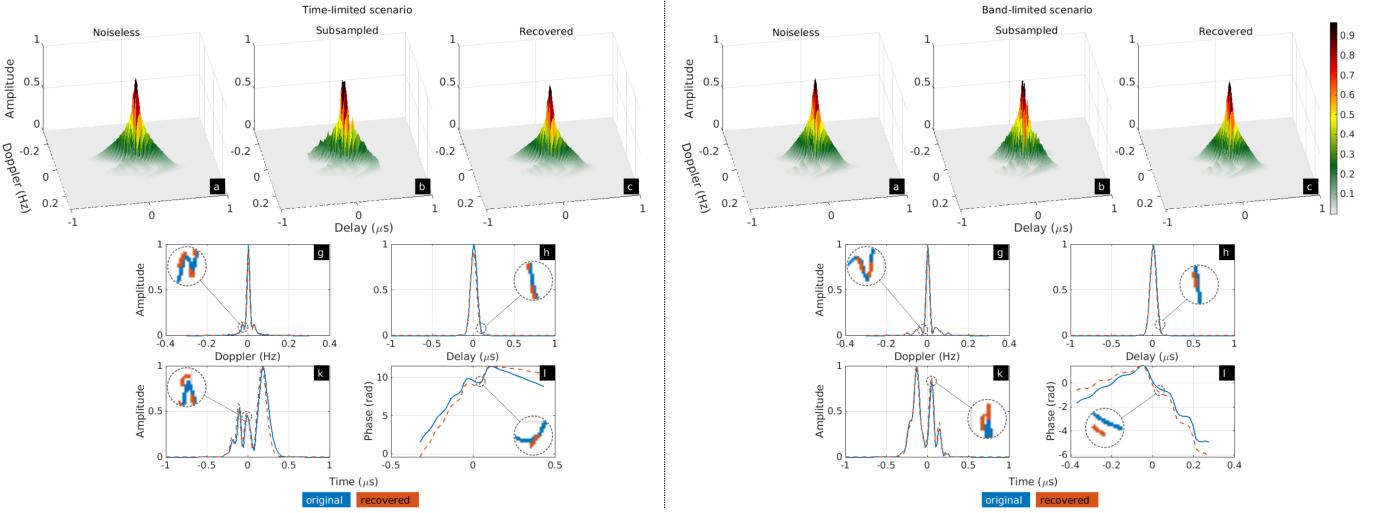


Fig. 9. Reconstructed time and band-limited signals when a 57% of the delays of their ambiguity functions are non-uniformly removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 9×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (l), (n) correspond to the recovered magnitude and phase of the estimated signals, respectively.

