Iterative Sparse Channel Estimation and Spatial Correlation Learning for Multichannel Acoustic OFDM Systems

Amir Tadayon , *Graduate Student Member, IEEE*, and Milica Stojanovic, *Fellow, IEEE*

Abstract—This article addresses the problem of coherent detection of acoustic orthogonal frequency division multiplexing (OFDM) signals using a sparse channel estimation method based on a physical model of multipath propagation. Unlike the conventional sample-spaced and subsample-spaced methods, such as least squares and orthogonal matching pursuit (OMP), which target the *taps* of an equivalent discrete-delay channel response, the path identification (PI) method targets the physical propagation *paths* in a continuous-delay domain, and focuses on explicit estimation of delays and complex amplitudes of the channel paths in an iterative fashion. When multiple receive elements are available, two situations are possible: one in which the array elements see uncorrelated channel responses, and another in which the channel responses are correlated. In the first case, channel estimation must be accomplished element-by-element. This is done simply by applying the PI algorithm to each element individually. In the second case, correlation between the elements can be exploited. In doing so, our goal is to reduce the signal processing complexity without compromising the performance. Toward this goal, an adaptive precombining method is proposed. Without requiring any *a priori* knowledge about the spatial distribution of received signals, the method exploits spatial coherence between receive channels by linearly combining them into fewer output channels so as to reduce the number of subsequent channel estimators. The algorithm learns the spatial coherence pattern recursively over the carriers, thus effectively achieving broadband beamforming. The reduced-complexity precombining method relies on differential encoding that keeps the receiver complexity at a minimum and requires a very low pilot overhead. Using synthetic data as well as 210 experimental signals transmitted over a 3–7-km shallow-water channel in the 10.5–15.5-kHz acoustic band during a 3.5-h experiment, we study the system performance in terms of data detection mean-squared error (MSE), symbol error rate, and bit error rate (BER), and show that the PI algorithm achieves excellent MSE performance while its complexity is considerably lower than that of the OMP algorithm. We also demonstrate that the receiver equipped with the proposed reduced-complexity precombining scheme requires three times fewer channel estimators while achieving the same MSE and BER performance as the full-complexity receiver.

Index Terms—Multichannel acoustic orthogonal frequency division multiplexing (OFDM), path identification (PI), recursive least squares (RLS), sparse channel estimation, spatial correlation learning, stochastic gradient descent.

I. INTRODUCTION

HIGH data rate coherent transmission over acoustic channels is a challenging problem due to the combined effects of long multipath and Doppler fluctuations. To account for these effects, we design a coherent receiver based on multi-carrier modulation in the form of orthogonal frequency division multiplexing (OFDM). OFDM is an attractive method for data transmission over frequency-selective channels due to its ability to achieve high bit rates at reasonably low computational loads. This fact motivates the use of OFDM in mobile acoustic communications where the channel exhibits long multipath delays but each narrowband carrier only experiences flat fading, thus eliminating the need for time-domain equalizers [1]–[4].

Reliable coherent data detection requires the channel state information (CSI) at the receiver. Pilot-assisted channel estimation is used as a standard method to obtain the necessary CSI for reliable coherent communications [5]–[11]. In [5], a pilot-aided channel estimator based on 2-D Wiener filtering, which is optimal in the mean-squared error (MSE) sense, is proposed. In [6], a low-rank channel estimator for OFDM systems was proposed based on the singular-value decomposition (SVD) or frequency-domain filtering. While these channel estimators are optimal in the MSE sense, they require *a priori* information about the channel statistics, which is not usually available in practice, and have high computational complexity. The conventional least squares (LS) algorithm targets estimation of sample-spaced channel taps, with sampling at the basic rate equal to the system bandwidth [7]. Although the LS method has low complexity, its performance suffers when the channel is not sample-spaced, as is the case in most practical situations.

Detection of OFDM signals requires the knowledge of channel coefficients in the frequency domain. However, channel is typically estimated in the impulse response domain (delay domain), where fewer delay-domain coefficients suffice to describe it [10]. With the advent of sparse estimation, focus has come to the sparse nature of physical multipath channels, as it can be leveraged to further improve channel estimation to either work with fewer pilot symbols or to achieve better noise suppression.

Sparse channel estimation for acoustic OFDM systems has been studied extensively in recent years [12]–[14]. These studies

have put forth a number of channel estimation algorithms that take advantage of this fact. In [12] and [13], a channel estimator based on the greedy matching pursuit and its orthogonal variant, orthogonal matching pursuit (OMP), using dictionaries with finer delay resolution (subsample spacing) has been addressed. This method can approximately reflect the fact that the physical path delays have a continuum of values. However, the improvement in the performance of channel estimation that results from using finer dictionaries comes at increased computational complexity, and further studies revealed an effect of strongly diminishing returns for finer dictionaries [13].

To address these issues, we focus on a physical, path-based channel model, which is amenable to explicit channel estimation, where the channel is parametrized by a number of distinct paths, each characterized by a delay and a complex amplitude. We target a continuum of path delays, eliminating the sample-spaced model and focusing instead on processing a transformed version of the signal observed over all the carriers spanning the system bandwidth. We draw on the ideas of the work in [15], where we proposed the basic approach of path identification (PI) for a single-element receiver. Unlike the sparse identification methods, the resolution and coverage in delay that the PI method provides can be increased arbitrarily without a prohibitive cost to complexity. In a digital implementation, the path delays are of course discretized, but the resolution can be arbitrary, i.e., it is not constrained by the algorithmic complexity. The PI algorithm focuses on explicit estimation of delays and complex amplitudes of the channel paths. Not only does it operate in a continuous estimation space, but it also eliminates the need to know the statistics of the channel, which is crucial for the channel estimation proposed in [5] and [6].

In addition to temporal fluctuations, spatial variability of the underwater channel presents a major problem for single-channel receivers and motivates the use of multiple spatially distributed receivers that offer robustness to fading [16]. However, multichannel processing of high rate underwater acoustic (UWA) communication signals requires computationally expensive receiver algorithms, and an increase in the number of receiving elements significantly increases the receiver complexity. When multiple receiving elements are available, two situations are possible: one in which the array elements see uncorrelated channel responses, and another in which the channels responses are correlated. In the first case, channel estimation must be accomplished element-by-element. This is done simply by applying the PI algorithm to each element individually. In the second case, correlation between the elements can be exploited so as to reduce the number of channel estimators.

In this paper, we propose a precombining method that exploits spatial coherence between receiving elements and provides the desired reduction in signal processing complexity. The approach is motivated by time-domain equalization in single-carrier systems [16], which we reformulate in light of multicarrier systems. The proposed scheme exploits spatial correlation between the input channels by linearly combining them so as to reduce the number of channel estimators in a coherent receiver. The method makes no assumptions about the spatial distribution of signals,

relying only on the fact that some coherence will exist between the signals received on different elements when the elements are spaced closely with respect to wavelength. The algorithm learns spatial coherence between the channels adaptively, and allows the precombiner coefficients to change from one carrier to another, thus effectively accomplishing broadband processing. The adaptive precombining method relies on differential encoding, which keeps the receiver complexity at minimum and requires only a very low pilot overhead. The proposed algorithm has lower complexity than the method proposed in [17] at the cost of negligible degradation in the MSE performance.

The techniques are demonstrated on experimental data from the Mobile Acoustic Communication Experiment (MACE'10), showing excellent results. In the MACE'10 experiment, the transmitter moved at a relative speed of 0.5–1.5 m/s with respect to the receiver, and OFDM blocks containing up to 2048 QPSK/8-phase shift keying (8-PSK) modulated carriers occupied the acoustic frequency range between 10.5 and 15.5 kHz. The proposed method achieves excellent performance in these challenging conditions.

The rest of this paper is organized as follows. In Section II, we introduce the system and channel model. Section III discusses the conventional (sub)sample-spaced channel estimation algorithms and details the PI algorithm. Section IV presents the precombining algorithm. Sections V and VI contain the results of synthetic and experimental data processing. Conclusions are summarized in Section VII.

II. System and Channel Model

A. OFDM System

We consider an OFDM system with M_r receivers and K carriers within a total bandwidth B. Let f_0 and $\Delta f = B/K$ denote the first carrier frequency and the carrier spacing, respectively. The transmitted OFDM block is then given by

$$s(t) = Re\left\{\sum_{k=0}^{K-1} d_k e^{2\pi i f_k t}\right\}, \quad t \in [0, T] \quad (1)$$

where $T = 1/\Delta f$ is the OFDM block duration. The data symbol d_k, which modulates the kth carrier of frequency $f_k = f_0 + k\Delta f$, belongs to a unit amplitude PSK alphabet.

After synchronization, carried out using the method proposed in [18],[1] and overlap-and-add procedure (or cyclic-prefix removal), the lowpass equivalent received signal on the mth receiving element is modeled as

$$v_m(t) = \sum_{k=0}^{K-1} H_k^m d_k e^{2\pi i k \Delta f t} + w_m(t), \quad t \in [0, T] \quad (2)$$

where H_k^m is the channel frequency response at the kth carrier of the mth receiving element, and $w_m(t)$ is the equivalent baseband

---

[1]Synchronization includes frame synchronization, initial resampling, and frequency offset compensation.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING 3
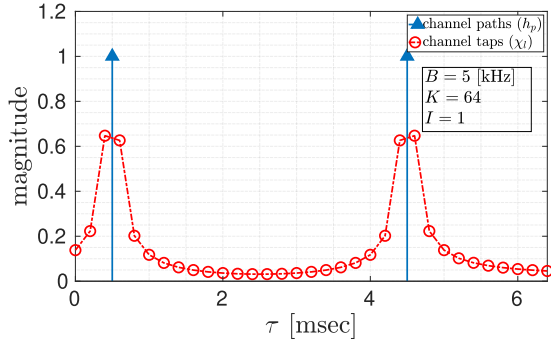

Fig. 1. Spillage between taps for the physical channel $h(\tau) = \delta(\tau - 2.5\frac{T}{K}) + \delta(\tau - 42.5\frac{T}{K})$. Note that the number of taps required to describe the channel is greater than the number of physical propagation paths.

noise. Fast Fourier transform (FFT) demodulation yields the observations

$$y_k^m = \frac{1}{T}\int_T v_m(t)e^{-2\pi ik\Delta ft}dt$$

$$= d_k H_k^m + z_k^m, \quad k = 0, 1, \ldots, K-1 \quad (3)$$

where z_k^m is the corresponding noise.

