in Proc. IEEE ICASSP-08
Las Vegas, NV, March–April 2008

A COMPRESSED SENSING TECHNIQUE FOR OFDM CHANNEL ESTIMATION IN MOBILE ENVIRONMENTS: EXPLOITING CHANNEL SPARSITY FOR REDUCING PILOTS

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: gtauboec@nt.tuwien.ac.at

ABSTRACT

We consider the estimation of doubly selective wireless channels within pulse-shaping multicarrier systems (which include OFDM systems as a special case). A new channel estimation technique using the recent methodology of *compressed sensing* (CS) is proposed. CS-based channel estimation exploits a channel's delay-Doppler sparsity to reduce the number of pilots and, hence, increase spectral efficiency. Simulation results demonstrate a significant reduction of the number of pilots relative to least-squares channel estimation.

***Index Terms*—** OFDM, multicarrier modulation, channel estimation, compressed sensing, sparse reconstruction, basis pursuit

1. INTRODUCTION

The recently introduced principle and methodology of *compressed sensing* (CS) allows the efficient reconstruction of sparse signals from a very limited number of measurements (samples) [1, 2]. CS has gained a fast-growing interest in applied mathematics. In this paper, we apply CS to pilot-based channel estimation in highly mobile environments. We consider pulse-shaping multicarrier (MC) systems, which include orthogonal frequency-division multiplexing (OFDM) as a special case [3]. Conventional methods for channel estimation (e.g., [4]) are not able to exploit the *inherent sparsity of the transmission channel* that is due to the sparse distribution of scatterers in space. As we will demonstrate, CS provides a constructive way for exploiting this sparsity in order to reduce the number of pilots and, hence, increase spectral efficiency.

This paper is organized as follows. In Section 2, some basic facts about CS are reviewed. The MC system model is described in Section 3. Section 4 studies the delay-Doppler sparsity of doubly selective channels. In Section 5, we present the CS-based channel estimation method. Finally, simulation results in Section 6 demonstrate performance gains relative to least-squares channel estimation.

2. REVIEW OF COMPRESSED SENSING

In a typical sparse reconstruction scenario, one tries to estimate a parameter vector $\mathbf{u} \in \mathbb{R}^M$ based on the linear model

$$\mathbf{v} = \mathbf{\Phi}\mathbf{u} + \mathbf{w}\,, \tag{1}$$

where $\mathbf{v} \in \mathbb{R}^N$ is an observation vector, $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ is a measurement matrix, and $\mathbf{w} \in \mathbb{R}^N$ is a noise vector. The reconstruction is subject to the constraint that $\mathbf{u}$ is S-sparse, i.e., at most S of its entries are nonzero. The positions of the nonzero entries are unknown. Typically, the number of parameters is much larger than the number of observations, i.e., $M \gg N$.

A key ingredient of CS is the *uniform uncertainty principle* [5], which essentially states that the measurement matrix $\mathbf{\Phi}$ obeys a "restricted isometry hypothesis." Let $\mathbf{\Phi}_{\mathcal{T}}$, $\mathcal{T} \subset \{1,\dots,M\}$ be the $N \times |\mathcal{T}|$ submatrix comprising those columns of $\mathbf{\Phi}$ that are indexed by the elements of $\mathcal{T}$. Then the *S-restricted isometry constant* δ_S of $\mathbf{\Phi}$ is defined as the smallest quantity δ_S such that

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

for all subsets $\mathcal{T}$ with $|\mathcal{T}| \le S$ and all vectors $\mathbf{a} \in \mathbb{R}^{|\mathcal{T}|}$.

We will use the estimator of $\mathbf{u}$ defined by the convex program

$$\hat{\mathbf{u}} \;\triangleq\; \arg\min_{\mathbf{u}\in\mathcal{U}_\epsilon} \|\mathbf{u}\|_1\,, \tag{2}$$

where $\mathcal{U}_\epsilon$ is the set of all $\mathbf{u} \in \mathbb{R}^M$ satisfying $\|\mathbf{\Phi}\mathbf{u} - \mathbf{v}\|_2 \le \epsilon$ for a given $\epsilon > 0$ [5]. This estimator—an extension of *basis pursuit* —is able to recover S-sparse parameter vectors according to the following result [5].

