This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

Compressive Shift Retrieval

Henrik Ohlsson, *Member, IEEE,* Yonina C. Eldar, *Fellow, IEEE*, Allen Y. Yang, *Senior Member, IEEE*, and S. Shankar Sastry, *Fellow, IEEE*

Abstract

The classical shift retrieval problem considers two signals in vector form that are related by a shift. This problem is of great importance in many applications and is typically solved by maximizing the cross-correlation between the two signals. Inspired by compressive sensing, in this paper, we seek to estimate the shift directly from compressed signals. We show that under certain conditions, the shift can be recovered using fewer samples and less computation compared to the classical setup. We also illustrate the concept of superresolution for shift retrieval. Of particular interest is shift estimation from Fourier coefficients. We show that under rather mild conditions only one Fourier coefficient suffices to recover the true shift.

I. INTRODUCTION

Shift retrieval between two given signals is a fundamental problem in many signal processing applications. For example, to map the ocean floor, an active sonar can be used. The sonar transmits reference sound pulses through the water, and the time it takes to receive the echoes of the pulses indicates the depth of the ocean floor. In target tracking using two acoustic sensors, the time shift when a sound wave of a vehicle reaches the microphones indicates the direction to the vehicle. In global positioning system (GPS) receivers, the correct alignment between CDMA (code division multiple access) codes is sought [1] in order to calculate the synchronization delay necessary for determining its position. In the case of a time shift, the shift retrieval problem is often referred to as *time delay*

Ohlsson, Yang and Sastry are with the Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA, e-mail: ohlsson@eecs.berkeley.edu.

Ohlsson is also with Dept. of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden.

Eldar is with the Dept. of Electrical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel.

Ohlsson is partially supported by the Swedish Research Council in the Linnaeus center CADICS, the European Research Council under the advanced grant LEARN, contract 267381, by a postdoctoral grant from the Sweden-America Foundation, donated by ASEA's Fellowship Fund, and by a postdoctoral grant from the Swedish Research Council. Ohlsson and Sastry are supported by FORCES (Foundations Of Resilient CybEr-physical Systems) which receives support from NSF (award number CNS-1239166). Eldar is supported in part by the Israel Science Foundation under Grant no. 170/10, in part by the SRC, in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), and in part by the Ollendorf Foundation. Yang is supported in part by ARO 63092-MA-II, DARPA FA8650-11-1-7153 and ONR N00014-13-1-0341.

A preliminary version of this paper was presented at the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 26-31, 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

estimation (TDE) [2]. In computer vision, the spatial shift relating two images is often sought and referred to as image registration or alignment [3], [4], [5].

Traditionally, the shift retrieval problem is solved by maximizing the cross-correlation between the two signals [6]. In this paper, we revisit this classical problem, and show how the basic premise of *compressive sensing* (CS) [7], [8], [9], [10] can be used in the context of shift retrieval. This allows to recover the shift from compressed data leading to computational and storage savings.

Compressive sensing is a sampling scheme that makes it possible to sample at the *information rate* instead of the classical Nyquist rate predicted by the bandwidth of the signal [11]. The majority of the results in compressive sensing discuss conditions and methods for guaranteed reconstruction from an under-sampled version of the signal. Therefore, the information rate is typically referred to as the one that guarantees recovery of the sparse signal.

However, for many applications such as the aforementioned examples in shift retrieval, obtaining the signal may not be needed. The goal is to recover some properties or statistics of the unknown signal. Taking the active sonar for example, one may wonder if it is really necessary to sample at a rate which is twice that of the bandwidth of the transmitted signal so that the received signal can be exactly reconstructed? Clearly the answer is no. Since the signal itself is not of interest to the application, we might consider an alternative sampling scheme to directly estimate the shift without first reconstructing the signal. These ideas have in fact been recently explored in the context of radar and ultrasound [12], [13], [14] with continuous time signals and multiple shifts. Here we consider a related problem and ask: *What is the minimal information rate for shift retrieval when two related discrete-time signals are under-sampled?*

It turns out that under rather mild conditions, we only need fractions of the signals. In fact, we will show that only one Fourier coefficient from each of the signals suffices to recover the true shift. We refer to this method as *compressive shift retrieval* (CSR). It should be made clear that CSR does not assume that any of the involved signals are sparse.

As the main contribution of the paper, we will show that when the sensing matrix is taken to be a partial Fourier matrix, under suitable conditions, the true shift can be recovered from both noise-free and noisy measurements using CSR. In fact, our results show that in some cases sampling as few as one Fourier coefficient is enough to perfectly recover the true shift. Furthermore, CSR also reduces the computational load. This is of particular interest since recent developments in sampling [15], [16], [13] have shown that Fourier coefficients can be efficiently obtained from space (or time) measurements by the use of an appropriate filter and by subsampling the output. From a user perspective, for example, reduced computational complexity could elongate battery life of GPS enabled devices. Finally, we introduce the concept of shift retrieval in superresolution, whereby shifts can be recovered at finer granularities than those of the received signals, based on the proposed CSR framework.

A. Prior Work

Compressive signal alignment problems have been addressed in only a few publications and, to the authors' best knowledge, not in the same setup studied in this paper. In [17], the authors considered alignment of images under random projection. The work was based on the Johnson-Lindenstrauss property of random projection and proposed an objective function that can be solved efficiently using *difference-of-two-convex programming* algorithms. In this paper, we instead focus on proving theoretical guarantees of exact shift recovery when the signal is subsampled by a partial Fourier basis. The theory developed in [17] does not apply to this setup.

The *smashed filter* [18] is another related technique. It is a general framework for maximum likelihood hypothesis testing and can be seen as a *matched filter* of reduced dimension. It can therefore be applied to the shift retrieval problem. The underlying idea of both the smashed filter and CSR are the same in that both approaches try to avoid reconstructing the signal and extract the sought descriptor, namely, the shift, from compressive measurements. However, the analysis and assumptions are very different. For CSR, we develop requirements for guaranteed recovery of the true shift using a given measurement matrix. For the smashed filter, the analysis focuses on random orthogonal projections and provides probabilities for correct recovery as a function of the number of projections.

Motivated by the GPS locking problem, [19], [20] studied computationally efficient algorithms for recovering shifts of a random code sequence, which may be corrupted by Gaussian noise. Their algorithms exploit the sparse nature of the signal matching problem, where the optimal signal alignment causes the cross-correlation between the source signal and the measurement to spike. The main limitation of the work is that their analysis assumes the source signals are sampled randomly with binary values in $\{-1, 1\}$. In this paper, our analysis of CSR and the conditions for guaranteed recovery is not restricted to signals with binary values. Motivated by the compressive sensing framework, the new algorithms are also sufficiently different from the ones in [19], [20].

B. Notation

We use normal fonts to represent scalars and bold fonts for vectors and matrices. The notation $|\cdot|$ represents the absolute value for scalars and it returns the cardinality of a set if the argument is a set. For both vectors and matrices, $\|\cdot\|_0$ is the ℓ_0-norm that counts the number of nonzero elements of its argument. Similarly, $\|\cdot\|_p$ represents the ℓ_p-norm: for a vector $\boldsymbol{x}$, $\|\boldsymbol{x}\|_p \triangleq (\sum_i |x_i|^p)^{1/p}$, where x_i is the ith element of $\boldsymbol{x}$, and for a matrix $\boldsymbol{X}$, $\|\boldsymbol{X}\|_p \triangleq (\sum_{i,j} |X_{i,j}|^p)^{1/p}$, where $X_{i,j}$ is the (i, j)-th element of $\boldsymbol{X}$. Furthermore, $\boldsymbol{X}^*$ denotes the complex conjugate transpose of $\boldsymbol{X}$. Let $\boldsymbol{I}_{n\times n}$ denote an $n \times n$ identity matrix, $\mathbf{0}_{m\times n}$ an $m \times n$ matrix of zeros, and $\mathbb{Z}$ be the set of integers. $\Re\{\cdot\}$ returns the real part of its argument and $\lfloor\cdot\rfloor$ denotes the floor function.

We say that two n-dimensional vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ are related by an l cyclic-shift if $\boldsymbol{y} = \boldsymbol{D}^l\boldsymbol{x}$, where $\boldsymbol{D}^l$ is defined as

$$\boldsymbol{D}^l = \begin{bmatrix} \mathbf{0}_{l\times(n-l)} & \boldsymbol{I}_{l\times l} \\ \boldsymbol{I}_{(n-l)\times(n-l)} & \mathbf{0}_{(n-l)\times l} \end{bmatrix}. \tag{1}$$

Throughout the paper, we assume that the (cyclic) shift is unique *up to a multiple of* n.

C. Organization

In Sections II and III, we study the CSR problem under the assumption that the measurements are noise free. We extend the results to noisy measurements in Section IV. As we are particularly interested in Fourier measurements, we tailor the results to this choice of sensing matrix. We illustrate the concept of superresolution for shift retrieval in Section V, and conclude in Section VI.

