COMPRESSIVE SENSING AND WAVEFORM DESIGN FOR THE IDENTIFICATION OF LINEAR TIME-VARYING SYSTEMS USING NOISY MEASUREMENTS

Jun Zhang[†], *Antonia Papandreou-Suppappola*[†*] *and Robin L. Murray*[‡]

[†]Department of Electrical Engineering, Arizona State University, Tempe, AZ
[‡]Naval Undersea Warfare Center, Newport, RI

ABSTRACT

The application of compressive sensing and waveform design on the estimation of linear time-varying system characteristics using noisy measurements is investigated in this paper. Due to the sparsity of the system's spreading function representation and the inherent noise in any real-world sensor or measurement device, we propose a new method based on our previous work for identifying narrowband, wideband and dispersive systems using a small set of measurements in the presence of noise. Through numerical simulations, we demonstrate the feasibility and the performance of compressive sensing to estimate the system spreading function.

Index Terms— compressive sensing, linear time-varying systems, system identification, waveform design, ℓ_1-norm minimization

1. INTRODUCTION

Due to parallel developments in separate fields, including radar, sonar and communications, linear time-varying (LTV) system processing has different mathematical characterizations. Specifically, an LTV system can be characterized using a kernel representation associated with a characteristic transform and an associated spreading function (SF). The characteristic transform describes how the propagating signal is affected by the system, and the spreading function describes how the signal energy diffuses during propagation.

LTV systems can be classified by different characteristic transform as narrowband, wideband or dispersive. Narrowband LTV systems have applications in multipath fast fading wireless communication channels, as well as in radar and sonar systems and can be represented using a kernel formulation and a narrowband spreading function (NSF) [1]. Similarly, wideband LTV systems have been represented using a kernel formulation and a wideband SF (WSF) [2]. Wideband LTV systems are characterized by time delay and Doppler scale changes, which describe the physical delay and dilation effect of the system on the analysis signal. Finally, dispersive systems induce a time-dependent frequency shift in the system output, which can be modeled as a superposition of instantaneous frequency shifts, weighted by a matched dispersive spreading function (DSF) [3]. We can also obtain a discrete equivalent representation under certain physical assumptions on the system, for an LTV system, which is useful in real system analysis.

In this paper, we define the LTV system identification as the estimation of the spreading function of the corresponding system. Previously [4], we investigated the identification problem assuming noiseless measurements. Exploiting physical restrictions on the real system, we obtained a sparse spreading function for the corresponding LTV system. However, ambient noise may corrupt measurements used to estimate the spreading function. Thus, the identification method must be robust in the presence of the noise. Recently, there has been a growing interest in recovering sparse signals from their projections onto a random vector using compressive sensing (CS) [5]. In this paper, we propose to use CS to identify LTV systems from a small set of noisy measurements. Signal recovery by CS in the presence of noise was investigated previously [6, 7, 8]; stable recovery conditions, recovery performance bounds, and recovery algorithms were investigated. Using these results, the feasibility and performance of LTV system identification using CS is discussed in this paper. Our technique has potential applications when many measurements cannot be collected or many signals cannot be transmitted, and the measurements are affected by additive noise.

2. LTV SYSTEM REPRESENTATIONS

Narrowband System Representations Assuming that $X(f)$, the Fourier transform of the input signal $x(t)$, is bandlimited to $[f_0, f_1]$ with bandwidth $W = f_1 - f_0$, and that the output signal $(\mathcal{L}x)(t)$ is time-limited to $[t_0, t_1]$ with duration $T = t_1 - t_0$, the output of a narrowband LTV system can be represented as

$$(\mathcal{L}x)(t) \quad = \quad \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \widehat{\mathrm{SF}}_{\mathcal{L}}\left(\frac{m}{W}, \frac{n}{T}\right) x_{m,n}(t). \qquad (1)$$

Here, $x_{m,n}(t) = (\mathcal{M}_{\frac{n}{T}}\mathcal{S}_{\frac{m}{W}}x)(t)$ is the time- and frequency-shifted version of $x(t)$, where $(\mathcal{S}_\tau x)(t) = x(t-\tau)$ is the time

*This work was supported by the NSF CAREER Award CCR-0134002 and the Department of Defense MURI Grant No. AFOSR FA9550-05-1-0443.

shift operation, $(\mathcal{M}_\nu x)(t) = x(t)\, e^{j2\pi\nu t}$ is the frequency shift operation, and $\widehat{\mathrm{SF}}_{\mathcal{L}}(\frac{m}{W}, \frac{n}{T})$ are two-dimensional (2-D) samples of a smoothed NSF [1].

