Compressive Spectral Estimation for Nonstationary Random Processes

Alexander Jung, Georg Tauböck, *Member, IEEE*, and Franz Hlawatsch, *Fellow, IEEE*

Abstract— Estimating the spectral characteristics of a nonstationary random process is an important but challenging task, which can be facilitated by exploiting structural properties of the process. In certain applications, the observed processes are underspread, i.e., their time and frequency correlations exhibit a reasonably fast decay, and approximately time-frequency sparse, i.e., a reasonably large percentage of the spectral values are small. For this class of processes, we propose a compressive estimator of the discrete Rihaczek spectrum (RS). This estimator combines a minimum variance unbiased estimator of the RS (which is a smoothed Rihaczek distribution using an appropriately designed smoothing kernel) with a compressed sensing technique that exploits the approximate time-frequency sparsity. As a result of the compression stage, the number of measurements required for good estimation performance can be significantly reduced. The measurements are values of the ambiguity function of the observed signal at randomly chosen time and frequency lag positions. We provide bounds on the mean-square estimation error of both the minimum variance unbiased RS estimator and the compressive RS estimator, and we demonstrate the performance of the compressive estimator by means of simulation results. The proposed compressive RS estimator can also be used for estimating other time-dependent spectra (e.g., the Wigner-Ville spectrum) since for an underspread process most spectra are almost equal.

Index Terms— Nonstationary random process, nonstationary spectral estimation, time-dependent power spectrum, Rihaczek spectrum, Wigner-Ville spectrum, compressed sensing, basis pursuit, cognitive radio.

I. INTRODUCTION

Estimating the spectral characteristics of a random process is an important task in many signal analysis and processing problems. Conventional spectral estimation based on the *power spectral density* is restricted to wide-sense stationary and, by extension, wide-sense cyclostationary processes [1], [2]. However, in many applications—including speech and audio, communications, image processing, computer vision, biomedical engineering, and machine monitoring—the signals of interest cannot be well modeled as wide-sense (cyclo)stationary processes. For example, in cognitive radio systems [3]–[5], the receiver has to infer from the received signal the location

Alexander Jung, Georg Tauböck, and Franz Hlawatsch are with the Institute of Telecommunications, Vienna University of Technology, A-1040 Vienna, Austria (e-mail: {ajung,gtauboec,fhlawats}@nt.tuwien.ac.at). This work was supported by the FWF under grant S10603 (Statistical Inference) within the National Research Network SISE and by the WWTF under grant MA 07-004 (SPORTS). Parts of this work were previously presented at IEEE SSP 2009, Cardiff, Wales, UK, Aug.–Sept. 2009.

of unoccupied frequency bands ("spectral holes") that can be used for data transmission. Here, modeling the received signal as a nonstationary process can be advantageous because it potentially allows a faster estimation of time-varying changes in band occupation [3].

For a general nonstationary process, a "power spectral density" that is nonnegative and extends all the essential properties of the conventional power spectral density is not available [6]–[10]. Several different definitions of a "time-dependent (or time-varying) power spectrum" have been proposed in the literature, see [6]–[26] and references therein. However, it has been shown [10], [24] that in the practically important case of nonstationary processes with fast decaying time-frequency (TF) correlations—so-called *underspread* processes [10], [24]–[30]—all major spectra yield effectively identical results, are (at least approximately) real-valued and nonnegative, and satisfy several other desirable properties at least approximately. Thus, in the underspread case, the specific choice of a spectrum is of secondary theoretical importance and can hence be guided by practical considerations such as computational complexity.

Once a specific definition of time-dependent spectrum has been adopted, an important problem is the estimation of the spectrum from a single observed realization of the process. This nonstationary spectral estimation problem is fundamentally more difficult than spectral estimation in the (cyclo)stationary case, because long-term averaging cannot be used to reduce the mean-square error (MSE) of the estimate. Formally, any estimator of a nonparametric time-dependent spectrum can also be viewed as a TF representation of the observed signal [7], [26], [31], [32]. Estimators have been previously proposed for several spectra including the Wigner-Ville spectrum and the Rihaczek spectrum (RS) (e.g., [7]–[9], [21], [26], [33]–[41]).

In this paper, extending our work in [42], we propose a "compressive" estimator of the RS that uses the recently introduced methodology of compressed sensing (CS) [43], [44]. The proposed estimator is suited to underspread processes that are *approximately TF sparse*. The latter property means that only a moderate percentage of the values of the discrete RS are significantly nonzero. Both assumptions—underspreadness and TF sparsity—are reasonably well satisfied in many applications, including, e.g., cognitive radio. We consider the RS because it is the simplest time-dependent spectrum from a computational viewpoint, especially in the discrete setting used. The proposed compressive estimator of the RS is obtained by augmenting a basic noncompressive estimator

(a smoothed version of the Rihaczek distribution (RD), cf. [7], [14], [26], [32], [33], [35], [38]–[41]) with a CS compression-reconstruction stage. Algorithmically, our estimator is similar to the compressive TF representation proposed in [45], [46]. In fact, both our estimator and the TF representation of [45], [46] are essentially based on a sparsity-regularized inversion of the Fourier transform relationship between a TF distribution and the values of the ambiguity function (AF) taken at randomly chosen time lag/frequency lag locations. The sparsity-regularization is achieved by requiring a small ℓ_1-norm of the resulting TF distribution. However, the setting of [45], [46] is that of deterministic TF signal analysis (more specifically, the goal is to improve the TF localization properties of the Wigner distribution), whereas we consider a stochastic setting, namely, spectral estimation for underspread, approximately TF sparse, nonstationary random processes.

Compressive spectral estimation methods have been proposed previously, also in the context of cognitive radio [47]–[50]. However, these methods are restricted to the estimation of the power spectral density of stationary or cyclostationary processes. Furthermore, they perform CS directly on the observed signal (process realization), whereas our method performs CS on an estimate of a TF autocorrelation function known as the expected ambiguity function (EAF). This EAF estimate is a quadratic time lag/frequency lag representation of the observed signal that is based on the signal's AF. It is an intermediate step in the calculation of the spectral estimator, somewhat similar to a sufficient statistic. In some sense, we perform a twofold compression, first by using only an EAF estimate (instead of the raw observed signal) for spectral estimation and secondly by "compressing" that estimate. This approach can be advantageous if dedicated hardware units for computing values of the EAF estimate (i.e., AF) from an observed continuous-time signal are employed [51]–[54], because fewer such units are required. It can also be advantageous if the values of the EAF estimate have to be transmitted over low-rate links—e.g., in wireless sensor networks [55]—or stored in a memory, because fewer such values need to be transmitted or stored.

The fact that we perform CS in the AF domain and not directly on the signal is a somewhat nonorthodox aspect of our method. Indeed, the objective of this paper is not to develop a sub-Nyquist sampling scheme in the spirit of, e.g., spectrum-blind sampling [56], [57]. Our work is based on the assumption that the original signal of interest is modeled as a continuous-time random process $X(t)$ that can be (approximately) represented by a finite-length, discrete-time random process $X[n]$. This discrete-time random process itself is not used in a practical application of our method; it is only used for the theoretical development of the method. A second assumption is that values of the AF of a continuous-time process realization $x(t)$ can be computed efficiently. The computation of the AF values from $x(t)$ using dedicated hardware is described in [51]–[54].

A major focus of our work is an analysis of the estimation accuracy of the proposed compressive estimator. Because finding a closed-form expression of the MSE is intractable, we derive upper bounds on the MSE. These bounds depend on two components: the first component is determined by the degree of "underspreadness," corresponding to the concentration of the EAF of the observed process; the second component is related to the TF sparsity properties of the observed process. As we will see below, there is a tradeoff between these components, since a well concentrated EAF of an underspread process tends to imply a poorly concentrated RS, which is disadvantageous in terms of TF sparsity.

The remainder of this paper is organized as follows. In Section II, we state our general setting and review some fundamentals of nonstationary random processes and their TF representation. In Section III, we describe a basic noncompressive estimator of the RS. In Section IV, we develop a compressive estimator by augmenting the noncompressive estimator with a CS compression-reconstruction stage. Bounds on the MSE of both the noncompressive and compressive estimators are derived in Section V. Finally, numerical results are presented in Section VI.

Notation. The modulus, complex conjugate, real part, and imaginary part of a complex number $a \in \mathbb{C}$ are denoted by $|a|$, a^*, $\Re\{a\}$, and $\Im\{a\}$, respectively. Boldface lowercase letters denote column vectors and boldface uppercase letters denote matrices. The kth entry of a vector $\mathbf{a}$ is denoted by $(\mathbf{a})_k$, and the entry of a matrix $\mathbf{A}$ in the ith row and jth column by $(\mathbf{A})_{i,j}$. The superscripts ^{T}, *, and ^{H} denote the transpose, conjugate, and Hermitian transpose, respectively, of a vector or matrix. The ℓ_1-norm of a vector $\mathbf{a} \in \mathbb{C}^L$ is denoted by $\|\mathbf{a}\|_1 \triangleq \sum_{k=1}^{L} |(\mathbf{a})_k|$, and the ℓ_2-norm by $\|\mathbf{a}\|_2 \triangleq \sqrt{\mathbf{a}^H \mathbf{a}}$. The number of nonzero entries is denoted by $\|\mathbf{a}\|_0$. The trace of a square matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$ is denoted by $\mathrm{tr}\{\mathbf{A}\} \triangleq \sum_{k=1}^{M} (\mathbf{A})_{k,k}$. Given a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, we denote by $\mathrm{vec}\{\mathbf{A}\} \in \mathbb{C}^{MN}$ the vector obtained by stacking all columns of $\mathbf{A}$. Given two matrices $\mathbf{A} \in \mathbb{C}^{M_1 \times N_1}$ and $\mathbf{B} \in \mathbb{C}^{M_2 \times N_2}$, we denote by $\mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{M_1 M_2 \times N_1 N_2}$ their Kronecker product [58]. The inner product of two square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{M \times M}$ is defined as $\langle \mathbf{A}, \mathbf{B} \rangle \triangleq \mathrm{tr}\{\mathbf{A}\mathbf{B}^H\}$. The Kronecker delta is denoted by $\delta[m]$, i.e., $\delta[m] = 1$ if $m = 0$ and $\delta[m] = 0$ otherwise. Finally, $[N] \triangleq \{0, 1, \ldots, N-1\}$.

II. EAF AND RS

In this section, we state our setting and review some fundamentals of the TF representation of nonstationary random processes. Let $X(t)$ be a bandlimited nonstationary continuous-time random process that can be equivalently represented by a nonstationary discrete-time random process $X[n]$. We assume that $X[n]$ is zero-mean, circularly symmetric complex, and defined for $n \in [N]$. (As mentioned above, the proposed compressive estimator does not presuppose that the discrete-time samples $X[n]$ are actually computed.) The autocorrelation function of the process $X[n]$ is given by $\gamma_X[n_1, n_2] \triangleq \mathrm{E}\{X[n_1]X^*[n_2]\}$, where $\mathrm{E}\{\cdot\}$ denotes expectation. Since $X[n]$ is only defined for $n \in [N]$, we consider $\gamma_X[n_1, n_2]$ only for $n_1, n_2 \in [N]$. This is justified for a process that is well concentrated in the interval $[N]$. An equivalent representation of $\gamma_X[n_1, n_2]$ is the correlation matrix $\mathbf{\Gamma}_X \triangleq \mathrm{E}\{\mathbf{x}\mathbf{x}^H\}$, where $\mathbf{x} \triangleq (X[0] \ X[1] \ \cdots \ X[N-1])^T \in \mathbb{C}^N$; note that $(\mathbf{\Gamma}_X)_{n_1+1, n_2+1} = \gamma_X[n_1, n_2]$ for $n_1, n_2 \in [N]$.

We assume that $X[n]$ is an *underspread* process [10], [24]–[30], which means that its correlation in time and frequency

decays reasonably fast. The underspread property is phrased mathematically in terms of the discrete EAF, which is defined as the following discrete Fourier transform (DFT) of the autocorrelation function [10], [24]–[27], [29], [59]:

$$\bar{A}_X[m,l] \triangleq \sum_{n \in [N]} \gamma_X[n, n-m]_N \, e^{-j\frac{2\pi}{N}ln} \,. \qquad (1)$$

Here, m and l denote discrete time lag and discrete frequency lag, respectively, and[1] $[n_1, n_2]_N \triangleq [n_1 \bmod N, n_2 \bmod N]$. Note that this definition of $\bar{A}_X[m,l]$ is N-periodic in both m and l. The EAF $\bar{A}_X[m,l]$ is a TF-lag representation of the second-order statistics of $X[n]$ that describes the TF correlation structure of $X[n]$. A nonstationary process $X[n]$ is said to be underspread if its EAF is well concentrated around the origin in the (m,l)-plane, i.e.,

$$\bar{A}_X[m,l] \approx 0, \quad \forall (m,l) \notin \mathcal{A}, \quad \text{with } \mathcal{A} \triangleq \{-M, \ldots, M\}_N$$
$$\times \{-L, \ldots, L\}_N \,,$$
$$\text{where } 0 \le M < \left\lfloor \frac{N}{2} \right\rfloor, \; 0 \le L < \left\lfloor \frac{N}{2} \right\rfloor, \text{ and } ML \ll N. \qquad (2)$$

Here, e.g., $\{-M, \ldots, M\}_N$ denotes the N-periodic continuation of the interval $\{-M, \ldots, M\}$. The concentration of the EAF around the origin can be measured by the *EAF moment* defined in Section V-A (see (47)). For later reference, we note that the EAF is the expectation of the AF [7], [31], [32]

$$A_X[m,l] \triangleq \sum_{n \in [N]} X[n] \, X^*[n-m]_N \, e^{-j\frac{2\pi}{N}ln}, \qquad (3)$$

i.e., $\bar{A}_X[m,l] = \mathrm{E}\{A_X[m,l]\}$.

Nonstationary spectral estimation is the problem of estimating a "time-dependent power spectrum" of the nonstationary process $X[n]$ from a single realization $x[n]$ observed for $n \in [N]$. As mentioned earlier, there is no definition of a "time-dependent power spectrum" that satisfies all desirable properties [6]–[10]. However, in the underspread case considered, most reasonable definitions of a time-dependent power spectrum are approximately equal, represent the mean energy distribution of the process over time and frequency, and approximately satisfy all desirable properties [10], [24]. Therefore, we use the simplest such definition, which is the RS [7], [9], [14], [36]. The discrete RS is defined as the following DFT of the autocorrelation function:

$$\bar{R}_X[n,k] \triangleq \sum_{m \in [N]} \gamma_X[n, n-m]_N \, e^{-j\frac{2\pi}{N}km} \,. \qquad (4)$$

Just as the EAF and AF, the RS $\bar{R}_X[n,k]$ is N-periodic in both its variables. Furthermore, the RS is complex-valued in general, but it is approximately real-valued and nonnegative in the underspread case [10], [24]. The RS is related to the EAF via a symplectic two-dimensional (2D) DFT:

---

[1]It will be convenient to consider length-N functions as periodic functions with period N.

$$\bar{R}_X[n,k] = \frac{1}{N} \sum_{m,l \in [N]} \bar{A}_X[m,l] \, e^{-j\frac{2\pi}{N}(km-nl)} \,, \qquad (5)$$

$$\bar{A}_X[m,l] = \frac{1}{N} \sum_{n,k \in [N]} \bar{R}_X[n,k] \, e^{j\frac{2\pi}{N}(mk-ln)} \,. \qquad (6)$$

Relation (5) extends the Fourier transform relation between the power spectral density and the autocorrelation function of a stationary process [1], [2] to the nonstationary case. It follows from (5) that the RS of an underspread process is a smooth function. Furthermore, the RS is the expectation of the RD defined as [7], [14], [31], [32], [36]

$$R_X[n,k] \triangleq \sum_{m \in [N]} X[n]_N X^*[n-m]_N \, e^{-j\frac{2\pi}{N}km}$$
$$= X[n]_N \hat{X}^*[k]_N \, e^{-j\frac{2\pi}{N}nk} \,,$$

where $\hat{X}[k] \triangleq \sum_{n \in [N]} X[n] e^{-j\frac{2\pi}{N}kn}$ is the DFT of $X[n]$. That is, $\bar{R}_X[n,k] = \mathrm{E}\{R_X[n,k]\}$. The 2D DFT relations (5), (6) hold also for the RD and AF, i.e.,

$$R_X[n,k] = \frac{1}{N} \sum_{m,l \in [N]} A_X[m,l] \, e^{-j\frac{2\pi}{N}(km-nl)} \,, \qquad (7)$$

$$A_X[m,l] = \frac{1}{N} \sum_{n,k \in [N]} R_X[n,k] \, e^{j\frac{2\pi}{N}(mk-ln)} \,. $$

Our central assumption, besides the underspread property, is that the nonstationary process $X[n]$ is "approximately TF sparse" in the sense that only a moderate percentage of the RS values $\bar{R}_X[n,k]$ within the fundamental (n,k)-region $[N]^2 = [N] \times [N]$ are significantly nonzero. For such approximately TF sparse processes, we will develop a compressive estimator of the RS by augmenting a basic RS estimator with a compression-reconstruction stage. We present the basic estimator first.

III. BASIC RS ESTIMATOR

In analogy to well-known estimators of the Wigner-Ville spectrum [7]–[9], [21], [26], [37], [38], a basic (noncompressive) estimator of the RS $\bar{R}_X[n,k]$ is given by the following smoothed version of the RD [36], [38]:

$$\hat{R}_X[n,k] \triangleq \frac{1}{N} \sum_{n',k' \in [N]} \Phi[n-n', k-k'] \, R_X[n',k'] \,. \qquad (8)$$

Here, $\Phi[n,k]$ is a smoothing function that is N-periodic in both arguments. Because of (6), the symplectic 2D inverse DFT of $\hat{R}_X[n,k]$,

$$\hat{A}_X[m,l] \triangleq \frac{1}{N} \sum_{n,k \in [N]} \hat{R}_X[n,k] \, e^{j\frac{2\pi}{N}(mk-ln)} \,, \qquad (9)$$

can be viewed as an estimator of the EAF $\bar{A}_X[m,l]$. Using (8) and (7) in (9), we obtain

$$\hat{A}_X[m,l] = \phi[m,l] \, A_X[m,l] \,, \qquad (10)$$

where the 2D window (weighting, taper) function $\phi[m,l]$ is related to the smoothing function $\Phi[n,k]$ through a 2D DFT,

i.e.,

$$\phi[m,l] \triangleq \frac{1}{N} \sum_{n,k \in [N]} \Phi[n,k] e^{j\frac{2\pi}{N}(mk-ln)}. \tag{11}$$

Note that $\phi[m,l]$ and $\hat{A}_X[m,l]$ are N-periodic in both m and l.

We now consider the choice of the smoothing function $\Phi[n,k]$ or, equivalently, of the window function $\phi[m,l]$. Our performance criterion is the MSE

$$\varepsilon \triangleq \mathrm{E}\{\|\hat{R}_X - \bar{R}_X\|_2^2\} = \sum_{n,k \in [N]} \mathrm{E}\{|\hat{R}_X[n,k] - \bar{R}_X[n,k]|^2\}.$$

The MSE can be decomposed as $\varepsilon = B^2 + V$ with the squared bias term $B^2 \triangleq \|\mathrm{E}\{\hat{R}_X\} - \bar{R}_X\|_2^2$ and the variance $V \triangleq \mathrm{E}\{\|\hat{R}_X - \mathrm{E}\{\hat{R}_X\}\|_2^2\}$. We will consider a *minimum variance unbiased* (MVU) design[2] of $\Phi[n,k]$. This means that $\hat{R}_X[n,k]$ is required to be unbiased, i.e., $B = 0$, and the variance V is minimized under this constraint. More specifically, we will adopt the MVU design proposed in [26], [38], which is based on the idealizing assumption that the EAF $\bar{A}_X[m,l]$ is supported on a periodized rectangular region $\mathcal{A} = \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$, i.e., $\bar{A}_X[m,l] = 0$ for all $(m,l) \notin \mathcal{A}$, with $0 \le M < \lfloor N/2 \rfloor$ and $0 \le L < \lfloor N/2 \rfloor$. This is somewhat similar to the underspread property (2); however, it is an exact, rather than approximate, support constraint. As a further difference from the underspread property, we do not require that $ML \ll N$. We note that this idealizing exact support constraint is only needed for the MVU interpretation of our design of $\Phi[n,k]$; in particular, it will not be used for our performance analysis in Section V. The size of $\mathcal{A}$— i.e., the choice of L and M—is a design parameter that can be chosen freely in principle. The resulting estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$ (cf. (17)) can be applied to any process $X[n]$, including, in particular, processes whose EAF $\bar{A}_X[m,l]$ is not exactly supported on $\mathcal{A}$.