B. Channel Model

In this section, we introduce the physical, path-based channel model as an alternative to the conventional discrete-delay (sample-spaced) channel model. In the path-based channel model, the channel is parameterized by a pair (τ_p, h_p), $p = 1, \ldots, N_p$ where the path delays τ_p have a continuum of values, the coefficients h_p represent the path amplitudes, and N_p is the number of channel paths. At the carrier frequency $f_k = f_0 + k\Delta f$, the channel frequency response is given by

$$H_k = H(f_k) = \sum_{p=1}^{N_p} h_p e^{-2\pi if_k\tau_p} = \sum_{p=1}^{N_p} c_p e^{-2\pi ik\Delta f\tau_p} \quad (4)$$

where $c_p = h_p e^{-2\pi if_0\tau_p}$.

A baseband discrete-delay model with delay spacing $\Delta\tau$ is given by

$$H_k = \sum_l \chi_l e^{-2\pi ik\Delta fl\Delta\tau} \quad (5)$$

where χ_l is the channel tap. If $\Delta\tau = T/K$ and $T = 1/\Delta f$, then χ_l is given by

$$\chi_l = \frac{1}{K}\sum_{p=1}^{N_p} c_p e^{-\pi i(K-1)\left(\frac{\tau_p}{T} - \frac{l}{K}\right)}$$

$$\times \frac{\sin\left(\pi K(\frac{\tau_p}{T} - \frac{l}{K})\right)}{\sin\left(\pi(\frac{\tau_p}{T} - \frac{l}{K})\right)}, \quad l = 0, \ldots, K-1. \quad (6)$$

From (6), it follows that if the channel path delay τ_p is not an integer multiple of the delay spacing, then the channel path amplitude h_p will leak to all the taps χ_l. Fig. 1 illustrates such leakage for a special case of a two-path equal-amplitude channel.

Alternatively, the "superresolution" models are based on using $\Delta\tau = T/KI$, where $I > 1$ accounts for an increased resolution in delay. Using a resolution higher than the basic T/K reduces the power spillage among adjacent taps of the estimated impulse response, ultimately aiming for a minimal channel representation where the number of taps tends to the number of physical propagation paths.

III. CHANNEL ESTIMATION

A. Channel Estimation Based on the Discrete-Delay Model

Conventionally, channel estimation is based on the discrete-delay model. When $\Delta\tau = T/K$, with $T = 1/\Delta f$ we have the usual discrete Fourier transform relationship between the vector $\mathbf{H}$ of channel coefficients H_k in the frequency domain and the vector $\boldsymbol{\chi}$ of channel taps χ_l in the impulse response domain

$$\mathbf{H} = \mathbf{F}_K\boldsymbol{\chi} \quad (7)$$

where $\mathbf{F}_K$ is the $K \times K$ Fourier matrix. If the multipath spread $T_{mp} < T$ is within LT/K where $L < K$, i.e., the support of $\boldsymbol{\chi}$ is limited to $L < K$, then we can also write

$$\mathbf{H} = \mathbf{F}_{K\times L}\boldsymbol{\chi}_L \quad (8)$$

where only the first L columns of the Fourier matrix $\mathbf{F}_K$ and the first L elements of the model vector $\boldsymbol{\chi}$ are kept. Alternatively, the superresolution discrete-delay model and attendant sparse estimators are based on using $\Delta\tau = T/KI$, where $I > 1$ accounts for an increased resolution in the delay domain.

Assuming without the loss of generality that all K data symbols are available for channel estimation (e.g., correct symbol decisions, or all-pilots in an initial block), the input to the channel estimator is given by $x_k = y_k/d_k$, $k = 0, \ldots, K-1$, i.e.,

$$\mathbf{x} = \mathbf{H} + \mathbf{z} \quad (9)$$

where $\mathbf{z}$ is the noise vector. If the data symbols are available only on pilot carriers, the vector $\mathbf{x}$ is formed using those carriers only.

The conventional LS estimate is then given by[2]

$$\hat{\boldsymbol{\chi}}_L = \frac{1}{K}\mathbf{F}'_{K\times L}\mathbf{x} \quad (10a)$$

$$\hat{\mathbf{H}} = \mathbf{F}_{K\times L}\hat{\boldsymbol{\chi}}_L. \quad (10b)$$

In many practical applications, including UWA communications, the multipath channel can be considered as sparse, as the number of significant paths is small even when the channel delay spread is long. Based on this assumption, greedy algorithms, such as OMP, can be used to estimate the channel response. The OMP algorithm identifies the dominant channel taps sequentially. At each iteration it selects one column of the dictionary that correlates best with the approximation residual from the previous iteration and recomputes the coefficients by solving a constraint LS problem to optimally fit the observations [19]. The OMP algorithm is summarized in Algorithm 1.

---

[2]Conjugate transpose is denoted by $(\cdot)'$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

IEEE JOURNAL OF OCEANIC ENGINEERING

Algorithm 1: OMP Algorithm With MGS.

Input: K-dimensional noisy channel observation vector $\mathbf{x}$, $K \times IL$ dictionary $\mathbf{F}_{K \times IL}$ ($\mathbf{F}_{K \times IL}$ obtained by keeping the first K rows and the first IL columns of a DFT matrix of size $IK \times IK$), and sparsity level $N_t \leq IL$ of χ (or threshold η)

Output: channel frequency response estimate $\hat{\mathbf{H}}$

1:   $\mathbf{\Phi}_0 = \emptyset$, $\mathbf{\Phi}_{\perp,0} = \emptyset$, $\mathcal{I}_0 = \emptyset$, and $t = 1$
2:   $\mathbf{r}_t = \mathbf{x}$
3:   **while** $t \leq N_t$ (or $|r_t(\tau)| > \eta \max_\tau |r(\tau)|$) **do**
4:     $i_t = \arg\max_{1 \leq i \leq IL} |\mathbf{F}'_{K \times IL} \mathbf{r}_t|$
5:     $\mathcal{I}_t = \mathcal{I}_{t-1} \cup \{i_t\}$
6:     $\phi_t = [\mathbf{F}_{K \times IL}]_{i_t}$
7:     $\mathbf{\Phi}_t = [\mathbf{\Phi}_{t-1} \ \phi_t]$
8:     **for** $k \in \{1, \ldots, t-1\}$ **do**
9:       $\phi_t = \phi_t - (([\mathbf{\Phi}_{\perp,t-1}]_k)' \phi_t) ([\mathbf{\Phi}_{\perp,t-1}]_k)'$
10:    **end for**
11:    $\phi_t = \frac{\phi_t}{\|\phi_t\|_2}$
12:    $\mathbf{\Phi}_{\perp,t} = [\mathbf{\Phi}_{\perp,t-1} \ \phi_t]$
13:    $\check{\chi}_t = \mathbf{\Phi}'_{\perp,t} \mathbf{x}$
14:    $\mathbf{r}_{t+1} = \mathbf{x} - \mathbf{\Phi}_{\perp,t} \check{\chi}_t$
15:    $t \leftarrow t + 1$
16:   **end while**
17:   $\check{\chi}_{N_t} = (\mathbf{\Phi}'_{N_t} \mathbf{\Phi}_{N_t})^{-1} \mathbf{\Phi}'_{N_t} \mathbf{x}$
18:   $[\hat{\chi}]_{i_t} = [\check{\chi}_{N_t}]_t$, $t = 1, \ldots, N_t$
19:   **return** $\hat{\mathbf{H}} = \mathbf{F}_{K \times IL} \hat{\chi}$

The running time of the OMP algorithm is dominated by step 1, whose total cost is $O(N_t ILK)$.[3] At iteration t, the LS problem can be solved with marginal cost $O(tIK)$. To do so, we maintain a QR factorization of $\mathbf{F}_{K \times IL}$. Our implementation uses the modified Gram–Schmidt (MGS) algorithm. Extensive details and a survey of alternate approaches can be found in [20].

B. Channel Estimation Based on the Physical Model

The physical model is based on the expression (4), which can be written as

$$\mathbf{H} = \sum_{p=1}^{N_p} c_p \mathbf{s}_K(2\pi \Delta f \tau_p) \tag{11}$$

where $\mathbf{s}_K(2\pi \Delta f \tau) = \begin{bmatrix} 1 & e^{-2\pi i \Delta f \tau} & \cdots & e^{-2\pi i (K-1) \Delta f \tau} \end{bmatrix}^T$ is referred to as the steering vector at an arbitrary delay τ.

Consider now the following operation performed on the noisy channel observation $\mathbf{x}$:

$$r(\tau) = \frac{1}{K} \mathbf{s}'_K(2\pi \Delta f \tau) \mathbf{x}$$

$$= \sum_{p=1}^{N_p} c_p g_K(2\pi \Delta f (\tau - \tau_p)) + w(\tau), \ \tau \in \tau_{obs} \tag{12}$$

where the interval τ_{obs} is a preset interval that captures the multipath spread,[4] $w(\tau)$ is the corresponding noise, and

$$g_K(\varphi) = \frac{1}{K} \sum_{k=0}^{K-1} e^{ik\varphi} \tag{13}$$

is a known signature function. This operation corresponds to steering across the carriers. Fig. 2(a) shows the signature waveform (magnitude) and Fig. 2(b) illustrates the signal $r(\tau)$ corresponding to the physical channel shown in Fig. 1.

The fact that the signature waveform is *known* can be exploited to estimate the channel parameters explicitly. When we say explicitly, we mean that we are targeting directly both the path gains c_p and the path delays τ_p, unlike in the conventional estimation where the delay axis is discretized to avoid the nonlinear problem of delay estimation.

Joint estimation of the parameters c_p and τ_p can be performed as follows. We start by setting

$$r_1(\tau) = r(\tau) \tag{14}$$

and evaluate this function for a preset range of delays τ with an arbitrary resolution $\Delta\tau$ in the delay domain. The range can be determined in accordance with the multipath spread T_{mp}.[5] An iterative procedure now follows over the path indices $p = 1, \ldots, N_p$. In the pth iteration, we estimate the path delay as

$$\hat{\tau}_p = \arg\max_\tau |r_p(\tau)| \tag{15}$$

and the path coefficient as

$$\hat{c}_p = r_p(\hat{\tau}_p). \tag{16}$$

We then subtract this path's contribution from the current signal, so as to form the signal for the next iteration (next path)

$$r_{p+1}(\tau) = r_p(\tau) - \hat{c}_p g_K(2\pi \Delta f(\tau - \hat{\tau}_p)). \tag{17}$$

The procedure ends according to a predefined criterion such as an *a priori* set number of paths N_p, or when the power in the residual reaches a certain threshold η or stops to change significantly.