Wideband System Representations For a wideband LTV system $\mathcal{B}$ defined on $L_2(\mathbb{R})$, the discrete wideband model was obtained in [2] assuming that $X(f)$ is bounded within $f \in [-W/2, W/2]$ where W is bandwidth of $X(f)$, and that its Mellin transform $\mathrm{MT}_x(\beta)$ is bounded within $\beta \in [-\beta_0/2, \beta_0/2]$, where β_0 is a constant representing the support of $\mathrm{MT}_x(\beta)$. Specifically, the output can be represented as

$$y(t) = \sum_{m\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} \hat{\chi}_{\mathcal{B}}(\frac{n}{a_0^m W}, a_0^m)\, a_0^{\frac{m}{2}}\, x(a_0^m t - \frac{n}{W}), \qquad (2)$$

where $\hat{\chi}_{\mathcal{B}}(\tau, a)$ is a 2-D smoothed version of the WSF, and the basic scaling factor is $a_0 = e^{1/\beta_0}$.

Dispersive System Representations If we assume a dispersive system $\mathcal{Z}$ changes the phase function of the input $x(t)$ by[1] $\xi(t/t_r)$, then $\nu(t) = \frac{d}{dt}\xi(t/t_r)$ is the dispersive (time-dependent) frequency shift. A DSP was developed to match the system propagation diffusion based on $\xi(t/t_r)$. A dispersive system can be treated as a unitary equivalent narrowband system, for which the input and output are temporally warped signals, $(\mathcal{U}_\xi x)(t)$ and $(\mathcal{U}_\xi \mathcal{Z} x)(t)$, respectively. Specifically, the relationship between the DSF and its equivalent NSF can be described as $\mathrm{DSF}_{\mathcal{Z}}(\zeta, \beta) = \mathrm{SF}_{\mathcal{U}_\xi \mathcal{Z}}\left(t_r\zeta, \frac{\beta}{t_r}\right)$. Here, $(\mathcal{Z}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{DSF}_{\mathcal{Z}}(\zeta, \beta)\, e^{-j\pi\zeta\beta} (\mathcal{D}_\beta^{(\xi)} \mathcal{G}_\zeta^{(\xi)} x)(t) d\zeta d\beta$, where $(\mathcal{G}_\zeta^{(\xi)} x)(t) = (\mathcal{U}_\xi^{-1} \mathcal{S}_{t_r\zeta} \mathcal{U}_\xi x)(t)$ and $(\mathcal{D}_\beta^{(\xi)} x)(t) = (\mathcal{U}_\xi^{-1} \mathcal{M}_{\beta/t_r} \mathcal{U}_\xi x)(t) = x(t) e^{-j2\pi\xi(\frac{t}{t_r})}$. The warping operator is defined as

$$(\mathcal{U}_\xi x)(t) = \left| t_r \nu\left(\xi^{-1}\left(\frac{t}{t_r}\right)\right)\right|^{-1/2} x\left(t_r \xi^{-1}\left(\frac{t}{t_r}\right)\right).$$

The corresponding discrete model can be found in [3].

Matrix Formulation of LTV System Outputs The discrete LTV system outputs can be represented in matrix formulation form. First, we provide the matrix formulation for the narrowband system in (1). Vectorizing D samples of the TF shifted waveform $x_{m,n}(t)$, we obtain the vector $\mathbf{x}_{m,n} = [x_{m,n}[1], x_{m,n}[2], \cdots, x_{m,n}[D]\,]^{\mathbb{T}}$. Then, the $D \times K$ signal basis matrix $\Phi = [\mathbf{x}_{0,0}, \cdots, \mathbf{x}_{m,n}, \cdots, \mathbf{x}_{M-1,N-1}]$ contains M time shifts and N frequency shifts of the TF shifted waveform with $K = MN$. Next, we concatenate the SF into the $K \times 1$ vector $\mathbf{H} = [\widehat{SF}_{\mathcal{L}}(0,0), \cdots, \widehat{SF}_{\mathcal{L}}(m,n), \cdots, \widehat{SF}_{\mathcal{L}}(M-1, N-1)]^{\mathbb{T}}$. Finally, the system output vector for a discrete LTV system can be expressed as $\mathbf{Y} = \Phi\mathbf{H}$. We can obtain a matrix formulation for wideband and dispersive systems following a similar procedure,.

[1]Here, $t_r > 0$ is a normalized time reference.