We briefly review the derivation of the MVU smoothing function presented in [26], [38]. Using (10) and $\mathrm{E}\{A_X[m,l]\} = \bar{A}_X[m,l]$, the bias term $B^2 = \|\mathrm{E}\{\hat{R}_X\} - \bar{R}_X\|_2^2 = \|\mathrm{E}\{\hat{A}_X\} - \bar{A}_X\|_2^2$ can be expressed as

$$B^2 = \sum_{m,l \in [N]} |(\phi[m,l] - 1)\bar{A}_X[m,l]|^2. \tag{12}$$

Thus, $B^2 = 0$ if and only if $\phi[m,l] = 1$ on the support of $\bar{A}_X[m,l]$, i.e., for all $(m,l) \in \mathcal{A}$. Under the constraint $B^2 = 0$, minimizing the variance of $\hat{R}_X[n,k]$ is equivalent to minimizing the mean power

$$P \triangleq \mathrm{E}\{\|\hat{R}_X\|_2^2\}$$
$$\overset{(9)}{=} \mathrm{E}\{\|\hat{A}_X\|_2^2\}$$
$$\overset{(10)}{=} \mathrm{E}\{\|\phi[m,l] A_X[m,l]\|_2^2\}$$
$$= \sum_{m,l \in [N]} |\phi[m,l]|^2 \mathrm{E}\{|A_X[m,l]|^2\}.$$

[2]The MVU design is analytically tractable and well established in TF spectrum estimation [26], [38]. An alternative design of $\Phi[n,k]$ could be based on the minimax rationale [60]; however, there does not seem to exist a simple solution to the minimax design problem.

Splitting this sum into a sum over $[N]^2 \cap \mathcal{A}$ (where $\phi[m,l] = 1$) and a sum over $[N]^2 \cap \bar{\mathcal{A}}$ (here, $\bar{\mathcal{A}}$ denotes the complement of $\mathcal{A}$), it is clear that P is minimized if and only if the latter sum is zero. This means that $\phi[m,l]$ must be zero for $(m,l) \in [N]^2 \cap \bar{\mathcal{A}}$, and further, due to the periodicity of $\phi[m,l]$, for $(m,l) \in \bar{\mathcal{A}}$. Thus, we conclude that the MVU window function (DFT of the MVU smoothing function) is the indicator function $I_{\mathcal{A}}[m,l]$ of the EAF support $\mathcal{A} = \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$:

$$\phi_{\mathrm{MVU}}[m,l] = I_{\mathcal{A}}[m,l] \triangleq \begin{cases} 1, & (m,l) \in \mathcal{A} \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The corresponding EAF estimator in (10) is obtained as

$$\hat{A}_{X,\mathrm{MVU}}[m,l] = \phi_{\mathrm{MVU}}[m,l] A_X[m,l]$$
$$= I_{\mathcal{A}}[m,l] A_X[m,l] \tag{14}$$
$$= \begin{cases} A_X[m,l], & (m,l) \in \mathcal{A} \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

Therefore, the MVU estimator of the RS is given by (see (9))

$$\hat{R}_{X,\mathrm{MVU}}[n,k] = \frac{1}{N} \sum_{m,l \in [N]} \hat{A}_{X,\mathrm{MVU}}[m,l] e^{-j\frac{2\pi}{N}(km-nl)} \tag{16}$$
$$= \frac{1}{N} \sum_{m=-M}^{M} \sum_{l=-L}^{L} A_X[m,l] e^{-j\frac{2\pi}{N}(km-nl)}, \tag{17}$$

where the periodicity of the summand with respect to m and l has been exploited in the last step.

IV. COMPRESSIVE RS ESTIMATOR

Next, we will augment the basic RS estimator presented in the previous section with a compression-reconstruction stage.

A. Basic DFT Relation

The proposed compressive RS estimator is based on a 2D DFT relation that will now be derived. We recall from (15) that the EAF estimate $\hat{A}_{X,\mathrm{MVU}}[m,l]$ is exactly zero outside the effective EAF support $\mathcal{A} = \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$, where $0 \le M < \lfloor N/2 \rfloor$ and $0 \le L < \lfloor N/2 \rfloor$. In what follows, we will denote by

$$S \triangleq |[N]^2 \cap \mathcal{A}| = (2M+1)(2L+1) \tag{18}$$

the size of one period of $\mathcal{A}$. Because $2M+1$ and $2L+1$ do not necessarily divide N, we furthermore define an "extended effective EAF support" as the periodized rectangular region $\mathcal{A}' \triangleq \{-M,\ldots,-M+\Delta M-1\}_N \times \{-L,\ldots,-L+\Delta L-1\}_N$. Here, ΔM and ΔL are chosen as the smallest integers such that $\Delta M \ge 2M+1$ and $\Delta L \ge 2L+1$ and, moreover, ΔM and ΔL divide N, i.e, there are integers Δn, Δk such that $\Delta n \Delta L = \Delta k \Delta M = N$ or, equivalently,

$$\Delta n = \frac{N}{\Delta L}, \qquad \Delta k = \frac{N}{\Delta M}. \tag{19}$$

The size of one period of $\mathcal{A}'$ is

$$S' \triangleq |[N]^2 \cap \mathcal{A}'| = \Delta M \Delta L.$$

Note that

$$\mathcal{A} \subseteq \mathcal{A}', \quad S \le S', \tag{20}$$

although typically $S \approx S'$. Let us arrange the values of one period of $\hat{A}_{X,\mathrm{MVU}}[m,l]$ that are located within $\mathcal{A}'$ into a matrix $\mathbf{A} \in \mathbb{C}^{\Delta M \times \Delta L}$, i.e.,

$$(\mathbf{A})_{m+1,l+1} \triangleq \hat{A}_{X,\mathrm{MVU}}[m-M, l-L],$$
$$m \in [\Delta M], \ l \in [\Delta L]. \tag{21}$$

Alternatively, we can represent $\hat{A}_{X,\mathrm{MVU}}[m,l]$ by the matrix $\mathbf{R} \in \mathbb{C}^{\Delta L \times \Delta M}$ whose entries are given by the following 2D DFT of dimension $\Delta L \times \Delta M$:

$$(\mathbf{R})_{p+1,q+1} \triangleq \sum_{m \in [\Delta M]} \sum_{l \in [\Delta L]} (\mathbf{A})_{m+1,l+1}$$
$$\times e^{-j2\pi \left(\frac{q(m-M)}{\Delta M} - \frac{p(l-L)}{\Delta L} \right)} \tag{22}$$

$$\overset{(21)}{=} \sum_{m=-M}^{-M+\Delta M-1} \sum_{l=-L}^{-L+\Delta L-1} \hat{A}_{X,\mathrm{MVU}}[m,l]$$
$$\times e^{-j2\pi \left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L} \right)}$$

$$\overset{(15)}{=} \sum_{m=-M}^{M} \sum_{l=-L}^{L} A_X[m,l] \, e^{-j2\pi \left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L} \right)},$$
$$p \in [\Delta L], \ q \in [\Delta M]. \tag{23}$$

It can be seen by comparing (23) and (17) that the matrix entries $(\mathbf{R})_{p+1,q+1}$ equal (up to a constant factor) a subsampled version of $\hat{R}_{X,\mathrm{MVU}}[n,k]$, i.e.,

$$(\mathbf{R})_{p+1,q+1} = N \hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k],$$
$$p \in [\Delta L], \ q \in [\Delta M], \tag{24}$$

with $\Delta n = N/\Delta L$ and $\Delta k = N/\Delta M$ as in (19). This subsampling does not cause a loss of information because $\hat{A}_{X,\mathrm{MVU}}[m,l]$ is supported in $\mathcal{A}$, and therefore, by (20), also in $\mathcal{A}' = \{-M, \ldots, -M+\Delta M-1\}_N \times \{-L, \ldots, -L+\Delta L-1\}_N$.

Inverting (22), we obtain

$$(\mathbf{A})_{m+1,l+1}$$
$$= \frac{1}{S'} \sum_{p \in [\Delta L]} \sum_{q \in [\Delta M]} (\mathbf{R})_{p+1,q+1} \, e^{j2\pi \left(\frac{(m-M)q}{\Delta M} - \frac{(l-L)p}{\Delta L} \right)},$$
$$m \in [\Delta M], \ l \in [\Delta L]. \tag{25}$$

This 2D DFT relation will constitute an important basis for our compressive RS estimator. It can be compactly written as

$$\mathbf{U}\mathbf{r} = \mathbf{a}, \tag{26}$$

where $\mathbf{r} \triangleq \mathrm{vec}\{\mathbf{R}\} \in \mathbb{C}^{S'}$, $\mathbf{a} \triangleq \mathrm{vec}\{\mathbf{A}^T\} \in \mathbb{C}^{S'}$, and

$$\mathbf{U} \triangleq \frac{1}{S'} \mathbf{F}_{\Delta M}^* \otimes \mathbf{F}_{\Delta L} \in \mathbb{C}^{S' \times S'}, \tag{27}$$

with $\mathbf{F}_{\Delta M}$ defined as $(\mathbf{F}_{\Delta M})_{q+1,m+1} \triangleq e^{-j2\pi \frac{q(m-M)}{\Delta M}}$, $q, m \in [\Delta M]$ and $\mathbf{F}_{\Delta L}$ defined as $(\mathbf{F}_{\Delta L})_{p+1,l+1} \triangleq e^{-j2\pi \frac{p(l-L)}{\Delta L}}$, $p, l \in [\Delta L]$.

Furthermore, using (21) in (16), we obtain

$$\hat{R}_{X,\mathrm{MVU}}[n,k] = \frac{1}{N} \sum_{m \in [\Delta M]} \sum_{l \in [\Delta L]} (\mathbf{A})_{m+1,l+1}$$
$$\times e^{-j\frac{2\pi}{N}[k(m-M) - n(l-L)]}.$$

Inserting (25), we see that the basic RS estimate $\hat{R}_{X,\mathrm{MVU}}[n,k]$ can be calculated from $\mathbf{r}$ (or, equivalently, from $\mathbf{R}$) according to

$$\hat{R}_{X,\mathrm{MVU}}[n,k] = \mathcal{L}\{\mathbf{r}\}[n,k]$$
$$\triangleq \frac{1}{NS'} \sum_{m \in [\Delta M]} \sum_{l \in [\Delta L]} \left[\sum_{p \in [\Delta L]} \sum_{q \in [\Delta M]} (\mathbf{R})_{p+1,q+1} \right.$$
$$\left. \times e^{j2\pi \left(\frac{(m-M)q}{\Delta M} - \frac{(l-L)p}{\Delta L} \right)} \right] e^{-j\frac{2\pi}{N}[k(m-M) - n(l-L)]}. \tag{28}$$

B. Measurement Equation and Sparse Reconstruction

The compressive RS estimator can be obtained by combining the results of the previous subsection with standard results from CS theory [44], [61]. To motivate our development, we assume that $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ is approximately K-sparse for some $K < S'$, i.e., at most K of the S' values of the basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$ on the subsampled grid $(n,k) = (p\Delta n, q\Delta k)$ are significantly nonzero. (Because $\hat{R}_{X,\mathrm{MVU}}[n,k]$ is an estimator of the RS, this assumption is consistent with our basic assumption that the RS $\bar{R}_X[n,k]$ itself is approximately sparse.) Due to (24), it follows that the matrix $\mathbf{R}$ and, equivalently, the vector $\mathbf{r} \equiv \mathrm{vec}\{\mathbf{R}\}$ are approximately K-sparse. Furthermore, according to (26), $\mathbf{r} \in \mathbb{C}^{S'}$ is related to the EAF estimate $\mathbf{a} \equiv \mathrm{vec}\{\mathbf{A}^T\} \in \mathbb{C}^{S'}$ as $\mathbf{U}\mathbf{r} = \mathbf{a}$, where $\mathbf{U}$ (see (27)) is an orthogonal (up to a factor) and equimodular matrix of size $S' \times S'$, i.e., $\mathbf{U}^H \mathbf{U} = \frac{1}{S'}\mathbf{I}$ and $|(\mathbf{U})_{i,j}| = \frac{1}{S'}$. Let us define $\mathbf{a}^{(P)} \in \mathbb{C}^P$ as the vector made up of P randomly selected entries of $\mathbf{a}$, for some $P < S'$ (typically, $P \ll S'$). Thus, recalling (21) and (14), the entries of $\mathbf{a}^{(P)}$ are P values of the masked AF $I_{\mathcal{A}}[m,l]A_X[m,l]$ randomly located within the region $[N]^2 \cap \mathcal{A}'$ or, equivalently,[3] the values of $I_{\mathcal{A}}[m,l]A_X[m,l]$ at P randomly chosen TF lag positions $(m,l) \in \{-M, \ldots, -M+\Delta M-1\} \times \{-L, \ldots, -L+\Delta L-1\}$. We have then from (26)

$$\mathbf{M}\mathbf{r} = \mathbf{a}^{(P)}, \tag{29}$$

where the matrix $\mathbf{M} \in \mathbb{C}^{P \times S'}$ is obtained by randomly selecting P rows from $\mathbf{U}$; the indices of these rows equal the indices of the entries selected from $\mathbf{a}$.

Equation (29) is an instance of a *measurement equation* as considered in CS theory. Because the "measurement matrix" $\mathbf{M}$ is formed by randomly selecting P rows from $\mathbf{U}$, and $\mathbf{U}$ is a unitary (up to a factor) and equimodular matrix, CS theory [44], [61] provides the following result: For

$$P \ge C\, (\log S')^4 K = C\, \big[\log(\Delta M) + \log(\Delta L) \big]^4 K, \tag{30}$$

---

[3]Typically, the region $[N]^2 \cap \mathcal{A}'$ is only slightly larger than the effective EAF support $[N]^2 \cap \mathcal{A}$. Thus, most of the P entries of $\mathbf{a}^{(P)}$ are values of $A_X[m,l]$ randomly located within $[N]^2 \cap \mathcal{A}$ or, equivalently, within $\{-M, \ldots, M\} \times \{-L, \ldots, L\}$. The remaining entries of $\mathbf{a}^{(P)}$ are zero.

where C is a positive constant that does not depend on $\mathbf{r}$, the result of Basis Pursuit [62] operating on $\mathbf{a}^{(P)}$, i.e.,

$$\hat{\mathbf{r}} \triangleq \underset{\mathbf{r}':\mathbf{M}\mathbf{r}'=\mathbf{a}^{(P)}}{\operatorname{argmin}} \|\mathbf{r}'\|_1, \qquad (31)$$

satisfies with overwhelming probability[4]

$$\|\hat{\mathbf{r}} - \mathbf{r}\|_2 \leq \frac{D}{\sqrt{K}} \|\mathbf{r} - \mathbf{r}^{\mathcal{G}}\|_1. \qquad (32)$$

Here, D is another positive constant that does not depend on $\mathbf{r}$, and $\mathbf{r}^{\mathcal{G}}$ denotes the vector that is obtained by zeroing all entries of $\mathbf{r}$ except the K entries whose indices are in a given index set $\mathcal{G} \subseteq \{1, \ldots, S'\}$ of size $|\mathcal{G}| = K$. Since $\mathbf{r}$ is approximately K-sparse, the index set $\mathcal{G}$ can be chosen such that the corresponding entries $\{(\mathbf{r})_k\}_{k \in \mathcal{G}}$ comprise, with high probability,[5] the significantly nonzero entries of $\mathbf{r}$, implying a small norm $\|\mathbf{r} - \mathbf{r}^{\mathcal{G}}\|_1$. The bound (32) then shows that the Basis Pursuit is capable of reconstructing $\mathbf{r}$—and, thus, the subsampled basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$—from the compressed AF vector $\mathbf{a}^{(P)}$ with a small reconstruction error $\|\hat{\mathbf{r}} - \mathbf{r}\|_2$. (We recall, at this point, that the entries of $\mathbf{r}$ equal the values of $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$.) The minimization in (31) can be implemented numerically using standard tools, e.g., the MATLAB toolbox CVX [63].

C. The Compressive RS Estimator

From the Basis Pursuit reconstruction result $\hat{\mathbf{r}}$ in (31), a compressive approximation of the basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[n, k]$ in (17) is finally obtained by substituting $\hat{\mathbf{r}}$ for $\mathbf{r}$ in (28):

$$\hat{R}_{X,\mathrm{CS}}[n, k] = \mathcal{L}\{\hat{\mathbf{r}}\}[n, k]$$

$$= \frac{1}{NS'} \sum_{m \in [\Delta M]} \sum_{l \in [\Delta L]} \left[\sum_{p \in [\Delta L]} \sum_{q \in [\Delta M]} (\hat{\mathbf{R}})_{p+1, q+1} \right.$$

$$\left. \times e^{j2\pi \left(\frac{(m-M)q}{\Delta M} - \frac{(l-L)p}{\Delta L} \right)} \right] e^{-j\frac{2\pi}{N}[k(m-M)-n(l-L)]}, \quad (33)$$

where $\hat{\mathbf{R}} = \mathrm{unvec}\{\hat{\mathbf{r}}\} \in \mathbb{C}^{\Delta L \times \Delta M}$ is the matrix corresponding to $\hat{\mathbf{r}}$. This defines the compressive RS estimator.

To summarize, the proposed compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n, k]$ is calculated by the following steps.

1) Choose $K < S'$ such that it reflects the prior intuition about the effective sparsity of the subsampled RS $\bar{R}_X[p\Delta n, q\Delta k]$, $(p, q) \in [\Delta L] \times [\Delta M]$. (Equivalently, KN^2/S' reflects the prior intuition about the effective sparsity of the RS $\bar{R}_X[n, k]$, $(n, k) \in [N]^2$.)

2) Acquire $P \geq C \left[\log(\Delta M) + \log(\Delta L) \right]^4 K$ values of the masked AF $I_{\mathcal{A}}[m, l] A_X[m, l]$ at randomly chosen

TF lag positions[6] $(m, l) \in \{-M, \ldots, -M + \Delta M - 1\} \times \{-L, \ldots, -L + \Delta L - 1\}$. Let $\mathbf{a}^{(P)}$ denote the vector containing these "compressive measurements." A compression has been achieved if $P < S' \equiv \Delta M \Delta L$; the "compression factor" is $S'/P \geq 1$. It is important to note that the AF values $A_X[m, l]$ can be equivalently obtained (up to small aliasing errors that are typically negligible) from the continuous-TF-lag AF of the underlying continuous-time process $X(t)$.[7]

3) Form the "measurement matrix" $\mathbf{M} \in \mathbb{C}^{P \times S'}$ comprising those rows of $\mathbf{U} \in \mathbb{C}^{S' \times S'}$ (see (27)) whose indices correspond to the TF lag positions (m, l) chosen in Step 2.

4) Compute an estimate $\hat{\mathbf{r}}$ of $\mathbf{r}$ from $\mathbf{a}^{(P)}$ by means of the Basis Pursuit (31), i.e., $\hat{\mathbf{r}} = \operatorname{argmin}_{\mathbf{r}':\mathbf{M}\mathbf{r}'=\mathbf{a}^{(P)}} \|\mathbf{r}'\|_1$.

5) From $\hat{\mathbf{r}}$, calculate $\hat{R}_{X,\mathrm{CS}}[n, k] = \mathcal{L}\{\hat{\mathbf{r}}\}[n, k]$ according to (33). This step can be implemented efficiently by two successive 2D FFT operations.

Based on the error bound (32) (with the index set $\mathcal{G}$ chosen as described below (32)), the compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n, k]$ can be expected to be close to the noncompressive basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[n, k]$ in (17) if the subsampled RS estimate $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ is approximately K-sparse. In Section V, we will derive an upper bound on the approximation error (MSE) that is formulated in terms of certain parameters depending on second-order statistics of the process $X[n]$, including the RS, $\bar{R}_X[n, k]$.

As previously mentioned in Section I, from an algorithmic viewpoint, our compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n, k]$ is similar to the compressive TF representation proposed in [45], [46]. However, the setting of [45], [46] is that of deterministic TF signal analysis (improving the TF localization of the Wigner distribution), rather than spectral estimation for nonstationary random processes.

D. An Improved Compressive RS Estimator

The compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n, k]$ in (33) is related to the compressive EAF estimator $\hat{A}_{X,\mathrm{CS}}[m, l]$ defined

---

[4]That is, the probability of (32) not being true decreases exponentially with P.

[5]Note that the index set $\mathcal{G}$ is deterministic and fixed, whereas the indices of the largest entries of $\mathbf{r}$ may vary with each realization of the random process. However, for the performance analysis in Section V, it is sufficient to assume that the index set $\mathcal{G}$ *approximately* contains the indices of the largest entries of $\mathbf{r}$ for each realization.