An extension to the above algorithm can also be applied to improve the quality of the estimates $\hat{c}_p$. Once the algorithm has been executed, the path coefficients $\hat{c}_p$ generated in the process are discarded, but the delay estimates $\hat{\tau}_p$ are kept. The delay estimates are used to form the matrix

$$\hat{\mathbf{S}} = \begin{bmatrix} \mathbf{s}_K(2\pi \Delta f \hat{\tau}_1) & \cdots & \mathbf{s}_K(2\pi \Delta f \hat{\tau}_{N_p}) \end{bmatrix}. \tag{18}$$

This matrix represents an estimate of the true matrix $\mathbf{S}$, which relates the vector $\mathbf{H}$ with the vector of path coefficients $\mathbf{c} = \begin{bmatrix} c_1 & \cdots & c_{N_p} \end{bmatrix}^T$ as $\mathbf{H} = \mathbf{S}\mathbf{c}$, i.e., it defines the observed signal as

$$\mathbf{x} = \mathbf{S}\mathbf{c} + \mathbf{z}. \tag{19}$$

---

[3]Note that step 4 can also be implemented using FFT whose complexity is $O(IK \log_2 IK)$. This will reduce the computational cost if $\log_2(IK) < L$ in limit.

[4]For example, the observation interval can be chosen as $\tau_{obs} = [0, T_g]$. To counteract the effect of time synchronization in non-minimum phase channels such as underwater channels [21], the observation interval τ_{obs} can be chosen as $\left[-\frac{1}{4} T_g, \frac{3}{4} T_g\right]$ assuming that multipath spread T_{mp} is within τ_{obs}.

[5]In a digital implementation, an arbitrary resolution is used, e.g., $\Delta\tau = T/IK$, where I represents the resolution factor, i.e., the increase in resolution over the standard sample spacing $1/B = T/K$. The total length of the observation interval is $L = \lceil T_g B \rceil$, where T_g is the guard interval which is at least as long as the multipath spread.
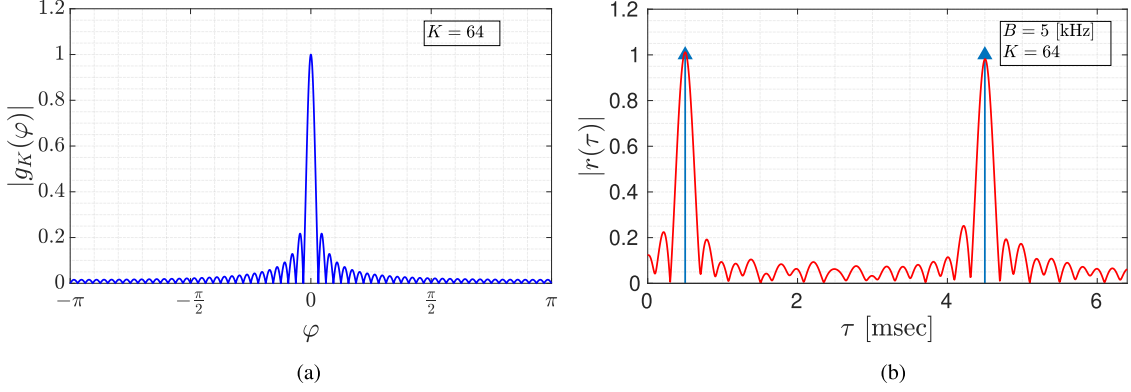
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING

5

Fig. 2.   (a) Signature waveform. (b) Signal for channel estimation corresponding to the physical channel shown in Fig. 1.

represents an improvement, which we will illustrate numerically in Section III-C.

Algorithm 2: PI Algorithm With LS Refinement Step.

Input: K-dimensional noisy channel observation vector $\mathbf{x}$, a preset interval τ_{obs}, and number of channel paths N_p (or threshold η)

Output: channel frequency response estimate $\hat{\mathbf{H}}$

1:   $p = 1$
2:   $r_p(\tau) = r(\tau) = \frac{1}{K}\mathbf{s}'_K(2\pi\Delta f \tau)\mathbf{x}, \ \tau \in \tau_{obs}$
3:   **while** $p \leq N_p$ (or $|r_p(\tau)| > \eta \max_\tau |r(\tau)|$) **do**
4:      $\hat{\tau}_p = \arg\max_\tau |r_p(\tau)|$
5:      $\hat{c}_p = r_p(\hat{\tau}_p)$
6:      $r_{p+1}(\tau) = r_p(\tau) - \hat{c}_p g_K\left(2\pi\Delta f(\tau - \hat{\tau}_p)\right)$
7:      $p \leftarrow p + 1$
8:   **end while**
9:   $\hat{\mathbf{S}} = \left[\mathbf{s}_K(2\pi\Delta f\hat{\tau}_1) \cdots \mathbf{s}_K(2\pi\Delta f\hat{\tau}_{N_p})\right]$
10: $\hat{\mathbf{c}} = \left(\hat{\mathbf{S}}'\hat{\mathbf{S}}\right)^{-1}\hat{\mathbf{S}}'\mathbf{x}$
11: **return** $\hat{\mathbf{H}} = \hat{\mathbf{S}}\hat{\mathbf{c}}$

The corresponding LS estimate is $\hat{\mathbf{c}} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{x}$. For lack of true $\mathbf{S}$, assuming that the delay estimates are accurate, we replace $\mathbf{S}$ by $\hat{\mathbf{S}}$. The channel coefficients are thus finally estimated as

$$\hat{\mathbf{c}} = (\hat{\mathbf{S}}'\hat{\mathbf{S}})^{-1}\hat{\mathbf{S}}'\mathbf{x}. \qquad (20)$$

Unlike the estimates (16), which are obtained sequentially (one after another), these estimates are obtained jointly, and hence offer a potential improvement. The formal steps of the PI algorithm are summarized in Algorithm 2.

The running time of the PI algorithm is dominated by the initialization step (line 2) which takes ILK floating point operations (flops) assuming that the observation interval τ_{obs} contains IL delay samples. Capitalizing on the fact that the signature function $g_K(\cdot)$ is known, the operations within the loop take only IL flops per iteration. The (optional) LS refinement step (line 10) takes KN_p^2 flops [22], [23]. The total complexity of the PI algorithm is thus $O(ILK) + O(N_pIL) + O(KN_p^2)$. Compared to $O(N_tILK) + O((5/2)KN_t^2)$ for the OMP algorithm, where N_t is the sparsity level of the channel tap vector, this

C. Multichannel Processing

In a conventional coherent receiver with M_r spatially distributed elements, one FFT demodulator is associated with each input channel. Using the model (3) then yields the M_r-element received signal vector

$$\mathbf{y}_k = d_k\mathbf{H}_k + \mathbf{z}_k, \qquad k = 0, \ldots, K-1 \qquad (21)$$

where $\mathbf{y}_k$ contains the demodulator outputs y_k^m, $m = 1, \ldots, M_r$, $\mathbf{H}_k$ and $\mathbf{z}_k$ contain the relevant channel and noise components, respectively. Using maximum ratio combining (MRC), the data symbols are then detected as

$$\hat{d}_k = \frac{\hat{\mathbf{H}}'_k\mathbf{y}_k}{\hat{\mathbf{H}}'_k\hat{\mathbf{H}}_k} \qquad (22)$$

where $\hat{\mathbf{H}}_k$ is an estimate of the channel coefficients $\mathbf{H}_k$. If there is no spatial coherence, the M_r channel estimates $\hat{H}_k^m$, $m = 1, \ldots, M_r$ are formed independently. However, if there exists a correlation between the M_r channels, this correlation can be exploited to reduce the receiver complexity.

In Fig. 3, we illustrate the PI algorithm operation on an OFDM block of experimental recordings with $K = 1024$ carrier modulated by QPSK data symbols. The signal for channel estimation, channel impulse response, and channel frequency response as well as the last block's scatter plot are shown.

IV. SPATIAL COHERENCE LEARNING

In this section, we propose a precombining method for multichannel OFDM systems. The proposed technique reduces the complexity of coherent OFDM receivers through linearly combining the M_r input channels into $Q \leq M_r$ output channels, so as to reduce the number of channel estimators needed. The technique makes no assumption about the spatial distribution of signals, and relies on differential encoding that keeps the receiver complexity at a minimum. Fig. 4 shows the block diagram of the receiver.
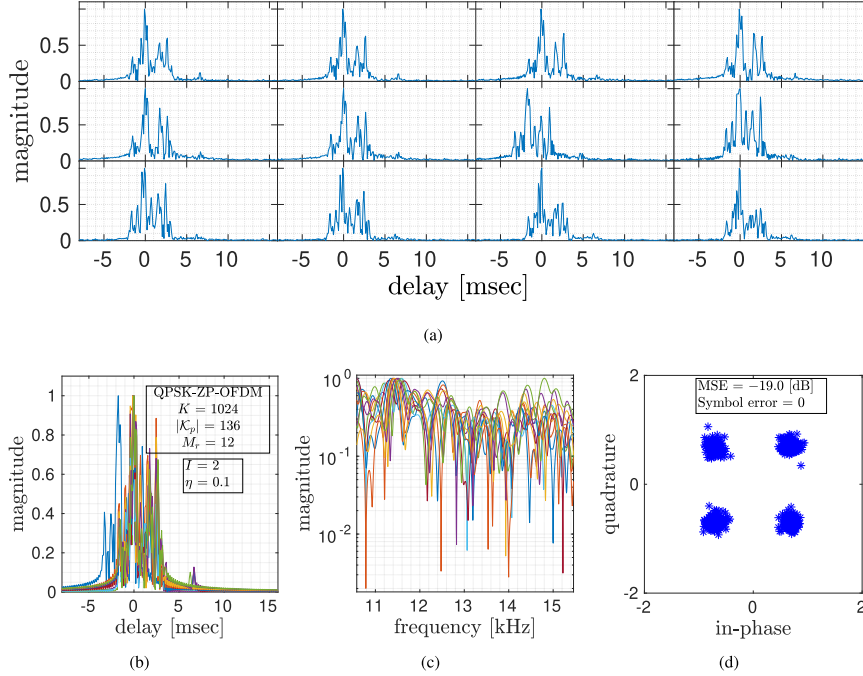
Fig. 3. Performance illustration for a ZP-OFDM block with $K = 1024$ carriers modulated by QPSK data symbols. Shown are the signals for channel estimation across the array with 12 receive elements spaced at 12 cm, estimates of channel impulse response and channel frequency response, and the scatter plot of detected data symbols. The elements are numbered left to right and top to bottom, and each signal is shown as a function of delay. Estimates are obtained by applying the PI algorithm terminated based on a predefined threshold $\eta = 0.1$ and the resolution factor $I = 2$. Channel estimation is based on 136 pilots. In the scatter plot, the data detection MSE is -19 dB and there are no symbols errors. (a) Signal for channel estimation in space and time. (b) Channel impulse response. (c) Channel frequency response. (d) Scatter plot.
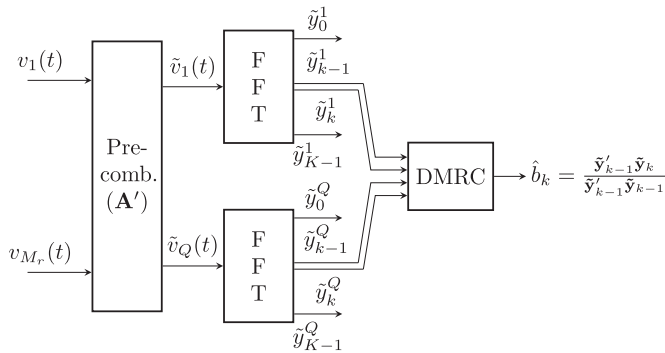

Fig. 4. Block diagram of the receiver with precombining. The precombiner $\mathbf{A}$ consisting of columns $\boldsymbol{\alpha}_q$, $q = 1, \ldots, Q$, linearly combines the M_r input channels into $Q \leq M_r$ output channels. FFT demodulation is then applied in each of the Q channels, followed by differential maximum ratio combining (DMRC). Alternatively, the order of precombining and FFT demodulation can be changed. FFT demodulation can be performed first on all M_r input channels, followed by precombining as indicated by the second equality in expression (23).

