in Proc. EUSIPCO 2008
Lausanne, Switzerland, Aug. 2008

COMPRESSED SENSING BASED ESTIMATION OF DOUBLY SELECTIVE CHANNELS USING A SPARSITY-OPTIMIZED BASIS EXPANSION

Georg Tauböck and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
http://www.nt.tuwien.ac.at/about-us/staff/georg-tauboeck/

ABSTRACT

We propose a technique for estimating doubly selective channels within multicarrier communication systems. The new channel estimation technique uses the methodology of *compressed sensing* for a reduction of the number of pilots, and it employs a *basis expansion* that is optimized with a criterion of maximum sparsity. Simulation results demonstrate that the optimized basis yields significant performance gains relative to a previously proposed technique.

1. INTRODUCTION

We consider the estimation of doubly selective channels within multicarrier (MC) communication systems. Conventional methods for channel estimation (e.g., [1]) do not take advantage of the inherent *sparsity* of the channel. A method that exploits channel sparsity for a reduction of the number of pilots (and, thus, an increase of spectral efficiency) was proposed by the authors in [2]. This method is based on the recently introduced methodology of *compressed sensing* (CS), which enables the efficient reconstruction of sparse signals from a very limited number of measurements (samples) [3, 4].

In this paper, we generalize the CS-based channel estimation method of [2] by including a *sparsity-improving basis expansion* of the channel's time-frequency coefficients. The improved sparsity is due to a reduction of leakage effects that limit the performance of the method of [2]. We propose an iterative basis optimization procedure that aims to maximize sparsity, and we demonstrate significant performance gains obtained with the optimized basis.

The paper is organized as follows. After a description of the MC system model in Section 2, Section 3 analyzes the sparsity of the channel's delay-Doppler representation and introduces the basis expansion. The generalized CS-based channel estimation method is developed in Section 4. Section 5 proposes an iterative procedure for optimizing the basis. Finally, simulation results in Section 6 assess the performance gains achieved with the optimized basis.

2. MULTICARRIER SYSTEM MODEL

We assume a *pulse-shaping* MC system for the sake of generality and because of its advantages over conventional cyclic-prefix (CP) OFDM [5, 6]; however, CP-OFDM is included as a special case. The complex baseband domain is considered throughout.

2.1 Modulator, Channel, Demodulator

The (discrete-time) transmit signal generated by the MC modulator is given by

$$s[n] = \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} a_{l,k}\, g_{l,k}[n], \tag{1}$$

where K and L denote the numbers of subcarriers and transmitted MC symbols, respectively; $a_{l,k}$ ($l=0,\dots,L-1$; $k=0,\dots,K-1$) denotes the data symbols; and $g_{l,k}[n] \triangleq g[n-lN]\, e^{j2\pi\frac{k}{K}(n-lN)}$ is a time-frequency shifted version of a transmit pulse $g[n]$ ($N \geq K$ is the symbol duration). Using an interpolation filter with impulse response $f(t)$, $s[n]$ is converted into the continuous-time signal $s(t) = \sum_{n=-\infty}^{\infty} s[n]\, f(t-nT_\mathrm{s})$, where T_s is the sampling period. At the output of the doubly selective wireless channel, we obtain the signal

$$r(t) = \int_{-\infty}^{\infty} h(t,\tau)\, s(t-\tau)\, d\tau + z(t).$$

Here, $h(t,\tau)$ is the channel's time-varying impulse response and $z(t)$ is white complex Gaussian noise. At the receiver, $r(t)$ is converted into the discrete-time signal $r[n] = \int_{-\infty}^{\infty} r(t)\, f^*(t-nT_\mathrm{s})\, dt$. Subsequently, the MC demodulator calculates

$$r_{l,k} = \langle r, \gamma_{l,k} \rangle = \sum_{n=-\infty}^{\infty} r[n]\, \gamma_{l,k}^*[n] \tag{2}$$

for $l=0,\dots,L-1$ and $k=0,\dots,K-1$. Here, $\gamma_{l,k}[n] \triangleq \gamma[n-lN]\, e^{j2\pi\frac{k}{K}(n-lN)}$ with a receive pulse $\gamma[n]$. Finally, the $r_{l,k}$ are equalized and quantized according to the data symbol alphabet. For CP-OFDM [7, 8], $g[n]$ is 1 on $[0,N-1]$ and 0 otherwise, and $\gamma[n]$ is 1 on $[N-K,N-1]$ and 0 otherwise ($N-K \geq 0$ is the CP length).

The following relation between the discrete-time signals $s[n]$ and $r[n]$ is easily established:

$$r[n] = \sum_{m=-\infty}^{\infty} h[n,m]\, s[n-m] + z[n], \tag{3}$$

with the discrete-time time-varying impulse response

$$h[n,m] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t+nT_\mathrm{s},\tau)\, f(t-\tau+mT_\mathrm{s})\, f^*(t)\, dt\, d\tau, \tag{4}$$

and complex Gaussian discrete-time noise $z[n]$.

2.2 System Channel

Combination of (2), (3), and (1) yields the *system channel* subsuming the MC modulator, the physical channel, and the MC demodulator. Neglecting intersymbol and intercarrier interference (which is justified if $g[n]$ and $\gamma[n]$ are properly designed), we obtain

$$r_{l,k} = H_{l,k}\, a_{l,k} + z_{l,k}, \tag{5}$$

for $l=0,\dots,L-1$ and $k=0,\dots,K-1$. Here, $z_{l,k} \triangleq \langle z, \gamma_{l,k} \rangle$, and the $H_{l,k}$ can easily be expressed in terms of $g[n]$, $h[n,m]$, and $\gamma[n]$ [5].