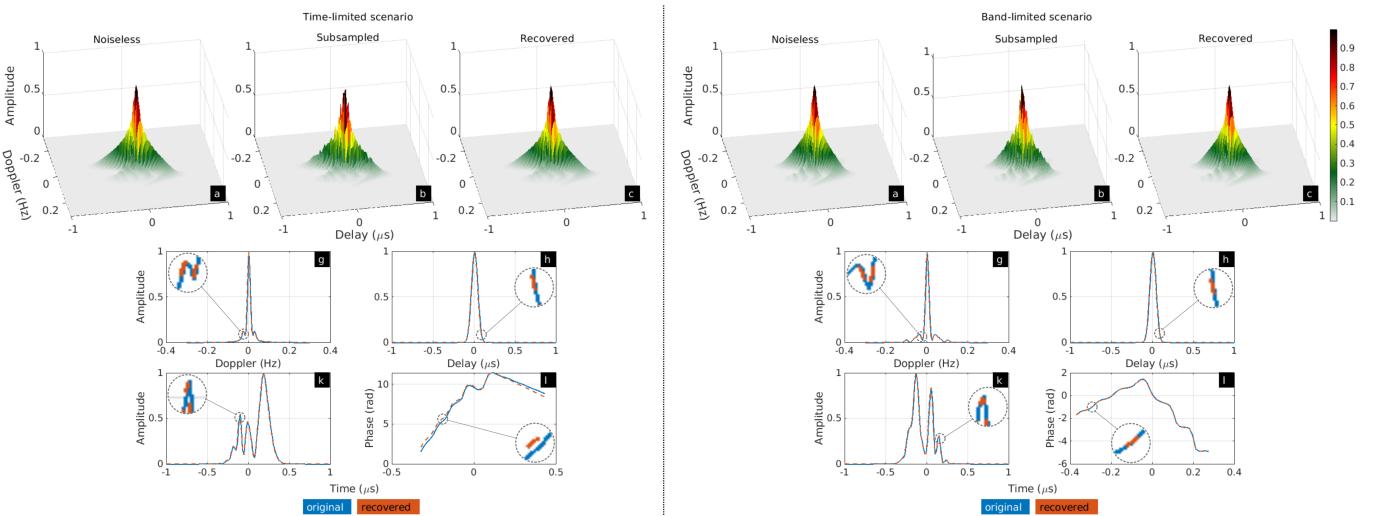


Fig. 10. Reconstructed time and band-limited signals when a 57% of the Fourier frequencies of their ambiguity functions are non-uniformly removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 6×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (l), (n) correspond to the recovered magnitude and phase of the estimated signals, respectively.

$a[n]$ for both kind of pulses model a rectangular envelope which is given by

$$a[n] = \begin{cases} 1 & 0 \leq \Delta tn \leq T \\ 0 & \text{otherwise} \end{cases}. \quad (92)$$

In this experiment two noisy scenarios are considered: first, the 50% of the delays are uniformly removed from the the AF, second, 19% of the first and last Fourier frequencies of the AF are removed. The results are summarized in Fig. 11, and 12, where $SNR = 20$ dB, and the attained relative error is also presented. These results suggest that Algorithm 1 is able to estimate accurately the phase of the pulses, while the reconstructed magnitudes present some artifacts. This limitation comes from the fact that their AF is significantly wide such that the removed information is enough to limit the reconstruction quality.

IX. CONCLUSION

This paper analytically demonstrates that time/band-limited signals can be estimated (up to trivial ambiguities) from its ambiguity function. We explore a trust region gradient method to estimate these kind of signals under complete/incomplete noisy and noiseless scenarios, and we verify that these signals can be estimated in a polynomial time with enough accuracy when the data is complete. Our algorithm consists of two steps: a spectral initialization followed by successive refinements based upon a sequence of gradient iterations. In the case of incomplete data, we found that although Proposition 1, and Corollary 1 suggest that the full AF is not required to guarantee uniqueness much more work can be done here in order to better estimate the pulses from incomplete data. In fact, numerical results suggest that pulses producing wide AF are not desire in order to