3. SPREADING FUNCTION ESTIMATION USING COMPRESSIVE SENSING

Based on the matrix representation for a discrete LTV system $\mathbf{Y} = \Phi\mathbf{H}$, we assume that $\mathbf{H}$ is an S-sparse vector, and $\Phi \in \mathbb{R}^{D\times K}$ the matrix representation for the signal basis. The identification goal is to determining $\mathbf{H}$ of the corresponding LTV system from a set of L available samples (which we refer to as measurements), where the number of measurements is less than the dimension of system output $\mathbf{Y}$ ($L < D$). In particular, we want to determine $\mathbf{H}$ from $\mathbf{s} = \mathbf{AY}$, using the appropriate measurement matrix $\mathbf{A} \in \mathbb{R}^{L\times D}$. Denoting $\Psi = A\Phi$, we also obtain $\mathbf{s} = \Psi\mathbf{H}$. For realistic applications, the measurement $\mathbf{s}$ is embedded in ambient noise $\mathbf{w}$, so that $\mathbf{s} = \Psi\mathbf{H} + \mathbf{w}$. Recently, Candès, Romberg and Tao proved the stable recovery theorem for the basis pursuit algorithm in [6]. Based on this algorithm, we are able to estimate the SF, $\mathbf{H}$, from the noisy measurements, $\mathbf{s}$.

Next, we briefly review some relevant definitions and principles of CS [5, 6]. The basis pursuit (BP) algorithm [6] can solve the convex programming problem: min $||\mathbf{H}||_{\ell_1}$ subject to $||\mathbf{s} - \Psi\mathbf{H}||_{\ell_2} \le \epsilon$, where $||\mathbf{H}||_{\ell_1} = \sum_i |H_i|$ denotes the ℓ_1-norm and ϵ is a positive constant. This can be recast as linear programming in the real case and cone programming in the complex case [6].

Denoting $\Lambda \subset \{1, \cdots, d\}$ and Ψ_Λ be the submatrix of Ψ consisting of the columns indexed by Λ. The local isometry constant $\delta_\Lambda = \delta_\Lambda(\Psi)$ is defined as the smallest number satisfying $(1 - \delta_\Lambda)||\mathbf{H}||_2^2 \le ||\Psi_\Lambda\mathbf{H}||_2^2 \le (1 + \delta_\Lambda)||\mathbf{H}||_2^2$ for all coefficient vectors supported on Λ [6]. Then the (global) restricted isometry constant is then defined as $\delta_S = \delta_S(\Psi) := \sup_{|\Lambda|=S} \delta_\Lambda(\Psi),\ S \in \mathbb{N}$.

Candès, Romberg and Tao proved the following recovery theorem for ℓ_1-norm programming can stably perform reconstruction from noise data in [6]:

Theorem 3.1 *Assume that Ψ satisfies $\delta_{3S}(\Psi) + 3\delta_{4S}(\Psi) < 2$ for some $S \in \mathbb{N}$. Let $\boldsymbol{H}$ be an S-sparse vector and assume we are given noisy data $\boldsymbol{s} = \Psi\boldsymbol{H} + \boldsymbol{w}$ with $||\boldsymbol{w}||_{\ell_2} \le \eta$. Then the solution $\widehat{\boldsymbol{H}}$ to the BP problem satisfies $||\widehat{\boldsymbol{H}} - \boldsymbol{H}||_{\ell_2} \le C\eta$. The constant C depends only on δ_{3S} and δ_{4S}. If $\delta_{4S} \le 1/3$ then $C \le 15.41$. In particular, if no noise is present, i.e., $\eta = 0$, then under the stated condition, BP recovers $\boldsymbol{H}$ exactly.*

In [6], it was proved that the minimizing ℓ_1-norm stably recovers the S-largest entries of an K-dimensional unknown vector $\mathbf{H}$ from L measurements only.