Let us suppose that $\gamma[n]=0$ outside $[0,L_\gamma]$. To compute $r_{l,k}$ in (2), we must know $r[n]$ for $n=0,\dots,N_r-1$, where $N_r \triangleq (L-1)N + L_\gamma + 1$. Assuming $h[n,m]$ to be causal with maximum delay at most $K-1$, the system channel coefficients $H_{l,k}$ can then be expressed as [2]

$$H_{l,k} = \sum_{m=0}^{K-1} \sum_{i=0}^{N_r-1} F[m,i]\, e^{-j2\pi(\frac{km}{K} - \frac{Nli}{N_r})}, \tag{6}$$

where

$$F[m,i] \triangleq S_h[m,i]\, A_{\gamma,g}^*\Big(m, \frac{i}{N_r}\Big) \tag{7}$$

with the *discrete-delay-Doppler spreading function* [9] $S_h[m,i] \triangleq \frac{1}{N_r}\sum_{n=0}^{N_r-1} h[n,m]\, e^{-j2\pi\frac{in}{N_r}}$ and the *cross-ambiguity function* [10]

This work was supported by the WWTF projects MOHAWI (MA 44) and SPORTS (MA 07-004).

$A_{\gamma,g}(m,\xi) \triangleq \sum_{n=-\infty}^{\infty} \gamma[n]\, g^*[n-m]\, e^{-j2\pi\xi n}$. Using the approximation $N_r \approx LN$ (which is exact for CP-OFDM), we can rewrite (6) as the 2-D DFT

$$H_{l,k} = \sum_{m=0}^{K-1} \sum_{i=0}^{L-1} \tilde{F}[m,i]\, e^{-j2\pi(\frac{km}{K}-\frac{li}{L})}\,, \tag{8}$$

with the "pre-aliased" version of $F[m,i]$

$$\tilde{F}[m,i] \triangleq \sum_{q=0}^{N-1} F[m,i+qL]\,, \quad i=0,\ldots,L-1\,. \tag{9}$$

2.3 Pilot-Based Channel Estimation

We assume that the channel's maximum (discrete) delay is smaller than D, i.e., $h[n,m]=0$ for $m \geq D$, where $D \leq K$ is chosen such that $\Delta K \triangleq K/D$ is an integer. This restricts the support of $\tilde{F}[m,i]$ to $[0,D-1]\times[0,L-1]$. Because of (8), $H_{l,k}$ can then be subsampled with respect to the frequency index k: it is uniquely specified by its values on the subsampled grid $\mathscr{G} \triangleq \{(l,k)=(l,k'\Delta K) : l=0,\ldots,L-1;\, k'=0,\ldots,D-1\}$, and (8) entails the relation

$$H_{l,k'\Delta K} = \sum_{m=0}^{D-1} \sum_{i=0}^{L-1} \tilde{F}[m,i]\, e^{-j2\pi(\frac{k'm}{D}-\frac{li}{L})}, \quad k'=0,\ldots,D-1\,. \tag{10}$$

Note that we also allow $D=K$, in which case $\Delta K=1$.

Suppose that pilot symbols $a_{l,k}=p_{l,k}$ are transmitted at time-frequency positions $(l,k)\in\mathscr{P}$, with $\mathscr{P}\subset\mathscr{G}$. From (5), $r_{l,k} = H_{l,k}\,p_{l,k} + z_{l,k}$ for $(l,k)\in\mathscr{P}$. The receiver calculates channel coefficient estimates $\hat{H}_{l,k}$ at the pilot positions according to

$$\hat{H}_{l,k} \triangleq \frac{r_{l,k}}{p_{l,k}} = H_{l,k} + \frac{z_{l,k}}{p_{l,k}}\,, \qquad (l,k)\in\mathscr{P}\,. \tag{11}$$

Thus, these $H_{l,k}$ are known up to the additive noise terms $z_{l,k}/p_{l,k}$. For calculating channel estimates $\hat{H}_{l,k}$ on the whole subsampled grid $\mathscr{G}$ from the $\hat{H}_{l,k}|_{(l,k)\in\mathscr{P}}$, some interpolation technique is usually employed. However, here we will extend the approach based on CS introduced in [2]. We first need to develop a sparse representation of the channel.

3. SPARSE CHANNEL REPRESENTATION

3.1 Delay-Doppler Sparsity

We model the doubly selective wireless channel by P propagation paths corresponding to P specular scatterers with delays τ_p and Doppler frequency shifts ν_p for $p=1,\ldots,P$. Thus,

$$h(t,\tau) = \sum_{p=1}^{P} \eta_p\, \delta(\tau-\tau_p)\, e^{j2\pi\nu_p t}, \tag{12}$$

where η_p is the complex amplitude factor of the pth path. The discrete-time impulse response (4) then becomes

$$\begin{aligned} h[n,m] &= \sum_{p=1}^{P} \eta_p\, e^{j2\pi\nu_p n T_s} \int_{-\infty}^{\infty} e^{j2\pi\nu_p t}\, f(t-\tau_p+mT_s)\, f^*(t)\, dt \\ &\approx \sum_{p=1}^{P} \eta_p\, e^{j2\pi\nu_p n T_s}\, \phi\Big(m-\frac{\tau_p}{T_s}\Big)\,, \end{aligned} \tag{13}$$

where $\phi(x) \triangleq f(t) * f^*(-t)|_{t=T_s x}$. The approximation (13) is good if the ν_p are not too large and if $f(t)$ decays sufficiently fast.

Furthermore, extending the analysis performed in [2] for an idealized (sinc-type) filter $f(t)$, the discrete-delay-Doppler spreading function is obtained as

$$S_h[m,i] = \sum_{p=1}^{P} \eta_p\, e^{j\pi(\nu_p T_s - \frac{i}{N_r})(N_r-1)}\, \Lambda\Big(m-\frac{\tau_p}{T_s}, i-\nu_p T_s N_r\Big)\,, \tag{14}$$

with

$$\Lambda(x,y) \triangleq \phi(x)\,\psi(y)\,, \tag{15}$$

where $\psi(y) \triangleq \frac{1}{N_r}\, e^{j\pi\frac{y}{N_r}(N_r-1)} \sum_{n=0}^{N_r-1} e^{-j2\pi\frac{y}{N_r}n} = \frac{\sin(\pi y)}{N_r \sin(\pi y/N_r)}$. The function $\Lambda(x,y)$ describes the *leakage effect* that is due to the finite bandwidth ($\approx 1/T_s$) and the finite blocklength ($N_r \approx LN$).

