MEASUREMENT MATRIX DESIGN FOR COMPRESSED SENSING BASED TIME DELAY ESTIMATION

Florian Roemer[1,2]*, Mohamed Ibrahim*[1,2]*, Norbert Franke*[2]*,*
Niels Hadaschik[2]*, Andreas Eidloth*[2]*, Benjamin Sackenreuter*[2]*, Giovanni Del Galdo*[1,2]
[1]Technische Universität Ilmenau, Institute for Information Technology
[2]Fraunhofer Institute for Integrated Circuits IIS, Ilmenau, Germany

***Abstract* – In this paper we study the problem of estimating the unknown delay(s) in a system where we receive a linear combination of several delayed copies of a known transmitted waveform. This problem arises in many applications such as timing-based localization or wireless synchronization. Since accurate delay estimation requires wideband signals, traditional systems need high-speed AD converters which poses a significant burden on the hardware implementation. Compressive sensing (CS) based system architectures that take measurements at rates significantly below the Nyquist rate and yet achieve accurate delay estimation have been proposed with the goal to alleviate the hardware complexity. In this paper, we particularly discuss the design of the measurement kernels based on a frequency-domain representation and show numerically that an optimized choice can outperform randomly chosen functionals in terms of the delay estimation accuracy.**

Keywords: *Compressive sensing, synchronization, delay estimation, measurement matrix design*

1. INTRODUCTION

Time delay estimation is a fundamental challenge arising in many applications such as wireless communications [1], radar [2], localization [3], sensor networks [4], and many others. A wide range of delay estimation techniques have been proposed, see, e.g., [5] for a survey for time delay estimation in linear dynamic systems. In this paper, we treat the special case of estimating the delay of a signal with a known pulse shape from a noisy superposition of several delayed copies. Such problems occur in synchronization of wireless sensor networks [6], TDOA-based localization [7] or wideband channel estimation [8] due to the multipath channel.

The complexity bottleneck in these systems is the high data rate from the high-speed AD converters that are required to sample the wideband signals which are used to ensure a high resolution in time. However, since the pulse shape is known, the actual rate of (unknown) information in the received signals is low. Therefore, it has been proposed to apply compressed sensing (CS) to reduce the hardware complexity while maintaining a high precision. In [9, 10], a framework for sampling time-delayed signals is presented based on a union of subspaces approach [11]. The authors derive sufficient conditions on the transmitted pulse and the sampling functions in order to ensure perfect recovery of the channel parameters in the absence of noise, which includes conditions on the minimal required sampling rate.

Considering concrete choices of the low rate sampling kernels, it is often suggested in the CS literature to use random kernels which are incoherent with any basis and so achieving informative measurements even at low rates. Though randomly chosen kernels represent a simple and generic approach, it is known that they do not provide the optimal robustness against noise. Only recently, the optimization of the measurement kernels has been investigated [12, 13, 14]. In particular, [12] studies the optimization of a discrete measurement matrix in time domain and [13] consider a continuous (sum-of-sincs) kernel in time domain whose output that is sampled at a sub-Nyquist rate. Both use criteria inspired by the Bayesian Cramér-Rao bound to optimize the kernels. In [14], the measurement matrix is optimized such that for a given (overcomplete) basis, the sensing matrix has a small average coherence.

In this paper, we propose another optimization based design for the measurement kernels of the compressed sampling based time delay estimation architectures. We consider an architecture based on a bank of K continuous-time periodic functionals that are sampled once per period. We show that their Fourier-domain representation allows to optimize these functions based on a finite number of coefficients. We demonstrate numerically that the optimized CS kernels outperform a randomly chosen one in terms of the delay estimation accuracy.

2. SYSTEM MODEL

The time delay estimation problem can be formalized as follows. A transmitter sends a signal $s(t)$ to allow a receiver (or multiple receivers) that know $s(t)$ to synchronize themselves. For simplicity, we assume that $s(t)$ is t_p-periodic, i.e., $s(t) = s(t + t_\mathrm{p})$. Periodic synchronization signals are common, e.g., in GNSS [15] and UWB Radar applications [16]. They allow receivers to keep track of the synchronization over time and to average over multiple periods. For the problem at hand, another advantage is that since these signals are also band-limited to a certain bandwidth B, they can be completely described by a finite number of $M = B \cdot t_p$ coefficients. Therefore, we can write $s(t)$ as

$$s(t) = \sum_{m \in \mathcal{M}} c_s[m] \mathrm{e}^{\jmath 2\pi m \frac{t}{t_\mathrm{p}}} \tag{1}$$

where $\mathcal{M}$ is the set of points in frequency where $s(t)$ is non-zero[1] with $|\mathcal{M}| = M$.