[6]More precisely, we choose uniformly at random a size-P subset of $\{-M, \ldots, -M + \Delta M - 1\} \times \{-L, \ldots, -L + \Delta L - 1\}$, containing P different TF lag positions (m, l).

[7]The continuous-TF-lag AF is defined as $A_X(\tau, \nu) \triangleq \int_{-\infty}^{\infty} X(t) X^*(t - \tau) e^{-j2\pi\nu t} dt$. If the process $X(t)$ is bandlimited to the frequency band $[0, 1/(2T_s)]$ and effectively localized within the time interval $[0, NT_s/2]$, we can use the approximation

$$A_X[m, l] \overset{(3)}{=} \sum_{n \in [N]} X[n] X^*[n - m]_N e^{-j\frac{2\pi}{N}ln}$$

$$\approx \sum_{n \in [N]} X(nT_s) X^*((n - m)T_s) e^{-j\frac{2\pi}{N}ln}$$

$$\approx \frac{1}{T_s} A_X\left(mT_s, \frac{l}{NT_s}\right), \quad \text{for } m, l \in \lfloor\lfloor N/2 \rfloor\rfloor.$$

Here, $X[n]$ is obtained from the continuous-time process $X(t)$ by regular sampling with period T_s, i.e., $X[n] = X(nT_s)$ for $n \in [N]$. Thus, $A_X[m, l]$ can be approximately calculated from the AF $A_X(\tau, \nu)$ of the continuous-time process $X(t)$.

as

$$
\hat{A}_{X,\mathrm{CS}}[m,l] \triangleq
\begin{cases}
\dfrac{1}{S'} \displaystyle\sum_{p\in[\Delta L]} \sum_{q\in[\Delta M]} (\hat{\mathbf{R}})_{p+1,q+1}\, e^{j2\pi\left(\frac{mq}{\Delta M}-\frac{lp}{\Delta L}\right)}, \\
\quad (m,l)\in\{-M,\ldots,-M+\Delta M-1\}_N \\
\qquad\qquad \times\{-L,\ldots,-L+\Delta L-1\}_N \\
0\,, \quad \text{otherwise.}
\end{cases}
\tag{34}
$$

This relation is given by the 2D DFT

$$
\hat{R}_{X,\mathrm{CS}}[n,k] = \frac{1}{N} \sum_{m=-M}^{-M+\Delta M-1} \sum_{l=-L}^{-L+\Delta L-1} \hat{A}_{X,\mathrm{CS}}[m,l] \\
\times e^{-j\frac{2\pi}{N}(km-nl)}. \tag{35}
$$

Now, although the AF and EAF satisfy the following symmetry property:

$$
A_X^*[-m,-l]\, e^{-j\frac{2\pi}{N}ml} = A_X[m,l] \tag{36a}
$$

$$
\bar{A}_X^*[-m,-l]\, e^{-j\frac{2\pi}{N}ml} = \bar{A}_X[m,l], \tag{36b}
$$

the EAF estimator $\hat{A}_{X,\mathrm{CS}}[m,l]$ does not exhibit this symmetry property in general. This fact suggests the following simple symmetrization modification (postprocessing) of the EAF estimator:

$$
\hat{A}_{X,\mathrm{CS}}^{(\mathrm{s})}[m,l] \triangleq \frac{1}{2}\left[\hat{A}_{X,\mathrm{CS}}[m,l] + \hat{A}_{X,\mathrm{CS}}^*[-m,-l]\, e^{-j\frac{2\pi}{N}ml}\right]. \tag{37}
$$

This, in turn, naturally leads to the definition of a "symmetrized" RS estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ via the 2D DFT transform in (35), i.e.,

$$
\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k] \triangleq \frac{1}{N} \sum_{m=-M}^{-M+\Delta M-1} \sum_{l=-L}^{-L+\Delta L-1} \hat{A}_{X,\mathrm{CS}}^{(\mathrm{s})}[m,l] \\
\times e^{-j\frac{2\pi}{N}(km-nl)}.
$$

The following explicit expression of the symmetrized RS estimator is easily shown:

$$
\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]
$$

$$
= \frac{1}{2NS'} \sum_{m\in[\Delta M]} \sum_{l\in[\Delta L]} \Bigg[\sum_{p\in[\Delta L]} \sum_{q\in[\Delta M]} \Big[(\hat{\mathbf{R}})_{p+1,q+1} \\
+ (\hat{\mathbf{R}})_{p+1,q+1}^* \, e^{-j\frac{2\pi}{N}(m-M)(l-L)} \Big] e^{j2\pi\left(\frac{(m-M)q}{\Delta M}-\frac{(l-L)p}{\Delta L}\right)} \Bigg] \\
\times e^{-j\frac{2\pi}{N}[k(m-M)-n(l-L)]}. \tag{38}
$$

This expression replaces (33). In Appendix A, we show that the MSE of the symmetrized RS estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ is always smaller than (or equal to) that of the original RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$, i.e.,

$$
\mathrm{E}\big\{\|\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})} - \bar{R}_X\|_2^2\big\} \leq \mathrm{E}\big\{\|\hat{R}_{X,\mathrm{CS}} - \bar{R}_X\|_2^2\big\}. \tag{39}
$$

Thus, the upper bound on the MSE of $\hat{R}_{X,\mathrm{CS}}[n,k]$ to be derived in Section V also applies to the MSE of $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$. To

summarize, by using instead of the compressive RS estimator in (33) the symmetrized compressive RS estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ given by (38), we can typically reduce the MSE.

Finally, we mention that in the case where no compression is performed, i.e., $S'/P = 1$, the basic (noncompressive) estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$, the compressive estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$, and the symmetrized compressive estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ all coincide, i.e., $\hat{R}_{X,\mathrm{MVU}}[n,k] \equiv \hat{R}_{X,\mathrm{CS}}[n,k] \equiv \hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$. The equivalence $\hat{R}_{X,\mathrm{MVU}}[n,k] \equiv \hat{R}_{X,\mathrm{CS}}[n,k]$ can be verified by observing that for $S'/P = 1$, the measurement matrix $\mathbf{M}$ in (29) coincides with the invertible matrix $\mathbf{U}$ in (26). Therefore, the vectors $\mathbf{r} = \mathrm{vec}\{\mathbf{R}\}$ in (26) and $\hat{\mathbf{r}} = \mathrm{vec}\{\hat{\mathbf{R}}\}$ in (31) coincide, and so do the corresponding RS estimators $\hat{R}_{X,\mathrm{MVU}}[n,k]$ and $\hat{R}_{X,\mathrm{CS}}[n,k]$ (cf. (28) and (33)). To verify that $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k] \equiv \hat{R}_{X,\mathrm{MVU}}[n,k]$ for $S'/P = 1$, note that because of (25) and (34), $\hat{R}_{X,\mathrm{CS}}[n,k] \equiv \hat{R}_{X,\mathrm{MVU}}[n,k]$ is equivalent to $\hat{A}_{X,\mathrm{CS}}[m,l] \equiv \hat{A}_{X,\mathrm{MVU}}[m,l]$. Since $\mathcal{A} = \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$ is symmetric, it follows from expression (15) that the basic EAF estimator $\hat{A}_{X,\mathrm{MVU}}[m,l]$ satisfies the symmetry relation (36), and hence $\hat{A}_{X,\mathrm{MVU}}^{(\mathrm{s})}[m,l] \triangleq \frac{1}{2}\big[\hat{A}_{X,\mathrm{MVU}}[m,l] + \hat{A}_{X,\mathrm{MVU}}^*[-m,-l]\, e^{-j\frac{2\pi}{N}ml}\big] = \hat{A}_{X,\mathrm{MVU}}[m,l]$. Thus, for $S'/P = 1$, we have $\hat{A}_{X,\mathrm{CS}}^{(\mathrm{s})}[m,l] = \hat{A}_{X,\mathrm{MVU}}^{(\mathrm{s})}[m,l] = \hat{A}_{X,\mathrm{MVU}}[m,l]$, and in turn $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k] = \hat{R}_{X,\mathrm{MVU}}[n,k]$.

V. MSE BOUNDS

In this section, we derive an upper bound on the MSE of the proposed compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$,

$$
\varepsilon_{\mathrm{CS}} \triangleq \mathrm{E}\big\{\|\hat{R}_{X,\mathrm{CS}} - \bar{R}_X\|_2^2\big\} \\
= \sum_{n,k\in[N]} \mathrm{E}\big\{|\hat{R}_{X,\mathrm{CS}}[n,k] - \bar{R}_X[n,k]|^2\big\},
$$

under the assumption that $X[n]$ is a circularly symmetric complex Gaussian nonstationary process. We do *not* assume that the EAF $\bar{A}_X[m,l]$ is exactly supported on some periodic lag rectangle $\mathcal{A} = \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$ with $0 \leq M < \lfloor N/2 \rfloor$ and $0 \leq L < \lfloor N/2 \rfloor$.

A. Parameters

Our MSE bound depends on three parameters of the second-order statistics of the process $X[n]$, which will be defined first.

1) As a measure (in the broad sense) of the sparsity of $\bar{R}_X[n,k]$, we define the *TF sparsity moment*

$$
\sigma_X^{(w)} \triangleq \frac{1}{\|\bar{R}_X\|_2^2} \left[\sum_{n,k\in[N]} w[n,k]\, |\bar{R}_X[n,k]| \right]^2, \tag{40}
$$

where $w[n,k] \geq 0$ is a suitably chosen weighting function and $\|\bar{R}_X\|_2^2 \triangleq \sum_{n,k\in[N]} |\bar{R}_X[n,k]|^2$ (i.e., the norm is taken over one period of $\bar{R}_X[n,k]$). In particular, for $w[n,k] \equiv 1$, $\sigma_X^{(w)} = \|\bar{R}_X\|_1^2/\|\bar{R}_X\|_2^2$.

2) For another way to measure the TF sparsity, let us first denote by

$$\widetilde{R}_{X,\text{MVU}}[n,k] \triangleq \text{E}\big\{\hat{R}_{X,\text{MVU}}[n,k]\big\} \qquad (41)$$

the expectation of the basic RS estimator $\hat{R}_{X,\text{MVU}}[n,k]$ in (17). It follows from (8) that $\widetilde{R}_{X,\text{MVU}}[n,k]$ is a smoothed version of the RS, i.e.,

$$\widetilde{R}_{X,\text{MVU}}[n,k]$$
$$= \frac{1}{N} \sum_{n',k'\in[N]} \Phi_{\text{MVU}}[n-n',k-k']\,\text{E}\big\{R_X[n',k']\big\}$$
$$= \frac{1}{N} \sum_{n',k'\in[N]} \Phi_{\text{MVU}}[n-n',k-k']\,\bar{R}_X[n',k'], \quad (42)$$

where $\text{E}\big\{R_X[n,k]\big\} = \bar{R}_X[n,k]$ has been used in the last step. Due to (11), the smoothing kernel is given by

$$\Phi_{\text{MVU}}[n,k] \triangleq \frac{1}{N} \sum_{m,l\in[N]} \phi_{\text{MVU}}[m,l]\,e^{-j\frac{2\pi}{N}(km-nl)}$$
$$\overset{(13)}{=} \frac{1}{N} \sum_{m,l\in[N]} I_{\mathcal{A}}[m,l]\,e^{-j\frac{2\pi}{N}(km-nl)} \quad (43)$$
$$= \frac{1}{N} \sum_{m=-M}^{M} \sum_{m=-L}^{L} e^{-j\frac{2\pi}{N}(km-nl)}.$$

Because of the smoothing, the number of significantly nonzero values of $\widetilde{R}_{X,\text{MVU}}[n,k]$ may be larger than the number of significantly nonzero values of the RS $\bar{R}_X[n,k]$. However, for an underspread process, the RS is inherently smooth, which implies that the smoothed RS is close to the RS. Therefore, for an underspread process with a small number of significantly nonzero RS values, we can expect that also the smoothed RS consists of only a small number of significantly nonzero values. Let us denote by $\mathcal{G}(K)$ the set of indices $(p,q) \in [\Delta L] \times [\Delta M]$ of the K largest (in magnitude) values of the subsampled expected RS estimator, $\widetilde{R}_{X,\text{MVU}}[p\Delta n,q\Delta k]$. Let $\overline{\mathcal{G}(K)} \triangleq ([\Delta L] \times [\Delta M]) \setminus \mathcal{G}(K)$, and note that $\big|\overline{\mathcal{G}(K)}\big| = S'-K$. We then define the *TF sparsity profile*[8]

$$\tilde{\sigma}_X(K) \triangleq \frac{1}{\|\bar{R}_X\|_2^2} \sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q}, \qquad (44)$$

with

$$h_{p,q} \triangleq \text{E}\big\{\big|(\mathbf{R})_{p+1,q+1}\big|^2\big\}$$
$$\overset{(24)}{=} N^2\,\text{E}\big\{\big|\hat{R}_{X,\text{MVU}}[p\Delta n,q\Delta k]\big|^2\big\}. \quad (45)$$

For later use, we note that

$$\sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q} = \text{E}\big\{\big\|\mathbf{r}^{\overline{\mathcal{G}(K)}}\big\|_2^2\big\} = \text{E}\big\{\big\|\mathbf{r}-\mathbf{r}^{\mathcal{G}(K)}\big\|_2^2\big\},$$
$$(46)$$

[8]We note that this definition is different from that in [64].

where $\mathbf{r}^{\overline{\mathcal{G}(K)}}$ (resp. $\mathbf{r}^{\mathcal{G}(K)}$) denotes the vector that is obtained from $\mathbf{r} \equiv \text{vec}\{\mathbf{R}\}$ by zeroing all entries except the $S'-K$ (resp. K) entries whose indices correspond to the indices[9] $(p,q) \in \overline{\mathcal{G}(K)}$ (resp. $(p,q) \in \mathcal{G}(K)$).

3) The "TF correlation width" of $X[n]$ can be measured by the *EAF moment* [10], [24]

$$m_X^{(\psi)} \triangleq \frac{1}{\|\bar{A}_X\|_2^2} \sum_{m,l\in[N]} \psi[m,l]\,\big|\bar{A}_X[m,l]\big|^2, \quad (47)$$

where $\psi[m,l]$ is some weighting function that is generally zero or small at the origin $(0,0)$ and increases with increasing $|m|$ and $|l|$, and $\|\bar{A}_X\|_2^2 \triangleq \sum_{m,l\in[N]} \big|\bar{A}_X[m,l]\big|^2 = \|\bar{R}_X\|_2^2$. For an underspread process $X[n]$ and a reasonable choice of $\psi[m,l]$, $m_X^{(\psi)}$ is small ($\ll 1$).

B. Bound on the MSE of the Basic RS Estimator

Our bound on the MSE $\varepsilon_{\text{CS}} = \text{E}\big\{\big\|\hat{R}_{X,\text{CS}} - \bar{R}_X\big\|_2^2\big\}$ is a combination of a bound on the MSE of the basic (non-compressive) RS estimator $\hat{R}_{X,\text{MVU}}[n,k]$ and a bound on the excess MSE introduced by the compression. First, we derive the bound on the MSE of the basic RS estimator,

$$\varepsilon \triangleq \text{E}\big\{\big\|\hat{R}_{X,\text{MVU}} - \bar{R}_X\big\|_2^2\big\}.$$

As in Section III, we use the decomposition

$$\varepsilon = B^2 + V, \qquad (48)$$

with the squared bias term $B^2 = \big\|\text{E}\{\hat{R}_{X,\text{MVU}}\} - \bar{R}_X\big\|_2^2$ and the variance $V = \text{E}\big\{\big\|\hat{R}_{X,\text{MVU}} - \text{E}\{\hat{R}_{X,\text{MVU}}\}\big\|_2^2\big\}$.

1) Bias: An expression of the bias term is obtained by setting $\phi[m,l] = \phi_{\text{MVU}}[m,l] = I_{\mathcal{A}}[m,l]$ in (12):

$$B^2 = \sum_{m,l\in[N]} \big|(I_{\mathcal{A}}[m,l]-1)\,\bar{A}_X[m,l]\big|^2$$
$$= \sum_{m,l\in[N]} I_{\overline{\mathcal{A}}}[m,l]\,\big|\bar{A}_X[m,l]\big|^2,$$

where $I_{\overline{\mathcal{A}}}[m,l] = 1 - I_{\mathcal{A}}[m,l]$ is the indicator function of the complement $\overline{\mathcal{A}}$ of the effective EAF support region $\mathcal{A} = \{-M,\dots,M\}_N \times \{-L,\dots,L\}_N$, i.e.,

$$I_{\overline{\mathcal{A}}}[m,l] = \begin{cases} 1, & (m,l) \notin \mathcal{A} \\ 0, & \text{otherwise.} \end{cases}$$

We can write B^2 in terms of the EAF moment (47) with weighting function $\psi[m,l] = I_{\overline{\mathcal{A}}}[m,l]$:

$$B^2 = \|\bar{A}_X\|_2^2\,m_X^{(I_{\overline{\mathcal{A}}})} = \|\bar{R}_X\|_2^2\,m_X^{(I_{\overline{\mathcal{A}}})}. \qquad (49)$$

[9]For convenience, though with an abuse of notation, we denote by $\mathcal{G}(K)$ both a set of indices k of $(\mathbf{r})_k$ and the corresponding set of 2D indices (p,q) of $(\mathbf{R})_{p+1,q+1} = (\text{unvec}\{\mathbf{r}\})_{p+1,q+1}$ or equivalently of $\hat{R}_{X,\text{MVU}}[p\Delta n,q\Delta k]$. Thus, depending on the context, we will write $k \in \mathcal{G}(K)$ or $(p,q) \in \mathcal{G}(K)$.

Note that $m_X^{(I_{\overline{\mathcal{A}}})} = 0$, and thus $B^2 = 0$, if and only if the EAF $\bar{A}_X[m,l]$ is exactly supported on $\mathcal{A}$.

2) Variance: In what follows, we will use the (scaled) discrete TF shift matrices $\mathbf{J}_{m,l}$ of size $N \times N$ whose action on $\mathbf{x} \in \mathbb{C}^N$ is given by

$$(\mathbf{J}_{m,l}\mathbf{x})_{n+1} = \frac{1}{\sqrt{N}} (\mathbf{x})_{(n-m)_N+1} e^{j\frac{2\pi}{N}ln}, \quad n \in [N],$$

with $(n)_N \triangleq n \bmod N$. Basic properties of the family of TF shift matrices $\{\mathbf{J}_{m,l}\}_{m,l \in [N]}$ are considered in Appendix B. Using $\mathbf{J}_{m,l}$, $\hat{R}_{X,\mathrm{MVU}}[n,k]$ can be written as a quadratic form in $\mathbf{x} = (X[0] \cdots X[N-1])^T$. In fact, starting from (17) and using (85), we can develop $\hat{R}_{X,\mathrm{MVU}}[n,k]$ as follows:

$$\hat{R}_{X,\mathrm{MVU}}[n,k] \overset{(17)}{=} \frac{1}{N} \sum_{m=-M}^{M} \sum_{l=-L}^{L} A_X[m,l] e^{-j\frac{2\pi}{N}(km-nl)}$$

$$\overset{(85)}{=} \frac{1}{\sqrt{N}} \sum_{m=-M}^{M} \sum_{l=-L}^{L} \langle \mathbf{x}\mathbf{x}^H, \mathbf{J}_{m,l} \rangle e^{-j\frac{2\pi}{N}(km-nl)}$$

$$= \left\langle \mathbf{x}\mathbf{x}^H, \frac{1}{\sqrt{N}} \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{m,l} \right\rangle.$$

Setting

$$\mathbf{C}_{n,k} \triangleq \frac{1}{\sqrt{N}} \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{m,l}, \qquad (50)$$

this becomes

$$\hat{R}_{X,\mathrm{MVU}}[n,k] = \langle \mathbf{x}\mathbf{x}^H, \mathbf{C}_{n,k} \rangle$$
$$= \mathrm{tr}\{\mathbf{x}\mathbf{x}^H \mathbf{C}_{n,k}^H\}$$
$$= \mathbf{x}^H \mathbf{C}_{n,k}^H \mathbf{x}. \qquad (51)$$

Note that the matrix $\mathbf{C}_{n,k}$ is not Hermitian in general.

