

Time-Delay Estimation From Low-Rate Samples: A Union of Subspaces Approach

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

Abstract—Time-delay estimation arises in many applications in which a multipath medium has to be identified from pulses transmitted through the channel. Previous methods for time delay recovery either operate on the analog received signal, or require sampling at the Nyquist rate of the transmitted pulse. In this paper, we develop a unified approach to time delay estimation from low-rate samples. This problem can be formulated in the broader context of sampling over an infinite union of subspaces. Although sampling over unions of subspaces has been receiving growing interest, previous results either focus on unions of finite-dimensional subspaces, or finite unions. The framework we develop here leads to perfect recovery of the multipath delays from samples of the channel output at the lowest possible rate, even in the presence of overlapping transmitted pulses, and allows for a variety of different sampling methods. The sampling rate depends only on the number of multipath components and the transmission rate, but not on the bandwidth of the probing signal. This result can be viewed as a sampling theorem over an infinite union of infinite dimensional subspaces. By properly manipulating the low-rate samples, we show that the time delays can be recovered using the well-known ESPRIT algorithm. Combining results from sampling theory with those obtained in the context of direction of arrival estimation, we develop sufficient conditions on the transmitted pulse and the sampling functions in order to ensure perfect recovery of the channel parameters at the minimal possible rate.

Index Terms—Sub-Nyquist sampling, time-delay estimation, union of subspaces.

I. INTRODUCTION

TIME-DELAY estimation is an important signal processing problem, arising in various applications such as radar [1], underwater acoustics [2], wireless communications [3], and more. In a typical scenario, pulses with *a priori* known shape are transmitted through a multipath medium, which consists of several propagation paths. As a result, the received signal is composed of delayed and weighted replicas of the transmitted pulses. In order to identify the medium, the time delay and gain coefficient of each multipath component has to be estimated from the received signal.

Manuscript received May 16, 2009; accepted January 27, 2010. Date of publication February 25, 2010; date of current version May 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiang-Gen Xia. This work was supported in part by the Israel Science Foundation under Grant 1081/07, the European Commission in the framework of the FP7 Network of Excellence in Wireless COMMunications NEWCOM++ (Contract 216715), and a Magnetron grant from the Israel Ministry of Industry and Trade.

The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: kfirge@technion.ac.il; yonina@ee.technion.ac.il).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2010.2044253

In this paper, we consider recovery of the parameters defining such a multipath medium from samples of the channel output. Specifically, we assume that pulses with known shape are transmitted at a constant rate through the medium, and our aim is to recover the time delays and time-varying gain coefficients of each multipath component, from samples of the received signal. We derive sampling schemes that ensure perfect recovery of the medium's parameters from samples taken at the minimal possible rate. These sampling schemes can be viewed in a broader context of sampling over a union of subspaces [4]–[6]. Sampling over unions of subspaces has been receiving growing interest in recent years. Some special cases include compressed sensing [7], compressed sensing of analog signals [6], [8]–[12] and finite-rate of innovation (FRI) sampling [13], [14]. All of these previous cases either treat finite unions (such as compressed sensing), or unions of finite-dimensional subspaces (such as FRI). Here we provide sampling theorems for a special case of analog signals with infinitely many degrees of freedom, lying in an infinite union of subspaces. Therefore, these results can be viewed as a systematic treatment of sampling over an infinite union of infinite dimensional subspaces.

The sampling rate of the schemes we develop is generally much lower than the traditional Nyquist rate, and depends only on the number of multipath components and the transmission rate, but not on the bandwidth of the transmitted pulse. This can lead to significant sampling rate reduction, comparing to the Nyquist rate, in applications where only a small number of propagation paths exists, or when the bandwidth of the transmitted pulse is relatively high. This reduction is desirable for practical implementation, since sampling at lower rates allows for analog-to-digital converters (ADCs) that are more precise (i.e., use more bits), and have a lower power consumption. In addition, lowering the sampling rate can reduce the load on both hardware and further digital processing units.

A classical solution for the time-delay estimation problem is based on correlation between the received signal and the transmitted pulse [1]. However, the time resolution of this method is limited by the inverse of the transmitted pulse bandwidth. Therefore, this technique is effective only when the multipath components are well separated in time, or when only one component is present. This approach was originally motivated in the analog domain, where the entire analog correlation is computed. Performing the correlation in the digital domain requires samples of the data at a high sampling rate, in order to approximate the continuous correlation.

In order to resolve closely spaced multipath components, various superresolution estimation algorithms were proposed. In [15], [16], the MUSIC [17] method was applied in the time

domain in order to estimate the time delay of each multipath component. Hou and Wu [18] were the first to convert the time estimation problem to model-based sinusoidal parameter estimation, and used an autoregressive method in order to estimate the model's parameters. Other works, such as [16], [19], [20], relied on the same principle, but different estimation algorithms were used: Tufts–Kumaresan SVD algorithm [21], TLS-ESPRIT method [22] and a modification of MUSIC, respectively.

In the above superresolution approaches, the sampling stage was typically not directly addressed. Most of these works rely on uniform pointwise samples of the received signal, at a high sampling rate. In [18] and [20] the required sampling rate is the Nyquist rate of the transmitted pulse. Since often the pulses are chosen to have small time support, the bandwidth can be quite large, corresponding to a high Nyquist rate. In [19] and in the frequency domain algorithm proposed in [16], the sampling of the received analog signal is not mentioned explicitly. Since these algorithms involve operations in the analog frequency domain, effectively they also require sampling at the Nyquist rate. The time domain algorithms proposed in [15] and [16] can theoretically recover the time delays from a low sampling rate, which depends on the number of multipath components. However, the sampling considerations were not directly addressed in these works, and no concrete conditions on the transmitted pulse and the samples were given, in order to ensure unique recovery of the delays from the samples.

Besides the sampling stage which is not studied in previous works, another assumption underlying all the methods above is that the receiver has access to a single experiment [18]–[20] or multiple nonoverlapping experiments [15], [16], [20] on the channel. In each experiment a pulse is transmitted through the multipath medium, and it is required that all the returns vanish before the next experiment is conducted. This imposes the constraint that the transmitted pulse is sufficiently time limited, which can be problematic in certain scenarios. For example, in wireless communications, modulated pulses are transmitted at a constant symbol rate through the medium. In this case, we cannot consider the observed signal over one symbol period as an independent experiment, since it will generally be affected by reflections caused by adjacent symbols.

In Section II, we propose a general signal model, that can describe the received signal from a time-varying multipath medium. An advantage of our model is that it does not require the assumption of nonoverlapping experiments, and allows for general pulse shapes. We then formulate the medium identification problem as a sampling problem, in which the set of parameters defining the medium have to be recovered from samples of the received signal at the lowest possible rate. To this end we develop a general sampling scheme, which consists of filtering the received signal with a bank of p sampling filters and uniformly sampling their outputs. This class of sampling schemes is common in sampling theory [23] and can accommodate a wide variety of sampling techniques, including pointwise uniform sampling. Given K multipath components, we show that in the worst case at least $2K$ sampling filters are required in order to perfectly recover the time delays. We then develop explicit sampling strategies that achieve this minimal rate. In

particular, we derive sufficient conditions on the transmitted pulse and the choice of sampling filters, which guarantee unique recovery of the channel parameters.

In order to recover the channel parameters from the given samples we combine results from standard sampling theory, with those of direction-of-arrival (DOA) algorithms [17], [22], [24], [25]. Specifically, by appropriate manipulation of the sampling sequences, we show that we can formulate our problem within the framework of DOA methods. We then rely on the Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [22], developed in that context. The unknown delays are recovered from the samples by first applying a digital correction filter bank, and then applying the ESPRIT algorithm. Once the time delays are identified, the gain coefficients are recovered using standard sampling tools.

This paper is organized as follows. In Section II, we describe our signal model. A general sampling scheme of the received signal is proposed in Section III. Section IV describes the recovery of the unknown delays from the samples, and provides sufficient conditions ensuring a unique recovery. Relation to previous work is discussed in detail in Section V. Section VI describes an application example of channel identification in wireless communications. Numerical experiments are described in Section VII.

II. SIGNAL MODEL AND PROBLEM FORMULATION

A. Notations and Definitions

Matrices and vectors are denoted by bold font, with lowercase letters corresponding to vectors and uppercase letters to matrices. The n th 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 complex conjugation, transposition and conjugate transposition, respectively. The Moore–Penrose pseudo-inverse of a matrix $\mathbf{A}$ is written as $\mathbf{A}^\dagger$. The identity matrix of size n is denoted by $\mathbf{I}_n$.

The Fourier transform of a continuous-time signal $x(t) \in L_2$ is defined by $X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$, and

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t)y^*(t)dt \quad (1)$$

denotes the inner product between two continuous-time signals. The discrete-time Fourier transform (DTFT) of a sequence $a[n] \in \ell_2$ is defined by

$$A(e^{j\omega T}) = \sum_{n \in \mathbb{Z}} a[n]e^{-j\omega nT} \quad (2)$$

and is periodic with period $2\pi/T$.

B. Signal Model

We consider the class of signals that can be written in the form

$$x(t) = \sum_{k=1}^K \sum_{n \in \mathbb{Z}} a_k[n]g(t - t_k - nT) \quad (3)$$

where $\tau = \{t_k\}_{k=1}^K$ is a set of K distinct unknown time delays in the continuous interval $[0, T)$, $a_k[n]$ is an arbitrary sequence in ℓ_2 , and $g(t) \in L_2$ is a known function. Each signal from this class can describe the propagation of a pulse with

known shape $g(t)$ which is transmitted at a constant rate of $1/T$ through a medium consisting of K paths. Each path has a constant delay t_k , and a time-varying gain, which is described by the sequence $a_k[n]$. In cases where the transmitted pulses are amplitude modulated, the sequence $a_k[n]$ describes the multiplication between the pulse amplitude and the gain coefficient of each path. In Section VI we will discuss more thoroughly an example of a communication signal transmitted through a multipath time-varying channel.

Our problem is to determine the delays τ and the gains $a_k[n]$ from samples of the received signal $x(t)$, at the minimal possible rate. Since these parameters uniquely define $x(t)$, our channel identification problem is equivalent to developing efficient sampling schemes for signals of the form (3), allowing perfect reconstruction of the signal from its samples.

The model (3) is more general than that described in previous work [15], [16], [18]–[20], which is based on single or multiple experiments on the medium. In each experiment the received signal is observed over a finite time window, which is synchronized to the transmission time of the pulse. More precisely, the received signal in the n th experiment is given by

$$x_n(t) = \sum_{k=1}^K a_{kn} g(t - t_k), \quad t \in [0, T] \quad (4)$$

where t_k is the delay of the k th multipath component which is constant in all the experiments, and a_{kn} is the gain coefficient of the k th multipath component at the n th experiment, which is generally varying from one experiment to another. This model can be seen as a special case of (3) with additional constraints on the pulse $g(t)$ and the transmission rate $1/T$. Indeed, we can write the received signal on the n th experiment as

$$x_n(t) = x(t - nT) \quad t \in [0, T] \quad (5)$$

if we require that

$$g(t - t_k - nT) = 0 \quad \text{for all } n \neq 0. \quad (6)$$

This requirement means that the pulse $g(t)$ has finite time support, and that the repetition period of the pulses T , is long enough such that all the reflections from one pulse vanish before the next pulse is transmitted. On the other hand, our signal model does not require these constraints, so that it can support infinite length pulses and allows interference between experiments.

