in Proc. IEEE ICASSP 2009
Taipei, Taiwan, R.O.C., April 2009, pp. 3029-3032

COMPRESSIVE SPECTRAL ESTIMATION FOR NONSTATIONARY RANDOM PROCESSES

Alexander Jung, Georg Tauböck, and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria; e-mail: ajung@nt.tuwien.ac.at

ABSTRACT

We propose a "compressive" estimator of the Wigner-Ville spectrum (WVS) for time-frequency sparse, underspread, nonstationary random processes. A novel WVS estimator involving the signal's Gabor coefficients on an undersampled time-frequency grid is combined with a compressed sensing transformation in order to reduce the number of measurements required. The performance of the compressive WVS estimator is analyzed via a bound on the mean square error and through simulations. We also propose an efficient implementation using a special construction of the measurement matrix.

Index Terms— Nonstationary spectral estimation, Wigner-Ville spectrum, Gabor expansion, compressed sensing, sparse reconstruction, basis pursuit

1. INTRODUCTION

The recently introduced methodology of *compressed sensing* (CS) enables the efficient reconstruction of sparse signals from a small number of measurements [1]. Here, we apply CS to nonstationary spectral estimation from a single observed process realization. We first present a spectral estimator for underspread nonstationary processes that is derived from the Gabor coefficients on an undersampled time-frequency (TF) grid. Based on a "TF sparsity" assumption, we then introduce a CS transformation that achieves a compression in the measurement space. This is useful if the measurements are transmitted over a channel or if dedicated measurement devices are available: one obtains a bit-rate reduction in the first case and a reduction of the number of measurement devices in the second.

Let $X(t)$ be a nonstationary, zero-mean, circularly symmetric complex random process with autocorrelation $r_X(t_1,t_2) \triangleq \mathrm{E}\{X(t_1)X^*(t_2)\}$ and finite mean energy $\bar{E}_X \triangleq \int_t r_X(t,t)\,dt$ (integrals and sums are from $-\infty$ to ∞ unless noted otherwise). We wish to estimate the Wigner-Ville spectrum (WVS) [2]

$$\overline{W}_X(t,f) \triangleq \int_\tau r_X\Big(t+\frac{\tau}{2}, t-\frac{\tau}{2}\Big)\, e^{-j2\pi f\tau} d\tau\,.$$

The WVS can be viewed as a "time-dependent power spectrum" if $X(t)$ is an *underspread* process, i.e., if process components that are not too close in the TF plane are approximately uncorrelated [3]. Existing WVS estimators for underspread processes include TF smoothed versions of the Wigner distribution of an observed realization $x(t)$ of $X(t)$ [2, 4] and estimators using a local cosine basis expansion of $x(t)$ [5]. However, these estimators do not perform a compression of the measurements.

Our contributions and the organization of this paper can be summarized as follows. In Section 2, we present a WVS estimator that involves the signal's Gabor coefficients on an undersampled TF grid. In Section 3, we introduce a CS extension of the estimator and discuss its efficient implementation. A bound on the mean square error (MSE) of the resulting compressive WVS estimator is provided in Section 4, and simulation results are presented in Section 5.

This work was supported by the FWF project "Statistical Inference" (S10603-N13) within the National Research Network SISE and by the WWTF project SPORTS (MA 07-004).

2. WVS ESTIMATION BASED ON GABOR ANALYSIS

We assume that the nonstationary process $X(t)$ is underspread, which means that it has small "TF correlation moments" [3, 6]

$$m_X^{(\phi)} \triangleq \frac{\int_\tau\int_\nu \phi(\tau,\nu)\,|\bar{A}_X(\tau,\nu)|\,d\tau\,d\nu}{\int_\tau\int_\nu |\bar{A}_X(\tau,\nu)|\,d\tau\,d\nu}\,,$$

$$M_X^{(\phi)} \triangleq \frac{\int_\tau\int_\nu \phi^2(\tau,\nu)\,|\bar{A}_X(\tau,\nu)|^2\,d\tau\,d\nu}{\int_\tau\int_\nu |\bar{A}_X(\tau,\nu)|^2\,d\tau\,d\nu}\,. \tag{1}$$

Here, $\bar{A}_X(\tau,\nu) \triangleq \int_t r_X\big(t+\frac{\tau}{2}, t-\frac{\tau}{2}\big)\, e^{-j2\pi\nu t} dt$ is the expected ambiguity function (EAF) of $X(t)$ and $\phi(\tau,\nu)$ is a suitable weighting function that generally increases for increasing values of delay τ and frequency lag ν. The EAF is related to the WVS by a 2-D Fourier transform, i.e., $\bar{A}_X(\tau,\nu) = \int_t\int_f \overline{W}_X(t,f)\, e^{-j2\pi(\nu t-\tau f)}\,dt\,df$. For an underspread process $X(t)$, $\bar{A}_X(\tau,\nu)$ is well concentrated around the origin of the (τ,ν) plane. This implies that $m_X^{(\phi)}$ and $M_X^{(\phi)}$ are small and, also, that the WVS is a smooth function.