II. Noise-Free Compressive Shift Retrieval

Consider two vectors $\boldsymbol{x} \in \mathbb{C}^n$ and $\boldsymbol{y} \in \mathbb{C}^n$ that are related by a cyclic shift, *i.e.,* $\boldsymbol{y} = \boldsymbol{D}^l \boldsymbol{x}$. The shift retrieval problem is a multi-hypothesis testing problem: Define the sth hypothesis $\mathcal{H}_s$, $s = 0, \dots, n-1$, as

$$\mathcal{H}_s : \ \boldsymbol{x} \text{ is related to } \boldsymbol{y} \text{ via a } s\text{-cyclic-shift},$$

and accept $\mathcal{H}_s$ if $\boldsymbol{y} = \boldsymbol{D}^s \boldsymbol{x}$ and otherwise reject. Since the true shift is assumed unique, only one hypothesis will be accepted.

To determine s we minimize the error $\|\boldsymbol{y} - \boldsymbol{D}^s \boldsymbol{x}\|_2$. Now,

$$\begin{aligned} \|\boldsymbol{y} - \boldsymbol{D}^s \boldsymbol{x}\|_2^2 =& \|\boldsymbol{y}\|_2^2 + \|\boldsymbol{D}^s \boldsymbol{x}\|_2^2 - \boldsymbol{y}^* \boldsymbol{D}^s \boldsymbol{x} - \boldsymbol{x}^* \boldsymbol{D}^s \boldsymbol{y} \\ =& \|\boldsymbol{y}\|_2^2 + \|\boldsymbol{x}\|_2^2 - 2\Re\{\langle \boldsymbol{y}, \boldsymbol{D}^s \boldsymbol{x} \rangle\} \end{aligned} \tag{2}$$

where we used the fact that $\|\boldsymbol{D}^s \boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2$. Therefore, minimizing $\|\boldsymbol{y} - \boldsymbol{D}^s \boldsymbol{x}\|_2^2$ is equivalent to maximizing the real part of the cross-correlation with respect to s:

$$\max_{s \in \{0,\dots,n-1\}} \Re\{\langle \boldsymbol{y}, \boldsymbol{D}^s \boldsymbol{x} \rangle\}. \tag{3}$$

The goal of CSR is to recover the shift l relating $\boldsymbol{x}$ and $\boldsymbol{y}$ from compressed measurements $\boldsymbol{z}$ and $\boldsymbol{v}$. The compressed measurement signals are assumed related to the ground-truth signals $\boldsymbol{x} \in \mathbb{C}^n$ and its shifted version $\boldsymbol{y} = \boldsymbol{D}^l \boldsymbol{x} \in \mathbb{C}^n$ via a sensing matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$, with $m \leq n$:

$$\boldsymbol{z} = \boldsymbol{A}\boldsymbol{y} \in \mathbb{C}^m \text{ and } \boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} \in \mathbb{C}^m. \tag{4}$$

Since only the compressed measurements $\boldsymbol{z}$ and $\boldsymbol{v}$ are assumed available, we cannot evaluate $\boldsymbol{y} = \boldsymbol{D}^s \boldsymbol{x}$ or maximize $\Re\{\langle \boldsymbol{y}, \boldsymbol{D}^s \boldsymbol{x} \rangle\}$ for each hypothesis $s = 0, \dots, n-1$. However, if $\boldsymbol{A}^* \boldsymbol{A}$ and $\boldsymbol{D}^s$ commute for all $s = 0, \dots, n-1$, then

$$\boldsymbol{y} = \boldsymbol{D}^s \boldsymbol{x} \quad \Rightarrow \quad \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{y} = \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{D}^s \boldsymbol{x} = \boldsymbol{D}^s \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{x}. \tag{5}$$

Therefore,

$$\boldsymbol{A}^* \boldsymbol{z} = \boldsymbol{D}^s \boldsymbol{A}^* \boldsymbol{v}. \tag{6}$$

Hence, in this case, we can consider the test:

$$\text{Accept } \mathcal{H}_s \text{ if } \boldsymbol{A}^* \boldsymbol{z} = \boldsymbol{D}^s \boldsymbol{A}^* \boldsymbol{v} \text{ and otherwise reject.} \tag{7}$$

It is clear that if s is such that $\boldsymbol{y} = \boldsymbol{D}^s \boldsymbol{x}$, then $\boldsymbol{A}^* \boldsymbol{z} = \boldsymbol{D}^s \boldsymbol{A}^* \boldsymbol{v}$ will also hold. However, the other way around might not be true. Therefore, we might erroneously accept a wrong hypotheses using (7). Theorem 1 below lists conditions under which testing (7) is guaranteed to accept the correct hypothesis.

Before stating the theorem, note that testing the condition $\boldsymbol{A}^*\boldsymbol{z} = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}$ is equivalent to minimizing $\|\boldsymbol{A}^*\boldsymbol{z} - \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\|_2^2$ with respect to s. Now,

$$\begin{aligned} &\|\boldsymbol{A}^*\boldsymbol{z} - \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\|_2^2 \\ &\quad = \|\boldsymbol{A}^*\boldsymbol{z}\|_2^2 + \|\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\|_2^2 - \boldsymbol{z}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v} - \boldsymbol{v}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{z} \\ &\quad = \|\boldsymbol{A}^*\boldsymbol{z}\|_2^2 + \|\boldsymbol{A}^*\boldsymbol{v}\|_2^2 - 2\Re\{\langle \boldsymbol{z}, \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\rangle\}. \end{aligned} \tag{8}$$

Since the only term depending on s is $2\Re\{\langle \boldsymbol{z}, \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\rangle\}$, seeking s satisfying $\boldsymbol{A}^*\boldsymbol{z} = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}$ is equivalent to maximizing $\Re\{\langle \boldsymbol{z}, \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\rangle\}$. Note that if $\boldsymbol{A}^*\boldsymbol{A} = \mathbf{I}_{n\times n}$, then the implication in (5) holds in both directions and maximizing $\Re\{\langle \boldsymbol{z}, \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{v}\rangle\}$ reduces to the classical test using uncompressed signals given in (3).

Theorem 1 (**Shift Recovery from Low-Rate Data**). *Let $\boldsymbol{X}$ be an $n \times n$ matrix with ith column equal to $\boldsymbol{D}^i\boldsymbol{x}$, $i = 1, \ldots, n$, and define $\bar{\boldsymbol{D}}^s = \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*$. If the sensing matrix $\boldsymbol{A}$ satisfies the following conditions:*

1) $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}$,
2) $\exists \alpha \in \mathbb{R}, \alpha\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{I}$ *and*
3) *all columns of $\boldsymbol{A}\boldsymbol{X}$ are different,*

then

$$\max_{s\in\{0,\ldots,n-1\}} \Re\{\langle \boldsymbol{z}, \bar{\boldsymbol{D}}^s\boldsymbol{v}\rangle\} \tag{9}$$

or equivalently the test (7) *recovers the true shift.*

Proof: See Appendix A. ■

The conditions of Theorem 1 may seem restrictive. However, as we will show in Lemma 3, if $\boldsymbol{A}$ is chosen as a partial Fourier matrix, then the first two conditions are trivially satisfied. The last condition is the only one that needs to be checked and will lead to a requirement on the sampled Fourier coefficients.

The conditions of Theorem 1 can be checked prior to estimating the shift. However, knowing the estimate of the shift, it is easy to see from the proof (see the proof of Lemma 9) that it is enough to check if the column of $\boldsymbol{A}\boldsymbol{X}$ associated with the estimate of the shift is different than all the other columns of $\boldsymbol{A}\boldsymbol{X}$. Hence, we do not need to check if all columns of $\boldsymbol{A}\boldsymbol{X}$ are different. This conclusion is formulated in the following corollary, which is less conservative than Theorem 1.

Corollary 2 (**Test for True Shift**). *Let $\boldsymbol{X}$ be an $n \times n$ matrix with the ith column equal to $\boldsymbol{D}^i\boldsymbol{x}$, $i = 1, \ldots, n$, and define $\bar{\boldsymbol{D}}^s = \boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*$. If the sensing matrix $\boldsymbol{A}$ satisfies the following conditions:*

1) $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}$, *and*
2) $\exists \alpha \in \mathbb{R}, \alpha\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{I}$,

then

$$s^* = \underset{s\in\{0,\ldots,n-1\}}{\arg\max}\ \Re\{\langle \boldsymbol{z}, \bar{\boldsymbol{D}}^s\boldsymbol{v}\rangle\} \tag{10}$$

is the true shift if the s^*th column of $\boldsymbol{A}\boldsymbol{X}$ is different than all the other columns of $\boldsymbol{A}\boldsymbol{X}$.