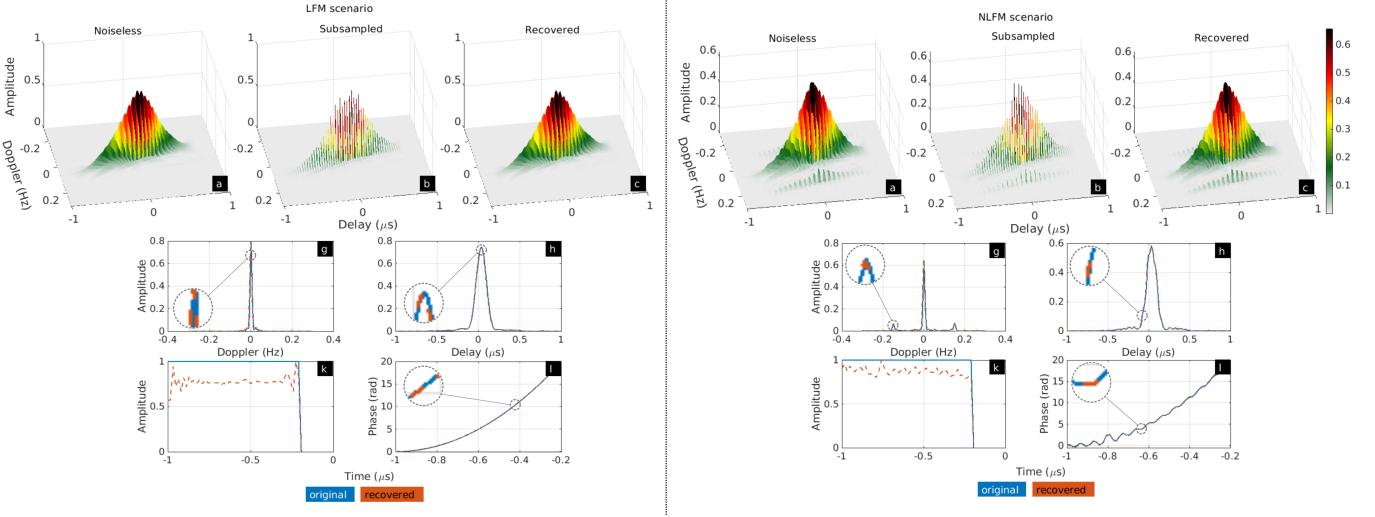


Fig. 11. Reconstructed time and band-limited signals when a 50% of the delays of their ambiguity functions are uniformly removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 6×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (j), (l) correspond to the recovered magnitude and phase of the estimated signals, respectively.

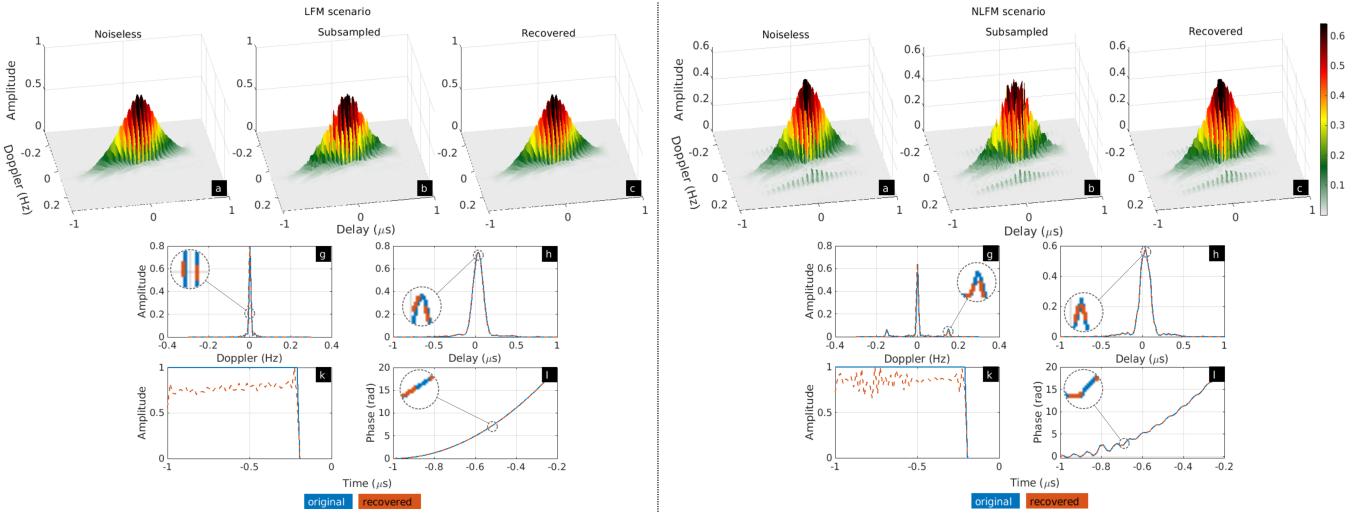


Fig. 12. Reconstructed time and band-limited signals when a 19% of the first and last Fourier frequencies of their ambiguity functions are removed. The incomplete AFs' were corrupted by noise with SNR = 20dB. The attained relative error as in (17) was 9×10^{-2} for both signals. (a),(d); (b),(e); and (c),(f) are the original, sub-sampled and recovered ambiguity functions, respectively. (g), (h), and (i), (j) are 1D slices of the ambiguity functions for the time and Doppler dimensions, respectively. (k), (m) and (j), (l) correspond to the recovered magnitude and phase of the estimated signals, respectively.

reduce the required data to be analyzed. Additionally, these result also validated Proposition 1, and Corollary 1 for three kind of signals.