Due to the multipath nature of wireless propagation channels, the receiver observes a weighted sum of delayed copies of $s(t)$, i.e.,

This work was partially supported by the Carl-Zeiss Foundation under the postdoctoral scholarship project "EMBiCoS".

[1]We employ a complex representation of all signals, which can refer to the complex low-pass domain where $\mathcal{M} = \{-\frac{M}{2} + 1, \ldots, \frac{M}{2}\}$ or analytic signals where only the positive half of the spectrum is considered.

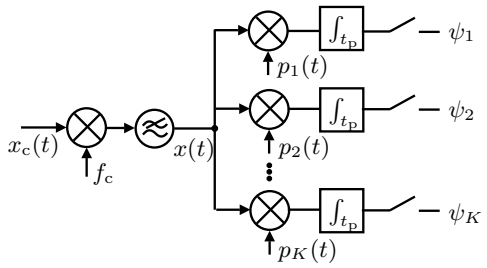

Fig. 1. Compressive receiver architecture: the signal is multiplied with K periodic waveforms $p_k(t)$ and the result is sampled once per period in each branch.

the received signal can be written as

$$x(t) = \sum_{\ell=1}^{L} \alpha_\ell s(t-\tau_\ell) + w(t), \tag{2}$$

where α_ℓ and τ_ℓ represent the complex amplitude and delay of the ℓ-th propagation path for $\ell = 1, 2, \ldots, L$, respectively, and $w(t)$ is the additive white Gaussian measurement noise. Note that since $s(t)$ is t_p-periodic so is $x(t)$ (besides for the additive noise).

We employ a compressive sampling architecture similar to [9], which is depicted in Figure 1. After an initial downconversion stage to remove the carrier frequency[2], instead of sampling the received signal $x(t)$ directly, the receiver performs a bank of analog multiplications with the functions $p_k(t)$, $k = 1, 2, \ldots, K$, followed by a sampling with a lower rate to yield the coefficients ψ_k. Since $x(t)$ is periodic, it is convenient to choose $p_k(t)$ such that they are also t_p-periodic. The overall sampling rate of the receiver is then equal to $\frac{K}{t_\mathrm{p}} = \frac{K}{M} \cdot B$ which means that the compression factor relative to Nyquist sampling is equal to $\frac{K}{M}$. Note that each of the K ADCs operates at a rate $\frac{B}{M}$, i.e., a factor M times lower than in a conventional system.

It is important to note that since our receiver is linear and $x(t)$ is a linear combination of delayed copies of $s(t)$, it is sufficient to describe the sampling of one delayed signal $s(t-\tau)$. Once a mathematical model for this has been obtained, the sampled version of $x(t)$ follows from a trivial linear combination.

Therefore, if we let $x(t) = s(t-\tau)$, the k-th sample ψ_k for $k = 1, 2, \ldots, K$ is given by

$$\psi_k(\tau) = \frac{1}{t_\mathrm{p}} \int_{t_\mathrm{p}} p_k(t)^* \cdot s(t-\tau)\mathrm{d}t. \tag{3}$$

As both, $p_k(t)$ and $s(t)$ are periodic, they can be expressed using a discrete representation in the Fourier domain, namely

$$p_k(t) = \sum_{m=-\infty}^{\infty} c_{p,k}[m] \mathrm{e}^{\jmath 2\pi m \frac{t}{t_\mathrm{p}}}. \tag{4}$$