2.1. Definition of the Basic WVS Estimator

A WVS estimator that uses Gabor coefficients of the observed realization $x(t)$ is motivated as follows. Since the WVS $\overline{W}_X(t,f)$ of an underspread process is a smooth function, it can be approximated by interpolating its samples $\big\{\overline{W}_X(kT,lF)\big\}_{k,l\in\mathbb{Z}}$ taken on an appropriate TF grid. That is, the TF function

$$\widetilde{W}_X(t,f) \triangleq \sum_k\sum_l \overline{W}_X(kT,lF)\,\phi(t-kT, f-lF) \tag{2}$$

will approximate $\overline{W}_X(t,f)$ if $\phi(t,f)$, T, and F are suitably chosen [6, Theorem 2.35]. In particular, if the EAF $\bar{A}_X(\tau,\nu)$ is zero for (τ,ν) outside a rectangle $[-\tau_{\max},\tau_{\max}]\times[-\nu_{\max},\nu_{\max}]$, the approximation will be exact provided $\phi(t,f)$ is the appropriate 2-D sinc kernel and the non-aliasing conditions $T \le 1/(2\nu_{\max})$ and $F \le 1/(2\tau_{\max})$ are satisfied.

We now define a WVS estimator $\widehat{W}_X(t,f)$ by replacing the WVS samples $\overline{W}_X(kT,lF)$ in (2) with estimates. Here, these estimates are chosen as the squared magnitudes of the Gabor coefficients [7] of an observed realization $x(t)$ using T,F as TF grid constants,

$$c_x^{(k,l)} \triangleq \int_t x(t)\,g^*(t-kT)\,e^{-j2\pi lFt}\,dt\,,$$

with a suitably chosen normalized window $g(t)$ (i.e., $\|g\|_2 = 1$). The WVS estimator is thus obtained as

$$\widehat{W}_X(t,f) \triangleq \sum_k\sum_l \big|c_x^{(k,l)}\big|^2\,\psi(t-kT, f-lF)\,, \tag{3}$$

where $\psi(t,f)$ may be different from $\phi(t,f)$ in (2). We note that for $X(t)$ underspread, $\mathrm{E}\big\{\big|c_X^{(k,l)}\big|^2\big\} \approx \overline{W}_X(kT,lF)$ [6]. Thus, $\big|c_x^{(k,l)}\big|^2$ is (approximately) an unbiased estimator of $\overline{W}_X(kT,lF)$. An expression of the MSE of $\widehat{W}_X(t,f)$ will be provided in Section 4.

2.2. Design of the Basic WVS Estimator

Next, we discuss the choice of the design parameters $\psi(t,f)$, T, F, and $g(t)$ involved in $\widehat{W}_X(t,f)$. We assume that $\bar{A}_X(\tau,\nu)=0$ for (τ,ν) outside a centered rectangular support region $\mathcal{S} = [-\tau_{\max},\tau_{\max}]\times[-\nu_{\max},\nu_{\max}]$, which is supposed known. The *TF correlation spread* is defined as $s_X \triangleq |\mathcal{S}| = 4\tau_{\max}\nu_{\max}$.

Choice of $\boldsymbol{\psi(t,f)}$. Inspired by [4], we design $\psi(t,f)$ such that $\widehat{W}_X(t,f)$ is approximately a minimum variance unbiased (MVU) estimator. The MSE $\varepsilon \triangleq \mathrm{E}\big\{\big\|\widehat{W}_X - \overline{W}_X\big\|_2^2\big\}$ can be written as $\varepsilon = B^2+V$ with the squared bias $B^2 \triangleq \big\|\mathrm{E}\{\widehat{W}_X\} - \overline{W}_X\big\|_2^2$ and the variance $V \triangleq \mathrm{E}\big\{\big\|\widehat{W}_X - \mathrm{E}\{\widehat{W}_X\}\big\|_2^2\big\}$. The MVU design minimizes V under the constraint $B^2=0$. Closed-form expressions of B^2 and V can be obtained under the following approximations: (A1) $\mathrm{E}\big\{\big|c_X^{(k,l)}\big|^2\big\} = \overline{W}_X(kT,lF)$; (A2) the $c_X^{(k,l)}$ are uncorrelated. Both approximations are better satisfied if $X(t)$ is more underspread (smaller s_X); approximation A2 will also be better satisfied for a larger value of TF. Using these approximations, one can show