We will now study the sparsity of $\Lambda(m-\tau_p/T_s, i-\nu_p T_s N_r)$. To this end, we first consider the energy of those samples of $\phi(m-\tau_p/T_s)$ whose distance from τ_p/T_s is greater than $\Delta m \in \{1,2,\ldots\}$. Let $\mathscr{M}$ denote the set of all integers $m\in\mathbb{Z}$ with $|m-\tau_p/T_s| > \Delta m$. We assume that $\phi(x)$ exhibits a polynomial decay,[1] i.e., $|\phi(x)| \leq C(1+|x/x_0|)^{-s}$ with $s\geq 1$, for some positive constants C and x_0. We then have the bound

$$\begin{aligned} \sum_{m\in\mathscr{M}} \Big|\phi\Big(m-\frac{\tau_p}{T_s}\Big)\Big|^2 &\leq C^2 \sum_{m\in\mathscr{M}} \Big(1+\Big|\frac{m-\tau_p/T_s}{x_0}\Big|\Big)^{-2s} \\ &\leq 2C^2 \sum_{m=\Delta m}^{\infty} \Big(1+\frac{m}{x_0}\Big)^{-2s} \leq 2C^2 \int_{\Delta m-1}^{\infty} \Big(1+\frac{x}{x_0}\Big)^{-2s} dx \\ &= \frac{2C^2 x_0}{2s-1}\Big(1+\frac{\Delta m-1}{x_0}\Big)^{-2s+1}, \end{aligned}$$

which shows that the energy of $\phi(m-\tau_p/T_s)$ outside the interval $[\lfloor \tau_p/T_s - \Delta m\rfloor, \lceil \tau_p/T_s + \Delta m\rceil]$ decays polynomially of order $2s-1$ with respect to Δm. In [2], an analogous result was obtained for the energy of $\psi(i-\nu_p T_s N_r)$ outside the interval $[\lfloor \nu_p T_s N_r - \Delta i\rfloor, \lceil \nu_p T_s N_r + \Delta i\rceil]$ (modulo N_r). There, the decay with respect to Δi is only linear (polynomial of order 1).

From these decay results, it follows that $\Lambda(m-\tau_p/T_s, i-\nu_p T_s N_r)$ can be considered as an *approximately sparse* function. We can therefore model $\Lambda(m-\tau_p/T_s, i-\nu_p T_s N_r)$ as N_Λ-sparse, i.e., at most N_Λ values of $\Lambda(m-\tau_p/T_s, i-\nu_p T_s N_r)$ are nonzero, with an appropriately chosen sparsity parameter N_Λ. It then follows from (14) that $S_h[m,i]$ is PN_Λ-sparse, and the same is true for $F[m,i]$ in (7) and for $\tilde{F}[m,i]$ in (9). The CS-based channel estimation method of [2] exploits this sparsity. Unfortunately, N_Λ cannot be chosen extremely small because of the strong leakage characterized by $\Lambda(x,y)$, which is typically due to the slowly decaying factor $\psi(i-\nu_p T_s N_r)$.

3.2 Basis Expansion

In order to improve the sparsity, we next introduce a 2-D basis expansion of the channel coefficients $H_{l,k'\Delta K}$ into orthonormal basis functions $u_{m,i}[l,k']$:

$$H_{l,k'\Delta K} = \sum_{m=0}^{D-1} \sum_{i=0}^{L-1} \alpha_{m,i}\, u_{m,i}[l,k']\,. \tag{16}$$

Clearly, (16) generalizes (10), which is a special case given by the orthonormal 2-D discrete Fourier basis

$$u_{m,i}[l,k'] = \frac{1}{\sqrt{DL}}\, e^{-j2\pi(\frac{k'm}{D}-\frac{li}{L})}\,. \tag{17}$$

In that case, the expansion coefficients are $\alpha_{m,i} = \sqrt{DL}\,\tilde{F}[m,i]$. The advantage of (16) is the possibility of using a basis $u_{m,i}[l,k']$ for which the coefficient function $\alpha_{m,i}$ is sparser than $\tilde{F}[m,i]$ (which uses the Fourier basis).

We will use a basis that is adapted to the channel model (12). We can conclude from (12) that the coefficient function $\alpha_{m,i}$ should be sparse for the single-scatterer channel

$$h^{(\tau_p,\nu_p)}(t,\tau) \triangleq \delta(\tau-\tau_p)\, e^{j2\pi\nu_p t}, \tag{18}$$

for all possible choices of $\tau_p \in [0,\tau_{\max}]$ and $\nu_p \in [-\nu_{\max},\nu_{\max}]$. Inserting (18) into (7) and using (9) and (10), we obtain

$$H_{l,k'\Delta K} = \sum_{m=0}^{D-1} \phi\Big(m-\frac{\tau_p}{T_s}\Big)\, B^{(\nu_p)}[m,l]\, e^{-j2\pi\frac{k'm}{D}} \tag{19}$$

[1] This includes (i) the ideal lowpass filter $f(t)=\sqrt{1/T_s}\,\mathrm{sinc}(t/T_s)$ with $\mathrm{sinc}(x) \triangleq \frac{\sin(\pi x)}{\pi x}$; here $\phi(x)=\mathrm{sinc}(x)$ and $s=1$; and (ii) a root-raised-cosine filter with roll-off factor ρ; here $\phi(x)=\mathrm{sinc}(x)\cos(\rho\pi x)/[1-(2\rho x)^2]$ and $s=3$.

with

$$B^{(\nu_p)}[m,l] \triangleq \sum_{i=0}^{L-1}\sum_{q=0}^{N-1} \psi^{(\nu_p)}(i+qL)\, A^*_{\gamma,g}\Big(m, \frac{i+qL}{N_r}\Big)\, e^{j2\pi\frac{li}{L}}, \quad (20)$$

where $\psi^{(\nu_p)}(i) \triangleq e^{j\pi(\nu_p T_s - \frac{i}{N_r})(N_r-1)}\psi\big(i-\nu_p T_s N_r\big)$. We now consider a 1-D basis expansion of $B^{(\nu_p)}[m,l]$ with respect to l, i.e.,

$$B^{(\nu_p)}[m,l] = \sum_{i=0}^{L-1} \tilde{\alpha}^{(\nu_p)}_{m,i}\, b_{m,i}[l], \quad (21)$$

with an m-dependent basis $\{b_{m,i}[l]\}_{i=0,\ldots,L-1}$ that is orthonormal (i.e., $\sum_{l=0}^{L-1} b_{m,i_1}[l]\, b^*_{m,i_2}[l] = \delta[i_1-i_2]$ for all m) and does not depend on the value of ν_p in $B^{(\nu_p)}[m,l]$. The idea is to choose this 1-D basis such that the coefficient vector $\big[\tilde{\alpha}^{(\nu_p)}_{m,0}\cdots\tilde{\alpha}^{(\nu_p)}_{m,L-1}\big]^T$ is sparse for all m and all $\nu_p \in [-\nu_{\max},\nu_{\max}]$.