Proof: See Appendix A. ■

III. Compressive Shift Retrieval using Fourier Coefficients

Of particular interest is the case in which $\boldsymbol{A}$ is made up of a partial Fourier basis. That is, $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ takes the form

$$\boldsymbol{A} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & e^{-\frac{2j\pi k_1}{n}} & e^{-\frac{4j\pi k_1}{n}} & \cdots & e^{-\frac{2(n-1)j\pi k_1}{n}} \\ 1 & e^{-\frac{2j\pi k_2}{n}} & & \ddots & e^{-\frac{2(n-1)j\pi k_2}{n}} \\ \vdots & \vdots & & & \\ 1 & e^{-\frac{2j\pi k_m}{n}} & e^{-\frac{4j\pi k_m}{n}} & \cdots & e^{-\frac{2(n-1)j\pi k_m}{n}} \end{bmatrix}$$

where $k_1, \ldots, k_m \in \{0, 1, 2, \ldots, n-1\}, m \leq n$. For this specific choice,

$$\boldsymbol{AX} = \frac{1}{\sqrt{n}} \begin{bmatrix} X_{k_1} & X_{k_1} e^{\frac{2j\pi k_1}{n}} & \cdots & X_{k_1} e^{\frac{2(n-1)j\pi k_1}{n}} \\ X_{k_2} & \ddots & & X_{k_2} e^{\frac{2(n-1)j\pi k_2}{n}} \\ \vdots & & & \\ X_{k_m} & X_{k_m} e^{\frac{2j\pi k_m}{n}} & \cdots & X_{k_m} e^{\frac{2(n-1)j\pi k_m}{n}} \end{bmatrix}$$

where X_r denotes the rth Fourier coefficient of the Fourier transform of $\boldsymbol{x}$.

For a sensing matrix made up by a partial Fourier basis, we have the following useful result:

Lemma 3. *Let $\boldsymbol{A}$ be a partial Fourier matrix. Then $\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s$ for all $s = 0, \ldots, n-1$.*

Proof: See Appendix A. ■

Applying this result to Theorem 1 gives the following corollary:

Corollary 4 (Shift Recovery from Low Rate Fourier Data). *Suppose $\boldsymbol{A}$ is chosen as a partial Fourier matrix with $k_1, \ldots, k_m \in \{0, 1, 2, \ldots n-1\}, m \leq n$. Let z_i and v_i be the ith elements of $\boldsymbol{z} = \boldsymbol{Ay}$ and $\boldsymbol{v} = \boldsymbol{Ax}$. Then* (9) *is simplified as*

$$\max_{s\in\{0,\ldots,n-1\}} \Re\left\{\sum_{i=1}^{m} z_i^* v_i e^{\frac{-2\pi j k_i s}{n}}\right\}, \tag{11}$$

and it recovers the true shift if there exists $p \in \{1, \ldots, m\}$ such that $X_{k_p} \neq 0$ and $\{1, \ldots, n-1\}\frac{k_p}{n}$ contains no integers. In particular, measuring only the first Fourier coefficients ($k_1 = 1$) of $\boldsymbol{x}$ and $\boldsymbol{y}$ would, as long as the coefficients are nonzero, suffice to recover the true shift.

Proof: See Appendix A. ■

Remarkably, in the extreme case when $m = 1$, the corollary states that all we need is two scalar measurements, z and v, to perfectly recover the true shift. The scalar measurements can be any nonzero Fourier coefficient of $\boldsymbol{x}$ and $\boldsymbol{y}$ as long as $\{1, \ldots, n-1\}\frac{k_1}{n}$ contains no integers. As noted in the corollary, the first Fourier coefficients ($k_1 = 1$) of $\boldsymbol{x}$ and $\boldsymbol{y}$ would suffice. Also note that only $2mn$ multiplications are required to evaluate the test. This should be compared with $\mathcal{O}(n \log n)$ multiplications to evaluate the cross-correlation for the full uncompressed signals $\boldsymbol{x}$ and $\boldsymbol{y}$ [21].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

Example 1 (**Noise Free Compressive Shift Retrieval).** *To validate the results, we carry out the following example. In each trial we let the sample dimension* m *and the shift* l *be random integers between* 1 *and* 9 *and generate* $\boldsymbol{x}$ *by sampling from an* n*-dimensional uniform distribution. We let* $n = 10$ *and generate a partial Fourier matrix by picking* $k_1, \dots, k_m$ *from* $\{0, 1, \dots, 9\}$ *at random without replacement. The coefficients* $k_1, \dots, k_m$ *are regenerated if the assumptions of Corollary 4 are not met. The true shift is successfully recovered in each trial by the simplified test* (11)*, namely, with 100% success rate. This is quite remarkable since when* $m = 1$*, we recover the true shift using only two scalar measurements* $\boldsymbol{z}$ *and* $\boldsymbol{v}$ *and one fifth of the multiplications that* (3) *would require.*

IV. Noisy Compressive Shift Retrieval

We now consider the noisy version of CSR, where the measurements $\boldsymbol{z}$ and $\boldsymbol{v}$ are perturbed by noise:

$$\tilde{\boldsymbol{z}} = \boldsymbol{z} + \boldsymbol{e}_z, \quad \tilde{\boldsymbol{v}} = \boldsymbol{v} + \boldsymbol{e}_v. \tag{12}$$

Similar to the noise-free case, we can guarantee recovery of the true shift. In particular, if the columns of the noisy version of $\boldsymbol{AX}$ are far enough apart with respect to the noise, then it can be shown that the columns of the noise free version of $\boldsymbol{AX}$ are distinct and the true shift is recovered. We note that a more natural scenario in some applications might be to assume that $\boldsymbol{v}$ is a known reference signal and noise free. Our derivations below also handle this case by setting $\boldsymbol{e}_v = 0$.

Our main result is given in the following theorem:

Theorem 5 (**Noisy Shift Recovery from Low-Rate Data).** *Let* $\tilde{\boldsymbol{x}}$ *be such that* $\tilde{\boldsymbol{v}} = \boldsymbol{A}\tilde{\boldsymbol{x}}$ *and let the* i*th column of* $\tilde{\boldsymbol{X}}$ *be* $\tilde{\boldsymbol{x}}$ *shifted by* i*, i.e.,* $\boldsymbol{D}^i\tilde{\boldsymbol{x}}$*. Assume that* $\boldsymbol{A}$ *is a partial Fourier matrix and that the noisy measurements are used in* (11) *to estimate the shift. If the* ℓ_2*-norm difference between any two columns of* $\boldsymbol{A}\tilde{\boldsymbol{X}}$ *is greater than*

$$\begin{aligned}\Delta_{\boldsymbol{zv}} \triangleq & \|\boldsymbol{e}_z\|_2 + \|\boldsymbol{e}_v\|_2 \\ & + \sqrt{\|\tilde{\boldsymbol{v}}\|_2^2 + \|\tilde{\boldsymbol{z}}\|_2^2 - 2 \max_{s \in \{0,\dots,n-1\}} \Re\{\langle \tilde{\boldsymbol{z}}, \bar{\boldsymbol{D}}^s \tilde{\boldsymbol{v}} \rangle\}},\end{aligned} \tag{13}$$

then the estimate of the shift is not affected by the noise.

Proof: See Appendix B. ■

The result of Theorem 5 is that, by requiring the difference between columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ to be greater than $\Delta_{\boldsymbol{zv}}$, we assure that the noise does not affect the outcome of the test (11).

Note that $\|\boldsymbol{e}_z\|_2$ and $\|\boldsymbol{e}_v\|_2$ might not be available in practice but could be replaced by an upper bound on the ℓ_2-norm of the noise if the noise is known to be bounded. Also note that the theorem only states that the noise does not affect the estimate of the shift. It does not state that the shift will be the true shift.

Example 2 (**Recovery of a Shift from Noisy Data).** *We illustrate the results by running a Monte Carlo simulation consisting of 10,000 trials for each sample dimension* $m = 1, \dots, 10$*, and for two different SNR levels. In Figure 1, 10 histograms are shown (corresponding to* $m = 1, \dots, 10$*) for SNR* $= \|\boldsymbol{z}\|_2^2 / \|\boldsymbol{e}_z\|_2^2$ *being 2 (low SNR) and in Figure 2, SNR* $= 10$ *(high SNR). The errors* $\boldsymbol{e}_z$ *and* $\boldsymbol{e}_v$ *are both generated by sampling from* $\mathcal{N}(0, \sigma^2) + j\mathcal{N}(0, \sigma^2)$*.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

We further use $n = 10$*,* $l = 5$ *and sample* $\boldsymbol{x}$ *from a uniform (0,1)-distribution. The conclusion from the simulation is that the smaller the* m*, the more the estimate of the shift is sensitive to noise. Notice that when* $m = 10$*, the test* (11) *reduces to the classical test of maximizing the cross-correlation.*

We can use Theorem 5 to check if the noise affectes the estimate of the shift or not in each of the trials. For $m = 2$ *and high SNR, 40% of the trials satisfy the conditions of Theorem 5 and the noise therefore does not affect the shift estimates in those cases. Of the trials that satisfy the conditions, all successfully predict the true shift and none of them are false shifts. Note however that Theorem 5 only states that if the conditions are satisfied, then the estimated shift is the same as if we would have used the noise free compressed measurements in the test* (11)*. It does not state that the estimate will be the true shift.*

