Study of Pilot Designs for Cyclic-Prefix OFDM on Time-Varying and Sparse Underwater Acoustic Channels

Christian R. Berger*, João Gomes†, and José M. F. Moura*
*Dept. of Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
†Institute for Systems and Robotics, Instituto Superior Técnico, 1049-001 Lisboa, Portugal

Abstract—Estimation of time-varying *and* sparse channels is a topic that has attracted considerable interest lately, especially for underwater acoustic communication. In the context of multicarrier transmission, several important questions remain unanswered: i) should pilots be placed in clusters as for time-varying, non-sparse channels, or randomly dispersed between the data as common in the Compressive Sensing literature; ii) as pilot and data subcarriers cannot be perfectly separated at the receiver due to intercarrier interference (ICI) should channel estimation be based on observations corresponding to pilot subcarriers only, or can the data subcarriers be used to practically extract additional information about the channel; and iii) how does the performance vary with the number of pilot symbols? We use data recorded at the recent CalCom'10 experiment to investigate these questions. We find that a sufficient number of adjacent observations is required to estimate the ICI, which can either be achieved by a pilot design that uses clustered pilots, or by using data subcarriers as additional observations. When using data subcarriers as observations, the effect of the unknown data should be modeled as increased noise. Finally, by varying the number of pilot subcarriers, we can trade off error performance for data rate.

Index Terms—Pilot-aided transmission, sparse channel estimation, basis pursuit, orthogonal matching pursuit, multicarrier transmission

I. INTRODUCTION

Underwater acoustic (UWA) channels are characterized by large delay spread and significant Doppler effects, qualifying UWA channels as doubly (time- and frequency-) spread channels. While the large delay spread motivates multicarrier transmission schemes that can use efficient frequency domain equalization (see, e.g., [1]–[3]), the significant Doppler spread will lead to loss of orthogonality between the subcarriers. Early multicarrier receiver designs ignored that the loss of orthogonality between the subcarriers would lead to intercarrier interference (ICI) [1]–[3], but now it is largely agreed that estimating and equalizing the ICI (or at least some of it) leads to significant performance improvement [4]–[7]. Unfortunately to estimate the ICI, e.g., using a basis expansion model (BEM) [6], [8], the number of unknown parameters explodes.

Although it has been recognized for a long time that UWA channels are sparse [9], so far most work to take advantage of this fact has focused only on reconstructing channels in the delay domain [3], [10]–[12]. It turns out that estimating the channel both in delay and Doppler dimension with sufficient accuracy is challenging even when using known transmitted signals [13], or needs increased pilot overhead [4].

The existing literature on pilot-assisted channel estimation is extensive see, e.g., the overview in [14]. For multicarrier transmission across time-invariant channels, it is well-known that equi-distantly spaced pilots are optimal under various criteria [14]. The challenges in pilot design for multicarrier transmission over time-varying channels are two-fold:

1) Sets of adjacent observations are needed to estimate the ICI coefficients.
2) Keeping pilot and data symbols orthogonal at the receiver is challenging due to the ICI.

For non-sparse channel estimation, equi-distantly spaced blocks of pilot symbols are therefore considered optimum under most BEMs see, e.g., [15], [16] and references therein. On the other hand, no consensus exists on whether pilots and data should be kept orthogonal at all cost; while in some works guard zeros are inserted around the pilots [15], in others the ICI caused by the unknown data symbols is treated as additional noise for purposes of channel estimation [16].

In existing work on sparse channel estimation for multicarrier transmission over time-varying channels various ad-hoc pilot schemes were reported. While in [17], [18] pilot symbols were placed on randomly selected subcarriers, in [4], [7] a pilot design was used that consisted of equi-distantly spaced blocks, where the block size alternated between a single pilot tone and groups of five pilots. In all cases the ICI between the pilot subcarriers was used to estimate the time-varying nature of the channel [4], [7], [17], [18], but the ICI between data and pilot subcarriers was neglected. Without guard zeros, ICI from the unknown data symbols is perceived as (some) additional noise on pilot subcarriers, but ICI originating from the known pilot symbols makes data subcarriers potential observations — although *very* noisy, due to the unknown data symbols.

In this work, we want to compare the random selection of pilot subcarriers as motivated by the compressive sensing literature [19], [20] with the systematic use of pilot blocks in regular intervals as seen in non-sparse channel estimation of time-varying channels. Furthermore, we are interested in how to address the ICI between data and pilot subcarriers. We do not consider the use of guard zeros, as it decreases spectral efficiency, but we focus on whether data symbol carrying subcarriers should be used as observations in channel

estimation, as they also contain ICI originating from the pilot symbols. Finally, we study the performance under varying amounts of pilot overhead. Specifically, we are looking for an optimum tradeoff between using more pilots or a more robust modulation scheme, to achieve the highest spectral efficiency.

We base our study on data recorded during the Calibration and Communications Sea Trial 2010 (CalCom'10), which was conducted about 30 km south of the port of Vilamoura, Portugal, on June 21–25, 2010. For sparse channel estimation we consider both orthogonal matching pursuit (OMP) and basis pursuit (BP) as in [4], [7]. We find the following:

- When limiting the observations to pilot subcarriers, the systematic approach that uses clusters of pilot subcarriers outperforms the approach that places pilot subcarriers randomly.
- When using all observations in channel reconstruction BP suffers from the additionally introduced noise, while OMP can take advantage of the additional observations under an appropriate stopping criteria.
- Using "observation masking", i.e., where observations based on pilot subcarriers are weighted more heavily than observations based on data subcarriers, both OMP and BP perform well, almost irrespectively of the chosen pilot pattern.

