in Proc. IEEE SPAWC 2010
Marrakech, Morocco, June 2010

GROUP SPARSITY METHODS FOR COMPRESSIVE CHANNEL ESTIMATION IN DOUBLY DISPERSIVE MULTICARRIER SYSTEMS

Daniel Eiwen[a], *Georg Tauböck*[b], *Franz Hlawatsch*[b], *and Hans Georg Feichtinger*[a]

[a]NuHAG, Faculty of Mathematics, University of Vienna; daniel.eiwen@univie.ac.at
[b]Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology; gtauboec@nt.tuwien.ac.at

ABSTRACT

We propose advanced compressive estimators of doubly dispersive channels within multicarrier communication systems (including classical OFDM systems). The performance of compressive channel estimation has been shown to be limited by leakage components impairing the channel's effective delay-Doppler sparsity. We demonstrate a group sparse structure of these leakage components and apply recently proposed recovery techniques for group sparse signals. We also present a basis optimization method for enhancing group sparsity. Statistical knowledge about the channel can be incorporated in the basis optimization if available. The proposed estimators outperform existing compressive estimators with respect to estimation accuracy and, in one instance, also computational complexity.

Index Terms**—** OFDM, multicarrier modulation, channel estimation, doubly dispersive channel, doubly selective channel, compressed sensing, group sparsity, block sparsity, union of subspaces.

1. INTRODUCTION

The methodology of *group sparse compressed sensing* (GSCS) [1] allows the efficient reconstruction of signals whose support is contained in the union of a small number of groups (sets) from a collection of predefined disjoint groups. GSCS is closely related to *block sparse compressed sensing* [2, 3], *model-based compressed sensing* [4], and *recovery of signals from a union of subspaces*, e.g., [2]. Conventional compressed sensing (CS) is a special case of GSCS, with each group containing only a single element. Several CS recovery algorithms like *basis pursuit denoising* (BPDN) [5] and *orthogonal matching pursuit* (OMP) [6] have been extended to the group sparse case [2, 3]. In this paper, we apply GSCS to the estimation of doubly dispersive (doubly selective, doubly spread) channels [7]. We consider pulse-shaping multicarrier (MC) systems, which include classical OFDM systems [8] as a special case.

Compressive channel estimation (e.g., [9–15]) uses CS recovery algorithms for channel estimation, exploiting the fact that many wireless channels tend to be dominated by a relatively small number of clusters of significant paths [16]. As observed in [9, 10], the performance of compressive channel estimation tends to be limited by leakage effects which are caused by the finite bandwidth and block-length, and which impair the effective delay-Doppler sparsity of doubly dispersive channels. We have previously shown [10] that leakage effects can be significantly mitigated by the use of a basis expansion that is optimized for maximum sparsity.

Here, we explore a new approach (which can still be combined with an optimized basis expansion). We demonstrate that the leakage components in the channel's delay-Doppler representation exhibit a *group sparse structure*, which can be exploited by the use of GSCS recovery techniques. GSCS techniques are interesting also because the propagation paths of typical wireless channels are often structured in clusters; this is an additional, physical reason why the channel's delay-Doppler components occur in groups. As we will demonstrate, GSCS-based channel estimation outperforms CS-based channel estimation. Furthermore, one specific GSCS technique is significantly less complex than its CS counterpart.

We also present a refined GSCS-based channel estimator using an optimized basis expansion. This is based on a group sparsity extension of the basis optimization method of [10], whose optimality criterion is maximum group sparsity instead of maximum sparsity. If available, statistical knowledge about the channel can be incorporated in the basis optimization.

This paper is organized as follows. In Section 2, we recall the MC system model. Some GSCS fundamentals are reviewed in Section 3. In Section 4, the GSCS-based channel estimator is presented. Section 5 investigates leakage effects and the group sparse structure of doubly dispersive channels. A basis optimization method for enhancing group sparsity is described in Section 6. Finally, simulation results demonstrating performance gains achieved by the proposed GSCS methods are presented in Section 7.

2. MULTICARRIER SYSTEM MODEL

We consider a pulse-shaping MC system [17, 18] with K subcarriers, symbol duration $N \geq K$, transmit pulse $g[n]$, and receive pulse $\gamma[n]$. A block of L MC symbols is transmitted. Let $a_{l,k} \in \mathcal{A}$ ($l = 0, \ldots, L-1$; $k = 0, \ldots, K-1$) denote the complex data symbols, drawn from a finite symbol alphabet $\mathcal{A}$. As described in [10], a discrete-time transmit signal $s[n]$ is calculated from the data symbols $a_{l,k}$ using the transmit pulse $g[n]$, and converted to a continuous-time transmit signal $s(t)$ via an interpolation filter with impulse response $f_1(t)$ and sampling period T_s. The continuous-time receive signal obtained at the output of a noisy, doubly dispersive (or doubly selective) channel is given by

$$r(t) \,=\, \int_{-\infty}^{\infty} h(t,\tau)\, s(t-\tau)\, d\tau \,+\, z(t)\,, \tag{1}$$

where $h(t, \tau)$ is the channel's time-varying impulse response and $z(t)$ is complex noise. At the receiver, $r(t)$ is converted to the discrete-time receive signal $r[n]$ by means of an anti-aliasing filter with impulse response $f_2(t)$ and sampling with period T_s. From $r[n]$, demodulated symbols $r_{l,k}$ ($l = 0, \ldots, L-1$; $k = 0, \ldots, K-1$) are calculated using the receive pulse $\gamma[n]$. Finally, the demodulated symbols $r_{l,k}$ are equalized and quantized according to the data symbol alphabet $\mathcal{A}$. Details about the MC modulator, interpolation filter, anti-aliasing filter, and MC demodulator can be found in [10].