Substituting (21) back into (19), we obtain

$$H_{l,k'\Delta K} = \sum_{m=0}^{D-1}\sum_{i=0}^{L-1} \phi\Big(m-\frac{\tau_p}{T_s}\Big)\, \tilde{\alpha}^{(\nu_p)}_{m,i}\, b_{m,i}[l]\, e^{-j2\pi\frac{k'm}{D}}.$$

This can now be identified with the 2-D basis expansion (16), i.e., $H_{l,k'\Delta K} = \sum_{m=0}^{D-1}\sum_{i=0}^{L-1} \alpha^{(\tau_p,\nu_p)}_{m,i} u_{m,i}[l,k']$, with the orthonormal 2-D basis

$$u_{m,i}[l,k'] \triangleq \frac{1}{\sqrt{D}}\, b_{m,i}[l]\, e^{-j2\pi\frac{k'm}{D}}$$

and the 2-D coefficient function $\alpha^{(\tau_p,\nu_p)}_{m,i} \triangleq \sqrt{D}\,\phi\big(m-\frac{\tau_p}{T_s}\big)\,\tilde{\alpha}^{(\nu_p)}_{m,i}$. The basis $\{u_{m,i}[l,k']\}$ agrees with the 2-D Fourier basis (17) with respect to k', but is different with respect to l because $e^{j2\pi\frac{li}{L}}$ is replaced by $b_{m,i}[l]$. This potentially improves the poor decay of $\psi(y)$ observed in Section 3.1. Thus, the coefficient functions $\alpha^{(\tau_p,\nu_p)}_{m,i}$ are potentially sparser than if the 2-D Fourier basis is used (this aspect will be pursued in Sections 5 and 6). When the channel comprises P scatterers as in (12), the coefficient function is $\alpha_{m,i} = \sum_{p=1}^{P}\eta_p\alpha^{(\tau_p,\nu_p)}_{m,i}$. If each $\alpha^{(\tau_p,\nu_p)}_{m,i}$ is S-sparse, $\alpha_{m,i}$ is PS-sparse.

4. CS-BASED CHANNEL ESTIMATION

We now combine pilot-based channel estimation and the basis expansion discussed above with CS.

4.1 The Measurement Equation

Consider the functions $H_{l,k'\Delta K}$ and $u_{m,i}[l,k']$ ($l=0,\ldots,L-1$; $k'=0,\ldots,D-1$) as $L\times D$ matrices and let $\mathbf{h} \triangleq \mathrm{vec}\big\{H_{l,k'\Delta K}\big\}$ and $\mathbf{u}_{m,i} \triangleq \mathrm{vec}\big\{u_{m,i}[l,k']\big\}$ denote the vectors of length DL obtained by stacking all columns of these matrices (e.g., $\mathbf{h} = [h_1\cdots h_{DL}]^T$ with $h_{k'L+l+1} = H_{l,k'\Delta K}$). We can then rewrite (16) as

$$\mathbf{h} = \sum_{m=0}^{D-1}\sum_{i=0}^{L-1} \alpha_{m,i}\,\mathbf{u}_{m,i} = \mathbf{U}\boldsymbol{\alpha}, \quad (22)$$

where $\boldsymbol{\alpha} \triangleq \mathrm{vec}\big\{\alpha_{m,i}\big\}$ and $\mathbf{U}$ is the $DL\times DL$ matrix whose $(iD+m+1)$th column is given by the vector $\mathbf{u}_{m,i}$. Because the vectors $\mathbf{u}_{m,i}$ are orthonormal, the matrix $\mathbf{U}$ is unitary.

According to Section 2.3, $|\mathscr{P}|$ of the DL entries of the channel vector $\mathbf{h}$ are given by the channel coefficients $H_{l,k}$ at the pilot positions $(l,k)\in\mathscr{P}$. Let $\mathbf{h}^{(\mathrm{p})}$ denote the corresponding length-$|\mathscr{P}|$ subvector of $\mathbf{h}$, and let $\mathbf{U}^{(\mathrm{p})}$ denote the $|\mathscr{P}|\times DL$ submatrix of $\mathbf{U}$ constituted by the corresponding $|\mathscr{P}|$ rows of $\mathbf{U}$. Furthermore, let $\boldsymbol{\Phi}_{\mathrm{c}} \triangleq \mathbf{U}^{(\mathrm{p})}\mathbf{D}$ and $\mathbf{x}_{\mathrm{c}} \triangleq \mathbf{D}^{-1}\boldsymbol{\alpha}$, where the diagonal matrix $\mathbf{D}$ is chosen such that all columns of $\boldsymbol{\Phi}_{\mathrm{c}}$ have unit l_2-norm. We then have

$$\mathbf{h}^{(\mathrm{p})} = \mathbf{U}^{(\mathrm{p})}\boldsymbol{\alpha} = \boldsymbol{\Phi}_{\mathrm{c}}\mathbf{x}_{\mathrm{c}}, \quad (23)$$

which is (22) reduced to the pilot positions.