$$B^2 = \int_\tau\int_\nu |\bar{A}_X(\tau,\nu)|^2 \left|1-\frac{1}{TF}\Psi(\tau,\nu)\right|^2 d\tau\, d\nu \qquad (4)$$

$$V \,=\, \frac{1}{TF}\, \|r_X\|_2^2 \,\|\Psi\|_2^2\,, \qquad (5)$$

where $\Psi(\tau,\nu) \triangleq \int_t\int_f \psi(t,f)\, e^{-j2\pi(\nu t - \tau f)}\, dt\, df$. For (5), we assumed $X(t)$ to be Gaussian and used Isserlis' formula [4]. Note that $r_X(t_1,t_2)$ and $\bar{A}_X(\tau,\nu)$ are unknown but $\mathcal{S}$ is assumed known.

To obtain $B^2=0$ in (4), $\Psi(\tau,\nu)$ must be equal to TF on $\mathcal{S}$. Then, V in (5) is minimized if $\Psi(\tau,\nu)=0$ outside $\mathcal{S}$, because this minimizes $\|\Psi\|_2^2$. Thus, the 2-D Fourier transform of the approximately MVU-optimum interpolation function $\psi_{\mathrm{MVU}}(t,f)$ is obtained as

$$\Psi_{\mathrm{MVU}}(\tau,\nu) \,=\, \begin{cases} TF, & (\tau,\nu)\in\mathcal{S} \\ 0\,, & \text{elsewhere,}\end{cases} \qquad (6)$$

which means that $\psi_{\mathrm{MVU}}(t,f)$ is a 2-D sinc-type function. The resulting approximate MSE is obtained as (recall that $B^2=0$)

$$\varepsilon \,=\, V \,=\, TF s_X \|r_X\|_2^2\,. \qquad (7)$$

Choice of $\boldsymbol{T,F}$. The TF grid constants T and F are constrained by the non-aliasing conditions $T\le 1/(2\nu_{\max})$ and $F\le 1/(2\tau_{\max})$. These conditions imply $TF\le 1/s_X$; thus, for a strongly underspread process (where $s_X \ll 1$), we may choose $TF>1$ ("undersampled" TF grid). The MSE expression (7) is minimized by choosing TF as small as possible. However, if T and/or F are too small, our assumption A2 (uncorrelated $c_X^{(k,l)}$) will be violated. To obtain small correlation of the $c_X^{(k,l)}$, the TF grid geometry defined by T,F should be matched to the correlation TF geometry of $X(t)$ as characterized by the EAF support $\mathcal{S} = [-\tau_{\max},\tau_{\max}]\times[-\nu_{\max},\nu_{\max}]$. This is achieved by letting $T/F = \tau_{\max}/\nu_{\max}$ [8].

Choice of $\boldsymbol{g(t)}$. To obtain small correlation of the $c_X^{(k,l)}$, the TF geometry of the Gabor analysis window $g(t)$ should be matched to the TF grid geometry defined by T,F. This is achieved by letting $T_g/F_g = T/F$ $(=\tau_{\max}/\nu_{\max})$, where T_g and F_g are the effective duration and bandwidth, respectively, of $g(t)$.

3. COMPRESSIVE WVS ESTIMATION

In what follows, we assume $X(t)$ to be "TF sparse," i.e., most WVS values within the total TF region considered are almost zero. In many applications, the WVS $\overline{W}_X(t,f)$ is effectively supported in a few relatively small regions of the TF plane (corresponding to TF localized signal components) and thus almost zero in the rest of the TF plane. We do not know which WVS values are effectively nonzero, only an estimate of their total number is assumed known.