Splitting $\hat{R}_{X,\mathrm{MVU}}[n,k]$ into its real and imaginary parts, we have

$$\mathrm{var}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\} = \mathrm{var}\{\Re\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}\}$$
$$+ \mathrm{var}\{\Im\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}\}. \quad (52)$$

It is easily shown that

$$\Re\{\hat{R}_{X,\mathrm{MVU}}[n,k]\} = \mathbf{x}^H \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{x}, \qquad (53)$$

$$\Im\{\hat{R}_{X,\mathrm{MVU}}[n,k]\} = \mathbf{x}^H \mathbf{C}_{n,k}^{(\mathrm{I})} \mathbf{x}, \qquad (54)$$

with the Hermitian matrices

$$\mathbf{C}_{n,k}^{(\mathrm{R})} \triangleq \frac{1}{2}\left(\mathbf{C}_{n,k}^H + \mathbf{C}_{n,k}\right), \quad \mathbf{C}_{n,k}^{(\mathrm{I})} \triangleq \frac{1}{2j}\left(\mathbf{C}_{n,k}^H - \mathbf{C}_{n,k}\right). \qquad (55)$$

Inserting (53) and (54) into (52) and using a standard result for the variance of a Hermitian form of a circularly symmetric complex Gaussian random vector [65], we obtain

$$\mathrm{var}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\} = \mathrm{tr}\{\mathbf{C}_{n,k}^{(\mathrm{R})}\mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})}\mathbf{\Gamma}_X\}$$
$$+ \mathrm{tr}\{\mathbf{C}_{n,k}^{(\mathrm{I})}\mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{I})}\mathbf{\Gamma}_X\}, \quad (56)$$

with $\mathbf{\Gamma}_X \triangleq \mathrm{E}\{\mathbf{x}\mathbf{x}^H\}$.

Using this expression, we next derive an upper bound on $V = \mathrm{E}\{\|\hat{R}_{X,\mathrm{MVU}} - \mathrm{E}\{\hat{R}_{X,\mathrm{MVU}}\}\|_2^2\}$. We have

$$V = \sum_{n,k \in [N]} \mathrm{E}\{|\hat{R}_{X,\mathrm{MVU}}[n,k] - \mathrm{E}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}|^2\}$$

$$= \sum_{n,k \in [N]} \mathrm{var}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}$$

$$\overset{(56)}{=} \sum_{n,k \in [N]} \mathrm{tr}\{\mathbf{C}_{n,k}^{(\mathrm{R})}\mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})}\mathbf{\Gamma}_X\}$$

$$+ \sum_{n,k \in [N]} \mathrm{tr}\{\mathbf{C}_{n,k}^{(\mathrm{I})}\mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{I})}\mathbf{\Gamma}_X\}. \qquad (57)$$

It is then shown in Appendix C that

$$V = \sum_{m,l \in [N]} |\bar{A}_X[m,l]|^2 \chi[m,l], \qquad (58)$$

with

$$\chi[m,l] = \frac{1}{N} \sum_{m',l' \in [N]} I_{\mathcal{A}}[m',l'] e^{j\frac{2\pi}{N}(lm'-ml')} \qquad (59)$$

$$= \frac{1}{N} \sum_{m'=-M}^{M} \sum_{l'=-L}^{L} e^{j\frac{2\pi}{N}(lm'-ml')}. \qquad (60)$$

We can bound the magnitude of $\chi[m,l]$ according to

$$|\chi[m,l]| \leq \frac{1}{N} \sum_{m'=-M}^{M} \sum_{l'=-L}^{L} |e^{j\frac{2\pi}{N}(lm'-ml')}|$$

$$= \frac{1}{N}(2M+1)(2L+1)$$

$$= \frac{S}{N}.$$

Combining with (58) leads to the following bound on V:

$$V \leq \sum_{m,l \in [N]} |\bar{A}_X[m,l]|^2 |\chi[m,l]|$$

$$\leq \frac{S}{N} \sum_{m,l \in [N]} |\bar{A}_X[m,l]|^2$$

$$\overset{(6)}{=} \frac{S}{N} \|\bar{R}_X\|_2^2. \qquad (61)$$

3) MSE: Finally, the desired bound on the MSE $\varepsilon = \mathrm{E}\{\|\hat{R}_{X,\mathrm{MVU}} - \bar{R}_X\|_2^2\}$ is obtained by inserting (49) and (61) into the expansion (48):

$$\varepsilon = B^2 + V$$

$$\leq \|\bar{R}_X\|_2^2 \, m_X^{(I_{\overline{\mathcal{A}}})} + \frac{S}{N} \|\bar{R}_X\|_2^2$$

$$= \|\bar{R}_X\|_2^2 \left(m_X^{(I_{\overline{\mathcal{A}}})} + \frac{S}{N}\right). \qquad (62)$$

This bound is small if $X[n]$ is underspread, i.e., if $m_X^{(I_{\overline{\mathcal{A}}})} \ll 1$ and $S \ll N$.

C. Bound on the Excess MSE Due to Compression

The excess MSE caused by the compression is given by

$$\Delta\varepsilon \triangleq \mathrm{E}\big\{\big\|\hat{R}_{X,\mathrm{CS}} - \hat{R}_{X,\mathrm{MVU}}\big\|_2^2\big\}.$$

Because of the Fourier transform relations (28) and (33), we have

$$\Delta\varepsilon = \frac{1}{S'}\,\mathrm{E}\big\{\|\hat{\mathbf{r}} - \mathbf{r}\|_2^2\big\}. \tag{63}$$

As in Section IV-B, let K denote a nominal sparsity degree that is chosen according to our intuition about the approximate sparsity of $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ and, equivalently, $\mathbf{r}$. We assume that the number P of randomly selected AF samples is sufficiently large so that (32) is satisfied, i.e.,

$$\|\hat{\mathbf{r}} - \mathbf{r}\|_2^2 \le \frac{D^2}{K}\,\|\mathbf{r} - \mathbf{r}^{\mathcal{G}}\|_1^2, \tag{64}$$

for any index set $\mathcal{G}$ of size $|\mathcal{G}| = K$. (A sufficient condition is (30).) An intuitively reasonable choice of K and $\mathcal{G}$ can be based on the smoothed RS $\widetilde{R}_{X,\mathrm{MVU}}[n,k] = \mathrm{E}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}$ in (41), (42): we choose K as the number of significantly nonzero values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$, and $\mathcal{G} = \mathcal{G}(K)$ of size K as the set of those indices of $\mathbf{r}$ that correspond to these significant values—equivalently, to the K largest (in magnitude) values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$. Thus, $\mathbf{r}^{\mathcal{G}(K)}$ comprises those K values $\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ for which the corresponding values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ are largest (in magnitude).

Based on this choice, we will now derive an approximate upper bound on the excess MSE $\Delta\varepsilon$. Inserting (64) into (63), we obtain

$$\Delta\varepsilon \le \frac{D^2}{S'K}\,\mathrm{E}\big\{\big\|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\big\|_1^2\big\}. \tag{65}$$

Using the inequality[10] $\|\cdot\|_1^2 \le \|\cdot\|_0 \|\cdot\|_2^2$, we have $\|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\|_1^2 \le \|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\|_0 \|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\|_2^2 \le (S'-K)\|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\|_2^2$, and thus (65) becomes further

$$\Delta\varepsilon \le \frac{(S'-K)D^2}{S'K}\,\mathrm{E}\big\{\big\|\mathbf{r} - \mathbf{r}^{\mathcal{G}(K)}\big\|_2^2\big\}$$

$$\stackrel{(46)}{=} \frac{(S'-K)D^2}{S'K}\sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q} \tag{66}$$

$$\stackrel{(44)}{=} \frac{(S'-K)D^2}{S'K}\,\|\bar{R}_X\|_2^2\,\tilde{\sigma}_X(K). \tag{67}$$

In what follows, we will derive an approximate expression of $h_{p,q} = \mathrm{E}\big\{\big|(\mathbf{R})_{p+1,q+1}\big|^2\big\}$ in terms of $\bar{R}_X[n,k]$; this expression will show under which condition $\tilde{\sigma}_X(K) \propto \sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q}$ is small. We have

$$h_{p,q} = \mathrm{E}\big\{\big|(\mathbf{R})_{p+1,q+1}\big|^2\big\}$$

$$= \mathrm{var}\big\{(\mathbf{R})_{p+1,q+1}\big\} + \big|\mathrm{E}\big\{(\mathbf{R})_{p+1,q+1}\big\}\big|^2$$

$$= \mathrm{var}\big\{\Re\big\{(\mathbf{R})_{p+1,q+1}\big\}\big\} + \mathrm{var}\big\{\Im\big\{(\mathbf{R})_{p+1,q+1}\big\}\big\}$$

$$\qquad + \big|\mathrm{E}\big\{(\mathbf{R})_{p+1,q+1}\big\}\big|^2. \tag{68}$$

---

[10]Indeed, the ℓ_1-norm of an arbitrary vector $\mathbf{z}$ can be expressed as $\|\mathbf{z}\|_1 = \mathbf{z}^H\mathbf{a}(\mathbf{z})$, where $\mathbf{a}(\mathbf{z})$ is given elementwise by $(\mathbf{a}(\mathbf{z}))_k \triangleq z_k/|z_k|$ for $z_k \ne 0$ and $(\mathbf{a}(\mathbf{z}))_k \triangleq 0$ for $z_k = 0$. Clearly, $\|\mathbf{a}(\mathbf{z})\|_2^2 = \|\mathbf{z}\|_0$, and thus $\|\mathbf{z}\|_1^2 = (\mathbf{z}^H\mathbf{a}(\mathbf{z}))^2 \le \|\mathbf{z}\|_2^2\|\mathbf{a}(\mathbf{z})\|_2^2 = \|\mathbf{z}\|_2^2\|\mathbf{z}\|_0$, where the Cauchy-Schwarz inequality has been used.

Using (23) and (85), we can express $(\mathbf{R})_{p+1,q+1}$ as a quadratic form:

$$(\mathbf{R})_{p+1,q+1} \stackrel{(23)}{=} \sum_{m=-M}^{M}\sum_{l=-L}^{L} A_X[m,l]\,e^{-j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)}$$

$$\stackrel{(85)}{=} \sqrt{N}\sum_{m=-M}^{M}\sum_{l=-L}^{L}\langle\mathbf{x}\mathbf{x}^H, \mathbf{J}_{m,l}\rangle\,e^{-j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)}$$

$$= \langle\mathbf{x}\mathbf{x}^H, \mathbf{T}_{p,q}\rangle$$

$$= \mathrm{tr}\big\{\mathbf{x}\mathbf{x}^H\mathbf{T}_{p,q}^H\big\}$$

$$= \mathbf{x}^H\mathbf{T}_{p,q}^H\mathbf{x}, \tag{69}$$

with

$$\mathbf{T}_{p,q} \triangleq \sqrt{N}\sum_{m=-M}^{M}\sum_{l=-L}^{L} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)}\mathbf{J}_{m,l}. \tag{70}$$

Note that the matrix $\mathbf{T}_{p,q}$ is not Hermitian in general. Inserting (69) into (68) then yields

$$h_{p,q} = \mathrm{var}\big\{\mathbf{x}^H\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{x}\big\} + \mathrm{var}\big\{\mathbf{x}^H\mathbf{T}_{p,q}^{(\mathrm{I})}\mathbf{x}\big\} + \big|\mathrm{E}\big\{\mathbf{x}^H\mathbf{T}_{p,q}^H\mathbf{x}\big\}\big|^2,$$

with the Hermitian matrices

$$\mathbf{T}_{p,q}^{(\mathrm{R})} \triangleq \frac{1}{2}\big(\mathbf{T}_{p,q}^H + \mathbf{T}_{p,q}\big), \quad \mathbf{T}_{p,q}^{(\mathrm{I})} \triangleq \frac{1}{2j}\big(\mathbf{T}_{p,q}^H - \mathbf{T}_{p,q}\big). \tag{71}$$

Using standard results for the variance and mean of a Hermitian form of a circularly symmetric complex Gaussian vector [65], we obtain further

$$h_{p,q} = \mathrm{tr}\big\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\big\} + \mathrm{tr}\big\{\mathbf{T}_{p,q}^{(\mathrm{I})}\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{I})}\mathbf{\Gamma}_X\big\}$$

$$\qquad + \big|\mathrm{tr}\big\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^H\big\}\big|^2. \tag{72}$$

There does not seem to exist a simple closed-form expression of (72) in terms of the EAF $\bar{A}_X[m,l]$ or the RS $\bar{R}_X[n,k]$. However, under the assumption that the process $X[n]$ is underspread and the effective EAF support dimensions M, L (cf. (18)) are accordingly chosen to be small, the following approximation is derived in Appendix D:

$$h_{p,q} \approx N\sum_{n,k\in[N]}\big|\bar{R}_X[n,k]\,\Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k]\big|^2$$

$$\qquad + \bigg|\sum_{n,k\in[N]}\bar{R}_X[n,k]\,\Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k]\bigg|^2, \tag{73}$$

where, as before, $\Delta n = N/\Delta L$ and $\Delta k = N/\Delta M$. Comparing with (42) and noting that $\Phi_{\mathrm{MVU}}[-n,-k] = \Phi_{\mathrm{MVU}}[n,k]$, it is seen that the second term on the right hand side of (73) is $N^2\big|\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]\big|^2$. Using the inequality $\|\cdot\|_2^2 \le \|\cdot\|_1^2$ [58] to bound the first term on the right-hand side of (73), and using a trivial upper bound on the second term, we obtain

$$h_{p,q} \lessapprox N\bigg[\sum_{n,k\in[N]}\big|\bar{R}_X[n,k]\,\Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k]\big|\bigg]^2$$

$$\qquad + \bigg[\sum_{n,k\in[N]}\big|\bar{R}_X[n,k]\,\Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k]\big|\bigg]^2$$

$$= (N+1) \left[\sum_{n,k\in[N]} \left| \bar{R}_X[n,k]\, \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \right]^2. \tag{74}$$

Here, $\sum_{n,k\in[N]} \left| \bar{R}_X[n,k]\, \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right|$ can be interpreted as a local average of the RS modulus $\left| \bar{R}_X[n,k] \right|$ about the TF point $(p\Delta n, q\Delta k)$. Thus, the (approximate) upper bound (74) shows that $h_{p,q}$ is small if $\bar{R}_X[n,k]$ is small within a neighborhood of $(p\Delta n, q\Delta k)$ or, said differently, if $(p\Delta n, q\Delta k)$ is located outside a broadened version of the effective support of $\bar{R}_X[n,k]$. The broadening is stronger for a larger spread of $\Phi_{\mathrm{MVU}}[n,k]$. According to (43), $\Phi_{\mathrm{MVU}}[n,k]$ is the 2D DFT of the indicator function $I_\mathcal{A}[m,l]$, and thus the broadening depends on the size of the effective EAF support $\mathcal{A}$; it will be stronger if $\mathcal{A}$ is smaller, i.e., if the process $X[n]$ is more underspread. Since a stronger broadening implies a poorer sparsity, this demonstrates an intrinsic tradeoff between the underspreadness and the TF sparsity of $X[n]$: better underspreadness implies a smaller effective EAF support $\mathcal{A}$, whereas better TF sparsity requires a larger $\mathcal{A}$.

With this "broadening" interpretation in mind, we reconsider $\tilde{\sigma}_X(K) \propto \sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q}$ in the bound (67). Recall that $\mathcal{G}(K)$ was defined as the set of those indices of $\mathbf{r}$ such that the corresponding values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ are the K largest (in magnitude). Therefore, a small $\tilde{\sigma}_X(K)$ requires that K is chosen such that $K\Delta n\Delta k$ is approximately equal to the area of the broadened effective support of $\bar{R}_X[n,k]$, because then $\sum_{n,k\in[N]} \left| \bar{R}_X[n,k]\, \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \approx 0$ for $(p,q) \in \overline{\mathcal{G}(K)}$ and thus, using (74), $\tilde{\sigma}_X(K) \propto \sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q} \approx 0$.

Using (74), we can upper-bound the MSE bound in (66), $\Delta\varepsilon \leq \frac{(S'-K)D^2}{S'K} \sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q}$, which results in a simpler (but generally looser) upper bound. Indeed, we have

$$\sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q} \overset{(74)}{\lessapprox} (N+1) \sum_{(p,q)\in\overline{\mathcal{G}(K)}} \left[\sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right. \right.$$
$$\left. \left. \times\, \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \right]^2$$
$$\overset{(*)}{\leq} (N+1) \left[\sum_{(p,q)\in\overline{\mathcal{G}(K)}} \left| \sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right. \right. \right.$$
$$\left. \left. \left. \times\, \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \right| \right]^2$$
$$= (N+1) \left[\sum_{(p,q)\in\overline{\mathcal{G}(K)}} \sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right| \right.$$
$$\left. \times \left| \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \right]^2$$
$$= (N+1) \left[\sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right| \right.$$
$$\left. \times \sum_{(p,q)\in\overline{\mathcal{G}(K)}} \left| \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right| \right]^2$$

$$= (N+1) \left[\sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right| w_\Phi[n,k] \right]^2, \tag{75}$$

where $\| \cdot \|_2^2 \leq \| \cdot \|_1^2$ was used in the step labeled with $(*)$ and

$$w_\Phi[n,k] \triangleq \sum_{(p,q)\in\overline{\mathcal{G}(K)}} \left| \Phi_{\mathrm{MVU}}[n-p\Delta n, k-q\Delta k] \right|. \tag{76}$$

Comparing with the definition of the TF sparsity moment $\sigma_X^{(w)}$ in (40), it is seen that the approximate bound (75) can be written as

$$\sum_{(p,q)\in\overline{\mathcal{G}(K)}} h_{p,q} \lessapprox (N+1)\, \|\bar{R}_X\|_2^2\, \sigma_X^{(w_\Phi)}. \tag{77}$$

Inserting (77) into (66) then gives the approximate MSE bound

$$\Delta\varepsilon \lessapprox \frac{(S'-K)D^2}{S'K}\, (N+1)\, \|\bar{R}_X\|_2^2\, \sigma_X^{(w_\Phi)}. \tag{78}$$

A small excess MSE $\Delta\varepsilon$ can be achieved if the TF sparsity moment $\sigma_X^{(w_\Phi)} \propto \left[\sum_{n,k\in[N]} \left| \bar{R}_X[n,k] \right| w_\Phi[n,k] \right]^2$ is small. This, in turn, is the case if the RS $\bar{R}_X[n,k]$ is negligible within the effective support of the TF weighting function $w_\Phi[n,k]$. Due to (76), the size of the effective support of $w_\Phi[n,k]$, which is concentrated around the points $\{(p\Delta n, q\Delta k)\}_{(p,q)\in\overline{\mathcal{G}(K)}}$, is not larger than $S'-K$ times the size of the effective support of $\Phi_{\mathrm{MVU}}[n,k]$ (recall that $|\overline{\mathcal{G}(K)}| = S'-K$). Because of the DFT expression (43) and the fact that $\left| [N]^2 \cap \mathcal{A} \right| = S$ (see (18)), the size of the effective support of $\Phi_{\mathrm{MVU}}[n,k]$ within one period $[N]^2$ can be estimated by N^2/S. Thus, for a small TF sparsity moment $\sigma_X^{(w_\Phi)}$, the RS $\bar{R}_X[n,k]$ should effectively vanish (within $[N]^2$) on a region of size at least $(S'-K)N^2/S \overset{(20)}{\geq} (S-K)N^2/S = N^2 - KN^2/S$. Since typically $S' \approx S$, implying that $(S'-K)N^2/S \approx N^2 - KN^2/S$, it follows that the size of the effective support (within $[N]^2$) of the RS $\bar{R}_X[n,k]$ should not be larger than KN^2/S. Note that K was defined as our prior intuition about the number of significantly nonzero values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$; furthermore, N^2/S is related to the TF undersampling in $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ because (for $S' \approx S$) it is approximately equal to the ratio of the number of samples $\{\widetilde{R}_{X,\mathrm{MVU}}[n,k]\}_{n,k\in[N]}$ (which is N^2) to the number of samples $\{\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]\}_{p\in[\Delta L], q\in[\Delta M]}$ (which is S').