Because $\mathbf{h}^{(\mathrm{p})}$ is unknown, we use its estimate $\hat{\mathbf{h}}^{(\mathrm{p})}$ instead, i.e., the length-$|\mathscr{P}|$ vector containing the estimates $\hat{H}_{l,k}|_{(l,k)\in\mathscr{P}}$. According to (11), $\hat{\mathbf{h}}^{(\mathrm{p})} = \mathbf{h}^{(\mathrm{p})} + \mathbf{z}_{\mathrm{c}}$, where $\mathbf{z}_{\mathrm{c}}$ is the vector of noise terms $z_{l,k}/p_{l,k}|_{(l,k)\in\mathscr{P}}$. Inserting (23) yields $\hat{\mathbf{h}}^{(\mathrm{p})} = \boldsymbol{\Phi}_{\mathrm{c}}\mathbf{x}_{\mathrm{c}} + \mathbf{z}_{\mathrm{c}}$. This can be reformulated as the real "measurement equation"

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x} + \mathbf{z}, \quad (24)$$

where $\mathbf{x} \triangleq \big[\Re\{\mathbf{x}_{\mathrm{c}}^T\}\ \Im\{\mathbf{x}_{\mathrm{c}}^T\}\big]^T$, $\mathbf{y} \triangleq \big[\Re\{\hat{\mathbf{h}}^{(\mathrm{p})T}\}\ \Im\{\hat{\mathbf{h}}^{(\mathrm{p})T}\}\big]^T$, $\mathbf{z} \triangleq \big[\Re\{\mathbf{z}_{\mathrm{c}}^T\}\ \Im\{\mathbf{z}_{\mathrm{c}}^T\}\big]^T$, and $\boldsymbol{\Phi} \triangleq \begin{bmatrix}\Re\{\boldsymbol{\Phi}_{\mathrm{c}}\} & -\Im\{\boldsymbol{\Phi}_{\mathrm{c}}\}\\ \Im\{\boldsymbol{\Phi}_{\mathrm{c}}\} & \Re\{\boldsymbol{\Phi}_{\mathrm{c}}\}\end{bmatrix}$ are real representations of $\mathbf{x}_{\mathrm{c}}, \hat{\mathbf{h}}^{(\mathrm{p})}, \mathbf{z}_{\mathrm{c}}$, and $\boldsymbol{\Phi}_{\mathrm{c}}$, respectively.

4.2 Sparse Linear Reconstruction via CS

Our task now is to recover the real vector $\mathbf{x}$ from the known real vector $\mathbf{y}$, based on the measurement equation (24) with known $2|\mathscr{P}|\times 2DL$ "measurement matrix" $\boldsymbol{\Phi}$. Once we have found $\mathbf{x}$ or, equivalently, $\mathbf{x}_{\mathrm{c}}$, we calculate channel coefficient estimates on $\mathscr{G}$ by using $\boldsymbol{\alpha} = \mathbf{D}\mathbf{x}_{\mathrm{c}}$ and (22). Estimates of all channel coefficients $H_{l,k}$ are then obtained by inversion of (10) and use of (8).

Unfortunately, this reconstruction problem is ill-posed because $2|\mathscr{P}| \ll 2DL$, i.e., the number of scalar equations in (24) is much smaller than the number of unknowns. However, assuming an appropriate choice of the basis $\{u_{m,i}[l,k']\}$ (which determines $\boldsymbol{\Phi}$), we can assume that $\boldsymbol{\alpha}$ is (approximately) sparse, as discussed in Section 3.2. Then $\mathbf{x}$ is sparse as well—it has at most twice as many nonzero entries as $\boldsymbol{\alpha}$. Our channel estimation problem is thus recognized to be a *sparse reconstruction problem*: we aim to recover $\mathbf{x}$ based on the underdetermined linear model (24), subject to the constraint that $\mathbf{x}$ is S-sparse, with a suitably chosen sparsity parameter S. Note that the positions of the nonzero entries of $\mathbf{x}$ are unknown.

This sparse reconstruction problem can be tackled by the CS methodology. A key condition of CS is that $\boldsymbol{\Phi}$ satisfies a "restricted isometry" property [11]. Let $\boldsymbol{\Phi}_{\mathscr{I}}$, $\mathscr{I}\subset\{1,\ldots,2DL\}$ be the $2|\mathscr{P}|\times|\mathscr{I}|$ submatrix comprising the columns of $\boldsymbol{\Phi}$ indexed by the elements of $\mathscr{I}$. Then the *S-restricted isometry constant* δ_S of $\boldsymbol{\Phi}$ is defined as the smallest quantity δ_S such that

$$(1-\delta_S)\|\mathbf{a}\|_2^2 \le \|\boldsymbol{\Phi}_{\mathscr{I}}\mathbf{a}\|_2^2 \le (1+\delta_S)\|\mathbf{a}\|_2^2$$

for all index sets $\mathscr{I}$ with $|\mathscr{I}|\le S$ and all vectors $\mathbf{a}\in\mathbb{R}^{|\mathscr{I}|}$.

For estimating $\mathbf{x}$, we consider an extension of the *basis pursuit* defined by the convex program [11]

$$\hat{\mathbf{x}} \triangleq \arg\min_{\mathbf{x}\in\mathscr{X}_\varepsilon} \|\mathbf{x}\|_1, \quad (25)$$

where $\mathscr{X}_\varepsilon$ is the set of all $\mathbf{x}\in\mathbb{R}^{2DL}$ satisfying $\|\boldsymbol{\Phi}\mathbf{x}-\mathbf{y}\|_2 \le \varepsilon$ for a given $\varepsilon>0$. This program can be solved efficiently by interior-point methods [12]. It is able to recover S-sparse parameter vectors according to the following result [11].

For a given S, assume that the $3S$- and $4S$-restricted isometry constants of $\boldsymbol{\Phi}$ satisfy

$$\delta_{3S} + 3\delta_{4S} < 2. \quad (26)$$

Let $\mathbf{y} = \boldsymbol{\Phi}\mathbf{x}+\mathbf{z}$ with $\|\mathbf{z}\|_2\le\varepsilon$, and let $\mathbf{x}_S\in\mathbb{R}^{2DL}$ contain the S components of $\mathbf{x}$ with largest absolute values, the remaining $2DL-S$ components being zero. Then the estimate $\hat{\mathbf{x}}$ in (25) satisfies

$$\|\hat{\mathbf{x}}-\mathbf{x}\|_2 \le C_1\varepsilon + C_2\frac{\|\mathbf{x}-\mathbf{x}_S\|_1}{\sqrt{S}},$$

where the constants C_1 and C_2 depend only on δ_{3S} and δ_{4S}.