Let $\boldsymbol{\alpha}_q$ denote the precombiner weights for the qth output channel, and let the vector $\mathbf{v}(t)$ contain the M_r input signals $v_m(t)$. FFT demodulation then yields

$$\tilde{y}_k^q = \frac{1}{T} \int_T \boldsymbol{\alpha}_q' \mathbf{v}(t) e^{-2\pi i k \Delta f t} dt = \boldsymbol{\alpha}_q' \mathbf{y}_k, \quad q = 1, \ldots, Q \tag{23}$$

where the vector $\mathbf{y}_k$, $k = 0, \ldots, K - 1$, contains the M_r FFT outputs corresponding to the M_r input channels. Expression (23)

above shows that the receiver structure of Fig. 4, which employs pre-FFT combining, is equivalent to one that would employ post-FFT combining by applying the Q weights $\boldsymbol{\alpha}_q$ to the FFT outputs $\mathbf{y}_k$. We use this fact to develop a method for computing the weights recursively over the carriers.

Stacking the FFT outputs $\tilde{y}_k^q$ into a column vector $\tilde{\mathbf{y}}_k$, the differentially encoded data symbols $b_k = d_{k-1}^* d_k$ are estimated by DMRC over the Q channels as[6] [24]

$$\hat{b}_k = \frac{\sum_{q=1}^{Q} (\tilde{y}_{k-1}^q)^* \tilde{y}_k^q}{\sum_{q=1}^{Q} (\tilde{y}_{k-1}^q)^* \tilde{y}_{k-1}^q} = \frac{\tilde{\mathbf{y}}_{k-1}' \tilde{\mathbf{y}}_k}{\tilde{\mathbf{y}}_{k-1}' \tilde{\mathbf{y}}_{k-1}}. \tag{24}$$

Here, we implicitly assume that the channel frequency response changes slowly from one carrier to the next.

A. Linear Error Algorithm (L-Algorithm)

Differentially coherent detection (24) introduces a particular type of nonlinearity into the data estimation problem, namely a quadratic dependence on the precombiner weights. In [17], we proposed a recursive linear squares method by treating $\tilde{\mathbf{y}}_{k-1}$ as independent of the precombiner weights, and retaining only the dependence of $\tilde{\mathbf{y}}_k$ on the precombiner weights.[7] Such an approach leads to linear minimum mean-squared error (MMSE) estimation, and also allows for explicit amplitude normalization [25].

---

[6]Complex conjugate is denoted by $(\cdot)^*$.

[7]This assumption is based on the fact that the channel frequency response does not change much from one carrier to the next.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING

7

The estimated data symbol is now expressed as

$$\hat{b}_k = \frac{\sum_{q=1}^{Q} (\tilde{y}_{k-1}^q)^* \tilde{y}_k^q}{\sum_{q=1}^{Q} |\tilde{y}_{k-1}^q|^2} = \sum_{q=1}^{Q} (\bar{\tilde{y}}_{k-1}^q)^* \boldsymbol{\alpha}_q' \mathbf{y}_k \qquad (25)$$

where $\bar{\tilde{y}}_{k-1}^q = \tilde{y}_{k-1}^q / \sum_{q=1}^{Q} |\tilde{y}_{k-1}^q|^2$. The expression (25) can be written in a vector form as

$$\hat{b}_k = \underbrace{\begin{bmatrix} \boldsymbol{\alpha}_1' & \cdots & \boldsymbol{\alpha}_Q' \end{bmatrix}}_{\mathbf{a}' = (\text{vec}(\mathbf{A}))'} \underbrace{\begin{bmatrix} (\bar{\tilde{y}}_{k-1}^1)^* \mathbf{y}_k \\ \vdots \\ (\bar{\tilde{y}}_{k-1}^Q)^* \mathbf{y}_k \end{bmatrix}}_{\mathbf{u}_k = \bar{\tilde{\mathbf{y}}}_{k-1}^* \otimes \mathbf{y}_k} = \mathbf{a}' \mathbf{u}_k \qquad (26)$$

where the vec($\mathbf{A}$) operator creates a column vector $\mathbf{a}$ from the matrix $\mathbf{A}$ by stacking columns of $\mathbf{A}$ one below another, $\bar{\tilde{\mathbf{y}}}_k = \begin{bmatrix} \bar{\tilde{y}}_k^1 & \cdots & \bar{\tilde{y}}_k^Q \end{bmatrix}^T$, and $\otimes$ denotes the Kronecker product.

To arrive at the weights $\mathbf{a}$ without *a priori* knowledge of the channel, we use the error

$$e_k = b_k - \hat{b}_k = b_k - \mathbf{a}' \mathbf{u}_k \qquad (27)$$

to formulate the MMSE solution for the precombiner coefficients $\mathbf{a}$, which are computed recursively over the carriers as

$$\mathbf{a}(k) = \mathbf{a}(k-1) + \mathcal{A}(\mathbf{u}_k, e_k) \qquad (28)$$

where $\mathcal{A}(\mathbf{u}_k, e_k)$ represents a particular algorithm's increment computed for the input $\mathbf{u}_k$ and the error e_k. Since the error is a linear function of the precombining weights, the algorithm is termed linear error algorithm. For instance, if LMS is used, then $\mathcal{A}(\mathbf{u}_k, e_k) = \mu \mathbf{u}_k e_k^*$, where μ is the LMS step size, and if the recursive LS (RLS) algorithm is used, then $\mathcal{A}(\mathbf{u}_k, e_k)$ is the RLS increment given in Algorithm 3, line 12. Clearly, the choice of the adaptive algorithm always presents a tradeoff between computational complexity and speed of convergence. In the RLS algorithm, the stochastic gradient $e_k^* \mathbf{u}_k$ (line 12) is premultiplied by an estimate of the inverse of the covariance matrix $\mathbf{P}_{k-1}$, which has the effect of decorrelating the inputs to the adaptive filter. This decorrelation, along with the learning rate μ_k (line 11), enhances the convergence rate of the algorithm, thus requiring fewer pilots then in the LMS algorithm at the cost of more computations. $O((QM_r)^2)$ operations are required for each weight update of the RLS, whereas only $O(QM_r)$ are necessary with LMS.

B. Quadratic Error Algorithm (Q-Algorithm)

In this approach, we take into account the fact that $\bar{\tilde{\mathbf{y}}}_k$ is also dependent upon the precombining weight. Assuming that the precombining weights remain unchanged from one carrier to the next, the estimated data symbols can be expressed as

$$\hat{b}_k = \frac{\sum_{q=1}^{Q} (\tilde{y}_{k-1}^q)^* \tilde{y}_k^q}{\sum_{q=1}^{Q} (\tilde{y}_{k-1}^q)^* \tilde{y}_{k-1}^q} = \frac{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{y}_k \mathbf{y}_{k-1}' \boldsymbol{\alpha}_q}{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{y}_{k-1} \mathbf{y}_{k-1}' \boldsymbol{\alpha}_q}$$

$$= \frac{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{Y}_k \boldsymbol{\alpha}_q}{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{Y}_{k-1} \boldsymbol{\alpha}_q}. \qquad (29)$$

Algorithm 3: L-Algorithm.

Input: Number of carriers K, number of blocks N_b, post-FFT observations $\mathbf{y}_k(n), \forall k = 0, \ldots, K-1, n = 1, \ldots, N_b$, pilot set $\mathcal{K}_p(n)$, forgetting factor λ, initial variance σ^2

Initialization: Initialize $\mathbf{A}_0(1)$ to the Q maximally spaced receiving elements, initialize the inverse covariance matrix $\mathbf{P}_0(1) = \sigma^{-2} \mathbf{I}_{QM_r}$ where $\mathbf{I}_{QM_r}$ is a $QM_r \times QM_r$ identity matrix.