III. SAMPLING SCHEME

A. Known Delays

Before we treat our sampling problem of signals of the form (3), we first consider a simpler setting in which the delays t_k are known. In this case our signal model is a special case of the more general class of signals that lie in a shift-invariant (SI) subspace. For such signals classes, there are well known sampling schemes that guarantee perfect recovery at the minimal possible rate [8], [23], [26]. Below, we review the main results in this setting which will serve as a basis for our development in the case in which the delays are unknown.

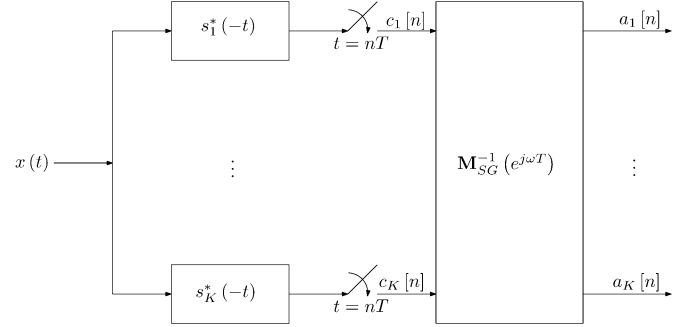


Fig. 1. Sampling and reconstruction scheme for the case of known delays.

A finitely-generated SI subspace of L_2 is defined as [27]–[29]

$$\mathcal{A} = \left\{ x(t) = \sum_{k=1}^K \sum_{n \in \mathbb{Z}} a_k[n] g_k(t - nT) : a_k[n] \in \ell_2 \right\}. \quad (7)$$

The functions $g_k(t)$ are referred to as the generators of $\mathcal{A}$. In order to guarantee a unique stable representation of any signal $x(t) \in \mathcal{A}$ by coefficients $a_k[n]$, the generators $g_k(t)$ are typically chosen to form a Riesz basis of $\mathcal{A}$ [27], [28]. Clearly, the signal model in (3) is a special case of (7) with generators $g_k(t)$, obtained from K delayed versions of $g(t)$:

$$g_k(t) = g(t - t_k), \quad 1 \leq k \leq K. \quad (8)$$

One way to sample a signal of the form (7) is to use K parallel sampling channels [8]. In each channel the signal is first filtered with an impulse response $s_\ell^*(-t)$, and then sampled uniformly at times $t = nT$ to produce the sampling sequence $c_\ell[n]$, as depicted in the left-hand side of Fig. 1. The sampling sequence at the output of the ℓ th channel can be written as

$$c_\ell[n] = \langle x(t), s_\ell(t - nT) \rangle, \quad 1 \leq \ell \leq K. \quad (9)$$

By analyzing the DTFTs $C_\ell(e^{j\omega T})$ of the sequences $c_\ell[n]$, $1 \leq \ell \leq K$, it was shown in [8] that the sequences $a_\ell[n]$, $1 \leq \ell \leq K$, which define the signal $x(t)$, can be recovered from $c_\ell[n]$ using an adequate multichannel filter bank. Specifically, let $\mathbf{c}(e^{j\omega T})$, $\mathbf{a}(e^{j\omega T})$ denote the length- K column vectors whose ℓ th elements are $C_\ell(e^{j\omega T})$, $A_\ell(e^{j\omega T})$ respectively. Then, it can be shown that

$$\mathbf{c}(e^{j\omega T}) = \mathbf{M}_{SG}(e^{j\omega T}) \mathbf{a}(e^{j\omega T}) \quad (10)$$

where $\mathbf{M}_{SG}(e^{j\omega T})$ is a $K \times K$ matrix with ℓk th element

$$\Phi_{S_\ell G_k}(e^{j\omega T}) = \frac{1}{T} \sum_{m \in \mathbb{Z}} S_\ell^* \left(\omega - \frac{2\pi}{T} m \right) G_k \left(\omega - \frac{2\pi}{T} m \right). \quad (11)$$

Here $G_k(\omega)$ and $S_\ell(\omega)$ denote the Fourier transforms of $g_k(t)$ and $s_\ell(t)$, respectively. If this matrix is stably invertible a.e. in ω , then the sequences $a_\ell[n]$ can be recovered from the samples using the multichannel filter bank whose frequency response is given by $\mathbf{M}_{SG}^{-1}(e^{j\omega T})$, as depicted in the right-hand side of Fig. 1.

The proposed sampling scheme achieves an average sampling rate of K/T since there are K sampling sequences each at rate $1/T$. Intuitively, this approach requires one sample per degree

of freedom of the signal $x(t)$: for a signal of the form (3), under the assumption that the time delays are known, or a signal of the form (7), in each time period T there are K new parameters.

B. Unknown Delays

We now address our original problem of designing a sampling scheme for signals of the form (3) with unknown delays. We propose a system similar to that used in the case of known delays, comprised of parallel sampling channels. Since now there are more degrees of freedom in the signal $x(t)$, intuitively we will require at least the same number of channels as in the case of known delays. Denoting the number of channels by p , this implies that $p \geq K$. As we will see, under certain conditions on the sampling filters and pulse $g(t)$, $p = 2K$ sampling filters are sufficient to guarantee perfect recovery of $x(t)$ from the samples. We will also show that this is the minimal possible rate achievable for all signals $x(t)$.

In each channel of our sampling system, the signal $x(t)$ is pre-filtered using the filter $s_\ell^*(-t)$ and sampled uniformly at times $t = nT$ resulting in the samples (9). In the Fourier domain, we can write (9) as

$$C_\ell(e^{j\omega T}) = \frac{1}{T} \sum_{m \in \mathbb{Z}} S_\ell^* \left(\omega - \frac{2\pi}{T}m \right) X \left(\omega - \frac{2\pi}{T}m \right). \quad (12)$$

From the definition of $x(t)$, its Fourier transform can be written as

$$\begin{aligned} X(\omega) &= \sum_{k=1}^K \sum_{n \in \mathbb{Z}} a_k[n] G(\omega) e^{-j\omega(t_k + nT)} \\ &= \sum_{k=1}^K A_k(e^{j\omega T}) G(\omega) e^{-j\omega t_k} \end{aligned} \quad (13)$$

where $A_k(e^{j\omega T})$ denotes the DTFT of the sequence $a_k[n]$, and $G(\omega)$ denotes the Fourier transform of $g(t)$. Substituting (13) into (12), we have

$$\begin{aligned} C_\ell(e^{j\omega T}) &= \sum_{k=1}^K A_k(e^{j\omega T}) e^{-j\omega t_k} \frac{1}{T} \sum_{m \in \mathbb{Z}} S_\ell^* \left(\omega - \frac{2\pi}{T}m \right) \\ &\quad \cdot G \left(\omega - \frac{2\pi}{T}m \right) e^{j\frac{2\pi}{T}mt_k} \end{aligned} \quad (14)$$

where we used the fact that $A_k(e^{j\omega T})$ is $2\pi/T$ -periodic.

In this sequel, we assume that $\omega \in [0, 2\pi/T)$, and that expressions in the DTFT domain are $2\pi/T$ periodic. Denoting by $\mathbf{c}(e^{j\omega T})$ the length- p column vector whose ℓ th element is $C_\ell(e^{j\omega T})$ and by $\mathbf{a}(e^{j\omega T})$ the length- K column vector whose k th element is $A_k(e^{j\omega T})$, we can write (14) in matrix form as

$$\mathbf{c}(e^{j\omega T}) = \mathbf{M}(e^{j\omega T}, \tau) \mathbf{D}(e^{j\omega T}, \tau) \mathbf{a}(e^{j\omega T}). \quad (15)$$

Here $\mathbf{M}(e^{j\omega T}, \tau)$ is a $p \times K$ matrix with ℓk th element

$$\begin{aligned} \mathbf{M}_{\ell k}(e^{j\omega T}, \tau) &= \frac{1}{T} \sum_{m \in \mathbb{Z}} S_\ell^* \left(\omega - \frac{2\pi}{T}m \right) G \left(\omega - \frac{2\pi}{T}m \right) e^{j\frac{2\pi}{T}mt_k} \end{aligned} \quad (16)$$

and $\mathbf{D}(e^{j\omega T}, \tau)$ is a diagonal matrix with k th diagonal element equal to $e^{-j\omega t_k}$. Defining the vector $\mathbf{b}(e^{j\omega T})$ as

$$\mathbf{b}(e^{j\omega T}) = \mathbf{D}(e^{j\omega T}, \tau) \mathbf{a}(e^{j\omega T}) \quad (17)$$

we can rewrite (15) in the form

$$\mathbf{c}(e^{j\omega T}) = \mathbf{M}(e^{j\omega T}, \tau) \mathbf{b}(e^{j\omega T}). \quad (18)$$

Our problem can then be reformulated as that of recovering $\mathbf{b}(e^{j\omega T})$ and the unknown delay set τ from the vectors $\mathbf{c}(e^{j\omega T})$, for all $\omega \in [0, 2\pi/T)$. Once these are known, the vectors $\mathbf{a}(e^{j\omega T})$ can be recovered using the relation in (17).

To proceed, we focus our attention on sampling filters $S_\ell(\omega)$ with finite support in the frequency domain, contained in the frequency range

$$\mathcal{F} = \left[\frac{2\pi}{T}\gamma, \frac{2\pi}{T}(p + \gamma) \right] \quad (19)$$

where $\gamma \in \mathbb{Z}$ is an index which determines the working frequency band $\mathcal{F}$. This choice should be such that it matches the frequency occupation of $g(t)$ (although $g(t)$ does not have to be bandlimited). This freedom allows our sampling scheme to support both complex and real valued signals. Under this choice of filters, each element $\mathbf{M}_{\ell k}(e^{j\omega T}, \tau)$ of (16) can be expressed as

$$\mathbf{M}_{\ell k}(e^{j\omega T}, \tau) = \sum_{m=1}^p \mathbf{W}_{\ell m}(e^{j\omega T}) \mathbf{N}_{mk}(\tau) \quad (20)$$

where $\mathbf{W}(e^{j\omega T})$ is a $p \times p$ matrix whose ℓm th element is given by

$$\begin{aligned} \mathbf{W}_{\ell m}(e^{j\omega T}) &= \frac{1}{T} S_\ell^* \left(\omega + \frac{2\pi}{T}(m - 1 + \gamma) \right) \\ &\quad \cdot G \left(\omega + \frac{2\pi}{T}(m - 1 + \gamma) \right) \end{aligned} \quad (21)$$

and $\mathbf{N}(\tau)$ is a $p \times K$ Vandermonde matrix with mk th element

$$\mathbf{N}_{mk}(\tau) = e^{-j\frac{2\pi}{T}(m-1+\gamma)t_k}. \quad (22)$$

Substituting (20) into (18),

$$\mathbf{c}(e^{j\omega T}) = \mathbf{W}(e^{j\omega T}) \mathbf{N}(\tau) \mathbf{b}(e^{j\omega T}). \quad (23)$$

If $\mathbf{W}(e^{j\omega T})$ is stably invertible, then we can define the modified measurement vector $\mathbf{d}(e^{j\omega T})$ as

$$\mathbf{d}(e^{j\omega T}) = \mathbf{W}^{-1}(e^{j\omega T}) \mathbf{c}(e^{j\omega T}). \quad (24)$$

This vector satisfies

$$\mathbf{d}(e^{j\omega T}) = \mathbf{N}(\tau) \mathbf{b}(e^{j\omega T}). \quad (25)$$

Since $\mathbf{N}(\tau)$ is independent of ω , from the linearity of the DTFT, we can express (25) in the time domain as

$$\mathbf{d}[n] = \mathbf{N}(\tau) \mathbf{b}[n], \quad n \in \mathbb{Z}. \quad (26)$$

The elements of the vectors $\mathbf{d}[n]$ and $\mathbf{b}[n]$ are the discrete time sequences, obtained from the inverse DTFT of the elements of the vectors $\mathbf{b}(e^{j\omega T})$ and $\mathbf{d}(e^{j\omega T})$ respectively.

