in Proc. IEEE ICASSP 2010
Dallas, TX, March 2010, pp. 3082-3085

MULTICHANNEL-COMPRESSIVE ESTIMATION OF DOUBLY SELECTIVE CHANNELS IN MIMO-OFDM SYSTEMS: EXPLOITING AND ENHANCING JOINT SPARSITY

Daniel Eiwen[a], *Georg Tauböck*[b], *Franz Hlawatsch*[b], *Holger Rauhut*[c], *and Nicolai Czink*[d]

[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
[c]Hausdorff Center for Mathematics and Institute for Numerical Simulation, University of Bonn; rauhut@hcm.uni-bonn.de
[d]FTW Telecommunications Research Center Vienna; czink@ftw.at

ABSTRACT

We propose a compressive estimator of doubly selective channels within pulse-shaping multicarrier MIMO systems (including MIMO-OFDM as a special case). The use of *multichannel compressed sensing* exploits the *joint* sparsity of the MIMO channel for improved performance. We also propose a multichannel basis optimization for enhancing joint sparsity. Simulation results demonstrate significant advantages over channel-by-channel compressive estimation.

Index Terms— MIMO-OFDM, multicarrier modulation, channel estimation, multichannel compressed sensing, joint sparsity

1. INTRODUCTION

The methodology of *multichannel compressed sensing* (MCS) or *distributed compressed sensing* [1, 2] allows the efficient *simultaneous* reconstruction of several *jointly* sparse signals. Here, "jointly sparse" means that all signals share the same small (effective) support. In this paper, we apply MCS to the estimation of doubly selective multiple-input multiple-output (MIMO) channels [3]. We consider pulse-shaping multicarrier MIMO (MC-MIMO) systems, which include MIMO-OFDM systems as a special case. MIMO-OFDM is part of several wireless standards [4].

Compressive channel estimation (e.g., [5–9]) exploits the sparsity resulting from the fact that wireless channels tend to be dominated by a relatively small number of clusters of significant paths [10]. However, MIMO channels also exhibit a *joint* sparsity across the component channels, because the antenna spacings are usually much smaller than the path lengths. Hence, we here propose the use of MCS to exploit this joint sparsity structure for a further performance improvement. We note that improved performance of MCS over conventional compressed sensing has been demonstrated in [11]. Our method differs from the compressive MIMO channel estimator proposed in [8] in that [8] uses conventional compressed sensing to exploit sparsity in the angular domain, whereas we use MCS to exploit joint sparsity. Also, [8] only considers frequency-selective MIMO channels, whereas we consider the doubly selective case.

A further contribution of this paper is a multichannel basis optimization technique that mitigates leakage effects. Such effects were observed in [5, 6, 9] to impair the effective delay-Doppler sparsity of doubly selective channels. The proposed technique extends the basis optimization techniques of [6, 9] to the MIMO case, in a way such that the *joint* sparsity is optimized.

This paper is organized as follows. The MC-MIMO system model is described in Section 2. In Section 3, we analyze the joint sparsity of the MIMO channel's delay-Doppler representation. The multichannel-compressive MIMO channel estimator and multichannel basis optimization technique are presented in Sections 4 and 5, respectively. Finally, performance gains achieved by the proposed multichannel techniques are demonstrated experimentally in Section 6.

This work was supported by WWTF grant MA 07-004 (SPORTS) and by FWF grants S10602-N13, S10603-N13, and S10607-N13 within NFN SISE. HR acknowledges support by the Hausdorff Center for Mathematics, University of Bonn. FTW is supported by the Austrian Government and the City of Vienna within the competence center program COMET.

2. MC-MIMO SYSTEM MODEL

We consider a pulse-shaping MC-MIMO system because of its advantages over cyclic-prefix (CP) MIMO-OFDM (e.g., [12]); however, CP MIMO-OFDM is included as a special case. The baseband domain is considered throughout. Let K, $N \geq K$, L, N_{T}, and N_{R} denote the number of subcarriers, symbol duration, number of symbol periods, and numbers of transmit and receive antennas, respectively. The modulator generates the discrete-time transmit signal vector

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

where $\mathbf{a}_{l,k} \triangleq \big(a_{l,k}^{(1)} \cdots a_{l,k}^{(N_{\mathrm{T}})}\big)^T$ denotes the data symbol vectors and $g_{l,k}[n] \triangleq g[n-lN]\, e^{j2\pi \frac{k}{K}(n-lN)}$ denotes time-frequency shifts of a transmit pulse $g[n]$. Subsequently, $\mathbf{s}[n]$ is converted into the continuous-time transmit signal vector $\mathbf{s}(t) = \sum_{n=-\infty}^{\infty} \mathbf{s}[n] f_1(t - nT_{\mathrm{s}})$, where $f_1(t)$ is an interpolation filter and T_{s} is the sampling period.