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

$$\delta_{3S} + 3\delta_{4S} < 2\,. \tag{3}$$

Let $\mathbf{v} = \mathbf{\Phi}\mathbf{u} + \mathbf{w}$ with $\|\mathbf{w}\|_2 \le \epsilon$, and let $\mathbf{u}_S \in \mathbb{R}^M$ contain the S components of $\mathbf{u}$ with largest absolute values, the remaining $M-S$ components being zero. Then the estimate $\hat{\mathbf{u}}$ in (2) satisfies

$$\|\hat{\mathbf{u}} - \mathbf{u}\|_2 \;\le\; C_1\epsilon + C_2\frac{\|\mathbf{u}-\mathbf{u}_S\|_1}{\sqrt{S}}\,, \tag{4}$$

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

For a zero-mean i.i.d. Gaussian noise vector, the condition $\|\mathbf{w}\|_2 \le \epsilon$ is satisfied with high probability for an appropriate ϵ. In the noiseless case ($\epsilon = 0$), (4) shows that $\hat{\mathbf{u}} = \mathbf{u}$ for any S-sparse vector $\mathbf{u}$.

Conditions under which the matrix $\mathbf{\Phi}$ satisfies (3) are obviously of interest. It has been shown [6] that if a complex-valued measurement matrix $\mathbf{\Phi}_{\mathrm{c}} \in \mathbb{C}^{N'\times M'}$ is constructed by selecting uniformly at random [1] N' rows from a unitary $M' \times M'$ matrix $\mathbf{U}$ and normalizing the columns (so that they have unit Euclidean norms), a sufficient condition for (3) to be true with overwhelming probability[1] is

$$N' \;\ge\; C_3\,(\ln M')^4 \mu^2 S\,. \tag{5}$$

Here, $\mu \triangleq \sqrt{M'}\max_{i,j}|U_{i,j}|$ (known as the *coherence* of $\mathbf{U}$) and C_3 is a constant. Later, we shall use the fact that under condition (5), the *real-valued* measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ (with $N = 2N'$, $M = 2M'$) that is obtained from $\mathbf{\Phi}_{\mathrm{c}}$ according to

$$\mathbf{\Phi} \;\triangleq\; \begin{bmatrix} \Re\{\mathbf{\Phi}_{\mathrm{c}}\} & -\Im\{\mathbf{\Phi}_{\mathrm{c}}\} \\ \Im\{\mathbf{\Phi}_{\mathrm{c}}\} & \Re\{\mathbf{\Phi}_{\mathrm{c}}\} \end{bmatrix} \tag{6}$$

also satisfies (3) with overwhelming probability. This follows from the special structure of $\mathbf{\Phi}$ (see, e.g., [7, Lemma 1]) and the fact that if a real-valued vector $\mathbf{u} = [u_1 \cdots u_{2M'}]^T$ is S-sparse, then so is the complex-valued vector $\mathbf{u}_{\mathrm{c}} \triangleq [u_1 \cdots u_{M'}]^T + j[u_{M'+1} \cdots u_{2M'}]^T$.

This work was supported by the WWTF projects MOHAWI (MA 44) and SPORTS (MA 07-004) and by the STREP project MASCOT (IST-026905) within the Sixth Framework Programme of the European Commission.

[1] "Overwhelming probability" means that the probability of (3) not being true decreases exponentially with an increasing number of selected rows, N'.

3. MULTICARRIER SYSTEM MODEL

3.1. Modulator, Channel, Demodulator

We consider a pulse-shaping MC system for the sake of generality and because of its advantages over conventional cyclic-prefix (CP) OFDM [3, 8]; however, CP-OFDM is included as a special case. The complex baseband domain is considered throughout. Let K, $N \geq K$, and L denote the number of subcarriers, the symbol duration, and the number of transmitted symbol periods, respectively. The *MC modulator* generates the discrete-time transmit signal

$$s[n] \;=\; \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} a_{l,k}\, g_{l,k}[n]\,, \qquad (7)$$

where $a_{l,k}$ with $l = 0,\ldots,L-1$ and $k = 0,\ldots,K-1$ 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 $g[n]$. Subsequently, $s[n]$ is converted into the continuous-time transmit signal $s(t) = \sum_{n=-\infty}^{\infty} s[n] f(t-nT_{\rm s})$, where $f(t)$ is an interpolation filter and $T_{\rm s}$ is the sampling period. For simplicity, we assume an ideal filter, i.e., $f(t) = \sqrt{1/T_{\rm s}}\,\mathrm{sinc}(\pi t/T_{\rm s})$ with $\mathrm{sinc}(x) \triangleq \frac{\sin x}{x}$.