The rest of this paper is organized as follows. In Section II we specify the signal model, and in Section III we review sparse channel estimation. Then we first study various pilot designs in Section IV; in Section V we investigate if data subcarriers can be practically used as additional observations in channel estimation; and lastly we show some results on varying pilot overhead in Section VI. We conclude in Section VII.

II. SIGNAL MODEL

A. Cyclic-Prefix OFDM

We consider cyclic-prefix (CP) orthogonal frequency division multiplexing (OFDM); let T denote the OFDM symbol duration and T_{CP} the length of the CP. The total duration of one OFDM block is $T' = T + T_{CP}$ and the subcarrier spacing is $1/T$. The kth subcarrier is at frequency

$$f_k = f_c + k/T, \quad k = -K/2, \ldots, K/2 - 1, \quad (1)$$

where f_c is the carrier frequency and K subcarriers are used, so that the nominal bandwidth is $B = K/T$. Let $s[k]$ denote the information symbol to be transmitted on the kth subcarrier, chosen from a constellation like quadrature phase shift keying (QPSK) or quadrature amplitude modulation (QAM). The non-overlapping sets of active subcarriers $\mathcal{S}_A$ and null subcarriers $\mathcal{S}_N$ satisfy $\mathcal{S}_A \cup \mathcal{S}_N = \{-K/2, \ldots, K/2 - 1\}$; the null subcarriers are used to facilitate Doppler compensation at the receiver [2], [21], and protect the band edges. A transmitted OFDM block in passband is given by

$$\tilde{x}(t) = 2\mathrm{Re}\left\{ \left[\sum_{k \in \mathcal{S}_A} s[k] e^{j2\pi \frac{k}{T} t} \right] e^{j2\pi f_c t} \right\}, \, t \in [-T_{CP}, T]. \quad (2)$$

symbol length	T	491.52 ms
cyclic prefix	T_{cp}	48 ms
block length	T'	539.52 ms
no. subcarriers	K	4096
subcarrier spacing	$1/T$	2.0345 Hz
baseband sampling rate	B	8.33 kHz
carrier frequency	f_c	12 kHz
band protection	N_P	128
no. null subcarriers	$|\mathcal{S}_N|$	$48 + 2 \times 128$
no. active subcarriers	$|\mathcal{S}_A|$	3792
effective bandwidth	B_{eff}	7.8125 kHz

TABLE II
DATA RATES USED IN CALCOM'10 EXPERIMENT.

pilot overhead	QPSK	16-QAM
1/8 pilots	6.28 kbit/s	
1/4 pilots	5.27 kbit/s	10.5 kbit/s
1/2 pilots	3.59 kbit/s	7.18 kbit/s

The parameters used in the CalCom'10 experiment are listed in Table I. While the baseband sampling rate is $B = 8.33$ kHz, factoring in the guard bands of $N_P = 128$ null subcarriers at each band edge, the effectively used bandwidth is only $B_{eff} = 7.8125$ kHz. Including also the 48 null subcarriers spread evenly among the active subcarriers, the total number of null subcarriers is $|\mathcal{S}_N| = 304$, leaving $|\mathcal{S}_A| = 3792$ active subcarriers to carry data and pilot symbols. We consider both QPSK and 16-QAM; OFDM blocks are encoded separately with a convolutional code of rate $1/2$ (constraint length 9), and pilot overheads of $1/8$, $1/4$, or $1/2$, leading to the data rates shown in Table II.

B. Underwater Acoustic Channel Model

The general time-varying multipath channel model commonly considered in UWA communications is,

$$h(\tau; t) = \sum_p A_p(t) \delta\big(\tau - \tau_p(t)\big). \quad (3)$$

As in [4], [7], we assume that for the duration of one OFDM block the path amplitudes are approximately constant $A_p(t) \approx A_p$, and that the delays can be modeled as linearly time-varying,

$$h(\tau; t) = \sum_p A_p \delta\big(\tau - (\tau_p - a_p t)\big). \quad (4)$$

Each arrival is thus characterized by an initial delay, τ_p, and rate-of-change, a_p; alternatively, the Doppler shift $f_{D,p} = a_p f_c$ may be used (which assumes a narrowband signal at f_c). The parameters are assumed constant for the duration of one OFDM block, T', and independent across blocks for purposes of channel estimation.

An example of a channel observed at the CalCom'10 experiment is shown in Fig. 1. The plots are created by correlating

Example of Time-Invariant Channel Estimate
Impulse response magnitude vs. delay [ms]
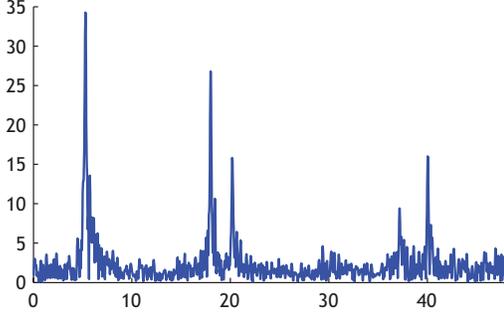
Example of Time-Varying Channel Estimate
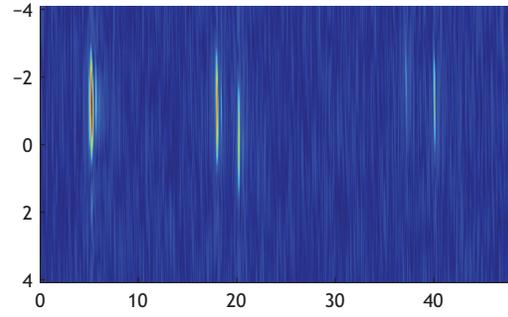Correlation magnitude vs. Doppler [Hz] vs. delay [ms]

Fig. 1. Estimates of a channel response $h(\tau; t)$ recorded at the CalCom'10 experiment; while the time-invariant estimate correlates the received signal only with different delays, the time-varying estimate additionally uses possible (wideband) Doppler shifts.

a received OFDM block with delayed and/or Doppler scaled versions of itself. Clearly the channel has a sparse structure, and the degree of time variation measured in terms of Doppler frequency is significant with respect to the OFDM block length $f_{D,\max} \times T \approx 0.5$.

