Identification of Parametric Underspread Linear Systems and Super-Resolution Radar

Waheed U. Bajwa, *Member, IEEE*, Kfir Gedalyahu, and Yonina C. Eldar, *Senior Member, IEEE*

Abstract—**Identification of time-varying linear systems, which introduce both time-shifts (delays) and frequency-shifts (Doppler-shifts), is a central task in many engineering applications. This paper studies the problem of identification of underspread linear systems (ULSs), whose responses lie within a unit-area region in the delay–Doppler space, by probing them with a known input signal. It is shown that sufficiently-underspread parametric linear systems, described by a finite set of delays and Doppler-shifts, are identifiable from a single observation as long as the time–bandwidth product of the input signal is proportional to the square of the total number of delay–Doppler pairs in the system. In addition, an algorithm is developed that enables identification of parametric ULSs from an input train of pulses in polynomial time by exploiting recent results on sub-Nyquist sampling for time delay estimation and classical results on recovery of frequencies from a sum of complex exponentials. Finally, application of these results to super-resolution target detection using radar is discussed. Specifically, it is shown that the proposed procedure allows to distinguish between multiple targets with very close proximity in the delay–Doppler space, resulting in a resolution that substantially exceeds that of standard matched-filtering based techniques without introducing leakage effects inherent in recently proposed compressed sensing-based radar methods.**

Index Terms—**Compressed sensing, Delay–Doppler estimation, rotational invariance techniques, super-resolution radar, system identification, time-varying linear systems**

I. INTRODUCTION

Physical systems arising in a number of application areas can often be described as linear and time varying [1], [2]. Identification of such systems may help improve overall performance, e.g., the bit-error rate in communications [1], or constitute an integral part of the overall system operation, e.g., target detection using radar or active sonar [2].

Mathematically, identification of a given time-varying linear system $\mathcal{H}$ involves probing it with a known input signal $x(t)$ and identifying $\mathcal{H}$ by analyzing the single system output $\mathcal{H}(x(t))$ [3], as illustrated in Fig. 1. Unlike time-invariant linear systems, however, a single observation of a time-varying linear system does not lead to a unique solution unless additional constraints on the system response are imposed.

W.U. Bajwa is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (Email: w.bajwa@duke.edu). K. Gedalyahu and Y.C. Eldar are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (Phone: +972-4-8293256, Fax: +972-4-8295757, E-mails: {kfirge@techunix,yonina@ee}.technion.ac.il. Y.C. Eldar is currently also a Visiting Professor at Stanford University, Stanford, CA 94305 USA.

This is due to the fact that such systems introduce both time-shifts (*delays*) and frequency-shifts (*Doppler-shifts*) to the input signal. It is now a well-established fact in the literature that a time-varying linear system $\mathcal{H}$ can only be identified from a single observation if $\mathcal{H}(\delta(t))$ is known to lie within a region $\mathcal{R}$ of the delay–Doppler space such that $\texttt{area}(\mathcal{R}) < 1$ [3]–[6]. Identifiable time-varying linear systems are termed *underspread*, as opposed to nonidentifiable *overspread* linear systems, which satisfy $\texttt{area}(\mathcal{R}) > 1$ [3], [6].[1]

In this paper, we study the problem of identification of underspread linear systems (ULSs) whose responses can be described by a finite set of delays and Doppler-shifts. That is,

$$\mathcal{H}(x(t)) = \sum_{k=1}^{K} \alpha_k x(t - \tau_k) e^{j 2\pi \nu_k t} \tag{1}$$

where (τ_k, ν_k) denotes a delay–Doppler pair and $\alpha_k \in \mathbb{C}$ is the complex attenuation factor associated with (τ_k, ν_k). Unlike most of the existing work in the literature, however, our goal in this paper is to explicitly characterize conditions on the bandwidth and temporal support of the input signal that ensure identification of such ULSs from single observations. The importance of this goal can be best put into perspective by realizing that ULSs of the form (1) tend to arise frequently in many applications. Consider, for example, a single-antenna transmitter communicating wirelessly with a single-antenna receiver in a mobile environment. Over a small-enough time interval, the channel between the transmitter and receiver is of the form (1) with each triplet $(\tau_k, \nu_k, \alpha_k)$ corresponding to a distinct physical path between the transmitter and receiver [7]. Identification of $\mathcal{H}$ can enable one to establish a relatively error-free communication link between the transmitter and receiver. But wireless systems also need to identify channels using signals that have as small *time–bandwidth product* as possible so that they can allocate the rest of the temporal degrees of freedom to communicating data [7], [8].

Similarly, in the case of target detection using radar or active sonar, the (noiseless, clutter-free) received signal is of the form (1) with each triplet $(\tau_k, \nu_k, \alpha_k)$ corresponding to an echo of the transmitted signal from a distinct target in the delay–Doppler space [2]. Identification of $\mathcal{H}$ in this case enables one to accurately obtain the radial position and velocity of the targets. Radar systems also strive to operate with signals (waveforms) that have as small temporal support

---

[1]It is still an open research question as to whether *critically-spread* linear systems, which correspond to $\texttt{area}(\mathcal{R}) = 1$, are identifiable or nonidentifiable [6]; see [3] for a partial answer to this question for the case when $\mathcal{R}$ is a rectangular region.
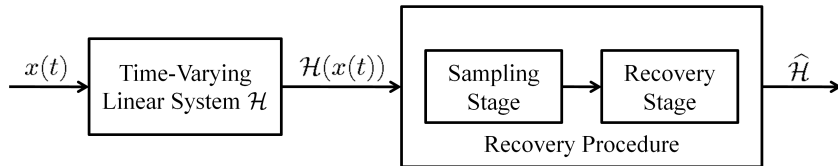
Fig. 1. Schematic representation of identification of a time-varying linear system $\mathcal{H}$ by probing it with a known input signal. Characterization of an identification scheme involves specification of the input probe, $x(t)$, and the accompanying sampling and recovery stages.

and bandwidth as possible. This is because the temporal support of the radar waveform is directly tied to the time it takes to identify all the targets while the bandwidth of the waveform—among other technical considerations—is tied to the sampling rate of the radar receiver [2].

Given the ubiquity of time-varying linear systems in engineering applications, there exists considerable amount of existing literature that studies identification of such systems in an abstract setting. Kailath was the first to recognize that the identifiability of a time-varying linear system $\mathcal{H}$ from a single observation is directly tied to the area of the region $\mathcal{R}$ that contains $\mathcal{H}(\delta(t))$ [4]. Kailath's seminal work in [4] laid the foundations for the future works of Bello [5], Kozek and Pfander [3], and Pfander and Walnut [6], which establish the nonidentifiability of overspread linear systems and provide constructive proofs for the identifiability of arbitrary ULSs. However, none of these results shed any light on the bandwidth and temporal support of the input signal needed to ensure identification of ULSs of the form (1). On the contrary, the constructive proofs provided in [3]–[6] require use of input signals with infinite bandwidth and temporal support.

In contrast, to the best of our knowledge, this is the first paper to develop a theory for identification of ULSs of the form (1), henceforth referred to as *parametric* ULSs, that parallels that of [3]–[6] for identification of arbitrary ULSs. One of the main contributions of this paper is that we establish using a constructive proof that sufficiently-underspread parametric linear systems are identifiable as long as the time–bandwidth product of the input signal is proportional to the square of the total number of delay–Doppler pairs in the system. Equally importantly, as part of our constructive proof, we concretely specify the nature of the input signal (a finite train of pulses) and the structure of a corresponding polynomial-time (in the number of delay–Doppler pairs) recovery procedure that enable identification of parametric ULSs. These ideas are also immediately applicable to super-resolution target detection using radar and we show in the latter part of the paper that our approach indeed results in a resolution that substantially exceeds that of standard matched-filtering based techniques without introducing leakage effects inherent in recently proposed compressed sensing-based radar methods [9].

The key developments in the paper leverage recent results on sub-Nyquist sampling for time-delay estimation [10] and classical results on direction-of-arrival (DOA) estimation [11]–[14]. Unlike the traditional DOA estimation literature, however, we do not assume that the system output is observed by an array of antennas. Additionally, in contrast to [10], our goal here is not a reduction in the sampling rate; rather, we are

interested in characterizing the minimum temporal degrees of freedom of the input signal needed to ensure identification of a parametric ULS $\mathcal{H}$. The connection to sub-Nyquist sampling can be understood by noting that the sub-Nyquist sampling results of [10] enable recovery of the delays associated with $\mathcal{H}$ using a small-bandwidth input signal. Further, the "train-of-pulses" nature of the input signal combined with results on recovery of frequencies from a sum of complex exponentials [14] allow recovery of the Doppler-shifts and attenuation factors using an input signal of small temporal support.

Several works in the past have considered identification of specialized versions of parametric ULSs. Specifically, [9], [15]–[18] treat parametric ULSs whose delays and Doppler-shifts lie on a quantized grid in the delay–Doppler space. On the other hand, [19] considers the case in which there are no more than two Doppler-shifts associated with the same delay. The proposed recovery procedures in [19] also have exponential complexity, since they require exhaustive searches in a K-dimensional space, and stable initializations of these procedures stipulate that the system output be observed by an M-element antenna array with $M \gtrsim K$.

While the insights of [9], [15]–[18] can be extended to arbitrary parametric ULSs by taking infinitesimally-fine quantization of the delay–Doppler space, this will require input signals with infinite bandwidth and temporal support. In contrast, our ability to avoid quantization of the delay–Doppler space is due to the fact that we treat the system-identification problem directly in the analog domain. This follows the philosophy in much of the recent work in analog compressed sensing, termed Xampling, which provides a framework for incorporating and exploiting structure in analog signals without the need for quantization [20]–[25]. This is in particular the key enabling factor that helps us avoid the catastrophic implications of the leakage effects in the context of radar target detection.

Before concluding this discussion, we note that responses of arbitrary ULSs can always be represented as (1) under the limit $K \to \infty$. Therefore, the main result of this paper can also be construed as an alternate constructive proof of the identifiability of sufficiently-underspread linear systems. Nevertheless, just like [3]–[6], this interpretation of the presented results again seem to suggest that identification of arbitrary ULSs requires use of input signals with infinite bandwidth and temporal support.

The rest of this paper is organized as follows. In Section II, we formalize the problem of identification of parametric ULSs along with the accompanying assumptions. In Section III, we propose a polynomial-time recovery procedure used for the identification of parametric ULSs, while Section IV specifies

the conditions on the input signal needed to guarantee unique identification using the proposed procedure. We compare the results of this paper to some of the related literature on identification of parametric ULSs in Section V and discuss an application of our results to super-resolution target detection using radar in Section VI. Finally, we present results of some numerical experiments in Section VII.

We make use of the following notational convention throughout this paper. Vectors and matrices are denoted by bold-faced lowercase and bold-faced uppercase letters, respectively. The nth element of a vector $\mathbf{a}$ is written as $\mathbf{a}_n$, and the (i, j)th element of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}_{ij}$. Superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent conjugation, transposition, and conjugate transposition, respectively. In addition, the Fourier transform of a continuous-time signal $x(t) \in L_2(\mathbb{C})$ is defined by $X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$, while $\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) y^*(t) dt$ denotes the inner product between two continuous-time signals in $L_2(\mathbb{C})$. Similarly, the discrete-time Fourier transform of a sequence $a[n] \in \ell_2(\mathbb{C})$ is defined by $A(e^{j\omega T}) = \sum_{n \in \mathbb{Z}} a[n] e^{-j\omega nT}$ and is periodic in ω with period $2\pi/T$. Finally, we use $\mathbf{A}^{\dagger}$ to write the Moore–Penrose pseudoinverse of a matrix $\mathbf{A}$.