The channel connecting transmit antenna $s \in \{1, \ldots, N_{\mathrm{T}}\}$ and receive antenna $r \in \{1, \ldots, N_{\mathrm{R}}\}$ is assumed doubly selective with time-varying impulse response $h^{(r,s)}(t, \tau)$. The MIMO channel output is thus given by $\mathbf{r}(t) = \int_{-\infty}^{\infty} \mathbf{H}(t,\tau)\mathbf{s}(t-\tau)\, d\tau + \mathbf{z}(t)$, where $\mathbf{H}(t,\tau)$ is the $N_{\mathrm{R}} \times N_{\mathrm{T}}$ matrix with entries $h^{(r,s)}(t,\tau)$ and $\mathbf{z}(t)$ is a noise vector. At the receiver, $\mathbf{r}(t)$ is converted into the discrete-time signal vector $\mathbf{r}[n] = \int_{-\infty}^{\infty} \mathbf{r}(t)\, f_2(nT_{\mathrm{s}} - t)\, dt$, where $f_2(t)$ is an anti-aliasing filter. Subsequently, the demodulator computes

$$\mathbf{x}_{l,k} = \sum_{n=-\infty}^{\infty} \mathbf{r}[n]\, \gamma_{l,k}^{*}[n] \,, \tag{2}$$

for $l = 0, \ldots, L-1$ and $k = 0 \ldots, K-1$, where $\gamma_{l,k}[n] \triangleq \gamma[n-lN]\, e^{j2\pi \frac{k}{K}(n-lN)}$ with some receive pulse $\gamma[n]$. We note that CP MIMO-OFDM [4] is a special case obtained for rectangular pulses $g[n]$ and $\gamma[n]$ that are 1 on $[0, N-1]$ and on $[N-K, N-1]$, respectively and 0 otherwise ($N-K \geq 0$ is the CP length).

By combining some of the above equations, we obtain

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

where $\mathbf{H}[n,m]$ depends on $\mathbf{H}(t,\tau)$, $f_1(t)$, and $f_2(t)$. Combining (2), (3), and (1) and neglecting intersymbol/intercarrier interference (which is justified if the channel dispersion is not too strong) yields

$$\mathbf{x}_{l,k} = \mathbf{H}_{l,k}\, \mathbf{a}_{l,k} + \mathbf{z}_{l,k} \,, \tag{4}$$

for $l = 0, \ldots, L-1$ and $k = 0, \ldots, K-1$. The channel coefficient matrices $\mathbf{H}_{l,k}$ can be expressed in terms of $g[n]$, $\mathbf{H}[n,m]$, and $\gamma[n]$.

Hereafter, we assume that $\gamma[n] = 0$ outside $[0, L_\gamma]$. To compute the $\mathbf{x}_{l,k}$, $\mathbf{r}[n]$ must be known for $n = 0, \ldots, N_0-1$, where $N_0 \triangleq (L-1)N + L_\gamma + 1$. For these n, we can write (3) as

$$\mathbf{r}[n] = \sum_{m=-\infty}^{\infty} \sum_{i=0}^{N_0-1} \mathbf{S}_h[m,i]\mathbf{s}[n-m]e^{j2\pi\frac{ni}{N_0}} + \mathbf{z}[n], \qquad (5)$$

with the *discrete-delay-Doppler spreading function matrix* [3]

$$\mathbf{S}_h[m,i] \triangleq \frac{1}{N_0}\sum_{n=0}^{N_0-1}\mathbf{H}[n,m]e^{-j2\pi\frac{in}{N_0}}. \qquad (6)$$

Assuming that $\mathbf{H}[n,m] = \mathbf{0}$ for $m \notin [0, K-1]$ and using (2), (5), and (1) and the approximation $N_0 \approx LN$ (which is exact for CP MIMO-OFDM), the $\mathbf{H}_{l,k}$ can be expressed as (L is assumed even)

$$\mathbf{H}_{l,k} = \sum_{m=0}^{K-1}\sum_{i=-L/2}^{L/2-1}\mathbf{F}_{m,i}\, e^{-j2\pi(\frac{km}{K}-\frac{li}{L})}. \qquad (7)$$

Here,

$$\mathbf{F}_{m,i} \triangleq \sum_{q=0}^{N-1}\mathbf{S}_h[m,i+qL]\, A^*_{\gamma,g}\Big(m, \frac{i+qL}{N_0}\Big) \qquad (8)$$

with the *cross-ambiguity function* $A_{\gamma,g}(m,\xi) \triangleq \sum_{n=0}^{L_\gamma}\gamma[n]\, g^*[n-m]\, e^{-j2\pi\xi n}$.

3. JOINT SPARSITY IN THE DELAY-DOPPLER DOMAIN

Next, we analyze the joint sparsity of the functions $F^{(r,s)}_{m,i} \triangleq [\mathbf{F}_{m,i}]_{r,s}$ in (8). We recall that a sequence F is called *S-sparse* if at most S of its values are nonzero, i.e., $|\operatorname{supp}\{F\}| \le S$ (here, $\operatorname{supp}\{F\}$ is the set of indices of all nonzero values of F). Furthermore, a collection of sequences $F^{(\theta)}$ is called *jointly S-sparse* if all $F^{(\theta)}$ share a common S-sparse support, i.e., $\big|\bigcup_\theta \operatorname{supp}\{F^{(\theta)}\}\big| \le S$.