C. Receiver Processing

According to the channel model in (4), the received signal is

$$\tilde{y}(t) = \tilde{x}(t) \star h(\tau; t) + \tilde{v}(t), \tag{5}$$

$$= \sum_p A_p \tilde{x}\big((1 + a_p)t - \tau_p\big) + \tilde{v}(t) \tag{6}$$

where $\tilde{v}(t)$ is the ambient noise. The signal is downconverted, then (over-) sampled and fed into an FFT block, including narrowband Doppler compensation ϵ (see [21] or [2] and references therein); the mth FFT output can be calculated as,

$$z_m = \frac{1}{T} \int_0^T \tilde{y}(t) e^{-j2\pi(f_c + \epsilon)t} e^{-j2\pi \frac{m}{T}t} \, dt, \tag{7}$$

$$= \sum_{k \in \mathcal{S}_A} H[m, k] s[k] + v_m. \tag{8}$$

OFDM subcarriers lose their orthogonality due to Doppler effects caused by the time-varying channel, leading to inter-carrier interference (ICI), where coefficient $H[m, k]$ specifies the contribution of the kth subcarrier to the mth FFT output.

Collecting the ICI coefficients into a matrix $\mathbf{H}$, and the z_m, $s[k]$, and v_m into vectors $\mathbf{z}$, $\mathbf{s}$, and $\mathbf{v}$, respectively, we can rewrite (8) as

$$\mathbf{z} = \mathbf{Hs} + \mathbf{v}. \tag{9}$$

If the channel were time-invariant, the matrix $\mathbf{H}$ would be diagonal. In time-varying channels with limited Doppler spread, $\mathbf{H}$ is often considered to be block-diagonal, or at least approximately so [15], [16].

D. Characterization of ICI

Based on the channel model in (4), and assuming that the delay spread is contained within the CP, $0 \leq \tau_p \leq T_{\mathrm{CP}}$, the

ICI coefficients can be calculated as,

$$H[m, k] = \sum_p \underbrace{A_p e^{-j2\pi f_c \tau_p}}_{\xi_p} \exp\left(-j2\pi \frac{k}{T} \tau_p\right)$$

$$\times \underbrace{\frac{1}{T} \int_0^T \exp\left(j2\pi \frac{k - m + a_p(k + f_c T) - \epsilon T}{T} t\right) dt}_{\varrho_{m,k}(a_p, \epsilon)}$$

$$= \sum_p \xi_p \, e^{-j2\pi \frac{k}{T} \tau_p} \varrho_{m,k}(a_p, \epsilon). \tag{10}$$

The integral constituting the $\varrho_{m,k}(a_p, \epsilon)$ can be identified as the (inverse) Fourier transform of the rectangular receiver window implicitly used in the CP-OFDM receiver at frequency

$$\omega_{m,k}(a, \epsilon) = \frac{k - m + a(k + f_c T) - \epsilon T}{T}, \tag{11}$$

hence

$$\varrho_{m,k}(a, \epsilon) = \mathrm{sinc}\left(\omega_{m,k}(a, \epsilon) T\right) e^{j\pi \omega_{m,k}(a, \epsilon) T}. \tag{12}$$

Clearly for $|k - m| \gg 1$ (and $a \ll 1$) the coefficients $\varrho_{m,k}(a_p, \epsilon)$ will be very small, justifying the assumption of an approximately block-diagonal matrix.

Defining the diagonal matrix $\mathbf{\Lambda}(\tau)$ with kth element

$$[\mathbf{\Lambda}(\tau)]_{k,k} = e^{-j2\pi \frac{k}{T} \tau}, \tag{13}$$

and the mixing matrix $\mathbf{\Gamma}(a, \epsilon)$ with (m, k)th element

$$[\mathbf{\Gamma}(a, \epsilon)]_{m,k} = \varrho_{m,k}(a, \epsilon), \tag{14}$$

we can use the following vector/matrix notation

$$\mathbf{H} = \sum_p \xi_p \mathbf{\Gamma}(a_p, \epsilon) \mathbf{\Lambda}(\tau_p), \tag{15}$$

which is equivalent to the zero-padded (ZP) OFDM formulation in [4], [7], but the matrices depending on delays, τ_p, and Doppler rate, a_p, have swapped order.

E. Equalization & Decoding

After the channel has been estimated as $\hat{\mathbf{H}}$ (which is discussed in detail next), minimum mean-square error (MMSE) equalization is applied,

$$\hat{\mathbf{s}} = \left(\hat{\mathbf{H}}^H \hat{\mathbf{H}} + N_0 \mathbf{I}_K\right)^{-1} \hat{\mathbf{H}}^H \mathbf{z}, \tag{16}$$

where we assumed that the noise is white with power N_0. As in UWA communications receiver arrays are common, this is also the stage where symbol estimates are combined, e.g., using maximum-ratio combining (MRC). For reasons of receiver complexity, the channel estimate $\hat{\mathbf{H}}$ is approximated as a strictly banded matrix as in [4], [7], with $D = 5$ off-diagonals on each side. The MMSE estimate of $\mathbf{s}$ is used to generate soft-input for error correction as described in [22].