3.1. Definition of the Compressive WVS Estimator

When $X(t)$ is both TF sparse and underspread, it follows from $\mathrm{E}\big\{\big|c_X^{(k,l)}\big|^2\big\} \approx \overline{W}_X(kT,lF)$ that the Gabor coefficients $c_X^{(k,l)}$ are sparse with high probability. Let us consider a process $X(t)$ with a finite duration and a finite effective bandwidth, corresponding to a rectangular TF region containing $N=KL$ Gabor coefficients $c_X^{(k,l)}$ $(k=1,\ldots,K,\ l=1,\ldots,L)$. We assume that at most $S\ll N$ of these Gabor coefficients are nonzero (note that we do not know which are nonzero). For an observed realization $x(t)$, let $\mathbf{c}\in\mathbb{C}^N$ be the vector of all N Gabor coeffients $c_x^{(k,l)}$, defined, e.g., as $[\mathbf{c}]_{k+(l-1)K} = c_x^{(k,l)}$ for $k=1,\ldots,K$ and $l=1,\ldots,L$. Then $\mathbf{c}$ is an S-*sparse* vector (at most S elements are nonzero), and CS theory tells us that we can "compress" $\mathbf{c}$ by multiplying it by a suitable matrix $\mathbf{\Phi}\in\mathbb{C}^{M\times N}$ with $M\ll N$:

$$\mathbf{z} = \mathbf{\Phi}\mathbf{c}\,. \qquad (8)$$

Note that $\mathbf{z}\in\mathbb{C}^M$ contains much fewer entries than $\mathbf{c}\in\mathbb{C}^N$. If the compressed dimension M and the "measurement matrix" $\mathbf{\Phi}$ are chosen as discussed in Section 3.2, we can recover $\mathbf{c}$ from $\mathbf{z}$ up to a small error by means of the convex program (*basis pursuit*) [9]

$$\hat{\mathbf{c}} \,\triangleq\, \arg\min_{\mathbf{c}'} \|\mathbf{c}'\|_1 \quad \text{subject to } \ \mathbf{\Phi}\mathbf{c}' = \mathbf{z}\,. \qquad (9)$$

From the estimated Gabor coefficients $\hat{c}_x^{(k,l)}$ contained in $\hat{\mathbf{c}}$, we finally obtain a *compressive WVS estimate* by substituting the $\hat{c}_x^{(k,l)}$ for the true Gabor coefficients $c_x^{(k,l)}$ in (3), i.e.,

$$\widehat{W}_{X,\mathrm{CS}}(t,f) \,\triangleq\, \sum_{k=1}^{K}\sum_{l=1}^{L} \big|\hat{c}_x^{(k,l)}\big|^2 \psi_{\mathrm{MVU}}(t-kT, f-lF)\,. \qquad (10)$$

Note that $\widehat{W}_{X,\mathrm{CS}}(t,f)$ is based on the $M\ll N$ "compressed measurements" $z_i = [\mathbf{z}]_i$, $i=1,\ldots,M$ that characterize $x(t)$ as far as is needed for WVS estimation (similar in spirit to a sufficient statistic). The z_i can be calculated directly from $x(t)$ as inner products $z_i = \int_t x(t)\, w_i^*(t)\, dt$ with the measurement functions $w_i(t) = \sum_{k=1}^{K}\sum_{l=1}^{L} [\mathbf{\Phi}]^*_{i,k+(l-1)K}\, g(t-kT)\, e^{j2\pi lFt}$. This is practically interesting if sensor devices implementing inner products are available; due to the compression, only $M\ll N$ such devices are needed.

3.2. Choice of $\mathbf{\Phi}$

CS theory [9] postulates that the measurement matrix $\mathbf{\Phi}\in\mathbb{C}^{M\times N}$ obeys a "restricted isometry hypothesis." Let $\mathbf{\Phi}_{\mathcal{T}}$, $\mathcal{T}\subset\{1,\ldots,N\}$ be the $M\times|\mathcal{T}|$ submatrix comprising the columns of $\mathbf{\Phi}$ indexed by the elements of the index set $\mathcal{T}$. The S-*restricted isometry constant* δ_S of $\mathbf{\Phi}$ is defined as the smallest $\delta>0$ such that

$$(1-\delta)\|\mathbf{a}\|_2^2 \,\le\, \|\mathbf{\Phi}_{\mathcal{T}}\mathbf{a}\|_2^2 \,\le\, (1+\delta)\|\mathbf{a}\|_2^2$$