D. Combining the Two MSE Bounds

We will now combine the bound (62) on $\varepsilon = \mathrm{E}\{\|\hat{R}_{X,\mathrm{MVU}} - \bar{R}_X\|_2^2\}$ and the bound (67) or (78) on $\Delta\varepsilon = \mathrm{E}\{\|\hat{R}_{X,\mathrm{CS}} - \hat{R}_{X,\mathrm{MVU}}\|_2^2\}$ into a bound on the MSE $\varepsilon_{\mathrm{CS}} = \mathrm{E}\{\|\hat{R}_{X,\mathrm{CS}} - \bar{R}_X\|_2^2\}$ of the proposed compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$. To this end, let us define the *norm* of a random process $Y[n,k]$ that is N-periodic in n and k as

$$\|Y\|_{\mathrm{R}} \triangleq \sqrt{\mathrm{E}\{\|Y\|_2^2\}} = \sqrt{\sum_{n,k\in[N]} \mathrm{E}\{|Y[n,k]|^2\}}.$$

The estimation error of the compressive RS estimator can be expanded as

$$\hat{R}_{X,\text{CS}}[n,k] - \bar{R}_X[n,k]$$
$$= \hat{R}_{X,\text{CS}}[n,k] - \hat{R}_{X,\text{MVU}}[n,k] + \hat{R}_{X,\text{MVU}}[n,k] - \bar{R}_X[n,k]$$
$$= Y_1[n,k] + Y_2[n,k],$$

where we have set $Y_1[n,k] \triangleq \hat{R}_{X,\text{MVU}}[n,k] - \bar{R}_X[n,k]$ and $Y_2[n,k] \triangleq \hat{R}_{X,\text{CS}}[n,k] - \hat{R}_{X,\text{MVU}}[n,k]$. Hence, the MSE of the compressive RS estimator can be rewritten as

$$\varepsilon_{\text{CS}} = \text{E}\{\|\hat{R}_{X,\text{CS}} - \bar{R}_X\|_2^2\} = \text{E}\{\|Y_1 + Y_2\|_2^2\} = \|Y_1 + Y_2\|_{\text{R}}^2.$$

Using the triangle inequality [58] $\|Y_1 + Y_2\|_{\text{R}} \leq \|Y_1\|_{\text{R}} + \|Y_2\|_{\text{R}}$, we obtain the bound $\varepsilon_{\text{CS}} \leq (\|Y_1\|_{\text{R}} + \|Y_2\|_{\text{R}})^2$. Recognizing that $\|Y_1\|_{\text{R}} = \sqrt{\text{E}\{\|\hat{R}_{X,\text{MVU}} - \bar{R}_X\|_2^2\}} = \sqrt{\varepsilon}$ and $\|Y_2\|_{\text{R}} = \sqrt{\text{E}\{\|\hat{R}_{X,\text{CS}} - \hat{R}_{X,\text{MVU}}\|_2^2\}} = \sqrt{\Delta\varepsilon}$, this bound can be rewritten as

$$\varepsilon_{\text{CS}} \leq (\sqrt{\varepsilon} + \sqrt{\Delta\varepsilon})^2.$$

Inserting the bounds (62) on ε and (67) on $\Delta\varepsilon$ then results in the following bound on ε_{CS}:

$$\varepsilon_{\text{CS}} \leq \|\bar{R}_X\|_2^2 \left[\sqrt{m_X^{(I_{\overline{\mathcal{A}}})} + \frac{S}{N}} + \sqrt{\frac{(S'-K)D^2}{S'K}\,\tilde{\sigma}_X(K)}\right]^2.$$

Alternatively, using the approximate bound (78) on $\Delta\varepsilon$ instead of (67), we obtain the simpler (but looser) approximate bound

$$\varepsilon_{\text{CS}} \lessapprox \|\bar{R}_X\|_2^2 \left[\sqrt{m_X^{(I_{\overline{\mathcal{A}}})} + \frac{S}{N}} + \sqrt{\frac{(S'-K)D^2}{S'K}(N+1)\,\sigma_X^{(w_\Phi)}}\right]^2.$$

We note that our bounds on $\Delta\varepsilon$ are based on the CS bound (32) together with (30), which is known to be very loose [61]. Thus, for a given nominal sparsity degree K and a given number of measurements P satisfying (30), our upper bounds on $\Delta\varepsilon$ and, in turn, on ε_{CS} will generally be quite pessimistic, i.e., too high. However, the bounds are still valuable theoretically in the sense of an asymptotic analysis, because they show that the MSE decreases with increasing underspreadness (expressed by a smaller moment $m_X^{(I_{\overline{\mathcal{A}}})}$ and a smaller ratio S/N) and with increasing TF sparsity (expressed by a smaller moment $\sigma_X^{(w_\Phi)}$).

VI. NUMERICAL STUDY

We will assess the performance of our compressive spectral estimator for two simple examples. The first example is inspired by a cognitive radio application; the second example concerns the analysis of chirp-like signals.

A. Orthogonal Frequency Division Multiplexing Symbol Process

1) Simulation Setting: In a cognitive radio system, a given transmitter/receiver node has to monitor a large overall frequency band and determine the unoccupied bands that it can use for its own transmission [3]–[5]. In our simulation, we consider a single active transmitter employing orthogonal frequency division multiplexing (OFDM) [66], [67], which is a modulation scheme employed, e.g., for wireless local area networks [67], [68], digital video broadcasting [69]–[71], and long term evolution cellular systems [72]. We use $Q = 64$ subcarriers and a cyclic prefix whose length is $1/8$ of the symbol length. Each subcarrier $i \in [Q]$ transmits a symbol s_i that is randomly selected from a quadrature phase-shift keying (QPSK) constellation with normalized symbol energy $|s_i|^2 = 1$. All QPSK symbols are equally likely, and the different subcarrier symbols s_i are statistically independent. The OFDM modulator uses an inverse DFT of length $Q = 64$ to map the frequency-domain transmit symbols s_i into the (discrete) time domain; this is followed by insertion of a cyclic prefix. Assuming an idealized, noise-free channel for simplicity, the resulting transmit signal is also observed by the receiver. However, we assume that our receiver monitors an overall bandwidth that is twice the nominal OFDM bandwidth, B. This corresponds to a twofold oversampling, i.e., a sampling period of $1/(2B)$, and can be easily realized by using an inverse DFT of length $N_{\text{s}} = 2Q = 128$. The lengths of an OFDM symbol and of the cyclic prefix are then given by $N_{\text{s}} = 128$ and $N_{\text{cp}} = 128/8 = 16$ samples, respectively. To keep the simulation complexity low, we assume that a single OFDM symbol is transmitted, with silent periods before and afterwards. Thus, the received time-domain signal (discrete-time baseband representation) is given by

$$X[n] = \begin{cases} \sum_{i \in [Q]} s_i\, e^{j\frac{2\pi}{2Q}(n-n_0)i}, & n \in \{n_0 - N_{\text{cp}}, \dots, \\ & \qquad n_0 + N_{\text{s}} - 1\}_N \\ 0, & \text{otherwise.} \end{cases} \tag{79}$$

Here, n_0 denotes an arbitrary but fixed time offset. In our simulation, we used $n_0 = N_{\text{cp}}$ and considered $X[n]$ for $n \in [N]$ with $N = 512$.

Because of the random s_i, $X[n]$ is a nonstationary random process. The RS and EAF of $X[n]$ are easily obtained from, respectively, (4) and (1) as

$$\bar{R}_X[n,k] = \begin{cases} \sum_{i \in [Q]} \text{dir}\left(N_{\text{s}} + N_{\text{cp}}, \dfrac{k}{N} - \dfrac{i}{2Q}\right) e^{-j\frac{2\pi}{N}nk}, \\ \qquad\qquad n \in \{n_0 - N_{\text{cp}}, \dots, n_0 + N_{\text{s}} - 1\}_N \\ 0, \qquad\qquad \text{otherwise;} \end{cases}$$

$$\bar{A}_X[m,l] = \begin{cases} \sum_{i \in [Q]} \text{dir}\left(N_{\text{s}} + N_{\text{cp}} - m, -\dfrac{l}{N}\right) e^{-j\frac{2\pi}{2Q}mi}, \\ \qquad\qquad m \in \{-N_{\text{cp}} - N_{\text{s}} + 1, \dots, 0\}_N \\ \bar{A}_X^*[-m,-l]_N\, e^{-j\frac{2\pi}{N}ml}, \\ \qquad\qquad m \in \{1, \dots, N_{\text{cp}} + N_{\text{s}} - 1\}_N \\ 0, \qquad\qquad \text{otherwise,} \end{cases}$$
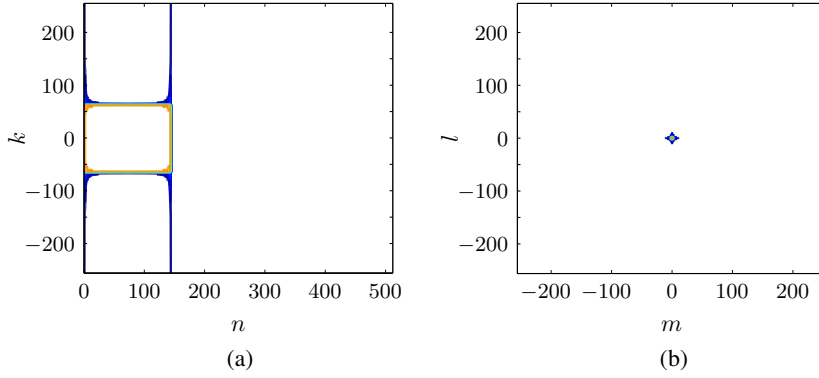
Fig. 1. TF representation of the OFDM process $X[n]$: (a) Real part of RS $\bar{R}_X[n,k]$, displayed for $(n,k) \in [N] \times \{-N/2, \ldots, N/2-1\}$, with $N = 512$; (b) magnitude of EAF $\bar{A}_X[m,l]$, displayed for $(m,l) \in \{-N/2, \ldots, N/2-1\}^2$.

where $\mathrm{dir}(n,\theta) \triangleq \sum_{n'=0}^{n-1} e^{j\pi\theta n'} = e^{j\pi\theta(n-1)} \frac{\sin(\pi\theta n)}{\sin(\pi\theta)}$. Note that the expression for $\bar{A}_X[m,l]$ requires that $N_{\mathrm{s}} + N_{\mathrm{cp}} < N/2$, a condition that is fulfilled in our simulation since $128 + 16 < 512/2$. The RS and EAF are shown in Fig. 1. From this figure, we can conclude that the process $X[n]$ is reasonably TF sparse but only moderately underspread (the latter observation follows from the fact that $\bar{R}_X[n,k]$ is not very smooth). Note that the TF sparsity could be further improved if we considered longer silent periods before and/or after the OFDM symbol, and if we considered a wider band (i.e., if we used an oversampling factor larger than 2).

For the design of the compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ in (33), we used $M = 3$, $L = 7$, $\Delta M = 8$, and $\Delta L = 16$. This corresponds to choosing the effective EAF support (see (2)) as $\mathcal{A} = \{-3, \ldots, 3\}_{512} \times \{-7, \ldots, 7\}_{512}$, of size $S \equiv (2M+1)(2L+1) = 105$; furthermore, the size of the extended effective EAF support $\mathcal{A}'$ is $S' \equiv \Delta M \Delta L = 128$. For an assessment of the TF sparsity of $X[n]$, we consider $h_{p,q} = N^2 \, \mathrm{E}\{|\hat{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]|^2\}$, which underlies the TF sparsity profile $\tilde{\sigma}_X(K)$ in (44). Let $(p,q)_r$ with $r \in \{1, \ldots, S'\}$ be the TF index of the rth largest (in magnitude) value of the set $\{\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]\}_{(p,q)\in[\Delta L]\times[\Delta M]}$, where, as before, $\widetilde{R}_{X,\mathrm{MVU}}[n,k] = \mathrm{E}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\} = \frac{1}{N}\sum_{n',k'\in[N]} \Phi_{\mathrm{MVU}}[n-n', k-k']\,\bar{R}_X[n',k']$ (see (41), (42)). In Fig. 2, we show the values $h_{(p,q)_r}$ along with the corresponding approximations (73)—here denoted $\tilde{h}_{(p,q)_r}$—as a function of the index r. It is seen that $h_{(p,q)_r}$ is close to zero for r larger than 15. Furthermore, we can conclude

that the ordering of the values $\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]$ according to decreasing magnitude matches the ordering of the values $h_{p,q}$ very well. Thus, for TF positions $(p\Delta n, q\Delta k)$ for which $|\widetilde{R}_{X,\mathrm{MVU}}[p\Delta n, q\Delta k]|$ is large, we can expect that also $h_{p,q}$ is large. Finally, it is seen that the curves representing $\tilde{h}_{(p,q)_r}$ and $h_{(p,q)_r}$ coincide, which shows that the approximation (73) is very accurate.

2) Simulation Results: We now consider the estimation of the RS $\bar{R}_X[n,k]$ from a single realization of $X[n]$ that is observed for $n \in [512]$. To evaluate the estimation performance, we generated 1000 realizations of the QPSK symbols $\{s_i\}_{i\in[64]}$ and computed the corresponding realizations of $X[n]$. In Fig. 3, we show the average of 1000 realizations of the compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ (obtained for the 1000 realizations of $X[n]$) as well as a single realization of $\hat{R}_{X,\mathrm{CS}}[n,k]$ for compression factors $S'/P = 1$, 2, and approximately 5 or, equivalently, $P = 128$, 64, and 25 randomly located AF measurements. The optimization in (31), which is required for the computation of $\hat{R}_{X,\mathrm{CS}}[n,k]$ in (33), was carried out using the MATLAB library CVX [63]. The true RS is also re-displayed for easy comparison with the estimates.

The case $S'/P = 1$ corresponds to the basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$ in (17) (cf. the discussion at the end of Section IV-D). We see that already in this case, even for the average $\hat{R}_{X,\mathrm{CS}}[n,k]$, there are noticeable deviations from the true RS. In fact, the average of the 1000 basic RS estimates $\hat{R}_{X,\mathrm{MVU}}[n,k]$ closely approximates the expected basic RS estimator $\widetilde{R}_{X,\mathrm{MVU}}[n,k] = \mathrm{E}\{\hat{R}_{X,\mathrm{MVU}}[n,k]\}$, which according to (42) is a smoothed version of the RS. This smoothing leads to a noticeable deviation from the RS, because the RS itself is not very smooth. The limited smoothness of the RS corresponds to the fact that the process $X[n]$ is only moderately underspread. For compression factor $S'/P = 2$, there is no visible degradation of the average estimate relative to the basic estimator. For $S'/P \approx 5$, a small degradation is visible. The results obtained for the individual realizations suggest a random variation and deviation from the true TF support of the RS that are higher for compression factor $S'/P \approx 5$. The results of the symmetrized compressive estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ in (38) are not shown in Fig. 3 because they can hardly be distinguished visually from
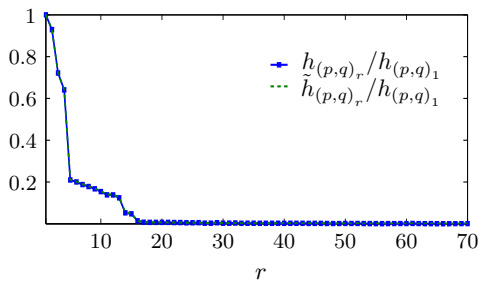

Fig. 2. $h_{(p,q)_r}$ normalized by $h_{(p,q)_1}$ and the corresponding normalized approximation $\tilde{h}_{(p,q)_r}/h_{(p,q)_1}$ according to (73) versus r.
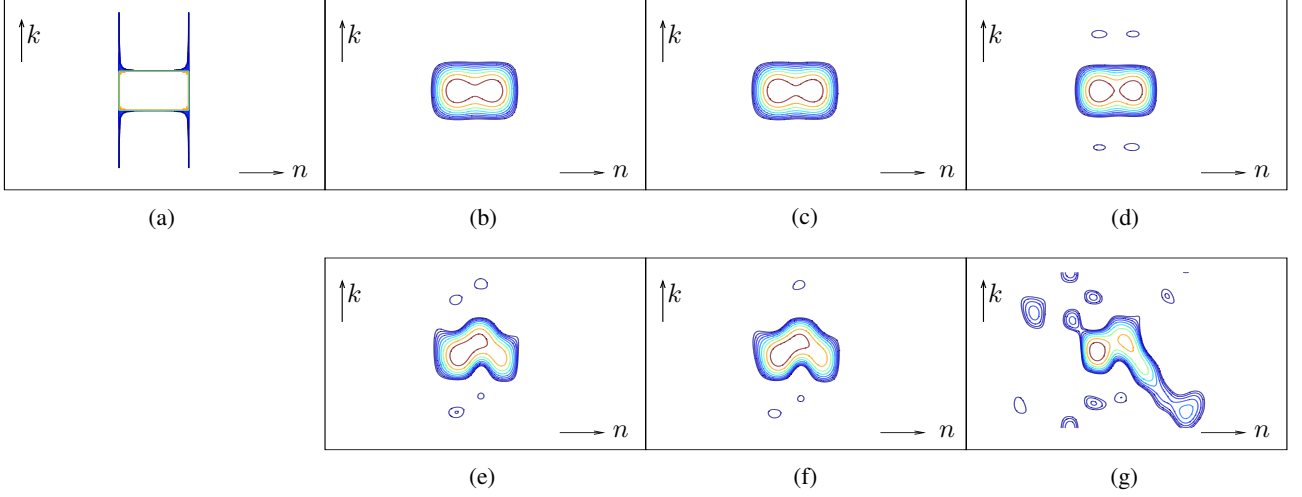
Fig. 3. Averages and single realizations of RS estimators: (a) RS of the OFDM process $X[n]$; (b) average of the noncompressive estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$ (compression factor $S'/P = 1$); (c) and (d) average of the compressive estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ for $S'/P = 2$ and $S'/P \approx 5$, respectively; (e) realization of $\hat{R}_{X,\mathrm{MVU}}[n,k]$; (f) and (g) realization of $\hat{R}_{X,\mathrm{CS}}[n,k]$ for $S'/P = 2$ and $S'/P \approx 5$, respectively. The real parts of all TF functions are shown for $(n,k) \in \{-150,\ldots,361\} \times \{-N/2,\ldots,N/2-1\}$, with $N = 512$.

those of $\hat{R}_{X,\mathrm{CS}}[n,k]$.

For a quantitative analysis of the degradation caused by the compression, we show in Fig. 4 the empirical normalized MSE (NMSE) of the compressive estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ versus the compression factor S'/P. The NMSE is an empirical, normalized version of the MSE $\varepsilon_{\mathrm{CS}} = \mathrm{E}\{\|\hat{R}_{X,\mathrm{CS}} - \bar{R}_X\|_2^2\}$, with the expectation replaced by the sample average over the 1000 process realizations and with normalization by $\|\bar{R}_X\|_2^2$. In the same figure, we also show the empirical normalized versions of the squared bias term $B_{\mathrm{CS}}^2 = \|\mathrm{E}\{\hat{R}_{X,\mathrm{CS}}\} - \bar{R}_X\|_2^2$ and of the variance $V_{\mathrm{CS}} = \mathrm{E}\{\|\hat{R}_{X,\mathrm{CS}} - \mathrm{E}\{\hat{R}_{X,\mathrm{CS}}\}\|_2^2\}$, again with normalization by $\|\bar{R}_X\|_2^2$. (Recall that $\varepsilon_{\mathrm{CS}} = B_{\mathrm{CS}}^2 + V_{\mathrm{CS}}$.) These results demonstrate a "graceful degradation" with increasing compression factor S'/P. Again, the result for $S'/P = 1$ corresponds to the basic RS estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$. Fig. 4 also shows the NMSE, normalized squared bias term, and normalized variance of the symmetrized compressive estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$. It is seen that the variance and MSE are reduced

by the symmetrization. We did not plot the MSE bounds derived in Section V because they are much larger than the empirical MSE. As mentioned in Section V-D, this lack of tightness is mostly due to the notoriously loose [61] CS error bound used in (64) (combined with (30)).

3) Comparison with a Reference Method: Next, we compare our compressive nonstationary spectral estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ with the compressive spectral estimation method proposed in [50], hereafter termed "reference estimator." The reference estimator was devised for estimating the power spectral density of a *stationary* random process; the underlying stationarity assumption allows the use of long-term averaging. However, for the nonstationary processes considered in this paper, long-term averaging is not an option and hence a deteriorated performance must be expected. We nevertheless chose the reference estimator for a performance comparison because we are not aware of any previously proposed compressive spectral estimation method for general nonstationary processes.