III. SPARSE CHANNEL ESTIMATION

A. Sparse Basis

As in [4], [7], to apply sparse reconstruction algorithms we discretize the delay and Doppler domain as follows:

$$\tau_p \in \left\{ \frac{T}{\alpha K}, \frac{2T}{\alpha K}, \ldots, T_{\text{CP}} \right\} \tag{17}$$

$$a_q \in \{-a_{\max}, -a_{\max} + \Delta a, \ldots, a_{\max}\} \tag{18}$$

Where the number of delay steps are $N_\tau = T_{\text{CP}}/(T/\alpha K)$ and Doppler steps $N_a = 2a_{\max}/(\Delta a) + 1$; we will denote the path weight that corresponds to a path with delay τ_p and rate-of-change a_q as $\xi_{p,q}$. With this we can write the channel model as

$$\mathbf{H} = \sum_{p=1}^{N_\tau} \sum_{q=1}^{N_a} \xi_{p,q} \boldsymbol{\Gamma}(a_q, \epsilon) \boldsymbol{\Lambda}(\tau_p), \tag{19}$$

where most $\xi_{p,q}$ will be close to zero. The channel model in (19) is not a basis expansion model (BEM) itself, but when assuming a narrowband signal, it reduces to the complex exponential basis expansion model (CE-BEM).

B. Selecting Observations

In all of [4], [7], [17], [18], only FFT outputs are used in channel estimation that correspond to pilot symbols. Therefore, we define the set of pilot subcarriers as $\mathcal{S}_P$ and the remaining active subcarriers as data subcarriers $\mathcal{S}_D$, accordingly $\mathcal{S}_D = \mathcal{S}_A \setminus \mathcal{S}_P$. We define the vector of pilot tones $\mathbf{p}$,

$$[\mathbf{p}]_k = \begin{cases} s[k] & k \in \mathcal{S}_P \\ 0 & k \in \mathcal{S}_N \cup \mathcal{S}_D, \end{cases} \tag{20}$$

and the vector of data symbols $\mathbf{d}$,

$$[\mathbf{d}]_k = \begin{cases} s[k] & k \in \mathcal{S}_D \\ 0 & k \in \mathcal{S}_N \cup \mathcal{S}_P. \end{cases} \tag{21}$$

Clearly $\mathbf{s} = \mathbf{p} + \mathbf{d}$.

If there are $M = |\mathcal{S}_P|$ pilot symbols we define the $M \times K$ selection matrix $\mathbf{S}$ that selects the elements in $\mathbf{z}$ that correspond to non-zero entries in $\mathbf{p}$,

$$[\mathbf{S}]_{m,k} = \begin{cases} 1 & k \in \mathcal{S}_P, \ m = |\{k' \in \mathcal{S}_P \mid k' \leq k\}| \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Given the model in (19) and assuming the vector $\mathbf{s}$ is fully known, the received signal $\mathbf{z}$ has a sparse representation as,

$$\mathbf{z_S} = \mathbf{Sz} \approx \sum_{p=1}^{N_\tau} \sum_{q=1}^{N_a} \xi_{p,q} \mathbf{S}\boldsymbol{\Gamma}(a_q, \epsilon)\boldsymbol{\Lambda}(\tau_p)\mathbf{p} + \mathbf{Sv}. \tag{23}$$

The approximation is that the ICI from the data symbols in $\mathbf{d}$ is negligible,

$$\mathbf{S}\boldsymbol{\Gamma}(a_q, \epsilon)\boldsymbol{\Lambda}(\tau_p)\mathbf{d} \approx \mathbf{0} \quad \forall p, q, \tag{24}$$

which amounts to approximating the matrix $\boldsymbol{\Gamma}(a_q, \epsilon)$ as diagonal, since $\boldsymbol{\Lambda}(\tau_p)$ is already diagonal, and $\mathbf{SDd} = \mathbf{0}$ for any diagonal matrix $\mathbf{D}$.

C. Sparse Reconstruction

As in [4], [7], [17], [18] the channel coefficients $\xi_{p,q}$ are estimated using the following linear model,

$$\mathbf{z_S} = \mathbf{A_S}\mathbf{x} + \mathbf{w} \tag{25}$$

$$\mathbf{A_S} = \mathbf{S}\left[\boldsymbol{\Gamma}(a_1, \epsilon)\boldsymbol{\Lambda}(\tau_1)\mathbf{p} \quad \cdots \quad \boldsymbol{\Gamma}(a_{N_a}, \epsilon)\boldsymbol{\Lambda}(\tau_{N_\tau})\mathbf{p}\right] \tag{26}$$

$$\mathbf{x} = \begin{bmatrix} \xi_{1,1} & \cdots & \xi_{N_\tau, N_a} \end{bmatrix}^T \tag{27}$$

which is (approximately) independent of the unknown data vector $\mathbf{d}$. The noise vector $\mathbf{w} = \mathbf{Sv}$ had of length M. We indicate the dependence of the observations on the selection matrix $\mathbf{S}$ as $\mathbf{A_S}$. As sparse reconstruction algorithms we consider orthogonal matching pursuit (OMP) [23] and basis pursuit (BP) [24] as in [4], [7].

Finally the channel estimate $\hat{\mathbf{H}}$ is reconstructed from the path weight estimates $\hat{\xi}_{p,q}$ according to (19). After outlining the general signal design and receiver structure, we will compare various pilot placement patterns $\mathcal{S}_P$ and observation selection matrices $\mathbf{S}$.