for all $\mathcal{T}$ with $|\mathcal{T}|\le S$ and all $\mathbf{a}\in\mathbb{C}^{|\mathcal{T}|}$. Using δ_S, the reconstruction error of the basis pursuit (9) is bounded as follows [9]. For a given S, assume that the $3S$- and $4S$-restricted isometry constants of $\mathbf{\Phi}\in\mathbb{C}^{M\times N}$ satisfy

$$\delta_{3S} + 3\delta_{4S} < 2\,. \qquad (11)$$

Let $\mathbf{c}\in\mathbb{C}^N$ (not necessarily sparse) and $\mathbf{z}=\mathbf{\Phi}\mathbf{c}$, and let $\mathbf{c}_S\in\mathbb{C}^N$ contain the S components of $\mathbf{c}$ with largest absolute values, the remaining $N-S$ components of $\mathbf{c}_S$ being zero (thus, $\mathbf{c}_S$ is an S-sparse approximation to $\mathbf{c}$). Then the estimate $\hat{\mathbf{c}}$ in (9) satisfies

$$\|\hat{\mathbf{c}}-\mathbf{c}\|_2 \,\le\, C\,\frac{\|\mathbf{c}-\mathbf{c}_S\|_1}{\sqrt{S}}\,, \tag{12}$$

where the constant C depends only on δ_{3S} and δ_{4S}. In particular, if $\mathbf{c}$ is almost S-sparse (i.e., $\mathbf{c} \approx \mathbf{c}_S$), then (12) shows that the reconstruction error $\hat{\mathbf{c}}-\mathbf{c}$ is small.

It is shown in [10] that if $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$ is constructed by randomly selecting M rows from a unitary $N \times N$ matrix $\mathbf{U}$ and normalizing the columns, a sufficient condition for (11) to be true with overwhelming probability[1] is

$$M \ge C'(\ln N)^4 \mu^2 S\,. \tag{13}$$

Here, $\mu \triangleq \sqrt{N}\max_{i,j}|U_{i,j}|$ (known as the *coherence* of $\mathbf{U}$) and C' is a constant. Thus, we will construct $\mathbf{\Phi}$ by randomly selecting M rows from a unitary matrix $\mathbf{U} \in \mathbb{C}^{N\times N}$. More specifically, we set $\mathbf{\Phi} = \mathbf{PUD}$ where $\mathbf{P} \in \{0,1\}^{M\times N}$ has M entries 1 located randomly with one 1 in each row and at most one 1 in each column, the remaining entries being zero, and $\mathbf{D}$ is a diagonal matrix normalizing the columns of $\mathbf{PU}$.

3.3. Efficient Implementation

An efficient discrete-time implementation of the CS measurement equation (8) can be based on the construction

$$\mathbf{U} = \frac{1}{\sqrt{KL}}\,\mathbf{F}_K \otimes \mathbf{F}_L\,,$$

where, e.g., $\mathbf{F}_K$ is the $K \times K$ IDFT matrix, i.e., $[\mathbf{F}_K]_{k,l} \triangleq e^{j2\pi\frac{(k-1)(l-1)}{K}}$, and $\otimes$ denotes the Kronecker product. This construction allows us to adapt an algorithm described in [11] to our context. Let $\mathbf{x}$ denote a discrete-time (sampled) segment of $x(t)$, obtained with a sufficiently high sampling rate f_s, and let the length-N_g vector $\mathbf{g}$ denote the discrete-time version of the (finite-length) analysis window $g(t)$. We partition $\mathbf{x}$ into overlapping blocks $\mathbf{x}[k]$, $k = 1,\ldots,K$ of length N_g, such that block $\mathbf{x}[k+1]$ is located $N_T = f_sT$ samples after block $\mathbf{x}[k]$. Next, each block $\mathbf{x}[k]$ is multiplied pointwise by the complex conjugate of $\mathbf{g}$ and fed into a pre-alisasing unit that outputs the length-L vector $\mathbf{y}[k]$ with entries $\big[\mathbf{y}[k]\big]_m = \sum_{i=0}^{\lceil N_g/L\rceil}\big[\mathbf{x}[k]\big]_{m+iL}[\mathbf{g}]^*_{m+iL}$, $m = 1,\ldots,L$ (vector entries with invalid indices are considered zero).