Let $\theta \triangleq (r,s)$ index the antenna pairs (or channels), with $\Theta \triangleq \{\theta = (r,s) \mid r = 1,\ldots,N_{\mathrm{R}}, s = 1,\ldots,N_{\mathrm{T}}\}$ the set of all θ. We assume in this section (not, however, for the compressive channel estimator presented in Section 4) that each channel comprises P propagation paths corresponding to the same set of P specular scatterers with channel-dependent delays $\tau_p^{(\theta)}$ and Doppler frequency shifts $\nu_p^{(\theta)}$ for $p = 1,\ldots,P$. For this channel model, the impulse response of channel θ is given by

$$h^{(\theta)}(t,\tau) = \sum_{p=1}^{P}\alpha_p^{(\theta)}\,\delta\big(\tau-\tau_p^{(\theta)}\big)\, e^{j2\pi\nu_p^{(\theta)}t}, \quad \theta\in\Theta,$$

with complex path gains $\alpha_p^{(\theta)}$. It can then be shown [9] that the delay-Doppler spreading function (see (6)) results as

$$S_h^{(\theta)}[m,i] = \sum_{p=1}^{P}\alpha_p^{(\theta)}\, e^{j\pi(\nu_p^{(\theta)}T_{\mathrm{s}}-\frac{i}{N_0})(N_0-1)}\,\Lambda_p^{(\theta)}[m,i], \qquad (9)$$

with the shifted *leakage kernels* $\Lambda_p^{(\theta)}[m,i] \triangleq \phi_p^{(\theta)}\big(m-\tau_p^{(\theta)}/T_{\mathrm{s}}\big) \cdot \psi\big(i-\nu_p^{(\theta)}T_{\mathrm{s}}N_0\big)$, where $\phi_p^{(\theta)}(x) \triangleq \int_{-\infty}^{\infty} e^{-j2\pi\nu_p^{(\theta)}t} f_1(T_{\mathrm{s}}x-t) \cdot f_2(t)\,dt$ and $\psi(x) \triangleq \sin(\pi x)/[N_0\sin(\pi x/N_0)]$. Each $\Lambda_p^{(\theta)}[m,i]$ is effectively supported in a rectangular region of some delay length $\Delta m \in \mathbb{N}$ and Doppler length $\Delta i \in \mathbb{N}$, centered about the delay-Doppler point $\eta_p^{(\theta)} \triangleq (\tau_p^{(\theta)}/T_{\mathrm{s}}, \nu_p^{(\theta)}T_{\mathrm{s}}N_0)$. Therefore, each $\Lambda_p^{(\theta)}[m,i]$ can be considered approximately N_Λ-sparse, with $N_\Lambda \triangleq \Delta m\,\Delta i$, where Δm and Δi can be chosen such that a desired approximation quality can be achieved. It then follows from (9) and (8) that each $F^{(\theta)}_{m,i} = [\mathbf{F}_{m,i}]_\theta$ is approximately PN_Λ-sparse.

We will now investigate the *joint* sparsity of two different channels. Let d_{T} and d_{R} be the maximum distances between any two transmit antennas and any two receive antennas, respectively. Furthermore, let $\mathbf{w}^{(s)}_{\mathrm{T},p}$ be the vector from transmit antenna s to a given scatterer p, $\mathbf{w}^{(r)}_{\mathrm{R},p}$ the vector from that scatterer to receive antenna r, and $w^{(s)}_{\mathrm{T},p} \triangleq \|\mathbf{w}^{(s)}_{\mathrm{T},p}\|_2$, $w^{(r)}_{\mathrm{R},p} \triangleq \|\mathbf{w}^{(r)}_{\mathrm{R},p}\|_2$. Then, the time delay for scatterer p is given by $\tau_p^{(\theta)} = (w^{(s)}_{\mathrm{T},p} + w^{(r)}_{\mathrm{R},p})/c$, where c is the speed of light. Let $\Delta\tau_p^{(\theta_1,\theta_2)} \triangleq \big|\tau_p^{(\theta_1)} - \tau_p^{(\theta_2)}\big|$ denote the difference in time delay between two different channels θ_1 and θ_2. Using some simple geometric arguments, one can show the bound $\Delta\tau_p^{(\theta_1,\theta_2)} \le \tau_{\mathrm{B}}$, with

$$\tau_{\mathrm{B}} = \frac{d_{\mathrm{T}} + d_{\mathrm{R}}}{c}.$$

Furthermore, let f_0 be the carrier frequency, $\mathbf{v}_{\mathrm{T},p}$ the velocity vector of scatterer p relative to the transmitter, $\mathbf{v}_{\mathrm{R},p}$ the velocity vector of the receiver relative to scatterer p, and $v_{\mathrm{T},p} \triangleq \|\mathbf{v}_{\mathrm{T},p}\|_2$, $v_{\mathrm{R},p} \triangleq \|\mathbf{v}_{\mathrm{R},p}\|_2$. Then, the Doppler shift for scatterer p is $\nu_p^{(\theta)} = \frac{f_0}{c}\big[\mathbf{v}^T_{\mathrm{T},p}\mathbf{w}^{(s)}_{\mathrm{T},p}/w^{(s)}_{\mathrm{T},p} + \mathbf{v}^T_{\mathrm{R},p}\mathbf{w}^{(r)}_{\mathrm{R},p}/w^{(r)}_{\mathrm{R},p}\big]$. Using the Cauchy-Schwarz inequality and some geometric arguments, the Doppler shift difference between two different channels θ_1 and θ_2, $\Delta\nu_p^{(\theta_1,\theta_2)} \triangleq \big|\nu_p^{(\theta_1)} - \nu_p^{(\theta_2)}\big|$, can be bounded as $\Delta\nu_p^{(\theta_1,\theta_2)} \le \nu_{\mathrm{B},p}^{(\theta_1,\theta_2)}$, with