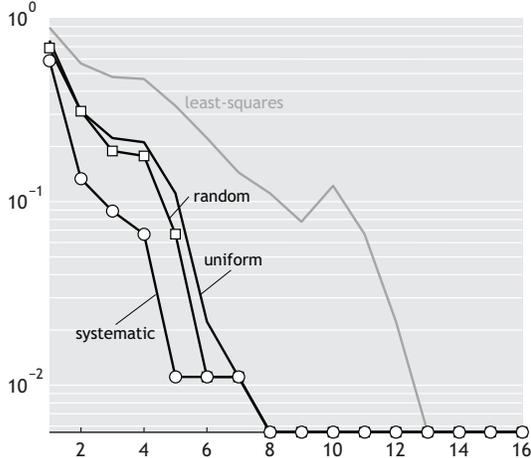## IV. COMPARISON OF PILOT DESIGNS

A. Considered Pilot Designs

a) Time-varying channel model: For non-sparse channel estimation, it is well known that for time-invariant channels, pilot designs with equi-distantly spaced pilot subcarriers are optimal [14]. While for time-varying channels it is largely agreed that equi-spaced clusters of equal size are optimal under most BEMs [14]–[16]. For sparse channel estimation, randomly placed pilots have been suggested in [17], [18], and a systematic design that consists of alternating clusters and single tones was considered in [4], [7]. We therefore compare two pilot designs for estimation of sparse *and* time-varying channels:

- *Systematic*: The pilot design used in [4], [7] places pilots in equi-spaced clusters of alternating size.
- *Random*: As used in [17], [18], the pilot tones are chosen randomly from the set of active subcarriers $\mathcal{S}_A$.

Each pilot design can be coupled with OMP or BP.

Comparison of Pilot Designs for OMP, QPSK
Block error rate (BLER) vs. number of phones

Comparison of Pilot Designs for BP, QPSK
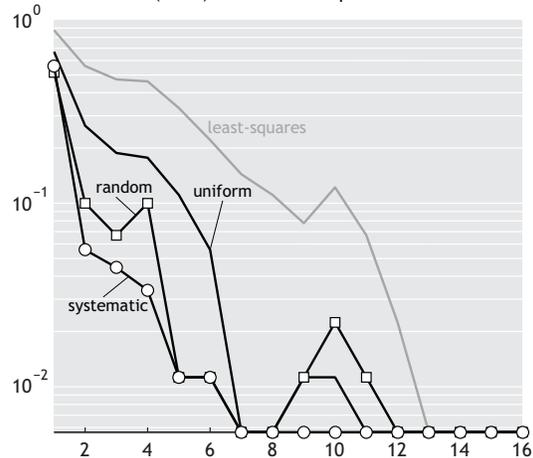Block error rate (BLER) vs. number of phones

Fig. 2. Performance comparison for different pilot designs for QPSK modulation and $1/4$ pilot overhead, leading to 5.27 kbit/s; pilot designs *systematic* and *random* assume a time-varying and sparse channel, while the *uniform* pilot design is based on a time-invariant channel model; *least-squares* is identical to *uniform*, but uses a conventional least-squares estimator instead of the sparse channel estimators OMP/BP.

b) Time-invariant channel model: For comparison we also consider channel estimators that approximate the channel as time-invariant. The pilot design is then based on equi-distantly spaced pilot subcarriers, and can be coupled with sparse channel estimation or non-sparse channel estimation:

- *Uniform*: The pilots are placed equi-distant between the data; OMP or BP can be used.
- *Least-squares*: For time-invariant and non-sparse channels, conventional least-squares channel estimation as, e.g., in [2], is used; the pilot pattern is the same as *uniform*.

For both the *uniform* and *least-squares* case, the channel estimate $\hat{\mathbf{H}}$ will be a diagonal matrix and demodulation in (16) reduces to per-subcarrier equalization. The observation model for all four cases is as in (23) and the matrix $\mathbf{S}$ only selects FFT outputs that correspond to pilot subcarriers.

B. Experimental Results

The presented experimental results are based on data recorded at the Calibration and Communications Sea Trial 2010 (CalCom'10), which was conducted about 30 km south of the port of Vilamoura, Portugal, on June 21-25, 2010. The water depth was approximately 100 m and the transmission distance about 1 km.

Channel estimation, demodulation, and decoding is performed for each OFDM block separately, and the considered performance metric is block error rate (BLER), i.e., the fraction of decoded OFDM blocks with one or more bit errors after error correction. The considered data set consists of nine transmissions of ten OFDM blocks at a time, leading to a total amount of 90 OFDM blocks for QPSK and 16-QAM each. Therefore, a BLER of less than $1/90$ in plotted results corresponds to error-free transmission, but is not plotted as zero BLER due to logarithmic scale. The x-axis in plots refers to multi-phone combining, where performance is increased by

incrementally combining phones in the receiver array, starting from the topmost one at 6 m depth; the considered vertical receiver array has a total of 16 hydrophones that are spaced in 4 m increments.

Results for varying pilot designs are shown in Fig. 2; we consider both OMP and BP recovery algorithms; the pilot overhead is set to $1/4$. For both recovery algorithms we find that the systematic pilot design leads to the best performance. As argued in [4], [7], we explain this with the need for a sufficient amount of adjacent observations to estimate the ICI. The random pilot placement will have some neighboring or closely spaced pilots to observe the ICI, but the channel estimation error seems to be larger. Comparing to the channel estimators that approximate the channel as time-invariant, there are significant gains, especially over the conventional least-squares estimator
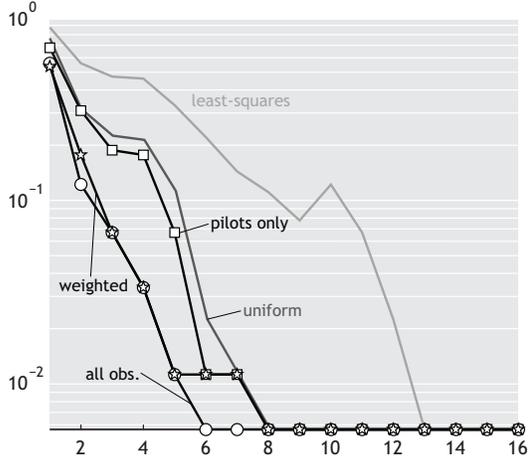
V. OBSERVATION MASKING