With the help of (4) and (1), we can now rewrite (3) in frequency domain [9]. For clarity, let us begin with $\tau = 0$ to obtain

$$\begin{aligned}
\psi_k(0) &= \frac{1}{t_\mathrm{p}} \int_{t_\mathrm{p}} \left(\sum_{m_1=-\infty}^{\infty} c_{p,k}^*[m_1] \mathrm{e}^{-\jmath 2\pi m_1 \frac{t}{t_\mathrm{p}}} \right) \\
&\qquad \cdot \left(\sum_{m_2 \in \mathcal{M}} c_s[m_2] \mathrm{e}^{\jmath 2\pi m_2 \frac{t}{t_\mathrm{p}}} \right) \mathrm{d}t \\
&= \sum_{m_1=-\infty}^{\infty} \sum_{m_2 \in \mathcal{M}} c_{p,k}^*[m_1] c_s[m_2] \frac{1}{t_\mathrm{p}} \underbrace{\int \mathrm{e}^{-\jmath 2\pi m_1 \frac{t}{t_\mathrm{p}}} \mathrm{e}^{\jmath 2\pi m_2 \frac{t}{t_\mathrm{p}}} \mathrm{d}t}_{t_p \cdot \delta[m_1 - m_2]} \\
&= \sum_{m \in \mathcal{M}} c_{p,k}^*[m] c_s[m]. \qquad (5)
\end{aligned}$$

Equation (5) shows, not surprisingly, that only the M coefficients of $p_k(t)$ that coincide with the spectral support of $s(t)$ contribute to the coefficients ψ_k. Therefore, if the hardware realization allows it, they should be chosen such that they are bandlimited to B as well[3].

Introducing the short hand notations $\boldsymbol{c}_s \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{c}_{p,k} \in \mathbb{C}^{M\times 1}$ for the M coefficients that $c_s[m]$ and $c_{p,k}[m]$ in (5) we then have $\psi_k(0) = \boldsymbol{c}_{p,k}^\mathrm{H} \cdot \boldsymbol{c}_s$. Moreover, it is now easy to carry out the same computation for $\tau \neq 0$. We obtain

$$\begin{aligned}
\psi_k(\tau) &= \sum_{m \in \mathcal{M}} c_{p,k}^*[m] \mathrm{e}^{-\jmath 2\pi m \frac{\tau}{t_\mathrm{p}}} c_s[m] \\
&= \boldsymbol{c}_{p,k}^\mathrm{H} \cdot \mathrm{diag}\{\boldsymbol{d}(\tau)\} \cdot \boldsymbol{c}_s, \qquad (6) \\
&= \boldsymbol{c}_{p,k}^\mathrm{H} \cdot \mathrm{diag}\{\boldsymbol{c}_s\} \cdot \boldsymbol{d}(\tau), \qquad (7)
\end{aligned}$$

where we have defined the vector $\boldsymbol{d}(\tau) = \left[\mathrm{e}^{-\jmath 2\pi m \frac{\tau}{t_\mathrm{p}}}\right]_{m \in \mathcal{M}}$.

Based on (7), the entire vector of observations $\boldsymbol{\psi}(\tau) \in \mathbb{C}^{K\times 1}$ can be described as

$$\boldsymbol{\psi}(\tau) = [\psi_1(\tau), \ \ldots, \ \psi_K(\tau)]^\mathrm{T} = \boldsymbol{C}_p^\mathrm{H} \cdot \mathrm{diag}\{\boldsymbol{c}_s\} \cdot \boldsymbol{d}(\tau), \tag{8}$$

where $\boldsymbol{C}_p = [\boldsymbol{c}_{p,1}, \ldots, \boldsymbol{c}_{p,K}] \in \mathbb{C}^{M\times K}$ contains the coefficients of all the K sequences $p_k(t)$. Equation (8) describes the observed output vector for a single delayed copy of $s(t)$. Since our receiver is linear, the observed vector $\boldsymbol{y} \in \mathbb{C}^{K\times 1}$ for the input signal $x(t)$ according to (2) is given by

$$\boldsymbol{y} = \sum_{\ell=1}^{L} \alpha_\ell \boldsymbol{\psi}(\tau_\ell) + \tilde{\boldsymbol{w}}, \tag{9}$$

where $\tilde{\boldsymbol{w}} \in \mathbb{C}^{K\times 1}$ is the effective noise vector.

3. DELAY ESTIMATION PROCEDURE

3.1. Gridded sparse recovery based estimator

Equation (9) shows that our observation vector can be modeled as a weighted sum of L terms $\boldsymbol{\psi}(\tau_\ell)$ under additive noise. Since $\boldsymbol{\psi}(\tau)$ is known to the receiver, this suggests that the delays can be recovered from $\boldsymbol{y}$ if $L < K$. The difficulty in estimating the delays lies in the fact that they can take any value from a continuous domain. A common and very simple approach to tackle such problems is to discretize the parameter space into an N-point sampling grid in τ

[2]The transmitter sends the bandwidth-B signal by modulating it onto a carrier frequency f_c, which is removed at this stage. The signal model is described entirely in the complex low-pass domain.