1: $\mathbf{x}_0(1) = \mathbf{A}_0'(1) \mathbf{y}_0(1)$
2: $\mathcal{C}_1 = \{1, 2, \ldots, K-1\}$,
    $\mathcal{C}_{-1} = \{K-2, K-3, \ldots, 0\}$
3: **for** $n = 1, \ldots, N_b$ **do**
4:     $d = (-1)^{n-1}$ ($d = \pm 1$ depending on the direction of iteration [See Fig. 5.])
5:     **for** $k \in \mathcal{C}_d$ **do**
6:        $\mathbf{u}_k(n) = \frac{\mathbf{x}_{k-d}^*(n)}{\|\mathbf{x}_{k-d}(n)\|^2} \otimes \mathbf{y}_k(n)$
7:        $\mathbf{a}_{k-d}(n) = \text{vec}(\mathbf{A}_{k-d}(n))$
8:        $\hat{b}_k(n) = \mathbf{a}_{k-d}'(n) \mathbf{u}_k(n)$
9:        $\tilde{b}_k(n) = \begin{cases} b_k(n) & k \in \mathcal{K}_p(n) \\ \text{decision}(\hat{b}_k(n)) & \text{otherwise} \end{cases}$
10:       $e_k(n) = \tilde{b}_k(n) - \hat{b}_k(n)$
11:       $\mu_k(n) = [\lambda + \mathbf{u}_k'(n) \mathbf{P}_{k-d}(n) \mathbf{u}_k(n)]^{-1}$
12:       $\mathbf{a}_k(n) =$
          $\mathbf{a}_{k-d}(n) + \underbrace{\mu_k(n) \mathbf{P}_{k-d}(n) \mathbf{u}_k(n)}_{\mathbf{G}_k(n)} e_k^*(n)$
13:       $\mathbf{P}_k(n) =$
          $\lambda^{-1} [\mathbf{I}_{QM_r} - \mathbf{G}_k(n) \mathbf{u}_k'(n)] \mathbf{P}_{k-d}(n)$
14:       $\mathbf{x}_k(n) = \mathbf{A}_k'(n) \mathbf{y}_k(n)$
15:    **end for**
16:    $\mathbf{A}_k(n+1) = \mathbf{A}_k(n)$
17:    $\mathbf{x}_k(n+1) = \mathbf{A}_k'(n) \mathbf{y}_k(n)$
18:    $\mathbf{P}_k(n+1) = \sigma^{-2} \mathbf{I}_{QM_r}$
19: **end for**

The error that is used to determine the precombining weight $\boldsymbol{\alpha}_q$ according to the MMSE criterion is then given by

$$e_k = b_k - \frac{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{Y}_k \boldsymbol{\alpha}_q}{\sum_{q=1}^{Q} \boldsymbol{\alpha}_q' \mathbf{Y}_{k-1} \boldsymbol{\alpha}_q} = b_k - \frac{N_k}{D_k}. \qquad (30)$$

Differentiating the MSE $E = \mathbb{E}\{|e_k|^2\}$ with respect to the precombining vector $\boldsymbol{\alpha}_q$ gives the error gradient

$$\frac{\partial E}{\partial \boldsymbol{\alpha}_q'}$$

$$= \mathbb{E}\left\{ -\frac{e_k^*}{D_k} \left[\mathbf{Y}_k - \hat{b}_k \mathbf{Y}_{k-1} \right] \boldsymbol{\alpha}_q - \frac{e_k}{D_k^*} \left[\mathbf{Y}_k' - \hat{b}_k^* \mathbf{Y}_{k-1}' \right] \boldsymbol{\alpha}_q \right\}$$

$$= \mathbb{E}\left\{ -2\Re \left\{ \frac{e_k^*}{D_k} \left[\mathbf{Y}_k - \hat{b}_k \mathbf{Y}_{k-1} \right] \right\} \boldsymbol{\alpha}_q \right\} = \mathbb{E}\{\boldsymbol{\gamma}_q(k)\}. \qquad (31)$$

Because the error (30) is a quadratic, rather than a linear function of the precombining weights, there is no apparent closed
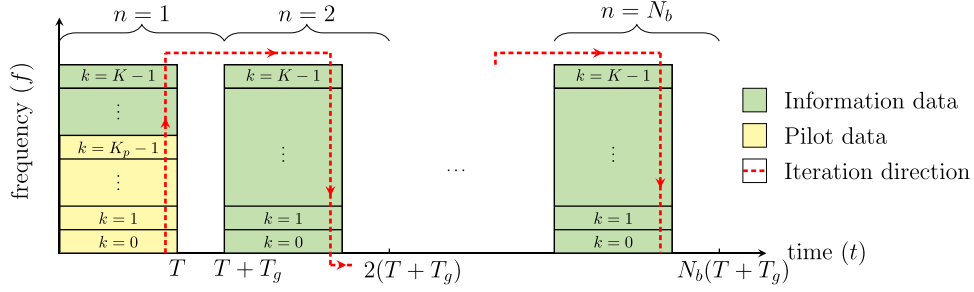
Fig. 5.   Progression of the algorithm. Shown is a frame containing N_b OFDM blocks. The algorithm forms the error using the pilot data symbols $b_k, k \in \mathcal{K}_p = \{0, \ldots, K_p - 1\}$ in the training mode, shown in yellow (lighter). Thereafter, it switches into decision-directed mode where the decisions are made on the composite estimate (24). The red-dashed line shows the direction of the algorithm's iteration. In the first block, the iteration goes from the lowest carrier to the highest carrier ($d = 1, k \in \mathcal{C}_{d=1}$). At the end of the first block, the obtained precombining weights for the highest carrier are used to initialize the precombining weights for the highest carrier of the second block, in which the iteration goes from the highest to the lowest carrier ($d = -1, k \in \mathcal{C}_{d=-1}$).

Algorithm 4: Q-Algorithm.

Input: Number of carriers K, number of blocks N_b,
post-FFT observations
$\mathbf{y}_k(n), \ \forall k = 0, \ldots, K - 1, \ n = 1, \ldots, N_b$, pilot set
$\mathcal{K}_p(n)$, step size $\mu_q, \ \forall q = 1, \ldots, Q$
Initialization: Initialize $\mathbf{A}_0(1)$ and $\mathbf{A}_1(1)$ to the Q
maximally spaced receiving elements
1: $\tilde{\mathbf{y}}_0(1) = \mathbf{A}_0'(1)\mathbf{y}_0(1)$
2: $\mathcal{C}_1 = \{1, 2, \ldots, K - 1\}$,
    $\mathcal{C}_{-1} = \{K - 2, K - 3, \ldots, 0\}$
3: **for** $n = 1, \ldots, N_b$ **do**
4:      $d = (-1)^{n-1}$ ($d = \pm 1$ depending on the direction
        of iteration [See Fig. 5.])
5:      **for** $k \in \mathcal{C}_d$ **do**
6:          $\tilde{\mathbf{y}}_k(n) = \mathbf{A}_k'(n)\mathbf{y}_k(n)$
7:          $N_k(n) = \tilde{\mathbf{y}}_{k-d}'(n)\tilde{\mathbf{y}}_k(n)$
8:          $D_k(n) = \tilde{\mathbf{y}}_{k-d}'(n)\tilde{\mathbf{y}}_{k-d}(n)$
9:          $\hat{b}_k(n) = \frac{N_k(n)}{D_k(n)}$
10:         $\tilde{b}_k(n) = \begin{cases} b_k(n) & k \in \mathcal{K}_p(n) \\ \text{decision}(\hat{b}_k(n)) & \text{otherwise} \end{cases}$
11:         $e_k(n) = \tilde{b}_k(n) - \hat{b}_k(n)$
12:         $\mathbf{Y}_k(n) = \mathbf{y}_k(n)\mathbf{y}_{k-d}'(n)$
13:         $\mathbf{Y}_{k-d}(n) = \mathbf{y}_{k-d}(n)\mathbf{y}_{k-d}'(n)$
14:         **for** $q = 1, \ldots, Q$ **do**
15:             $\boldsymbol{\gamma}_q(k; n) = -2\Re\{\frac{1}{D_k(n)} [\mathbf{Y}_k(n) - \hat{b}_k(n)\mathbf{Y}_{k-d}(n)]e_k^*(n)\}\boldsymbol{\alpha}_q(k; n)$
16:             $\boldsymbol{\alpha}_q(k + 1; n) = \boldsymbol{\alpha}_q(k; n) + \mu_q\boldsymbol{\gamma}_q(k; n)$
17:         **end for**
18:     **end for**
19:     $\mathbf{A}_k(n + 1) = \mathbf{A}_k(n)$
20:     $\mathbf{A}_{k-d}(n + 1) = \mathbf{A}_k(n)$
21:     $\tilde{\mathbf{y}}_k(n + 1) = \mathbf{A}_k'(n)\mathbf{y}_k(n)$
22: **end for**

---

form solution for the vector which sets the gradient to zero. Nonetheless, the solution can be obtained numerically, using the stochastic gradient approach. Specifically, the precombining weights are obtained recursively over the carriers as

$$\boldsymbol{\alpha}_q(k + 1) = \boldsymbol{\alpha}_q(k) + \mu_q\boldsymbol{\gamma}_q(k), \quad \forall q = 1, \ldots, Q \qquad (32)$$

where μ_q is the step size. Algorithm 4 summarizes the Q-Algorithm.

The L/Q algorithms form the error using the pilot data symbols $b_k, k \in \mathcal{K}_p = \{0, \ldots, K_p - 1\}$ in the training mode. Thereafter, they switch into decision-directed mode where the decisions are made on the composite estimate (24). As shown in Fig. 5, the process can continue into the next block, where recursion will evolve in reverse order (from the highest carrier to the lowest) and require fewer (or no) pilots [4].

The initial value $\mathbf{A}_0 = \begin{bmatrix} \boldsymbol{\alpha}_1(0) & \cdots & \boldsymbol{\alpha}_Q(0) \end{bmatrix}$ needs to be chosen carefully. A possible choice corresponds to selecting Q out of M_r channels and passing them through the precombiner intact. For example, the initial channels can be chosen as equally spaced, starting with the first channel. It is important to set the initial conditions such that the algorithm is allowed enough freedom to form the best Q outputs that have as little correlation as possible.

Recursive computation of the combiner coefficients across carriers requires M_r FFTs but allows the precombiner to change from one carrier to another, thus effectively accomplishing broadband processing. If the channels are slowly varying, where the precombiner computed in one block could be preserved for use in the next block, precombining could be implemented before FFT demodulation, thus simultaneously reducing the total number of channel estimators as well as the total number of FFT operations. If the spatial coherence does not change significantly from one block to the next within an OFDM frame, the precombining weights calculated in the first block can be used for linearly combining the input channels in the rest of the blocks in the frame; thus, further reduction in the receiver complexity can be accomplished.