Numerical experiments were conducted to evaluate the performance of the proposed method. The results show improvements that both the magnitude and the phase of the signal can be estimated even from noisy incomplete data. Additionally, the numerical results suggest the effectiveness of the proposed initialization under both noiseless and noisy scenarios with incomplete data. Future work should include implementing the proposed method on real data to further validate its performance.

There are several limitations of our current reconstruction algorithm. First, the initialization strategy employed is simple and can be improved in several ways. The optimization problem that our initialization pursues to solve is highly non-convex, and since we

use an alternating approach the present of saddle points and local minimum should be avoided. Second, we currently fix the parameters by simply cross-validation, however they can be learned from the kind of signals.

REFERENCES

- [1] N. Levanon, "Radar principles," *wi*, 1988.
- [2] M. Soumekh, *Synthetic aperture radar signal processing*. New York: Wiley, 1999, vol. 7.
- [3] M. Davis, "Foliage penetration radar: Detection and characterization of objects under trees, scitech. pub," *Inc., New York*, 2011.
- [4] F. Giovannesci, K. V. Mishra, M. A. Gonzalez-Huici, Y. C. Eldar, and J. H. Ender, "Dictionary learning for adaptive gpr landmine classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 10 036–10 055, 2019.
- [5] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, pp. 1–19, 2016.
- [6] J. Ville, "Theorie et application dela notion de signal analytique," *Câbles et transmissions*, vol. 2, no. 1, pp. 61–74, 1948.
- [7] P. M. Woodward, *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*. Elsevier, 2014, vol. 3.
- [8] ———, "Radar ambiguity analysis," ROYAL RADAR ESTABLISHMENT MALVERN (UNITED KINGDOM), Tech. Rep., 1967.
- [9] M. Jankiraman, *Design of multi-frequency CW radars*. SciTech Publishing, 2007, vol. 2.
- [10] G. San Antonio, D. R. Fuhrmann, and F. C. Robey, "Mimo radar ambiguity functions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 167–177, 2007.
- [11] B. Boashash, *Time-frequency signal analysis and processing: a comprehensive reference*. Academic Press, 2015.
- [12] A. W. Rihaczek, "Principles of high-resolution radar," *Norwood, MA: Artech House*, 1996., 1996.
- [13] R. de Buda, "Signals that can be calculated from their ambiguity function," *IEEE Transactions on Information Theory*, vol. 16, no. 2, pp. 195–202, 1970.
- [14] S. Sen and A. Nehorai, "Adaptive design of ofdm radar signal with improved wideband ambiguity function," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 928–933, 2009.
- [15] B. Nuss, J. Mayer, S. Marahrens, and T. Zwick, "Frequency comb ofdm radar system with high range resolution and low sampling rate," *IEEE Transactions on Microwave Theory and Techniques*, 2020.
- [16] F. Rodriguez-Morales, C. Leuschen, C. Carabajal, J. Paden, J. A. Wolf, S. Garrison, and J. McDaniel, "An improved uwb microwave radar for very long-range measurements of snow cover," *IEEE Transactions on Instrumentation and Measurement*, 2020.
- [17] P. Jamming, "The phase retrieval problem for the radar ambiguity function and vice versa," in *2010 IEEE Radar Conference*. IEEE, 2010, pp. 230–235.
- [18] ———, "Uniqueness results for the phase retrieval problem of fractional fourier transforms of variable order," *arXiv preprint arXiv:1009.3418*, 2010.
- [19] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "STFT phase retrieval: Uniqueness guarantees and recovery algorithms," *J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 770–781, 2016.
- [20] S. Nawab, T. Quatieri, and J. Lim, "Signal reconstruction from short-time fourier transform magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 986–998, 1983.
- [21] Y. C. Eldar, P. Sidorenko, D. G. Mixon, S. Barel, and O. Cohen, "Sparse phase retrieval from short-time fourier measurements," *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 638–642, 2014.