This work was supported by WWTF grant MA 07-004 (SPORTS) and by FWF grants S10602-N13 and S10603-N13 within NFN SISE.

Neglecting intersymbol and intercarrier interference—which is justified if the channel dispersion is not too strong—and assuming a causal channel with maximum delay at most $(K-1)T_{\mathrm{s}}$, the data symbols $a_{l,k}$ and demodulated symbols $r_{l,k}$ are approximately related as

$$r_{l,k} \,=\, H_{l,k}\, a_{l,k} + z_{l,k}\,, \tag{2}$$

for $l=0,\ldots,L-1$ and $k=0,\ldots,K-1$ [17]. Here, the time-frequency channel coefficients $H_{l,k}$ describe an equivalent "system channel" that subsumes the MC modulator, interpolation filter, physical channel, anti-aliasing filter, and MC demodulator, and $z_{l,k}$ is an equivalent discrete noise sequence. It can be shown [10] that

$$H_{l,k} \,=\, \sum_{m=0}^{K-1} \sum_{i=-L/2}^{L/2-1} F[m,i]\, e^{-j2\pi(\frac{km}{K} - \frac{li}{L})} \tag{3}$$

(L is assumed even for mathematical convenience), with

$$F[m,i] \,\triangleq\, \sum_{q=0}^{N-1} S_h[m,i+qL]\, A^*_{\gamma,g}\Big(m, \frac{i+qL}{LN}\Big)\,. \tag{4}$$

Here, $S_h[m,i]$ denotes the *discrete-delay-Doppler spreading function* [7] (which depends on $f_1(t)$, $h(t,\tau)$, and $f_2(t)$ as discussed in [10]) and $A_{\gamma,g}(m,\xi) \triangleq \sum_{n=-\infty}^{\infty} \gamma[n]\, g^*[n-m]\, e^{-j2\pi\xi n}$ is the *cross-ambiguity function* [19] of $\gamma[n]$ and $g[n]$. The function $F[m,i]$ is a representation of the system channel in terms of the discrete-delay variable m and the discrete-Doppler variable i.

We note that traditional cyclic-prefix (CP) OFDM is a special case of the pulse-shaping MC framework, corresponding to a rectangular transmit pulse $g[n]$ that is 1 for $n=0,\ldots,N-1$ and 0 otherwise, and a rectangular receive pulse $\gamma[n]$ that is 1 for $n = N-K,\ldots,N-1$ and 0 otherwise ($N-K \geq 0$ is the CP length).

3. REVIEW OF COMPRESSED SENSING FOR GROUP SPARSE SIGNALS

To prepare the ground for the compressive channel estimators proposed in Sections 4 and 6, we next review some GSCS fundamentals. Let $\mathcal{J} \triangleq \{\mathcal{I}_b\}_{b=1}^{B}$ be a partition of the set $\{1,\ldots,M\}$, i.e., $\bigcup_{b=1}^{B} \mathcal{I}_b = \{1,\ldots,M\}$ and $\sum_{b=1}^{B} |\mathcal{I}_b| = M$. For a vector $\mathbf{x} \in \mathbb{C}^M$, we define $\mathbf{x}[b] \in \mathbb{C}^{|\mathcal{I}_b|}$ as the subvector of $\mathbf{x}$ that comprises all entries x_i whose indices i are in the *group* $\mathcal{I}_b$. Then, $\mathbf{x} \in \mathbb{C}^M$ is called *group S-sparse over* $\mathcal{J}$, with $0 \leq S \leq B$, if at most S group subvectors $\mathbf{x}[b]$ are nonzero (i.e., not identically zero).

A typical GSCS reconstruction task is to estimate from a measured vector $\mathbf{y} \in \mathbb{C}^Q$ an unknown vector $\mathbf{x} \in \mathbb{C}^M$ that is (approximately) group S-sparse over $\mathcal{J}$, based on the linear model

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{z}\,. \tag{5}$$

Here, $\mathbf{\Phi} \in \mathbb{C}^{Q\times M}$ is a known measurement matrix and $\mathbf{z} \in \mathbb{C}^Q$ is an unknown noise vector. The "group sparsity" S is known but the group indices b for which the $\mathbf{x}[b]$ are nonzero are unknown. Typically, $Q \ll M$. Several GSCS reconstruction algorithms have been proposed recently. Here, we will consider the group variants of basis pursuit denoising (G-BPDN) [2] and orthogonal matching pursuit (G-OMP, also known as block orthogonal matching pursuit) [3].