The *channel* is assumed doubly selective/dispersive with time-varying impulse response $h(t,\tau)$. The channel output is

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

where $z(t)$ is zero-mean, stationary, white, rotationally invariant, complex Gaussian noise with power spectral density N_0.

At the receiver, the channel output $r(t)$ is converted into the discrete-time signal $r[n] = \int_{-\infty}^{\infty} r(t)\, f(t-nT_{\rm s})\, dt$, where $f(t)$ now serves as an anti-aliasing filter. Subsequently, the *MC demodulator* computes the inner products of $r[n]$ with time-frequency shifted versions $\gamma_{l,k}[n] \triangleq \gamma[n-lN]\, e^{j2\pi \frac{k}{K}(n-lN)}$ of a receive pulse $\gamma[n]$, i.e.,

$$x_{l,k} \;=\; \langle r, \gamma_{l,k}\rangle \;=\; \sum_{n=-\infty}^{\infty} r[n]\, \gamma_{l,k}^*[n]\,, \qquad (8)$$

for $l = 0,\ldots,L-1$ and $k = 0,\ldots,K-1$. The $x_{l,k}$ are finally equalized and quantized according to the data symbol alphabet.

In practice, CP-OFDM is typically used [9, 10]. This is a special case of our pulse-shaping MC setting; it is obtained for a rectangular transmit pulse $g[n]$ that is 1 on $[0, N-1]$ and 0 otherwise, and a rectangular receive pulse $\gamma[n]$ that is 1 on $[N-K, N-1]$ and 0 otherwise ($N-K \geq 0$ is the CP length).

By combining some of the equations presented earlier, a relation between the discrete-time signals $s[n]$ and $r[n]$ is obtained as

$$r[n] \;=\; \sum_{m=-\infty}^{\infty} h[n,m]\, s[n-m] \;+\; z[n]\,, \qquad (9)$$

with the discrete-time time-varying impulse response

$$h[n,m] \;=\; \int_{-\infty}^{\infty} h(nT_{\rm s},\tau)\, \mathrm{sinc}\Big(\pi\Big(m - \frac{\tau}{T_{\rm s}}\Big)\Big)\, d\tau\,. \qquad (10)$$

The discrete-time noise $z[n]$ is zero-mean, stationary, white, rotationally invariant, complex Gaussian with variance $\sigma_z^2 = N_0$.

3.2. System Channel

Next, we consider the *system channel* subsuming the MC modulator, the physical channel, and the MC demodulator. Combining (8), (9), and (7) and neglecting intersymbol and intercarrier interference (which is justified even in highly mobile environments if $g[n]$ and $\gamma[n]$ are properly designed), we obtain

$$x_{l,k} \;=\; H_{l,k}\, a_{l,k} \;+\; z_{l,k}\,, \qquad (11)$$

for $l = 0,\ldots,L-1$ and $k = 0,\ldots,K-1$. Here, $z_{l,k} = \langle z, \gamma_{l,k}\rangle$, and the system channel coefficients $H_{l,k}$ can easily be expressed in terms of $g[n]$, $h[n,m]$, and $\gamma[n]$ [3].

We will need a "delay-Doppler-domain expression" of the channel coefficients $H_{l,k}$. Let us assume that the receive pulse $\gamma[n]$ is zero outside $[0, L_\gamma]$. To compute $x_{l,k}$ in (8) for $l = 0,\ldots,L-1$, $r[n]$ must then be known for $n = 0,\ldots,N_r-1$, where $N_r \triangleq (L-1)N + L_\gamma + 1$. In this interval, we can express $r[n]$ as

$$r[n] \;=\; \sum_{m=-\infty}^{\infty} \sum_{i=0}^{N_r-1} S_h[m,i]\, s[n-m]\, e^{j2\pi \frac{in}{N_r}} \;+\; z[n]\,, \qquad (12)$$

with the *discrete-delay-Doppler spreading function* [11]

$$S_h[m,i] \;\triangleq\; \frac{1}{N_r} \sum_{n=0}^{N_r-1} h[n,m]\, e^{-j2\pi \frac{in}{N_r}}. \qquad (13)$$