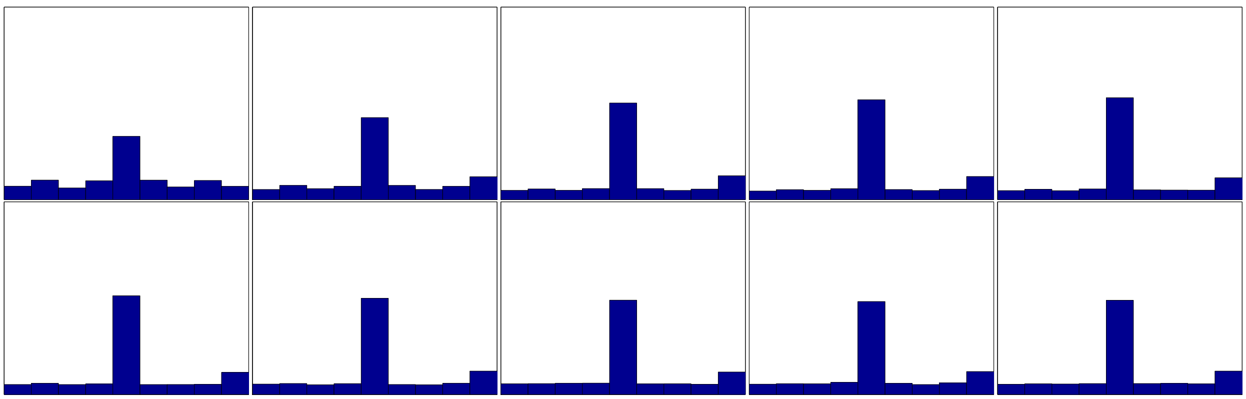

Fig. 1. Histogram plots for the estimated shift and low SNR. From left to right, top to bottom, $m = 1, \ldots, 10$. The true shift is set to 5 in all trials.

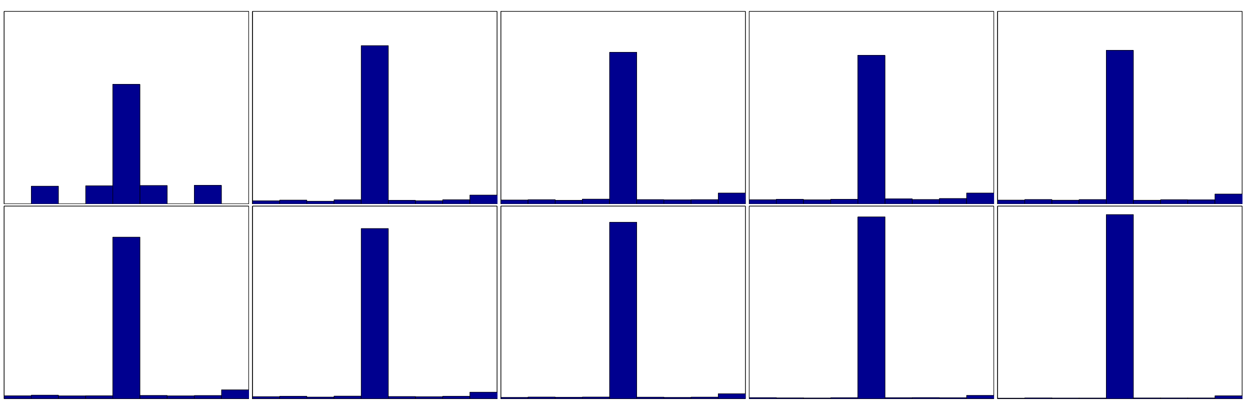

Fig. 2. Histogram plots for the estimated shift and high SNR. From left to right, top to bottom, $m = 1, \ldots, 10$. The true shift is set to 5 in all trials.

Example 3 (Varying SNR). *To further illustrate the ability of CSR to handle different SNR levels, in this example, we vary SNR and study the recovery rate predicted by the theory and obtained in simulations. Let* $n = 100$*,* $m = 10$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

and generate $\boldsymbol{x}$ *by sampling from a uniform distribution between* 0 *and* 1 *(we have also sampled from a standard Gaussian distribution but the results were essentially the same and therefore not shown here). We generate the true shift by sampling an integer randomly between 0 and 99, add Gaussian complex noise to* $\boldsymbol{z}$ *and* $\boldsymbol{v}$ *and repeat the experiment* $1,000$ *times for each SNR level. The results are shown in Figure 3. The solid curve shows the rate of recovery seen in simulations. The dashed curve shows rate at which the condition* (13) *of Theorem 5 holds. The dashed-dotted line shows the rate at which the difference between columns of* $\boldsymbol{A}\tilde{\boldsymbol{X}}$ *is greater than* $2\|\boldsymbol{e}_v\|$*. This test, as shown in Corollary 6, is relevant for guaranteeing that the estimated shift is the true one.*

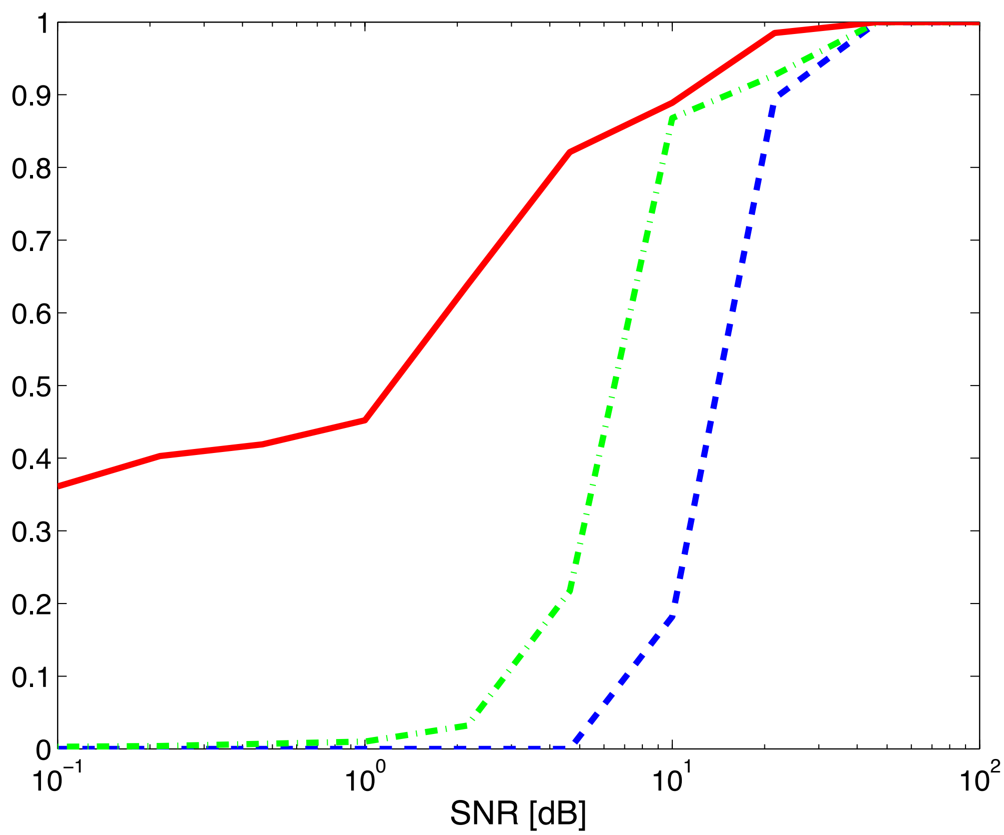

Fig. 3. The recovery rate (solid line), the rate at which the condition of Theorem 5 holds (dashed curve), the rate at which the difference between columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ are greater than $2\|\boldsymbol{e}_v\|$ (dashed-dotted line) for SNRs between 0.1 and 100dB.

Theorem 5 provides conditions under which the noise does not affect the estimate of the shift. A better result would be to guarantee the recovery of the true shift. We saw in the first part of this paper that if the columns of $\boldsymbol{AX}$ are distinct, then the true shift is recovered. To guarantee the recovery of the true shift from noisy measurements we need:

1) that (11) gives the same shift estimate for the noisy measurements as for the noise free measurement (Theorem 5), and in addition,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

2) that the columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ are far enough apart so that if the noise would be removed, the columns would still be distinct.

The details are given by the following corollary.

Corollary 6 (**Recovery of the True Shift from Noisy Low-Rate Data).** *If the ℓ_2-norm difference between any two columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ is greater than $2\|\boldsymbol{e}_v\|_2$ and the conditions of Theorem 5 are fulfilled, then* (11) *recovers the true shift.*

Proof: See Appendix B. ■

Note that the result is not independent of the noise $\boldsymbol{e}_z$ since the conditions of Theorem 5 depend on it.