The Gabor coefficients $c_x^{(k,l)}$ could be obtained as the length-L DFT of $\mathbf{y}[k]$. However, we can directly calculate the compressed measurement vector $\mathbf{z} = \mathbf{\Phi c}$, with $\mathbf{\Phi} = \mathbf{PUD} = \frac{1}{\sqrt{M}}\mathbf{P}(\mathbf{F}_K \otimes \mathbf{F}_L)$ (note that $[\mathbf{D}]_{i,i} = \sqrt{KL/M}$), according to

$$\mathbf{z} = \frac{L}{\sqrt{M}}\,\mathbf{P}\,\mathrm{vec}\big\{\mathbf{YF}_K\big\}\,. \tag{14}$$

Here, $\mathbf{Y}$ is the $L \times K$ matrix whose columns are the vectors $\mathbf{y}[k]$, $k = 1,\ldots,K$ and $\mathrm{vec}\{\cdot\}$ stacks all columns of a matrix into a vector. Computing (14) requires L 1-D DFTs of length K and a random selection stage (implementing multiplication by $\mathbf{P}$).

4. PERFORMANCE ANALYSIS

MSE of the compressive estimator. We develop an upper bound on the MSE $\varepsilon_{\mathrm{CS}} \triangleq \mathrm{E}\big\{\big\|\widehat{W}_{X,\mathrm{CS}} - \overline{W}_X\big\|_2^2\big\}$ of the compressive WVS estimator $\widehat{W}_{X,\mathrm{CS}}(t,f)$ in (10). An expression for the MSE of the basic WVS estimator $\widehat{W}_X(t,f)$ in (3), $\varepsilon = \mathrm{E}\big\{\big\|\widehat{W}_X - \overline{W}_X\big\|_2^2\big\}$, will be obtained as a by-product. We emphasize that we do not use the approximations A1 and A2 that were used for estimator design in Section 2.2. We merely assume the following: (i) $\bar{A}_X(\tau,\nu) = 0$ for (τ,ν) outside $\mathcal{S} = [-\tau_{\max},\tau_{\max}] \times [-\nu_{\max},\nu_{\max}]$; (ii) $T \le 1/(2\nu_{\max})$ and $F \le 1/(2\tau_{\max})$ (non-aliasing conditions); (iii) $\psi(t,f) = \psi_{\mathrm{MVU}}(t,f)$ (see (6)).

Let us define the excess MSE due to compression as $\Delta\varepsilon \triangleq \mathrm{E}\big\{\big\|\widehat{W}_{X,\mathrm{CS}} - \widehat{W}_X\big\|_2^2\big\}$. We also define a norm $\|\cdot\|_{\mathrm{R}}$ for 2-D random processes $Y(t,f)$ by $\|Y\|_{\mathrm{R}} \triangleq \sqrt{\mathrm{E}\{\|Y\|_2^2\}}$. Then, applying the triangle inequality $\|Y_1+Y_2\|_{\mathrm{R}} \le \|Y_1\|_{\mathrm{R}} + \|Y_2\|_{\mathrm{R}}$ to $Y_1(t,f) = \widehat{W}_{X,\mathrm{CS}}(t,f) - \widehat{W}_X(t,f)$ and $Y_2(t,f) = \widehat{W}_X(t,f) - \overline{W}_X(t,f)$, we obtain the upper bound $\sqrt{\varepsilon_{\mathrm{CS}}} \le \sqrt{\Delta\varepsilon} + \sqrt{\varepsilon}$ or equivalently

$$\varepsilon_{\mathrm{CS}} \le (\sqrt{\varepsilon} + \sqrt{\Delta\varepsilon})^2\,. \tag{15}$$

The MSEs ε and $\Delta\varepsilon$ will be considered next.

MSE of the basic estimator. It is possible to show the expression

$$\varepsilon = \|r_X\|_2^2\bigg[M_X^{(\phi)} + TF\,s_X\sum_k\sum_l M_X^{(\phi_{k,l})}\,\mathrm{sinc}(2k\nu_{\max}T) \cdot \mathrm{sinc}(2l\tau_{\max}F)\bigg]\,, \tag{16}$$

where $\phi(\tau,\nu) = |1 - A_g(\tau,\nu)|$ and $\phi_{k,l}(\tau,\nu) = |A_g(\tau - kT, \nu - lF)|$, with $A_g(\tau,\nu) \triangleq \int_t g\big(t+\frac{\tau}{2}\big)g^*\big(t-\frac{\tau}{2}\big)e^{-j2\pi\nu t}dt$ being the ambiguity function of $g(t)$, and $\mathrm{sinc}(x) \triangleq \sin(\pi x)/(\pi x)$. Thus, ε will be smaller for a process $X(t)$ that is more underspread (smaller $M_X^{(\phi)}$ and $M_X^{(\phi_{k,l})}$, $(k,l) \ne (0,0)$, see (1)).