[3]Compressive sampling architectures that use PN-sequences for the $p_k(t)$ have been proposed, e.g., the modulated wideband converter [9]. Although these are not strictly bandlimited to B, their practical advantage is that they can be realized in hardware up to very high switching speeds.

referred to as $\tau_n^{(\mathrm{G})}$, $n = 1, 2, \ldots, N$. In the special case of a uniform sampling grid, we have $\tau_n^{(\mathrm{G})} = (n-1) \cdot \Delta\tau$, $n = 1, 2, \ldots, N$ where $\Delta\tau = \frac{t_\mathrm{p}}{N}$. Based on the sampled delays, we can define a basis $\boldsymbol{\Psi} \in \mathbb{C}^{K \times N}$ according to

$$\boldsymbol{\Psi} = \left[\boldsymbol{\psi}\left(\tau_1^{(\mathrm{G})}\right), \quad \boldsymbol{\psi}\left(\tau_2^{(\mathrm{G})}\right), \quad \ldots, \quad \boldsymbol{\psi}\left(\tau_N^{(\mathrm{G})}\right)\right]. \tag{10}$$

Since the delays τ_ℓ can take any value, they will not be on any predefined sampling grid almost surely. However, it has been shown that if the sampling of the grid is not too coarse, one can still use the fact that $\boldsymbol{y}$ is approximately sparse in $\boldsymbol{\Psi}$ and apply suitable grid offset estimation procedures to correct for the mismatch between the grid points and the actual delays, cf. e.g., [17] for a comparison of interpolation strategies in this setting. Therefore, to facilitate the further explanations we now assume that all the delays τ_ℓ were exactly on the sampling grid, i.e., $\tau_\ell = \tau_{d_\ell}^{(\mathrm{G})}$ for some $d_\ell \in \{1, 2, \ldots, N\}$. This allows us to write $\boldsymbol{y}$ as

$$\boldsymbol{y} = \boldsymbol{\Psi} \cdot \boldsymbol{\alpha} + \tilde{\boldsymbol{w}}, \tag{11}$$

where $\boldsymbol{\alpha} \in \mathbb{C}^{N \times 1}$ contains α_ℓ at the indices d_ℓ for $\ell = 1, 2, \ldots, L$ and zeros otherwise. In other words, besides for the noise, $\boldsymbol{y}$ is L-sparse in $\boldsymbol{\Psi}$. The delay estimation problem can then be cast as a sparse recovery problem, e.g.,

$$\min \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \quad \|\boldsymbol{y} - \boldsymbol{\Psi} \cdot \boldsymbol{\alpha}\|_2^2 \leq \sigma^2, \tag{12}$$

where σ^2 is an estimate of the noise power. Note that it has been shown in [9] that $K \geq 2L$ is sufficient to recover all the delays from $\boldsymbol{y}$ in the noise-free case.

3.2. Correlation based estimator

A simpler estimator is inspired by traditional, Nyquist-sampling based systems which simply correlate the observed signal with the known waveform and then estimate the location of the peak. Note that if there is only one path ($L = 1$) or if paths are very well separated (by more than the width of the correlation peak), this method is in fact optimal. Along these lines, the correlation based estimator in our setting is defined as

$$\hat{\tau} = \arg\max_\tau \left| \frac{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \boldsymbol{y}}{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau)} \right|, \tag{13}$$

where the peak search is not limited to a prespecified grid. It is instructive to expand (13) for the special case $L = 1$ where we obtain

$$\hat{\tau} = \arg\max_\tau \ \alpha_1 \cdot \underbrace{\frac{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau_1)}{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau)}}_{\rho^{(\mathrm{c})}(\tau)} + \frac{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \tilde{\boldsymbol{w}}}{\boldsymbol{\psi}(\tau)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau)}. \tag{14}$$

Equation (14) shows that we are essentially finding the peak in the "compressed" correlation function $\rho^{(\mathrm{c})}(\tau)$.