If the estimate of the shift has been computed, a less conservative test can be used to check if the computed estimate has been affected by noise and if it is the true one. We summarize our conclusion in the following corollary.

Corollary 7 (**Test for True Shift in the Presence of Noise).** *Assume that* (11) *gives s^* as an estimate of the shift. If the ℓ_2 difference between any column and the s^*-column of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ is greater than $2\|\boldsymbol{e}_v\|_2$ and $\Delta_{\boldsymbol{zv}}$, then s^* is the true shift.*

Proof: See Appendix B. ■

Note that $\|\boldsymbol{e}_z\|_2$ and $\|\boldsymbol{e}_v\|_2$ might not be available in practice but can be replaced by an upper bound if the noise is known to be bounded. This holds for both Corollaries 6 and 7.

V. Shift Retrieval in Superresolution

The resolution of any classical electromagnetic sensing system is limited by the wavelength of the measured electromagnetic wave, and details finer than a wavelength cannot be observed [22]. This limitation can be improved to some extent by imposing some structural information about the image to enhance its resolution, also known as *superresolution* [23], [24]. In this section, we study an analogue of superresolution for shift retrieval. The goal is to recover shifts in higher resolution/precision than those by maximizing the cross-correlation in the source signal resolution.

First, we observe that under the conditions of Theorem 1, CSR recovers the shift with a resolution defined by n, the dimension of $\boldsymbol{x}$ and $\boldsymbol{y}$. Hence, the resolution is independent of the number of measurements acquired. It implies that in the noise free case, the correct shift can be recovered up to any accuracy without increasing the number of measurements, as long as the conditions of Theorem 1 are satisfied.

Also note that the above observation would not be practical if we first had to sample $\boldsymbol{x}$ and $\boldsymbol{y}$ to compute their partial Fourier transforms $\boldsymbol{z}$ and $\boldsymbol{v}$. The concept of superresolution is more meaningful if the signals $\boldsymbol{z}$ and $\boldsymbol{v}$ are measured directly. The details are given in the following corollary:

Corollary 8 (**Shift Retrieval in Superresolution).** *Let $x(t)$, $t \in [0, T)$ be a continuous time signal, $y(t) = x(t-l)$, $t \in [l, T)$ and $y(t) = x(t-l+T)$, $t \in [0, l)$. Assume that l is a multiple of T/n for some n. Let the ith* *element of* $\boldsymbol{x} \in \mathbb{R}^n$ *be* $x((i-1)T/n)$, $i = 1, \dots, n$, *and define* $\boldsymbol{y}$ *similarly by sampling* $y(t)$. *Let* $\boldsymbol{X}$ *be an* $n \times n$ *matrix with* i*th column equal to* $\boldsymbol{D}^i\boldsymbol{x}$, $i = 1, \dots, n$, *and let* $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ *be a partial Fourier matrix. Suppose that we are given* $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}$ *and* $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{y}$. *Then the true shift* l *can be recovered from* $\boldsymbol{v}$ *and* $\boldsymbol{z}$ *as* $l = Tl'/n$ *with* l' *solving* (11), *as long as* $\boldsymbol{A}\boldsymbol{X}$ *has distinct columns.*

Proof: From the construction of $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{v}$, $\boldsymbol{z}$, we have that $\boldsymbol{y} = \boldsymbol{D}^{l'}\boldsymbol{x}$ and $\boldsymbol{z} = \bar{\boldsymbol{D}}^{l'}\boldsymbol{v}$, for some l'. Using Theorem 1 it follows that $Tl'/n = l$ as long as $\boldsymbol{A}\boldsymbol{X}$ has distinct columns. ■

Note that l has to be a multiple of T/n for Corollary 8 to hold. If this is not the case, and the grid does not include the true shift, we may ask under what conditions (11) recovers the shift estimate $l' \in \{0, \dots, n-1\}$ that minimizes the error $|l - l'T/n|$. To answer this question, consider the following setup: Let $x(t)$, $t \in [0, T)$ be a continuous time signal, $y(t) = x(t-l)$, $t \in [l, T)$ and $y(t) = x(t-l+T)$, $t \in [0, l)$. Introduce $\tilde{\boldsymbol{x}} \in \mathbb{R}^n$ by stacking the (possibly noisy) samples $x(iT/n), i = 0, \dots, n-1$. Define $\tilde{\boldsymbol{y}} \in \mathbb{R}^n$ accordingly. Let $\tau \in (-T/(2n), T/(2n)]$ be the smallest offset such that $l - \tau$ is a multiple of T/n and introduce $\boldsymbol{x} \in \mathbb{R}^n$ by stacking the samples $x(iT/n + \tau), i = 0, \dots, n-1$; define $\boldsymbol{y} \in \mathbb{R}^n$ accordingly. Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ be a partial Fourier matrix and generate $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}$, $\tilde{\boldsymbol{v}} = \boldsymbol{A}\tilde{\boldsymbol{x}}$, $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{y}$ and $\tilde{\boldsymbol{z}} = \boldsymbol{A}\tilde{\boldsymbol{y}}$. By identifying $\boldsymbol{e}_z = \tilde{\boldsymbol{z}} - \boldsymbol{z}$ and $\boldsymbol{e}_v = \tilde{\boldsymbol{v}} - \boldsymbol{v}$, we can view the misalignment in the grid as noise and use the theory developed for noisy compressive shift retrieval to give guarantees for recovery. We demonstrate this through an example.

Example 4 (**Superresolution).** *The aim of this example is to illustrate superresolution for both the noise free and noisy cases. Let the continuous time signal* $x(t)$ *be a realization of a fractional Brownian motion (a continuous-time Gaussian process) on the time interval* $[0, 1600)$, $y(t) = x(t-l)$, $t \in [l,\ 1600)$, $y(t) = x(t-l+1600)$, $t \in [0,\ l)$ *and sample the delay* l *randomly from* $\{1, 3, 5 \dots, 1599\}$. *Assume that we are given 10 noise-free Fourier transform measurements of* x *and* y *at frequencies randomly chosen from* $\{0, 1/1664, \dots, 12/1664\}$. *We stack these 10 measurements in* $\tilde{\boldsymbol{z}}$ *and* $\tilde{\boldsymbol{v}}$, *respectively.*

To recover the shift we grid the time interval $[0, 1600)$. *We start by a rather coarse grid:*

$$\{0, 2^k, 2 \times 2^k, \dots, \lfloor 1600/2^k \rfloor 2^k\}, \tag{14}$$

with $k = 7$. *The grid is successively refined for* $k = 6, 5, 4, 3, 2, 1, 0$, *each time using the same 10 Fourier measurements. Note that by construction, the true shift does not match any grid points for* $k = 7, 6, 5, 4, 3, 2, 1$. *One example of the estimated shifts in the different resolutions is illustrated in Figure 4.*

For $k = 7$, *the two grid points closest to the true shift are shown with blue circles and the true shift by the red vertical line. The filled circle shows the estimate of the shift given by* (11) *using the 10 Fourier samples. As seen, the grid point closest to the true time delay is correctly found. This was also verified by checking that the column difference of* $\boldsymbol{A}\tilde{\boldsymbol{X}}$ *exceeded* Δ_{zv} *and* $2\|\boldsymbol{e}_v\|$. *Since the conditions of Corollary 6 are satisfied, the true shift must be within* $\pm 2^6$ *of the estimate.*

For each of $k = 6, 5, 4, 3$, *the closest grid point to the true shift is found. For this particular example, since all column differences exceeded* Δ_{zv} *and* $2\|\boldsymbol{e}_v\|$, *the true shift is guaranteed to be within* $\pm 2^{k-1}$ *of the estimates by*