II. PROBLEM FORMULATION AND MAIN RESULTS

In this section, we formalize the problem of identification of a parametric ULS $\mathcal{H}$ whose response is described by a total of K *arbitrary* delay–Doppler-shifts of the input signal. The task of identification of $\mathcal{H}$ essentially requires specifying two distinct but highly intertwined steps. First, we need to specify the conditions on the bandwidth and temporal support of the input signal $x(t)$ that ensure identification of $\mathcal{H}$ from a single observation. Second, we need to provide a polynomial-time recovery procedure that takes as input $\mathcal{H}(x(t))$ and provides an estimate $\widehat{\mathcal{H}}$ of the system response by exploiting the properties of $x(t)$ specified in the first step. We begin by detailing our system model and the accompanying assumptions.

In (1), some of the delays, τ_k, might be repeated. Expressing $\mathcal{H}$ in terms of $K_\tau \leq K$ *distinct* delays in this case leads to

$$\mathcal{H}(x(t)) = \sum_{i=1}^{K_\tau} \sum_{j=1}^{K_{\nu,i}} \alpha_{ij} x(t - \tau_i) e^{j2\pi \nu_{ij} t} \qquad (2)$$

where ν_{ij} denotes the jth Doppler-shift associated with the ith distinct delay τ_i, $\alpha_{ij} \in \mathbb{C}$ denotes the attenuation factor associated with the delay–Doppler pair (τ_i, ν_{ij}), and $K = \sum_{i=1}^{K_\tau} K_{\nu,i}$. We choose to express $\mathcal{H}(x(t))$ as in (2) so as to facilitate the forthcoming analysis. Throughout the rest of the paper, we use $\boldsymbol{\tau} = \{\tau_i, \ i = 1, \ldots, K_\tau\}$ to denote the set of K_τ distinct delays associated with $\mathcal{H}$. The first main assumption that we make concerns the footprint of $\mathcal{H}$ in the delay–Doppler space:

[A1] The response $\mathcal{H}(\delta(t))$ of $\mathcal{H}$ lies within a rectangular region of the delay–Doppler space; in other words, $(\tau_i, \nu_{ij}) \in [0, \tau_{max}] \times [-\nu_{max}/2, \nu_{max}/2]$. This is indeed the case in many engineering applications (see, e.g., [1], [2]). The parameters τ_{max} and ν_{max} are termed in the parlance of linear systems as the *delay spread* and the *Doppler spread* of the system, respectively.

Next, we use $\mathcal{T}$ and $\mathcal{W}$ to denote the temporal support and the two-sided bandwidth of the known input signal $x(t)$ used to probe $\mathcal{H}$, respectively. We propose using input signals that correspond to a finite train of pulses:

$$x(t) = \sum_{n=0}^{N-1} x_n g(t - nT), \ 0 \leq t \leq \mathcal{T} \qquad (3)$$

where $g(t)$ is a prototype pulse of bandwidth $\mathcal{W}$ that is (essentially) temporally supported on $[0, T]$ and is assumed to have unit energy ($\int |g(t)|^2 dt = 1$), and $\{x_n \in \mathbb{C}\}$ is an N-length probing sequence. The parameter N is proportional to the time–bandwidth product of $x(t)$, which roughly defines the number of temporal degrees of freedom available for estimating $\mathcal{H}$ [8]: $N = \mathcal{T}/T \propto \mathcal{T}\mathcal{W}$.[2] The final two assumptions that we make concern the relationship between the delay spread τ_{max} and the Doppler spread ν_{max} of $\mathcal{H}$, and the temporal support T and bandwidth $\mathcal{W}$ of $g(t)$:

[A2] The delay spread of $\mathcal{H}$ is strictly smaller than the temporal support of $g(t)$: $\tau_{max} < T$, and

[A3] The Doppler spread of $\mathcal{H}$ is much smaller than the bandwidth of $g(t)$: $\nu_{max} \ll \mathcal{W}$.

Note that, since $\mathcal{W} \propto 1/T$, [A3] equivalently imposes that $\nu_{max}T \ll 1$. This assumption states that the temporal scale of variations in $\mathcal{H}$ is large relative to the temporal scale of variations in $x(t)$. It is worth pointing out that linear systems that are sufficiently underspread in the sense that $\tau_{max}\nu_{max} \ll 1$ can always be made to satisfy [A2] and [A3] for any given budget of the time–bandwidth product.

Remark 1. In order to elaborate on the validity of [A2] and [A3], note that there exist many communication applications where underlying linear systems tend to be highly underspread [1, § 14.2]. Similarly, [A2] and [A3] in the context of radar target detection simply mean that the targets are not too far away from the radar and that their velocities are not too high. Consider, for example, an L-band radar (operating frequency of 1.3 GHz) that transmits a pulse $g(t)$ of bandwidth $\mathcal{W} = 10$ MHz after every $T = 50$ μs. Then both [A2] and [A3] are satisfied when the distance between the radar and any target is at most 7.5 km ($\tau_{max} \approx 50$ μs) and the radial velocity of any target is at most 185 km/h ($\nu_{max} \approx 445$ Hz) [2].

The following theorem summarizes our key result concerning identification of parametric ULSs.

Theorem 1 (Identification of Parametric Underspread Linear Systems). *Suppose that $\mathcal{H}$ is a parametric ULS that is completely described by a total of $K = \sum_{i=1}^{K_\tau} K_{\nu,i}$ triplets $(\tau_i, \nu_{ij}, \alpha_{ij})$. Then, irrespective of the distribution of $\{(\tau_i, \nu_{ij})\}$ within the delay–Doppler space, $\mathcal{H}$ can be identified in polynomial-time from a single observation $\mathcal{H}(x(t))$ as long as it satisfies [A1]–[A3], the probing sequence $\{x_n\}$ remains bounded away from zero in the sense that $|x_n| > 0$ for every $n = 0, \ldots, N - 1$, and the time–bandwidth product of the known input signal $x(t)$ satisfies the condition*

$$\mathcal{T}\mathcal{W} \geq 8\pi K_\tau K_{\nu,max} \qquad (4)$$

---

[2]Recall that the temporal support and the bandwidth of an arbitrary pulse $g(t)$ are related to each other as $\mathcal{W} \propto 1/T$.

where $K_{\nu,max} = \max_i K_{\nu,i}$ is the maximum number of Doppler-shifts associated with any one of the distinct delays. In addition, the time–bandwidth product of $x(t)$ is guaranteed to satisfy (4) as long as $\mathcal{TW} \geq 2\pi(K+1)^2$.

The rest of this paper is devoted to providing a proof of Theorem 1. In terms of a general roadmap for the proof, we first exploit the sub-Nyquist sampling results of [10] to argue that $x(t)$ with small bandwidth suffices to recover the delays associated with $\mathcal{H}$. We then exploit the "train-of-pulses" structure of $x(t)$ and classical results on recovery of frequencies from a sum of complex exponentials [14] to argue that $x(t)$ with small temporal support suffices to recover the Doppler-shifts and attenuation factors associated with $\mathcal{H}$. The statement of Theorem 1 will then follow by a simple combination of the two claims concerning the bandwidth and temporal support of $x(t)$. We will make use of (2) and (3) in the following to describe:

[1] The polynomial-time recovery procedure used for the identification of $\mathcal{H}$ (cf. Section III), and

[2] The accompanying conditions on $x(t)$ needed to guarantee identification of $\mathcal{H}$ (cf. Section IV).

III. POLYNOMIAL-TIME IDENTIFICATION OF ULSs

In this section, we characterize the polynomial-time recovery procedure used for identification of ULSs of the form (2). In order to facilitate understanding of the proposed algorithm, shown in Fig. 2, we conceptually partition the method into two stages: sampling and recovery. The rest of this section is devoted to describing these two steps in detail. Before proceeding further, however, it is instructive to first make use of (2) and (3) and rewrite the output of $\mathcal{H}$ as

$$
\begin{aligned}
\mathcal{H}(x(t)) &= \sum_{i=1}^{K_\tau}\sum_{j=1}^{K_{\nu,i}}\sum_{n=0}^{N-1}\alpha_{ij}x_n e^{j2\pi\nu_{ij}t}g(t-\tau_i-nT)\\
&\overset{(a)}{\approx} \sum_{i=1}^{K_\tau}\sum_{j=1}^{K_{\nu,i}}\sum_{n=0}^{N-1}\alpha_{ij}x_n e^{j2\pi\nu_{ij}nT}g(t-\tau_i-nT)\\
&= \sum_{i=1}^{K_\tau}\sum_{n=0}^{N-1}a_i[n]g(t-\tau_i-nT)
\end{aligned} \tag{5}
$$

where (a) follows from the assumption $\nu_{max}T \ll 1$, which implies that $e^{j2\pi\nu_{ij}t} \approx e^{j2\pi\nu_{ij}nT}$ for all $t \in [(n-1)T, nT)$, and the sequences $\{a_i[n]\}, i = 1, \ldots, K_\tau$, are defined as

$$
a_i[n] = \sum_{j=1}^{K_{\nu,i}}\alpha_{ij}x_n e^{j2\pi\nu_{ij}nT}, \ n = 0, \ldots, N-1. \tag{6}
$$

A. The Sampling Stage

We leverage the ideas of [10] on time-delay estimation from sub-Nyquist samples to describe the sampling stage of our recovery procedure. While the primary objective in [10] is time-delay estimation from low-rate samples, the development here is carried out with an eye towards identification of parametric ULSs regardless of the distribution of system parameters within the delay–Doppler space—the so-called *super-resolution identification*. In [10], a general multi-channel

sub-Nyquist sampling scheme was introduced for the purpose of recovering a set of unknown delays from signals of the form (5). Here, we focus on one special case of that scheme, which consists of a low-pass filter (LPF) followed by a uniform sampler. This architecture may be preferable from an implementation viewpoint since it requires only one sampling channel, thereby simplifying analog front-end of the sampling hardware. The LPF, besides being required by the sampling stage, also serves as the front-end of the system-identification hardware and rejects noise and interference outside the working spectral band.

Our sampling stage first passes the system output $y(t) = \mathcal{H}(x(t))$ through a LPF whose impulse response is given by $s^*(-t)$ and then uniformly samples the LPF output at times $\{t = nT/p\}$. We assume that the frequency response, $S^*(\omega)$, of the LPF is contained in the spectral band $\mathcal{F}$, defined as

$$
\mathcal{F} = \left[-\frac{\pi}{T}p, \frac{\pi}{T}p\right], \tag{7}
$$

and is zero for frequencies $\omega \notin \mathcal{F}$. Here, the parameter p is assumed to be even and satisfies the condition $p \geq 2K_\tau$; exact requirements on p to ensure identification of $\mathcal{H}$ will be given in Section IV. In order to relate the sampled output of the LPF with the multi-channel sampling formulation of [10], we define p sampling (sub)sequences $\{c_\ell[n]\}$ as

$$
c_\ell[n] = \langle y(t), s(t-nT-(\ell-1)T/p)\rangle, \ \ell = 1, \ldots, p. \tag{8}
$$

These sequences correspond to periodically splitting the samples at the output of the LPF, which is generated at a rate of p/T, into p slower sequences at a rate of $1/T$ each using a serial-to-parallel converter; see Fig. 2 for a schematic representation of this splitting.