4. MEASUREMENT DESIGN

The previous sections have shown how the compressive-sensing based receiver can be employed for delay estimation. As we have seen, it provides a sparsifying basis $\boldsymbol{\psi}(\tau)$ for the signal and the atoms $\boldsymbol{\psi}(\tau)$ depend on the choice of the signal $s(t)$ (through the vector $\boldsymbol{c}_s$) and the sampling functions $p_k(t)$ (through the matrix $\boldsymbol{C}_p$). In this section we shed some light on their proper choice to obtain a good synchronization performance.

An "ideal" choice of $\boldsymbol{\psi}(\tau)$ would satisfy the conditions

$$\boldsymbol{\psi}(\tau_1)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau_2) \approx \begin{cases} 0 & \tau_1 \neq \tau_2 \\ \text{const} & \tau_1 = \tau_2, \end{cases} \tag{15}$$

where the first condition asks for good cross-correlation properties between different delays and the second condition guarantees that the measurement is equally sensitive to all possible delays. This choice will ensure that the compressed correlation function $\rho^{(\mathrm{c})}(\tau)$ introduced in (14) is close to the ideal delta function. At the same time, it is also beneficial for the gridded sparse recovery based estimator since the first condition asks for low correlation between the columns of the sensing matrix $\boldsymbol{\Psi}$ (i.e., a low coherence) and the second condition guarantees that all columns have similar norms (to achieve a uniform sensitivity for all possible delays).

To measure how well a given matrix $\boldsymbol{\Phi}$ satisfies (15), we can formulate an error measure of the form $e(\boldsymbol{\Phi}, \tau_1, \tau_2) = (\boldsymbol{\psi}(\tau_1)^\mathrm{H} \cdot \boldsymbol{\psi}(\tau_2) - C \cdot \delta[\tau_1 - \tau_2]) \cdot \omega_{\tau_1,\tau_2}$ where $\omega_{\tau_1,\tau_2} \in \mathbb{R}_{\geq 0}$ is a weight function which allows to trade the weight between certain (τ_1, τ_2) regions, e.g., between uniform sensitivity ($\tau_1 = \tau_2$) and low cross-correlation ($\tau_1 \neq \tau_2$). However, since it is difficult to minimize the error over the continuous variables τ_1 and τ_2 we consider it only on the N-point sampling grid introduced earlier. This leads to an error matrix $\boldsymbol{E}(\boldsymbol{\Phi}) \in \mathbb{R}^{N \times N}$ given by

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{\Phi}) &= \left[e(\boldsymbol{\Phi}, \tau_{n_1}^{(\mathrm{G})}, \tau_{n_2}^{(\mathrm{G})})\right]_{(n_1,n_2=1,2,\ldots,N)} \\ &= \left(\boldsymbol{\Psi}^\mathrm{H} \cdot \boldsymbol{\Psi} - C \cdot \boldsymbol{I}_N\right) \odot \boldsymbol{\Omega}, \\ &= \left(\boldsymbol{D}^\mathrm{H} \cdot \mathrm{diag}\{\boldsymbol{c}_s^*\} \cdot \boldsymbol{C}_p \cdot \boldsymbol{C}_p^\mathrm{H} \cdot \mathrm{diag}\{\boldsymbol{c}_s\} \cdot \boldsymbol{D} - C \cdot \boldsymbol{I}_N\right) \odot \boldsymbol{\Omega}, \\ &= \left(\boldsymbol{D}^\mathrm{H} \cdot \left(\left[\boldsymbol{C}_p \cdot \boldsymbol{C}_p^\mathrm{H}\right] \odot \left[\boldsymbol{c}_s^* \cdot \boldsymbol{c}_s^\mathrm{T}\right]\right) \cdot \boldsymbol{D} - C \cdot \boldsymbol{I}_N\right) \odot \boldsymbol{\Omega}, \end{aligned} \tag{16}$$

where we have used (8) to rewrite $\boldsymbol{\Psi}$ and defined $\boldsymbol{D} = [\boldsymbol{d}(\tau_1^{(\mathrm{G})}), \ldots, \boldsymbol{d}(\tau_N^{(\mathrm{G})})] \in \mathbb{C}^{M \times N}$. Moreover, $\odot$ in (16) denotes the Hadamard-Schur (elementwise) product and $\boldsymbol{\Omega}$ contains the weights ω_{τ_1,τ_2}. Based on (16), the quality of $\boldsymbol{\Psi}$ can be measured via an appropriate norm of $\boldsymbol{E}$. For instance, minimizing the (weighted) average squared error corresponds to minimizing the squared Frobenius norm of $\boldsymbol{E}$, whereas minimizing the maximal error is achieved by minimizing $\|\mathrm{vec}\{\boldsymbol{E}\}\|_\infty$. For simplicity, let us consider a Frobenius norm, leading to the following criterion for $\boldsymbol{C}_p$