Combining (8), (12), and (7), we reobtain the system channel relation (11) with $H_{l,k}$ expressed as

$$H_{l,k} \;=\; \sum_{m=-\infty}^{\infty} \sum_{i=0}^{N_r-1} F[m,i]\, e^{-j2\pi(\frac{km}{K} - \frac{Nli}{N_r})}, \qquad (14)$$

where

$$F[m,i] \;\triangleq\; S_h[m,i]\, A_{\gamma,g}^*\Big(m, \frac{i}{N_r}\Big) \qquad (15)$$

with the *cross-ambiguity function* [12] $A_{\gamma,g}(m,\xi) \triangleq \sum_{n=-\infty}^{\infty} \gamma[n]\, g^*[n-m]\, e^{-j2\pi\xi n}$. Using the approximation $N_r \approx LN$ (which is exact for CP-OFDM), we can write (14) as the two-dimensional discrete Fourier transform

$$H_{l,k} \;=\; \sum_{m=0}^{K-1} \sum_{i=0}^{L-1} \tilde{F}[m,i]\, e^{-j2\pi(\frac{km}{K} - \frac{li}{L})}\,, \qquad (16)$$

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

$$\tilde{F}[m,i] \;\triangleq\; \sum_{q=0}^{N-1} F[m, i+qL]\,, \quad i = 0,\ldots,L-1\,. \qquad (17)$$

4. DELAY-DOPPLER SPARSITY

We assume that the channel comprises P propagation paths corresponding to P specular (point) scatterers with fixed delays τ_p and Doppler frequency shifts ν_p for $p = 1,\ldots,P$. This simple structure is often a good approximation to real mobile radio channels. The channel impulse response then has the form

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

where η_p characterizes the attenuation and initial phase of path p. The discrete-time impulse response (10) becomes

$$h[n,m] \;=\; \sum_{p=1}^{P} \eta_p\, e^{j2\pi\nu_p n T_{\rm s}}\, \mathrm{sinc}\Big(\pi\Big(m - \frac{\tau_p}{T_{\rm s}}\Big)\Big)\,. \qquad (18)$$

Furthermore, inserting (18) into (13) and applying the geometric sum formula, the delay-Doppler spreading function is obtained as

$$S_h[m,i] \;=\; \sum_{p=1}^{P} \eta_p\, e^{j\pi(\nu_p T_{\rm s} - \frac{i}{N_r})(N_r-1)}\, \Lambda_p[m,i]\,, \qquad (19)$$

with

$$\Lambda_p[m,i] \;\triangleq\; \mathrm{sinc}\Big(\pi\Big(m - \frac{\tau_p}{T_{\rm s}}\Big)\Big)\, \mathrm{dir}_{N_r}\big(\pi(i - \nu_p T_{\rm s} N_r)\big)\,, \qquad (20)$$

where $\mathrm{dir}_N(x) \triangleq \frac{1}{N}\, e^{j\frac{x}{N}(N-1)} \sum_{n=0}^{N-1} e^{-j2\frac{x}{N}n} = \frac{\sin(x)}{N\sin(x/N)}$.

In the following, we investigate the sparsity of $S_h[m,i]$. In view of (19), we first consider the sparsity of $\Lambda_p[m,i]$. Using the appropriate versions of Parseval's theorem, it can be readily shown that $\sum_{m=-\infty}^{\infty}\sum_{i=0}^{N_r-1}\Lambda_p^2[m,i]=1$, i.e., the total energy of $\Lambda_p[m,i]$ is 1. Next, we calculate an upper bound on the percentage of the energy of $\Lambda_p[m,i]$ that is located outside a rectangular neighborhood of the delay-Doppler point of the pth scatterer, $(\tau_p/T_\mathrm{s}, \nu_p T_\mathrm{s} N_r)$. To this end, we first consider the energy of those samples of $\mathrm{sinc}(\pi(m-\tau_p/T_\mathrm{s}))$ whose distance from τ_p/T_s is greater than $\Delta m \in \{2,3,\dots\}$. Let $\mathcal{M}$ denote the set of all integers $m\in\mathbb{Z}$ except those with $|m-\tau_p/T_\mathrm{s}|\le\Delta m$. We have the bound