The reference estimator uses as its input an observed realization $\mathbf{x}$ of a block of a stationary discrete random process $X[n]$ and calculates a reduced number of compressive measurements, for some compression factor c. From these measurements, it derives an estimate $\hat{P}_X(\omega)$ of the power spectral density $P_X(\omega) \triangleq \sum_{m \in \mathbb{Z}} r_X[m]\, e^{-j\omega m}$ (here, $r_X[m] \triangleq \mathrm{E}\{X[m]X^*[0]\}$ is the autocorrelation function of $X[n]$). In our case, however, $X[n]$ is the nonstationary OFDM process of length N defined in (79). In order to impart a time dependence (time resolution) to the reference estimator, we consecutively apply it to a sequence of overlapping length-N_b blocks $\mathbf{x}^{(b)} \triangleq (x[b\Delta N]\ x[b\Delta N+1] \cdots x[b\Delta N + N_\mathrm{b}-1])^T$, $b \in \{0,\ldots,B-1\}$ of the observed realization $x[n]$ of $X[n]$. Here, $B = \lfloor \frac{N-N_\mathrm{b}+1}{\Delta N} \rfloor + 1$ and $N_\mathrm{b} \geq \Delta N$; note that $N_\mathrm{b} - \Delta N$ is the overlap length. For each block $\mathbf{x}^{(b)}$, we thus obtain a (discrete-frequency) local power spectrum
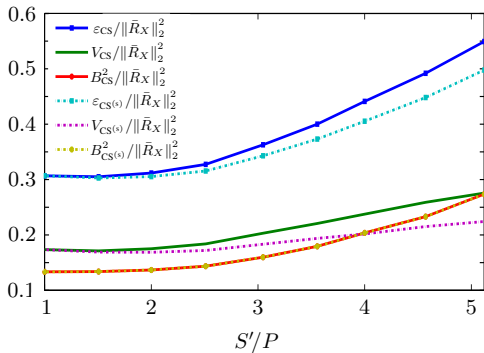

Fig. 4. Empirical NMSE, normalized squared bias, and normalized variance of the compressive RS estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ (solid curves) and of the symmetrized compressive RS estimator $\hat{R}_{X,\mathrm{CS}}^{(\mathrm{s})}[n,k]$ (dash-dotted curves) versus the compression factor S'/P.
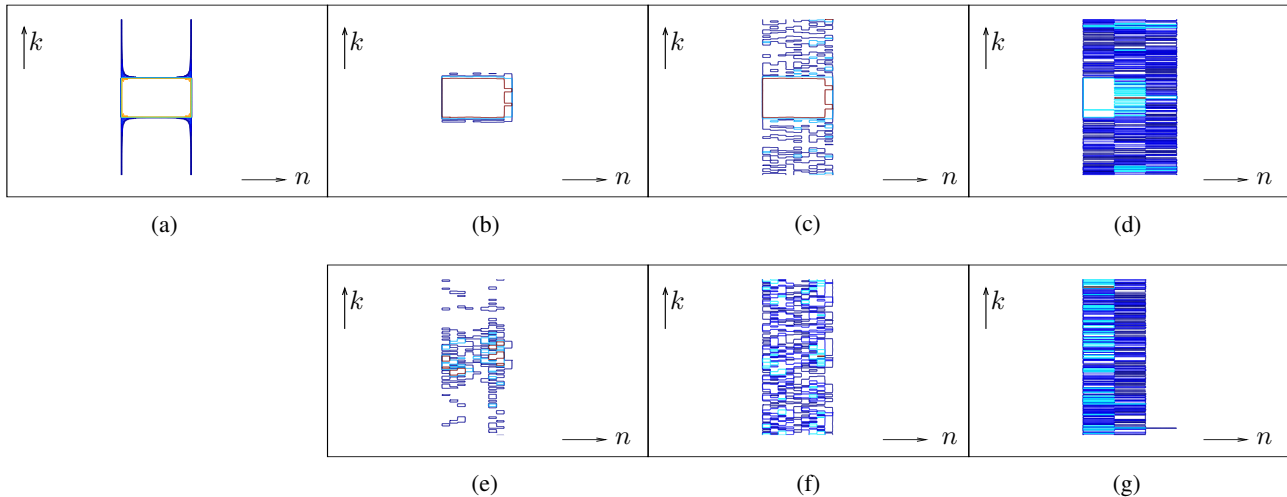
Fig. 5. Averages and single realizations of the reference estimator $\hat{R}_X^{(\text{ref})}[n,k]$: (a) real part of the RS of the OFDM process $X[n]$; (b), (c), and (d) magnitude of the average of $\hat{R}_X^{(\text{ref})}[n,k]$ for $c=1$, $c=2$, and $c\approx 5$, respectively; (e), (f), and (g) magnitude of a realization of $\hat{R}_X^{(\text{ref})}[n,k]$ for $c=1$, $c=2$, and $c\approx 5$, respectively. All TF functions are shown for $(n,k) \in \{-150,\dots,361\} \times \{-N/2,\dots,N/2-1\}$, with $N=512$.

estimate[11] $\hat{P}_X^{(b)}(2\pi k/N)$, $k \in [N]$. From the sequence of local power spectrum estimates $\hat{P}_X^{(b)}(2\pi k/N)$, $b=0,\dots,B-1$, we then construct a *time-dependent* (more specifically, piecewise constant) compressive power spectral density estimate $\hat{R}_X^{(\text{ref})}[n,k]$ by setting $\hat{R}_X^{(\text{ref})}[n,k] \triangleq \hat{P}_X^{(b)}(2\pi k/N)$ for $n \in \{b\Delta N, b\Delta N+1,\dots,(b+1)\Delta N-1\}$, with $b \in \{0,\dots,B-1\}$.

In Fig. 5, we show the average of $\hat{R}_X^{(\text{ref})}[n,k]$ obtained for 1000 realizations of the OFDM process $X[n]$ as well as a single realization of $\hat{R}_X^{(\text{ref})}[n,k]$ for compression factors $c=1$, 2, and approximately 5. We used block length $N_\text{b}=32$ and time increment $\Delta N=16$ for $c=1$ and $c=2$, and[12] $N_\text{b}=128$ and $\Delta N=64$ for $c\approx 5$. For convenience, the true RS is again re-displayed in part (a). A comparison of Fig. 5 with Fig. 3 shows that the proposed estimator $\hat{R}_{X,\text{CS}}[n,k]$ clearly outperforms the reference estimator $\hat{R}_X^{(\text{ref})}[n,k]$, especially when single realizations are considered and in the compressive case ($c>1$), which are the cases of greatest relevance in our context. This result is not surprising and should not be interpreted as evidence of poor performance of the estimator proposed in [50]. In fact, as noted previously, that estimator was devised for stationary random processes where long-term averaging can be used, and it was not intended for our straightforward and somewhat naïve adaptation to nonstationary processes.

B. Chirp Process

Next, we apply our compressive RS estimator to a two-component chirp process. Chirp signals arise, e.g., in the context of engine diagnosis [73]–[75], system identification and radar [76]–[79], and the study of bat echolocation [80], [81].

---

[11] We note the following details of our implementation of the reference estimator (cf. [50] for background and notation). The maximum correlation lag was chosen as $L=1$. The second-order statistics (cross-correlation functions) $r_{y_i,y_j}[k]$ were estimated by time-averages over blocks of length $L+1=2$. The weights $c_i[n]$ were randomly drawn from the set $\{-1,1\}$ with equal probabilities.

[12] These different choices of N_b and ΔN for $c=1$, 2 and for $c\approx 5$ are due to the condition $N_\text{b} \geq 2(2c-1)c$ that is required by the reference estimator [50].

In our simulation, we construct a finite-length, nonstationary, discrete-time process as $X[n] \triangleq X(nT_\text{s})$, $n \in [512]$, where T_s is some sampling period. The continuous-time process $X(t)$ is given by

$$X(t) = a_1 s(t-t_1) + a_2 s(t-t_2),$$

where $t_1 = 128\,T_\text{s}$ and $t_2 = 384\,T_\text{s}$; a_1 and a_2 are independent zero-mean Gaussian random variables with unit variance; and $s(t)$ is a chirp pulse defined as $s(t) = \exp\left(-(t/T_0)^2/2\right)\exp(-j\pi\beta t^2)$, with pulse width parameter $T_0 = 60\,T_\text{s}$ and chirp rate $\beta = 1/(600\,T_\text{s}^2)$. The RS and EAF of the discrete-time process $X[n]$ are shown in Fig. 6. We see that the process $X[n]$ is only moderately TF sparse and not very underspread, i.e., the underspread approximation used, e.g., in Section III can be hardly justified.

We implemented the compressive RS estimator $\hat{R}_{X,\text{CS}}[n,k]$ in (33) as well as the symmetrized compressive estimator $\hat{R}_{X,\text{CS}}^{(\text{s})}[n,k]$ in (38) using $M=L=15$ and $\Delta M = \Delta L = 32$. This corresponds to the effective EAF support $\mathcal{A} = \{-15,\dots,15\}_{512} \times \{-15,\dots,15\}_{512}$, of size $S \equiv (2M+1)(2L+1) = 961$. The size of the extended effective EAF support $\mathcal{A}'$ is $S' \equiv \Delta M \Delta L = 1024$. Fig. 7 shows the average of 1000 realizations of $\hat{R}_{X,\text{CS}}[n,k]$ and $\hat{R}_{X,\text{CS}}^{(\text{s})}[n,k]$ (obtained for 1000 realizations of $X(t)$) as well as a single realization of $\hat{R}_{X,\text{CS}}[n,k]$ for compression factors $S'/P \approx 5$ and 10 or, equivalently, $P = 204$ and 102 randomly located AF measurements. We see that already in the noncompressive case $S'/P = 1$, where $\hat{R}_{X,\text{CS}}[n,k]$ and $\hat{R}_{X,\text{CS}}^{(\text{s})}[n,k]$ coincide with the basic RS estimator $\hat{R}_{X,\text{MVU}}[n,k]$, there are noticeable deviations from the true RS; these differences are again due to the smoothing employed by $\hat{R}_{X,\text{MVU}}[n,k]$. However, the proposed compressive RS estimator $\hat{R}_{X,\text{CS}}[n,k]$ still performs well in the sense that it indicates the main characteristics of the two chirp signal components—the TF locations and the chirp rate—up to a compression factor of 10, i.e., based on the observation of a significantly reduced number of AF samples. In this sense, our estimator appears to be robust
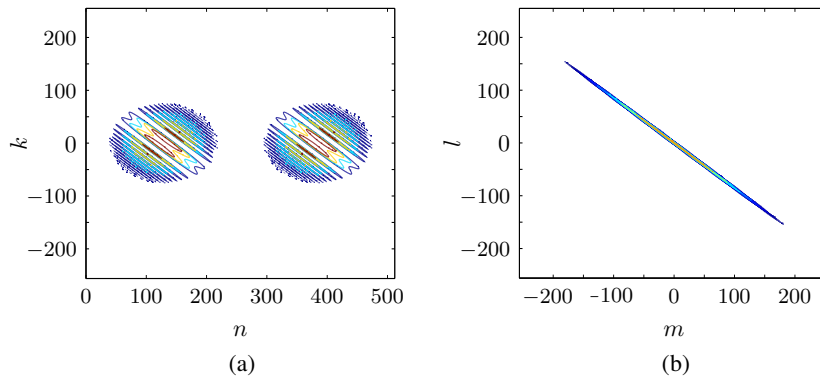
Fig. 6.   TF representation of the two-component chirp process $X[n]$: (a) Real part of RS $\bar{R}_X[n,k]$, displayed for $(n,k) \in [N] \times \{-N/2, \ldots, N/2 - 1\}$, with $N = 512$; (b) magnitude of EAF $\bar{A}_X[m,l]$, displayed for $(m,l) \in \{-N/2, \ldots, N/2 - 1\}^2$.
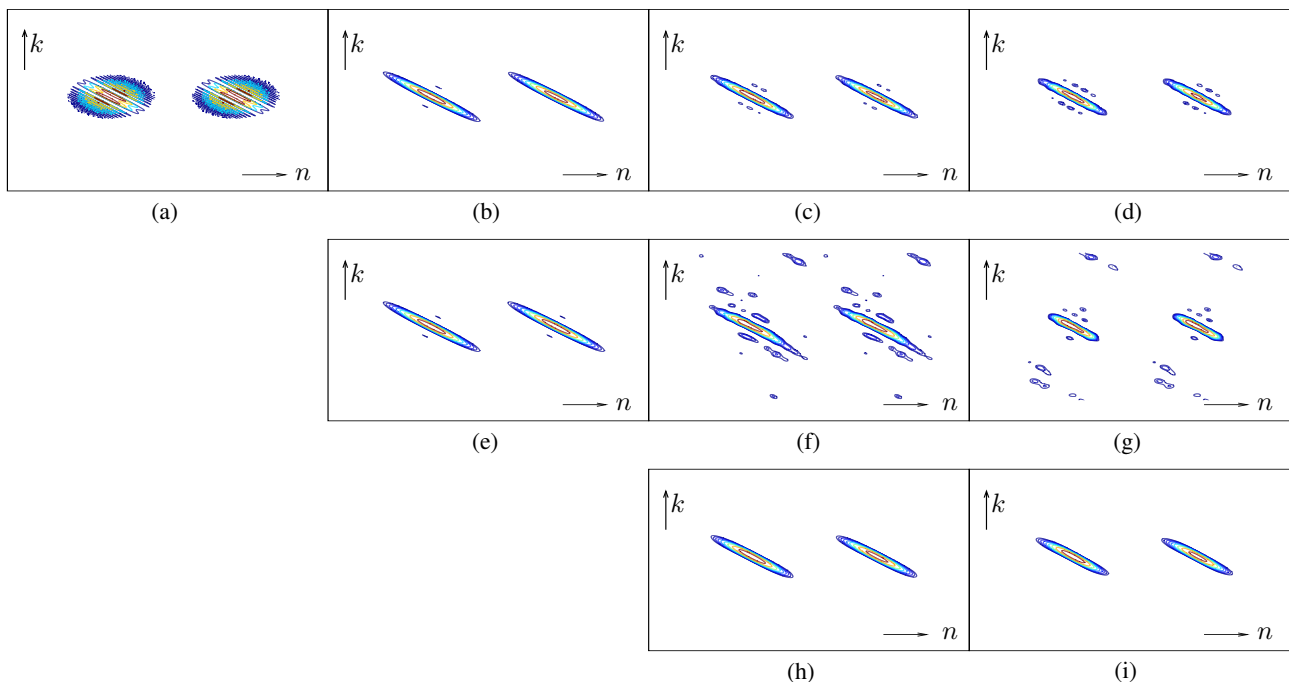

Fig. 7.   Averages and single realizations of RS estimators: (a) RS of the chirp process $X[n]$; (b) average of the noncompressive estimator $\hat{R}_{X,\mathrm{MVU}}[n,k]$ (compression factor $S'/P = 1$); (c) and (d) average of the compressive estimator $\hat{R}_{X,\mathrm{CS}}[n,k]$ for $S'/P \approx 5$ and 10, respectively; (e) realization of $\hat{R}_{X,\mathrm{MVU}}[n,k]$; (f) and (g) realization of $\hat{R}_{X,\mathrm{CS}}[n,k]$ for $S'/P \approx 5$ and 10, respectively; (h) and (i) average of the symmetrized compressive estimator $\hat{R}_{X,\mathrm{CS}}^{(s)}[n,k]$ for $S'/P \approx 5$ and 10, respectively. The real parts of all TF functions are shown for $(n,k) \in [N] \times \{-N/2, \ldots, N/2 - 1\}$, with $N = 512$.

to deviations from the assumed properties of approximate TF sparsity and underspreadness. More specifically, the main deviation from the true RS is due to the fact that the oscillatory structures (*inner interference terms* [82]) contained in the RS are suppressed by the smoothing; this result is in fact desirable in most applications. It is furthermore seen that the average results of the symmetrized estimator $\hat{R}_{X,\mathrm{CS}}^{(s)}[n,k]$ are similar to those of $\hat{R}_{X,\mathrm{CS}}[n,k]$.

VII. Conclusion

For estimating a time-dependent spectrum of a nonstationary random process, long-term averaging cannot be used as this would smear out the time-dependence of the spectrum.

However, if the spectrum as a function of time and frequency is sufficiently smooth, which amounts to an *underspread* assumption, a local TF smoothing can be used. In particular, the RS of an underspread nonstationary process can be estimated by a local smoothing of a TF distribution known as the RD.

In this paper, we have considered the practically relevant case of underspread processes that are approximately TF sparse in the sense that only a moderate percentage of the RS values are significantly nonzero. For such processes, we have proposed a "compressive" RS estimator that exploits the TF sparsity structure for a significant reduction of the number of measurements required for good estimation performance. The measurements are values of the AF of the observed signal

at randomly chosen time lag/frequency lag positions. Our overall approach is advantageous if dedicated hardware units for computing values of the AF from the original continuous-time signal are employed, and/or if the AF values have to be transmitted over low-rate links or stored in a memory. The proposed compressive RS estimator extends a conventional RS estimator for underspread processes (a smoothed RD using an MVU design of the smoothing function) by a CS reconstruction technique. For the latter, we used the Basis Pursuit because it is supported by a convenient performance guarantee (a bound on the ℓ_2-norm of the reconstruction error); however, other CS reconstruction techniques can be used as well.

We provided upper bounds on the MSE of both the MVU RS estimator and its compressive extension. The MSE bound for the compressive estimator is based on the error bound of the Basis Pursuit, which is known to be quite loose. Therefore, the MSE bound for the compressive estimator is usually quite pessimistic. However, it is still useful theoretically, since it reveals the asymptotic dependence of the estimation accuracy on the underspreadness and TF sparsity properties of the process. Numerical experiments demonstrated the good performance of our compressive estimator for two typical scenarios.

We considered the RS because in the discrete setting used, it is the simplest time-dependent spectrum from a computational viewpoint. However, for underspread processes, the RS is very close to other important time-dependent spectra such as the Wigner-Ville spectrum and the evolutionary spectrum. Therefore, the proposed RS estimator can also be used for estimating other time-dependent spectra if the process is sufficiently underspread. Finally, the proposed RS estimator can also be used for estimating the EAF and the autocorrelation function, which are related to the RS via DFTs.

APPENDIX A: MSE OF THE SYMMETRIZED COMPRESSIVE RS ESTIMATOR

We will prove the MSE inequality (39). Let us define the symmetrization operator corresponding to (37), i.e.,

$$\mathbf{P}_s A[m,l] = \frac{1}{2} \left[A[m,l] + A^*[-m,-l] e^{-j\frac{2\pi}{N}ml} \right]$$

and note that (see (36b), (37))

$$\mathbf{P}_s \bar{A}_X[m,l] = \bar{A}_X[m,l], \quad \mathbf{P}_s \hat{A}_{X,\text{CS}}[m,l] = \hat{A}_{X,\text{CS}}^{(s)}[m,l]. \tag{80}$$

Furthermore, let us consider the estimation error of the compressive EAF estimator $\hat{A}_{X,\text{CS}}[m,l]$, $E[m,l] \triangleq \hat{A}_{X,\text{CS}}[m,l] - \bar{A}_X[m,l]$. Using the triangle inequality [83], $\|\mathbf{P}_s E\|_2 = \frac{1}{2} \|E[m,l] + E^*[-m,-l]e^{-j\frac{2\pi}{N}ml}\|_2 \leq \frac{1}{2} [\|E\|_2 + \|E\|_2] = \|E\|_2$. Using $\mathbf{P}_s E = \mathbf{P}_s \hat{A}_{X,\text{CS}} - \mathbf{P}_s \bar{A}_X$ and (80), it is seen that the above inequality is equivalent to

$$\left\| \hat{A}_{X,\text{CS}}^{(s)} - \bar{A}_X \right\|_2^2 \leq \left\| \hat{A}_{X,\text{CS}} - \bar{A}_X \right\|_2^2. \tag{81}$$

Furthermore, since $\bar{A}_X[m,l]$, $\hat{A}_{X,\text{CS}}[m,l]$, and $\hat{A}_{X,\text{CS}}^{(s)}[m,l]$ are related to $\bar{R}_X[n,k]$, $\hat{R}_{X,\text{CS}}[n,k]$, and $\hat{R}_{X,\text{CS}}^{(s)}[n,k]$, respectively via the 2D DFT in (35), which is norm-preserving, the

inequality (81) implies that

$$\left\| \hat{R}_{X,\text{CS}}^{(s)} - \bar{R}_X \right\|_2^2 \leq \left\| \hat{R}_{X,\text{CS}} - \bar{R}_X \right\|_2^2.$$

Finally, taking the expectation on both sides yields the MSE inequality (39).