$$\nu_{\mathrm{B},p}^{(\theta_1,\theta_2)} = \frac{f_0}{c}\left[\frac{v_{\mathrm{T},p}\,d_{\mathrm{T}}}{w^{(s_1,s_2)}_{\mathrm{T},p,\min}} + \frac{v_{\mathrm{R},p}\,d_{\mathrm{R}}}{w^{(r_1,r_2)}_{\mathrm{R},p,\min}}\right],$$

where $w^{(s_1,s_2)}_{\mathrm{T},p,\min} \triangleq \min\{w^{(s_1)}_{\mathrm{T},p}, w^{(s_2)}_{\mathrm{T},p}\}$ and $w^{(r_1,r_2)}_{\mathrm{R},p,\min} \triangleq \min\{w^{(r_1)}_{\mathrm{R},p}, w^{(r_2)}_{\mathrm{R},p}\}$.

From these bounds on $\Delta\tau_p^{(\theta_1,\theta_2)}$ and $\Delta\nu_p^{(\theta_1,\theta_2)}$, it follows that the center points $\eta_p^{(\theta_1)}$ and $\eta_p^{(\theta_2)}$ of the shifted leakage kernel supports differ at most by $\lceil\tau_{\mathrm{B}}/T_{\mathrm{s}}\rceil$ in the m-direction and by $\lceil\nu_{\mathrm{B},p}^{(\theta_1,\theta_2)}T_{\mathrm{s}}N_0\rceil$ in the i-direction. Since this is true for any pair of channels (θ_1,θ_2), the set of all shifted leakage kernels $\Lambda_p^{(\theta)}[m,i]$, $\theta\in\Theta$ is (approximately) *jointly* $N_{\Lambda,p}$-sparse, where $N_{\Lambda,p}$ is bounded as

$$N_{\Lambda,p} \le (\Delta m + \Delta m')(\Delta i + \Delta i'_p), \qquad (10)$$

with $\Delta m' \triangleq \lceil\tau_{\mathrm{B}}/T_{\mathrm{s}}\rceil$ and $\Delta i'_p \triangleq \max_{\theta_1\neq\theta_2}\{\lceil\nu_{\mathrm{B},p}^{(\theta_1,\theta_2)}T_{\mathrm{s}}N_0\rceil\}$. It then follows from (9) that the spreading functions $S_h^{(\theta)}[m,i]$, $\theta\in\Theta$ are jointly S-sparse with $S = \sum_{p=1}^{P}N_{\Lambda,p}$. Finally, the same is true for the functions $F^{(\theta)}_{m,i} = [\mathbf{F}_{m,i}]_\theta$ in (8). If the antenna spacings (characterized by d_{T} and d_{R}) are much smaller than the path lengths (characterized by the $w^{(s)}_{\mathrm{T},p}$ and $w^{(r)}_{\mathrm{R},p}$) and if the velocities (characterized by the $v_{\mathrm{T},p}$ and $v_{\mathrm{R},p}$) are not too large, $\Delta m'$ and $\Delta i'_p$ will be small, and thus the *joint* sparsity parameter S will not be much larger than the *individual* sparsity parameters of the functions $F^{(r,s)}_{m,i}$.

4. THE MULTICHANNEL-COMPRESSIVE ESTIMATOR

For practical (underspread [3]) channels as well as practical transmit and receive pulses, the functions $F^{(\theta)}_{m,i}$ in (8) are effectively supported in a small rectangular region about the origin. (Within this region, these functions share a common, sparse support as shown in Section 3.) Thus, we assume that the support of $\mathbf{F}_{m,i}$ is contained in $[0, D-1] \times [-J/2, J/2-1]$, where $D \le K$, $J \le L$ is assumed even, and D and J are chosen such that $\Delta K \triangleq K/D$ and $\Delta L \triangleq L/J$ are integers. Because of (7), the function $\mathbf{H}_{l,k}$ is then uniquely specified by its values on the subsampled grid $\mathcal{G} \triangleq \{(l,k) = (\lambda\,\Delta L, \kappa\,\Delta K)\,|\,\lambda = 0,\ldots,J-1, \kappa = 0,\ldots,D-1\}$, and there is

$$\mathbf{H}_{\lambda\,\Delta L,\kappa\,\Delta K} = \sum_{m=0}^{D-1}\sum_{i=-J/2}^{J/2-1} \mathbf{F}_{m,i}\, e^{-j2\pi(\frac{\kappa m}{D}-\frac{\lambda i}{J})}. \qquad (11)$$