Next, we define the vector $\mathbf{c}\left(e^{j\omega T}\right)$ as the p-length vector whose ℓth element is $C_\ell\left(e^{j\omega T}\right)$, which denotes the *discrete-time Fourier transform* (DTFT) of $c_\ell[n]$. In a similar fashion, we define $\mathbf{a}\left(e^{j\omega T}\right)$ as the K_τ-length vector whose ith element is given by $A_i\left(e^{j\omega T}\right)$, the DTFT of $a_i[n]$. It can be shown following the developments carried out in [10] that these two vectors are related to each other by

$$
\mathbf{c}\left(e^{j\omega T}\right) = \mathbf{W}\left(e^{j\omega T}\right)\mathbf{N}\left(\boldsymbol{\tau}\right)\mathbf{D}\left(e^{j\omega T}, \boldsymbol{\tau}\right)\mathbf{a}\left(e^{j\omega T}\right). \tag{9}
$$

Here, $\mathbf{W}\left(e^{j\omega T}\right)$ is a $p \times p$ matrix with (ℓ, m)th element

$$
\mathbf{W}_{\ell m}\left(e^{j\omega T}\right) = e^{j\omega(\ell-1)T/p}\frac{1}{T}S^*\left(\omega + \frac{2\pi}{T}m'\right)G\left(\omega + \frac{2\pi}{T}m'\right)e^{j\frac{2\pi}{p}(\ell-1}
$$
(10)

where $m' = m - p/2 - 1$, $\mathbf{N}\left(\boldsymbol{\tau}\right)$ is a $p \times K_\tau$ Vandermonde matrix with (m, i)th element

$$
\mathbf{N}_{mi}\left(\boldsymbol{\tau}\right) = e^{-j\frac{2\pi}{T}m'\tau_i}, \tag{11}
$$

and $\mathbf{D}\left(e^{j\omega T}, \boldsymbol{\tau}\right)$ is a $K_\tau \times K_\tau$ diagonal matrix whose ith diagonal element is given by $e^{-j\omega\tau_i}$. Assuming for the time being that $\mathbf{W}\left(e^{j\omega T}\right)$ is a stably-invertible matrix, we define the *modified measurement vector* $\mathbf{d}\left(e^{j\omega T}\right) = \mathbf{W}^{-1}\left(e^{j\omega T}\right)\mathbf{c}\left(e^{j\omega T}\right)$. Denoting

$$
\mathbf{b}\left(e^{j\omega T}\right) = \mathbf{D}\left(e^{j\omega T}, \boldsymbol{\tau}\right)\mathbf{a}\left(e^{j\omega T}\right), \tag{12}
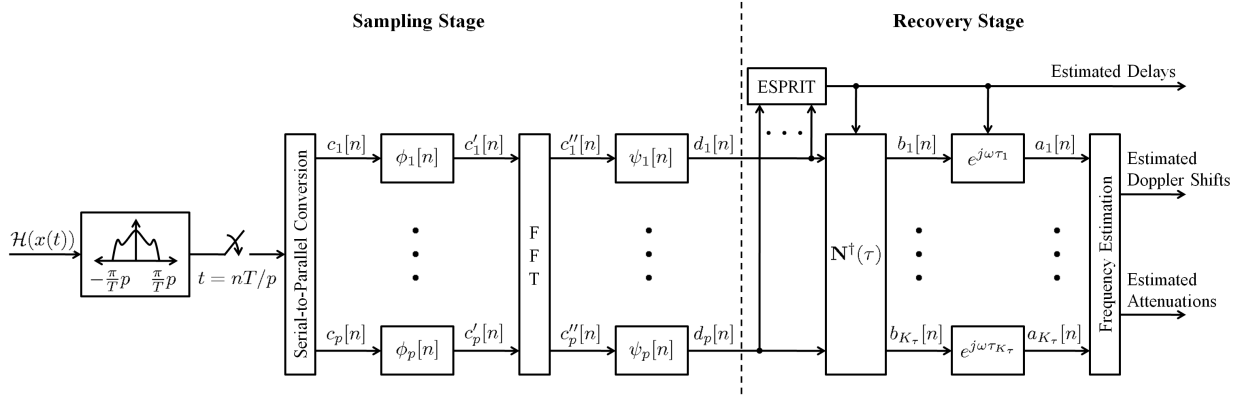$$

Fig. 2. Schematic representation of the polynomial-time recovery procedure for identification of parametric underspread linear systems from single observations.

we see from (10) that

$$\mathbf{d}\left(e^{j\omega T}\right) = \mathbf{N}\left(\boldsymbol{\tau}\right)\mathbf{b}\left(e^{j\omega T}\right). \qquad (13)$$

Since $\mathbf{N}\left(\boldsymbol{\tau}\right)$ is not a function of ω, (13) can be expressed in the discrete-time domain using the linearity of the DTFT as

$$\mathbf{d}[n] = \mathbf{N}\left(\boldsymbol{\tau}\right)\mathbf{b}[n], \quad n \in \mathbb{Z}. \qquad (14)$$

Here, the elements of the vectors $\mathbf{d}[n]$ and $\mathbf{b}[n]$ are discrete-time sequences that are given by the inverse DTFT of the elements of $\mathbf{d}\left(e^{j\omega T}\right)$ and $\mathbf{b}\left(e^{j\omega T}\right)$, respectively.

The key insight to be drawn here is that (13), and its time-domain equivalent (14), can be viewed as an infinite ensemble of modified measurement vectors in which each element corresponds to a distinct matrix $\mathbf{N}\left(\boldsymbol{\tau}\right)$ that, in turn, depends on the set of (distinct) delays $\boldsymbol{\tau}$. Linear measurement models of the form (14)—in which the measurement matrix is completely determined by a set of (unknown) parameters—have been studied extensively in a number of research areas such as system identification [26] and direction-of-arrival and spectrum estimation [14], [27]. One specific class of methods that has proven to be quite useful in these areas in efficiently recovering the parameters that describe the measurement matrix are the so-called *subspace methods* [27]. Consequently, our approach in the recovery stage will be to first use subspace methods in order to recover the set $\boldsymbol{\tau}$ from $\mathbf{d}[n]$. Afterwards, since $\mathbf{N}^{\dagger}\left(\boldsymbol{\tau}\right)\mathbf{N}\left(\boldsymbol{\tau}\right) = \mathbf{I}$ because of the assumption that $p \geq 2K_{\tau}$, we will recover $\mathbf{a}\left(e^{j\omega T}\right)$ from $\mathbf{d}[n]$ using linear filtering operations as follows [cf. (12), (13)]

$$\mathbf{a}\left(e^{j\omega T}\right) = \mathbf{D}^{-1}\left(e^{j\omega T}, \boldsymbol{\tau}\right)\mathbf{N}^{\dagger}\left(\boldsymbol{\tau}\right)\mathbf{d}\left(e^{j\omega T}\right). \qquad (15)$$

Finally, the Doppler-shifts and attenuation factors associated with $\mathcal{H}$ are determined from the vector $\mathbf{a}\left(e^{j\omega T}\right)$ by an additional use of the subspace methods.

Before proceeding to the recovery stage, we justify the assumption that $\mathbf{W}\left(e^{j\omega T}\right)$ can be stably inverted. To this end, observe from (10) that $\mathbf{W}\left(e^{j\omega T}\right)$ can be decomposed as

$$\mathbf{W}\left(e^{j\omega T}\right) = \boldsymbol{\Phi}\left(e^{j\omega T}\right)\mathbf{F}^{H}\boldsymbol{\Psi}\left(e^{j\omega T}\right), \qquad (16)$$

where $\boldsymbol{\Phi}\left(e^{j\omega T}\right)$ is a $p \times p$ diagonal matrix with ℓth diagonal element

$$\boldsymbol{\Phi}_{\ell\ell}\left(e^{j\omega T}\right) = \sqrt{p}(-1)^{\ell-1}e^{j\omega(\ell-1)T/p}, \qquad (17)$$

$\mathbf{F}$ is a p-point discrete Fourier transform (DFT) matrix with (ℓ, m)th element equal to

$$\mathbf{F}_{\ell m} = \frac{1}{\sqrt{p}}e^{-j\frac{2\pi}{p}(\ell-1)(m-1)}, \qquad (18)$$

and $\boldsymbol{\Psi}\left(e^{j\omega T}\right)$ is a $p \times p$ diagonal matrix whose mth diagonal element is given by

$$\boldsymbol{\Psi}_{mm}\left(e^{j\omega T}\right) = \frac{1}{T}S^{*}\left(\omega + \frac{2\pi}{T}(m - p/2 - 1)\right)G\left(\omega + \frac{2\pi}{T}(m - p/2 - \right. \qquad (19)$$

It can now be easily seen from the decomposition in (16) that, in order for $\mathbf{W}\left(e^{j\omega T}\right)$ to be stably invertible, each of the above three matrices has to be stably invertible. By construction, both $\boldsymbol{\Phi}\left(e^{j\omega T}\right)$ and $\mathbf{F}^{H}$ are stably invertible. The invertibility requirement on the diagonal matrix $\boldsymbol{\Psi}\left(e^{j\omega T}\right)$ leads to the following conditions on the pulse $g(t)$ and the kernel $s^{*}(-t)$ of the LPF.

Condition 1. *In order for $\boldsymbol{\Psi}\left(e^{j\omega T}\right)$ to be stably invertible, the continuous-time Fourier transform of $g(t)$ has to satisfy*

$$a \leq |G(\omega)| \leq b \quad a.e.\ \omega \in \mathcal{F} \qquad (20)$$

for some positive constants $a > 0$ and $b < \infty$.

Condition 2. *In order for $\boldsymbol{\Psi}\left(e^{j\omega T}\right)$ to be stably invertible, the continuous-time Fourier transform of the LPF $s^{*}(-t)$ has to satisfy*

$$c \leq |S(\omega)| \leq d \quad a.e.\ \omega \in \mathcal{F} \qquad (21)$$

for some positive constants $c > 0$ and $d < \infty$.

Condition 1 requires that the bandwidth $\mathcal{W}$ of the prototype pulse $g(t)$ has to satisfy

$$\mathcal{W} \geq \frac{2\pi}{T}p. \qquad (22)$$

In Section IV, we will derive additional conditions on the parameter p and combine them with (22) to obtain equivalent requirements on the time–bandwidth product of the input signal $x(t)$ that will ensure invertibility of the matrix $\mathbf{W}\left(e^{j\omega T}\right)$.

We conclude discussion of the sampling stage by pointing out that the decomposition in (16) also provides an

TABLE I
DELAY RECOVERY ALGORITHM

(i) Construct the matrix

$$\overline{\mathbf{R}}_{dd} = \frac{1}{M}\sum_{m=1}^{M}\sum_{n\in\mathbb{Z}}\mathbf{d}_m[n]\mathbf{d}_m^H[n],$$

where $\mathbf{d}_m$ is the $M = p/2$ length subvector which is given by

$$\mathbf{d}_m[n] = \begin{bmatrix} d_m[n] & d_{m+1}[n] & \dots & d_{m+M}[n] \end{bmatrix}^T.$$

(ii) Recover K_τ as the rank of $\overline{\mathbf{R}}_{dd}$.
(iii) Perform a *singular value decomposition* (SVD) of $\overline{\mathbf{R}}_{dd}$ and construct the matrix $\mathbf{E}_s$ consisting of the K_τ singular vectors corresponding to the K_τ nonzero singular values of $\overline{\mathbf{R}}_{dd}$ as its columns.
(iv) Compute the matrix $\mathbf{\Phi} = \mathbf{E}_{\mathbf{s}\downarrow}^\dagger \mathbf{E}_{s\uparrow}$, where $\mathbf{E}_{s\uparrow}$ and $\mathbf{E}_{s\downarrow}$ denote the submatrices extracted from $\mathbf{E}_s$ by removing its first row and its last row, respectively.
(v) Compute the eigenvalues of $\mathbf{\Phi}$, $\lambda_i, i = 1, 2, \dots, K_\tau$.
(vi) Recover the unknown delays as follows: $\tau_i = -\frac{T}{2\pi}\arg(\lambda_i)$.

efficient way to implement the digital-correction filter bank $\mathbf{W}^{-1}\left(e^{j\omega T}\right)$. This is because (16) implies that

$$\mathbf{W}^{-1}\left(e^{j\omega T}\right) = \mathbf{\Psi}^{-1}\left(e^{j\omega T}\right)\mathbf{F}\mathbf{\Phi}^{-1}\left(e^{j\omega T}\right). \quad (23)$$

Therefore the implementation of $\mathbf{W}^{-1}\left(e^{j\omega T}\right)$ can be done in three stages, where each stage corresponds to one of the three matrices in (23). Specifically, define the set of digital filters $\{\phi_\ell[n]\}$ and $\{\psi_\ell[n]\}$ as

$$\phi_\ell[n] = \text{IDTFT}\left\{\mathbf{\Phi}_{\ell\ell}^{-1}\left(e^{j\omega T}\right)\right\}[n], \quad 1 \le \ell \le p \quad (24)$$

and

$$\psi_\ell[n] = \text{IDTFT}\left\{\mathbf{\Psi}_{\ell\ell}^{-1}\left(e^{j\omega T}\right)\right\}[n], \quad 1 \le \ell \le p, \quad (25)$$

where IDTFT denotes the inverse DTFT operation. The first correction stage involves filtering the sequences $\{c_\ell[n]\}$ using the set of filters $\{\phi_\ell[n]\}$. Next, multiplication with the DFT matrix $\mathbf{F}$ can be efficiently implemented by applying the fast Fourier transform (FFT) to the outputs of the filters in the first stage. Finally, the third correction stage involves filtering the resulting sequences using the set of filters $\{\psi_\ell[n]\}$ to get the desired sequences $\{d_\ell[n]\}$. This last correction stage can be interpreted as an equalization step that compensates for the non-flatness of the frequency responses of the prototype pulse and the kernel of the LPF. A detailed schematic representation of the sampling stage, which is based on the preceding interpretation of $\mathbf{W}^{-1}\left(e^{j\omega T}\right)$, is provided in Fig. 2.