For a zero-mean Gaussian noise vector $\mathbf{z}$, the condition $\|\mathbf{z}\|_2\le\varepsilon$ is satisfied with high probability for ε suitably chosen. Regarding the restricted isometry condition (26), the following result has been shown in [13]. If $\boldsymbol{\Phi}_{\mathrm{c}}\in\mathbb{C}^{N_1\times N_2}$ with $N_1\le N_2$ is constructed by selecting uniformly at random [3] N_1 rows from a unitary $N_2\times N_2$ matrix

$\mathbf{U}$ and normalizing the columns, a sufficient condition for (26) to be true with overwhelming probability[2] is

$$N_1 \geq C_3\,\mu^2\,(\ln N_2)^4\,S. \tag{27}$$

Here, $\mu \triangleq \sqrt{N_2}\,\max_{i,j}|U_{i,j}|$ (known as the *coherence* of $\mathbf{U}$) and C_3 is a constant. It has been pointed out in [2] that under condition (27), the corresponding real measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{2N_1 \times 2N_2}$ also satisfies (26) with overwhelming probability.

According to Section 4.1, the pilot positions $(l,k) \in \mathscr{P}$ correspond to $|\mathscr{P}|$ entries of the channel vector $\mathbf{h}$. For consistency with the CS framework, these positions are selected uniformly at random [3]. The number of pilots is chosen to satisfy condition (27), which becomes (note that $N_1 = |\mathscr{P}|$ and $N_2 = DL$)

$$|\mathscr{P}| \geq C_3\,\mu^2\,\big(\ln(DL)\big)^4\,S.$$

For the Fourier basis, $S = 2PN_\Lambda$ and $\mu = 1$. In practice, the pilot positions will be randomly chosen only once before the beginning of data transmission. With high probability, they will lead to good performance for arbitrary channels with at most P paths.

5. SPARSITY-OPTIMIZED BASIS EXPANSION

In this section, we propose a procedure for optimizing the "sparsity-improving" 1-D basis $\{b_{m,i}[l]\}$ in (21).

5.1 The Optimization Problem

Ideally, the m-dependent orthonormal basis $\{b_{m,i}[l]\}_{i=0,\ldots,L-1}$ should be such that the coefficient vector $\big[\tilde{\alpha}_{m,0}^{(\nu_p)} \cdots \tilde{\alpha}_{m,L-1}^{(\nu_p)}\big]^T$ in (21) is sparse for all m and all $\nu_p \in [-\nu_{\max}, \nu_{\max}]$ (the maximum Doppler $\nu_{\max}$ is assumed known). For our optimization, we slightly relax this requirement in that we only require a sparse coefficient vector for a finite number of uniformly spaced Doppler frequencies $\nu_p \in \mathscr{D}$ where $\mathscr{D} \triangleq \big\{\nu_\Delta d,\ d = -\lceil \nu_{\max}/\nu_\Delta \rceil, \ldots, \lceil \nu_{\max}/\nu_\Delta \rceil\big\}$.

Regarding the choice of the Doppler frequency spacing ν_Δ, it is interesting to note that for the "canonical spacing" $\nu_\Delta = 1/(T_s N_r)$, $B^{(\nu_p)}[m,l]$ in (20) collapses into

$$B^{(\nu_\Delta d)}[m,l] = \sum_{i=0}^{L-1} e^{j\pi\frac{d-i}{N_r}(N_r-1)}\,\delta[i-d]\,A^*_{\gamma,g}\Big(m, \frac{i}{N_r}\Big)\,e^{j2\pi\frac{li}{L}}.$$

This is an expansion of $B^{(\nu_\Delta d)}[m,l]$ into the 1-D Fourier basis $b_{m,i}[l] \propto e^{j2\pi\frac{li}{L}}$ with an ideally sparse coefficient function $\tilde{\alpha}_{m,i}^{(\nu_\Delta d)} \propto \delta[i-d]$ (no leakage effect). Hence, the resulting optimal basis would trivially be the Fourier basis. We therefore choose a Doppler spacing that is twice as dense, i.e., $\nu_\Delta = 1/(2T_s N_r)$. In this case, $\mathscr{D}$ includes also the Doppler frequencies located midway between the canonical sampling points, for which the Fourier basis results in maximum leakage (these frequencies are obtained for odd d).

The expansion coefficients defined by (21) can be calculated as $\tilde{\alpha}_{m,i}^{(\nu_p)} = \sum_{l=0}^{L-1} B^{(\nu_p)}[m,l]\,b^*_{m,i}[l]$, $i = 0,\ldots,L-1$. Equivalently,

$$\tilde{\boldsymbol{\alpha}}_m^{(\nu_p)} = \mathbf{B}_m\,\boldsymbol{\beta}_m^{(\nu_p)},$$

with the length-L vectors $\tilde{\boldsymbol{\alpha}}_m^{(\nu_p)} \triangleq \big[\tilde{\alpha}_{m,0}^{(\nu_p)} \cdots \tilde{\alpha}_{m,L-1}^{(\nu_p)}\big]^T$ and $\boldsymbol{\beta}_m^{(\nu_p)} \triangleq \big[B^{(\nu_p)}[m,0] \cdots B^{(\nu_p)}[m,L-1]\big]^T$ and the unitary $L \times L$ matrix $\mathbf{B}_m$ with elements $(\mathbf{B}_m)_{i,l} = b^*_{m,i}[l]$. We can now state the basis optimization problem as follows: for given vectors $\boldsymbol{\beta}_m^{(\nu_p)}$, $m = 0,\ldots,D-1$, find $L \times L$ unitary matrices $\mathbf{B}_m$ not dependent on ν_p such that the vectors $\tilde{\boldsymbol{\alpha}}_m^{(\nu_p)} = \mathbf{B}_m\,\boldsymbol{\beta}_m^{(\nu_p)}$ are maximally sparse for all $\nu_p \in \mathscr{D}$. As is usual in the CS framework, we measure the sparsity of $\tilde{\boldsymbol{\alpha}}_m^{(\nu_p)}$ by the l_1 norm (more precisely, the l_1-norm averaged over all $\nu_p \in \mathscr{D}$), i.e., by $\frac{1}{|\mathscr{D}|}\sum_{\nu_p\in\mathscr{D}} \big\|\tilde{\boldsymbol{\alpha}}_m^{(\nu_p)}\big\|_1 = \frac{1}{|\mathscr{D}|}\sum_{\nu_p\in\mathscr{D}} \big\|\mathbf{B}_m\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1$. Thus, the optimization problem reads

$$\hat{\mathbf{B}}_m = \operatorname*{arg\,min}_{\mathbf{B}_m \in \mathscr{U}} \sum_{\nu_p\in\mathscr{D}} \big\|\mathbf{B}_m\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1, \quad m = 0,\ldots,D-1, \tag{28}$$

where $\mathscr{U}$ denotes the set of all unitary $L \times L$ matrices.