A. Data Subcarriers in Channel Estimation

Due to the ICI caused by time-varying channels, most or all FFT outputs in $\mathbf{z}$ will be affected by both pilot tones and unknown data. Although there are some approaches that jointly estimate the channel and the data symbols (see references in [14]), this usually leads to significantly more complex receivers. When the matrix $\mathbf{H}$ is treated as strictly banded, pilot and data can be decoupled by inserting guard bands of null subcarriers [15], then only observations on subcarriers corresponding to pilot symbols are used. Since the matrix $\mathbf{H}$ is not strictly banded and the use of null subcarriers wastes significant spectral resources, we focus on pilot designs without guard bands. The key question we are interested in is: *Can the FFT outputs corresponding to data subcarriers be used to improve channel estimation?*

To test this, the selection matrix $\mathbf{S}$ in (22) is exchanged with a generalized observation masking matrix. We consider three
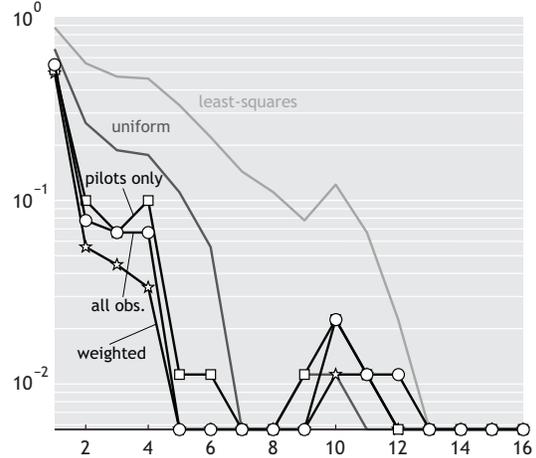
Fig. 3. Considering only the *random* pilot design, performance varies considerably depending on whether only FFT outputs corresponding to pilot subcarriers are used, or all FFT outputs are considered as observations in channel estimation. In the latter case weighting the observations corresponding to their perceived noise plus interference level improves the BP performance significantly; $1/4$ pilots leads to 5.27 kbit/s.

cases:

- *Pilots only*: As before we assume that the ICI between data and pilot subcarriers is negligible, only FFT outputs that correspond to pilot subcarriers are used as observations and $\mathbf{S}$ is as defined in (22).
- *All observations*: We assume that the ICI is significant, so all FFT outputs are used as observations, $\mathbf{S} = \mathbf{I}_K$; the unknown data symbols are simply replaced with their mean of zero, i.e., their effect on channel estimation is neglected.
- *Weighted:* Again, all FFT outputs are used as observations, but the unknown data is treated as additional noise; the observations are weighted to achieve a uniform noise model with diagonal matrix $\mathbf{S} = \mathbf{W}$, as explained next.

B. Modeling Unknown Data

Without the selection matrix (23) becomes

$$\mathbf{z} = \sum_{p=1}^{N_\tau} \sum_{q=1}^{N_a} \xi_{p,q} \mathbf{\Gamma}(a_q, \epsilon) \mathbf{\Lambda}(\tau_p) (\mathbf{p} + \mathbf{d}) + \mathbf{v}. \quad (28)$$

Using all FFT outputs as observations without any further changes is equivalent to modeling the unknown vector $\mathbf{d}$ by only its first moment $E[\mathbf{d}] = \mathbf{0}$,

$$\mathbf{z}_\mathbf{I} \approx \sum_{p=1}^{N_\tau} \sum_{q=1}^{N_a} \xi_{p,q} \mathbf{\Gamma}(a_q, \epsilon) \mathbf{\Lambda}(\tau_p) \mathbf{p} + \mathbf{v}. \quad (29)$$

This leads to the following model

$$\mathbf{z}_\mathbf{I} = \mathbf{A}_\mathbf{I} \mathbf{x} + \mathbf{v} \quad (30)$$

$$\mathbf{A}_\mathbf{I} = \begin{bmatrix} \mathbf{\Gamma}(a_1, \epsilon) \mathbf{\Lambda}(\tau_1) \mathbf{p} & \cdots & \mathbf{\Gamma}(a_{N_a}, \epsilon) \mathbf{\Lambda}(\tau_{N_\tau}) \mathbf{p} \end{bmatrix}, \quad (31)$$

where $\mathbf{x}$ and $\mathbf{v}$ remain as previously defined.

To model the unknown data symbols as additional noise based on their second moment, we start with

$$\mathbf{z} = \sum_{p=1}^{N_\tau} \sum_{q=1}^{N_a} \xi_{p,q} \mathbf{\Gamma}(a_q, \epsilon) \mathbf{\Lambda}(\tau_p) \mathbf{p} + (\mathbf{H} \mathbf{d} + \mathbf{v}). \quad (32)$$

The new noise term now depends on the unknown channel matrix $\mathbf{H}$. To proceed, we disregard ICI caused by data subcarriers and approximate the channel for this purpose as $\mathbf{H} \approx \sqrt{\gamma} \mathbf{I}_K$, where γ is the average received signal power. With this we can characterize the noise covariance as

$$\mathrm{Cov}\{\mathbf{H} \mathbf{d} + \mathbf{v}\} = \gamma \mathbf{D}_\mathbf{d} + N_0 \mathbf{I}_K, \quad (33)$$

where $\mathbf{D}_\mathbf{d}$ is a diagonal matrix that accounts for the unknown data symbols

$$[\mathbf{D}_\mathbf{d}]_{k,k} = \begin{cases} 1 & k \in \mathcal{S}_D \\ 0 & k \in \mathcal{S}_N \cup \mathcal{S}_P. \end{cases} \quad (34)$$