B. The Recovery Stage

We conclude this section by describing in detail the recovery stage, which—as noted earlier—consists of two steps. In the first step, we rely on subspace methods to recover the delays τ from $\mathbf{d}[n]$ [cf. (14)]. In the second step, we make use of the recovered delays to obtain the Doppler-shifts and attenuation factors associated with each of the delays.

1) Recovery of the Delays: In order to recover τ from $\mathbf{d}[n]$, we rely on the approach advocated in [10] and make use of the well-known ESPRIT algorithm [28] together with an additional smoothing stage [29]. The exact algorithm is given in Table I; we refer the reader to [10], [28] for details.

2) Recovery of the Doppler-Shifts and Attenuation Factors: Once the unknown delays are found, we can recover the vectors $\mathbf{a}[n]$ through the frequency relation (15). Next, define for each delay τ_i, the set of corresponding Doppler-shifts

$$\boldsymbol{\nu}_i = \{\nu_{ij}, \ j = 1, \dots, K_{\nu,i}\} \quad (26)$$

and recall that the ith element of $\mathbf{a}[n]$ is given by (6). We can therefore write the N-length sequence $\{a_i[n]\}$ for each index i in the following matrix–vector form

$$\mathbf{a}_i = \mathbf{X}\mathbf{R}(\boldsymbol{\nu}_i)\boldsymbol{\alpha}_i, \quad (27)$$

where $\mathbf{a}_i$ is a length-N vector whose nth element is $a_i[n]$, $\mathbf{X}$ is an $N \times N$ diagonal matrix whose nth diagonal element is given by x_n, $\mathbf{R}(\boldsymbol{\nu}_i)$ is an $N \times K_{\nu,i}$ Vandermonde matrix with (n, j)th element $e^{j2\pi\nu_{ij}nT}$, and $\boldsymbol{\alpha}_i$ is length-$K_{\nu,i}$ vector with jth element α_{ij}. The matrix $\mathbf{X}$ in (27) can be inverted under the assumption that the sequence $\{x_n\}$ satisfies $|x_n| > 0$ for every $n = 0, \dots, N - 1$, resulting in

$$\tilde{\mathbf{a}}_i = \mathbf{R}(\boldsymbol{\nu}_i)\boldsymbol{\alpha}_i, \quad (28)$$

where $\tilde{\mathbf{a}}_i = \mathbf{X}^{-1}\boldsymbol{a}_i$. From a simple inspection, we can express the elements of $\tilde{\mathbf{a}}_i$ as

$$\tilde{a}_i[n] = \sum_{j=1}^{K_{\nu,i}} \alpha_{ij}e^{j2\pi\nu_{ij}nT}, \quad 0 \le n \le N - 1. \quad (29)$$

It is now easy to see from this representation that recovery of the Doppler-shifts from the sequences $\{\tilde{a}_i[n]\}$ is equivalent to the problem of recovering distinct frequencies from a (weighted) sum of complex exponentials. In the context of our problem, for each fixed index i, the frequency of the jth exponential is $\omega_{ij} = 2\pi\nu_{ij}nT$ and its amplitude is α_{ij}.

Fortunately, the problem of recovering frequencies from a sum of complex exponentials has been studied extensively in the literature and various strategies exist for solving this problem (see [14] for a review). One of these techniques that has gained interest recently, especially in the literature on finite rate of innovation [30]–[33], is the *annihilating-filter* method. The annihilating-filter approach, in contrast to some of the other techniques, allows the recovery of the frequencies associated with the ith index even at the critical value of $N = 2K_{\nu,i}$. On the other hand, subspace methods such as ESPRIT [12], matrix-pencil algorithm [13], and the Tufts and Kumaresan approach [11] are generally more robust to noise but also require more samples than $2K_{\nu,i}$. Once the Doppler-shifts for each index i have been recovered then, since $\mathbf{R}^\dagger(\boldsymbol{\nu}_i)\mathbf{R}(\boldsymbol{\nu}_i) = \mathbf{I}$ because of the requirement that $N \ge 2K_{\nu,i}$, the attenuation factors $\{\alpha_{ij}\}$ are determined as

$$\boldsymbol{\alpha}_i = \mathbf{R}^\dagger(\boldsymbol{\nu}_i)\tilde{\mathbf{a}}_i, \quad i = 1, \dots, K_\tau. \quad (30)$$

IV. SUFFICIENT CONDITIONS FOR IDENTIFICATION

Our focus in Section III was on developing a recovery algorithm for the identification of ULSs. We now turn to specify conditions that guarantee that the proposed procedure recovers the set of triplets $\{(\tau_i, \nu_{ij}, \alpha_{ij})\}$ that describe the parametric ULS $\mathcal{H}$. We present these requirements in terms of equivalent conditions on the time–bandwidth product $\mathcal{TW}$

of the input signal $x(t)$. This is a natural way to describe the performance of system identification schemes since $\mathcal{T}\mathcal{W}$ roughly defines the number of temporal degrees of freedom available for estimating $\mathcal{H}$ [8].

To begin with, recall that the recovery stage involves first determining the unknown delays $\boldsymbol{\tau}$ from the set of equations given by (14) (cf. Section III-B). Therefore, to ensure that our algorithm successfully returns the parameters of $\mathcal{H}$, we first need to provide conditions that guarantee a unique solution to (14). To facilitate the forthcoming analysis, we let $\mathbf{d}\,[\Lambda] = \{\mathbf{d}\,[n]\,, n \in \mathbb{Z}\}$ and $\mathbf{b}\,[\Lambda] = \{\mathbf{b}\,[n]\,, n \in \mathbb{Z}\}$ denote the set of all vectors $\mathbf{d}\,[n]$ and $\mathbf{b}\,[n]$, respectively. Using this notation, we can rewrite (14) as

$$\mathbf{d}\,[\Lambda] = \mathbf{N}\,(\boldsymbol{\tau})\,\mathbf{b}\,[\Lambda]\,. \tag{31}$$

We now leverage the analysis carried out in [10] to provide sufficient conditions for a unique solution to (31); see [10] for a formal proof.

Proposition 1. *If $\left(\bar{\boldsymbol{\tau}}, \bar{\mathbf{b}}\,[\Lambda] \neq \mathbf{0}\right)$ solves (31) and if*

$$p > 2K_\tau - dim\left(span\left(\bar{\mathbf{b}}\,[\Lambda]\right)\right) \tag{32}$$

then $\left(\bar{\boldsymbol{\tau}}, \bar{\mathbf{b}}\,[\Lambda]\right)$ is the unique solution of (31). Here, $span\left(\bar{\mathbf{b}}\,[\Lambda]\right)$ is used to denote the subspace of minimal dimensions that contains $\bar{\mathbf{b}}\,[\Lambda]$.

Proposition 1 suggests that a unique solution to (31)—and, by extension, unique recovery of the set of delays $\boldsymbol{\tau}$—is guaranteed through a proper selection of the parameter p. In particular, since $dim\left(span\left(\bar{\mathbf{b}}\,[\Lambda]\right)\right)$ is a positive number in general, we have from Proposition 1 that $p \geq 2K_\tau$ is a sufficient condition for unique recovery of $\boldsymbol{\tau}$ and $\mathbf{b}\,[\Lambda]$. From Condition 1 in Section III, we have seen that the parameter p effectively determines the minimum bandwidth $\mathcal{W}$ of the prototype pulse [cf. (22)]. Combining the condition $p \geq 2K_\tau$ and the one obtained earlier in (22) leads to the following sufficient condition on the bandwidth of the input signal for unique recovery of $\boldsymbol{\tau}$ and $\mathbf{b}\,[\Lambda]$:

$$\mathcal{W} \geq \frac{4\pi K_\tau}{T}. \tag{33}$$

The second step in the recovery stage involves recovering the Doppler-shifts and attenuation factors (cf. Section III-B). We now investigate the conditions required for unique recovery of the Doppler-shifts. Recall that the vectors $\mathbf{b}[n]$ and $\mathbf{a}[n]$ are related to each other by the invertible frequency relation (12). Therefore, since the diagonal matrix $\mathbf{D}\left(e^{j\omega T}, \boldsymbol{\tau}\right)$ is invertible and completely specified by $\boldsymbol{\tau}$, the condition given in (33) also guarantees unique recovery of the vectors $\mathbf{a}[n]$ from $\mathbf{b}[n]$. Further, it can be easily verified that the matrix $\mathbf{R}(\boldsymbol{\nu}_i)$ in (28) has the same parametric structure as that required by Proposition 1. We can therefore once again appeal to the results of Proposition 1 in providing conditions for unique recovery of the Doppler-shifts $\{\boldsymbol{\nu}_i\}$ from the vectors $\{\tilde{\mathbf{a}}_i\}$ [cf. (28)]. To that end, we interchange p with N and K_τ with $K_{\nu,i}$ in Proposition 1 and use the fact that $dim\left(span\left(\boldsymbol{a}_i\right)\right) = 1$ (since it is a nonzero vector). Therefore, by making use of Proposition 1, we conclude that a sufficient condition for

unique recovery of $\boldsymbol{\nu}_i$ from (28) is

$$N \geq 2K_{\nu,i}. \tag{34}$$

Condition (34) is intuitive in the sense that there are $2K_{\nu,i}$ unknowns in (28) ($K_{\nu,i}$ unknown Doppler-shifts and $K_{\nu,i}$ unknown attenuation factors) and therefore at least $2K_{\nu,i}$ equations are required to solve for these unknown parameters. Finally, since we need to ensure unique recovery of the Doppler-shifts and attenuation factors for each distinct delay τ_i, we have the condition

$$N \geq 2 \max_i K_{\nu,i} \tag{35}$$

which trivially ensures that (34) holds for every $i = 1, \dots, K_\tau$. We summarize these results in the following theorem.

Theorem 2 (Sufficient Conditions for System Identification). *Suppose that $\mathcal{H}$ is a parametric ULS that is completely described by a total of $K = \sum_{i=1}^{K_\tau} K_{\nu,i}$ triplets $(\tau_i, \nu_{ij}, \alpha_{ij})$. Then, irrespective of the distribution of $\{(\tau_i, \nu_{ij})\}$ within the delay–Doppler space, the recovery procedure specified in Section III with samples taken at $\{t = 2n\pi/\mathcal{W}\}$ uniquely identifies $\mathcal{H}$ from a single observation $\mathcal{H}(x(t))$ as long as the system satisfies [A1]–[A3], the probing sequence $\{x_n\}$ remains bounded away from zero in the sense that $|x_n| > 0$ for every $n = 0, \dots, N - 1$, and the time–bandwidth product of the (known) input signal $x(t)$ satisfies the condition*

$$\mathcal{T}\mathcal{W} \geq 8\pi K_\tau K_{\nu,max} \tag{36}$$

where $K_{\nu,max} = \max_i K_{\nu,i}$ is the maximum number of Doppler-shifts associated with any one of the distinct delays.