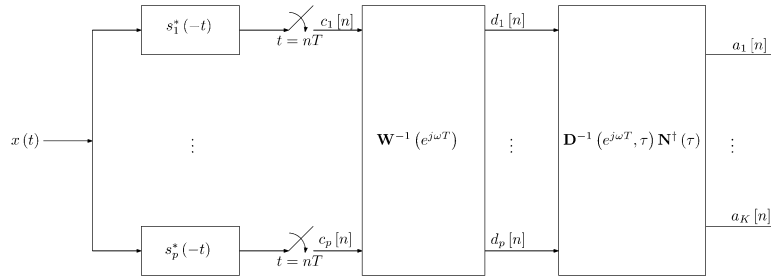


Fig. 2. Sampling and reconstruction scheme for the case of unknown delays.

Equation (25) and its equivalent time domain representation (26), describe an infinite set of measurement vectors, each obtained by the same measurement matrix $\mathbf{N}(\tau)$, which depends on the unknown delays τ . This problem is reminiscent of the type of problems that arise in the field of DOA estimation [25], as we discuss in the next section. One class of efficient methods for DOA recovery, are known as subspace methods [25]. These techniques have subsequently been applied to many other problems such as spectral estimation [30], system identification [31] and more. Our approach is to rely on these methods in order to first recover τ from the measurements. After τ is known, the vectors $\mathbf{b}(e^{j\omega T})$ and $\mathbf{a}(e^{j\omega T})$ can be found using linear filtering relations by

$$\mathbf{b}(e^{j\omega T}) = \mathbf{N}^\dagger(\tau) \mathbf{d}(e^{j\omega T}). \quad (27)$$

Under the assumption that τ is a set of distinct time delays, $\mathbf{N}(\tau)$ is a Vandermonde matrix with linearly independent columns, and consequently $\mathbf{N}^\dagger(\tau) \mathbf{N}(\tau) = \mathbf{I}_K$. Using (17)

$$\mathbf{a}(e^{j\omega T}) = \mathbf{D}^{-1}(e^{j\omega T}, \tau) \mathbf{N}^\dagger(\tau) \mathbf{d}(e^{j\omega T}). \quad (28)$$

The resulting sampling and reconstruction scheme is depicted in Fig. 2.

We point out that when two or more delays are close, the matrix $\mathbf{N}^\dagger(\tau)$ can be ill conditioned. This will affect the stability of the recovery of $\mathbf{a}(e^{j\omega T})$ in the presence of noise in the system. We examine this issue in more detail in Section VII.

Our last step, is to derive conditions on the filters $s_\ell^*(-t)$ and the function $g(t)$ ensuring that the matrix $\mathbf{W}(e^{j\omega T})$ is stably invertible. To this end, we decompose the matrix $\mathbf{W}(e^{j\omega T})$ as

$$\mathbf{W}(e^{j\omega T}) = \mathbf{S}(e^{j\omega T}) \mathbf{G}(e^{j\omega T}) \quad (29)$$

where $\mathbf{S}(e^{j\omega T})$ is a $p \times p$ matrix with ℓm th element

$$\mathbf{S}_{\ell m}(e^{j\omega T}) = \frac{1}{T} S_\ell^* \left(\omega + \frac{2\pi}{T} (m-1+\gamma) \right) \quad (30)$$

and $\mathbf{G}(e^{j\omega T})$ is a $p \times p$ diagonal matrix whose m th diagonal element is given by

$$\mathbf{G}_{mm}(e^{j\omega T}) = G \left(\omega + \frac{2\pi}{T} (m-1+\gamma) \right). \quad (31)$$

Each one of these matrices needs to be stably invertible. Therefore, from (31) the condition that the function $g(t)$ needs to satisfy is that

$$0 < a \leq |G(\omega)| \leq b < \infty \text{ a.e } \omega \in \mathcal{F}. \quad (32)$$

In addition the filters $s_\ell^*(-t)$ should be chosen in such a way that they form a stably invertible matrix $\mathbf{S}(e^{j\omega T})$. Examples of such filters are given in the next subsection.

We note here that the conditions given above guarantee a stable digital correction filter bank $\mathbf{W}^{-1}(e^{j\omega T})$; however, generally it will be comprised of infinite length digital filters. Practical implementation of these filters can be achieved by truncating the impulse response. The length of the resulting filters will affect the total delay of our proposed method, which will generally be longer than that of the methods described in Section I, due to the additional digital filtering stage.

We summarize the results so far in the following proposition.

Proposition 1: Let $c_\ell[n] = \langle x(t), s_\ell(t-nT) \rangle$, $1 \leq \ell \leq p$ be a set of p sequences obtained by filtering the signal $x(t)$ defined by (3) with p filters $s_\ell^*(-t)$ and sampling their outputs at times nT . Let $S_\ell(\omega)$ be supported on $\mathcal{F} = [2\pi/T\gamma, 2\pi/T(p+\gamma)]$, and let $\Omega = [0, 2\pi/T)$. If the function $g(t)$ satisfies the condition in (32) and the matrix $\mathbf{S}(e^{j\omega T})$, defined by (30), is stably invertible a.e. $\omega \in \Omega$, then the delays τ and vector $\mathbf{b}(e^{j\omega T})$ can be found from the set of equations

$$\mathbf{d}(e^{j\omega T}) = \mathbf{N}(\tau) \mathbf{b}(e^{j\omega T}) \quad (33)$$

using subspace methods, described in the next section. Here $\mathbf{N}(\tau)$ is a $p \times K$ Vandermonde matrix with $m k$ th element $e^{-j2\pi/T(m-1+\gamma)\tau_k}$, and

$$\mathbf{d}(e^{j\omega T}) = \mathbf{W}^{-1}(e^{j\omega T}) \mathbf{c}(e^{j\omega T}) \quad (34)$$

with $\mathbf{W}(e^{j\omega T})$ defined by (21). The sequences $a_\ell[n]$ can then be recovered by

$$\mathbf{a}(e^{j\omega T}) = \mathbf{D}^{-1}(e^{j\omega T}, \tau) \mathbf{N}^\dagger(\tau) \mathbf{d}(e^{j\omega T}), \quad \omega \in \Omega \quad (35)$$

where $\mathbf{D}(e^{j\omega T}, \tau)$ is a diagonal matrix with diagonal elements $e^{-j\omega \tau_k}$.

C. Examples of Filters

We now provide some examples of filters $s_\ell(t)$ satisfying the required conditions.

1) Complex Bandpass Filter-Bank: The first example is a set of complex bandpass filters. We assume that the working band is $\mathcal{F} = [0, 2\pi/T]$ ($\gamma = 0$), and the function $g(t)$ satisfies (32) on that frequency range. We choose the filters $s_\ell^*(-t)$ as ideal bandpass filters, covering consecutive frequency bands

$$S_\ell(\omega) = \begin{cases} T, & \omega \in [(\ell-1)\frac{2\pi}{T}, \ell\frac{2\pi}{T}] \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

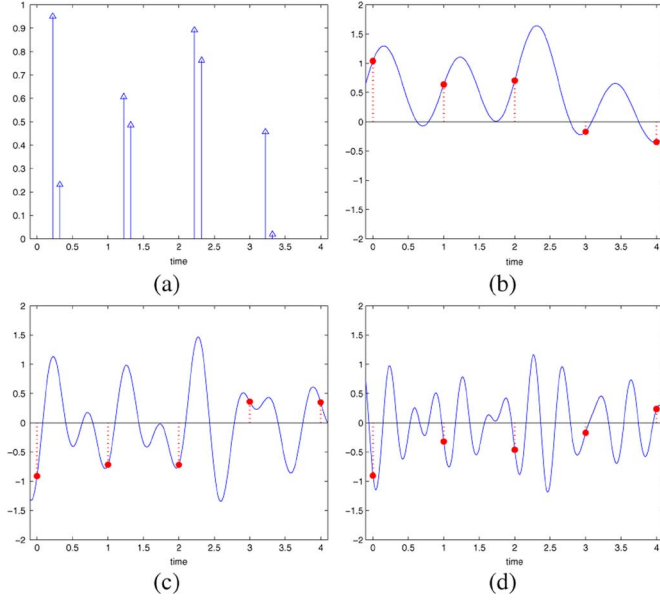


Fig. 3. Stream of diracs. (a) $K = 2$ diracs per period $T = 1$. (b)–(d) The outputs of the first three sampling channels. The dashed lines denote the sampling instants.

The resulting matrix $\mathbf{S}(e^{j\omega T})$ is diagonal, and stably invertible. This example generalizes to any valid working band, given by (19), by shifting the frequency response of the filters.

We now provide an example demonstrating the importance of the sampling filter.

Example 1: We consider the case where $g(t) = \delta(t)$ and there are $K = 2$ diracs per period of $T = 1$, as illustrated in Fig. 3(a). The sampling scheme described above, consisting of a complex bandpass filter-bank, is used. In Fig. 3(b)–(d), we show the outputs of the first three sampling channels. This example demonstrates the need for the sampling filters when sampling short-length pulses at a low sampling rate. The sampling kernels have the effect of smoothing the short pulses (diracs in this example). Consequently, even when the sampling rate is low, the samples contain information about the signal. In contrast, if we were to sample the signal in Fig. 3(a) directly at a low rate, then we would often obtain only zero samples which contain no information about the signal.

2) *Delayed Channels:* In this example we assume p is even and define the working band as

$$\mathcal{F} = \left[-\frac{\pi}{T}p, \frac{\pi}{T}p\right] \quad (37)$$

($\gamma = -p/2$). We also assume that $g(t)$ satisfies (32). We choose the ℓ th filter as a delay of $\Delta_\ell \in [0, T)$ followed by an ideal low pass filter. Thus,

$$S_\ell(\omega) = \begin{cases} T e^{-j\omega \Delta_\ell}, & \omega \in \mathcal{F} \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

With this choice of filters, the ℓ mth element of the matrix $\mathbf{S}(e^{j\omega T})$ defined in (30) is given by

$$\begin{aligned} \mathbf{S}_{\ell m}(e^{j\omega T}) &= e^{j(\omega + \frac{2\pi}{T}(m-1-p/2))\Delta_\ell} \\ &= e^{j(\omega - \frac{2\pi}{T}(p/2))\Delta_\ell} e^{j\frac{2\pi}{T}(m-1)\Delta_\ell}. \end{aligned} \quad (39)$$

The matrix $\mathbf{S}(e^{j\omega T})$ can be expressed as

$$\mathbf{S}(e^{j\omega T}) = \mathbf{\Phi}(e^{j\omega T})\mathbf{F} \quad (40)$$

where $\mathbf{\Phi}(e^{j\omega T})$ is a diagonal $p \times p$ matrix whose ℓ th diagonal element is $e^{j(\omega - 2\pi/T(p/2))\Delta_\ell}$, and $\mathbf{F}$ is a Vandermonde matrix whose ℓ mth element is given by $e^{j2\pi/T(m-1)\Delta_\ell}$. From (40), it can be seen that $\mathbf{S}(e^{j\omega T})$ is invertible for all $\omega \in \Omega$, when the delays in each channel Δ_ℓ are distinct.