To mitigate the leakage effects affecting $\mathbf{F}_{m,i}$ according to Section 3 and thereby enhance the joint sparsity, we generalize (11) to an orthonormal 2-D basis expansion

$$\mathbf{H}_{\lambda\,\Delta L,\kappa\,\Delta K} = \sum_{m=0}^{D-1}\sum_{i=-J/2}^{J/2-1} \mathbf{\Gamma}_{m,i}\,\beta_{m,i}[\lambda,\kappa]\,. \qquad (12)$$

We again assume that the coefficient functions $\gamma_{m,i}^{(\theta)} \triangleq [\mathbf{\Gamma}_{m,i}]_\theta$ share a common S-sparse support that is contained in $[0, D-1] \times [-J/2, J/2-1]$ (a suitable construction of the basis $\{\beta_{m,i}[\lambda,\kappa]\}$ will be presented in Section 5). Evidently, the 2-D DFT (11) is a special case of (12) with $\mathbf{\Gamma}_{m,i} = \sqrt{JD}\,\mathbf{F}_{m,i}$.

Let $\mu \triangleq (l,k)$. We suppose that for each $s \in \{1,\ldots,N_\mathrm{T}\}$, Q identical pilot vectors $\mathbf{a}_{\mu_q^{(s)}} = \mathbf{p}^{(s)} = \big(p_1^{(s)} \cdots p_{N_\mathrm{T}}^{(s)}\big)^T$, $q = 1,\ldots,Q$ are transmitted at Q time-frequency positions $\mu_q^{(s)} \in \mathcal{P}^{(s)}$, where the $\mathcal{P}^{(s)}$ are disjoint subsets of the subsampled grid $\mathcal{G}$. The pilot vectors $\mathbf{p}^{(s)}$ are chosen linearly independent for $s = 1,\ldots,N_\mathrm{T}$. Note that $|\mathcal{P}^{(s)}| = Q$, that there are in total $N_\mathrm{T}Q$ pilot vectors, and that different entries of $\mathbf{a}_{\mu_q^{(s)}} = \mathbf{p}^{(s)}$ correspond to pilots at the same time-frequency position transmitted from different antennas. Writing $\mu_q^{(s)} = (\lambda_q^{(s)}\Delta L, \kappa_q^{(s)}\Delta K)$ and inserting relation (12) restricted to the pilot positions into (4), we obtain

$$x_{\mu_q^{(s)}}^{(r)} = \sum_{m=0}^{D-1}\sum_{i=-J/2}^{J/2-1} G_{m,i}^{(\theta)}\,\beta_{m,i}[\lambda_q^{(s)},\kappa_q^{(s)}] \,+\, z_{\mu_q^{(s)}}^{(r)} \qquad (13)$$

for all $\mu_q^{(s)} \in \mathcal{P}^{(s)}$ and all $s = 1,\ldots,N_\mathrm{T}$, $r = 1,\ldots,N_\mathrm{R}$. Here, $x_\mu^{(r)} \triangleq [\mathbf{x}_\mu]_r$, $z_\mu^{(r)} \triangleq [\mathbf{z}_\mu]_r$, and $G_{m,i}^{(\theta)} \triangleq \boldsymbol{\gamma}_{m,i}^{(r)T}\mathbf{p}^{(s)}$, with $\boldsymbol{\gamma}_{m,i}^{(r)T}$ denoting the rth row of $\mathbf{\Gamma}_{m,i}$. Let $\mathbf{B}$ be the unitary $JD\times JD$ matrix whose $((i+J/2)D+m+1)$th column is given by $\mathrm{vec}_{\lambda,\kappa}\{\beta_{m,i}[\lambda,\kappa]\}$, which denotes the columnwise stacking with respect to λ,κ of the $J\times D$ "matrix" $\beta_{m,i}[\lambda,\kappa]$ into a JD-dimensional vector. Furthermore, let $\mathbf{\Phi}^{(s)} \triangleq \mathbf{B}^{(s)}\mathbf{D}^{(s)}$, where $\mathbf{B}^{(s)}$ denotes the $Q\times JD$ submatrix of $\mathbf{B}$ composed of the Q rows corresponding to the pilot positions $\mathcal{P}^{(s)}$, and $\mathbf{D}^{(s)}$ is a diagonal matrix such that the columns of $\mathbf{B}^{(s)}$ have unit ℓ_2-norm. We also define the JD-dimensional vectors $\mathbf{u}^{(\theta)} \triangleq (\mathbf{D}^{(s)})^{-1}\mathrm{vec}_{m,i}\{G_{m,i}^{(\theta)}\}$. Finally, let $\mathbf{x}^{(\theta)} \triangleq \big(x_1^{(\theta)}\cdots x_Q^{(\theta)}\big)^T$, where $x_q^{(\theta)} \triangleq x_{\mu_q^{(s)}}^{(r)}$ with $\mu_q^{(s)} \in \mathcal{P}^{(s)}$. We can then rewrite (13) as

$$\mathbf{x}^{(\theta)} = \mathbf{\Phi}^{(s)}\mathbf{u}^{(\theta)} + \mathbf{z}^{(\theta)}, \qquad \theta \in \Theta\,. \qquad (14)$$