Let $\mathbf{x} \in \mathbb{C}^M$, not necessarily sparse or group sparse. G-BPDN solves the convex problem

$$\min_{\mathbf{x}' \in \mathbb{C}^M} \|\mathbf{x}'\|_{2,\mathcal{J}} \quad \text{subject to } \|\mathbf{\Phi}\mathbf{x}' - \mathbf{y}\|_2 \leq \epsilon\,, \tag{6}$$

where $\|\mathbf{x}'\|_{2,\mathcal{J}} \triangleq \sum_{b=1}^{B} \|\mathbf{x}'[b]\|_2$. Comparing BPDN with G-BPDN, we see that the ℓ_1-norm is replaced by $\|\cdot\|_{2,\mathcal{J}}$. Let $\mathbf{x}^S$ denote the best group S-sparse approximation to $\mathbf{x}$ [2]. It is shown in [2] that the solution of (6), denoted $\hat{\mathbf{x}}$, satisfies

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \,\leq\, C_1 S^{-1/2}\, \|\mathbf{x} - \mathbf{x}^S\|_{2,\mathcal{J}} + C_2\, \epsilon\,,$$

provided that $\|\mathbf{\Phi}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon$ and $\delta_{2S|\mathcal{J}} < \sqrt{2}-1$. Here, the *group restricted isometry constant* $\delta_{2S|\mathcal{J}}$ is defined as the smallest constant δ such that the measurement matrix $\mathbf{\Phi}$ satisfies $(1-\delta)\,\|\mathbf{x}'\|_2^2 \leq \|\mathbf{\Phi}\mathbf{x}'\|_2^2 \leq (1+\delta)\,\|\mathbf{x}'\|_2^2$ for all $\mathbf{x}' \in \mathbb{C}^M$ that are group $2S$-sparse over $\mathcal{J}$. Furthermore, C_1 and C_2 decrease with decreasing $\delta_{2S|\mathcal{J}}$. Since every group $2S$-sparse vector over $\mathcal{J}$ is certainly also $2S'$-sparse, where S' is the sum of the S largest $|\mathcal{I}_b|$, it follows that $\delta_{2S|\mathcal{J}} \leq \delta_{2S'}$, where $\delta_{2S'}$ is the conventional (non-group) restricted isometry constant of $\mathbf{\Phi}$. We can expect that $\delta_{2S|\mathcal{J}} < \delta_{2S'}$, even though this has not been formally proven for most of the typical measurement matrices. This would imply the existence of improved performance guarantees for G-BPDN compared to BPDN.

G-OMP, on the other hand, extends OMP in that the index set corresponding to the nonzero components of the iterated $\mathbf{x}'$ is augmented not only by a single index in each recursion but by all indices of the best matching group. Performance guarantees are here phrased in terms of *group-coherence* and *sub-coherence* rather than in terms of the group restricted isometry constant $\delta_{2S|\mathcal{J}}$ [3]. An important advantage of G-OMP over OMP is its reduced computational complexity, especially for larger groups, since the search space for the best matching indices is substantially reduced.

4. COMPRESSIVE CHANNEL ESTIMATION EXPLOITING GROUP SPARSITY

We now present a GSCS-based channel estimator that is able to exploit group sparsity. This estimator generalizes the CS-based estimator of [9, 10]; as a consequence, the following development is largely parallel to (but less detailed than) that in [10].

For practical (underspread [7]) wireless channels and practical transmit and receive pulses, $F[m,i]$ in (4) is effectively supported in $\{0,\ldots,D-1\} \times \{-J/2,\ldots,J/2-1\}$ (within the fundamental i period), with some D and J that divide K and L, respectively (J is assumed even). Because of the 2-D DFT relation (3), the $H_{l,k}$ are then uniquely specified by their values on the *subsampled time-frequency grid* $\mathcal{G} = \{(l,k) = (\lambda\,\Delta L, \kappa\,\Delta K) : \lambda = 0,\ldots,J-1\,, \kappa = 0,\ldots,D-1\}$, where $\Delta L \triangleq L/J$ and $\Delta K \triangleq K/D$. Due to (3), these subsampled values are given by

$$H_{\lambda\,\Delta L,\kappa\,\Delta K} \,=\, \sum_{m=0}^{D-1} \sum_{i=-J/2}^{J/2-1} F[m,i]\, e^{-j2\pi(\frac{\kappa m}{D} - \frac{\lambda i}{J})}\,. \tag{7}$$

To mitigate leakage effects in $F[m,i]$ (see Section 5) and, thus, to enhance group sparsity, we generalize (7) to an orthonormal 2-D basis expansion:

$$H_{\lambda\,\Delta L,\kappa\Delta K} \,=\, \sum_{m=0}^{D-1} \sum_{i=-J/2}^{J/2-1} \alpha_{m,i}\, u_{m,i}[\lambda,\kappa]\,. \tag{8}$$

A suitable construction of the basis $\{u_{m,i}[\lambda,\kappa]\}$ will be presented in Section 6. Evidently, the 2-D DFT (7) is a special case of the general orthonormal 2-D basis expansion (8), with $\alpha_{m,i} = \sqrt{JD}\, F[m,i]$ and $u_{m,i}[\lambda,\kappa] = \frac{1}{\sqrt{JD}}\, e^{-j2\pi(\frac{\kappa m}{D} - \frac{\lambda i}{J})}$. The functions $H_{\lambda\,\Delta L,\kappa\,\Delta K}$ and $u_{m,i}[\lambda,\kappa]$ are defined for $\lambda = 0,\ldots,J-1$ and $\kappa = 0,\ldots,D-1$

and may thus be considered as $J \times D$ matrices. Let us form the vector $\mathbf{h} \triangleq \mathrm{vec}_{\lambda,\kappa}\{H_{\lambda\,\Delta L,\kappa\,\Delta K}\}$ of length JD by stacking all columns of the respective matrix (i.e., $\mathbf{h} = [h_1 \cdots h_{JD}]^T$ with $h_{\kappa J+\lambda+1} = H_{\lambda\,\Delta L,\,\kappa\,\Delta K}$) and the unitary $JD \times JD$ matrix $\mathbf{U}$ whose $\big((i + J/2)D + m + 1\big)$th column is given by the vector $\mathrm{vec}_{\lambda,\kappa}\{u_{m,i}[\lambda,\kappa]\}$. We can then rewrite (8) as

$$\mathbf{h} = \mathbf{U}\boldsymbol{\alpha}\,, \tag{9}$$

where $\boldsymbol{\alpha} \triangleq \mathrm{vec}_{m,i}\{\alpha_{m,i}\}$.