One special case of this choice of sampling filters is when the delays are uniformly spaced, i.e., $\Delta_\ell = (\ell - 1)T/p$. In this case our sampling scheme can be implemented by an ideal low pass filter with cutoff $\pi p/T$, followed by a uniform sampler at a rate of p/T .

IV. RECOVERY OF THE UNKNOWN DELAYS

We have seen in the previous section that perfect reconstruction of a signal $x(t)$ of the form (3), is equivalent to that of recovering the delays τ from the modified measurements of (26). As we now show, this problem is similar to that arising in DOA estimation.

A. Relation to Direction of Arrival Estimation

In DOA estimation [17], [22], [24], [25], K narrow band sources impinge on an array, composed of p sensors, from distinct DOAs. The goal is to estimate the DOAs of the sources from a set of M measurements, obtained from the sensors outputs at distinct time instants.

The DOA estimation problem can be formulated using the following measurement model:

$$\mathbf{X} = \mathbf{A}(\Theta)\mathbf{S} \quad (41)$$

where $\mathbf{X}$ is a $p \times M$ matrix, composed of the measurements in its columns, $\mathbf{S}$ is a $K \times M$ matrix consisting of the sources signals in its columns and $\mathbf{A}(\Theta)$ is a $p \times K$ matrix which depends on the set of unknown DOAs Θ . The structure of $\mathbf{A}(\Theta)$ is such that its k th column, denoted $\mathbf{a}(\theta_k)$, depends only on the DOA θ_k of the k th source. The vector $\mathbf{a}(\theta_k)$ is referred to as the steering vector of the array toward direction θ_k . The set containing all possible steering vectors, i.e., $\{\mathbf{a}(\theta), \theta \in [0, 2\pi)\}$ is referred to as the array manifold. Given $\mathbf{X}$, the problem is to recover the DOAs θ_k , and the sources $\mathbf{S}$.

The set of equations in (26) has the same form as (41). The k th column of the matrix $\mathbf{N}(\tau)$ depends only on the k th unknown delay t_k , and can be described by the vector $\mathbf{n}(t_k)$, where

$$\mathbf{n}(t) = [e^{-j\frac{2\pi}{T}\gamma t} \quad e^{-j\frac{2\pi}{T}(1+\gamma)t} \quad \dots \quad e^{-j\frac{2\pi}{T}(p-1+\gamma)t}]^T. \quad (42)$$

The array manifold in our setting is the set of vectors $\{\mathbf{n}(t), t \in [0, T)\}$. Therefore, we can adopt DOA methods to estimate the unknown delays. The only difference between the two problems is that in our setting, we have infinitely many measurement vectors, in contrast to the DOA problem in which $\mathbf{X}$ has finitely many columns. This will require several adjustments, which we will detail in the ensuing subsections.

Two prominent methods used for DOA estimation are MUSIC (Multiple Signal Classification) [17] and ESPRIT [22]. These algorithms belong to a class of techniques, known as subspace methods, which are based on separating the space containing the measurements into two subspaces, the signal

and noise subspaces. Estimating the unknown set of parameters using MUSIC involves a continuous one-dimensional search over the parameter range, which can be costly computationally. The ESPRIT approach estimates the unknown set of parameters more efficiently, by imposing the additional requirement that the measurement matrix is rotationally invariant. We describe this property in Section IV-C and show that in our case $\mathbf{N}(\tau)$ satisfies this condition, and therefore we use the ESPRIT approach.

B. Sufficient Conditions for Perfect Recovery

We now rely on results obtained in the context of DOA estimation in order to develop sufficient conditions for a unique solution to (26). Such a solution consists of the infinite set of vectors $\mathbf{b}[n]$, $n \in \mathbb{Z}$ and the unknown delays τ .

Conditions for a unique solution $(\Theta, \mathbf{S})$ for (41) were derived in [32]. Since [32] deals with a finite number of measurements, we have to extend the results to our case, which consists of an infinite number of measurements. The analysis in [32] requires a preliminary condition that any subset of p distinct steering vectors from the array manifold is linearly independent. In our case this condition translates into the requirement that any set of p vectors $\mathbf{n}(t_i)$, $1 \leq i \leq p$ associated with distinct delays $t_i \in [0, T]$, $1 \leq i \leq p$ are linearly independent. From (42), any such set forms a $p \times p$ Vandermonde matrix, and are therefore linearly independent. Therefore, this condition automatically holds in our problem without forcing any additional constraints.

To derive sufficient conditions for a unique solution of the set of infinite equations (26) we introduce some notation. We define the measurement set $\mathbf{d}[\Lambda]$ as the set containing all measurement vectors $\mathbf{d}[\Lambda] = \{\mathbf{d}[n], n \in \mathbb{Z}\}$. Similarly, we define the unknown vector set $\mathbf{b}[\Lambda]$ as $\mathbf{b}[\Lambda] = \{\mathbf{b}[n], n \in \mathbb{Z}\}$. We may then rewrite (26) as

$$\mathbf{d}[\Lambda] = \mathbf{N}(\tau)\mathbf{b}[\Lambda]. \quad (43)$$

The following proposition provides sufficient conditions for a unique solution to (43).

Proposition 2: If $(\bar{\tau}, \bar{\mathbf{b}}[\Lambda])$ is a solution to (43),

$$p > 2K - \dim(\text{span}(\bar{\mathbf{b}}[\Lambda])) \quad (44)$$

and

$$\dim(\text{span}(\bar{\mathbf{b}}[\Lambda])) \geq 1 \quad (45)$$

then $(\bar{\tau}, \bar{\mathbf{b}}[\Lambda])$ is the unique solution of (43).

The notation $\text{span}(\bar{\mathbf{b}}[\Lambda])$ is used for the minimal dimension subspace containing the unknown vector set $\bar{\mathbf{b}}[\Lambda]$. The condition (45) is needed to avoid the case where $\bar{\mathbf{b}}[\Lambda] = \mathbf{0}$. In this case clearly the set τ can not be recovered uniquely.

Proof: We denote $r = \dim(\text{span}(\bar{\mathbf{b}}[\Lambda]))$. From (45) $r \geq 1$. Therefore, there exist a finite subset $\tilde{\Lambda} = \{n_i\}_{i=1}^r \subset \Lambda$, such that the vector set $\bar{\mathbf{b}}[\tilde{\Lambda}]$ spans an r -dimensional subspace:

$$\dim(\text{span}(\bar{\mathbf{b}}[\tilde{\Lambda}])) = r. \quad (46)$$

By defining the matrices $\mathbf{B}$ and $\mathbf{D}$ as the matrices consisting of the vector sets $\bar{\mathbf{b}}[\tilde{\Lambda}]$ and $\bar{\mathbf{d}}[\tilde{\Lambda}]$, we can write

$$\mathbf{D} = \mathbf{N}(\bar{\tau})\mathbf{B}. \quad (47)$$

Clearly, from its construction, the rank of the matrix $\mathbf{B}$ is r . From (44),

$$p > 2K - r. \quad (48)$$

According to Theorem 1 in [32], the solution $(\bar{\tau}, \mathbf{B})$ is the unique solution to (47) under the condition (48).

Since the set of unknown delays $\bar{\tau}$ is the unique solution to the finite set of (47), it is also a unique solution to the infinite set of (43). Once $\bar{\tau}$ is uniquely determined, the matrix $\mathbf{N}(\bar{\tau})$ is known. Since every vector of the set $\bar{\mathbf{b}}[\tilde{\Lambda}]$ is contained in $\mathbb{C}^K$,

$$\dim(\text{span}(\bar{\mathbf{b}}[\tilde{\Lambda}])) \leq K. \quad (49)$$

Therefore, according to (44), $p > K$. The matrix $\mathbf{N}(\bar{\tau})$ is a $p \times K$ Vandermonde matrix which consist of K linearly independent vectors. Therefore, for every $n \in \Lambda$, if $\bar{\mathbf{b}}[n]$ is a solution to

$$\mathbf{d}[n] = \mathbf{N}(\bar{\tau})\mathbf{b}[n] \quad (50)$$

then it is the unique solution. ■

Proposition 2 suggests that a unique solution to the set of (26) is guaranteed, under proper selection of the number of sampling channels p . This parameter, in turn, determines the average sampling rate of our sampling scheme, which is given by p/T . Condition (44) depends on the value of $\dim(\text{span}(\bar{\mathbf{b}}[\Lambda]))$, which is generally not known in advance. According to our assumption $\dim(\text{span}(\bar{\mathbf{b}}[\Lambda])) \geq 1$, therefore in order to satisfy the uniqueness condition (44) for every signal of the form (3), we must have $p > 2K - 1$ sampling channels or a minimal sampling rate of $2K/T$. Comparing this result to the minimal sampling rate in the case when the delays are known in advance, there is a penalty of 2 in the minimal rate.

In Section V-A, we show that our signal model, described in (3), can be considered as part of a more general framework of signals that lie in a union of SI subspaces [4]. It was shown in [4] that the theoretical minimum sampling rate required for perfect recovery of such a signal from its samples is $2K/T$. Therefore, according to the results of Proposition 2, our sampling scheme can achieve the minimal sampling rate required for signals of the form (3).

The minimal sampling rate of $2K/T$, which is achieved by our scheme, does not depend on the bandwidth of the pulse $g(t)$, but only on the number of propagation paths K . In applications where the number of propagation paths is relatively small, or the bandwidth of the transmitted pulse is high, our approach can provide a sampling rate lower than the Nyquist rate. More precisely, when $2K/T < W$, where W is the bandwidth of the transmitted pulse, our method can reduce the sampling rate relatively to the Nyquist rate. For example, the setup in [33], used for characterization of ultra-wideband (UWB) wireless indoor channels, consists of pulses with bandwidth $W = 1$ GHz transmitted at a rate of $1/T = 2$ MHz. Under the assumption that there are 32 significant multipath components, our method can reduce the sampling rate down to 128 MHz compared with the 2 GHz Nyquist rate.

Besides the theoretical interest, sampling rate reduction is also important for implementation considerations. For practical ADCs, which perform the sampling process, there is a tradeoff

between sampling rate and resolution [34]. Therefore, reducing the sampling rate allows the use of more precise ADCs, which can improve the time-delay estimation. The power consumption of an ADC can also be reduced by lowering the sampling rate [34]. In addition, a lower rate leads to more efficient digital processing hardware, since a smaller number of samples has to be processed. This also allows performing the digital processing operations in real time.

C. Recovering the Unknown Delays

According to Proposition 2, in order to be able to perfectly reconstruct every signal of the form (3), our sampling scheme must have $p \geq 2K$ sampling channels. We assume throughout that this condition holds.