Proof: Recall from the previous discussion that the delays, Doppler-shifts, and attenuation factors associated with $\mathcal{H}$ can be uniquely recovered as long as $N \geq 2K_{\nu,max}$, $\mathcal{W} \geq \frac{4\pi K_\tau}{T}$, and $p \geq 2K_\tau$. Now take $N = \frac{\mathcal{T}\mathcal{W}}{4\pi K_\tau}$ and note that under the assumption $\mathcal{T}\mathcal{W} \geq 8\pi K_\tau K_{\nu,max}$, we trivially have $N \geq 2K_{\nu,max}$. Further, since $\mathcal{T} = NT$ and since sampling rate of $2\pi/\mathcal{W}$ implies $p = 2\pi/(\mathcal{W}T)$, we also have that $\mathcal{W} = \frac{4\pi K_\tau}{T} \Rightarrow p = 2K_\tau$, completing the proof. ∎

Theorem 2 implicitly assumes that K_τ (or an upper bound on K_τ) and $K_{\nu,max}$ (or an upper bound on $K_{\nu,max}$) are known at the transmitter side. We explore this point in further detail in Section V and numerically study the effects of "model-order mismatch" on the robustness of the proposed recovery procedure. It is also instructive (especially for comparison purposes with related work such as [9], [17]) to present a weaker version of Theorem 2 that only requires knowledge of the total number of delay–Doppler pairs K.

Corollary 1 (Weaker Sufficient Conditions for System Identification). *Suppose that the assumptions of Theorem 2 hold. Then the recovery procedure specified in Section III with samples taken at $\{t = 2n\pi/\mathcal{W}\}$ uniquely identifies $\mathcal{H}$ from a single observation $\mathcal{H}(x(t))$ as long as the time–bandwidth product of the known input signal $x(t)$ satisfies the condition $\mathcal{T}\mathcal{W} \geq 2\pi(K + 1)^2$.*

Proof: This corollary is a simple consequence of Theorem 2 and the fact that $K_\tau K_{\nu,max} \leq \frac{(K+1)^2}{4}$. To prove the

latter fact, note that for any fixed K and K_τ, we always have $K_{\nu,max} \leq K - (K_\tau - 1)$. Indeed, if $K_{\nu,max}$ were greater than $K - (K_\tau - 1)$ then either $\sum_{i=1}^{K_\tau} K_{\nu,i} > K$ or there exists an i such that $K_{\nu,i} = 0$, both of which are contradictions. Consequently, for any fixed K, we have that

$$K_\tau K_{\nu,max} \leq -K_\tau^2 + (K+1)K_\tau \qquad (37)$$

and since the maximum of $-K_\tau^2 + (K+1)K_\tau$ occurs at $K_\tau = \frac{K+1}{2}$, we get $K_\tau K_{\nu,max} \leq \frac{(K+1)^2}{4}$. ∎

V. DISCUSSION

In Sections III and IV, we proposed and analyzed a polynomial-time recovery procedure that ensures identification of parametric ULSs under certain conditions. In particular, one of the key contributions of the preceding analysis is that it parlays a key sub-Nyquist sampling result of [10] into conditions on the time–bandwidth product, $\mathcal{TW}$, of the input signal $x(t)$ that guarantee identification of arbitrary linear systems as long as they are sufficiently underspread. Specifically, in the parlance of system identification, Corollary 1 states that the recovery procedure of Section III achieves *infinitesimally-fine resolution* in the delay–Doppler space as long as the temporal degrees of freedom available to excite a ULS are on the order of $\Omega(K^2)$. In addition, we carry out extensive numerical experiments in Section VII, which confirm that—as long as the condition $\mathcal{TW} \geq 2\pi(K+1)^2$ is satisfied—the ability of the proposed procedure to distinguish between (resolve) closely spaced delay–Doppler pairs is primarily a function of the signal-to-noise ratio (SNR) and its performance degrades gracefully in the presence of noise. In order to best put the significance of our results into perspective, it is instructive to compare them with corresponding results in recent literature. We then discuss an application of these results to super-resolution target detection using radar in Section VI.

There exists a large body of existing work—especially in the communications and radar literature—treating identification of parametric ULSs; see, e.g., [2], [9], [15]–[19]. One of the approaches that is commonly taken in many of these works, such as in [9], [15]–[18], is to quantize the delay–Doppler space (τ, ν) by assuming that both τ_i and ν_{ij} lie on a grid. The following theorem is representative of some of the known results in this case.[3]

Theorem 3 ([9], [17])**.** *Suppose that $\mathcal{H}$ is a parametric ULS that is completely described by a total of $K = \sum_{i=1}^{K_\tau} K_{\nu,i}$ triplets $(\tau_i, \nu_{ij}, \alpha_{ij})$. Further, let the delays and the Doppler-shifts of the system be such that $\tau_i = r_i \mathcal{W}^{-1}$ and $\nu_{ij} = \ell_{ij} \mathcal{T}^{-1}$ for $r_i \in \mathbb{Z}_+$ and $\ell_{ij} \in \mathbb{Z}$. Then $\mathcal{H}$ can be identified in polynomial-time from a single observation $\mathcal{H}(x(t))$ as long as the system satisfies [A1]–[A3] and the time–bandwidth product of the input signal $x(t)$ satisfies $\mathcal{TW} = \Omega(K^2 / \log \mathcal{TW})$.*

There are two conclusions that can be immediately drawn from Theorem 3. First, both [9], [17] require about the same scaling of the temporal degrees of freedom as that required by Corollary 1: $\mathcal{TW} \approx \Omega(K^2)$. Second, the resolution of the

recovery procedures proposed in [9], [17] is limited to $\mathcal{W}^{-1}$ in the delay space and $\mathcal{T}^{-1}$ in the Doppler space because of the assumption that $\tau_i = r_i \mathcal{W}^{-1}$ and $\nu_{ij} = \ell_{ij} \mathcal{T}^{-1}$.[4] Similarly, in another related recent paper [18], two recovery procedures are proposed that have been numerically shown to uniquely identify $\mathcal{H}$ as long as $\mathcal{TW} \gg 1$ and each (τ_i, ν_{ij}) corresponds to one of the points in the quantized delay–Doppler space with resolution proportional to $\mathcal{W}^{-1}$ and $\mathcal{T}^{-1}$ in the delay space and the Doppler space, respectively. Note that the assumption of a quantized delay–Doppler space can have unintended consequences in certain applications and we carry out a detailed discussion of this issue in the next section in the context of radar target detection.

Finally, the work in [19] leverages some of the results in DOA estimation to propose a scheme for the identification of linear systems of the form (2) without requiring that $\tau_i = r_i \mathcal{W}^{-1}$ and $\nu_{ij} = \ell_{ij} \mathcal{T}^{-1}$. Nevertheless, our results differ from those in [19] in three important respects. First, we explicitly state the relationship between the time–bandwidth product $\mathcal{TW}$ of the input signal $x(t)$ and the number of delay–Doppler pairs $K = \sum_{i=1}^{K_\tau} K_{\nu,i}$ that guarantees recovery of the system response by studying the sampling and recovery stages of our proposed recovery procedure. On the other hand, the method proposed in [19] assumes the sampling stage to be given and, as such, fails to make explicit the connection between the time–bandwidth product of $x(t)$ and the number of delay-Doppler pairs. Second, the algorithms proposed in [19] have exponential complexity, since they require exhaustive searches in a K-dimensional space, which can be computationally prohibitive for large-enough values of K. Last, but not the least, recovery methods proposed in [19] are guaranteed to work as long as there are no more than two delay–Doppler pairs having the same delay, $\max_i K_{\nu,i} \leq 2$, and the system output is observed by an M-element antenna array with $M \gtrsim K$. In contrast, our recovery algorithm does not impose any restrictions on the distribution of $\{(\tau_i, \nu_{ij})\}$ within the delay-Doppler space and is guaranteed to work with a single observation of the system output.

VI. APPLICATION: SUPER-RESOLUTION RADAR

We have established in Section IV that the polynomial-time recovery procedure of Section III achieves infinitesimally-fine resolution in the delay–Doppler space under mild assumptions on the temporal degrees of freedom of the input signal. This makes the proposed algorithm extremely useful for application areas in which the system performance depends critically on the ability to resolve closely spaced delay–Doppler pairs. In particular, our method can be used for super-resolution target detection using radar. This is because the noiseless, clutter-free received signal in the case of monostatic radars is exactly of the form (1) with each triplet $(\tau_k, \nu_k, \alpha_k)$ corresponding to an echo of the radar waveform $x(t)$ from a distinct target

---

[3]It is worth mentioning here that a somewhat similar result was also obtained independently in [34] in an abstract setting.

[4]Note that there is also a Bayesian variant of Theorem 3 in [9] that requires $\mathcal{TW} \approx \Omega(K)$ under the assumption that $\mathcal{H}$ has a uniform statistical prior over the quantized delay–Doppler space. A somewhat similar Bayesian variant of Corollary 1 can also be obtained by trivially extending the results of this paper to the case when $\mathcal{H}$ is assumed to have a uniform statistical prior over the *non-quantized* delay–Doppler space.
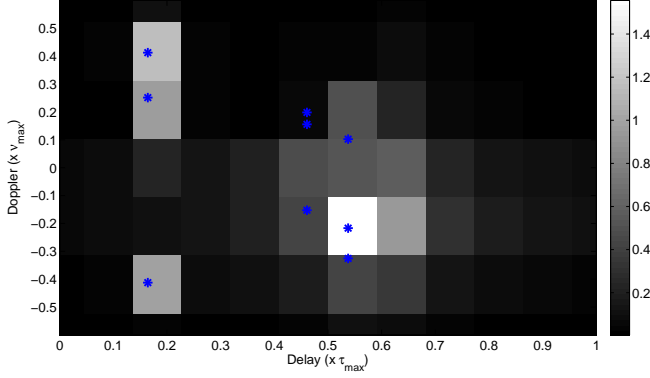
Fig. 3. Quantized representation of nine targets (represented by ∗) in the delay–Doppler space with $\tau_{max} = 10$ μs and $\nu_{max} = 10$ kHz. The quantized delay–Doppler approximation of the targets corresponds to $\mathcal{W} = 1.2$ MHz and $\mathcal{T} = 0.48$ ms.

[2].[5] The fact that our recovery procedure allows to identify arbitrary parametric ULSs, therefore, enables us to distinguish between multiple targets even if their radial positions are quite close to each other and/or their radial velocities are similar—the so-called super-resolution detection of targets.