For pilot-aided channel estimation, we assume that pilot symbols $a_{l,k} = p_{l,k}$ are transmitted at time-frequency positions $(l,k) \in \mathcal{P}$, where $\mathcal{P} \subset \mathcal{G}$. Based on the known $p_{l,k}$, the receiver calculates estimates $\hat{H}_{l,k}$ of the channel coefficients $H_{l,k}$ at the pilot positions according to $\hat{H}_{l,k} \triangleq r_{l,k}/p_{l,k}$, $(l,k) \in \mathcal{P}$. Using (2), it follows that

$$\hat{H}_{l,k} = H_{l,k} + \frac{z_{l,k}}{p_{l,k}}\,, \quad (l,k) \in \mathcal{P}. \tag{10}$$

Thus, $|\mathcal{P}|$ of the JD entries of $\mathbf{h}$ are known up to noise. Let us reduce (9) to these pilot positions, i.e., we consider the reduced system $\mathbf{h}^{(\mathrm{p})} = \mathbf{U}^{(\mathrm{p})}\boldsymbol{\alpha}$ that is formed by the corresponding length-$|\mathcal{P}|$ subvector $\mathbf{h}^{(\mathrm{p})}$ of $\mathbf{h}$ and the corresponding $|\mathcal{P}| \times JD$ submatrix $\mathbf{U}^{(\mathrm{p})}$ of $\mathbf{U}$. Furthermore, we scale the columns of $\mathbf{U}^{(\mathrm{p})}$ so that they have unit ℓ_2-norm, i.e., $\boldsymbol{\Phi} \triangleq \mathbf{U}^{(\mathrm{p})}\mathbf{D}$ with a nonsingular diagonal scaling matrix $\mathbf{D}$. This yields $\mathbf{h}^{(\mathrm{p})} = \boldsymbol{\Phi}\mathbf{x}$, where $\mathbf{x} \triangleq \mathbf{D}^{-1}\boldsymbol{\alpha}$ equals $\boldsymbol{\alpha}$ up to (known) scaling factors. Finally, we consider instead of the unknown subvector $\mathbf{h}^{(\mathrm{p})}$ its pilot-based estimate, denoted as $\mathbf{y}$. Since the entries of $\mathbf{y}$ are given by the $\hat{H}_{l,k}$ in (10), we have

$$\mathbf{y} = \mathbf{h}^{(\mathrm{p})} + \mathbf{z} = \boldsymbol{\Phi}\mathbf{x} + \mathbf{z}\,, \tag{11}$$

where $\mathbf{z}$ is the vector with entries $z_{l,k}/p_{l,k}\big|_{(l,k)\in\mathcal{P}}$. Thus, we have obtained a *measurement equation* of the form (5), with dimensions $Q = |\mathcal{P}|$ and $M = JD$. In practice, $|\mathcal{P}| \ll JD$.

We will show in Section 5 that for a channel with a moderate number of scatterers, and using the 2-D DFT basis $u_{m,i}[\lambda,\kappa] = \frac{1}{\sqrt{JD}}\, e^{-j2\pi(\frac{\kappa m}{D} - \frac{\lambda i}{J})}$, $\mathbf{x}$ is (approximately) group sparse over a certain partition $\mathcal{J}$. The same will be true if $u_{m,i}[\lambda,\kappa]$ is constructed using the basis optimization algorithm presented in Section 6. We can therefore use G-BPDN, G-OMP, or any other GSCS recovery algorithm to obtain an estimate of $\mathbf{x}$ from $\mathbf{y}$, based on (11), and, in turn, to obtain an estimate of $\boldsymbol{\alpha} = \mathbf{D}\mathbf{x}$. Subsequently, using (8), we obtain estimates of the subsampled channel coefficients $H_{\lambda\,\Delta L,\kappa\,\Delta K}$. Finally, inverting (7) and using (3) yields estimates of all $H_{l,k}$.

For consistency with the standard construction of measurement matrices [20], the pilot positions $\mathcal{P}$ are randomly (uniformly) chosen from the subsampled grid $\mathcal{G}$. Accordingly, the measurement matrix $\boldsymbol{\Phi}$ is constructed by randomly (uniformly) drawing $|\mathcal{P}|$ rows from the unitary matrix $\mathbf{U}$ and normalizing the columns to unit ℓ_2-norm. For $|\mathcal{P}|$ chosen sufficiently large, this guarantees small restricted isometry constants $\delta_{2S|\mathcal{J}} \le \delta_{2S'}$ (see Section 3) with overwhelming probability [20]. As pointed out in Section 3, we expect an improved performance of GSCS recovery algorithms relative to their CS counterparts. This will be verified experimentally in Section 7. Furthermore, a specific GSCS recovery algorithm (G-OMP) is significantly less complex than its CS counterpart (OMP).

5. DELAY-DOPPLER GROUP SPARSITY

In this section, we analyze the group sparse structure of $\mathbf{x}$, assuming that the 2-D DFT basis $u_{m,i}[\lambda,\kappa] = \frac{1}{\sqrt{JD}} e^{-j2\pi(\frac{\kappa m}{D} - \frac{\lambda i}{J})}$ is used.