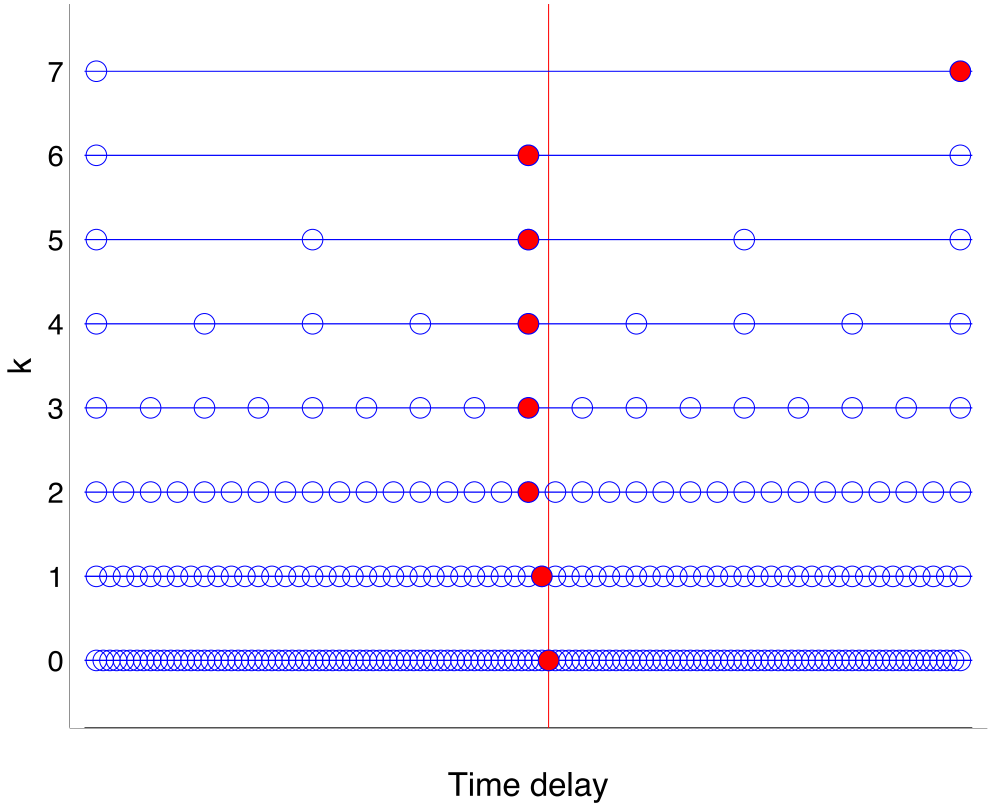

Fig. 4. An illustration of the grid for different k's. The true shift is illustrated by a red vertical line and the shift predicted by CSR marked by a red bullet.

Corollary 6. For $k = 2$, CSR does not return the grid point closest to the true shift and Δ_{zv} exceeds the smallest distances between two columns. For $k = 1$, CSR returns the grid point closest to the true shift but Δ_{zv} exceeds the smallest distances between two columns. When $k = 0$, one grid point is aligned with the true shift and this grid point is correctly identified by CSR. It can be also verified that $\boldsymbol{A}\tilde{\boldsymbol{X}}$ has distinct columns.

The above experiment is further repeated 100 times and the results summarized in Table I. We also run the above Monte Carlo simulation with noise added to $\boldsymbol{z}$ and $\boldsymbol{v}$ (SNR= 10dB). The results are also reported in Table I.

VI. Conclusion

To recover the cyclic shift relating two signals, the cross-correlation is usually evaluated for all possible shifts. Recent advances in hardware, signal acquisition and signal processing have made it possible to sample or compute Fourier coefficients of a signal efficiently. It is therefore of particular interest to see under what conditions the true shift can be recovered from the Fourier coefficients. We have proposed a criterion that is computationally more efficient than using the time samples, and we have shown that the true shift can be recovered using as few as one Fourier coefficient. We have also derived bounds for perfect recovery for both noise free and noisy measurements and introduced the concept of superresolution for shift retrieval.

TABLE I

SUCCESS RATES OF SHIFT RETRIEVAL IN SUPERRESOLUTION IN A MONTE CARLO SIMULATION. FOR BOTH THE NOISE-FREE CASE AND THE NOISY CASE, THE PERCENTAGE OF SUCCESSFUL TRIALS IN WHICH THE OPTIMAL SHIFT ESTIMATES (CLOSEST TO THE TRUE SHIFT) IN DIFFERENT RESOLUTIONS $k = 7, 6, \ldots, 0$ ARE RECOVERED BY CSR AND THE PERCENTAGE PREDICTED BY THE THEORY ARE SHOWN.

	$k = 7$	6	5	4	3	2	1	0
Noise-free (numerical)	95	62	53	59	63	60	100	100
Noise-free (theoretical)	86	51	45	51	55	55	0	100
Noisy (numerical)	79	43	26	30	10	4	4	0
Noisy (theoretical)	28	7	2	0	0	0	0	0

APPENDIX A
PROOFS: NOISE-FREE COMPRESSIVE SHIFT RETRIEVAL

Before proving Theorem 1, we state two lemmas.

Lemma 9 (**Recovery of a Shift using Projections**). *Let $\boldsymbol{X}$ be the $n \times n$-matrix made up of cyclically shifted versions of $\boldsymbol{x}$ as columns. If the columns of $\boldsymbol{AX}$ are distinct, then the true shift can be recovered by*

$$\min_{\boldsymbol{q} \in \{0,1\}^n} \|\boldsymbol{Ay} - \boldsymbol{AXq}\|_2^2 \quad \textit{s.t.} \quad \|\boldsymbol{q}\|_0 = 1. \tag{15}$$

Proof of Lemma 9: Since the shift relating $\boldsymbol{x}$ and $\boldsymbol{y}$ is assumed unique, it is clear that the true shift is recovered by

$$\min_{\boldsymbol{q} \in \{0,1\}^n} \|\boldsymbol{y} - \boldsymbol{Xq}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{q}\|_0 = 1. \tag{16}$$

Assume that the solution of (15) is not equivalent to that of (16). Namely, assume that (16) gives $\boldsymbol{q}$, (15) gives $\tilde{\boldsymbol{q}}$ and $\boldsymbol{q} \neq \tilde{\boldsymbol{q}}$. Since $\boldsymbol{q}$ will give a zero objective value in (15), so must $\tilde{\boldsymbol{q}}$. We therefore have that $\boldsymbol{Ay} = \boldsymbol{AX}\tilde{\boldsymbol{q}} = \boldsymbol{AXq}$ and hence

$$\boldsymbol{AX}\tilde{\boldsymbol{q}} - \boldsymbol{AXq} = \boldsymbol{AX}(\tilde{\boldsymbol{q}} - \boldsymbol{q}) = 0. \tag{17}$$

Since $\boldsymbol{q}$, $\tilde{\boldsymbol{q}} \in \{0,1\}^n$, $\|\tilde{\boldsymbol{q}}\|_0 = \|\boldsymbol{q}\|_0 = 1$, and $\boldsymbol{q} \neq \tilde{\boldsymbol{q}}$, $\boldsymbol{AX}(\tilde{\boldsymbol{q}} - \boldsymbol{q}) = 0$ implies that two columns of $\boldsymbol{AX}$ are identical. This is a contradiction and we therefore conclude that both (15) and (16) recover the true shift. ■

Lemma 10 (**From** (15) **to** (9)). *Under conditions 1) and 2) of Theorem 1, the shifts recovered by* (15) *and* (9) *are the same.*

Proof of Lemma 10: Consider the objective of (15):

$$\begin{aligned}\|\boldsymbol{Ay} - \boldsymbol{AXq}\|_2^2 =& (\boldsymbol{Ay})^*\boldsymbol{Ay} + (\boldsymbol{AXq})^*\boldsymbol{AXq} \\ & - (\boldsymbol{Ay})^*\boldsymbol{AXq} - (\boldsymbol{AXq})^*\boldsymbol{Ay}.\end{aligned} \tag{18}$$

Notice that we can write

$$\boldsymbol{Xq} = \boldsymbol{D}^s\boldsymbol{x}, \tag{19}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

for some s. This follows from the construction of $\boldsymbol{X}$ as a matrix with delayed versions of $\boldsymbol{x}$ as its columns and from the fact that $\boldsymbol{q}$ selects exactly one of these columns. Problem (15) is then equal to

$$\max_{s\in\{0,\ldots,n-1\}} 2\Re\{(\boldsymbol{A}\boldsymbol{y})^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}\} - (\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x})^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}. \tag{20}$$

Using the assumption $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}$ and that $(\boldsymbol{D}^s)^*\boldsymbol{D}^s = \boldsymbol{I}$ for a shift matrix, we have

$$(\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x})^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x} = \boldsymbol{x}^*(\boldsymbol{D}^s)^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x} = \|\boldsymbol{A}\boldsymbol{x}\|_2^2, \tag{21}$$

which is independent of s. Therefore, the shift recovered by (20) is the same as that of

$$\max_{s\in\{0,\ldots,n-1\}} \Re\{(\boldsymbol{A}\boldsymbol{y})^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}\}. \tag{22}$$

Lastly, if we again use that $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}$ and $\alpha\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{I}$, then (9) follows from

$$\begin{aligned}\Re\{(\boldsymbol{A}\boldsymbol{y})^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}\} &= \Re\{\boldsymbol{y}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}\}\\ &= \alpha\Re\{\boldsymbol{y}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{x}\}\\ &= \alpha\Re\{\boldsymbol{y}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}\}\\ &= \alpha\Re\{\langle\boldsymbol{z}, \bar{\boldsymbol{D}}^s\boldsymbol{v}\rangle\}\end{aligned} \tag{23}$$

where $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{y}$ and $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}$. ■

We are now ready to prove Theorem 1.

Proof of Theorem 1: The assumptions of Theorem 1 imply that the requirements of both Lemmas 9 and 10 are satisfied. The theorem therefore follows trivially. ■

We next prove Corollary 2.