On the other hand, note that apart from the fact that none of the methods based on the assumption of a quantized delay–Doppler space can ever carry out super-resolution target detection, a major drawback of the radar target detection approach in works such as [9], [18] is that targets in the real-world do not in fact correspond to points in the quantized delay–Doppler space, which causes *leakage* of their energies in the quantized space. In order to elaborate further on this point, define $L = \lceil W\tau_{max} \rceil$ and $M = \lceil T\nu_{max}/2 \rceil$ and note that (canonical) quantization corresponds to transforming the $\mathcal{C} = [0, \tau_{max}] \times [-\nu_{max}/2, \nu_{max}/2]$ continuous delay–Doppler space into a $\mathcal{Q} = \{0, \ldots, L\} \times \{-M/2, \ldots, M/2\}$ two-dimensional quantized grid, which in turn transforms the received signal $\mathcal{H}(x(t))$ at the radar into [35], [36, Chapter 4]

$$\tilde{\mathcal{H}}(x(t)) \approx \sum_{\ell=0}^{L} \sum_{m=-M}^{M} \tilde{\alpha}_{\ell m} x(t - \tilde{\tau}_\ell) e^{j2\pi\tilde{\nu}_m t} \quad (38)$$

where $\tilde{\alpha}_{\ell m} = \sum_{i=1}^{K_\tau} \sum_{j=1}^{K_{\nu,i}} \alpha_{ij} e^{j\pi(m-\mathcal{T}\nu_{ij})} \text{sinc}(m - \mathcal{T}\nu_{ij})\text{sinc}(\ell - TW\tau_i)$ and the quantized delay–Doppler pairs $(\tilde{\tau}_\ell, \tilde{\nu}_m) \in \mathcal{Q}$. It is now easy to conclude from (38) that, unless the original targets (delay–Doppler pairs) happen to lie in $\mathcal{Q}$, most of the attenuation factors $\{\tilde{\alpha}_{\ell m}\}$ will be nonzero because of the sinc kernels—the so-called "leakage effect." This has catastrophic implications for target detection using radar since leakage makes it impossible to reliably identify the original set of delays and Doppler-shifts. This limitation of target-detection methods that are based on the assumption of a quantized delay–Doppler space is also depicted in Fig. 3 for the case of nine hypothetical targets. The figure illustrates that each of the nine non-quantized targets not only
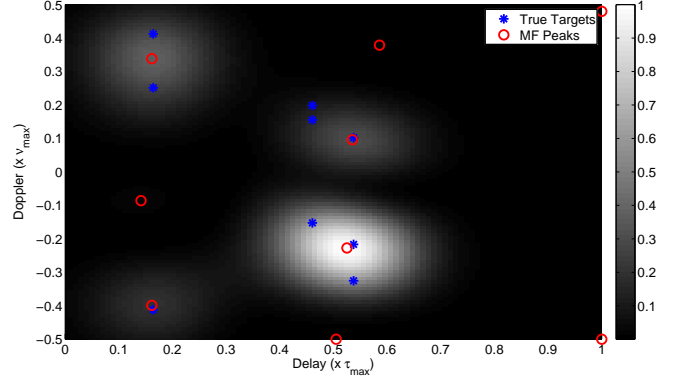
(a)
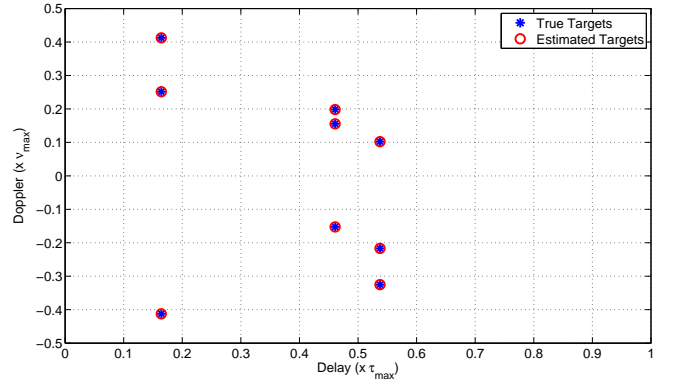
(b)

Fig. 4. Comparison between the target-detection performance of matched-filtering and our proposed recovery procedure for the case of nine targets (represented by ∗) in the delay–Doppler space with $\tau_{max} = 10$ μs, $\nu_{max} = 10$ kHz, $\mathcal{W} = 1.2$ MHz, and $\mathcal{T} = 0.48$ ms. The sequence $\{x_n\}$ corresponds to a random binary (± 1) sequence with $N = 48$, the pulse $g(t)$ is designed to have a nearly-flat frequency response in the working spectral band $\mathcal{F}$, and the pulse repetition interval is taken to be $T = 10$ μs. (a) Target detection by matched-filtering the received signal $\mathcal{H}(x(t))$ with the input signal $x(t)$. (b) Target detection using the proposed recovery procedure with $p = 12$.

contributes energy to its own $(\tilde{\tau}_\ell, \tilde{\nu}_m)$ in $\mathcal{Q}$ but also leaks its energy to the nearby points in the quantized space.

Owing to the fact that leakage can cause missed detections and false alarms, conventional radar literature in fact tends to focus only on recovery procedures that do no impose any structure on the distribution of $\{(\tau_i, \nu_{ij})\}$ within the delay–Doppler space. The most commonly used approach in the radar signal processing literature corresponds to matched-filtering (MF) the received signal with the input signal $x(t)$ in the delay–Doppler space [2]. The MF output $\chi(\tau, \nu)$ takes the form

$$\chi(\tau, \nu) = \int_{\mathbb{R}} \mathcal{H}(x(t))x^*(t - \tau)\exp(-j2\pi\nu t)dt = \sum_{i=1}^{K_\tau} \sum_{j=1}^{K_{\nu,i}} \alpha_{ij} \mathcal{A}(\tau - \tau_i, \nu \quad (39)$$

where $\mathcal{A}(\tau, \nu) = \int_{\mathbb{R}} x(t)x^*(t - \tau)\exp(-j2\pi\nu t)dt$ is termed the Woodward's *ambiguity function* of $x(t)$. It can be easily deduced from (39) that the resolution of the MF-based recovery procedure is tied to the *support* of the ambiguity
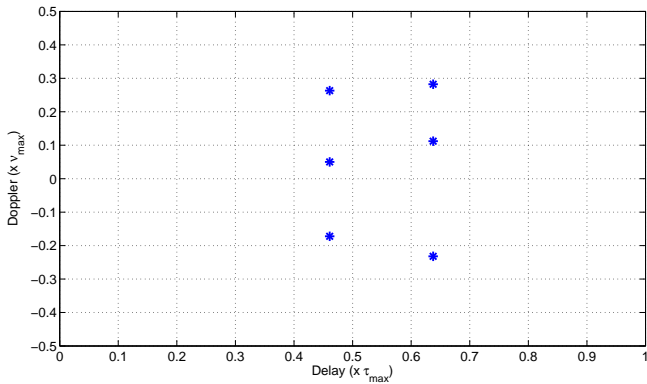
Fig. 5. Delay–Doppler representation of a parametric ULS $\mathcal{H}$ corresponding to $K = 6$ delay–Doppler pairs with $\tau_{max} = 10$ μs and $\nu_{max} = 10$ kHz.

Fig. 6. Mean-squared error of the estimated delays and Doppler-shifts as a function of the signal-to-noise ratio.

function in the delay–Doppler space. Ideally, one would like to have $\mathcal{A}(\tau, \nu) = \delta(\tau)\delta(\nu)$ for super-resolution detection of targets but two fundamental properties of ambiguity functions, namely, $|\mathcal{A}(0,0)|^2 = \int |x(t)|^2 dt$ and $\iint |\mathcal{A}(\tau, \nu)|^2 d\tau d\nu = \int |x(t)|^2 dt$, dictate that no real-world signal $x(t)$ can yield infinitesimally-fine resolution in this case either [2]. In fact, the resolution of MF-based recovery techniques also tends to be on the order of $\mathcal{W}^{-1}$ and $\mathcal{T}^{-1}$ in the delay space and the Doppler space, respectively, which severely limits their ability to distinguish between two closely-spaced targets in the delay–Doppler space. This inability of MF-based methods to resolve closely-spaced delay–Doppler pairs is depicted in Fig. 4. This figure compares the target-detection performance of MF and the recovery procedure proposed in this paper for the case of nine closely-spaced targets. It is easy to see from Fig. 4(a) that matched-filtering the received signal $\mathcal{H}(x(t))$ with the input signal $x(t)$ gives rise to peaks that are not centered at the true targets for a majority of the targets. On the other hand, Fig. 4(b) illustrates that our recovery procedure correctly identifies the locations of all nine of the targets in the delay–Doppler space.

VII. NUMERICAL EXPERIMENTS

In this section, we explore various issues using numerical experiments that were not treated theoretically earlier in the paper. These include robustness of our method in the presence of noise and the effects of truncated digital filters, use of finite number of samples, choice of probing sequence $\{x_n\}$, and model-order mismatch on the recovery performance. Throughout this section, the numerical experiments correspond to a parametric ULS $\mathcal{H}$ that is described by a total of $K = 6$ delay–Doppler pairs with $K_\tau = 2$ and $K_{\nu,1} = K_{\nu,2} = 3$. The locations of these pairs in the delay–Doppler space are given by Fig. 5, while the attenuation factors associated with each of the six delay–Doppler pairs are taken to have unit amplitudes and random phases.

In order to identify $\mathcal{H}$, we design the prototype pulse $g(t)$ to have a constant frequency response over the working spectral band $\mathcal{F} = \left[-\frac{\pi}{T}p, \frac{\pi}{T}p\right]$ with $p = 4$ and $T = 10$ μs, that is, $G(\omega) \approx 1$ when $\omega \in \mathcal{F}$ and $G(\omega) \approx 0$ when $\omega \notin \mathcal{F}$. In
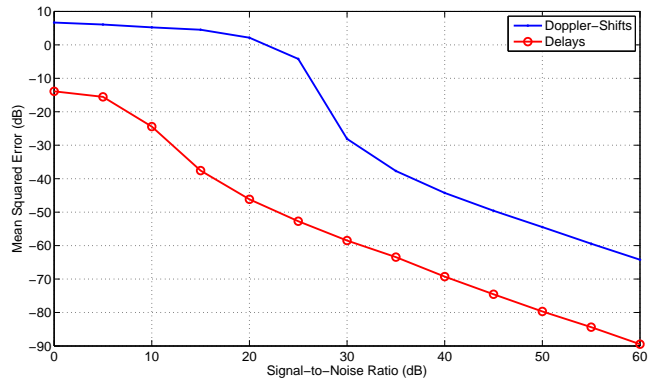
other words, the input signal $x(t)$ is chosen to have bandwidth $\mathcal{W} = \frac{8\pi}{T}$. In addition, unless otherwise noted, we use a probing sequence $\{x_n\}$ corresponding to a random binary (± 1) sequence with $N = 30$, which leads to a time–bandwidth product of $\mathcal{T}\mathcal{W} \approx 240\pi$. Note that the chosen time–bandwidth product here is more than the lower bound of Theorem 2 by a factor of 5 so as to increase the robustness to noise. Also, unless otherwise stated, all experiments in the following use an ideal (flat) LPF as the sampling filter (cf. Fig. 2). We use the ESPRIT method described in Section III for recovery of the delays and the matrix-pencil method [13] for recovery of the corresponding Doppler-shifts. Given the rich history of these two subspace methods, there exist many standard techniques in the literature (see, e.g., [37], [38]) for providing them with reliable estimates of the model orders in the presence of noise. As such, we assume in the following that both these methods have access to correct values of K_τ and $K_{\nu,i}$'s. Finally, the performance metrics that we use in this section are the (normalized) mean-squared error (MSE) of the estimated delays and Doppler-shifts (averaged over 100 noise realizations), defined as

$$e^2_{delay} = \frac{1}{2}\sum_{i=1}^{2}\left[(\hat{\tau}_i - \tau_i)/\tau_{max}\right]^2, \tag{40}$$

and

$$e^2_{Doppler} = \frac{1}{6}\sum_{i=1}^{2}\sum_{j=1}^{3}\left[(\hat{\nu}_{ij} - \nu_{ij})/\nu_{max}\right]^2, \tag{41}$$

where $\hat{\tau}_i$ and $\hat{\nu}_{ij}$ denote the estimated delays and Doppler-shifts, respectively.