5.2 Iterative Optimization Algorithm

Because the minimization problem (28) is nonconvex (since $\mathscr{U}$ is not a convex set), we propose an approximate iterative algorithm that relies on the following facts [14]: (i) every unitary $L \times L$ matrix $\mathbf{B}$ can be represented in terms of a Hermitian $L \times L$ matrix $\mathbf{A}$ as $\mathbf{B} = e^{j\mathbf{A}}$; (ii) the matrix exponential $\mathbf{B} = e^{j\mathbf{A}}$ can be approximated by its first-order Taylor expansion, i.e.,

$$\mathbf{B} \approx \mathbf{I}_L + j\mathbf{A}, \tag{29}$$

where $\mathbf{I}_L$ is the $L \times L$ identity matrix. Even though $\mathbf{I}_L + j\mathbf{A}$ is not a unitary matrix, the approximation (29) will be good if $\mathbf{A}$ is small. Because of this condition, we construct $\mathbf{B}_m$ iteratively: starting with the Fourier basis, we perform a *small* update at each iteration, using the approximation (29) in the optimization criterion but not for actually updating $\mathbf{B}_m$ (thus, the iterated $\mathbf{B}_m$ is always unitary).

More specifically, at the ith iteration, the idea is to update the unitary matrix $\mathbf{B}_m^{(i)}$ obtained at the $(i-1)$th iteration as $\mathbf{B}_m^{(i+1)} = e^{j\mathbf{A}_m}\,\mathbf{B}_m^{(i)}$—which is again unitary because $\mathbf{A}_m$ is Hermitian—and optimize $\mathbf{A}_m$ according to the criterion (28), i.e., by minimizing $\sum_{\nu_p\in\mathscr{D}} \big\|e^{j\mathbf{A}_m}\,\mathbf{B}_m^{(i)}\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1$. Since this problem is still nonconvex, we use the approximation (29), and thus the final minimization problem at the ith iteration is

$$\hat{\mathbf{A}}_m = \operatorname*{arg\,min}_{\mathbf{A}\in\mathscr{A}_i} \sum_{\nu_p\in\mathscr{D}} \big\|(\mathbf{I}_L + j\mathbf{A})\,\mathbf{B}_m^{(i)}\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1. \tag{30}$$

Here, $\mathscr{A}_i$ is the set of all Hermitian $L \times L$ matrices $\mathbf{A}$ that are small in the sense that $\|\mathbf{A}\|_\infty \leq \lambda_i$, where $\|\mathbf{A}\|_\infty$ denotes the largest modulus of all elements of $\mathbf{A}$ and λ_i is a positive constraint level (a small λ_i ensures a good approximation accuracy in (29) and that the unitary matrix $e^{j\hat{\mathbf{A}}_m}$ is close to $\mathbf{I}_L$). The problem (30) is convex and thus can be solved by standard convex optimization techniques [12].

Next, we test whether the cost function is smaller for the new unitary matrix $e^{j\hat{\mathbf{A}}_m}\,\mathbf{B}_m^{(i)}$, i.e., if

$$\sum_{\nu_p\in\mathscr{D}} \big\|e^{j\hat{\mathbf{A}}_m}\,\mathbf{B}_m^{(i)}\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1 < \sum_{\nu_p\in\mathscr{D}} \big\|\mathbf{B}_m^{(i)}\,\boldsymbol{\beta}_m^{(\nu_p)}\big\|_1.$$

In the positive case, we actually perform the update, i.e., set $\mathbf{B}_m^{(i+1)} = e^{j\hat{\mathbf{A}}_m}\,\mathbf{B}_m^{(i)}$, and we propagate the constraint level, i.e., $\lambda_{i+1} = \lambda_i$. Otherwise, we reject the update, i.e., set $\mathbf{B}_m^{(i+1)} = \mathbf{B}_m^{(i)}$, and we reduce the constraint level as $\lambda_{i+1} = \lambda_i/2$. This iteration process is terminated either if $\lambda_i < \varepsilon$ for a prescribed accuracy parameter ε or if the number of iterations exceeds a certain value. The iteration process is initialized by the $L \times L$ DFT matrix $\mathbf{B}_m^{(0)} = \mathbf{F}_L$. As mentioned in Section 3, the Fourier basis already yields a relatively sparse coefficient vector, and thus it provides a reasonable initial choice. We note that efficient algorithms for computing the matrix exponentials $e^{j\hat{\mathbf{A}}_m}$ exist [14].

For classical CP-OFDM with CP length $N-K \geq D-1$, we have $A_{\gamma,g}(m,\xi) = A_{\gamma,g}(0,\xi)$ for all $m = 1,\ldots,D-1$, so $B^{(\nu_p)}[m,l] = B^{(\nu_p)}[0,l]$ (see (20)) and thus $\boldsymbol{\beta}_m^{(\nu_p)} = \boldsymbol{\beta}_0^{(\nu_p)}$. Because $\boldsymbol{\beta}_m^{(\nu_p)}$ no longer depends on m, only one basis $\mathbf{B}$ (instead of D different bases $\mathbf{B}_m$, $m = 0,\ldots,D-1$) has to be optimized.

6. SIMULATION RESULTS

We compare the performance of the proposed CS-based channel estimation method using an optimized basis, of the CS-based channel

[2] "Overwhelming probability" means that the probability of (26) not being true decreases exponentially with a growing number N_1 of selected rows.