$$\min_{\boldsymbol{C}_p} \left\| \left[\boldsymbol{D}^\mathrm{H} \cdot \left(\left[\boldsymbol{C}_p \cdot \boldsymbol{C}_p^\mathrm{H}\right] \odot \left[\boldsymbol{c}_s^* \cdot \boldsymbol{c}_s^\mathrm{T}\right]\right) \cdot \boldsymbol{D} - C \cdot \boldsymbol{I}_N\right] \odot \boldsymbol{\Omega} \right\|_\mathrm{F}. \tag{17}$$

Note that the problem (17) belongs to the class of weighted low-rank approximation problems which have been shown to be NP-hard [18] and do not admit a closed-form solution in general. However, iterative methods with some performance guarantees exist [19].

Note that in the special case $\boldsymbol{\Omega} = \mathbf{1}_{N \times N}$, problem (17) is equivalent to coherence minimization which has been studied, e.g., in [14]. Moreover, due to its structure we can actually solve it in closed form. From [20], we have the following theorem:

Theorem 1. *For a row-orthogonal matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ and a square Hermitian matrix $\boldsymbol{T} \in \mathbb{C}^{N \times N}$, consider the following optimization problem over matrices $\boldsymbol{\Phi} \in \mathbb{C}^{m \times M}$ with $m < M$*

$$\arg\max_{\boldsymbol{\Phi}} \left\| \boldsymbol{A}^\mathrm{H} \cdot \boldsymbol{\Phi}^\mathrm{H} \cdot \boldsymbol{\Phi} \cdot \boldsymbol{A} - \boldsymbol{T} \right\|_\mathrm{F}^2. \tag{18}$$

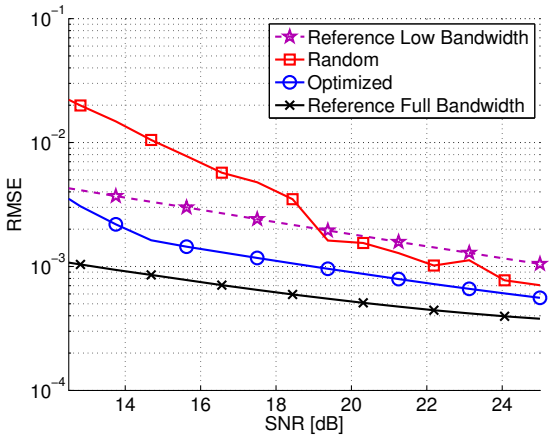

Fig. 2. Relative RMSE of the delay τ_1 (in units of t_p) vs. SNR for $K = 5$ branches and $M = 20$ spectral lines.

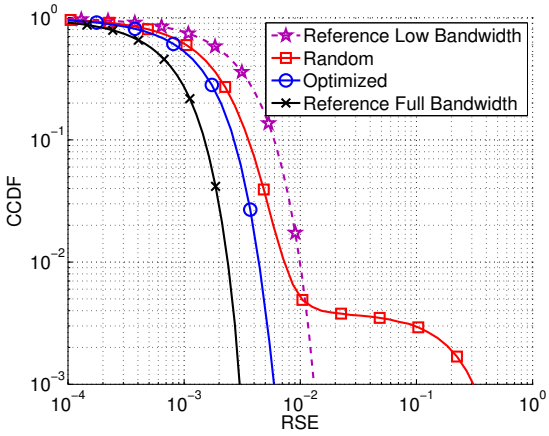

Fig. 3. Histogram of the relative root squared error of τ_1 for $K = 5$, $M = 20$, and an SNR of 13 dB.