We now describe an algorithm for the recovery of the unknown delays from the measurement set $\mathbf{d}[\Lambda]$, which is based on the ESPRIT [22] algorithm. One of the conditions needed in order to use ESPRIT is that the correlation matrix

$$\mathbf{R}_{bb} = \sum_{n \in \mathbb{Z}} \mathbf{b}[n] \mathbf{b}^H[n] \quad (51)$$

is positive definite. In order to relate this condition to our problem, we state the following proposition from [11].

Proposition 3: If the sum (51) exists, then every matrix $\mathbf{V}$ satisfying $\mathbf{R}_{bb} = \mathbf{V}\mathbf{V}^H$ has column span equal to $\text{span}(\mathbf{b}[\Lambda])$.

An immediate corollary from Proposition 3 is that $\mathbf{R}_{bb} \succ 0$ is equivalent to the condition $\dim(\text{span}(\mathbf{b}[\Lambda])) = K$. In this case, which we refer to as the *uncorrelated case*, we can apply the ESPRIT algorithm on the measurement set $\mathbf{d}[n]$ in order to recover the unknown delays. The case $\dim(\text{span}(\mathbf{b}[\Lambda])) < K$, will be referred to as the *correlated case*. In this setting the condition $\mathbf{R}_{bb} \succ 0$ does not hold, and the ESPRIT algorithm cannot be applied directly. Instead, we will use an additional stage originally proposed in [35] and [36].

Note that

$$\begin{aligned} \mathbf{R}_{dd} &= \sum_{n \in \mathbb{Z}} \mathbf{d}[n] \mathbf{d}^H[n] \\ &= \mathbf{N}(\tau) \left(\sum_{n \in \mathbb{Z}} \mathbf{b}[n] \mathbf{b}^H[n] \right) \mathbf{N}^H(\tau) \\ &= \mathbf{N}(\tau) \mathbf{R}_{bb} \mathbf{N}^H(\tau). \end{aligned} \quad (52)$$

Since for any set of delays t_k , the matrix $\mathbf{N}(\tau)$ has full column-rank, the ranks of the matrices $\mathbf{R}_{dd}$ and $\mathbf{R}_{bb}$ are equal. Therefore, the decision whether we are in the uncorrelated or correlated case can be made directly from the given measurements by forming the matrix $\mathbf{R}_{dd}$.

1) *Uncorrelated Case:* From (52), under the assumption that the matrix $\mathbf{R}_{bb}$ is positive definite, it can be shown that the rank of the matrix $\mathbf{R}_{dd}$ is K . Moreover, the matrices $\mathbf{R}_{dd}$ and $\mathbf{N}(\tau)$ have the same column span which is referred as the signal subspace. By performing a singular value decomposition (SVD) of the matrix $\mathbf{R}_{dd}$, we can obtain K vectors, which span the signal subspace, by taking the K left singular vectors associated to the nonzero singular values of $\mathbf{R}_{dd}$. We define the $p \times K$ matrix $\mathbf{E}_s$ as the matrix containing those vectors in its columns.

Next, we exploit the special structure of the Vandermonde matrix. We denote the matrix $\mathbf{N}_{\downarrow}(\tau)$ as the sub matrix extracted from $\mathbf{N}(\tau)$ by deleting its last row. In the same way we define $\mathbf{N}_{\uparrow}(\tau)$ as the sub matrix extracted from $\mathbf{N}(\tau)$ by deleting its first row. The Vandermonde matrix $\mathbf{N}(\tau)$ satisfies the following rotational invariance property:

$$\mathbf{N}_{\uparrow}(\tau) = \mathbf{N}_{\downarrow}(\tau) \mathbf{R}(\tau) \quad (53)$$

where $\mathbf{R}(\tau)$ is a diagonal $K \times K$ matrix, whose k th diagonal element is given by $\mathbf{R}_{kk}(\tau) = e^{-j2\pi t_k/T}$. Since the matrices $\mathbf{N}(\tau)$ and $\mathbf{E}_s$ have the same column span, there exists an invertible $K \times K$ matrix $\mathbf{T}$ such that

$$\mathbf{N}(\tau) = \mathbf{E}_s \mathbf{T}. \quad (54)$$

By deleting the last row in (54) we get

$$\mathbf{N}_{\downarrow}(\tau) = \mathbf{E}_{s\downarrow} \mathbf{T}. \quad (55)$$

Similarly, deleting the first row in (54) and using the rotational invariance property (53), we have

$$\mathbf{N}_{\uparrow}(\tau) \mathbf{R}(\tau) = \mathbf{E}_{s\uparrow} \mathbf{T}. \quad (56)$$

Combining (55) and (56) leads to the following relation between the matrices $\mathbf{E}_{s\uparrow}$ and $\mathbf{E}_{s\downarrow}$:

$$\mathbf{E}_{s\uparrow} = \mathbf{E}_{s\downarrow} \mathbf{T} \mathbf{R}(\tau) \mathbf{T}^{-1}. \quad (57)$$

The matrix $\mathbf{E}_{s\uparrow}$ is a $(p-1) \times K$ ($p > K$) matrix with full column rank. Therefore, $\mathbf{E}_{s\downarrow}^\dagger \mathbf{E}_{s\uparrow} = \mathbf{I}_K$. Using (57) we define the following $K \times K$ matrix Φ as

$$\Phi = \mathbf{E}_{s\downarrow}^\dagger \mathbf{E}_{s\uparrow} = \mathbf{T} \mathbf{R}(\tau) \mathbf{T}^{-1}. \quad (58)$$

From (58) it is clear that the diagonal matrix $\mathbf{R}(\tau)$ can be obtained from the matrix Φ by performing an eigenvalue decomposition. Once the matrix $\mathbf{R}(\tau)$ is known, the unknown delays can be retrieved from its diagonal elements as

$$t_k = -\frac{T}{2\pi} \arg(\mathbf{R}_{kk}(\tau)). \quad (59)$$

In summary, our algorithm consist of the following steps.

- 1) Construct the correlation matrix $\mathbf{R}_{dd} = \sum_{n \in \mathbb{Z}} \mathbf{d}[n] \mathbf{d}^H[n]$.
- 2) Perform an SVD decomposition of $\mathbf{R}_{dd}$ and construct the matrix $\mathbf{E}_s$ consisting of the K singular vectors associated with the nonzero singular values in its columns.
- 3) Compute the matrix $\Phi = \mathbf{E}_{s\downarrow}^\dagger \mathbf{E}_{s\uparrow}$.
- 4) Compute the eigenvalues of Φ , $\lambda_i, i = 1, 2, \dots, K$.
- 5) Retrieve the unknown delays by $t_i = -T/(2\pi) \arg(\lambda_i)$.

2) *Correlated Case:* When the condition $\mathbf{R}_{bb} \succ 0$ is not satisfied the ESPRIT algorithm cannot be applied directly on the vector set $\mathbf{d}[\Lambda]$. In this case the rank of $\mathbf{R}_{dd}$ is smaller than K , and therefore its column span is no longer equal to the entire signal subspace. To accommodate this setting, we perform an additional stage before applying the ESPRIT method, based on the spatial smoothing technique proposed in [35], [36].

To proceed, we define $M = p - K$ length- $(K+1)$ sub vectors

$$\mathbf{d}_i[n] = [d_i[n] \ d_{i+1}[n] \ \dots \ d_{i+K}[n]]^T. \quad (60)$$

We define the smoothed correlation matrix $\bar{\mathbf{R}}_{dd}$ as

$$\bar{\mathbf{R}}_{dd} = \frac{1}{M} \sum_{i=1}^M \sum_{n \in \mathbb{Z}} \mathbf{d}_i[n] \mathbf{d}_i^H[n]. \quad (61)$$

Under our assumptions $p \geq 2K$, therefore $M \geq K$. According to [36], when $M \geq K$ the rank of the smoothed correlation matrix is K regardless of the rank of $\mathbf{R}_{bb}$. We may therefore refer now to column rank of $\bar{\mathbf{R}}_{dd}$ as the signal subspace, and apply the ESPRIT algorithm on this matrix.

V. RELATED SAMPLING PROBLEMS

In this section, we explore in more detail the relationship between our sampling problem and previous related setups treated in the sampling literature: sampling signals from a union of subspaces [4], [5], compressed sensing of analog signals [6], [8]–[10], and FRI sampling [13], [14].

A. Sampling Signals From a Union of Subspaces

A model which received growing interest recently is that of signals that lie in a union of subspaces [4]–[6], [9], [10], [12]. Under this model each signal $x(t)$ can be described as [4]

$$x(t) \in \bigcup_{\gamma \in \Gamma} \mathcal{S}_\gamma \quad (62)$$

where $\mathcal{S}_\gamma$ are subspaces of a given Hilbert space and Γ is an index set. The signal $x(t)$ lies in one of the subspaces $\mathcal{S}_\gamma$, however it is not known in advance in which one. Thus, effectively, to determine $x(t)$, we first need to find the active subspace, or the index γ .

Our signal model, given by (3), can be formulated as in (62). As described in Section III-A once the time delays are fixed, each signal $x(t)$ lies in a SI subspace spanned by K generators. Therefore, the set of all signals of the form (3) constitute an infinite union of SI subspaces, where γ is the set of delays τ , which can take on any continuous value in the interval $[0, T]$, and $\mathcal{S}_\gamma$ is the corresponding SI subspace.

In [4] and [5], necessary and sufficient conditions are derived for a sampling operator to be invertible over a union of subspaces. For the case of a union of SI subspaces, [4] suggests a sampling scheme, similar to that used in [8] and in this paper, comprised of parallel sampling channels. Conditions on the sampling filters are then given in order to ensure reconstruction of the signals from its samples. In addition, the minimal number of sampling channels allowing perfect recovery of the signal from its samples is shown to be $2K$. This leads to a minimal sampling rate of $2K/T$ which is achieved by our scheme. However, in [4] no concrete reconstruction algorithms were given that can achieve this rate. Furthermore, although conditions for invertibility were provided, these do not necessarily imply that there exists an efficient recovery algorithm, which can recover the signal from its samples at the minimal

rate. Our aim in this work, is to provide concrete recovery techniques, that are simple to implement, for signals over an infinite union of SI subspaces.

In summary, in this work we focus on a special case of signals that lie in an infinite union of SI subspaces. For this case, in contrast to [4], we provide a concrete reconstruction method. This method achieves the minimal theoretical sampling rate derived in [4]. In addition, while other works [6], [8]–[10], [12] provided reconstruction algorithms only for signals defined over a finite union of subspaces, here we provide a systematic sampling and reconstruction method for signals in an infinite union of SI subspaces.

B. Compressed Sensing of Analog Signals

The results of [8] provide recovery algorithms for signals that lie in a finite union of SI subspaces. In contrast, here there are an infinite number of possible subspaces.

The signal model in [8] is described in terms of N generating functions $a_\ell(t)$ as

$$x(t) = \sum_{|\ell|=K} \sum_{n \in \mathbb{Z}} d_\ell[n] a_\ell(t - nT), \quad (63)$$

where the notation $|\ell| = K$ means a sum over at most K elements. Thus, for each signal there are only K active generating functions out of N total possible functions, but we do not know in advance which generators are active. In principle, such signals can be sampled and recovered using the paradigm described in Section III corresponding to N generating functions. Indeed, any signal of the form (63) clearly also lies in the SI subspace spanned by the N generators $a_\ell(t)$, where some of the sequences $d_\ell[n]$ are identically 0. However, this would require a sampling rate of N/T , obtained by N sampling filters. Since only K of the generators are active, intuitively, we should be able to reduce the rate and still be able to recover the signal. The main contribution of [8] is a sampling scheme consisting of $2K$ filters that is sufficient in order to recover $x(t)$ exactly.