1) Robustness to Noise: We first examine the robustness of our method when the received signal $\mathcal{H}(x(t))$ is corrupted by additive noise. The results of this experiment are shown in Fig. 6, which plots the MSE of the estimated delays and Doppler-shifts as a function of the SNR. It can be seen from the figure that the ability of the proposed procedure to resolve delay–Doppler pairs is primarily a function of the SNR and its performance degrades gracefully in the presence of noise.
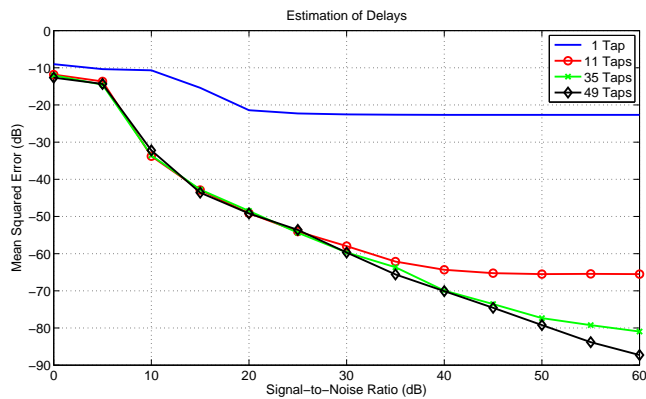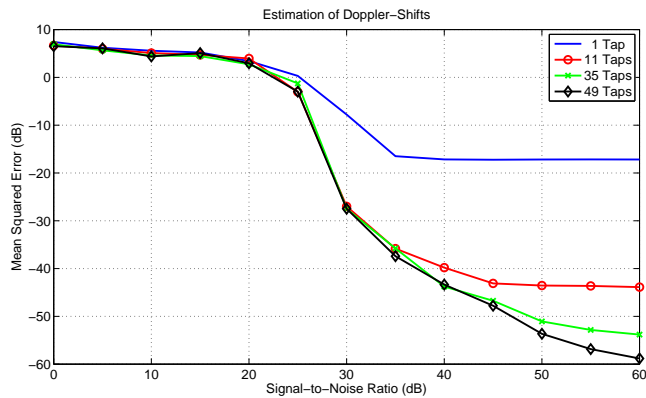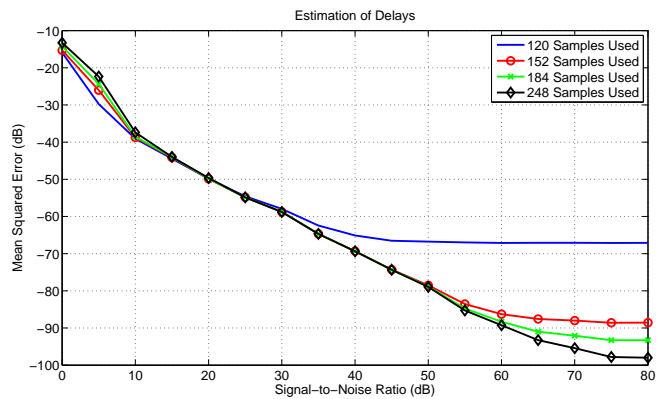
Fig. 7. Mean-squared error of the estimated delays and Doppler-shifts as a function of the signal-to-noise ratio for various lengths of the impulse responses of the filters.
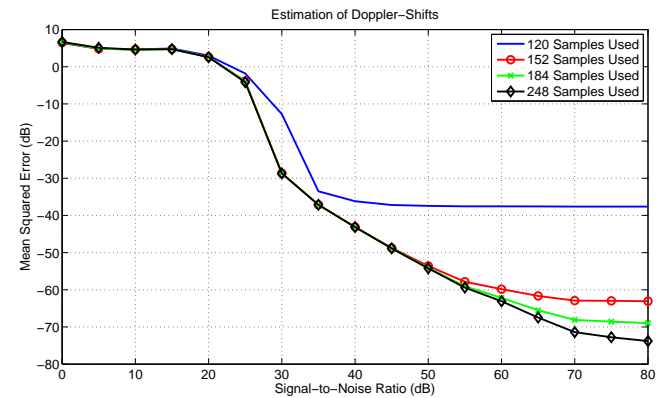
Fig. 8. Mean-squared error of the estimated delays and Doppler-shifts as a function of the signal-to-noise ratio for different numbers of samples collected at the output of the sampling filter (corresponding to an ideal low-pass filter).

2) Effects of Truncated Digital-Correction Filter Banks: Recall from Section III that our recovery method is composed of various digital-correction stages (see also Fig. 2). The filters used in these stages, which include $\{\phi_\ell[n]\}$ and $\{\psi_\ell[n]\}$, have infinite impulse responses in general so that their practical implementation requires truncation of their impulse responses. The truncated lengths of these filters also determine the (computational) delay and the computational load of the proposed procedure. Fig. 7 plots the MSE of the estimated delays (Fig. 7(a)) and Doppler-shifts (Fig. 7(b)) as a function of the SNR for various lengths of the impulse responses of the filters. There are two important insights that can be drawn from the results of this experiment. First, for a fixed length of the impulse responses, there is always some SNR beyond which the estimation error caused by the truncation of the impulse responses becomes more dominant than the error caused by the additive noise (as evident by the error floors in Fig. 7). Second, and perhaps most importantly, filters with 49 taps seem to provide good estimation accuracy up to an SNR of 60 dB, whereas filters with even just 35 taps yield good estimates at SNRs below 50 dB.

3) Effects of Finite Number of Samples: The sampling filter used at the front-end in Fig. 2 is bandlimited in nature and, therefore, has infinite support in the time domain. Conse-

quently, our sampling method theoretically requires collecting an infinite number of samples at the back-end of this filter. The next numerical experiment examines the effect of collecting a finite number of samples on the estimation performance. The results are reported in Fig. 8, which depicts the MSE of the estimated delays (Fig. 8(a)) and Doppler-shifts (Fig. 8(b)) as a function of SNR for different numbers of samples collected at the output of the sampling filter (corresponding to an ideal LPF). As in the case of truncation of digital-correction filter banks, there is always some SNR for every fixed number of samples beyond which the estimation error caused by the finite number of samples becomes more dominant than the error due to additive noise. Equally importantly, however, note that $x(t)$ in these experiments corresponds to a train of $N = 30$ prototype pulses. Therefore, under the assumption of $p = 4$ samples per pulse period T, it is clear that we require at least $N \cdot p = 120$ samples in total to represent just the input signal $x(t)$. On the other hand, Fig. 8 shows that collecting 248 samples, which is roughly twice the minimum number of samples required, provides good (delay and Doppler) estimation accuracy for SNRs up to 70 dB.

Finally, it is worth noting here that making use of an ideal LPF as the sampling filter requires collecting relatively more samples at the filter back-end due to the slowly-decaying
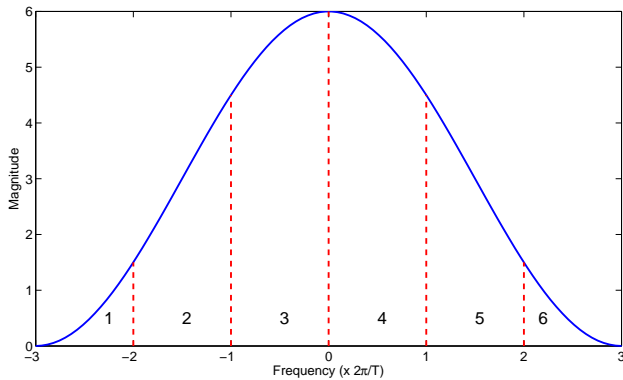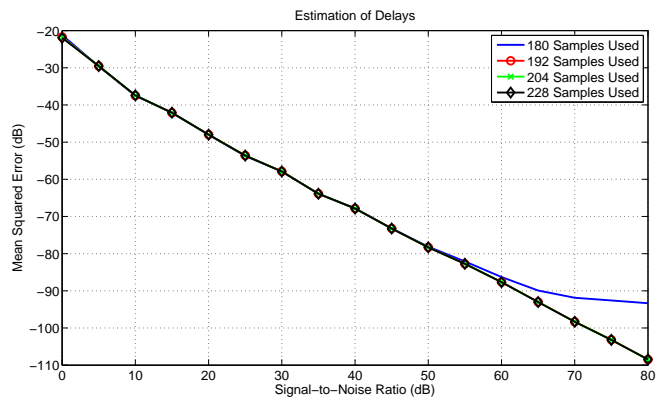
Fig. 9. Frequency response of a raised-cosine filter with roll-off factor 1.

(a)

(b)

Fig. 10. Mean-squared error of the estimated delays and Doppler-shifts as a function of the signal-to-noise ratio for different numbers of samples collected at the output of a raised-cosine sampling filter with roll-off factor 1.
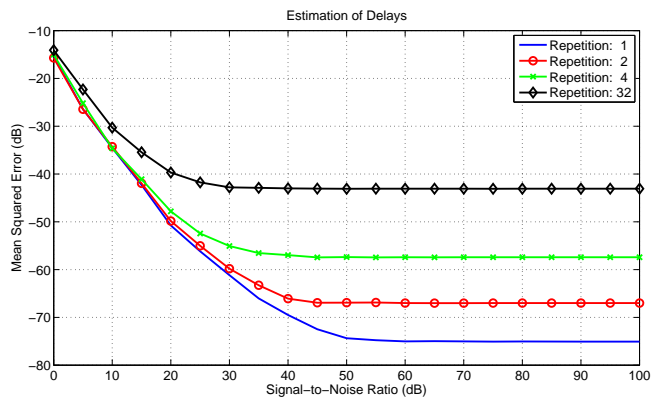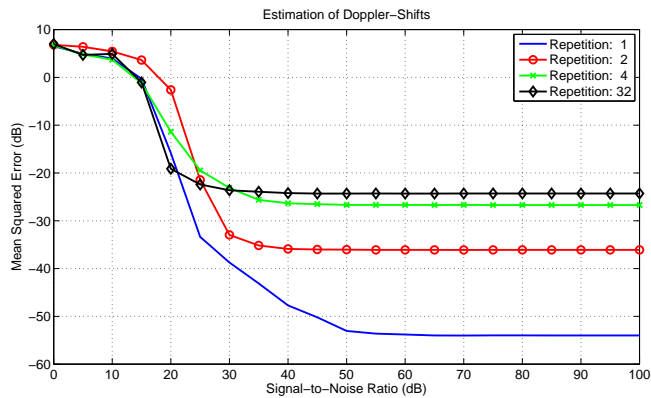
nature of the sinc kernel. Therefore, in order to reduce the number of samples required at the back-end of the sampling filter for reasonable estimation accuracy, we can instead make use of sampling filters whose (time-domain) kernels decay faster than the sinc kernel. One such possible choice is a raised-cosine filter with roll-off factor equal to 1, whose frequency response is given by $S(\omega) = \frac{p}{2T}\left(1 + \cos(\frac{T}{p}\omega)\right)$ when $\omega \in \mathcal{F}$ and $S(\omega) = 0$ when $\omega \notin \mathcal{F}$. It is a well-known fact (and can be easily checked) that this filter decays faster in the time domain than the sinc kernel. However, the main issue here is that raised-cosine filter does not satisfy Condition 2 in Section III, since its frequency response is not bounded away from zero at the ends of the spectral band $\mathcal{F}$ (see, e.g., Fig. 9).

However, we now show that this problem can be overcome by slightly increasing the sampling rate and the bandwidth requirement stated in Section IV. Specifically, note that Proposition 1 requires that the parameter p, which controls the minimal bandwidth of $x(t)$ and the sampling rate of our proposed procedure, satisfies $p \geq 4$ under the current simulation setup (since $K_\tau = 2$). We now instead choose $p = 6$ and argue that raised-cosine filter can be successfully used under this choice of p. To this end, recall from Section III that the function of the digital-correction filters $\psi_1[n]$ and $\psi_6[n]$ is to invert the frequency response of the sampling kernel corresponding to the frequency bands denoted by 1 and 6 in Fig. 9, respectively (under the assumption that the pulse $g(t)$ has a flat frequency response). In the case of a raised-cosine filter, however, we cannot compensate for the non-flat nature of these two bands since they are not bounded away from zero. Nevertheless, because of the fact that we are using $p = 6$, we can simply disregard channels 1 and 6 after the first digital-correction stage and work with the rest of the four channels (2-4) only. We make use of this insight to repeat the last numerical experiment using a raised-cosine filter and report the results in Fig. 10. It is easy to see from Fig. 10 that, despite increasing p to 6, raised-cosine filter performs better than an ideal LPF using fewer samples.