Relying on differential encoding allows the L/Q algorithms not only to be used as stand-alone, but also to be used in a coherent receiver. In a coherent receiver, the precombined FFT outputs $\tilde{y}_k^q, \ k = 0, \ldots, K - 1$ given by

$$\tilde{y}_k^q = \boldsymbol{\alpha}_q'(k)\mathbf{y}_k = d_k \underbrace{\boldsymbol{\alpha}_q'(k)\mathbf{H}}_{\tilde{H}_k^q} + \tilde{z}_k^q, \quad q = 1, \ldots, Q \qquad (33)$$

are fed into the PI channel estimator, described in Section III, to form the $Q \leq M_r$ estimates $\hat{\tilde{H}}_k^q$, of the precombined channel

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING
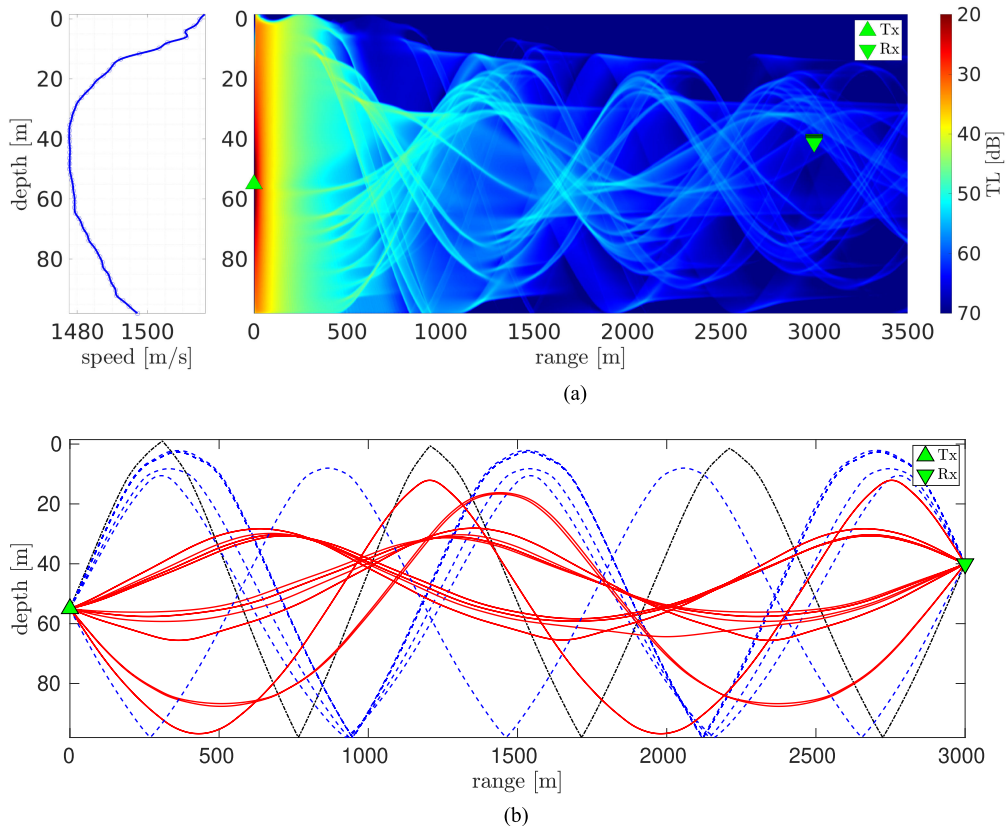
9

Fig. 6. (a) Sound-speed profile measured during the MACE'10 experiment (left). Gaussian beam incoherent acoustic transmission loss (TL) of a 13-kHz signal emitted at a depth of 55 m (right). (b) Geometric eigenrays for the MACE'10 sound-speed profile with the source at 55 m and the receiver at 40-m depth. In (b), the solid red curves show the "direct paths" while dotted black curves and dashed blue curves illustrate rays that hit both top and bottom boundaries, and the bottom boundary only, respectively. The amplitude and length of the swelling surface waves are 1.5 and 350 m, respectively. The seabed boundary is modeled as an acoustic-elastic half-space having a sound speed of 1400 m/s (soft seabed) and density of 1.8 g/m^3. This plot provides an insight into the propagation paths over the course of the MACE'10 experiment.

coefficients $\tilde{H}_k^q$. Using a Q-channel configuration instead of the full-complexity configuration reduces the total number of channel estimators by M_r/Q. This feature brings a significant reduction in complexity when the value of Q at which saturation is reached is low. Using MRC, the data symbol estimates are then obtained as

$$\hat{d}_k = \frac{\hat{\tilde{\mathbf{H}}}_k' \tilde{\mathbf{y}}_k}{\hat{\tilde{\mathbf{H}}}_k' \hat{\tilde{\mathbf{H}}}_k} \qquad (34)$$

where $\hat{\tilde{\mathbf{H}}} = \left[\hat{\tilde{H}}_k^1 \cdots \hat{\tilde{H}}_k^Q \right]^T$.

V. SIMULATION RESULTS

In this section, we assess the average performance of the PI algorithm through simulation and compare it to the conventional channel estimation methods. We use the Bellhop ray-tracing program [26] to obtain the nominal state of a UWA channel for a given geometry and signal frequency. The Bellhop program provides an accurate deterministic channel impulse response over a 2-D grid in space, which covers the region of the receiver. However, it does not take into account random channel variation. To consider the effects of inevitable random fluctuations of the

TABLE I
SIMULATION PARAMETERS

water depth [m]	100	bandwidth B [kHz]	5
distance between tx and rx [km]	3	lowest carrier frequency f_0 [kHz]	10.5
transmitter depth [m]	55	number of carriers K	1024
receiver depth [m]	40	carrier spacing Δf [Hz]	4.9

The guard interval is $T_g = 16$ ms. The number of pilots is $|\mathcal{K}_p| = 128$.

environment, as well as changes in the receiver position, we feed the nominal channel impulse response obtained from the Bellhop program into the statistical channel simulator based on the model proposed in [27].

We choose simulation geometries similar to those of MACE'10, which are summarized in Table I (see [17] for further details). Specifically, we use the recorded sound-speed profile shown on the left of Fig. 6(a), and a moored vertical receive antenna array with four equally spaced receiving elements that are spaced $d_{rx} = 12$ cm apart and submerged at 40-m depth from the surface. The height of the swelling wave is 3 m peak-to-peak and the receiver is 3 km away from the transmitter. Fig. 6 also

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10   IEEE JOURNAL OF OCEANIC ENGINEERING

(a)

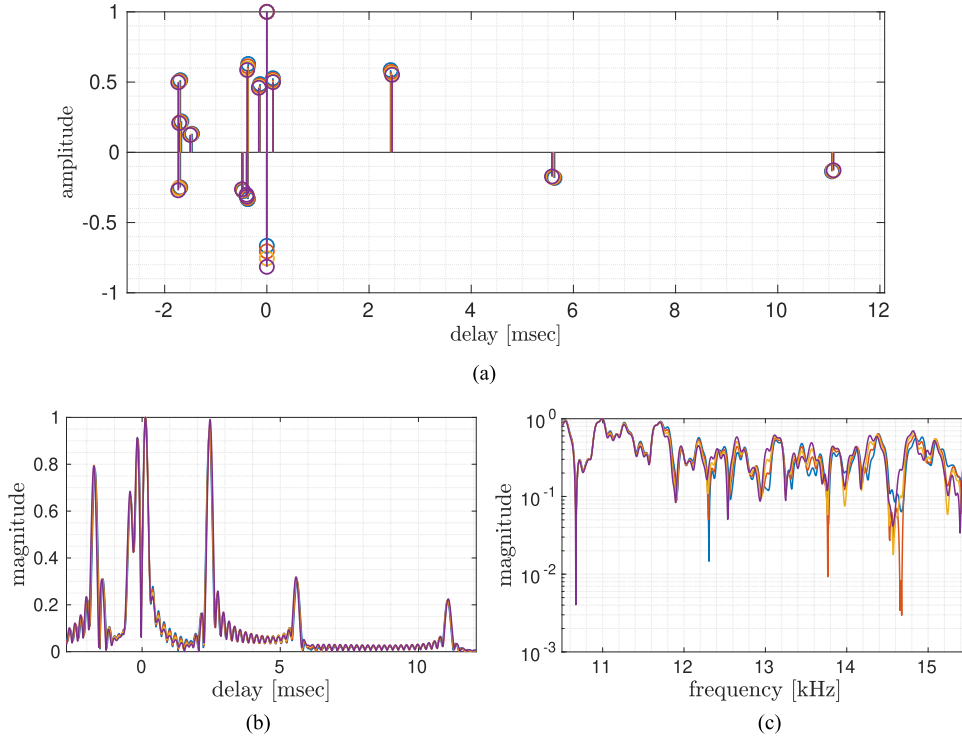(b)   (c)

Fig. 7.   (a) Bellhop nominal multipath structure for the channel with geometry given in Table I. (b) Impulse response and (c) transfer function for the nominal channel geometry corresponding to the multipath structures shown in (a), as seen in $B = 5$ kHz of bandwidth centered at $f_c = 13$ kHz. In (a), those paths whose absolute value is above one tenth of the strongest paths are shown. As shown in (b), there are multiple arrivals of similar amplitude that span approximately 12 ms.

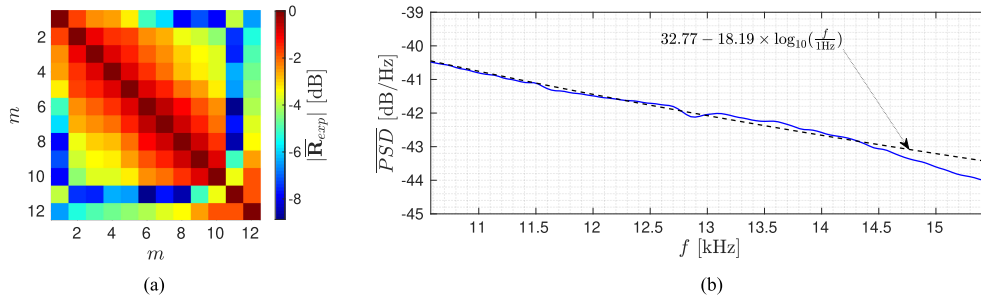(a)   (b)

Fig. 8.   (a) Average correlation coefficient between noise from receiver elements. (b) Noise power spectral density (PSD) averaged over all measurements and receive antenna elements. The result in (a) illustrates significant correlation between adjacent elements while the element spacing $d_{rx} = 12$ cm is on the order of the wavelength $\lambda_c = 11.5$ cm at the operating frequency $f_c = 13$ kHz. As shown by the black-dashed line in (b), the noise PSD decays at a rate of approximately 18 dB/decade [21]. This plot also suggests a highly colored noise spectrum (more than 4-dB change over the frequency band).

shows the acoustic TL at 13 kHz, as well as the eigenrays connecting the source to the receiver. Clearly, it is difficult to identify a single ray as the direct path as shown in Fig. 6(b) [28].

Fig. 7(a) shows the nominal multipath structure seen by the four receive antenna elements. In an acoustic communication system, the channel can be observed only in a limited band of frequencies. As a result, the observable response will be smeared, as illustrated in Fig. 7(b). It is important to make the distinction between the propagation paths [see Fig. 7(a)], and the samples of the observable response [see Fig. 7(b)], which are the so-called channel taps. To simulate the UWA channel, we utilize the multipath structure shown in Fig. 7(a) as an input to the statistical channel simulator [29] that generates independent complex Gaussian channel path amplitudes.

In the simulation, one OFDM block with $K = 1024$ carriers modulated by data symbols from three constellations, namely QPSK, 8-PSK, and 16-quadrature amplitude modulation (16-QAM),[8] is transmitted through the synthesized UWA channel with unit power and is corrupted by an additive Gaussian noise generated based on the statistical characteristics of the recorded noise signal during the MACE'10 experiment (see Fig. 8).