Proof of Corollary 2: In the proof of Lemma 9, $\boldsymbol{A}\boldsymbol{X}(\tilde{\boldsymbol{q}} - \boldsymbol{q}) = 0$ leads to $\tilde{\boldsymbol{q}} - \boldsymbol{q} = 0$ if the columns of $\boldsymbol{A}\boldsymbol{X}$ were all distinct. Now, if

$$s^* = \arg\max_{s\in\{0,\ldots,n-1\}} \Re\{\langle\boldsymbol{z}, \bar{\boldsymbol{D}}^s\boldsymbol{v}\rangle\}, \tag{24}$$

then the s^*th element of $\tilde{\boldsymbol{q}}$ is one and all other elements zero. Hence, Lemma 9 can be made less conservative if s^* is known by requiring that only the s^*th column of $\boldsymbol{A}\boldsymbol{X}$ is different than all other columns. ■

Proof of Lemma 3: Let $\boldsymbol{M} = \boldsymbol{A}\boldsymbol{D}^s$ and $\boldsymbol{Q} = \boldsymbol{A}(\boldsymbol{D}^s)^*$. By the definition of $\boldsymbol{D}^s$, $\boldsymbol{M}$ is a column permutation of $\boldsymbol{A}$ where the columns are shifted s times to the right. Thus, the rth column of $\boldsymbol{M}$ is equal to the tth column of $\boldsymbol{A}$ where $t = (r - s) \bmod n$. It is also easy to see that $(\boldsymbol{D}^s)^*$ permutes the columns of $\boldsymbol{A}$ by s to the left so that the rth column of $\boldsymbol{Q}$ is equal to the qth column of $\boldsymbol{A}$ where $q = (r + s) \bmod n$. Now, the prth element of $\boldsymbol{A}^*\boldsymbol{M} = \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{D}^s$ is given by

$$(\boldsymbol{A}_{:,p})^*\boldsymbol{M}_{:,r} = (\boldsymbol{A}_{:,p})^*\boldsymbol{A}_{:,r-s} = \frac{1}{n}\sum_{i=1}^{m} e^{2j\pi k_i(p-r+s)}, \tag{25}$$

where $\boldsymbol{A}_{:,p}$ is used to denote the pth column of $\boldsymbol{A}$ and $\boldsymbol{M}_{:,r}$ the rth column or $\boldsymbol{M}$. On the other hand, the (p, r)-th element of $\boldsymbol{Q}^*\boldsymbol{A} = \boldsymbol{D}^s\boldsymbol{A}^*\boldsymbol{A}$ is given by

$$(\boldsymbol{Q}_{:,p})^*\boldsymbol{A}_{:,r} = (\boldsymbol{A}_{:,p+s})^*\boldsymbol{A}_{:,r} = \frac{1}{n}\sum_{i=1}^{m} e^{2j\pi k_i(p+s-r)}. \tag{26}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

Clearly, the two are equivalent. ■

We are now ready to prove Corollary 4.

Proof of Corollary 4: Lemma 3 gives that Condition 1) of Theorem 1 is satisfied. Since a full Fourier matrix is orthonormal, a matrix made up of a selection of rows of a Fourier matrix satisfies Condition 2). The last condition of Theorem 1 requires columns of $\boldsymbol{AX}$ to be distinct. A sufficient condition is that there exists a row with all distinct elements. As shown previously, the (p, r)-th element of $\boldsymbol{AX}$ is $X_{k_p} e^{\frac{2j\pi k_p (r-1)}{n}}$. If X_{k_p} is assumed nonzero, then a sufficient condition for $\boldsymbol{AX}$ to have distinct columns is that $e^{\frac{2j\pi k_p r_1}{n}} \neq e^{\frac{2j\pi k_p r_2}{n}}$, $r_1, r_2 \in \{0, \dots, n-1\}, r_1 \neq r_2$. This condition can be simplified to $\frac{k_p r_1}{n} \neq \frac{k_p r_2}{n} + \gamma$, $\gamma \in \mathbb{Z}$. By realizing that $r_1 - r_2$ takes values in $\{-n+1, \dots, -1, 1, \dots, n-1\}$ we get that the condition is equivalent to requiring that there are no integers in $\{-n+1, \dots, -1, 1, \dots, n-1\}\frac{k_p}{n}$. Due to symmetry, a sufficient condition for distinct columns is that there exists a $p \in \{1, \dots, m\}$ such that $X_{k_p} \neq 0$ and $\{1, \dots, n-1\}\frac{k_p}{n}$ contains no integers.

Lastly, writing out $\boldsymbol{AD}^s\boldsymbol{A}^*$ we get that the prth element is equal to $\delta_{p,r} e^{-\frac{2j\pi k_p s}{n}}/n$, leading to the simplified test proposed in (11). ■

APPENDIX B
PROOFS: NOISY COMPRESSIVE SHIFT RETRIEVAL

Proof of Theorem 5: From Lemma 10 we can see that seeking s that maximizes $\Re\{\langle \tilde{\boldsymbol{z}}, \bar{\boldsymbol{D}}^s \tilde{\boldsymbol{v}} \rangle\}$ is equivalent to seeking $\boldsymbol{q}$ that solves

$$\min_{\boldsymbol{q} \in \{0,1\}^n} \|\tilde{\boldsymbol{z}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\boldsymbol{q}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{q}\|_0 = 1, \tag{27}$$

where the first column of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ is equal to $\tilde{\boldsymbol{v}}$ (which defines the first column of $\tilde{\boldsymbol{X}}$) and the ith column of $\tilde{\boldsymbol{X}}$ is a circular shift of the first column of $\tilde{\boldsymbol{X}}$ $i-1$ steps.

Assume that $\hat{\boldsymbol{q}}$ solves (27). Since our measurements are noisy, we cannot expect a zero loss. The loss can be shown to be given by

$$\|\tilde{\boldsymbol{z}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2^2 = \|\tilde{\boldsymbol{v}}\|_2^2 + \|\tilde{\boldsymbol{z}}\|_2^2 - \max_{s \in \{0,\dots,n-1\}} 2\Re\{\tilde{\boldsymbol{z}}^* \bar{\boldsymbol{D}}^s \tilde{\boldsymbol{v}}\}. \tag{28}$$

Now, consider $\|\tilde{\boldsymbol{z}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2$. Assume that $\boldsymbol{q}_0$ solves the noise-free version of (27) and let $\tilde{\boldsymbol{X}} = \boldsymbol{X} + \boldsymbol{H}$. We have the following inequality:

$$\begin{aligned}
\|\tilde{\boldsymbol{z}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 &= \|\boldsymbol{z} + \boldsymbol{e}_z - \boldsymbol{z} + \boldsymbol{A}\boldsymbol{X}\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 \\
&= \|\boldsymbol{e}_z + \boldsymbol{A}\boldsymbol{X}\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 \\
&= \|\boldsymbol{e}_z + \boldsymbol{A}(\tilde{\boldsymbol{X}} - \boldsymbol{H})\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 \\
&\geq \|\boldsymbol{A}\tilde{\boldsymbol{X}}\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 - \|\boldsymbol{e}_z\|_2 - \|\boldsymbol{e}_v\|_2,
\end{aligned}$$

where we used the fact that $\boldsymbol{AH}\boldsymbol{q}_0 = \boldsymbol{e}_v$ and the reverse triangle inequality. Therefore

$$\|\boldsymbol{A}\tilde{\boldsymbol{X}}\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2 \leq \Delta_{\boldsymbol{zv}}. \tag{29}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

Since $\|\hat{\boldsymbol{q}}\|_0 = \|\boldsymbol{q}_0\|_0 = 1$, $\|\boldsymbol{A}\tilde{\boldsymbol{X}}\boldsymbol{q}_0 - \boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}}\|_2$ equals the ℓ_2 difference between two columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$. It is hence sufficient to require that the ℓ_2 difference between any two columns of $\boldsymbol{A}\tilde{\boldsymbol{X}}$ is greater than $\Delta_{\boldsymbol{vz}}$ for $\boldsymbol{q}_0 = \hat{\boldsymbol{q}}$. ∎

Proof of Corollary 6: Let $\tilde{\boldsymbol{q}}$ and $\hat{\boldsymbol{q}}$ be any vectors such that $\|\hat{\boldsymbol{q}}\|_0 = \|\tilde{\boldsymbol{q}}\|_0 = 1$, $\hat{\boldsymbol{q}} \neq \tilde{\boldsymbol{q}}$ and $\hat{\boldsymbol{q}}, \tilde{\boldsymbol{q}} \in \{0,1\}^n$. Using the triangle inequality and the fact that $\boldsymbol{H}\hat{\boldsymbol{q}}$ and $\boldsymbol{H}\tilde{\boldsymbol{q}}$ are shifted versions of the same vector, we have that

$$\|\boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{q}}\|_2 = \|\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{H})(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2 \tag{30}$$

$$\leq \|\boldsymbol{A}\boldsymbol{X}(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2 + \|\boldsymbol{A}\boldsymbol{H}(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2 \tag{31}$$

$$\leq \|\boldsymbol{A}\boldsymbol{X}(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2 + 2\|\boldsymbol{e}_v\|_2. \tag{32}$$