Excess MSE due to compression. Let us define the *TF sparsity profile*

$$\sigma_X(S) \triangleq \frac{1}{\bar{E}_X}\bigg|\sum_{(k,l)\in\mathcal{G}_S}\overline{W}_X(kT,lF)\bigg|,$$

where $S \in \mathbb{N}$ is a nominal sparsity and $\mathcal{G}_S$ is the set of (k,l) for which $|\overline{W}_X(kT,lF)|$ is *not* among the S largest elements of the set $\big\{|\overline{W}_X(kT,lF)|\big\}_{(k,l)\in\mathbb{Z}^2}$. Assuming that $\mathbf{\Phi}$ satisfies (11) for some S, so that (12) holds, one can show (using a CS error bound in [9])

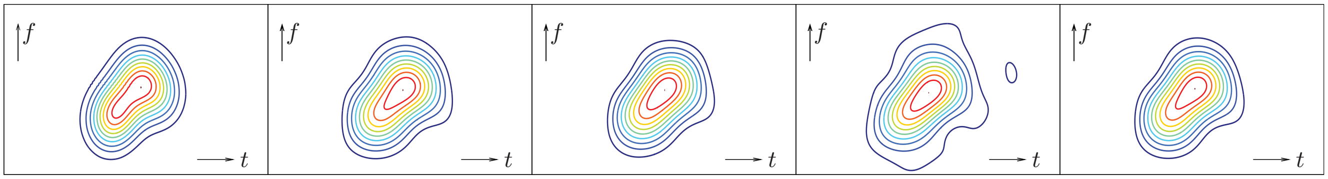
$$\begin{aligned}\Delta\varepsilon \le 16\,\alpha(S)\bar{E}_X^2\bigg[&4\,TF\alpha(S)\big[\sigma_X^2(S) + \big((N{-}S)s_X m_X^{(\phi)}\big)^2\big]\\ &+ 2\big[\sigma_X(S) + (N{-}S)s_X m_X^{(\phi)}\big](1 + NTF s_X m_X^{(\phi)})\\ &+ \frac{\alpha(S)+1}{\bar{E}_X^2}\sum_{(k,l)\in\mathbb{Z}^2\setminus(0,0)} M_X^{(\phi_{k,l})}\bigg]\,,\end{aligned} \tag{17}$$

with $\alpha(S) \triangleq \frac{N-S}{S}C^2$ and $\phi(\tau,\nu)$, $\phi_{k,l}(\tau,\nu)$ as before. For a larger S, $\sigma_X(S)$ tends to be smaller but the number M of measurements could be higher according to (13). Furthermore, the bound (17) will be smaller for a process that is more underspread (smaller $M_X^{(\phi)}$, $M_X^{(\phi_{k,l})}$ for $(k,l) \ne (0,0)$, s_X, $m_X^{(\phi)}$) and that has a better effective TF sparsity (smaller $\sigma_X(S)$). Using the approximations A1 and A2, (17) simplifies to $\Delta\varepsilon \le 32\,\alpha(S)\bar{E}_X^2\big[TF\alpha(S)\sigma_X^2(S) + \sigma_X(S)\big]$.

5. SIMULATION RESULTS

Using the TF synthesis method of [12], we generated 1000 realizations of a discrete-time random process of length 256 whose WVS is shown in Fig. 1(a). From the TF sparsity profile and EAF plotted in Fig. 2, we conclude that the process is only moderately TF sparse and underspread. In Fig. 1(b)–(e), we depict the results of the following WVS estimators averaged over the 1000 realizations: the basic (noncompressive) Gabor-based WVS estimator (3), the com-

[1] "Overwhelming probability" means that the probability of (11) not being true decreases exponentially with an increasing number of selected rows, M.