Then, a matrix $\boldsymbol{\Phi}$ *maximizes* (18) *if and only if* $\boldsymbol{\Phi}^\mathrm{H} \cdot \boldsymbol{\Phi} = \boldsymbol{S}_m$*, where* $\boldsymbol{S}_m$ *is a rank-*m *approximation of the matrix* $\boldsymbol{S} = \boldsymbol{A} \cdot \boldsymbol{T} \cdot \boldsymbol{A}^\mathrm{H}$*, obtained by setting all but the dominant* m *eigenvalues of* $\boldsymbol{S}$ *to zero.*

Applying Theorem 1 to (17) shows that $\boldsymbol{C}_p$ is optimal with respect to (17) if and only if it can be written as

$$\boldsymbol{C}_p = \sqrt{C} \cdot \mathrm{diag}\{\boldsymbol{c}_s^*\}^{-1} \cdot \boldsymbol{Q}, \tag{19}$$

where $\boldsymbol{Q} \in \mathbb{C}^{M \times K}$ is an arbitrary unitary matrix. However, this result is not very useful since it leaves a large space of possible solutions (any properly scaled row-orthogonal matrix) that achieve the same minimum in the cost function.

Therefore, instead of using (19), we choose to solve (17) using numerical optimization methods. Note that since the measurement matrix design is performed off-line once only, the computational complexity of solving this problem is not a critical issue.

5. NUMERICAL RESULTS

In this section we present some numerical results to show the performance of the proposed CS-based architecture for delay estimation. The observed signal is generated according to (9). The transmit signal is chosen according to $c_s[m] = \mathrm{e}^{\jmath\varphi_m}$ where φ_m is drawn from a uniform distribution in $[0, 2\pi)$. In other words, the transmitter distributes its power evenly across frequency, which corresponds to a sinc-like pulse in the time domain. The noise samples in the noise vector $\boldsymbol{w}$ are drawn from a zero-mean circularly symmetric complex Gaussian distribution with $N_0 = \frac{1}{\mathrm{SNR}}$.

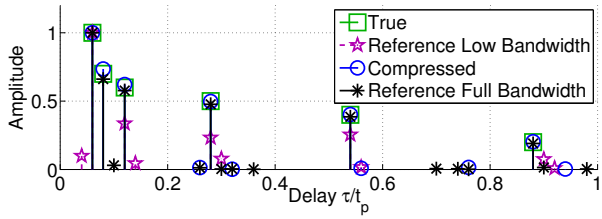

Fig. 4. Estimation of a multipath channel with $L = 6$ paths for $K = 10$, $M = 20$, SNR = 30 dB

The coefficients $\boldsymbol{C}_p$ are modeled as $\mathrm{e}^{\jmath\varphi_{m,k}}$ where $\varphi_{m,k}$ are drawn from a uniform distribution in $[0, 2\pi)$ for the random approach and used as optimization variables in the proposed optimized choice of $\boldsymbol{C}_p$. The latter is found by solving (17) via Matlabs numerical optimization toolbox. The weight matrix is chosen according to $\boldsymbol{\Omega} = (1 - \rho) \cdot \boldsymbol{I}_N + \rho \cdot \boldsymbol{1}_N$, where $\rho \in \mathbb{R}_{[0,1]}$ allows to adjust this trade-off: values close to zero put more weight on the main diagonal (for uniform sensitivity) whereas values closer to one shift the weight to the off-diagonal elements (for low cross-correlation). The results shown here are obtained for $\rho = 0.5$. Finally, the constant C is set to[4] $K \cdot M$.

Figure 2 depicts the result for a system with $K = 5$ branches that uses $M = 20$ spectral lines (i.e., the compression rate is 4). The relative root mean square delay estimation error of the estimator (14) (in units of t_p) is estimated over 100'000 Monte Carlo trials. A single path scenario is chosen where $\alpha_1 = 1$ and τ_1 is drawn randomly from $[0, t_\mathrm{p})$. Since we have a single path, the correlation-based estimator according to (13) is used to estimate the delays. For comparison, the achievable accuracy of a traditional system that uses no compression but a full-rate ADC is also shown ("Reference Full Bandwidth") as well as the performance of a system that uses 1/4 of the bandwidth ("Reference Low Bandwidth"). The results demonstrate that the optimized measurement kernel outperforms the randomly chosen one, in particular for lower SNRs. This behavior is mainly due to the outliers that occur due to the sidelobes in the compressed correlation function. Moreover, the CS-based system achieves an accuracy better than a Nyquist system operating with a reduced bandwidth / sampling rate.