$$\sum_{m\in\mathcal{M}}\Big|\mathrm{sinc}\Big(\pi\Big(m-\frac{\tau_p}{T_\mathrm{s}}\Big)\Big)\Big|^2 \le 2\sum_{i=\Delta m}^{\infty}\frac{1}{(\pi m)^2} \le \frac{2}{\pi^2}\int_{\Delta m-1}^{\infty}\frac{dx}{x^2} = \frac{2}{\pi^2(\Delta m-1)}, \tag{21}$$

where $\sin^2 x\le 1$ and some monotonicity arguments have been used. In a similar manner, we consider the energy of those samples of $\mathrm{dir}_{N_r}\big(\pi(i-\nu_p T_\mathrm{s} N_r)\big)$ whose distance (up to the modulo-N_r operation, see below) from $\nu_p T_\mathrm{s} N_r$ is greater than $\Delta i\in\{2,\dots,N_r/2\}$ (N_r is assumed even). Let $\mathcal{I}$ denote the set $\{0,\dots,N_r-1\}$ with the exception of all $i=i_\mathbb{Z} \bmod N_r$, where $i_\mathbb{Z}$ is any integer with $|i_\mathbb{Z}-\nu_p T_\mathrm{s} N_r|\le\Delta i$. We have the bound

$$\begin{aligned}\sum_{i\in\mathcal{I}}\big|\mathrm{dir}_{N_r}\big(\pi(i-\nu_p T_\mathrm{s} N_r)\big)\big|^2 &\le \frac{2}{N_r^2}\sum_{i=\Delta i}^{N_r/2}\frac{1}{\sin^2\big(\frac{\pi}{N_r}i\big)}\\ &\le \frac{2}{N_r^2}\int_{\Delta i-1}^{N_r/2}\frac{dx}{\sin^2\big(\frac{\pi}{N_r}x\big)} = \frac{2}{N_r\pi}\cot\Big(\frac{\pi}{N_r}(\Delta i-1)\Big)\\ &\le \frac{1}{\pi(\Delta i-1)},\end{aligned} \tag{22}$$

where we have used $\sin^2 x\le 1$, some monotonicity arguments, and $\cot x\le\frac{\pi}{2x}$ within $[0,\frac{\pi}{2}]$. Combining (20)–(22), we see that

$$\sum_{(m,i)\in\mathcal{M}\times\mathcal{I}}\Lambda_p^2[m,i] \le \frac{2}{\pi^3(\Delta m-1)(\Delta i-1)}. \tag{23}$$

Thus, at most a percentage of $\frac{2}{\pi^3(\Delta m-1)(\Delta i-1)}\cdot 100\,\%$ of the energy of $\Lambda_p[m,i]$ is located outside a rectangular neighborhood of $(\tau_p/T_\mathrm{s}, \nu_p T_\mathrm{s} N_r)$ comprising $(2\Delta m+1)(2\Delta i+1)$ samples. Since the domain of $\Lambda_p[m,i]$ is $\mathbb{Z}\times\{0,\dots,N_r-1\}$, with N_r typically large, we conclude that $\Lambda_p[m,i]$ is an *approximately sparse* function.

Hereafter, therefore, we will consider $\Lambda_p[m,i]$ as N_Λ-sparse, with an appropriately chosen number N_Λ of nonzero samples. The energy bound (23) allows us to choose N_Λ such that a prescribed approximation quality can be guaranteed. From (19), it then follows that $S_h[m,i]$ is PN_Λ-sparse, and the same is true for $F[m,i]$ in (15) and, in turn, for $\tilde{F}[m,i]$ in (17).

5. CS-BASED CHANNEL ESTIMATION

To develop the new channel estimation method, we first show that pilot-based channel estimation can be formulated as a sparse reconstruction problem of the type discussed in Section 2.

For practical channels as well as transmit and receive pulses, the function $F[m,i]$ in (15) is effectively supported in a subdomain of the delay-Doppler plane. This allows us to perform a subsampling in the time-frequency domain. Thus, we assume that the support of $\tilde{F}[m,i]$ is contained in $[0,D-1]\times\big([0,I/2-1]\cup[L-I/2,L-1]\big)$, where $D\le K$, I is even, and D and I are chosen such that $\Delta K\triangleq K/D$ and $\Delta L\triangleq L/I$ are integers. Because of (16), the channel coefficient function $H_{l,k}$ is then uniquely specified by its values on the subsampled grid $(l,k)=(l'\Delta L, k'\Delta K)$ with $l'=0,\dots,I-1$, $k'=0,\dots,D-1$. Furthermore, (16) entails the relation

$$H_{l'\Delta L,k'\Delta K} = (-1)^{l'}\sum_{m=0}^{D-1}\sum_{i=0}^{I-1}F'[m,i-I/2]\,e^{-j2\pi(\frac{k'm}{D}-\frac{l'i}{I})}, \tag{24}$$

where $F'[m,i]$ is a cyclically extended version of $\tilde{F}[m,i]$ that equals $\tilde{F}[m,i]$ for $i=0,\dots,L-1$ and $\tilde{F}[m,i+L]$ for $i=-L,\dots,-1$.