Hence, if $\|\boldsymbol{A}\tilde{\boldsymbol{X}}\hat{\boldsymbol{q}} - \boldsymbol{A}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{q}}\|_2 - 2\|\boldsymbol{e}_v\|_2 > 0$, then $\|\boldsymbol{A}\boldsymbol{X}(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2$ is greater than zero. Now since Theorem 5 gives that (11) recovers the same shift as if the measurements would have been noise-free, and since Theorem 1 gives that the noise-free estimate is equal to the true shift if $\|\boldsymbol{A}\boldsymbol{X}(\hat{\boldsymbol{q}} - \tilde{\boldsymbol{q}})\|_2$ is greater than zero (or equivalently that all columns of $\boldsymbol{A}\boldsymbol{X}$ are distinct), we can guarantee the recovery of the true shift also in the noisy case. ∎

Proof of Corollary 7: The corollary follows trivially by setting the s^*th element of $\hat{\boldsymbol{q}}$ to one and all other elements to zero in the proofs of Theorem 5 and Corollary 6. ∎

REFERENCES

[1] J. B.-Y. Tsui, *Fundamentals of Global Positioning System Receivers: A Software Approach*, ser. Wiley Series in Microwave and Optical Engineering. Wiley, 2005.

[2] G. Carter and C. Knapp, "Time delay estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1976.

[3] B. Lucas and T. Kanade, "An iterative image registration technique with an application in stereo vision," in *Int'l J. Conf. Artificial Intelligence*, 1981.

[4] G. Hager and P. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 1025–1039, 1998.

[5] Y. Peng, A. Ganesh, J. Wright, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, 2010.

[6] A. Hero, H. Messer, J. Goldberg, D. Thomson, M. Amin, G. Giannakis, A. Swami, J. Tugnait, A. Nehorai, A. Swindlehurst, J.-F. Cardoso, L. Tong, and J. Krolik, "Highlights of statistical signal and array processing," *IEEE Signal Processing Magazine*, vol. 15, no. 5, pp. 21–64, 1998.

[7] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, 2005.

[8] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, Feb. 2006.

[9] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[10] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[11] M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 98–124, 2011.

[12] T. Chernyakova and Y. C. Eldar, "Fourier domain beamforming: The path to compressed ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Jul. 2013, submitted.

[13] E. Baransky, G. Itzhak, I. Shmuel, N. Wagner, E. Shoshan, and Y. C. Eldar, "A sub-Nyquist radar prototype: Hardware and algorithms," *IEEE Transactions on Aerospace and Electronic Systems, special issue on Compressed Sensing for Radar*, Aug. 2012, to appear.

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

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

[15] R. Tur, Y. C. Eldar, and Z. Friedman, “Innovation rate sampling of pulse streams with application to ultrasound imaging,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011.

[16] K. Gedalyahu, R. Tur, and Y. C. Eldar, “Multichannel sampling of pulse streams at the rate of innovation,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1491–1504, Apr. 2011.

[17] E. Kokiopoulou, D. Kressner, and P. Frossard, “Optimal image alignment with random measurements,” in *EUSIPCO*, 2009.

[18] M. A. Davenport, M. F. Duarte, M. Wakin, J. N. Laska, D. Takhar, K. Kelly, and R. G. Baraniuk, “The smashed filter for compressive classification and target recognition,” in *Proc. Computational Imaging V at SPIE Electronic Imaging*, 2007.

[19] H. Hassanieh, F. Adib, D. Katabi, and P. Indyk, “Faster GPS via the sparse Fourier transform,” in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, ser. Mobicom ’12. New York, NY, USA: ACM, 2012, pp. 353–364.

[20] A. Andoni, P. Indyk, D. Katabi, and H. Hassanieh, “Shift finding in sub-linear time,” in *SODA*, 2013, pp. 457–465.

[21] G. E. P. Box and G. Jenkins, *Time Series Analysis, Forecasting and Control (3rd ed.)*. Holden-Day, Incorporated, 1994.

[22] L. Rayleigh, “On pin-hole photography,” *The London, Edinburg and Dublin philosophical magazine and journal of science*, vol. 31, no. 5, Feb. 1891.

[23] S. Gazit, A. Szameit, Y. C. Eldar, and M. Segev, “Super-resolution and reconstruction of sparse sub-wavelength images,” *Opt. Express*, vol. 17, no. 26, pp. 23 920–23 946, Dec 2009.

[24] E. J. Candès and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Communications on Pure and Applied Mathematics*, 2013.

Henrik Ohlsson was born in Sweden in 1981. He received the M.Sc. degree in Applied Physics and Electrical Engineering in Oct. 2006 and his Ph.D. degree in Automatic Control in Nov. 2010, all from Linköping University, Sweden. He has held visiting positions at the University of Cambridge and the University of Massachusetts. His research interests are mainly within the areas of system identification and machine learning. Henrik is currently a visiting scholar at University of California at Berkeley and an Assistant Professor at Linköping University.

Yonina C. Eldar (S'98-M'02-SM'07-F'12) 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 2002.

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 and holds the The Edwards Chair in Engineering. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT and a Visiting Professor at Stanford University, Stanford, CA. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, in 2007 the Henry Taub Prize for Excellence in Research, in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion Outstanding Lecture Award, in 2009 the Technion's Award for Excellence in Teaching, in 2010 the Michael Bruno Memorial Award from the Rothschild Foundation, and in 2011 the Weizmann Prize for Exact Sciences. In 2012 she was elected to the Young Israel Academy of Science and to the Israel Committee for Higher Education, and elected an IEEE Fellow. In 2013 she received the Technion's Award for Excellence in Teaching, the Hershel Rich Innovation Award, and the IEEE Signal Processing Technical Achievement Award, and in 2014 the IEEE/AESS Fred Nathanson Memorial Radar Award. She received several best paper awards together with her research students and colleagues. She is the Editor in Chief of Foundations and Trends in Signal Processing and a member of the IEEE Sensor Array and Multichannel Technical Committee. In the past, she was a Signal Processing Society Distinguished Lecturer, member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing technical committees, and served as 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.

Allen Y. Yang serves as the CTO of Atheer Inc., an immersive computing startup based in Mountain View, CA. Currently he is also a principal investigator at UC Berkeley. His primary research areas include high-dimensional pattern recognition, computer vision, image processing, and applications in motion segmentation, image segmentation, face recognition, and sensor networks. He has published three books/chapters, 13 journal papers and more than 30 conference papers. He is also an inventor of 12 US patents/applications.

Allen received his BEng degree in Computer Science from the University of Science and Technology of China (USTC) in 2001. From the University of Illinois at Urbana-Champaign (UIUC), he received two MS degrees in Electrical Engineering and Mathematics in 2003 and 2005, respectively, and a PhD in Electrical and Computer Engineering in 2006. Among the awards he received are a Best Bachelor's Thesis Award from USTC in 2001, a Henry Ford II Scholar Award from UIUC in 2003, a Best Paper Award from the International Society of Information Fusion and a Best Student Paper Award from Asian Conference on Computer Vision in 2009. He is a Senior Member of IEEE.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2014.2332974, IEEE Transactions on Signal Processing

S. Shankar Sastry received his Ph.D. degree in 1981 from the University of California, Berkeley. He was on the faculty of MIT as Assistant Professor from 1980 to 82 and Harvard University as a chaired Gordon Mc Kay professor in 1994. He is currently the Dean of Engineering at University of California, Berkeley. His areas of personal research are embedded control especially for wireless systems, cybersecurity for embedded systems, critical infrastructure protection, autonomous software for unmanned systems (especially aerial vehicles), computer vision, nonlinear and adaptive control, control of hybrid and embedded systems, and network embedded systems and software. He has supervised over 60 doctoral students and over 50 M.S. students to completion. His students now occupy leadership roles in several places and on the faculties of many major universities. He has coauthored over 450 technical papers and 9 books. Dr. Sastry served on the editorial board of numerous journals, and is currently an Associate Editor of the IEEE Proceedings.

Dr. Sastry was elected into the National Academy of Engineering in 2001 and the American Academy of Arts and Sciences (AAAS) in 2004. He also received the President of India Gold Medal in 1977, the IBM Faculty Development award for 19831985, the NSF Presidential Young Investigator Award in 1985 and the Eckman Award of the American Automatic Control Council in 1990, the Ragazzini Award for Distinguished Accomplishments in teaching in 2005, an M. A. (honoris causa) from Harvard in 1994, Fellow of the IEEE in 1994, the distinguished Alumnus Award of the Indian Institute of Technology in 1999, and the David Marr prize for the best paper at the International Conference in Computer Vision in 1999, an honorary doctorate from the Royal Swedish Institute of Technology in 2007 and the C.L. Tien Award for Academic Leadership in 2010.