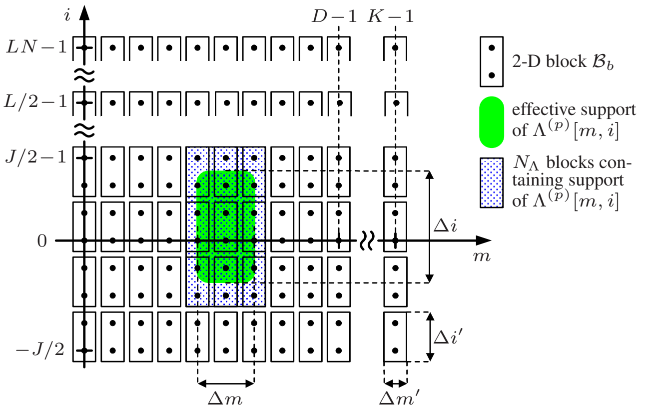


Fig. 1. Illustration of the 2-D block tiling $\{\mathcal{B}_b\}_{b=1}^{B}$, the effective support of a shifted leakage kernel $\Lambda^{(p)}[m,i]$, and the N_Λ blocks jointly containing the effective support of $\Lambda^{(p)}[m,i]$. In this figure, $\Delta m = 2$, $\Delta i = 4$, $\Delta m' = 1$, $\Delta i' = 2$, and $N_\Lambda = 9$.

This amounts to studying the group sparse structure of the delay-Doppler representation of the system channel, $F[m,i]$, because $\mathbf{x}$ is identical to $F[m,i]$ up to scalings and a stacking. We assume[1] that the physical channel comprises P paths corresponding to P specular scatterers characterized by time delay τ_p and Doppler shift ν_p for each $p = 1,\ldots,P$. The channel impulse response in (1) thus 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 the pth propagation path and $\delta(\cdot)$ is the Dirac delta. The discrete-delay-Doppler spreading function involved in (4) then becomes [10]

$$S_h[m,i] = \sum_{p=1}^{P} \eta_p\, e^{j\pi(\nu_p T_{\mathrm{s}} - \frac{i}{LN})(LN-1)} \Lambda^{(p)}[m,i]\,, \tag{12}$$

with the *shifted leakage kernels*

$$\Lambda^{(p)}[m,i] \triangleq \phi^{(\nu_p)}\Big(m - \frac{\tau_p}{T_{\mathrm{s}}}\Big)\, \psi(i - \nu_p T_{\mathrm{s}} LN)\,,$$

where $\phi^{(\nu)}(x) \triangleq \int_{-\infty}^{\infty} f_1(T_{\mathrm{s}}x - t)\, f_2(t)\, e^{-j2\pi\nu t}\, dt$ and $\psi(y) \triangleq \sin(\pi y)/[LN \sin(\pi y/(LN))]$. It is shown in [10] that, under suitable assumptions on $f_1(t)$ and $f_2(t)$, each $\Lambda^{(p)}[m,i]$ within $\{0,\ldots,K-1\} \times \{0,\ldots,LN-1\}$ is effectively supported in a box of size $\Delta m \times \Delta i$, where Δm and Δi can be chosen such that the energy of $\Lambda^{(p)}[m,i]$ outside the box is guaranteed to be below a prescribed threshold.

We now consider the tiling of $\mathbb{Z}^2$ into 2-D blocks $\mathcal{B}_b$ of equal size $\Delta m' \times \Delta i'$, where $\Delta m'$ and $\Delta i'$ divide D and $J/2$, respectively, and the tiling is such that $\{0,\ldots,\Delta m'-1\} \times \{0,\ldots,\Delta i'-1\}$ is one of these blocks. Such a tiling is depicted in Fig. 1. Then, the effective support of $\Lambda^{(p)}[m,i]$ within $\{0,\ldots,K-1\} \times \{0,\ldots,LN-1\}$ will be contained in at most N_Λ blocks $\mathcal{B}_b$, where $N_\Lambda = (\lceil \Delta m/\Delta m' \rceil + 1)(\lceil \Delta i/\Delta i' \rceil + 1)$. Furthermore, due to (12), the number of blocks jointly containing the effective support of $S_h[m,i]$ is at most PN_Λ.

[1] This assumption is only made for analyzing group sparsity and motivating the basis optimization method in this and the next section; it is not required for the proposed channel estimation methods.

Since $\Delta i'$ divides L, the summation in (4) defining $F[m,i]$ only adds whole blocks and therefore does not create any "new" nonzero blocks. Also, since $\Delta m'$ divides D and $\Delta i'$ divides $J/2$, the support restriction of $F[m,i]$ to $\{0,\dots,D-1\}\times\{-J/2,\dots,J/2-1\}$ (see Section 4) is compatible with the block boundaries (note that $S_h[m,i]$ is LN-periodic in i [10]). Therefore, the number of blocks jointly containing the effective support of $F[m,i]$ within $\{0,\dots,D-1\}\times\{-J/2,\dots,J/2-1\}$ is at most PN_Λ.