As the noise covariance matrix is kept diagonal by the various assumptions and approximations, a white noise model can be easily achieved by premultiplying with the diagonal weighting matrix

$$[\mathbf{W}]_{k,k} = \begin{cases} \sqrt{\frac{N_0}{\gamma + N_0}} & k \in \mathcal{S}_D \\ 1 & k \in \mathcal{S}_N \cup \mathcal{S}_P. \end{cases} \quad (35)$$
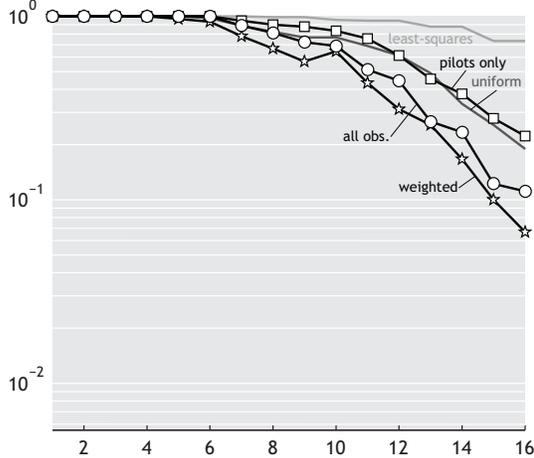
Accordingly, the linear model is

$$\mathbf{z}_\mathbf{W} = \mathbf{A}_\mathbf{W} \mathbf{x} + \mathbf{W} (\mathbf{H} \mathbf{d} + \mathbf{v}) \quad (36)$$

$$\mathbf{A}_\mathbf{W} = \mathbf{W} \begin{bmatrix} \mathbf{\Gamma}(a_1, \epsilon) \mathbf{\Lambda}(\tau_1) \mathbf{p} & \cdots & \mathbf{\Gamma}(a_{N_a}, \epsilon) \mathbf{\Lambda}(\tau_{N_\tau}) \mathbf{p} \end{bmatrix}. \quad (37)$$
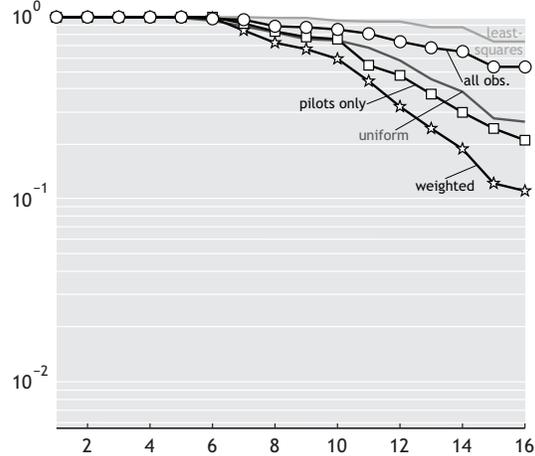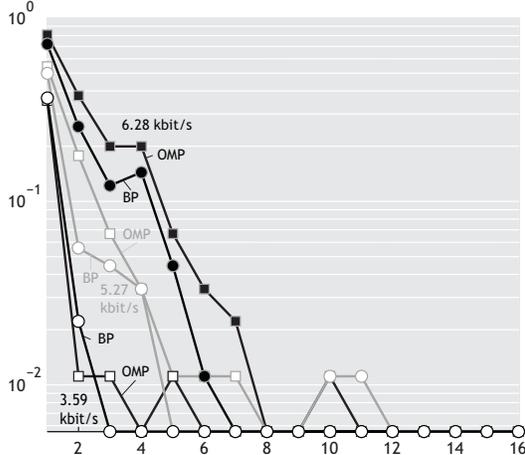
Fig. 4. Performance results for 16-QAM modulation, using some or all FFT outputs as observations; 1/4 randomly placed pilots lead to 10.5 kbit/s.
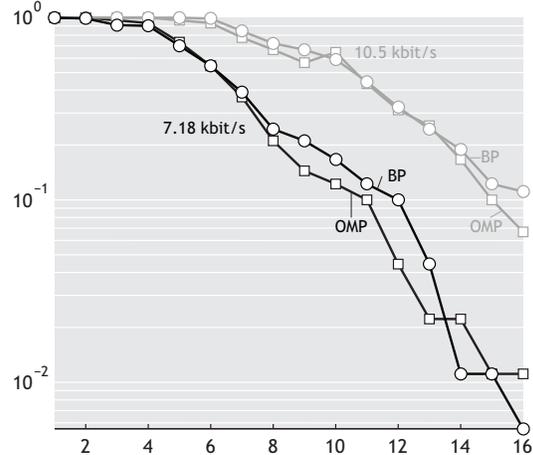

Fig. 5. Performance comparison for different levels of pilot overhead; for QPSK the pilot overhead is decreasing form left-to-right as 1/2-1/4-1/8, for 16-QAM modulation 1/2 and 1/4 are used; all results are based on random pilot placement with with weighted observations.

C. Experimental Results

As we found that the random pilot placement suffered from too few adjacent observations, we now focus on this pilot design, but subcarriers corresponding to data subcarriers can now be included in channel estimation. We show results in Fig. 3 for QPSK data and in Fig. 4 for 16-QAM.