4) Effects of the Probing Sequence: Theorem 2 in Section IV stipulates that the choice of the probing sequence $\{x_n\}$ has no impact on the noiseless performance of the proposed recovery procedure as long as $|x_n| > 0$ for every
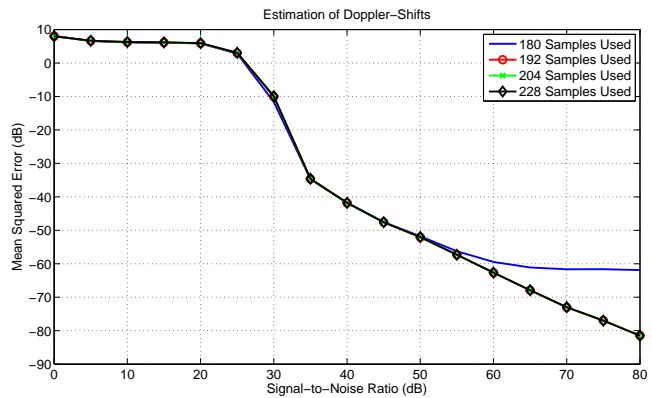
$n = 0, \ldots, N-1$. However, it is quite expected that $\{x_n\}$ will have an effect on the performance in the presence of noise and implementation issues related to truncated digital filters and use of finite number of samples. The next experiment examines this effect for four different choices of binary probing sequences of length $N = 32$ that periodically alternate between $+1$ and -1 every r entries. The results are reported in Fig. 11, which depicts the MSE of the estimated delays (Fig. 11(a)) and Doppler-shifts (Fig. 11(b)) as a function of the SNR for probing sequences with $r = 1, 2, 4$, and 32. We can draw two immediate conclusions from observing the results of this experiment. First, faster alternating probing sequences (in other words, sequences with higher frequency content) appear to provide better resilience against the truncation of digital filters and the use of finite number of samples. Second, the effect of the choice of probing sequence is less pronounced at low SNRs, since the error due to noise at low SNRs dominates the errors caused by other implementation imperfections.

5) Effects of Model-Order Mismatch: Our final numerical experiment studies the situation where the conditions of Theorem 2 do not exactly hold. To this end, we simulate identification of a parametric ULS with $K_\tau = 4$ delays. For the first 3 delays we take $K_{\nu,i} = 2$, $i = 1, 2, 3$, whereas we choose $K_{\nu,4} = 8$ for the last delay. Finally, we take

(a)

(b)

Fig. 11. Mean-squared error of the estimated delays and Doppler-shifts as a function of the signal-to-noise ratio for various probing sequences.

Fig. 12. Effects of model-order mismatch on the performance of the proposed recovery procedure corresponding to $\mathcal{H}$ with $K = 14$ delay–Doppler pairs.

VIII. CONCLUSION

In this paper, we revisited the problem of identification of parametric underspread linear systems that are completely described by a finite set of delays and Doppler-shifts. We established that sufficiently-underspread parametric linear systems are identifiable as long as the time–bandwidth product of the input signal is proportional to the square of the total number of delay–Doppler pairs. In addition, we concretely specified the nature of the input signal and the structure of a corresponding polynomial-time recovery procedure that enable identification of parametric underspread linear systems. Extensive simulation results confirm that—as long as the time–bandwidth product of the input signal satisfies the requisite conditions—the proposed recovery procedure is quite robust to noise and other implementation issues. This makes our algorithm extremely useful for application areas in which the system performance depends critically on the ability to resolve closely spaced delay–Doppler pairs. In particular, our proposed identification method can be used for super-resolution target detection using radar.

REFERENCES

[1] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, 2001.
[2] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York, NY: McGraw-Hill, 2001.
[3] W. Kozek and G. E. Pfander, "Identification of operators with bandlimited symbols," *SIAM J. Math. Anal.*, vol. 37, no. 3, pp. 867–888, 2005.
[4] T. Kailath, "Measurements on time-variant communication channels," *IRE Trans. Inform. Theory*, vol. 8, no. 5, pp. 229–236, Sep. 1962.
[5] P. Bello, "Measurement of random time-variant linear channels," *IEEE Trans. Inform. Theory*, vol. 15, no. 4, pp. 469–475, Jul. 1969.
[6] G. E. Pfander and D. F. Walnut, "Measurement of time-variant linear channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 4808–4820, Nov. 2006.
[7] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
[8] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, no. 3, pp. 292–300, Mar. 1976.
[9] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2275–2284, Jun. 2009.
[10] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3017–3031, Jun. 2010.

the prototype pulse $g(t)$ as at the start of this section (with bandwidth $\mathcal{W} = \frac{8\pi}{T}$), but we use a probing sequence $\{x_n\}$ corresponding to a random (± 1) sequence with $N = 8$. Clearly, this does not satisfy the conditions of Theorem 2 because of the large number of Doppler-shifts associated with the last delay ($K_{\nu,4} = 8$).

The results of this numerical experiment are reported in Fig. 12. It can be easily seen from the figure that, despite the fact that $x(t)$ does not satisfy the conditions of Theorem 2, our algorithm successfully recovers the first three delays and the corresponding Doppler-shifts. In addition, the fourth delay is correctly recovered but (as expected) the Doppler-shifts associated with the last delay are not properly identified. Note that in addition to demonstrating the robustness of our procedure in the presence of model-order mismatch, this experiment also highlights the advantage of the sequential nature of our approach where we first recover the delays and then estimate the Doppler-shifts and attenuation factors associated with the recovered delays. The main advantage of this being that if the input signal does not satisfy $N \geq 2K_{\nu,i}$ for some i then recovery fails only for the Doppler-shifts associated with the ith delay. Moreover, recovery of the ith delay itself does not suffer from the mismodeling and it will be recovered correctly as long as the bandwidth of $x(t)$ is not too small.
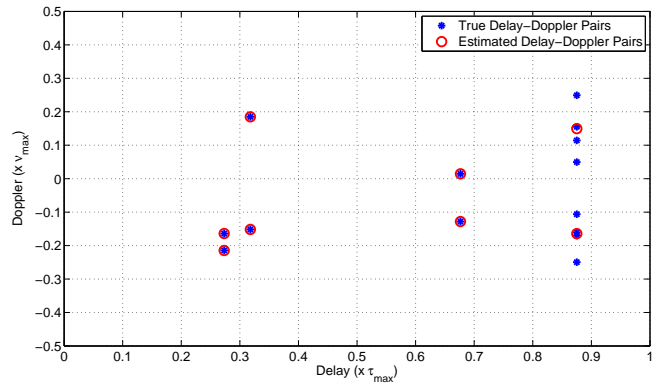
[11] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975–989, Sep. 1982.

[12] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 5, pp. 1340–1342, Oct. 1986.

[13] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.

[14] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[15] A. W. Habboosh, R. J. Vaccaro, and S. Kay, "An algorithm for detecting closely spaced delay/Doppler components," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97)*, Munich, Germany, Apr. 1997, pp. 535–538.

[16] I. C. Wong and B. L. Evans, "Low-complexity adaptive high-resolution channel prediction for OFDM systems," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM'06)*, San Francisco, CA, Dec. 2006, pp. 1–5.

[17] W. U. Bajwa, A. M. Sayeed, and R. Nowak, "Learning sparse doubly-selective channels," in *Proc. 45th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2008, pp. 575–582.

[18] X. Tan, W. Roberts, J. Li, and P. Stoica, "Range-Doppler imaging via a train of probing pulses," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1084–1097, Mar. 2009.

[19] A. Jakobsson, A. L. Swindlehurst, and P. Stoica, "Subspace-based estimation of time delays and Doppler shifts," *IEEE Trans. Signal Processing*, vol. 46, no. 9, pp. 2472–2483, Sep. 1998.

[20] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," CCIT Report # 751, EE Dept., Technion—Israel Institute of Technology, Dec. 2009. [Online]. Available: arXiv:0912.2495

[21] M. Mishali, Y. C. Eldar, and A. Elron, "Xampling: Signal acquisition and processing in union of subspaces," CCIT Report # 747, EE Dept., Technion—Israel Institute of Technology, Oct. 2009. [Online]. Available: arXiv:0911.0519

[22] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Select. Topics Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[23] ——, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[24] Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 2986–2997, Aug. 2009.

[25] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[26] E. Moulines, P. Duhamel, J. Cardoso, S. Mayrargue, and T. Paris, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–25, Feb. 1995.

[27] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[28] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[29] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 4, pp. 806–811, Aug. 1985.

[30] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1741–1757, May 2007.

[31] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[32] R. Tur, K. Gedalyahu, and Y. C. Eldar, "Multichannel sampling of pulse streams at the rate of innovation," submitted to *IEEE Trans. Signal Processing*. [Online]. Available: arXiv:1004.5070

[33] R. Tur, Y. C. Eldar, and Z. Friedman, "Low rate sampling of pulse streams with application to ultrasound imaging," submitted to *IEEE Trans. Signal Processing*. [Online]. Available: arXiv:1003.2822

[34] G. E. Pfander and H. Rauhut, "Sparsity in time-frequency representations," *J. Fourier Anal. Appl.*, Aug. 2009.

[35] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[36] W. U. Bajwa, "New information processing theory and methods for exploiting sparsity in wireless systems," Ph.D. dissertation, University of Wisconsin-Madison, Madison, WI, 2009.

[37] M. Wax, "Detection and estimation of superimposed signals," Ph.D. dissertation, Stanford University, Stanford, CA, 1985.

[38] R. H. Roy, "ESPRIT—estimation of signal parameters via rotational invariance techniques," Ph.D. dissertation, Stanford University, Stanford, CA, 1987.

Waheed U. Bajwa received BE (with Honors) degree in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan in 2001, and MS and PhD degrees in electrical engineering from the University of Wisconsin-Madison, Madison, WI in 2005 and 2009, respectively. He was a Postdoctoral Research Associate in the Program in Applied and Computational Mathematics at Princeton University, Princeton, NJ from 2009 to 2010. He is currently a Research Scientist in the Department of Electrical and Computer Engineering at Duke University, Durham, NC. His research interests include high-dimensional inference and inverse problems, statistical signal processing, wireless communications, and applications in biological sciences, complex networked systems, and radar & image processing.

Dr. Bajwa was affiliated with Communications Enabling Technologies, Islamabad, Pakistan—the research arm of Avaz Networks Inc., Irvine, CA (now Quartics LLC)—from 2000-2003, with the Center for Advanced Research in Engineering, Islamabad, Pakistan during 2003, and with the RF and Photonics Lab of GE Global Research, Niskayuna, NY during the summer of 2006. He received the Best in Academics Gold Medal and President's Gold Medal in Electrical Engineering from the National University of Sciences and Technology (NUST) in 2001, and the Morgridge Distinguished Graduate Fellowship from the University of Wisconsin-Madison in 2003. He was Junior NUST Student of the Year (2000), Wisconsin Union Poker Series Champion (Spring 2008), and President of the University of Wisconsin-Madison chapter of Golden Key International Honor Society (2009). He currently serves as a Guest Associate Editor for Elsevier Physical Communication Journal and is a member of the IEEE, Pakistan Engineering Council, and Golden Key International Honor Society.

Kfir Gedalyahu received the B.Sc. degree in Electrical Engineering in 2008 (*Summa Cum Laude*) from the Technion–Israel Institute of Technology, Haifa, Israel. He is currently pursuing the M.Sc. degree in Electrical Engineering at the Technion.

From 2005 to 2008, he was a research engineer at RAFAEL Research Laboratories, Israel Ministry of Defense, Haifa. His research interests include signal-processing algorithms, and sampling theory with applications in radar and communication.

Yonina C. Eldar received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering both from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2001. From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group at MIT. She is currently a Professor in the Department of Electrical Engineering at the Technion-Israel Institute of Technology, Haifa. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT and a Visiting Professor at Stanford University, Stanford, CA. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, in 2007 the Henry Taub Prize for Excellence in Research, in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion Outstanding Lecture Award, in 2009 the Technion's Award for Excellence in Teaching, and in 2010 the Michael Bruno Memorial Award from the Rothschild Foundation. She is a member of the IEEE Signal Processing Theory and Methods technical committee and the Bio Imaging Signal Processing technical committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the EURASIP Journal of Signal Processing, the SIAM Journal on Matrix Analysis and Applications, and the SIAM Journal on Imaging Sciences, and on the Editorial Board of Foundations and Trends in Signal Processing.