Theorem 3.2 *Suppose that $\boldsymbol{H}_0$ is an arbitrary vector in $\mathbb{R}^K$ and let $\boldsymbol{H}_{0,S}$ be the truncated vector corresponding to the S largest values of $\boldsymbol{H}_0$ (in absolute value). Assume that Ψ satisfies $\delta_{3S}(\Psi) + 3\delta_{4S}(\Psi) < 2$ for some $S \in \mathbb{N}$. Let $\boldsymbol{H}$ be an S-sparse vector and assume we are given noisy data $\boldsymbol{s} = \Psi\boldsymbol{H} + \boldsymbol{w}$ with $||\boldsymbol{w}||_{\ell_2} \le \eta$, then the solution $\boldsymbol{H}^{\#}$ to the ℓ_1-norm minimization obeys $||\boldsymbol{H}^{\#} - \boldsymbol{H}_0||_{\ell_2} \le C_{1,S}\epsilon + C_{2,S}\frac{||\boldsymbol{H}_0 - \boldsymbol{H}_{0,S}||_{\ell_1}}{\sqrt{S}}$,* *where for reasonable values of δ_{4S} the constants are well behaved; e.g. $C_{1,S} \approx 12.04$ and $C_{1,S} \approx 8.77$ for $\delta_{4S} = \frac{1}{5}$.*

The authors argued in [6] that for the class of *compressible* signals, e.g. the wavelet coefficients of a piecewise smooth signal, there are no fundamentally better estimates available than ℓ_1-norm minimization estimates.

3.1. The Measurement Matrix A

For the identification of LTV systems, we need to choose measurement matrix $\mathbf{A}$ in $\Psi = \mathbf{A}\Phi$ to satisfy the *concentration inequality*, specifically, $\mathbb{P}(\left| ||\mathbf{A}\mathbf{v}||^2 - ||\mathbf{v}||^2 \right| \geq \varepsilon ||\mathbf{v}||) \leq 2e^{-c\frac{L}{2}\varepsilon^2}, \varepsilon \in (0, \frac{1}{3})$, for all $\mathbf{v} \in \mathbb{R}^D$ and some constant $c > 0$ [9], where $\mathbb{P}(\cdot)$ denotes probability. This inequality is satisfied by the Gaussian ensemble random matrix, the Bernoulli ensemble random matrix, etc [9]. Specifically for Gaussian ensemble random matrix case, if the entries of $\mathbf{A}$ are independent normal variables with mean zero and variance L^{-1}, then the *concentration inequality* holds with $c = 7/18$.

3.2. Stable Recovery Condition Using Basis Pursuit

In [9], the stable recovery condition and isometry constants of $\Psi = \mathbf{A}\Phi$ were investigated using the following theorem.

Theorem 3.3 *Let $\Psi \in \mathbb{R}^{D\times K}$ be a dictionary with coherence μ. Assume that $S - 1 \leq \frac{1}{16}\mu^{-1}$. Let $\mathbf{A} \in \mathbb{R}^{L\times D}$ be a random matrix that satisfies the concentration inequality. Assume that $L \geq C_1(S\log(K/S) + C_2 + t)$, where t is a positive real number. Then with probability at least $1 - e^{-t}$, the composed matrix $\Psi = \mathbf{A}\Phi$ has restricted isometry constant $\delta_S(\Psi) \leq 1/3$. The constants satisfy $C_1 \leq 138.51c^{-1}$ and $C_2 \leq \log(1250/13) + 1 \approx 5.57$.*

If each column of the matrix Φ is normalized to 1, the coherence μ of the matrix Φ can be defined as $\mu = \max_{i\neq j} | \langle \varphi_i, \varphi_j \rangle |$, where φ_i and φ_j are the ith and jth column of Φ, respectively. We can obtain the stable recovery condition for the spreading function vector $\mathbf{H}$ for BP by combining Theorem 3.1 and Theorem 3.3. We can notice that the number of necessary measurements L is on the order of $S\log(K/S)$.

3.3. Improved Basis Pursuit Algorithm

In [7], the author proposed a variation of the *Dantzig selector* using the shrinkage technique [10]. Here, we use a similar technique to improve identification performance. The algorithm is stated as follows and the stability is given by Theorem 3.2.

(1) Estimate $\hat{\mathbf{H}}$ by solving min $||\hat{\mathbf{H}}||_{\ell_1}$ subject to $\mathbf{s} = \Psi\hat{\mathbf{H}}$, by basis pursuit;

(2) Obtain the index set $\hat{\mathbf{I}}_S$ corresponding to S-largest coefficients in estimated coefficients $\hat{\mathbf{H}}$, or choose the index set as $\hat{\mathbf{I}}_T = \{i : |\hat{H}|_i > \alpha\sigma\}$ for some $\alpha \geq 0$.