Suppose that pilot symbols $a_{l,k}=p_{l,k}$ are transmitted at time-frequency positions $(l,k)\in\mathcal{P}$, where $\mathcal{P}$ is a subset of the subsampled grid $(l'\Delta L, k'\Delta K)$, $l'=0,\dots,I-1$, $k'=0,\dots,D-1$. The $|\mathcal{P}|$ pilots $p_{l,k}$ and their time-frequency positions are known to the receiver. From (11), $x_{l,k}=H_{l,k}\,p_{l,k}+z_{l,k}$ for $(l,k)\in\mathcal{P}$. The receiver calculates channel coefficient estimates $\hat{H}_{l,k}$ at the pilot positions $(l,k)\in\mathcal{P}$ according to

$$\hat{H}_{l,k}\triangleq\frac{x_{l,k}}{p_{l,k}} = H_{l,k}+\frac{z_{l,k}}{p_{l,k}}, \qquad (l,k)\in\mathcal{P}. \tag{25}$$

Thus, these $H_{l,k}$ are known up to the additive noise terms $z_{l,k}/p_{l,k}$.

Next, we use the fact that all $H_{l'\Delta L,k'\Delta K}$ can be expressed via the Fourier transform relation (24). In vector notation, (24) reads

$$\mathbf{h} = \sqrt{ID}\,\mathbf{U}\mathbf{u}_\mathrm{c}, \tag{26}$$

where we defined (i) the ID-dimensional complex "parameter" vector $\mathbf{u}_\mathrm{c}\triangleq\big[\mathbf{u}_\mathrm{c}^T(0)\cdots\mathbf{u}_\mathrm{c}^T(D-1)\big]^T$ with $\mathbf{u}_\mathrm{c}(m)\triangleq\big[F'[m,-I/2]\cdots F'[m,I/2-1]\big]^T$; (ii) the ID-dimensional complex channel vector $\mathbf{h}\triangleq\big[\mathbf{h}(0)\cdots\mathbf{h}(D-1)\big]^T$ with $\mathbf{h}(k')\triangleq\big[H_{0,k'\Delta K}\ H_{\Delta L,k'\Delta K}\cdots H_{(I-1)\Delta L,k'\Delta K}\big]$; and (iii) the $ID\times ID$ block matrix $\mathbf{U}$ with $I\times I$ blocks $\mathbf{U}_{k',m}\triangleq\frac{1}{\sqrt{D}}e^{-j2\pi\frac{(k'-1)(m-1)}{D}}\mathbf{SF}$ for $k',m=1,\dots,D$, where $\mathbf{F}$ is the $I\times I$ IDFT matrix with entries $(\mathbf{F})_{l',i}=\frac{1}{\sqrt{I}}e^{j2\pi\frac{(l'-1)(i-1)}{I}}$ for $l',i=1,\dots,I$ and $\mathbf{S}$ is the diagonal $I\times I$ matrix with diagonal entries $1,-1,1,-1,\dots$. We note that $\mathbf{U}$ is unitary with coherence $\mu=1$. Furthermore, according to the previous section, $\mathbf{u}_\mathrm{c}$ is modeled as PN_Λ-sparse.

$|\mathcal{P}|$ specific entries of the channel vector $\mathbf{h}$ are given by the channel coefficients $H_{l,k}$ at the $|\mathcal{P}|$ pilot positions $(l,k)\in\mathcal{P}$. Let $\mathbf{h}^{(\mathrm{p})}$ denote the corresponding $|\mathcal{P}|$-dimensional subvector of $\mathbf{h}$. Furthermore, let $\mathbf{\Phi}_\mathrm{c}$ denote the $|\mathcal{P}|\times ID$ submatrix of $\mathbf{U}$ obtained by selecting the corresponding $|\mathcal{P}|$ rows of $\mathbf{U}$ and multiplying the resulting matrix by $\sqrt{ID/|\mathcal{P}|}$ (this implies that the columns of $\mathbf{\Phi}_\mathrm{c}$ have unit Euclidean norm). We then have $\sqrt{1/|\mathcal{P}|}\,\mathbf{h}^{(\mathrm{p})}=\mathbf{\Phi}_\mathrm{c}\mathbf{u}_\mathrm{c}$, which, up to a constant factor, is (26) reduced to the pilot positions.