(a) (b) (c) (d) (e)

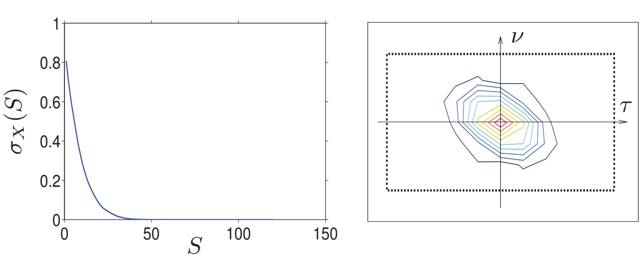
Fig. 1. Average performance of WVS estimators: (a) WVS of a discrete-time process of length 256, (b)–(e) averaged WVS estimates obtained with (b) the basic/noncompressive estimator (compression factor $N/M = 1$) (c) the compressive estimator with $N/M = 2$, (d) the compressive estimator with $N/M = 5$, and (e) the conventional estimator (smoothed Wigner distribution [4]).

pressive Gabor-based WVS estimator (10) for compression factors $N/M = 2$ and 5, and a conventional WVS estimator (smoothed Wigner distribution with MVU-type smoothing function [4]). For the basic Gabor-based estimator, the number of Gabor coefficients was chosen as $N = KL = 112$ ($K = 8$, $L = 14$), corresponding to $TF = 1.125$. It is seen that, on average, the basic Gabor-based estimator performs equally well as the conventional estimator, even though it uses only 112 Gabor coefficients on an undersampled grid. The result of the compressive estimator with compression factor $N/M = 2$ (using 56 Gabor coefficients) is practically equal to the results of the basic and conventional estimators, while small deviations are observed for $N/M = 5$ (22 Gabor coefficients).

Fig. 3 shows the empirical normalized MSE (NMSE) of the compressive estimator (10) versus the compression factor N/M, for $TF = 1.125$. The NMSE of the conventional estimator (which, of course, does not depend on N/M) is shown as a reference. It is seen that the performance loss of the compressive estimator compared to the basic (i.e., $N/M = 1$) and conventional estimators is small up to about $N/M = 3$ but increases beyond that point. The NMSE decrease for N/M between 1 and 2 may be due to a regularization effect of the CS recovery stage (such an effect is reported in [13] in a different context). We did not plot the MSE bound (15)–(17) because it is much larger than the empirical MSE (this lack of tightness is mostly due to the notoriously loose [10] CS error bound used in (17) and the moderate TF sparsity of our process). However, the bound is still valuable from a theoretical viewpoint.

6. CONCLUSION

We presented a "compressive" WVS estimator for TF sparse, underspread, nonstationary random processes. Our estimator is based on the signal's Gabor coefficients on an undersampled TF grid and uses a CS transformation to reduce the number of measurements. The measurements are inner products of the signal with randomly chosen linear combinations of Gabor functions; they can be computed efficiently by means of fast Fourier transforms. We provided a bound on the mean square error of the compressive WVS estimator and simulation results illustrating the estimator's performance.



(a) (b)

Fig. 2. (a) TF sparsity profile $\sigma_X(S)$, (b) EAF magnitude $|\bar{A}_X(\tau,\nu)|$ (the dotted reference rectangle has area 1).

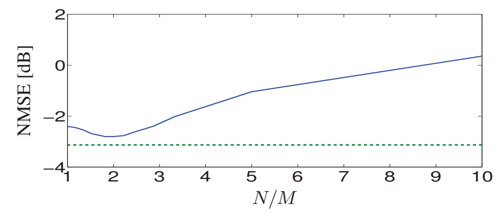


Fig. 3. Empirical NMSE of the compressive WVS estimator versus the compression factor N/M, for $TF = 1.125$. The dotted horizontal line indicates the NMSE of the conventional estimator.

7. REFERENCES

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

[2] 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.

[3] 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.

[4] 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.

[5] S. G. Mallat, G. Papanicolaou, and Z. Zhang, "Adaptive covariance estimation of locally stationary processes," *Ann. Stat.*, vol. 26, pp. 1–47, Feb. 1998.

[6] G. Matz, *A Time-Frequency Calculus for Time-varying Systems and Nonstationary Processes with Applications*. PhD thesis, Vienna University of Technology, Nov. 2000.

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

[8] W. Kozek, *Matched Weyl-Heisenberg Expansions of Nonstationary Environments*. PhD thesis, Vienna University of Technology, March 1997.

[9] 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, March 2006.

[10] 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.

[11] G. Matz, D. Schafhuber, K. Gröchenig, M. Hartmann, and F. Hlawatsch, "Analysis, optimization, and implementation of low-interference wireless multicarrier systems," *IEEE Trans. Wireless Comm.*, vol. 6, pp. 1921–1931, May 2007.

[12] F. Hlawatsch and W. Kozek, "Second-order time-frequency synthesis of nonstationary random processes," *IEEE Trans. Inf. Theory*, vol. 41, pp. 255–267, Jan. 1995.

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