We will now show that the vector $\mathbf{x}$ in (11) is (effectively) group PN_Λ-sparse over a suitably defined partition $\mathcal{J}$. Let $\mathcal{B}_b$, $b=1,\dots,B$ (with $B\triangleq JD/(\Delta m'\Delta i')$) denote the 2-D blocks of size $\Delta m'\times\Delta i'$ tiling $\{0,\dots,D-1\}\times\{-J/2,\dots,J/2-1\}$ as described above. Furthermore, let us define the one-to-one mapping $\mathsf{S}:\{0,\dots,D-1\}\times\{-J/2,\dots,J/2-1\}\rightarrow\{1,\dots,JD\}$ given by $\mathsf{S}(m,i)=(i+J/2)D+m+1$. Note that S corresponds to the stacking operation carried out in Section 4, i.e., $[\boldsymbol{\alpha}]_{\mathsf{S}(m,i)}=[\boldsymbol{\alpha}]_{(i+J/2)D+m+1}=F[m,i]$ for the 2-D DFT basis. We now convert the 2-D blocks $\mathcal{B}_b$ of size $\Delta m'\times\Delta i'$ into the 1-D *groups*

$$\mathcal{I}_b \triangleq \mathsf{S}(\mathcal{B}_b)\subseteq\{1,\dots,JD\},$$

of size $|\mathcal{I}_b|=\Delta m'\Delta i'$. Then, $\mathcal{J}\triangleq\{\mathcal{I}_b\}_{b=1}^B$ constitutes a partition of $\{1,\dots,JD\}$. From the fact that the number of blocks containing the effective support of $F[m,i]$ within $\{0,\dots,D-1\}\times\{-J/2,\dots,J/2-1\}$ is at most PN_Λ, it then follows that $\boldsymbol{\alpha}$—and, in turn, $\mathbf{x}=\mathbf{D}^{-1}\boldsymbol{\alpha}$—is (effectively) group PN_Λ-sparse over $\mathcal{J}$.

6. GROUP SPARSITY OPTIMIZED BASIS EXPANSION

Next, we propose a group sparsity extension of the basis optimization algorithm presented in [10]. The goal is to enhance group sparsity by mitigating the leakage effects in $F[m,i]$. As before, we will use the partition $\mathcal{J}=\{\mathcal{I}_b\}_{b=1}^B$ of $\{1,\dots,JD\}$ defined by the groups $\mathcal{I}_b=\mathsf{S}(\mathcal{B}_b)$.

Let us consider the single-scatterer channel $h^{(\tau_1,\nu_1)}(t,\tau)=\delta(\tau-\tau_1)\,e^{j2\pi\nu_1 t}$, with random τ_1 and ν_1 that are distributed according to a known probability density function (pdf) $p(\tau_1,\nu_1)$. (A nonstatistical design that does not require knowledge of $p(\tau_1,\nu_1)$ is easily obtained by formally using a uniform pdf $p(\tau_1,\nu_1)$.) Clearly, the vector $\boldsymbol{\alpha}$ of expansion coefficients $\alpha_{m,i}$ in (8), (9) is now a random vector. In what follows, we will use a 2-D basis of the form

$$u_{m,i}[\lambda,\kappa] \triangleq \frac{1}{\sqrt{D}}\,v_{m,i}[\lambda]\,e^{-j2\pi\frac{\kappa m}{D}},$$

for $m=0,\dots,D-1$, $i=-J/2,\dots,J/2-1$, where $\{v_{m,i}[\lambda]\}_{m=0}^{D-1}$ is a family of D orthonormal 1-D bases (see [10] for a motivation of this construction). Our goal is to find 1-D basis functions $v_{m,i}[\lambda]$ such that the support of $\boldsymbol{\alpha}=\boldsymbol{\alpha}(\{v_{m,i}\})$ is contained in the union of a minimum number of groups $\mathcal{I}_b$ on average, i.e., such that $\boldsymbol{\alpha}$ is *maximally group sparse over $\mathcal{J}$ on average*. Note that this implies that also $\mathbf{x}=\mathbf{D}^{-1}\boldsymbol{\alpha}$ is maximally group sparse over $\mathcal{J}$ on average. Motivated by GSCS theory—see Section 3—we measure the group sparsity of $\boldsymbol{\alpha}$ by $\|\boldsymbol{\alpha}\|_{2,\mathcal{J}}$. Thus, we have to minimize $\mathsf{E}\{\|\boldsymbol{\alpha}(\{v_{m,i}\})\|_{2,\mathcal{J}}\}$ (with E denoting expectation with respect to the random variables τ_1 and ν_1) over the set of all families of orthonormal 1-D bases $\{v_{m,i}[\lambda]\}$.

For a convenient reformulation of this minimization problem, let $\mathbf{V}\triangleq\mathrm{diag}\{\mathbf{V}_0,\dots,\mathbf{V}_{D-1}\}\in\mathbb{C}^{JD\times JD}$ denote the block diagonal, unitary matrix whose blocks $\mathbf{V}_m\in\mathbb{C}^{J\times J}$ have entries $(\mathbf{V}_m)_{i+1,\lambda+1}\triangleq v^*_{m,i-J/2}[\lambda]$. Furthermore, we define a second partition $\mathcal{J}'\triangleq\{\mathcal{I}'_b\}_{b=1}^B$ of $\{1,\dots,JD\}$ by $\mathcal{I}'_b\triangleq\mathsf{S}'(\mathcal{B}_b)$ with $\mathsf{S}'(m,i)\triangleq mJ+i+J/2+1$. Note that whereas S corresponds to column-wise stacking, S' corresponds to row-wise stacking. One can then show the identity

$$\|\boldsymbol{\alpha}(\{v_{m,i}\})\|_{2,\mathcal{J}} = \|\mathbf{V}\mathbf{c}(\tau_1,\nu_1)\|_{2,\mathcal{J}'},$$

where $\mathbf{c}(\tau,\nu)\triangleq\big[\mathbf{c}_0^T(\tau,\nu)\cdots\mathbf{c}_{D-1}^T(\tau,\nu)\big]^T$ with

$$\mathbf{c}_m(\tau,\nu) \triangleq \sqrt{D}\,\phi^{(\nu)}\Big(m-\frac{\tau}{T_\mathrm{s}}\Big)\big[C^{(\nu)}[m,0]\cdots C^{(\nu)}[m,J-1]\big]^T.$$