Fig. 9(a) illustrates the performances of the channel estimation methods in terms of data detection MSE and symbol error rate (SER) as a function of the SNR at the input to the receiver ranging

---

[8]In developing (9), we consider PSK without loss of generality, but the treatment applies to QAM as well (any linear modulation).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING 11
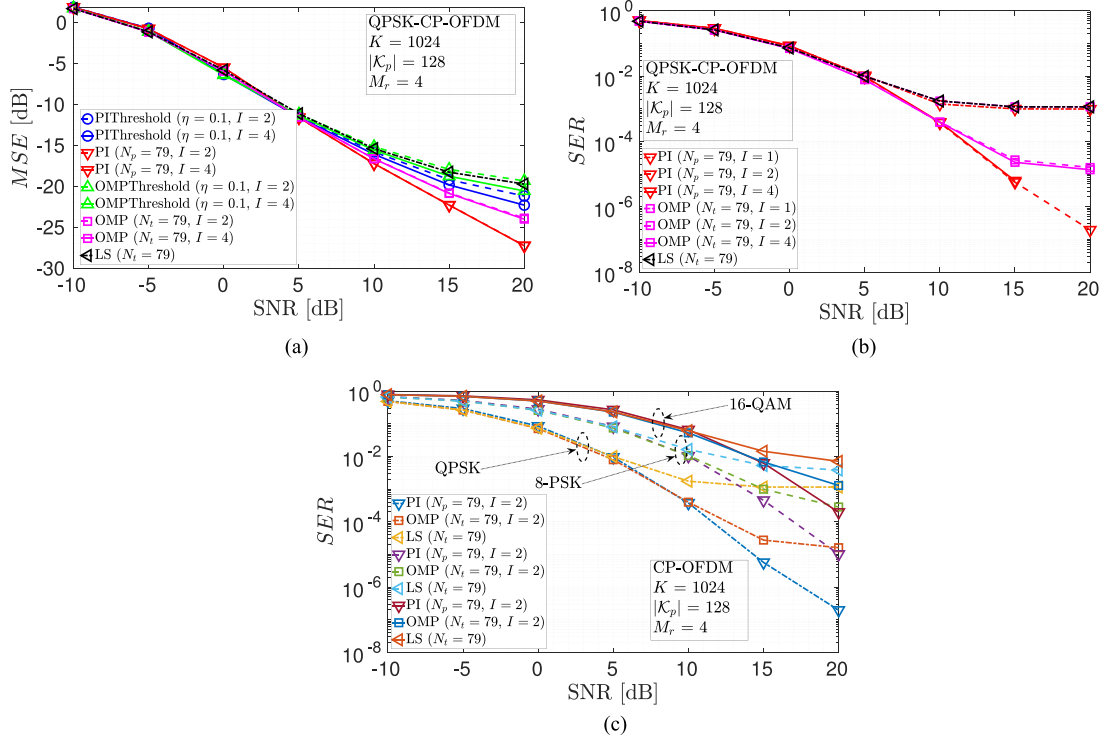

Fig. 9. Performance of the PI channel estimation and the conventional methods for 315 000 realizations of the MACE'10-like simulated UWA channel and noise. (a) and (b) Average MSE and SER, respectively, as a function of input SNR for CP-OFDM blocks containing QPSK symbols. (c) Average SER versus input SNR for OFDM blocks containing QPSK, 8-PSK, and 16-QAM symbols. The resolution factors I used in the PI and OMP algorithms are 1, 2, and 4. For the case with $I = 4$ and $\eta = 0.1$, the PI and OMP algorithms return about 7 delay-domain coefficients. Using the PI algorithm, we achieve uncoded SER as low as 2×10^{-7} at the SNR of 20 dB for QPSK symbols. Four receivers are used to perform MRC (22).

from -10 to 20 dB. The MSE is measured as follows:

$$\text{MSE} = \frac{1}{N_r} \sum_{i=1}^{N_r} \frac{1}{K} \sum_{k=0}^{K-1} |d_k^i - \hat{d}_k^i|^2 \tag{35}$$

where $\hat{d}_k^i$ is the estimate of the kth data symbol d_k^i in the ith realization and N_r is the number of channel and noise realizations. For the OMP and PI algorithms, data detection MSE performances are shown for two stopping criteria: one is based on a prespecified threshold $\eta = 0.1$ and the other on a prespecified sparsity level, which is denoted by N_p (number of paths) and N_t (number of taps) for the PI and OMP algorithms, respectively. Each point in Fig. 9 is obtained by averaging over $N_r = 5000$ realizations of noise and channel.

Fig. 9 clearly shows that the PI algorithm, which terminates based on a predefined sparsity level, outperforms the LS and OMP algorithms by 3 and 7 dB, respectively, when the input SNR is 20 dB and the resolution factor is $I = 4$. Fig. 9(a) and (b) also shows that increasing the resolution factor I from 1 to 2 improves the performance of both the PI and OMP algorithms. Further increase from 2 to 4 does not bring as much improvement. In Fig. 9(c), we demonstrate the performance of the system used for three linear modulations schemes, namely QPSK, 8-PSK, and 16-QAM, in terms of the average SER. Using the PI algorithm with $I = 2$, we obtain a 16-QAM CP-ODFM system with data rate and bandwidth efficiency as high as 16 kb/s and 3.2 b/s/Hz, respectively, while achieving SER as low as

2×10^{-4}. Note that these results correspond to an uncoded system, while channel coding will further reduce the BER. We will quantify the effect of coding in Section VI.

VI. EXPERIMENTAL RESULT

To assess the system performance, we focus on the experimental data from the Mobile Acoustic Communication Experiment (MACE'10) which took place off the coast of Martha's Vineyard, MA, USA, in June 2010. The experimental signals, whose parameters are given in Table II, were transmitted using the acoustic frequency range between 10.5 and 15.5 kHz. The receiver array of 12 equally-spaced elements spanning a total linear aperture of 1.32 m was deployed at the depth of 40 m, and the transmitter was towed at the depth of 40–60 m. The water depth was approximately 100 m, and the transmission distance varied between 3 and 7 km. More details about the experiment can be found in [18].

The experiment consisted of multiple repeated transmissions, each containing all the OFDM signals listed in Table II. There was a total of 210 transmissions spanning 3.5 h of recording. During this time, the transmitting station moved away and toward the receiving station, at varying speeds ranging from 0.5 to 1.5 m/s. The results provided in this section are obtained from all 210 transmissions included the three different configurations; namely, QPSK ZP-OFDM, QPSK CP-OFDM, and 8-PSK ZP-OFDM blocks.

TABLE II
MACE'10 SIGNAL PARAMETERS

number of carriers K	256	512	1024	2048
number of blocks per frame N_b	32	16	8	4
carrier spacing Δf [Hz]	19.5	9.8	4.9	2.4
bit rate [kbps]	11.4	13	13.9	14.4
bandwidth efficiency [bps/Hz]	1.1	1.9	2.4	2.7

The guard interval is $T_g = 16$ ms. The total bandwidth is $B = 5$ kHz and the lowest carrier frequency is $f_0 = 10.5$ kHz. The uncoded bit rate and bandwidth efficiency are calculated for 8-PSK modulation. The values of the bit rate and bandwidth efficiency for QPSK modulation are 2/3 of the corresponding 8-PSK values. The bandwidth efficiency is obtained assuming 136 pilots per block.

We demonstrate the performance of the proposed PI algorithm for channel estimation and that of the Q-Algorithm for learning spatial coherence between elements of the receiver array in terms of data detection MSE and average execution time $\bar{T}_{\text{exe}}$ which is deemed a practical indicator of the algorithm complexity. We also report on the estimated cumulative density function (CDF) of the MSE measured in each signal frame. Furthermore, we show the bit error rate (BER) and block error rate (BLER) of the system when low-density parity check (LDPC) codes are used with various code rates.

We compare the performance of the PI algorithm with that of the conventional LS and OMP algorithms and show that the PI algorithm outperforms the OMP in both aforementioned criteria. Furthermore, we make a comparison between the Q-Algorithm proposed in this paper and the L-Algorithm proposed in [17], showing that the Q-Algorithm outperforms the L-Algorithm in terms of average execution time at the cost of negligible degradation in data detection MSE and BLER, thus achieving further reduction in complexity.

The MSE corresponding to an OFDM signal with K carriers is measured in the nth block of the ith frame as

$$\text{MSE}^i(n, K) = \frac{1}{K} \sum_{k=1}^{K} |d_k^i(n) - \hat{d}_k^i(n)|^2 \qquad (36)$$

and the MSE per frame is obtained as

$$\text{MSE}^i(K) = \frac{1}{N_b} \sum_{n=1}^{N_b} \text{MSE}^i(n, K). \qquad (37)$$

The average over all N_f frames is

$$\overline{\text{MSE}}(K) = \frac{1}{N_f} \sum_{i=1}^{N_f} \text{MSE}^i(K). \qquad (38)$$

Note that due to the random channel variation and a finite number of measurements, each of these quantities is a random variable.

Fig. 10 illustrates the average MSE and the average execution time of the LS, OMP, and PI algorithms as a function of the number of carriers K (log scale). This result clearly shows that the PI algorithm terminated based on a predefined sparsity level outperforms the LS and OMP algorithms by 2 and 1 dB, respectively,

when $\log_2(K) = 8, 9$, and 10. When $\log_2(K) = 11$, both the PI and OMP experience a deterioration in performance, which can be explained by the increased block duration that nudges the temporal coherence of the channel. Fig. 10(b) compares the running time of the PI algorithm with the LS, OMP algorithm and shows that the PI algorithm has lower complexity than the OMP method. It thus enables operation with a greater number of carriers, effectively increasing the bandwidth efficiency at a lower computational complexity. Lower computational complexity and better MSE performance make the PI algorithm a good practical candidate for channel estimation in acoustic OFDM systems. Fig. 10(c) provides the MSE performance for the ZP-OFDM frames containing 8-PSK symbols. Fig. 10(d) illustrates data detection MSE for frames containing ZP-OFDM and CP-OFDM blocks. As expected, CP-OFDM exhibits slightly better performance at the cost of more transmission energy.

Fig. 11 illustrates the estimated CDF of the MSE per block for the three different scenarios (QPSK ZP-OFDM, QPSK CP-OFDM, and 8-PSK ZP-OFDM). This result refers to $K = 1024$ carries and includes all the frames, transmitted over 3.5 h. Systems equipped with the threshold-based PI and OMP algorithms deliver MSE below -14 dB for 87% and 75% of the OFDM blocks, respectively, for the ZP-OFDM blocks conveying QPSK symbols. The same performance is observed for the QPSK CP-OFDM and 8-PSK ZP-OFDM configurations.