In the case of BP and QPSK modulation, we see that using all active subcarriers as observations leads to some limited improvement, while weighting the observations based on data subcarriers relative to observations based on pilot subcarriers gains significantly more. In contrast, for OMP both schemes that use the data subcarriers as observations gain significantly over using only pilot symbols. We explain this by the fact that without weighting, the unknown data symbols will lead to a large ℓ_2 reconstruction error; while BP is directly affected by the ℓ_2 fit as part of its cost function, we configured OMP with

a stopping criterion that focuses on the decrease in its ℓ_2 fit, which is relatively independent of its absolute value.

Considering the same setup with 16-QAM in Fig. 4, the superiority of weighting observations is more pronounced. Especially for BP, we found that the increased ℓ_2 error leads to heavily biased channel estimates, i.e., the channel estimates seemed to be scaled down, leading in turn to too large symbol estimates. As QPSK is invariant with respect to scaling errors, this was less pronounced in this case.

VI. PILOT OVERHEAD

Finally we present results based on varying the pilot overhead with *random* pilot design and *weighted* observations. In Fig. 5 we show performance results for all data rates listed in Table II, comparing BP with OMP. Increasing the number of pilots decreases the data rate, but improves the BLER performance; to make up for the decreased number of data

symbols we can transmit more bits per symbol by switching from QPSK to 16-QAM.

As can be seen in the results, the BLER performance improves significantly when increasing the number of pilots. Unfortunately, using only these two modulation schemes we do not observe a tradeoff, since the achievable data rates and BLER performance curves do not overlap. Still, it is interesting to note that in time-varying channels adaptive modulation and coding can include a third tradeoff by changing the number of pilots. We explain this gradual performance tradeoff with two facts: first, since pilots are scarce and the parameter space of the channel is large it is not possible to disregard channel estimation error with respect to BLER performance; and second, since UWA channels are approximately sparse, more pilots will lead to a less sparse estimate, revealing more of the channel details and leading to a gradual decrease in channel estimation error.

VII. Conclusion

In this paper, we have considered sparse channel estimation for time-varying UWA channels using multicarrier transmission. Due to the time-variation the orthogonality between subcarriers is lost, making it costly in terms of spectral efficiency to use pilot designs that keep pilots and data symbols fully orthogonal at the receiver. We found that observations on adjacent subcarriers are required to estimate the time-varying channel, which matches existing results for estimation of non-sparse time-varying channels. This contradicts some early notions that for sparse channel estimation, observations should simply be placed randomly among the data. Furthermore, when pilots and data symbols are not separated at the receiver, we found that subcarriers corresponding to data symbols can be used to observe the ICI caused by the neighboring pilot subcarriers. In this case the effect of the unknown data symbols should be modeled as additional colored noise, which is especially needed for BP, while OMP seems to be less affected by an unknown noise spectrum. Finally, we also considered varying the number of pilots to improve BLER performance at the cost of reduced data rate. We found that this leads to an almost continuous tradeoff that can be included as an additional dimension in any adaptive modulation and coding approach.

References

[1] M. Stojanovic, "Low complexity OFDM detector for underwater channels," in *Proc. of MTS/IEEE OCEANS Conf.*, Boston, MA, Sept. 18-21, 2006.

[2] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 198–209, Apr. 2008.

[3] T. Kang and R. A. Iltis, "Iterative carrier frequency offset and channel estimation for underwater acoustic OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 9, pp. 1650–1661, Dec. 2008.

[4] C. R. Berger, S. Zhou, J. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.

[5] K. Tu, D. Fertonani, T. M. Duman, and P. Hursky, "Mitigation of intercarrier interference in OFDM systems over underwater acoustic channels," in *Proc. of MTS/IEEE OCEANS Conf.*, Bremen, Germany, May 2009.

[6] G. Leus and P. A. van Walree, "Multiband OFDM for covert acoustic communications," *IEEE J. Select. Areas Commun.*, vol. 26, no. 9, pp. 1662–1673, Dec. 2008.

[7] C. R. Berger, Z. Wang, J.-Z. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 1–11, Nov. 2010.

[8] S.-J. Hwang and P. Schniter, "Efficient communication over highly spread underwater acoustic channels," in *Proc. of the ACM Intl. Workshop on Underwater Networks (WUWNet)*, Montréal, Quèbec, Sep. 2007.

[9] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[10] C. Carbonelli and U. Mitra, "A simple sparse channel estimator for underwater acoustic channels," in *Proc. of MTS/IEEE OCEANS Conf.*, Vancouver, Canada, Oct. 2007.

[11] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE J. Ocean. Eng.*, vol. 32, no. 4, pp. 927–939, Oct. 2007.

[12] M. Stojanovic, "OFDM for underwater acoustic communications: Adaptive synchronization and sparse channel estimation," in *Proc. of Intl. Conf. on Acoustics, Speech and Signal Proc.*, Las Vegas, NV, Apr. 2008.

[13] E. Zamanizadeh, J. Gomes, and J. Bioucas-Dias, "Identification and matching of sparse delay-Doppler spread functions from high-frequency communications signals," in *Proc. of MTS/IEEE OCEANS Conf.*, Seattle, WA, Sep. 2010.

[14] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–26, Nov. 2004.

[15] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[16] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 2226–2238, May 2007.

[17] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Select. Topics Signal Proc.*, vol. 4, no. 2, pp. 255–271, Apr. 2010.

[18] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. of the IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[19] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[20] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[21] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: Identifiability, algorithms, and performance," *IEEE J. Select. Areas Commun.*, vol. 19, no. 12, pp. 2504–2515, Dec. 2001.

[22] M. Tuechler, A. C. Singer, and R. Koetter, "Minimum mean square error equalization using a priori information," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 673–683, Mar. 2002.

[23] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[24] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.