Since the coefficients $\gamma_{m,i}^{(\theta)}$ were assumed jointly S-sparse, the functions $G_{m,i}^{(\theta)} = \boldsymbol{\gamma}_{m,i}^{(r)T}\mathbf{p}^{(s)}$ are jointly S-sparse, too. Therefore, (14) is a *simultaneous sparse reconstruction problem*: we have to recover the $N_\mathrm{T}N_\mathrm{R}$ vectors $\mathbf{u}^{(\theta)} \in \mathbb{C}^{JD}$, $\theta \in \Theta$ from the $N_\mathrm{T}N_\mathrm{R}$ vectors $\mathbf{x}^{(\theta)} \in \mathbb{C}^Q$, $\theta \in \Theta$, based on the known "measurement matrices" $\mathbf{\Phi}^{(s)} \in \mathbb{C}^{Q\times JD}$. Typically, $Q \ll JD$; however, the $\mathbf{u}^{(\theta)}$ are (approximately) jointly S-sparse (e.g., $S = \sum_{p=1}^{P} N_{\Lambda,p}$ for the DFT basis in (11), with $N_{\Lambda,p}$ bounded according to (10)). Hence, we can use an MCS recovery algorithm such as DCS-SOMP [2] to obtain estimates of the $\mathbf{u}^{(\theta)}$ and, after rescaling with $\mathbf{D}^{(s)}$, of the $G_{m,i}^{(\theta)}$. Writing $\mathbf{g}_{m,i}^{(r)} \triangleq \big(G_{m,i}^{(r,1)}\cdots\, G_{m,i}^{(r,N_\mathrm{T})}\big)^T$ and $\mathbf{P} \triangleq \big(\mathbf{p}^{(1)}\cdots\mathbf{p}^{(N_\mathrm{T})}\big)$, the relations $G_{m,i}^{(\theta)} = \boldsymbol{\gamma}_{m,i}^{(r)T}\mathbf{p}^{(s)}$ yield $\boldsymbol{\gamma}_{m,i}^{(r)} = \mathbf{P}^{-T}\mathbf{g}_{m,i}^{(r)}$ (note that $\mathbf{P}$ is nonsingular because the $\mathbf{p}^{(s)}$ are linearly independent). From the estimated $\boldsymbol{\gamma}_{m,i}^{(r)}$, we obtain estimates of the subsampled channel coefficients $\mathbf{H}_{\lambda\,\Delta L,\kappa\,\Delta K}$ via (12). Finally, inverting (11) and using (7) yields estimates of all channel coefficients $\mathbf{H}_{l,k}$.

For consistency with the standard construction of measurement matrices [13], the pilot positions $\mathcal{P}^{(s)}$ are randomly chosen from the subsampled grid $\mathcal{G}$. Accordingly, the measurement matrices $\mathbf{\Phi}^{(s)}$ are constructed by randomly drawing $|\mathcal{P}^{(s)}| = Q$ rows from the unitary matrix $\mathbf{B}$ and normalizing the columns to unit ℓ_2-norm.

Alternatively, since $\mathbf{\Phi}^{(s)}$ in (14) does not depend on r, we can also apply the SOMP algorithm [1] for $s = 1,\ldots,N_\mathrm{T}$. This is a special case of DCS-SOMP for identical measurement matrices. Applying SOMP N_T times is less complex than using DCS-SOMP, or using orthogonal matching pursuit (OMP) [14] $N_\mathrm{T}N_\mathrm{R}$ times.

5. MULTICHANNEL BASIS OPTIMIZATION

We now present an algorithm for designing the basis $\{\beta_{m,i}[\lambda,\kappa]\}$ in (12) such that the *joint* sparsity of the channel coefficients $\gamma_{m,i}^{(\theta)} = [\mathbf{\Gamma}_{m,i}]_\theta$ is optimized. This algorithm is an extension of our basis optimization algorithms in [6, 9] to the MIMO (multichannel) case.

Consider the single-scatterer channels $h^{(\theta)}(t,\tau) \triangleq \delta\big(\tau-\tau_1^{(\theta)}\big) \cdot e^{j2\pi\nu_1^{(\theta)}t}$, $\theta \in \Theta$, where $\tau_1^{(\theta)}$ and $\nu_1^{(\theta)}$ are random and distributed according to the probability density function (pdf) $p(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1) \triangleq p\big(\tau_1^{(1,1)},\ldots,\tau_1^{(N_\mathrm{R},N_\mathrm{T})},\nu_1^{(1,1)},\ldots,\nu_1^{(N_\mathrm{R},N_\mathrm{T})}\big)$. (A nonstatistical design that does not require knowledge of $p(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)$ is easily obtained by constructing $p(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)$ from uniform distributions.) We set $\beta_{m,i}[\lambda,\kappa] \triangleq \frac{1}{\sqrt{D}}\,b_{m,i}[\lambda]\,e^{-j2\pi\frac{\kappa m}{D}}$ (see [6, 9] for a motivation of this construction), where $\{b_{m,i}[\lambda]\}_{i=-J/2}^{J/2-1}$, $m = 0,\ldots,D-1$ is a family of D orthonormal 1-D bases that will be chosen such that $\mathbf{\Gamma}_{m,i}$ is maximally *jointly* sparse on average. As in [11], the joint sparsity is measured by $\|\mathbf{\Gamma}\|_{2,1} \triangleq \sum_{m=0}^{D-1}\sum_{i=-J/2}^{J/2-1}\|\mathbf{\Gamma}_{m,i}\|_\mathrm{F}$, where $\|\mathbf{\Gamma}_{m,i}\|_\mathrm{F}$ denotes the Frobenius norm of $\mathbf{\Gamma}_{m,i}$. One can show the following expression for the expectation of $\|\mathbf{\Gamma}\|_{2,1}$:

$$\mathrm{E}\big\{\|\mathbf{\Gamma}\|_{2,1}\big\} = \sum_{m=0}^{D-1}\sum_{i=0}^{J-1}\mathrm{E}\big\{\|\mathbf{B}_m\mathbf{C}_m(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)\|_2^{(i+1)}\big\}\,,$$

where $\mathbf{B}_m$ is the $J\times J$ matrix with entries $(\mathbf{B}_m)_{i+1,\lambda+1} \triangleq b_{m,i-J/2}^*[\lambda]$ and $\mathbf{C}_m(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)$ is the $J\times N_\mathrm{T}N_\mathrm{R}$ matrix with columns $\sqrt{D}\,\phi_1^{(\theta)}(m-\tau_1^{(\theta)}/T_\mathrm{s})\,\mathbf{c}_m^{(\nu_1^{(\theta)})}$, $\theta \in \Theta$. Here, $\mathbf{c}_m^{(\nu)} \triangleq \big(C_{m,0}^{(\nu)}\cdots C_{m,J-1}^{(\nu)}\big)^T$ with

$$C_{m,\lambda}^{(\nu)} \triangleq \sum_{i=-J/2}^{J/2-1}\sum_{q=0}^{N-1}\psi^{(\nu)}(i+qL)\,A_{\gamma,g}^*\Big(m,\frac{i+qL}{N_0}\Big)\,e^{j2\pi\frac{\lambda i}{J}},$$

where $\psi^{(\nu)}(i) \triangleq e^{j\pi(\nu T_\mathrm{s}-\frac{i}{N_0})(N_0-1)}\,\psi(i-\nu T_\mathrm{s}N_0)$. Finally, $\|\mathbf{A}\|_2^{(i)}$ denotes the ℓ_2 norm of the ith row of $\mathbf{A}$. Hence, the joint sparsity can be optimized by solving the D optimization problems $\min_{\mathbf{B}_m\in\mathcal{U}}\sum_{i=0}^{J-1}\mathrm{E}\big\{\|\mathbf{B}_m\mathbf{C}_m(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)\|_2^{(i+1)}\big\}$, $m = 0,\ldots,D-1$, where $\mathcal{U}$ denotes the set of all unitary $J\times J$ matrices. Using a Monte-Carlo approximation, we redefine these optimization problems as

$$\hat{\mathbf{B}}_m \triangleq \underset{\mathbf{B}_m\in\mathcal{U}}{\arg\min}\sum_\rho\sum_{i=0}^{J-1}\|\mathbf{B}_m\mathbf{C}_m(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)_\rho\|_2^{(i+1)}\,, \qquad (15)$$

for $m = 0,\ldots,D-1$, in which the $(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)_\rho$ denote samples of the random vector $(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)$ drawn from its pdf $p(\boldsymbol{\tau}_1,\boldsymbol{\nu}_1)$.

An approximate solution of (15) can be obtained by an iterative algorithm in which each iteration constitutes a convex minimization for which efficient methods exist (see [6, 9] for details). This algorithm is initialized by the DFT matrix.

6. SIMULATION RESULTS

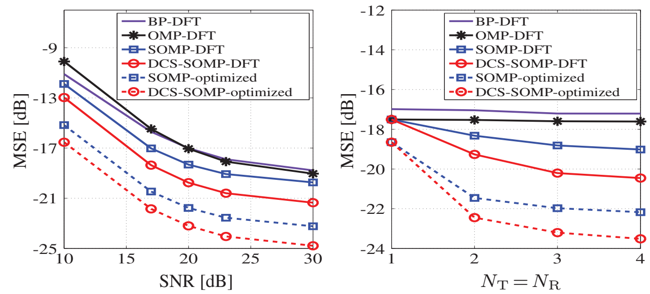
We simulated CP MIMO-OFDM systems with $K = 512$ subcarriers, CP length $N - K = 128$, QPSK symbols, carrier frequency $f_0 = 5\,\text{GHz}$, and bandwidth $B = 1/T_s = 5\,\text{MHz}$. The interpolation/anti-aliasing filters $f_1(t) = f_2(t)$ were root-raised-cosine filters with roll-off factor $1/4$. During blocks of $L = 32$ OFDM symbols, we generated a doubly selective MIMO channel using the simulation tool IlmProp [15]. 7 clusters of 10 specular scatterers each were randomly distributed between transmitter and receiver, which were separated by about 1500 m. Scatterers and receiver had random velocity vectors with uniformly distributed directions and velocities of up to 50 m/s for the scatterers and 50 m/s for the receiver. We used up to four transmit/receive antennas spaced $c/(2f_0)$ apart. The noise $\mathbf{z}[n]$ in (3) was complex white Gaussian with component variance σ_z^2 adjusted to achieve a prescribed receive signal-to-noise ratio (SNR). The SNR is defined as the mean received signal power averaged over one receive block of length N_0 and all receive antennas, divided by σ_z^2. The pilot position sets $\mathcal{P}^{(s)}$ were randomly chosen from a subsampled grid $\mathcal{G}$ with spacings $\Delta L = 1$ and $\Delta K = 4$. We used $|\mathcal{P}^{(s)}| = 1024$, which implies that $6.25 \cdot N_T\,\%$ of all symbols are pilots. The pilot vectors $\mathbf{p}^{(s)}$, $s = 1, \ldots, N_T$ were randomly chosen, orthogonalized, and normalized such that the power of each $\mathbf{p}^{(s)}$ equals the power of N_T data symbols.