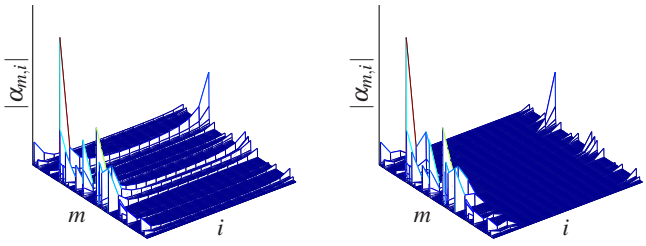

Figure 1: Modulus of the basis expansion coefficients $\alpha_{m,i}$ for the Fourier basis (left) and for the optimized basis (right).

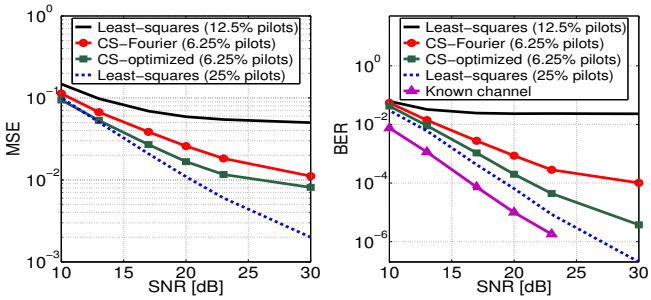

Figure 2: Performance of CS-based and least-squares channel estimation. Left: MSE versus SNR, right: BER versus SNR.

estimation method of [2] (using the implicit Fourier basis), and of classical least-squares (LS) channel estimation. In accordance with the DVB-T standard [8], we simulated a CP-OFDM system with $K=2048$ subcarriers and CP length $N-K=512$, whence $N=2560$. The system employed 4-QAM symbols with Gray labeling, a rate-$1/2$ convolutional code (generator polynomials $(13_8, 15_8)$), and 32×16 row-column interleaving. The interpolation filter $f(t)$ was chosen as a root-raised-cosine filter with roll-off factor $\rho = 1/4$.

During blocks of $L=16$ transmitted OFDM symbols, we simulated a noisy doubly selective channel whose discrete-delay-Doppler spreading function $S_h[m,i]$ was computed from (14) and (15). We assumed $P=20$ propagation paths with $(\tau_p/T_s, \nu_p T_s)$ randomly chosen within $[0,511]\times[-0.03/K, 0.03/K]$ for each block of 16 OFDM symbols (hence, the maximum Doppler normalized by the subcarrier spacing is $\pm 3\%$). The complex scatterer amplitudes η_p were randomly chosen from zero-mean, complex Gaussian distributions with three different variances (3 strong scatterers of equal power, 7 medium scatterers with 10 dB less power, and 10 weak scatterers with 20 dB less power).

Fig. 1 compares the expansion coefficients $\alpha_{m,i}$ (see (22)) obtained with the Fourier basis and the optimized basis for one channel realization. The minimization of (30) was carried out using the convex optimization package `CVX` [15]. It is seen that the basis optimization yields a significant improvement of sparsity.

For LS channel estimation, we used two different rectangular pilot constellations: (i) a pilot on every fourth subcarrier for each OFDM symbol, corresponding to 8192 pilots or 25% of all transmit symbols, and (ii) a pilot on every fourth subcarrier and every second OFDM symbol, corresponding to 4096 pilots or 12.5% of all symbols. For CS-based channel estimation, we placed uniformly at random 2048 pilots on a subsampled grid $\mathcal{G}$ with $\Delta K = 4$, corresponding to 6.25% of all symbols—one quarter of the number of pilots used in constellation (i) and half that used in constellation (ii). The same pilot constellation was used for both CS-based methods. CS-based estimation was implemented using the classical basis pursuit, i.e., (25) with $\varepsilon = 0$; this allows a faster implementation than extended basis pursuit with $\varepsilon > 0$. The MATLAB function `l1eq_pd()` from the toolbox ℓ_1-MAGIC [16] was used.

Fig. 2 depicts the mean square error (MSE) of the channel estimates and the bit error rate (BER) versus the channel signal-to-noise ratio (SNR). It is seen that both CS-based methods (with 6.25% pilots) significantly outperform the LS method with 12.5% pilots. The extremely poor performance of the LS method with 12.5% pilots is due to the fact that the Shannon sampling theorem is violated by the pilot grid. In contrast, the CS-based methods are able to produce reliable channel estimates even far below the Shannon sampling rate. Compared with the LS method with 25% pilots, we observe only a relatively small performance degradation of both CS-based methods with 6.25% pilots, especially in the low-to-medium SNR regime. The CS-based method using the optimized basis is seen to clearly outperform the CS-based method using the Fourier basis. This performance gain is due to the improved sparsity, and it is obtained in spite of the greater coherence of the optimized basis ($\mu = 2.237$) compared to that of the Fourier basis ($\mu = 1$).

7. CONCLUSION

CS-based channel estimation makes it possible to exploit the sparsity of wireless channels for a reduction of the number of pilots. We demonstrated that the inclusion of an optimized sparsity-improving basis expansion in the CS-based channel estimation method of [2] yields significant performance gains. The additional numerical complexity is moderate; in particular, the basis optimization has to be performed only once before the start of data transmission.

REFERENCES

[1] E. G. Larsson, G. Liu, J. Li, and G. B. Giannakis, "Joint symbol timing and channel estimation for OFDM based WLANs," *IEEE Comm. Letters*, vol. 5, pp. 325–327, Aug. 2001.

[2] G. Tauböck and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots," in *Proc. IEEE ICASSP-2008*, (Las Vegas, NV), pp. 2885–2888, March/Apr. 2008.

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

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

[5] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Comm.*, vol. 16, pp. 1579–1589, Oct. 1998.

[6] 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.

[7] IEEE P802 LAN/MAN Committee, "The working group for wireless local area networks (WLANs)." http://grouper.ieee.org/groups/802/11/index.html.

[8] 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).

[9] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

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

[11] 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.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge (UK): Cambridge Univ. Press, Dec. 2004.

[13] 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.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore: Johns Hopkins University Press, 2nd ed., 1989.

[15] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming (web page and software), Stanford University, CA (http://stanford.edu/~boyd/cvx).

[16] E. J. Candès and J. Romberg. Toolbox ℓ_1-MAGIC, California Inst. of Technol., Pasadena, CA (http://www.acm.caltech.edu/l1magic/).