- [22] T. Bendory, Y. C. Eldar, and N. Boumal, "Non-convex phase retrieval from STFT measurements," *IEEE Trans. on Inf. Theory*, vol. 64, no. 1, pp. 467–484, 2018.
- [23] T. Bendory, D. Edidin, and Y. C. Eldar, "On signal reconstruction from FROG measurements," *Appl. and Compu. Harmon. Anal.*, 2018.
- [24] T. Bendory, P. Sidorenko, and Y. C. Eldar, "On the uniqueness of FROG methods," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 722–726, 2017.
- [25] S. Pinilla, T. Bendory, Y. C. Eldar, and H. Arguello, "Frequency-resolved optical gating recovery via smoothing gradient," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 6121–6132, 2019.
- [26] 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," *IEEE signal process. mag.*, vol. 32, no. 3, pp. 87–109, 2015.
- [27] M. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 2, pp. 140–154, 1982.
- [28] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals: Uniqueness conditions," *arXiv preprint arXiv:1308.3058*, 2013.
- [29] H. Ohlsson and Y. C. Eldar, "On conditions for uniqueness in sparse phase retrieval," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 1841–1845.
- [30] T. Bendory, D. Edidin, and Y. C. Eldar, "Blind phaseless short-time fourier transform recovery," *IEEE Transactions on Information Theory*, 2019.
- [31] G. Wang, L. Zhang, G. B. Giannakis, M. Akçakaya, and J. Chen, "Sparse phase retrieval via truncated amplitude flow," *IEEE Trans. on Signal Process.*, vol. 66, no. 2, pp. 479–491, Jan 2018.
- [32] S. Pinilla, J. Bacca, and H. Arguello, "Phase retrieval algorithm via nonconvex minimization using a smoothing function," *IEEE Trans. on Signal Process.*, vol. 66, no. 17, pp. 4574–4584, 2018.
- [33] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Appl. and Comput. Harmon. Anal.*, vol. 39, no. 2, pp. 277–299, 2015.
- [34] ———, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. on Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.
- [35] I. P. Theron, E. K. Walton, S. Gunawan, and L. Cai, "Ultrawide-band noise radar in the vhf/uhf band," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 6, pp. 1080–1084, 1999.
- [36] C.-Y. Chen and P. P. Vaidyanathan, "Mimo radar space-time adaptive processing using prolate spheroidal wave functions," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 623–635, 2008.
- [37] R. Beinert and G. Plonka, "Enforcing uniqueness in one-dimensional phase retrieval by additional signal information in time domain," *Applied and Computational Harmonic Analysis*, vol. 45, no. 3, pp. 505–525, 2018.
- [38] T. Bendory, R. Beinert, and Y. C. Eldar, "Fourier phase retrieval: Uniqueness and algorithms," in *Compressed Sensing and its Applications*. Springer, 2017, pp. 55–91.
- [39] E. J. R. Pauwels, A. Beck, Y. C. Eldar, and S. Sabach, "On Fienup methods for sparse phase retrieval," *IEEE Trans. on Signal Process.*, vol. 66, no. 4, pp. 982–991, 2018.
- [40] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Trans. on Inf. Theory*, vol. 64, no. 2, pp. 773–794, 2018.
- [41] H. Zhang and Y. Liang, "Reshaped Wirtinger flow for solving quadratic system of equations," in *Adv. in Neural Inf. Process. Systems*, 2016, pp. 2622–2630.
- [42] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [43] R. Hunger, *An introduction to complex differentials and complex differentiability*. Munich University of Technology, Inst. for Circuit Theory and Signal Processing, 2007.
- [44] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley & Sons, 2005, vol. 65.
- [45] N. Parikh and S. Boyd, "Proximal algorithms," *Foundat. and Trends® in Optim.*, vol. 1, no. 3, pp. 127–239, 2014. [Online]. Available: <http://dx.doi.org/10.1561/2400000003>
- [46] S. Ghadimi and G. Lan, "Stochastic first-and zeroth-order methods for nonconvex stochastic programming," *SIAM J. on Opti.*, vol. 23, no. 4, pp. 2341–2368, 2013.
- [47] C. Zhang and X. Chen, "Smoothing projected gradient method and its application to stochastic linear complementarity problems," *SIAM J. on Opti.*, vol. 20, no. 2, pp. 627–649, 2009.
- [48] J. C. Duchi and F. Ruan, "Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval," *arXiv preprint arXiv:1705.02356*, 2017.