APPENDIX B: TF SHIFT MATRICES

We consider the family of (scaled) discrete TF shift matrices $\{\mathbf{J}_{m,l}\}_{m,l\in[N]}$ of size $N \times N$ whose action on $\mathbf{x} \in \mathbb{C}^N$ is given by

$$(\mathbf{J}_{m,l}\mathbf{x})_{n+1} = \frac{1}{\sqrt{N}} (\mathbf{x})_{(n-m)_N+1} e^{j\frac{2\pi}{N}ln}, \quad n \in [N], \tag{82}$$

with $(n)_N \triangleq n \bmod N$. These matrices can be written $\mathbf{J}_{m,l} = \frac{1}{\sqrt{N}} \mathbf{M}_l \mathbf{T}_m$, where $\mathbf{M}_l$ is the diagonal $N \times N$ matrix with diagonal elements $1, e^{j\frac{2\pi}{N}l}, \ldots, e^{j\frac{2\pi}{N}l(N-1)}$ and $\mathbf{T}_m$ is the circulant $N \times N$ matrix whose entries $(\mathbf{T}_m)_{n,n'}$ are given by 1 if $(n-n')_N = (m)_N$ and 0 otherwise. It can be easily verified that the set $\{\mathbf{J}_{m,l}\}_{m,l\in[N]}$ forms an orthonormal basis for the linear space of matrices $\mathbb{C}^{N \times N}$ equipped with inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}\{\mathbf{A}\mathbf{B}^H\}$, i.e.,

$$\langle \mathbf{J}_{m,l}, \mathbf{J}_{m',l'} \rangle = \delta[m-m']_N \, \delta[l-l']_N \tag{83}$$

and

$$\mathbf{A} = \sum_{m,l\in[N]} \langle \mathbf{A}, \mathbf{J}_{m,l} \rangle \mathbf{J}_{m,l}, \quad \text{for all } \mathbf{A} \in \mathbb{C}^{N \times N}. \tag{84}$$

It can furthermore be shown that the EAF in (1) and the AF in (3) can be written as

$$\bar{A}_X[m,l] = \sqrt{N} \langle \mathbf{\Gamma}_X, \mathbf{J}_{m,l} \rangle$$
$$A_X[m,l] = \sqrt{N} \langle \mathbf{x}\mathbf{x}^H, \mathbf{J}_{m,l} \rangle, \tag{85}$$

where $\mathbf{\Gamma}_X = \text{E}\{\mathbf{x}\mathbf{x}^H\}$ with $\mathbf{x} = (X[0] \cdots X[N-1])^T$. Thus, according to (84), we have the expansions

$$\mathbf{\Gamma}_X = \frac{1}{\sqrt{N}} \sum_{m,l\in[N]} \bar{A}_X[m,l] \mathbf{J}_{m,l} \tag{86}$$

$$\mathbf{x}\mathbf{x}^H = \frac{1}{\sqrt{N}} \sum_{m,l\in[N]} A_X[m,l] \mathbf{J}_{m,l}.$$

Finally, from (82), one can deduce the following relations:

$$\mathbf{J}_{m,l}\mathbf{J}_{m',l'} = \frac{1}{\sqrt{N}} \mathbf{J}_{m+m',l+l'} e^{-j\frac{2\pi}{N}ml'} \tag{87}$$

$$\mathbf{J}_{m,l}^H = \mathbf{J}_{-m,-l} e^{-j\frac{2\pi}{N}ml}, \tag{88}$$

and, in turn,

$$\mathbf{J}_{n,k}\mathbf{J}_{m,l}\mathbf{J}_{n,k}^H = \frac{1}{N} \mathbf{J}_{m,l} e^{-j\frac{2\pi}{N}(nl-km)}. \tag{89}$$

APPENDIX C: DERIVATION OF EXPRESSIONS (58) AND (59)

We will derive (58) and (59) from (57). Our derivation will be based on expansions of $\mathbf{C}_{n,k}^{(\mathrm{R})}$ and $\mathbf{C}_{n,k}^{(\mathrm{I})}$ into the TF shift matrices $\mathbf{J}_{m,l}$. Using (55), (50), and (88), we have

$$
\mathbf{C}_{n,k}^{(\mathrm{R})} \overset{(55),(50)}{=} \frac{1}{2\sqrt{N}} \Bigg[\sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{-j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{m,l}^{H}
$$
$$
+ \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{m,l} \Bigg]
$$
$$
\overset{(88)}{=} \frac{1}{2\sqrt{N}} \Bigg[\sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{-j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{-m,-l}\, e^{-j\frac{2\pi}{N}ml}
$$
$$
+ \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j\frac{2\pi}{N}(km-nl)} \mathbf{J}_{m,l} \Bigg]
$$
$$
= \frac{1}{2\sqrt{N}} \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j\frac{2\pi}{N}(km-nl)} \left(e^{-j\frac{2\pi}{N}ml} + 1 \right) \mathbf{J}_{m,l}
$$
$$
= \sum_{m,l\in[N]} e^{j\frac{2\pi}{N}(km-nl)} c_{m,l}^{(\mathrm{R})} \mathbf{J}_{m,l}, \tag{90}
$$

with

$$
c_{m,l}^{(\mathrm{R})} \triangleq \frac{1}{2\sqrt{N}}\, I_{\mathcal{A}}[m,l] \left(e^{-j\frac{2\pi}{N}ml} + 1 \right). \tag{91}
$$

In a similar way, we obtain from (55)

$$
\mathbf{C}_{n,k}^{(\mathrm{I})} = \sum_{m,l\in[N]} e^{j\frac{2\pi}{N}(km-nl)} c_{m,l}^{(\mathrm{I})} \mathbf{J}_{m,l}, \tag{92}
$$

with

$$
c_{m,l}^{(\mathrm{I})} \triangleq \frac{1}{2j\sqrt{N}}\, I_{\mathcal{A}}[m,l] \left(e^{-j\frac{2\pi}{N}ml} - 1 \right). \tag{93}
$$

The first term in (57) can then be written as

$$
\sum_{n,k\in[N]} \mathrm{tr}\big\{ \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \big\} = \mathrm{tr}\{\mathbf{D}\mathbf{\Gamma}_X\} = \mathrm{tr}\{\mathbf{D}\mathbf{\Gamma}_X^{H}\}, \tag{94}
$$

with

$$
\mathbf{D} \triangleq \sum_{n,k\in[N]} \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})}
$$
$$
= \sum_{n,k\in[N]} \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})H}
$$
$$
\overset{(90)}{=} \sum_{m,l,m',l'\in[N]} c_{m,l}^{(\mathrm{R})} c_{m',l'}^{(\mathrm{R})*}
$$
$$
\times \underbrace{\left[\sum_{n,k\in[N]} e^{j\frac{2\pi}{N}[k(m-m')-n(l-l')]} \right]}_{N^2\,\delta[m-m']_N\,\delta[l-l']_N} \mathbf{J}_{m,l} \mathbf{\Gamma}_X \mathbf{J}_{m',l'}^{H}
$$
$$
= N^2 \sum_{m,l\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \mathbf{J}_{m,l} \mathbf{\Gamma}_X \mathbf{J}_{m,l}^{H}
$$
$$
\overset{(86)}{=} N\sqrt{N} \sum_{m,l,m'l'\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \bar{A}_X[m',l'] \mathbf{J}_{m,l} \mathbf{J}_{m',l'} \mathbf{J}_{m,l}^{H}
$$

$$
\overset{(89)}{=} \sqrt{N} \sum_{m,l,m'l'\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \bar{A}_X[m',l'] \mathbf{J}_{m',l'}
$$
$$
\times e^{-j\frac{2\pi}{N}(ml'-lm')} \tag{95}
$$

and

$$
\mathbf{\Gamma}_X^{H} \overset{(86)}{=} \frac{1}{\sqrt{N}} \sum_{m,l\in[N]} \bar{A}_X^{*}[m,l] \mathbf{J}_{m,l}^{H}. \tag{96}
$$

Inserting (95) and (96) into (94) then yields

$$
\sum_{n,k\in[N]} \mathrm{tr}\big\{ \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{R})} \mathbf{\Gamma}_X \big\}
$$
$$
= \mathrm{tr}\Bigg\{ \sum_{m,l,m',l',m'',l''\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \bar{A}_X[m',l'] \mathbf{J}_{m',l'}
$$
$$
\times e^{-j\frac{2\pi}{N}(ml'-lm')} \bar{A}_X^{*}[m'',l''] \mathbf{J}_{m'',l''}^{H} \Bigg\}
$$
$$
= \sum_{m,l,m',l',m'',l''\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \bar{A}_X[m',l'] \bar{A}_X^{*}[m'',l'']
$$
$$
\times e^{-j\frac{2\pi}{N}(ml'-lm')} \underbrace{\mathrm{tr}\big\{ \mathbf{J}_{m',l'} \mathbf{J}_{m'',l''}^{H} \big\}}_{\langle \mathbf{J}_{m',l'}, \mathbf{J}_{m'',l''}\rangle \overset{(83)}{=} \delta[m'-m'']_N\,\delta[l'-l'']_N}
$$
$$
= \sum_{m,l,m',l'\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \big| \bar{A}_X[m',l'] \big|^2 e^{-j\frac{2\pi}{N}(ml'-lm')}
$$
$$
= \sum_{m,l,m',l'\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \big| \bar{A}_X[m',l'] \big|^2 e^{j\frac{2\pi}{N}(ml'-lm')}, \tag{97}
$$

where the symmetry relation (36b) has been used in the last step.

In a similar manner, using (92), we obtain for the second term in (57)

$$
\sum_{n,k\in[N]} \mathrm{tr}\big\{ \mathbf{C}_{n,k}^{(\mathrm{I})} \mathbf{\Gamma}_X \mathbf{C}_{n,k}^{(\mathrm{I})} \mathbf{\Gamma}_X \big\}
$$
$$
= \sum_{m,l,m',l'\in[N]} \big| c_{m,l}^{(\mathrm{I})} \big|^2 \big| \bar{A}_X[m',l'] \big|^2 e^{j\frac{2\pi}{N}(ml'-lm')}. \tag{98}
$$

Inserting (97) and (98) into (57) then gives (58):

$$
V = \sum_{m,l,m',l'\in[N]} \big| c_{m,l}^{(\mathrm{R})} \big|^2 \big| \bar{A}_X[m',l'] \big|^2 e^{j\frac{2\pi}{N}(ml'-lm')}
$$
$$
+ \sum_{m,l,m',l'\in[N]} \big| c_{m,l}^{(\mathrm{I})} \big|^2 \big| \bar{A}_X[m',l'] \big|^2 e^{j\frac{2\pi}{N}(ml'-lm')}
$$
$$
= \sum_{m',l'\in[N]} \big| \bar{A}_X[m',l'] \big|^2 \chi[m',l'],
$$

with

$$
\chi[m,l] \triangleq \sum_{m',l'\in[N]} \left(\big| c_{m',l'}^{(\mathrm{R})} \big|^2 + \big| c_{m',l'}^{(\mathrm{I})} \big|^2 \right) e^{j\frac{2\pi}{N}(m'l-l'm)}. \tag{99}
$$

Using (91) and (93), we have

$$
\big| c_{m,l}^{(\mathrm{R})} \big|^2 + \big| c_{m,l}^{(\mathrm{I})} \big|^2
$$
$$
= \frac{1}{N} I_{\mathcal{A}}[m,l] \left[\left| \frac{e^{-j\frac{2\pi}{N}ml} + 1}{2} \right|^2 + \left| \frac{e^{-j\frac{2\pi}{N}ml} - 1}{2j} \right|^2 \right]
$$
$$
= \frac{1}{N} I_{\mathcal{A}}[m,l] \left[\cos^2\!\Big(\frac{\pi}{N}ml\Big) + \sin^2\!\Big(\frac{\pi}{N}ml\Big) \right]
$$

$$= \frac{1}{N} I_{\mathcal{A}}[m,l].$$

Inserting this into (99) yields (59).

APPENDIX D: DERIVATION OF EXPRESSION (73)

We will derive (73) from (72).

1) Expansions of $\mathbf{T}_{p,q}^{(\mathrm{R})}$ and $\mathbf{T}_{p,q}^{(\mathrm{I})}$: Our derivation will be based on the underspread assumption and on expansions of $\mathbf{T}_{p,q}^{(\mathrm{R})}$ and $\mathbf{T}_{p,q}^{(\mathrm{I})}$ into the TF shift matrices $\mathbf{J}_{m,l}$. Inserting (70) into the definition of $\mathbf{T}_{p,q}^{(\mathrm{R})}$ in (71) yields

$$
\mathbf{T}_{p,q}^{(\mathrm{R})} = \frac{\sqrt{N}}{2} \left[\sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{-j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} \mathbf{J}_{m,l}^{H} \right.
$$
$$
\left. + \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} \mathbf{J}_{m,l} \right]
$$
$$
\stackrel{(88)}{=} \frac{\sqrt{N}}{2} \left[\sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{-j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} \mathbf{J}_{-m,-l}\, e^{-j\frac{2\pi}{N}ml} \right.
$$
$$
\left. + \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} \mathbf{J}_{m,l} \right]
$$
$$
= \frac{\sqrt{N}}{2} \sum_{m=-M}^{M} \sum_{l=-L}^{L} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} \left(e^{-j\frac{2\pi}{N}ml} + 1\right) \mathbf{J}_{m,l}
$$
$$
= \sum_{m,l\in[N]} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} t_{m,l}^{(\mathrm{R})} \mathbf{J}_{m,l}, \tag{100}
$$

with

$$
t_{m,l}^{(\mathrm{R})} = \frac{\sqrt{N}}{2} I_{\mathcal{A}}[m,l] \left(e^{-j\frac{2\pi}{N}ml} + 1\right). \tag{101}
$$

In a similar manner, we obtain the expansion

$$
\mathbf{T}_{p,q}^{(\mathrm{I})} = \sum_{m,l\in[N]} e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} t_{m,l}^{(\mathrm{I})} \mathbf{J}_{m,l}, \tag{102}
$$

with

$$
t_{m,l}^{(\mathrm{I})} = \frac{\sqrt{N}}{2j} I_{\mathcal{A}}[m,l] \left(e^{-j\frac{2\pi}{N}ml} - 1\right). \tag{103}
$$

For an underspread process $X[n]$, the effective EAF support $\mathcal{A} \equiv \{-M,\ldots,M\}_N \times \{-L,\ldots,L\}_N$ is a small region about the origin of the (m,l) plane (plus its periodically continued replicas, which are irrelevant to our argument and will hence be disregarded). Looking at the expressions of $t_{m,l}^{(\mathrm{R})}$ and $t_{m,l}^{(\mathrm{I})}$ in (101) and (103), we can then conclude from the presence of the factor $I_{\mathcal{A}}[m,l]$ that $t_{m,l}^{(\mathrm{R})}$ and $t_{m,l}^{(\mathrm{I})}$ can be nonzero only for $|ml| \ll N$. This means that in (101) and (103), we can approximate $e^{-j\frac{2\pi}{N}ml}$ by 1, yielding

$$
t_{m,l}^{(\mathrm{R})} \approx \sqrt{N}\, I_{\mathcal{A}}[m,l] \tag{104}
$$
$$
t_{m,l}^{(\mathrm{I})} \approx 0. \tag{105}
$$

Using (105) in (102) yields

$$
\mathbf{T}_{p,q}^{(\mathrm{I})} \approx \mathbf{0}, \tag{106}
$$

and thus (72) approximately simplifies to

$$
h_{p,q} \approx \operatorname{tr}\{\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X\} + \left|\operatorname{tr}\{\mathbf{\Gamma}_X \mathbf{T}_{p,q}^{H}\}\right|^2. \tag{107}
$$

2) First term in (107)*:* We will now develop the two terms on the right-hand side of (107). The first term can be written as

$$
\operatorname{tr}\{\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X \mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X\} = \left\langle \mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X, \left(\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X\right)^{H} \right\rangle. \tag{108}
$$

In order to find an approximation for this inner product, we use the following general result for the product $\mathbf{C} = \mathbf{AB}$ of two $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$. The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be expanded into the orthonormal basis $\{\mathbf{J}_{m,l}\}_{m,l\in[N]}$, with respective expansion coefficients $a_{m,l}$, $b_{m,l}$, and $c_{m,l}$, e.g., $\mathbf{A} = \sum_{m,l\in[N]} a_{m,l} \mathbf{J}_{m,l}$. Then the $c_{m,l}$ are related to the $a_{m,l}$ and $b_{m,l}$ by the "twisted convolution" [24], [29], [84], [85]

$$
c_{m,l} = \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} a_{m',l'}\, b_{m-m',l-l'}\, e^{-j\frac{2\pi}{N}m'(l-l')}. \tag{109}
$$

This expression can be verified using (87). Let us apply it to the matrix product $\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X$. We have the expansion

$$
\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X = \sum_{m,l\in[N]} d_{p,q;m,l} \mathbf{J}_{m,l}. \tag{110}
$$

The expansion coefficients of $\mathbf{T}_{p,q}^{(\mathrm{R})}$ are $e^{j2\pi\left(\frac{qm}{\Delta M} - \frac{pl}{\Delta L}\right)} t_{m,l}^{(\mathrm{R})}$ (see (100)); those of $\mathbf{\Gamma}_X$ are $\frac{1}{\sqrt{N}} \bar{A}_X[m,l]$ (see (86)). Inserting these expressions into (109) yields

$$
d_{p,q;m,l} = \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} \left[e^{j2\pi\left(\frac{qm'}{\Delta M} - \frac{pl'}{\Delta L}\right)} t_{m',l'}^{(\mathrm{R})} \right]
$$
$$
\times \left[\frac{1}{\sqrt{N}} \bar{A}_X[m-m',l-l'] \right] e^{-j\frac{2\pi}{N}m'(l-l')}
$$
$$
\approx \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} e^{j2\pi\left(\frac{qm'}{\Delta M} - \frac{pl'}{\Delta L}\right)} I_{\mathcal{A}}[m',l']
$$
$$
\times \bar{A}_X[m-m',l-l']\, e^{-j\frac{2\pi}{N}m'(l-l')}, \tag{111}
$$

where the approximate expression (104) was used in the last step. For an underspread process, because of the support of $I_{\mathcal{A}}[m,l]$ and the effective support of $\bar{A}_X[m,l]$, the terms in the sum (111) are significantly nonzero only for $|m'(l-l')| \ll N$. We can thus use the approximation $e^{-j\frac{2\pi}{N}m'(l-l')} \approx 1$, which yields

$$
d_{p,q;m,l} \approx \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\, \bar{A}_X[m-m',l-l']
$$
$$
\times e^{j2\pi\left(\frac{qm'}{\Delta M} - \frac{pl'}{\Delta L}\right)}. \tag{112}
$$

Next, we consider

$$
\left(\mathbf{T}_{p,q}^{(\mathrm{R})} \mathbf{\Gamma}_X\right)^{H} \stackrel{(110)}{=} \sum_{m,l\in[N]} d_{p,q;m,l}^{*} \mathbf{J}_{m,l}^{H}
$$
$$
\stackrel{(88)}{=} \sum_{m,l\in[N]} d_{p,q;m,l}^{*} \mathbf{J}_{-m,-l}\, e^{-j\frac{2\pi}{N}ml}
$$
$$
= \sum_{m,l\in[N]} d_{p,q;-m,-l}^{*} \mathbf{J}_{m,l}\, e^{-j\frac{2\pi}{N}ml}, \tag{113}
$$

where the N-periodicity of $d_{p,q;m,l}$ with respect to m and l was used in the last step. For an underspread process, again

because of the support of $I_{\mathcal{A}}[m,l]$ and the effective support of $\bar{A}_X[m,l]$, it follows from (112) that the coefficients $d_{p,q;m,l}$ are significantly nonzero only for $|ml| \ll N$. Hence, we can set $e^{-j\frac{2\pi}{N}ml} \approx 1$ in (113), which gives

$$\left(\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\right)^H \approx \sum_{m,l\in[N]} d_{p,q;-m,-l}^* \, \mathbf{J}_{m,l}\,. \tag{114}$$

In a similar way, we obtain from (36b) the following approximation:

$$\bar{A}_X^*[-m,-l] \approx \bar{A}_X[m,l]\,, \quad \text{for } |ml| \ll N\,. \tag{115}$$

We now insert (110) and (114) into (108), and obtain

$$\operatorname{tr}\left\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\right\}$$
$$\approx \left\langle \sum_{m,l\in[N]} d_{p,q;m,l}\,\mathbf{J}_{m,l}\,,\, \sum_{m',l'\in[N]} d_{p,q;-m',-l'}^*\,\mathbf{J}_{m',l'} \right\rangle$$
$$\stackrel{(83)}{=} \sum_{m,l\in[N]} d_{p,q;m,l}\,d_{p,q;-m,-l}\,. \tag{116}$$

From the underspread approximations (112) and (115), it readily follows that $d_{p,q;-m,-l} \approx d_{p,q;m,l}^*$. Indeed,

$$d_{p,q;-m,-l} \stackrel{(112)}{\approx} \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\,\bar{A}_X[-m-m',-l-l']$$
$$\times e^{j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$\stackrel{(115)}{\approx} \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\,\bar{A}_X^*[m+m',l+l']$$
$$\times e^{j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$\stackrel{(*)}{=} \frac{1}{\sqrt{N}} \sum_{m'=-M}^{M}\sum_{l'=-L}^{L} \bar{A}_X^*[m+m',l+l']$$
$$\times e^{j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$= \frac{1}{\sqrt{N}} \sum_{m'=-M}^{M}\sum_{l'=-L}^{L} \bar{A}_X^*[m-m',l-l']$$
$$\times e^{-j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$\stackrel{(*)}{=} \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\,\bar{A}_X^*[m-m',l-l']$$
$$\times e^{-j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$\stackrel{(112)}{\approx} d_{p,q;m,l}^*\,, \tag{117}$$

where the periodicity of the summand with respect to m' and l' has been exploited in the steps labeled with $(*)$. Using (117) in (116) then gives

$$\operatorname{tr}\left\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\right\} \approx \sum_{m,l\in[N]} |d_{p,q;m,l}|^2\,. \tag{118}$$