(3) Construct the estimator $\hat{\mathbf{H}}_{\hat{\mathbf{I}}_S} = (\Psi^{\mathbb{H}}_{\hat{\mathbf{I}}_S}\Psi_{\hat{\mathbf{I}}_S})^{-1}\Psi^{\mathbb{H}}_{\hat{\mathbf{I}}_S}\mathbf{s}$, or similarly for $\hat{\mathbf{I}}_T$ as $\hat{\mathbf{H}}_{\hat{\mathbf{I}}_T} = (\Psi^{\mathbb{H}}_{\hat{\mathbf{I}}_T}\Psi_{\hat{\mathbf{I}}_T})^{-1}\Psi^{\mathbb{H}}_{\hat{\mathbf{I}}_T}\mathbf{s}$.

4. WAVEFORM DESIGN

A very low coherence of Φ is required to satisfy the condition $S - 1 \leq \frac{1}{16}\mu$ for compressed sensing according to [9]. Moreover, [9] showed that given Φ with orthogonal columns, the upper bound of the required number of samples for stable recovery decreases. As a result, we want to design the transmission waveforms to form orthogonal basis in Φ for different types of LTV systems.

If the system can be characterized as narrowband system, we consider the direct sequence code division multiple access (DS-CDMA) signal as the basic transmission waveform. The DS-CDMA signal is generated using a pseudo noise (PN) sequence. For a PN code with length N_c, the waveform is represented as $x(t) = \sum_{p=0}^{N_c-1} c_n v(t - pT_c)$ where c_n is the nth bit (or chip) of the PN sequence, and $v(t)$ is the PN chip waveform with duration T_c. Different time-frequency shifted versions of these waveforms are mutually orthogonal [1]. This property can be represented in mathematic form as $\langle x_{m,n}, x_{m',n'}\rangle = \int_0^T x_{m,n}(t), x^*_{m',n'}(t)dt \approx C\delta[m - m']\delta[n - n']$, where $\delta[\cdot]$ denotes the Kronecker delta function, C is a constant, and $T = N_c T_c$ is the duration of the DS-CDMA signal waveform. Note that in [11], the author proposed a similar analysis window for signal recovery from Gabor time-frequency representations.

If the system can be characterized as wideband systems, we propose a wavelet-based waveform design scheme [2] due to the similarity of $x_{n,m}(t)$ in (2) to orthonormal wavelets. Specifically, we let $\psi(t)$ be a basis wavelet function, i.e. $\psi_{n,m}(t) = 2^{-\frac{m}{2}}\psi(\frac{t}{2^m} - n),\ n, m \in \mathbb{Z}$, constitute an orthonormal basis on $L^2(\mathbb{R})$, and then we can choose the signaling waveform in (2) to be $x(t) = \frac{1}{\sqrt{T_w}}\psi(\frac{t}{T_w})$, where T_w is some positive real number. With the assumption that $W \approx 1/T_w$ and letting $a_0 = \frac{1}{2}$ in (2), the delayed and dilated signal $x_{n,m}(t)$ also constitutes an orthonormal basis since $\forall\, n, n', m, m' \in \mathbb{Z}$, we have $\frac{1}{T_w}\int_{-\infty}^{\infty}\psi_{m,n}(\frac{t}{T_w})\psi^*_{m',n'}(\frac{t}{T_w})\,dt = \delta[n - n']\ \delta[m - m']$.

A dispersive system is unitarily equivalent to the narrowband system, thus we can use the corresponding warped version of the aforementioned narrowband waveform design scheme. For dispersive system, we design the waveform as a form of DS-CDMA signaling as $x(t) = \sum_{p=0}^{N_c-1} c_n (\mathcal{U}_\xi^{-1}v)(t - pT_c)$. Using the unitary relationship, it can be shown that the corresponding basis functions $x_{m,n}(t)$ are mutually orthogonal. As a result, Φ will have orthogonal columns to satisfy the condition for using CS.

5. NUMERICAL RESULTS

Numerical results are displayed for a narrowband system in Fig. 1. We denote the estimated spreading function as $\widehat{\mathbf{H}}$, and define the estimation error rate as $\epsilon = \mathrm{E}\left(\frac{||\widehat{\mathbf{H}}-\mathbf{H}||_{\ell_1}}{||\mathbf{H}||_{\ell_1}}\right)$, where $\mathrm{E}(\cdot)$ is the expectation operator. In the following simulations, ϵ is obtained by averaging the results from 100 simulations. In Fig. 1(a), we plot the true narrowband SF. A SF estimate using BP with $\epsilon = 0.8762$ and a 6dB SNR is shown in Fig. 1(b) and using an improved BP algorithm with $\epsilon = 0.05$ is shown in Fig. 1(c). Fig. 1(d) compares the performance of the three estimators: the BP algorithm, the improved BP algorithm, and the ideal least square (LS) estimator. Although the LS estimator is not feasible in realistic applications, it provides a lower bound on performance. From numerical results, the S-sparse vector $\mathbf{H}$ can be recovered from 238 measurements (out of $D = 1043$ possible measurements); the improved BP algorithm outperforms the BP, and in the high SNR region approaches nearly optimal performance. We obtain similar results for wideband system. In Fig. 1(e), we demonstrate the true WSF. The estimation error using BP for a 6dB signal-to-noise ratio is $\epsilon = 0.6609$. The estimation error using an improved BP algorithm with $\epsilon = 0.07$ is shown in Fig. 1(f).