We can formulate our problem as a finite union of SI spaces of the form (63) if we assume that the K unknown delays are taken from a discrete grid containing N possible time delays. Under this assumption the generating functions in (63) are expressible as

$$a_\ell(t) = g(t - t_\ell), \quad 1 \leq \ell \leq N. \quad (64)$$

Therefore, assuming a discrete setting, the method of [8] can provide a sampling and reconstruction scheme for a signal of the form (3) with rate $2K/T$.

Similar to our approach here, the sampling scheme in [8] is based on $2K$ parallel channels, each comprised of a filter and a uniform sampler at rate $1/T$. However, in order to achieve this minimum sampling rate, the reconstruction in [8] involves brute-force solving an optimization problem with combinatorial complexity. The complexity of the reconstruction stage can be reduced by increasing the number of channels, which entails a price in terms of sampling rate. In contrast, our reconstruction algorithm is based on the ESPRIT algorithm and can obtain the minimal sampling rate of $2K/T$ in polynomial complexity. Furthermore, we do not require discretization of the time delays

but rather can accommodate any continuous set of delays. In this sense we can view our sampling paradigm as a special case of compressed sensing for an infinite union of SI spaces. Since previous work in this area has focused on sampling methods for finite unions, this appears to be a first systematic example of a sampling theory where the subspace is chosen over an infinite union.

Another difference with the approach of [8] is the design of the sampling filters. In our method, simple sampling filters can be used, such as low pass filter or bandpass filter-bank. In contrast, the scheme of [8] requires proper design of the sampling filters, which is obtained in two stages. In the first stage, N filters $h_\ell(t)$, $1 \leq \ell \leq N$ are chosen that satisfy some conditions with respect to the N possible generating functions. At the second stage, a smaller set of $p \geq 2K$ filters $s_i(t)$ is constructed from $h_\ell(t)$, via

$$s_i(t) = \sum_{\ell=1}^N \sum_{m=1}^N \sum_{n \in \mathbb{Z}} \mathbf{A}_{i\ell}^* c_{im}[n] h_m(t - nT) \quad (65)$$

where $\mathbf{A}$ is a $p \times N$ matrix that satisfies the requirements of compressed sensing in the appropriate dimension [7], and $c_{im}[n]$ are a set of sequences given explicitly in [8]. In order to arrive at filters that are easy to implement, a careful choice of the parameters is needed, which may be difficult to obtain.

C. Signals With Finite Rate of Innovation

Another interesting class of signals that has been treated recently in the sampling literature are FRI signals [13], [14]. Such signals have a finite number of degrees of freedom per unit time, referred to as the rate of innovation. Examples of FRI signals include streams of diracs, nonuniform splines, and piecewise polynomials. A general form of an FRI signal is given by [13]

$$x(t) = \sum_{n \in \mathbb{Z}} c_n \phi(t - t_n), \quad (66)$$

where $\phi(t)$ is a known function, t_n are unknown time shifts and c_n are unknown weighing coefficients. Recovery of such signals from their samples is equivalent to the recovery of the delays t_n and the weights c_n .

Our signal model (3) can be seen as a special case of (66), where additional shift invariant structure is imposed. This means that in each period T the time delays are constant relative to the beginning of the period, whereas in a general FRI signal the time delay can vary from period to period. Our method is designed in such a way that it utilizes this extra structure to reduce the rate, while still guaranteeing perfect recovery.

The FRI signals dealt with in [13] and [14] are divided into three main classes: periodic, finite length and infinite length. If we address our signal model as an FRI signal it will generally fall into the third category of infinite length FRI signals. Some special classes of finite (and periodic) FRI signals were treated in [13], such as streams of diracs. For these special settings sampling theorems were derived with very specific kernels, that achieve the minimal rate (the rate of innovation). However, these methods are not adapted to the general model (66).

Sampling and reconstruction of infinite length FRI signals was treated in [14]. The method in [14] is based on the use of

specific sampling kernels which have finite time support: kernels that can reproduce polynomials or exponentials. In addition the function $\phi(t)$ is limited to diracs, differentiated diracs, or short pulses with compact support. The reconstruction algorithm proposed in [14] is local, namely it recovers the signal's parameters in each time interval separately. Naive use of this approach in our context has two main disadvantages. First, in our method the unknown delays are recovered from all the samples of the signal $x(t)$. A local algorithm is less robust to noise and does not take the shared information into account. In addition, in terms of computational complexity, in our method all the samples are collected to form a finite size correlation matrix, and then the ESPRIT algorithm is applied once. Using the local algorithm requires applying the annihilating filter method, used for FRI recovery, on each time interval over again.

A final disadvantage of the FRI approach is the higher sampling rate required. In order to discuss the sampling rate achieved by the local algorithm proposed in [14], we limit our discussion to the case where the function $\phi(t)$ is a dirac, which is the main case dealt with in [14]. The theorems for unique recovery of the signal from its samples in [14] require that in each interval of size $2KLT_s$ there are at most K diracs, where L is the support of the sampling kernel and T_s is the sampling period. Since in each interval of size T we have K diracs, it can be easily shown that the minimal sampling rate is $2KL/T$, which is a factor of L larger than the rate achieved by our scheme. For example, when using a B-spline kernel, which is the function with the shortest time support that can reproduce polynomials of a certain order, an order of at least $N = 2K - 1$ is needed, which has time support $L = 2K$. Thus, the sampling rate is $2K$ times larger than our approach.

As discussed in [4], FRI signals can also be viewed as an example of signals that lie in an infinite union of subspaces. We point out here the main difference between our model to FRI signals in that respect. Although a general FRI signal of the form (66) has an infinite number of degrees of freedom, the recovery methods in [14] are local and handle each finite time interval separately. Hence, effectively, these recovery methods treat signals with a finite number of degrees of freedom, and for each interval the signal can be seen as lying in a union of finite dimensional subspaces. This is in contrast to our signal model, which consists of signals that lie in a union of SI subspaces, which have an infinite number of degrees of freedom.

VI. APPLICATION

In this section we describe an application of the proposed signal model to the problem of channel estimation in wireless communication [37]. We assume that the transmitted signal passes through a multipath time-varying channel, and the aim of the receiver is to estimate the channel's parameters from samples of the received signal.

We consider a baseband communication system operating in a multipath fading environment with pulse amplitude modulation (PAM). The data symbols are transmitted at a symbol rate of $1/T$, modulated by a known pulse $g(t)$. The transmitted signal $x_t(t)$ is then given by

$$x_T(t) = \sum_{n=1}^{N_{\text{sym}}} d[n] g(t - nT) \quad (67)$$

where $d[n]$ are the data symbols taken from a finite alphabet, and N_{sym} is the total number of transmitted symbols.

The transmitted signal $x_T(t)$ passes through a baseband time-varying multipath channel whose impulse response is modeled as [38]

$$h(\tau, t) = \sum_{k=1}^K \alpha_k(t) \delta(\tau - \tau_k) \quad (68)$$

where $\alpha_k(t)$ is the path time varying complex gain for the k th multipath propagation path and τ_k is the corresponding time delay. The total number of paths is denoted by K . We assume that the channel is slowly varying relative to the symbol rate, so that the path gains are considered to be constant over one symbol period:

$$\alpha_k(t) = \alpha_k[nT] \quad \text{for } t \in [nT, (n+1)T]. \quad (69)$$

In addition, we assume that the propagation delays are confined to one symbol, i.e., $\tau_k \in [0, T)$. Under these assumptions, the received signal is given by

$$x_R(t) = \sum_{k=1}^K \sum_{n=1}^{N_{\text{sym}}} a_k[n] g(t - \tau_k - nT) + n(t) \quad (70)$$

where

$$a_k[n] = \alpha_k[nT] d[n] \quad (71)$$

and $n(t)$ denotes the channel noise.

The received signal $x_R(t)$ fits the signal model described in (3). Therefore, if the pulse shape $g(t)$ satisfies (32) with $p = 2K$, then our sampling scheme can recover the time delays of the propagation paths. In addition, if the transmitted symbols are known to the receiver, then the time varying path gains are recoverable from the sequences $a_k[n]$.

As a result our sampling scheme can estimate the channel's parameters from samples of the output at a low rate, proportional to the number of paths. As an example, we can look at the channel estimation problem in code-division multiple-access (CDMA) communication. This problem was handled using subspace techniques in [39] and [40]. In these works, the sampling is done at the chip rate $1/T_c$ or above, where T_c is the chip duration given by $T_c = T/N$ and N is the spreading factor which is usually high (1023, for example, in GPS applications). In contrast, our sampling scheme can provide recovery of the channel's parameters at a sampling rate of $2K/T$. For a channel with a small number of paths, this sampling rate can be significantly lower than the chip rate.

Another example is UWB [41] communications which has gained popularity recently. In this technology, the bandwidth of the transmitted pulse can be up to several gigahertz. Current technology commercial ADCs cannot operate at these sampling rates. For example, the highest sampling rate ADC device, manufactured by National Semiconductor, supports sampling rates of up to 3 GHz at a relatively low resolution of 8 bits and high power consumption. Our proposed method, has the potential of reducing the sampling rate, to rates which can be achieved by lower rate ADCs with better resolution and lower power.

VII. NUMERICAL EXPERIMENTS

We now provide several experiments in which we examine various aspects of our method. The numerical experiments are divided into five parts:

- 1) demonstration of a channel estimation application;
- 2) evaluation of performance in the presence of noise;
- 3) effects of approximation of the matrix $\mathbf{R}_{dd}$ using only a finite number of measurement vectors;
- 4) effects of the distance between delays on the performance;
- 5) effects of imperfect digital correction filtering, using finite length filters.

In all the simulations, except for the one in Section VII-E, we use the sampling scheme described in Section III-C-1), which consists of a bank of ideal bandpass filters. We assume that the working band is $\mathcal{F} = [0, 2\pi/Tp]$, and that the function $g(t)$ has constant frequency response on that frequency range, i.e., $G(\omega) = T, \omega \in \mathcal{F}$. In order to improve the robustness to noise in the delay recovery stage, we use the total least-squares (TLS) version of the ESPRIT algorithm described in [22]. All the results are based on averaging 1000 experiments.

A. Channel Estimation

In the first simulation we demonstrate a channel estimation application. We consider a time-varying channel of the form (68), with $K = 4$ paths. In order to simulate a time varying channel, the channel's gain coefficients $\alpha_k[nT]$ are modeled according to the Jakes' model [42] as a zero-mean complex-valued Gaussian stationary process with the classical U-shape power spectral density. In such a model the varying rate of each gain coefficient depends on the maximal Doppler shift f_d . In order to simulate a slow varying channel, relatively to the symbol rate $1/T$, we used for each path a maximal Doppler shift of $f_d = 0.05/T$. The energy of each time-varying path gain coefficient was normalized to $(1/2)^{-k+1}$. The path delays were drawn uniformly in the range $[0, T]$. For the estimation $N_{\text{sym}} = 100$ symbols were used where the data symbols are assumed to be known. The samples at the output of each of the sampling channels were corrupted by complex-valued Gaussian white noise with an SNR of 20 dB.