To shed further light on this aspect, Figure 3 depicts the estimated complementary cumulative distribution function (CCDF) of the RMSE for the same simulation at an SNR of 13 dB. We can see that the random choice of the measurement kernels is more prone to outliers. This behavior becomes even more pronounced for lower SNRs.

Figure 4 shows the estimation result for a $L = 6$ path channel for $K = 10$, $M = 20$, and an SNR of 30 dB, using the estimator from (12). The true value of delays and amplitudes are indicated by the markers labeled "True" and compared to the proposed compressed approach (using an optimized matrix for $\rho = 0.5$) and the low/full bandwidth reference. The result shows that the compressed approach finds all the taps while the low bandwidth version misses some peaks (and finds comparably strong spurious ones).

[4] Due to the assumption that $\boldsymbol{c}_s$ and $\boldsymbol{C}_p$ have constant magnitude, we can show that $\mathrm{trace}\{\boldsymbol{\Psi}^\mathrm{H}\boldsymbol{\Psi}\} = M\,K\,N$ regardless of the angles in $\boldsymbol{c}_s$ and $\boldsymbol{C}_p$.

6. CONCLUSIONS

In this paper we investigate a system architecture for delay estimation via Compressed Sensing (CS). We propose to use a bank of K periodic functions $p_k(t)$ that multiply the received signal and are sampled once per period. Thereby, the effective sampling rate is reduced by a factor which depends on the period, the bandwidth, and the number of channels K. We then discuss the design of the functions $p_k(t)$ based on their frequency domain representation. We propose an approach that directly optimizes the shape of the autocorrelation function in terms of the choice of $p_k(t)$ and demonstrate that it outperforms a random choice in terms of the delay estimation accuracy. Note that other criteria could be considered as well, such as Cramér-Rao Bounds. This approach will be further studied in future work.

A. REFERENCES

[1] 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.

[2] 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.

[3] T. Grün, N. Franke, D. Wolf, N. Witt, and A. Eidloth, "A real-time tracking system for football match and training analysis," in *Microelectronic Systems*, A. Heuberger, G. Elst, and R. Hanke, Eds. Springer Berlin Heidelberg, 2011, pp. 199–212.

[4] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: a survey," *IEEE Network*, vol. 18, no. 4, pp. 45–50, July 2004.

[5] S. Björklund, "A survey and comparison of time-delay estimation methods in linear systems," Ph.D. dissertation, Linköping University, 2003.

[6] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," *SIGOPS Oper. Syst. Rev.*, vol. 36, no. SI, pp. 147–163, Dec 2002.

[7] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, Jul. 2009.

[8] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 383–395, 2007.

[9] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, 2010.

[10] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3017–3031, June 2010.

[11] Y. M. Lu and M. N. Do, "A theory for sampling signals from a union of subspaces," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2334–2345, June 2008.

[12] Y. Gu and N. A. Goodman, "Compressive sensing kernel optimization for time delay estimation," in *IEEE Radar Conference*, May 2014, pp. 1209–1213.

[13] S. Bernhardt, R. Boyer, S. Marcos, Y. Eldar, P. Larzabal, and R. Boyer, "Sampling FRI Signals with the SoS Kernel: Bounds and Optimal Kernel," in *European Signal Processing Conference (EUSIPCO'15)*, Nice, France, Aug. 2015. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01159937

[14] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, "On projection matrix optimization for compressive sensing systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2887–2898, June 2013.

[15] P. Misra and P. Enge, *Global positioning system: signals, measurements, and performance*. Ganga-Jamuna Press, 2006.

[16] D. J. Daniels, *Ground penetrating radar*. IET, 2004, vol. 1.

[17] K. Fyhn, S. H. Jensen, and M. F. Duarte, "Compressive time delay estimation using interpolation," in *IEEE Global Conference on Signal and Information Processing*, Dec 2013, pp. 624–624.

[18] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *ICML*, vol. 3, 2003, pp. 720–727.

[19] W. Rey, "On weighted low-rank approximation," *arXiv preprint arXiv:1302.0360*, 2013.

[20] M. Ibrahim, F. Roemer, and G. D. Galdo, "On the design of the measurement matrix for compressed sensing based DOA estimation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Sig. Proc. (ICASSP 2015)*, Brisbane, Australia, Apr. 2015.