Here,

$$C^{(\nu)}[m,\lambda] \triangleq \sum_{i=-J/2}^{J/2-1}\sum_{q=0}^{N-1}\psi^{(\nu)}(i+qL)\,A^*_{\gamma,g}\Big(m,\frac{i+qL}{LN}\Big)\,e^{j2\pi\frac{\lambda i}{J}}$$

with $\psi^{(\nu)}(i)\triangleq e^{j\pi(\nu T_\mathrm{s}-\frac{i}{LN})(LN-1)}\,\psi(i-\nu T_\mathrm{s}LN)$. Hence, we obtain the equivalent minimization problem

$$\hat{\mathbf{V}} = \underset{\mathbf{V}\in\mathcal{U}}{\arg\min}\ \mathsf{E}\big\{\|\mathbf{V}\mathbf{c}(\tau_1,\nu_1)\|_{2,\mathcal{J}'}\big\},$$

where $\mathcal{U}$ is the set of all block diagonal, unitary $JD\times JD$ matrices with blocks of equal size $J\times J$ on the diagonal. Using a Monte-Carlo approximation, we redefine this minimization problem as

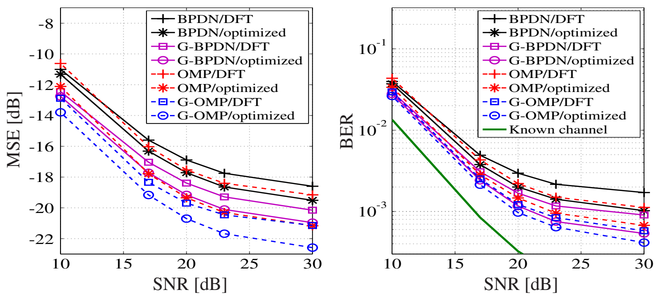
$$\hat{\mathbf{V}} \triangleq \underset{\mathbf{V}\in\mathcal{U}}{\arg\min}\sum_\rho\|\mathbf{V}\mathbf{c}(\tau_1,\nu_1)_\rho\|_{2,\mathcal{J}'}, \qquad (13)$$

in which the $(\tau_1,\nu_1)_\rho$ denote samples of the random vector (τ_1,ν_1) drawn from its pdf $p(\tau_1,\nu_1)$.

The minimization problem (13) is not convex, since $\mathcal{U}$ is not a convex set. However, extending [10], we can use an iterative algorithm that is based on the matrix exponential representation $\mathbf{V}=e^{j\mathbf{A}}$, where $\mathbf{A}$ is a block diagonal Hermitian matrix (note that the matrix exponential function preserves the block diagonal structure of a matrix). At each iteration, $\mathbf{A}$ is updated, using at a certain point the first-order Taylor approximation to the matrix exponential function. This update is a convex minimization problem, for which efficient methods exist. The $\mathbf{V}$ used for initializing this iterative algorithm is the block diagonal, unitary matrix in which each block on the diagonal is the $J\times J$ DFT matrix. Details of this algorithm (without the block diagonal structure) can be found in [10]. The computational complexity is significantly reduced by the fact that (13) can be decomposed into $D/\Delta m'$ separate minimization problems, each of dimension only $J\Delta m'\times J\Delta m'$ instead of $JD\times JD$, where $\Delta m'$ is typically quite small (due to the fast decay of $\phi^{(\nu)}(m-\tau/T_\mathrm{s})$, see Section 5 and [10]). Note also that $\mathbf{V}$ (or, equivalently, $\{v_{m,i}[\lambda]\}$) needs to be computed only once before the start of data transmission.

7. SIMULATION RESULTS

We simulated a CP-OFDM system with $K=512$ subcarriers, CP length $N-K=128$, carrier frequency 5 GHz, and bandwidth $1/T_\mathrm{s}=5\,\mathrm{MHz}$. The system used QPSK symbols with Gray labeling, a rate-1/2 convolutional code, and 32×16 row-column interleaving. The interpolation/anti-aliasing filters $f_1(t)=f_2(t)$ were root-raised-cosine filters with roll-off factor $1/4$. We employed the geometry-based channel simulation tool IlmProp [21] to generate doubly dispersive channels during blocks of $L=32$ OFDM symbols. Between transmitter and receiver, which were spaced about 1500 m apart, 7 clusters of 10 specular scatterers each were randomly distributed. Each cluster and the receiver had a random velocity vector with a uniformly distributed direction and a velocity that was uniformly distributed up to 50 m/s for the scatterers, and exactly 50 m/s



(a) (b)

Fig. 2. Performance of the proposed GSCS-based channel estimators and of previously proposed CS-based channel estimators: (a) MSE versus SNR, (b) BER versus SNR.

for the receiver. The bandlimited version of the noise $z(t)$ in (1) was complex bandlimited white Gaussian with variance σ_z^2 adjusted to achieve a prescribed receive signal-to-noise ratio (SNR). The SNR is defined as the mean receive signal power averaged over one block of length LN, divided by σ_z^2. The pilot set $\mathcal{P}$ was randomly chosen from a subsampled grid $\mathcal{G}$ with spacings $\Delta L = 1$ and $\Delta K = 4$. We used $|\mathcal{P}| = 1024$ pilots, corresponding to 6.25% of all symbols.