Samuel Pinilla (S'17) received the B.S. degree (*cum laude*) in Computer Science in 2014, the B.S. degree in Mathematics, and the M.S. degree in Mathematics from Universidad Industrial de Santander, Bucaramanga, Colombia in 2016 and 2017, respectively. He is currently working toward the Ph.D. degree in the Department of the Electrical and Computer Engineering, Universidad Industrial de Santander, Bucaramanga, Colombia. His research interests focuses on the areas of high-dimensional structured signal processing and (non)convex optimization methods.

Kumar Vijay Mishra (S'17-M'15-SM'18) received the B.Tech. degree (*summa cum laude*) (Hons.) in electronics and communications engineering from the National Institute of Technology, Hamirpur, India, in 2003, the M.S. degree in electrical and computer engineering from Colorado State University, Fort Collins, CO, USA, in 2012, and the Ph.D. degree in electrical and computer engineering and the M.S. degree in mathematics from The University of Iowa, Iowa City, IA, USA, in 2015, while working on the NASA Global Precipitation Mission (GPM) Ground Validation (GV) weather radars.

He has over 15 years of experience in research and development of various radar systems. From 2003 to 2007, he was involved in military surveillance radars as a Research Scientist at the Electronics and Radar Development Establishment (LRDE), Defence Research and Development Organisation (DRDO), Bengaluru, India. From 2015 to 2017, he was an Andrew and Erna Finci Viterbi and Lady Davis Post-Doctoral Fellow with the Viterbi Faculty of Electrical Engineering, Technion–Israel Institute of Technology, Haifa, Israel. His research interests include radar systems, signal processing, remote sensing, and electromagnetics.

Dr. Mishra was a recipient of the Royal Meteorological Society Quarterly Journal Editor's Prize in 2017, the Lady Davis Fellowship from 2016 to 2017, the Andrew and Erna Finci Viterbi Fellowship (twice awarded) in 2015 and 2016, the Technion Excellent Undergraduate Adviser Award in 2017, the DRDO LRDE Scientist of the Year Award in 2006, the NITH Director's Gold Medals in 2003, and the NITH Best Student Award in 2003.

Brian M. Sadler (S'81-M'81-SM'02-F'07) received the B.S. and M.S. degrees in electrical engineering from the University of Maryland, College Park,

MD, USA, and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, USA. He is the US Army Senior Scientist for Intelligent Systems, and a Fellow of the Army Research Laboratory (ARL), Adelphi, MD, USA. His research interests include information science, and networked and autonomous intelligent systems. He was an IEEE Signal Processing Society Distinguished Lecturer from 2017 to 2018, and the General Co-Chair of IEEE GlobalSIP'16. He has been an Associate Editor for the IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, and EURASIP Signal Processing, and a Guest Editor for several journals including the IEEE JSTSP, IEEE JSAC, IEEE SP Magazine, Autonomous Robots, and International Journal of Robotics Research. He was the recipient of the Best Paper Awards from the IEEE Signal Processing Society in 2006 and 2010, several ARL and Army R&D awards, and a 2008 Outstanding Invention of the Year Award from the University of Maryland.

Henry Arguello (S'11) received the B.Sc. Eng. degree in electrical engineering and the M.Sc. degree in electrical power from the Universidad Industrial de Santander (UIS), Colombia, in 2000 and 2003, respectively, and the Ph.D. degree in electrical engineering from the University of Delaware, USA, in 2013. He is currently an Associate Professor with the Department of Systems Engineering, Universidad Industrial de Santander, Colombia. In 2020, he was a visiting professor at Stanford University funded by Fulbright. His research interests include high-dimensional signal processing, optical imaging, compressed sensing, hyperspectral imaging, and computational imaging. He has published one book chapter, more than 100 journal papers, 150 conference papers, and holds one international patent. Web: <http://hdspgroup.com/>