Next, we define $\mathbf{v}_\mathrm{c}$ as the vector $\sqrt{1/|\mathcal{P}|}\,\mathbf{h}^{(\mathrm{p})}$ but with the entries $H_{l,k}$ of $\mathbf{h}^{(\mathrm{p})}$ replaced by their estimates $\hat{H}_{l,k}$. Then, because of (25), $\mathbf{v}_\mathrm{c}=\sqrt{1/|\mathcal{P}|}\,\mathbf{h}^{(\mathrm{p})}+\mathbf{w}_\mathrm{c}$, where $\mathbf{w}_\mathrm{c}$ comprises the scaled noise terms $\sqrt{1/|\mathcal{P}|}\,z_{l,k}/p_{l,k}$, $(l,k)\in\mathcal{P}$. Inserting $\sqrt{1/|\mathcal{P}|}\,\mathbf{h}^{(\mathrm{p})}=\mathbf{\Phi}_\mathrm{c}\mathbf{u}_\mathrm{c}$, we obtain $\mathbf{v}_\mathrm{c}=\mathbf{\Phi}_\mathrm{c}\mathbf{u}_\mathrm{c}+\mathbf{w}_\mathrm{c}$. This is equivalently stated as

$$\mathbf{v} = \mathbf{\Phi}\mathbf{u}+\mathbf{w}, \tag{27}$$

where $\mathbf{u}\triangleq\big[\Re\{\mathbf{u}_\mathrm{c}^T\}\ \ \Im\{\mathbf{u}_\mathrm{c}^T\}\big]^T$, $\mathbf{v}\triangleq\big[\Re\{\mathbf{v}_\mathrm{c}^T\}\ \Im\{\mathbf{v}_\mathrm{c}^T\}\big]^T$, $\mathbf{w}\triangleq\big[\Re\{\mathbf{w}_\mathrm{c}^T\}\ \Im\{\mathbf{w}_\mathrm{c}^T\}\big]^T$, and $\mathbf{\Phi}$ (defined in (6)) are the real versions of $\mathbf{u}_\mathrm{c}$ etc. According to our construction, $\mathbf{u}$ is $2PN_\Lambda$-sparse. Thus, (27) is seen to be a sparse reconstruction problem of the form (1), with dimensions $M=2ID$ and $N=2|\mathcal{P}|$ and sparsity $S=2PN_\Lambda$.

We can hence use the extended basis pursuit (2) to obtain an estimate of $\mathbf{u}$ or, equivalently, of $F'[m,i]$. From this estimate, estimates of all channel coefficients $H_{l,k}$ are finally obtained via (16). The convex program (2) is a simple instance of a second-order cone program, which can be solved efficiently by interior-point methods [13]. Thus, CS-based channel estimation is computationally feasible, albeit more complex than classical least-squares channel estimation.

Regarding the choice of the pilot positions $(l,k) \in \mathcal{P}$, we recall that these positions correspond to $|\mathcal{P}|$ indices within the index range $\{1,\ldots,ID\}$ of the channel vector $\mathbf{h}$. To be consistent with the CS framework of Section 2, we select these positions uniformly at random [1]. For good approximation quality in the sense of (4), the number of pilots should satisfy condition (5), which becomes

$$|\mathcal{P}| \;\geq\; C_3 \left(\ln(ID)\right)^4 \cdot 2PN_\Lambda$$

(recall that $\mu = 1$). This bound is not useful for actually determining $|\mathcal{P}|$ because of the constant C_3. However, the bound suggests that the required number of pilots scales only linearly with the number P of channel paths (scatterers) and the sparsity parameter N_Λ, and poly-logarithmically with the system design parameters I and D. In practice, the pilot positions will be randomly chosen (and communicated to the receiver) 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.