For compressive channel estimation, we compared the GSCS recovery algorithms G-BPDN and G-OMP with their CS counterparts BPDN and OMP. Both the 2-D DFT basis and an optimized basis were used. The latter was group sparsity optimized—designed according to Section 6—in the GSCS case (G-BPDN and G-OMP), and sparsity optimized—designed according to [10]—in the CS case (BPDN and OMP), in both cases using a pdf $p(\tau_1, \nu_1)$ that was uniform in a rectangular delay-Doppler region determined by the maximum delay and Doppler. The size of the blocks $\mathcal{B}_b$ in Sections 5 and 6 was optimized experimentally and found to be $\Delta m' = 2$ and $\Delta i' = 4$ in both cases.

In Fig. 2, we show the normalized mean-square error (MSE) and the bit error rate (BER) obtained with the various compressive channel estimators as a function of the SNR. It is seen that, for a comparable basis, GSCS-based channel estimation significantly outperforms CS-based estimation. Moreover, for both GSCS methods (G-BPDN and G-OMP), the group sparsity optimized basis yields a large additional performance gain relative to the DFT basis. Finally, each GSCS-based method outperforms its CS-based counterpart even if the former uses the DFT basis and the latter uses the sparsity (not group sparsity) optimized basis. It is also noteworthy that the best GSCS recovery algorithm (G-OMP) has a substantially lower complexity than its CS counterpart (OMP).

8. CONCLUSION

We have proposed compressive estimators of doubly dispersive channels within multicarrier communication systems (including classical OFDM systems as a special case). These channel estimators employ *group sparse* recovery methods in order to take advantage of the group sparse structure of wireless channels. Simulations using a geometry-based channel simulator demonstrated that the proposed methods consistently outperform existing compressive estimators. Moreover, one of the group sparse recovery methods (G-OMP) has a lower complexity than its classical counterpart (OMP). We also proposed a basis optimization method that enhances group sparsity and was seen to yield large additional performance gains. The optimized basis can be precomputed before the start of data transmission.

9. ACKNOWLEDGMENT

The authors would like to thank Dr. Nicolai Czink and Prof. Holger Rauhut for fruitful discussions.

10. REFERENCES

[1] E. van den Berg, M. Schmidt, M. P. Friedlander, and K. Murphy, “Group sparsity via linear-time projection,” Technical Report TR-2008-09, University of British Columbia, 2008.

[2] Y. C. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[3] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, “Block-sparse signals: Uncertainty relations and efficient recovery,” to appear in *IEEE Trans. Sig. Process.*, 2010.

[4] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, April 2010.

[5] Z. Ben-Haim, Y. C. Eldar, and M. Elad, “Coherence-based performance guarantees for estimating a sparse vector under random noise,” submitted to *IEEE Trans. Sig. Process.*, 2009.

[6] J. A. Tropp, “Greed is good: Algorithmic results for sparse approximation,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

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

[8] Ming Jiang and L. Hanzo, “Multiuser MIMO-OFDM for next-generation wireless systems,” *Proc. IEEE*, vol. 95, no. 7, pp. 1430–1469, July 2007.

[9] 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, March/Apr. 2008, pp. 2885–2888.

[10] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, “Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing,” *IEEE J. Sel. Top. Sig. Process.*, vol. 4, no. 2, pp. 255–271, April 2010.

[11] W. U. Bajwa, A. M. Sayeed, and R. Nowak, “Learning sparse doubly-selective channels,” in *Proc. 46th Annu. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Sept. 2008, pp. 575–582.

[12] W. U. Bajwa, A. M. Sayeed, and R. Nowak, “Compressed sensing of wireless channels in time, frequency, and space,” in *Proc. 42nd Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, Oct. 2008, pp. 2048–2052.

[13] M. Sharp and A. Scaglione, “Application of sparse signal recovery to pilot-assisted channel estimation,” in *Proc. IEEE ICASSP-2008*, Las Vegas, NV, April 2008, pp. 3469–3472.

[14] Jun Zhang and A. Papandreou-Suppappola, “Compressive sensing and waveform design for the identification of linear time-varying systems,” in *Proc. IEEE ICASSP-2008*, Las Vegas, NV, April 2008, pp. 3865–3868.

[15] D. Eiwen, G. Tauböck, F. Hlawatsch, H. Rauhut, and N. Czink, “Multichannel-compressive estimation of doubly selective channels in MIMO-OFDM systems: Exploiting and enhancing joint sparsity,” in *Proc. IEEE ICASSP-2010*, Dallas, TX, March 2010, pp. 3082–3085.

[16] V. Raghavan, G. Hariharan, and A. M. Sayeed, “Capacity of sparse multipath channels in the ultra-wideband regime,” *IEEE J. Sel. Top. Sig. Process.*, vol. 1, no. 3, pp. 357–371, Oct. 2007.

[17] W. Kozek and A. F. Molisch, “Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels,” *IEEE J. Sel. Areas Comm.*, vol. 16, no. 8, pp. 1579–1589, Oct. 1998.

[18] 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, no. 5, pp. 1921–1931, May 2007.

[19] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[20] M. Rudelson and R. Vershynin, “Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements,” in *Proc. 40th Annu. Conf. Inform. Sci. Syst. (CISS'06)*, Princeton, NJ, March 2006, pp. 207–212.

[21] `http://tu-ilmenau.de/ilmprop`