The number of sampling channels is taken to be $p = 5$, which is only one more than the number of unknown delays. Although we have seen that $2K$ sampling channels are required for perfect recovery of every signal of the form (3), for some signals lowering the number of sampling channels is possible. Indeed, according to Proposition 2, for signals with $\dim(\text{span}(\mathbf{b}[\Lambda])) = K$, the minimal number of sampling channels required is $K+1$. We will demonstrate that for this example, $K+1$ channels are sufficient.

In Fig. 4 the original and estimated channels are shown. Since the gain coefficients of the channel are time-varying, only their averaged energy over time is shown in the figure. In Fig. 5, we plot the magnitude of the original and estimated gains of the first path versus time. From Figs. 4 and 5 it is evident that our method can provide a good estimate of the channel's parameters, even when the samples are noisy.

B. Performance in the Presence of Noise

In the next simulations we examine the effect of SNR and the number of sampling channels on the error in the delays es-

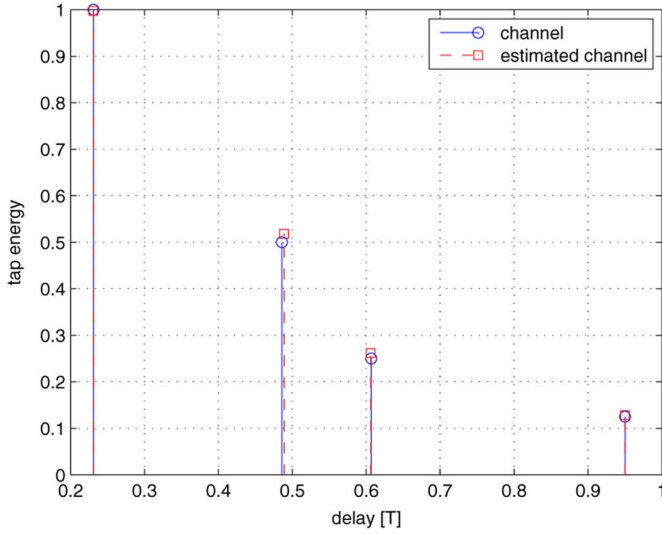


Fig. 4. Channel estimation with $p = 5$ sampling channels, and SNR = 20 dB.

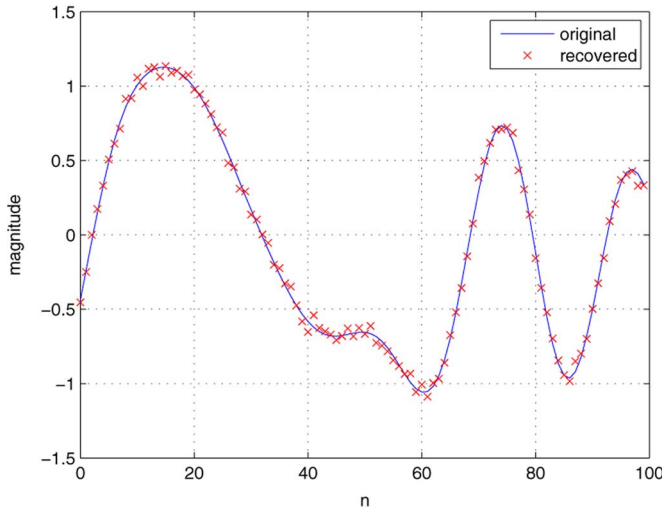


Fig. 5. Estimation of the time-varying gain coefficient of the first path, $p = 5$, SNR = 20 dB.

timization. We choose $K = 2$ close delays, $t_1 = 0.4352T$ and $t_2 = 0.521T$. The sequences $a_k[n]$, $k = 1, 2$, $n = 1, 2, \dots, 100$ are finite length sequences with unit power chosen according to Jakes' model with $f_d = 0.05/T$.

Under the setting of the simulation, which consists of a pulse with constant frequency response and ideal bandpass filters, from (23) it can be verified that the sampling sequences satisfy the following relation in the time domain:

$$\mathbf{c}[n] = \mathbf{N}(\tau)\mathbf{b}[n], \quad n \in \mathbb{Z}. \quad (72)$$

The Cramér–Rao bound (CRB) for unbiased estimators of $\theta_k = 2\pi/Tt_k$ from the data $\mathbf{c}[n]$, was derived in [43] for this data model. The TLS-ESPRIT algorithm, used for the delays estimation in our method, is known to be asymptotically unbiased [44]. Experimentally we verified that, under the simulation setup, the bias of the delays estimation is low enough for SNRs above 15 dB. Therefore, in this range of SNRs, the CRB derived in [43] can give a lower bound on the MSE of the delays estimation (up to factor of $2\pi/T$), assuming our specific sampling scheme.

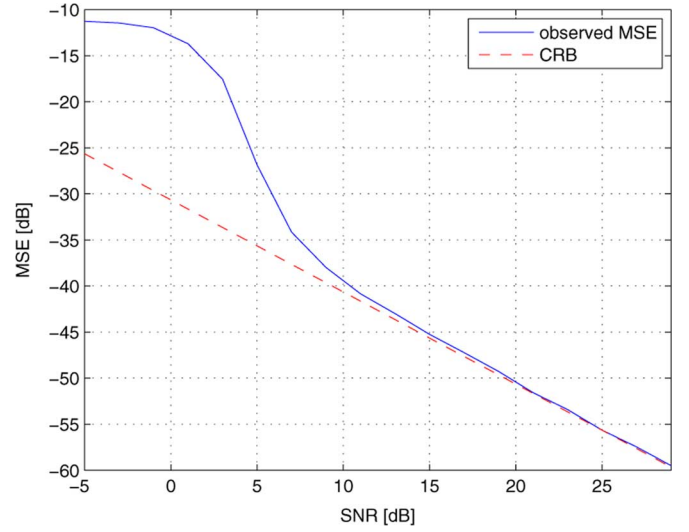


Fig. 6. MSE of the delays estimation versus SNR, for $K = 2$ and $p = 4$.

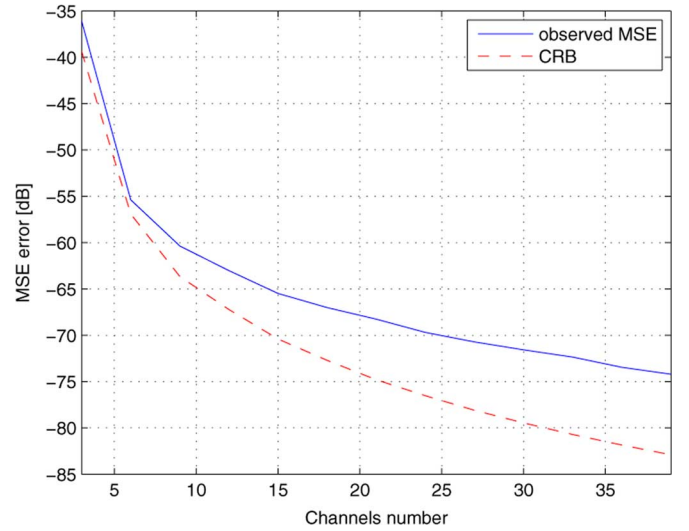


Fig. 7. MSE of the delays estimation versus the number of sampling channels p , for $K = 2$ and SNR = 10 dB.

In Fig. 6, the mean-squared error (MSE) of the time-delays estimation is shown versus the SNR, when using $p = 4$ sampling channels. For comparison we also plot the CRB. The figure demonstrates that our method achieves the CRB for SNR > 15 dB, which is the range that delays estimation can be considered as unbiased.

In Fig. 7, the MSE of the estimation of the time delays is shown versus the number of sampling channels, for a constant SNR of 10 dB. The results demonstrate that the estimation error can be improved by increasing the number of channels. Therefore, oversampling improves the robustness of our method to noise.

C. Effects of Imperfect Approximation of $\mathbf{R}_{dd}$

Next, we investigate the influence of estimating the matrix $\mathbf{R}_{dd}$ using only a finite number of measurement vectors $\mathbf{d}[n]$. This number affects the total delay of our method, since reconstruction of the sequences $a_k[n]$ is performed only after the unknown delays τ are recovered. In Fig. 8 the MSE of the delays

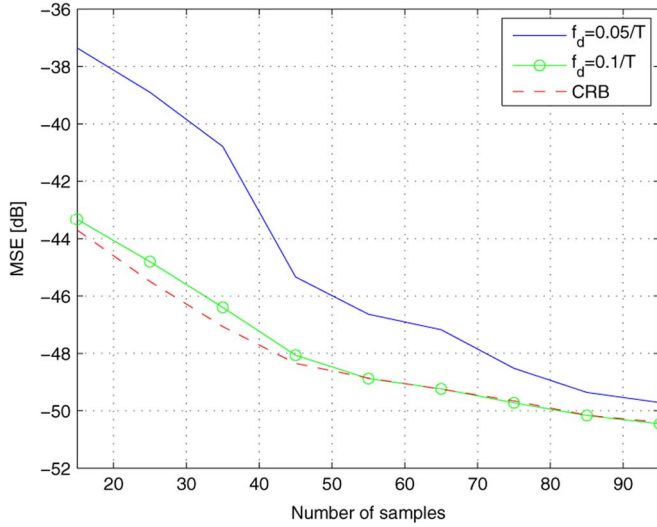


Fig. 8. MSE of the delays estimation versus the number of samples used, for $K = 2$, $p = 4$, and $\text{SNR} = 20$ dB.

estimation is shown versus the number of measurement vectors used for estimation of $\mathbf{R}_{dd}$. A constant SNR of 20 dB and $p = 4$ sampling channels are used. Two cases are illustrated: in the first, the sequences $a_k[n]$ are taken according to the Jakes' model with parameter $f_d = 0.05/T$ and in the second case $f_d = 0.1/T$ is used, which corresponds to sequences with faster variation rate. Fig. 8 demonstrates that the MSE depends on the variation rate of the sequences. Intuitively the faster the sequences vary, the more information each new measurement vector $\mathbf{d}[n]$ contains, improving the estimation of $\mathbf{R}_{dd}$. In addition, it can be seen that using only 50 measurement vectors, yields a reasonable estimation of the delays in the case of $f_d = 0.1/T$. The same estimation error is achieved using 80 measurement vectors, when using slow varying sequences. This result can be further improved by increasing the SNR or the number of sampling channels.

D. Effects of the Distance Between Delays

In the following simulation we examine the effects of the distance between the delays on the performance of our method. We choose $K = 2$ delays with varying distance, $p = 4$ sampling channels and a constant SNR of 25 dB. In Fig. 9 the MSE of the delays estimation is shown as a function of the delays distance. The figure shows that the estimation error decreases as the distance between delays grows. It can also be shown that good performance is achieved even for very close delays with a distance of $0.05T$.