For a MIMO system with $N_T = N_R = 3$ antennas, Fig. 1(a) shows the normalized mean-square error (MSE) versus the SNR for the proposed MCS-based estimator and for conventional channel-by-channel compressed-sensing (CS) based estimation [5]. In Fig. 1(b), the normalized MSE is plotted versus the number of antennas, $N_T = N_R \in \{1, \ldots, 4\}$, at an SNR of 20 dB. The recovery algorithms were DCS-SOMP [2] and SOMP [1] for MCS-based estimation and OMP [14] and BP (denoising) [16] for CS-based estimation. We used the DFT basis (see (11)) and, for MCS-based channel estimation, also an optimized basis designed according to Section 5. The pdf for basis optimization was chosen as $p(\boldsymbol{\tau}_1, \boldsymbol{\nu}_1) = p\big(\tau_1^{(1,1)}, \nu_1^{(1,1)}\big)\, p\big(\tau_1^{(1,2)}, \nu_1^{(1,2)} | \tau_1^{(1,1)}, \nu_1^{(1,1)}\big) \cdots p\big(\tau_1^{(N_R,N_T)}, \nu_1^{(N_R,N_T)} | \tau_1^{(1,1)}, \nu_1^{(1,1)}\big)$, where each factor is uniform in a rectangular region. The region for the first factor was determined by the maximum delay and Doppler, and the regions for the remaining factors were determined by $\Delta\tau_p^{(\theta_i,\theta_j)}$ and $\Delta\nu_p^{(\theta_i,\theta_j)}$ (cf. Section 3).

It is seen from Fig. 1 that MCS-based estimation using the DFT basis significantly outperforms CS-based estimation, due to its ability to exploit joint sparsity. Furthermore, use of the optimized basis in MCS-based estimation yields a large additional performance gain, because of the enhanced joint sparsity. DCS-SOMP performs better than SOMP, since it exploits the joint sparsity of a larger number of channels. By the same reason, the performance gains mentioned above increase with the number of antennas (see Fig. 1(b)).

7. CONCLUSION

The application of recently proposed *multichannel compressed sensing* methods to MIMO channel estimation makes it possible to take advantage of the *joint* sparsity of MIMO wireless channels, in addition to the sparsity of the individual component channels. Simulation results demonstrated significant performance gains and (for SOMP recovery) reduced computational complexity of multichannel-compressive MIMO channel estimation relative to channel-by-channel compressive estimation. A further substantial performance gain was achieved by a *multichannel* basis optimization that mitigates leakage effects and enhances joint sparsity. The optimized basis can be precomputed before the start of data transmission.



(a) (b)

Fig. 1. Performance of MCS-based and CS-based MIMO channel estimation: (a) MSE versus SNR, (b) MSE versus $N_T = N_R$.

8. ACKNOWLEDGMENT

The authors would like to thank G. Del Galdo, J. Lotze, and M. Haardt for their assistance regarding the channel simulator IlmProp.

9. REFERENCES

[1] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, pp. 572–588, March 2006.

[2] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," *preprint*, 2005.

[3] G. Matz, "Characterization and analysis of doubly dispersive MIMO channels," in *Proc. 40th Asilomar Conf. Sig., Syst., Comp.*, Pacific Grove, CA, pp. 946–950, Oct. 2006.

[4] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," *Proc. IEEE*, vol. 95, pp. 1430–1469, July 2007.

[5] 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.

[6] G. Tauböck and F. Hlawatsch, "Compressed sensing based estimation of doubly selective channels using a sparsity-optimized basis expansion," in *Proc. EUSIPCO 2008*, Lausanne, Switzerland, Aug. 2008.

[7] 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, pp. 575–582, Sept. 2008.

[8] 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, pp. 2048–2052, Oct. 2008.

[9] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing." to appear in *IEEE J. Sel. Topics Signal Process.*, 2010.

[10] V. Raghavan, G. Hariharan, and A. M. Sayeed, "Capacity of sparse multipath channels in the ultra-wideband regime," *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp. 357–371, Oct. 2007.

[11] Y. C. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. Inf. Theory*, vol. 56, pp. 505–519, Jan. 2010.

[12] 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.

[13] 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, pp. 207–212, March 2006.

[14] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, Oct. 2004.

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

[16] 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.