6. SIMULATION RESULTS

We next present numerical results to compare the performance of the proposed CS-based channel estimation method with that of classical least-squares channel estimation. In accordance with the DVB-T standard [10], we simulated a CP-OFDM system with $K = 2048$ subcarriers and CP length $N - K = 512$, whence $N = 2560$. The system employed a 4-QAM symbol alphabet.

During blocks of $L = 2$ transmitted OFDM symbols, we simulated a noisy doubly selective/dispersive channel whose discrete-delay-Doppler spreading function $S_h[m,i]$ was computed from (19) and (20). We assumed $P = 20$ propagation paths whose (continuous-valued) delays τ_p/T_s and Doppler frequencies $\nu_p T_\mathrm{s}$ were randomly chosen within $[0,511] \times [-0.02/K, 0.02/K]$ for each block of 2 OFDM symbols (hence, the maximum allowed normalized Doppler frequency—maximum Doppler divided by the subcarrier spacing—is $\pm 2\%$). The complex scatterer amplitudes η_p were randomly chosen from zero-mean, rotationally invariant, complex Gaussian distributions with three different variances. More precisely, we assumed 3 strong scatterers of equal power, 7 medium scatterers with power 10 dB below that of the strong scatterers, and 10 weak scatterers with power 20 dB below that of the strong scatterers.

For least-squares channel estimation, we used two different rectangular pilot constellations, namely, (i) with spacings $\Delta L = 1$ and $\Delta K = 4$, corresponding to 1024 pilots or 25% of all available transmit symbols, and (ii) with spacings $\Delta L = 1$ and $\Delta K = 8$, corresponding to only 512 pilots or 12.5% of all symbols. For CS-based channel estimation, we selected uniformly at random 511 pilots from the rectangular pilot constellation (i), corresponding to 12.48% of all symbols—i.e., about the number of pilots used in constellation (ii) and about half that used in constellation (i). These three pilot constellations were held fixed during all simulation runs. The CS-based method used the classical basis pursuit, i.e., (2) with $\epsilon = 0$; this allows a faster implementation than extended basis pursuit with $\epsilon > 0$. The MATLAB function `l1eq_pd()` from the toolbox ℓ_1-MAGIC [14] was employed.

Fig. 1 depicts the mean square error (MSE) and the symbol error rate (SER) versus the signal-to-noise ratio (SNR). It is seen that the CS-based method (with 511 pilots) significantly outperforms the least-squares method with 512 pilots. The extremely poor performance of the least-squares method is due to the fact that the Shannon sampling theorem is violated by the pilot grid. In contrast, the CS-based method is able to produce reliable channel estimates even far below the Shannon sampling rate. Compared with the least-squares method with 1024 pilots, we observe only a relatively small performance degradation of the CS-based method with 511 pilots, especially in the low-to-medium SNR regime (up to 20 dB).

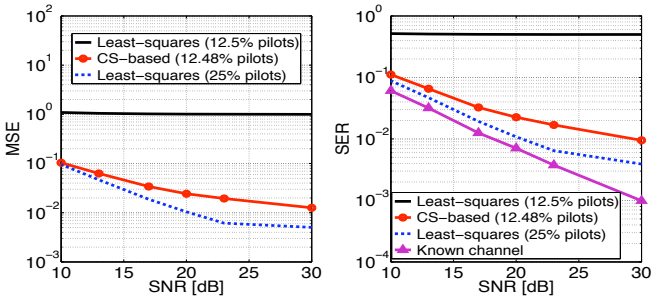

Fig. 1. Performance of CS-based and conventional least-squares channel estimation. Left: MSE versus SNR, right: SER versus SNR.

7. CONCLUSION

We have proposed a channel estimation technique based on the recently introduced principle of *compressed sensing* (CS). Our results demonstrate that CS makes it possible to exploit the "delay-Doppler sparsity" of wireless channels for a reduction of the number of pilots required for channel estimation within multicarrier systems. We conjecture that the performance of the proposed technique can be further improved by using more sophisticated CS methods.

8. ACKNOWLEDGMENT

The authors would like to thank G. Matz, H. Rauhut, and P. Fertl for fruitful discussions.

9. REFERENCES

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

[6] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and gaussian measurements," in *Proc. 40th Annual Conf. Information Sciences and Systems*, pp. 207–212, March 2006.

[7] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, pp. 585–596, Nov. 1999.

[8] 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.

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

[10] 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).

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

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

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

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