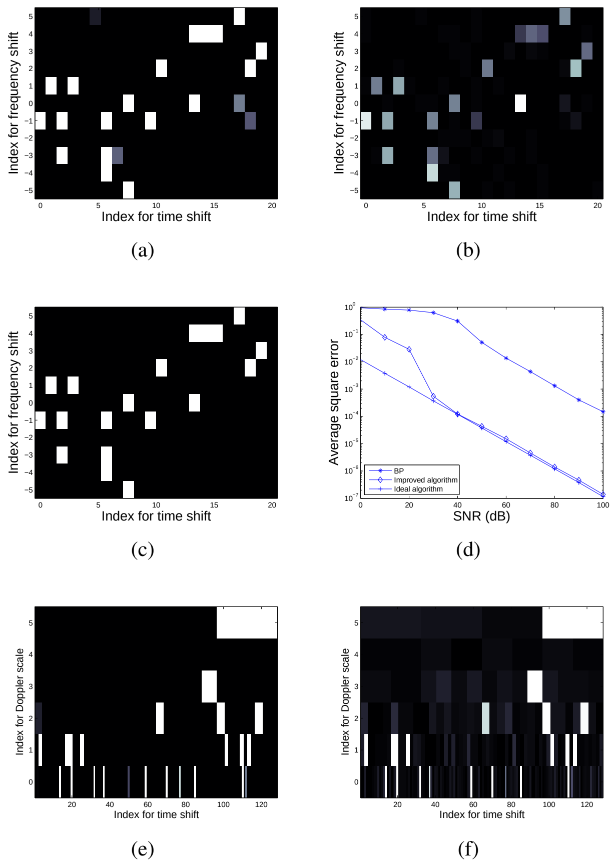

Fig. 1. (a) The true NSF and the recovered NSF estimates using (b) basis pursuit and (c) improved basis pursuit. (d) Comparison of the performance of three estimators: the BP, the improved BP, and the ideal LS estimator. (e) The true WSF and (f) recovered WSF using improved BP algorithm.

6. CONCLUSION

We investigated the application of CS and waveform design for linear time-varying system identification in the presence of measurement noise. Specifically, we proposed a method to estimate the spreading function of a LTV system using a small set of noisy measurements. We also suggested an improved compressive sensing algorithm based on shrinkage which greatly improves the estimation performance. Through numerical simulations, we successfully demonstrated the feasibility and the performance of CS system identification.

7. REFERENCES

[1] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, pp. 123–132, Jan. 1999.

[2] Y. Jiang and A. Papandreou-Suppappola, "Discrete time-scale characterization of wideband time-varying systems," *IEEE Trans. Signal Processing*, vol. 54, no. 4, pp. 1364–1375, April 2006.

[3] Y. Jiang and A. Papandreou-Suppappola, "Discrete time-frequency characterizations of dispersive linear time-varying systems," *IEEE Trans. Signal Processing*, vol. 55, pp. 2066–2076, May 2007.

[4] J. Zhang and A. Papandreou-Suppappola, "Compressive sensing and waveform design for the identification of linear time-varying systems," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Processing*, accepted, 2008.

[5] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–121, 2007.

[6] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[7] Emmanuel Candés and Terence Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *ANNALS OF STATISTICS*, vol. 35, pp. 2313, 2007.

[8] Yaakov Tsaig and David L. Donoho, "Extensions of compressed sensing," *Signal Process.*, vol. 86, no. 3, pp. 549–571, 2006.

[9] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *homepage.univie.ac.at/holger.rauhut/CompSensing.pdf*, 2006.

[10] Michael Elad, "Why simple shrinkage is still relevant for redundant representations?," *IEEE transactions on information theory*, vol. 52, no. 12, pp. 5559–5569, 2006.

[11] Götz Pfander and Holger Rauhut, "Sparsity in time-frequency representations," *Preprint (2007), submitted for publication.*