3) Second term in (107)*:* Next, we consider the second term on the right-hand side of (107). We have

$$\operatorname{tr}\left\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^H\right\} \stackrel{(71)}{=} \operatorname{tr}\left\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{R})}\right\} + j\operatorname{tr}\left\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{I})}\right\}$$
$$\stackrel{(106)}{\approx} \operatorname{tr}\left\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^{(\mathrm{R})}\right\}$$
$$= \operatorname{tr}\left\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\right\}\,. \tag{119}$$

Using $\mathbf{J}_{0,0} = \frac{1}{\sqrt{N}}\mathbf{I}$, where $\mathbf{I}$ denotes the $N \times N$ identity matrix, we obtain further

$$\operatorname{tr}\left\{\mathbf{\Gamma}_X\mathbf{T}_{p,q}^H\right\} \stackrel{(119)}{\approx} \operatorname{tr}\left\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\mathbf{I}\right\}$$
$$= \sqrt{N}\,\operatorname{tr}\left\{\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\mathbf{J}_{0,0}^H\right\}$$
$$= \sqrt{N}\,\left\langle\mathbf{T}_{p,q}^{(\mathrm{R})}\mathbf{\Gamma}_X\,,\,\mathbf{J}_{0,0}\right\rangle$$
$$\stackrel{(110)}{=} \sqrt{N}\,\left\langle \sum_{m,l\in[N]} d_{p,q;m,l}\,\mathbf{J}_{m,l}\,,\,\mathbf{J}_{0,0}\right\rangle$$
$$\stackrel{(83)}{=} \sqrt{N}\,d_{p,q;0,0}\,. \tag{120}$$

4) Approximation of $h_{p,q}$: Inserting (118) and (120) into (107), we obtain the following approximation of $h_{p,q}$:

$$h_{p,q} \approx \sum_{m,l\in[N]} |d_{p,q;m,l}|^2 + N|d_{p,q;0,0}|^2\,.$$

This can be expressed as

$$h_{p,q} \approx \sum_{n,k\in[N]} |\hat{d}_{p,q;n,k}|^2 + N\left|\frac{1}{N}\sum_{n,k\in[N]} \hat{d}_{p,q;n,k}\right|^2, \tag{121}$$

where $\hat{d}_{p,q;n,k}$ is the 2D DFT of $d_{p,q;m,l}$ with respect to (m,l). We have

$$\hat{d}_{p,q;n,k}$$
$$= \frac{1}{N} \sum_{m,l\in[N]} d_{p,q;m,l}\,e^{-j\frac{2\pi}{N}(km-nl)}$$
$$\stackrel{(112)}{\approx} \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\left[\frac{1}{N}\sum_{m,l\in[N]} \bar{A}_X[m-m',l-l']\right.$$
$$\left.\times e^{-j\frac{2\pi}{N}(km-nl)}\right]e^{j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$\stackrel{(5)}{=} \frac{1}{\sqrt{N}} \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']\,\bar{R}_X[n,k]$$
$$\times e^{-j\frac{2\pi}{N}(km'-nl')}\,e^{j2\pi\left(\frac{qm'}{\Delta M}-\frac{pl'}{\Delta L}\right)}$$
$$= \frac{1}{\sqrt{N}}\,\bar{R}_X[n,k] \sum_{m',l'\in[N]} I_{\mathcal{A}}[m',l']$$
$$\times e^{-j\frac{2\pi}{N}\left[\left(k-\frac{N}{\Delta M}q\right)m'-\left(n-\frac{N}{\Delta L}p\right)l'\right]}$$
$$\stackrel{(43)}{=} \sqrt{N}\,\bar{R}_X[n,k]\,\Phi_{\mathrm{MVU}}[n-p\Delta n,k-q\Delta k]\,, \tag{122}$$

where, as before, $\Delta n = N/\Delta L$ and $\Delta k = N/\Delta M$. Finally, inserting (122) into (121) yields (73).

ACKNOWLEDGMENT

REFERENCES

[1] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs (NJ): Prentice Hall, 1988.

[2] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs (NJ): Prentice Hall, 1997.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communication," *IEEE J. Sel. Areas Comm.*, vol. 23, pp. 201–220, Feb. 2005.

[4] D. Cabric, I. O'Donnell, M.-W. Chen, and R. Brodersen, "Spectrum sharing radios," *IEEE Circuits and Systems Magazine*, vol. 6, no. 2, pp. 30–45, 2006.

[5] A. Sahai, S. Mishra, R. Tandra, and K. Woyach, "Cognitive radios for spectrum sharing [applications corner]," *IEEE Signal Processing Magazine*, vol. 26, pp. 140–145, Jan. 2009.

[6] R. M. Loynes, "On the concept of the spectrum for non-stationary processes," *J. Roy. Stat. Soc. Ser. B*, vol. 30, no. 1, pp. 1–30, 1968.

[7] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. San Diego (CA): Academic Press, 1999.

[8] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 1461–1470, Dec. 1985.

[9] P. Flandrin and W. Martin, "The Wigner-Ville spectrum of nonstationary random signals," in *The Wigner Distribution — Theory and Applications in Signal Processing* (W. Mecklenbräuker and F. Hlawatsch, eds.), pp. 211–267, Amsterdam (The Netherlands): Elsevier, 1997.

[10] G. Matz and F. Hlawatsch, "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1067–1086, March 2006.

[11] C. H. Page, "Instantaneous power spectra," *J. Appl. Phys.*, vol. 23, pp. 103–106, Jan. 1952.

[12] M. J. Levin, "Instantaneous spectra and ambiguity functions," *IEEE Trans. Inf. Theory*, vol. 10, pp. 95–97, Jan. 1964.

[13] M. B. Priestley, "Evolutionary spectra and non-stationary processes," *J. Roy. Stat. Soc. Ser. B*, vol. 27, no. 2, pp. 204–237, 1965.

[14] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Inf. Theory*, vol. 14, pp. 369–374, May 1968.

[15] W. D. Mark, "Spectral analysis of the convolution and filtering of non-stationary stochastic processes," *J. Sound Vib.*, vol. 11, no. 1, pp. 19–63, 1970.

[16] M. B. Priestley, "Some notes on the physical interpretation of spectra of nonstationary stochastic processes," *J. Sound Vib.*, vol. 17, pp. 51–54, 1971.

[17] D. Tjøstheim, "Spectral generating operators for non-stationary processes," *Adv. Appl. Prob.*, vol. 8, pp. 831–846, 1976.

[18] G. Mélard, "Propriétés du spectre évolutif d'un processus non-stationnaire," *Ann. Inst. H. Poincaré B*, vol. XIV, no. 4, pp. 411–424, 1978.

[19] O. D. Grace, "Instantaneous power spectra," *J. Acoust. Soc. Amer.*, vol. 69, no. 1, pp. 191–198, 1981.

[20] G. Mélard and A. de Schutter-Herteleer, "Contributions to evolutionary spectral theory," *J. Roy. Stat. Soc. Ser. B*, vol. 10, pp. 41–63, 1989.

[21] P. Flandrin, "Time-dependent spectra for nonstationary stochastic processes," in *Time and Frequency Representation of Signals and Systems* (G. Longo and B. Picinbono, eds.), pp. 69–124, Vienna, Austria: Springer, 1989.

[22] C. S. Detka and A. El-Jaroudi, "The transitory evolutionary spectrum," in *Proc. IEEE ICASSP-94*, (Adelaide, Australia), pp. 289–292, April 1994.

[23] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 45, pp. 1520–1534, June 1997.

[24] G. Matz, *A time-frequency calculus for time-varying systems and nonstationary processes with applications*. PhD thesis, Vienna University of Technology, Nov. 2000.

[25] G. Matz and F. Hlawatsch, "Time-varying power spectra of nonstationary random processes," in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference* (B. Boashash, ed.), ch. 9.4, pp. 400–409, Oxford (UK): Elsevier, 2003.

[26] G. Matz and F. Hlawatsch, "Time-frequency methods for non-stationary statistical signal processing," in *Time-Frequency Analysis: Concepts and Methods* (F. Hlawatsch and F. Auger, eds.), pp. 279–320, London, UK: ISTE/Wiley, 2008.

[27] W. Kozek, F. Hlawatsch, H. Kirchauer, and U. Trautwein, "Correlative time-frequency analysis and classification of nonstationary random processes," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Philadelphia, PA), pp. 417–420, Oct. 1994.

[28] W. Kozek, "On the underspread/overspread classification of nonstationary random processes," in *Proc. Int. Conf. Industrial and Applied Mathematics* (K. Kirchgässner, O. Mahrenholtz, and R. Mennicken, eds.), vol. 3 of *Mathematical Research*, (Berlin, Germany), pp. 63–66, Akademieverlag, 1996.

[29] W. Kozek, *Matched Weyl-Heisenberg expansions of nonstationary environments*. PhD thesis, Vienna University of Technology, March 1997.

[30] W. Kozek, "Adaptation of Weyl-Heisenberg frames to underspread environments," in *Gabor Analysis and Algorithms: Theory and Applications* (H. G. Feichtinger and T. Strohmer, eds.), ch. 10, pp. 323–352, Boston (MA): Birkhäuser, 1998.

[31] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs (NJ): Prentice Hall, 1995.

[32] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[33] A. M. Sayeed and D. L. Jones, "Optimal kernels for nonstationary spectral estimation," *IEEE Trans. Signal Processing*, vol. 43, pp. 478–491, Feb. 1995.

[34] L. L. Scharf, B. Friedlander, P. Flandrin, and A. Hanssen, "The Hilbert space geometry of the stochastic Rihaczek distribution," in *Proc. 35th Asilomar Conf. Sig., Sys., Comput.*, (Pacific Grove, CA), pp. 720–725, Nov. 2001.

[35] L. L. Scharf, B. Friedlander, and D. J. Thomson, "Covariant estimators of time-frequency descriptors for nonstationary random processes," in *Proc. 32nd Asilomar Conf. Sig., Sys., Comput.*, (Pacific Grove, CA), pp. 808–811, Nov. 1998.

[36] L. L. Scharf, P. J. Schreier, and A. Hanssen, "The Hilbert space geometry of the Rihaczek distribution for stochastic analytic signals," *IEEE Signal Processing Letters*, vol. 12, pp. 297–300, April 2005.

[37] M. G. Amin, "Time-frequency spectrum analysis and estimation for non-stationary random processes," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), pp. 208–232, Melbourne, Australia: Longman Cheshire, 1992.

[38] W. Kozek and K. Riedel, "Quadratic time-varying spectral estimation for underspread processes," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Philadelphia, PA), pp. 460–463, Oct. 1994.

[39] M. Bayram and R. G. Baraniuk, "Multiple window time-frequency analysis," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Paris, France), pp. 173–176, June 1996.

[40] M. Bayram and R. G. Baraniuk, "Multiple window time-varying spectrum estimation," in *Nonlinear and Nonstationary Signal Processing* (W. Fitzgerald, R. Smith, A. Walden, and P. Young, eds.), pp. 292–316, Cambridge, UK: Cambridge Univ. Press, 2000.

[41] F. Cakrak and P. J. Loughlin, "Multiple window time-varying spectral analysis," *IEEE Trans. Signal Processing*, vol. 49, pp. 448–453, Feb. 2001.

[42] A. Jung, G. Tauböck, and F. Hlawatsch, "Compressive nonstationary spectral estimation using parsimonious random sampling of the ambiguity function," in *Proc. IEEE-SSP 2009*, (Cardiff, Wales, UK), pp. 642–645, Aug.–Sep. 2009.

[43] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, April 2006.

[44] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, pp. 1207–1223, Aug. 2006.

[45] P. Borgnat and P. Flandrin, "Time-frequency localization from sparsity constraints," in *Proc. IEEE ICASSP-2008*, (Las Vegas, NV), pp. 3785–3788, Apr. 2008.

[46] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Trans. Signal Processing*, vol. 58, pp. 2974–2982, June 2010.

[47] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *Proc. IEEE GLOBECOM 2008*, (New Orleans, LA), pp. 1–5, Dec. 2008.

[48] Y. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wideband spectrum sensing," in *Proc. IEEE ICASSP-2009*, (Taipei, Taiwan), pp. 2337–2340, Apr. 2009.

[49] Z. Tian, Y. Tafesse, and B. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE J. Sel. Topics Sig. Proc.*, vol. 6, pp. 58–69, Feb. 2012.

[50] G. Leus and D. Ariananda, "Power spectrum blind sampling," *IEEE Signal Processing Letters*, vol. 18, pp. 443–446, Aug. 2011.

[51] C. Tu, B. Liu, and H. Chen, "An analog correlator for ultra-wideband receivers," *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 455–461, Jan. 2005.

[52] D. Casasent and B. V. K. V. Kumar, "Optical image plane correlator for ambiguity surface computation," *Appl. Opt.*, vol. 18, pp. 1673–1678, May 1979.

[53] R. J. Marks II, J. F. Walkup, and T. F. Krile, "Ambiguity function display: an improved coherent processor," *Appl. Opt.*, vol. 16, pp. 746–750, Mar 1977.

[54] S. Padin, J. K. Cartwright, M. C. Shepherd, J. K. Yamasaki, and W. L. Holzapfel, "A wideband analog correlator for microwave background observations," *IEEE Trans. Instr. Meas.*, vol. 50, pp. 1234–1240, Oct. 2001.

[55] F. Zhao and L. J. Guibas, *Wireless Sensor Networks: An Information Processing Approach*. Amsterdam, The Netherlands: Morgan Kaufmann, 2004.

[56] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," in *Proc. Information Theory and Applications Workshop, 2008*, (Urbana, IL), pp. 547–554, Jan.-Feb. 2008.

[57] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Sig. Proc.*, vol. 4, no. 2, pp. 375–390, 2010.

[58] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 3rd ed., 1996.

[59] M. Jachan, G. Matz, and F. Hlawatsch, "Time-frequency ARMA models and parameter estimators for underspread nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 55, pp. 4366–4381, Sep. 2007.

[60] E. L. Lehmann and G. Casella, *Theory of Point Estimation*. New York: Springer, 2nd ed., 1998.

[61] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements," in *Proc. 40th Annual Conf. Inform. Sci. Syst.*, (Princeton, NJ), pp. 207–212, March 2006.

[62] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scient. Comput.*, vol. 20, pp. 33–61, 1998.

[63] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming, version 2.0," 2012.

[64] A. Jung, G. Tauböck, and F. Hlawatsch, "Compressive spectral estimation for nonstationary random processes," in *Proc. IEEE ICASSP-2009*, (Taipei, Taiwan, R.O.C.), pp. 3029–3032, April 2009.

[65] G. G. Tziritas, "On the distribution of positive-definite Gaussian quadratic forms," *IEEE Trans. Inf. Theory*, vol. 33, pp. 895–906, Nov. 1987.

[66] J. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Comm. Mag.*, vol. 28, pp. 5–14, May 1990.

[67] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. IEEE Press and Wiley, 2003.

[68] IEEE, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Standard 802.11 [Online].

[69] U. Reimers, "Digital video broadcasting," *IEEE Comm. Mag.*, vol. 36, pp. 104–110, June 1998.

[70] J. Stott, "The DVB terrestrial (DVB-T) specification and its implementation in a practical modem," in *Proc. International Broadcasting Convention (Conf. Publ. No. 428)*, (Amsterdam, The Netherlands), pp. 255–260, Sep. 1996.

[71] ETSI, "Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television." EN 300 744, V1.4.1, 2001 (http://www.etsi.org).

[72] H. Holma and A. Toskala, eds., *LTE for UMTS – OFDMA and SC-FDMA Based Radio Access*. Chichester, UK: Wiley, 2009.

[73] L. Stankovic and J. F. Böhme, "Time-frequency analysis of multiple resonances in combustion engine signals," *Signal Processing*, vol. 79, pp. 15–28, Nov. 1999.

[74] S. Carstens-Behrens, M. Wagner, and J. F. Böhme, "Detection of multiple resonances in noise," *Int. J. Electron. Commun. (AEÜ)*, vol. 52, no. 5, pp. 285–292, 1998.

[75] G. Matz and F. Hlawatsch, "Time-frequency subspace detectors and application to knock detection," *Int. J. Electron. Commun. (AEÜ)*, vol. 53, no. 6, pp. 379–385, 1999.

[76] X.-G. Xia, "System identification using chirp signals and time-variant filters in the joint time-frequency domain," *IEEE Trans. Signal Processing*, vol. 45, pp. 2072–2084, Aug. 1997.

[77] S. Salous, N. Nikandrou, and N. Bajj, "Digital techniques for mobile radio chirp sounders," *IEE Proc. Commun.*, vol. 145, pp. 191–196, June 1998.

[78] A. W. Rihaczek, *Principles of High-Resolution Radar*. New York: McGraw Hill, 1969.

[79] M. Skolnik, *Radar Handbook*. New York: McGraw-Hill, 1984.

[80] Y. Grenier, "Non-stationary signal modelling with application to bat echolocation calls," *ACUSTICA*, vol. 61, pp. 156–165, Sept. 1986.

[81] P. Flandrin, "Time-frequency processing of bat sonar signals," in *Animal Sonar: Processes and Performance* (P. Nachtigall and P. Moore, eds.), pp. 797–802, New York: Plenum Press, 1988.

[82] F. Hlawatsch and P. Flandrin, "The interference structure of the Wigner distribution and related time-frequency signal representations," in *The Wigner Distribution — Theory and Applications in Signal Processing* (W. Mecklenbräuker and F. Hlawatsch, eds.), pp. 59–133, Amsterdam, The Netherlands: Elsevier, 1997.

[83] W. Rudin, *Real and Complex Analysis*. New York: McGraw-Hill, 3rd ed., 1987.

[84] G. B. Folland, *Harmonic Analysis in Phase Space*, vol. 122 of *Annals of Mathematics Studies*. Princeton (NJ): Princeton University Press, 1989.

[85] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Boston, MA: Birkhäuser, 2001.

Alexander Jung received the Diplom-Ingenieur and Dr. techn. degrees in electrical engineering from Vienna University of Technology, Vienna, Austria, in 2008 and 2011, respectively. Since 2008, he has been a Research Assistant with the Institute of Telecommunications, Vienna University of Technology. His research interests are in statistical signal processing with emphasis on sparse estimation problems. He received several national awards and a Best Student Paper Award at IEEE ICASSP 2011.

Georg Tauböck (S'01–M'07) received the Dipl.-Ing. degree and the Dr. techn. degree (with highest honors) in electrical engineering and the Dipl.-Ing. degree in mathematics (with highest honors) from Vienna University of Technology, Vienna, Austria in 1999, 2005, and 2008, respectively. He also received the diploma in violoncello from the Conservatory of Vienna, Vienna, Austria, in 2000.

From 1999 to 2005, he was with the FTW Telecommunications Research Center Vienna, Vienna, Austria, and since 2005, he has been with the Institute of Telecommunications, Vienna University of Technology, Vienna, Austria. From February to August 2010, he was a visiting researcher with the Communication Technology Laboratory/Communication Theory Group at ETH Zurich, Zurich, Switzerland.

His research interests include wireline and wireless communications, compressed sensing, signal processing, and information theory.

Franz Hlawatsch (S'85–M'88–SM'00–F'12) received the Diplom-Ingenieur, Dr. techn., and Univ.-Dozent (habilitation) degrees in electrical engineering/signal processing from Vienna University of Technology, Vienna, Austria in 1983, 1988, and 1996, respectively.

Since 1983, he has been with the Institute of Telecommunications, Vienna University of Technology, where he is currently an Associate Professor. During 1991–1992, as a recipient of an Erwin Schrödinger Fellowship, he spent a sabbatical year with the Department of Electrical Engineering, University of Rhode Island, Kingston, RI, USA. In 1999, 2000, and 2001, he held one-month Visiting Professor positions with INP/ENSEEIHT, Toulouse, France and IRCCyN,

Nantes, France. He (co)authored a book, two review papers that appeared in the IEEE SIGNAL PROCESSING MAGAZINE, about 200 refereed scientific papers and book chapters, and three patents. He coedited three books. His research interests include signal processing for wireless communications and sensor networks, statistical signal processing, and compressive signal processing.

Prof. Hlawatsch was Technical Program Co-Chair of EUSIPCO 2004 and served on the technical committees of numerous IEEE conferences. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2003 to 2007 and for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2008 to 2011. From 2004 to 2009, he was a member of the IEEE SPCOM Technical Committee. He is coauthor of papers that won an IEEE Signal Processing Society Young Author Best Paper Award and a Best Student Paper Award at IEEE ICASSP 2011.