Next, we investigate the affect of the distance on the recovery of the sequences $a_k[n]$. Since this recovery depends also on the estimation of delays, it suffers from two sources of errors: errors in the estimation of $\mathbf{N}(\tau)$ and errors caused by the noise. To distinguish between these two errors, we plot the MSE in estimating the sequences when using the estimated delays, and when using the true delay values. The results are shown in Fig. 10. For small distances between the delays, the error in the estimation of the matrix $\mathbf{N}(\tau)$ is dominant. For distances above $0.1T$ the two graphs almost coincide. In addition, as can be seen from the plot that uses the true delay values, for

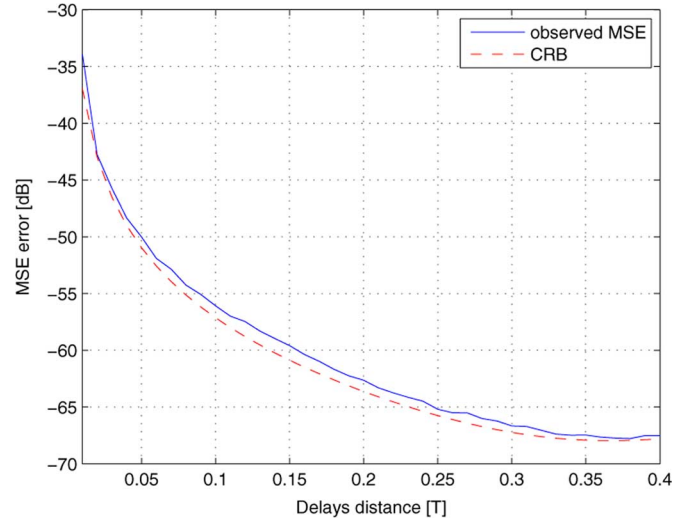


Fig. 9. MSE of the delays estimation versus the distance between delays, for $K = 2$, $p = 4$ and $\text{SNR} = 25$ dB.

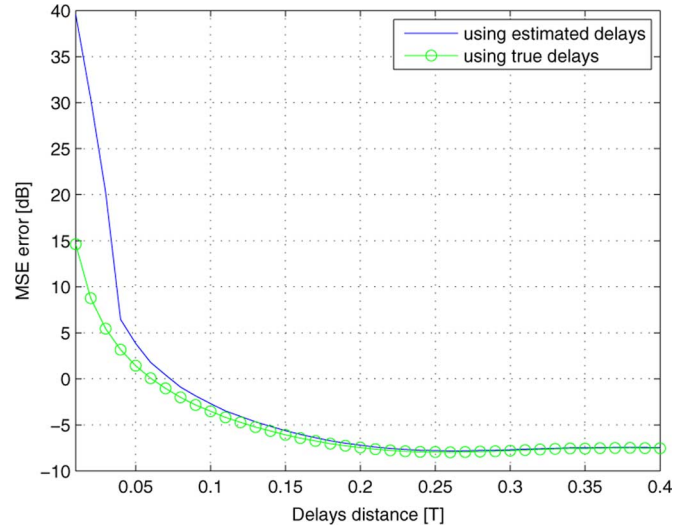


Fig. 10. MSE of the sequences estimation versus the distance between delays, for $K = 2$, $p = 4$ and $\text{SNR} = 25$ dB.

close delays there is a noise enhancement caused by the ill conditioning of the matrix $\mathbf{N}(\tau)$. This enhancement reduces as the distance between the delays grows, and for distances above $0.15T$ this phenomenon is negligible.

E. Effects of Imperfect Digital Filtering Correction

In the next simulation we examine the effects of approximating the digital correction filter bank $\mathbf{W}^{-1}(e^{j\omega T})$ by finite length digital filters. The length of the filters affects the delay of our scheme. To demonstrate this point, we arbitrarily choose a sampling scheme composed of three non-ideal bandpass filters with a frequency response given by

$$S_\ell(\omega) = \begin{cases} 1.1 - (1 - 0.4\ell) \cos(\omega - \frac{2\pi}{T}\ell), & \omega \in \mathcal{F}_\ell \\ 0, & \text{otherwise} \end{cases} \quad (73)$$

where

$$\mathcal{F}_\ell = \left[(\ell - 1) \frac{2\pi}{T}, \ell \frac{2\pi}{T} \right]. \quad (74)$$

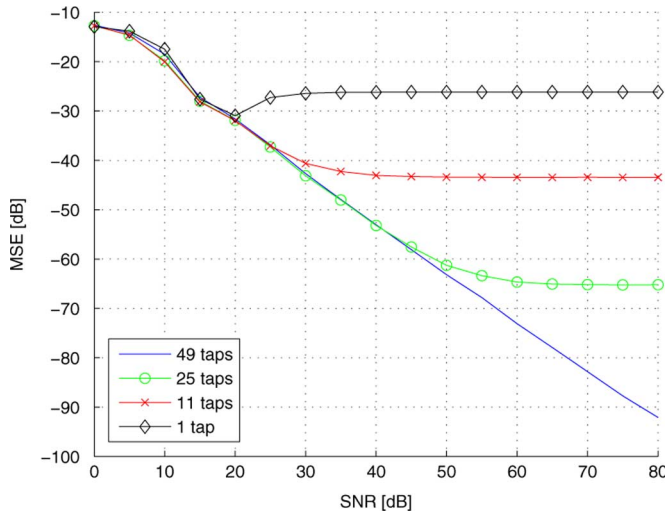


Fig. 11. MSE of the delays estimation versus SNR for different lengths of digital correction filter bank approximations.

These filters satisfy the conditions of Proposition 1 and can model realistic sampling filters with nonflat frequency response. In this case a non trivial digital correction filter bank is required, whose coefficients are calculated using the inverse DTFT of $\mathbf{W}^{-1}(e^{j\omega T})$.

In Fig. 11, the MSE of the delays estimation versus the SNR is plotted for different lengths of filters. At low SNRs the dominant error is caused by the noise, while for high SNRs the error is mostly a result of the correction filter approximation. It can be seen that a 49 taps filter provides a good approximation to the correction filter bank, resulting in a delay of 24 samples. When working at SNRs below 40 dB, filters with 11 taps seem reasonable.

VIII. CONCLUSION

In this paper, we considered the problem of estimating the time delays and time varying coefficients of a multipath channel, from low-rate samples of the received signal. We showed that this problem can be formulated within the broader context of sampling theory, in which our goal is to recover an analog signal $x(t)$ that lies in a SI subspace, spanned by K generators with unknown delays. This class of problems can be viewed as an infinite union of subspaces.

We showed that if the channel has K multipath components, or equivalently, if the SI subspace is generated by K generators, than under appropriate conditions on the sampling filters, a sampling rate of $2K/T$ is necessary and sufficient to guarantee perfect recovery of any signal $x(t)$. Here T is the transmission rate, or the period of the generators. We developed sufficient conditions on the generators and the sampling filters in order to guarantee perfect recovery at the minimal possible rate. To recover the unknown time delays, we showed that our problem can be formulated within the context of DOA estimation. Using this relationship, we proposed an ESPRIT-type algorithm to determine the unknown delays from the given low-rate samples. Once the delays are properly identified, the time varying coefficients can be found by digital filtering.

Besides the application to time-delay estimation, the problem we treated here can be seen as a first example of a sampling

theory for analog signals defined over an infinite union of SI subspaces. Recently, there has been growing interest in sampling theorems for signals over a union of subspaces [4]–[6], [8]–[10], [12]. However, previous work addressing stability issues and concrete recovery algorithms have focused on finite unions or finite dimensional subspaces. Here, we take a first step in the direction of extending these ideas to a broader setting that treats analog signals lying in an infinite union of SI subspaces.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

- [1] A. Quazi, "An overview on the time delay estimate in active and passive systems for target localization," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 527–533, 1981.
- [2] R. J. Urick, *Principles of Underwater Sound*. New York: McGraw-Hill, 1983.
- [3] G. L. Turin, "Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio," *Proc. IEEE*, vol. 68, no. 3, pp. 328–353, Mar. 1980.
- [4] Y. M. Lu and M. N. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2334–2345, Jun. 2008.
- [5] T. Blumensath and M. E. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Inf. Theory*, vol. 55, pp. 1872–1882, Apr. 2009.
- [6] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.
- [7] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [8] Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 2986–2997, Aug. 2009.
- [9] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, pp. 993–1009, Mar. 2009.
- [10] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
- [11] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.
- [12] Y. C. Eldar, "Uncertainty relations for shift-invariant analog signals," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5742–5757, Dec. 2009.
- [13] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
- [14] 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 Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [15] A. Bruckstein, T. J. Shan, and T. Kailath, "The resolution of overlapping echos," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1357–1367, 1985.
- [16] M. A. Pallas and G. Jourdain, "Active high resolution time delay estimation for large BT signals," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 781–788, Apr. 1991.
- [17] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [18] Z. Q. Hou and Z. D. Wu, "A new method for high resolution estimation of time delay," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 1982, vol. 7, pp. 420–423.
- [19] H. Saarnisaari, "TLS-ESPRIT in a time delay estimation," in *Proc. IEEE 47th Veh. Technol. Conf.*, May 1997, vol. 3, pp. 1619–1623.
- [20] F.-X. Ge, D. Shen, Y. Peng, and V. O. K. Li, "Super-resolution time delay estimation in multipath environments," *IEEE Trans. Circuits Syst.*, vol. 54, no. 9, pp. 1977–1986, Sep. 2007.

- [21] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 1, pp. 134–139, Jan. 1983.
- [22] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
- [23] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 48–68, May 2009.
- [24] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [25] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
- [26] A. Aldroubi and K. Gröchenig, "Non-uniform sampling and reconstruction in shift-invariant spaces," *SIAM Rev.*, vol. 43, pp. 585–620, 2001.
- [27] C. de Boor, R. A. DeVore, and A. Ron, "The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$," *J. Funct. Anal.*, vol. 119, no. 1, pp. 37–78, 1994.
- [28] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *J. Approx. Theory*, vol. 78, no. 3, pp. 373–401, 1994.
- [29] O. Christensen and Y. C. Eldar, "Generalized shift-invariant systems and frames for subspaces," *J. Fourier Anal. Appl.*, vol. 11, no. 3, pp. 299–313, 2005.
- [30] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.
- [31] E. Moulines, P. Duhamel, J. Cardoso, S. Mayrargue, and T. Paris, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.
- [32] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 996–1000, Jul. 1989.
- [33] M. Z. Win and R. A. Scholtz, "Characterization of ultra-wide bandwidth wireless indoor channels: A communication-theoretic view," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 9, pp. 1613–1627, Dec. 2002.
- [34] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, Nov. 2005.
- [35] J. E. Evans, D. F. Sun, and J. R. Johnson, "Application of advanced signal processing techniques to angle of arrival estimation in ATC Navigation and surveillance systems," MIT, Lincoln Lab., Lexington, MA, Tech. Rep., 1982.
- [36] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, Aug. 1985.
- [37] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York: Wiley, 1997.
- [38] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
- [39] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 1009–1020, 1996.
- [40] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten, "DS-CDMA synchronization in time-varying fading channels," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 8, pp. 1636–1642, Oct. 1996.
- [41] L. Yang and G. B. Giannakis, "Ultra-wideband communications: An idea whose time has come," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 26–54, 2004.
- [42] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
- [43] P. Stoica and N. A. Aas, "Music, maximum likelihood, and Cramer–Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.
- [44] B. Ottersten, M. Viberg, and T. Kailath, "Performance analysis of the total least squares ESPRIT algorithm," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1122–1135, May 1991.



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 (S'98–M'02–SM'07) 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 with the Electrical and Statistics departments at Stanford University, Stanford, CA.

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, and in 2009 the Technion's Award for Excellence in Teaching. 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*.