Fig. 12 illustrates the MSE performance as a function of the number of receiving elements M_r, which are chosen maximally and equally spaced among the 12 available elements. Evidently, a significant improvement is observed as the number of elements increases and spatial diversity gain is extracted. Although the best performance (-16 dB of MSE) is achieved by using all the 12 elements, using 6 elements also provides an excellent performance (-15 dB of MSE). Increasing the number of elements exhibits the effect of diminishing returns as the total array aperture remains the same. In Fig. 14, we show that using the precombining the same performance as the full complexity receiver can be achieved by combining the 12 available receivers into to 4 output channels.

In Fig. 13, we demonstrate the performance of the system used for the three configurations of signals in terms of average BER and average BLER using regular LDPC codes with various code rates ranging from 0.1 to 1. The codeword length is $N = 2K$ for QPSK or $N = 3K$ for 8-PSK, respectively; thus, each codeword constitutes an OFDM block. The column weight of the $M \times N$ parity check matrix, where M is the number of parity bits, is $w_c = 3$ for all the code rates considered, and the row weight $w_r = w_c M/N$ varies from 3.3 to 30 corresponding to code rates from 0.1 to 0.9 [30]. We use soft decision decoding that takes the likelihood ratio for each code-bit as an input [31]. Decoding is performed based on the probability propagation algorithm which can be seen as an instance of the sum–product algorithm [32]. Employing the PI algorithm for channel estimation enables LDPC to work to its full potential. Using the PI algorithm and code rate as high as 1, we achieve BER and BLER as low as 2×10^{-4} and 1.5×10^{-1}, respectively, for ZP-OFDM frames containing QPSK symbols. Code rates below 1 result in low

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING   13
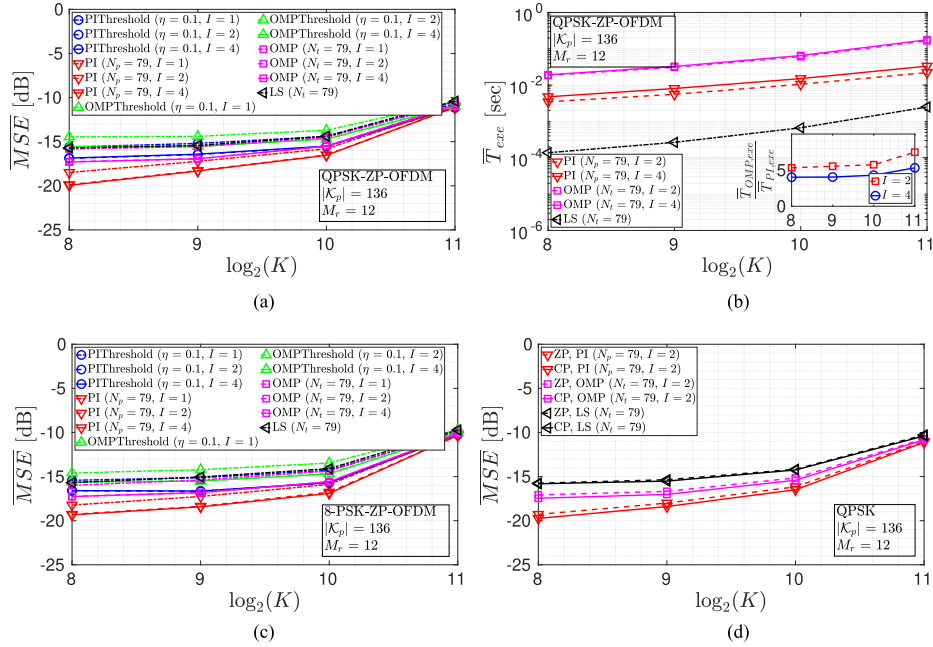

Fig. 10. (a) Average MSE versus the number of carriers K for ZP-OFDM blocks containing QPSK symbols. (b) Average execution time as a function of the number of carriers (log scale). (c) Average MSE as a function of the number of carriers for ZP-OFDM blocks containing 8-PSK symbols. (d) Average MSE versus the number of carriers for ZP-OFDM and CP-OFDM. The results in (a) and (b) are obtained by averaging over 104 transmissions of ZP-OFDM frames containing QPSK symbols. Each point in (c) and (d) is obtained from 52 transmissions made over 3.5 h of QPSK CP-OFDM and 8-PSK ZP-OFDM signals. In (a), (c), and (d), the performance of the systems equipped with the PI algorithm is compared to that of the systems with LS and OMP algorithms. For both the PI and OMP algorithms, two stopping criteria have been chosen, one based on the sparsity level (for PI $N_p = 79$ and for OMP $N_t = 79$) and another based on the threshold $\eta = 0.1$. The resolution factors I used in the PI and OMP algorithms are 1, 2, and 4. For the case with $K = 1024$, $I = 4$, and $\eta = 0.1$, the PI and OMP algorithms return about 23 delay-domain coefficients. The inset in (b) shows the ratio of the execution time of OMP algorithm to that of PI algorithm as a function of the number of carriers for $I = 2, 4$. The results shown in (b) rely on the FFT implementation of the step 4 in the OMP algorithm (Algorithm 1). All 12 receivers are used to perform MRC (22).
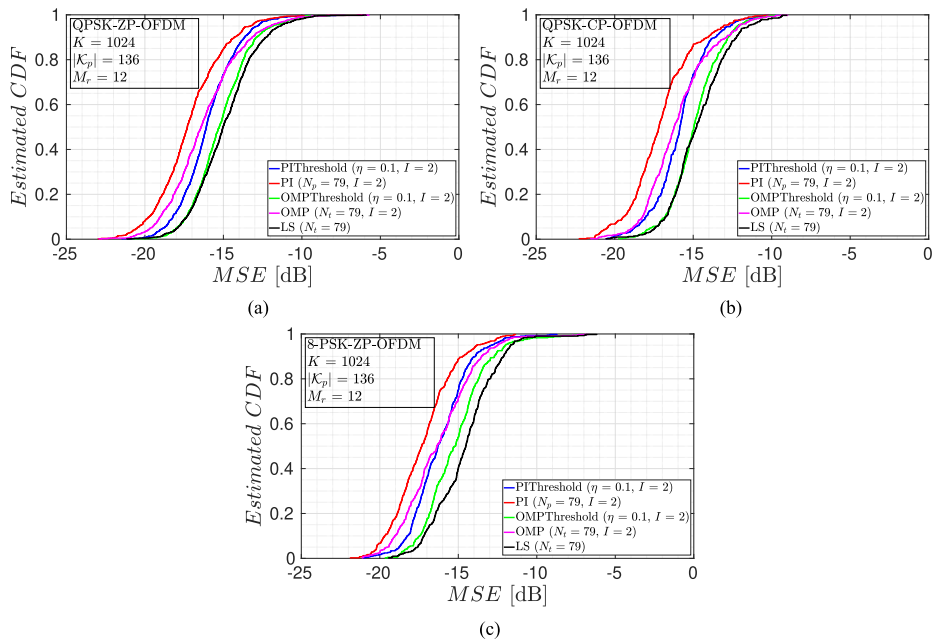

Fig. 11. Estimated CDF of the MSE for the LS, OMP, and PI algorithms applied to the three configurations of transmitted signals, (a) QPSK ZP-OFD, (b) QPSK CP-OFDM, and (c) 8-PSK ZP-OFDM. The CDF in (a) reflect all 104 transmissions with $K = 1024$ carriers during MACE'10. The CDFs in (b) and (c) are obtained from 52 transmissions of QPSK CP-OFDM and 8-PSK ZP-OFDM signals with the same number of carriers. The resolution factor for the PI and OMP algorithms is $I = 2$. The conventional LS technique delivers MSE below -14 dB for only 68% of OFDM blocks. The OMP algorithm with $I = 2$ improves the performance by delivering MSE below -14 dB for 83% of the blocks. The PI algorithm outperforms both by delivering MSE below -14 dB for 93% of the blocks at lower complexity compared to the OMP algorithm.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

14   IEEE JOURNAL OF OCEANIC ENGINEERING
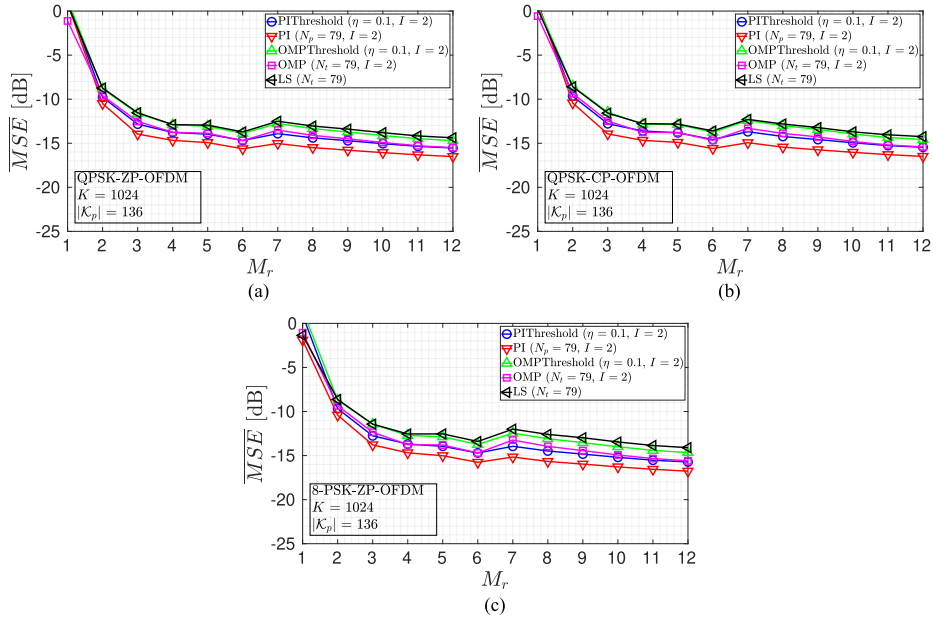

Fig. 12. Average MSE versus the number of receiving elements M_r. The number of carriers and pilots are 1024 and 136, respectively. The PI algorithm outperforms the LS and OMP algorithms for all the OFDM configurations considered. The M_r elements are chosen out of the 12 available elements as maximally space. The spacing between the elements is thus equal for $M_r = 2, \ldots, 6$ and 12, but not otherwise (hence the visible discontinuity at $M_r = 7$).
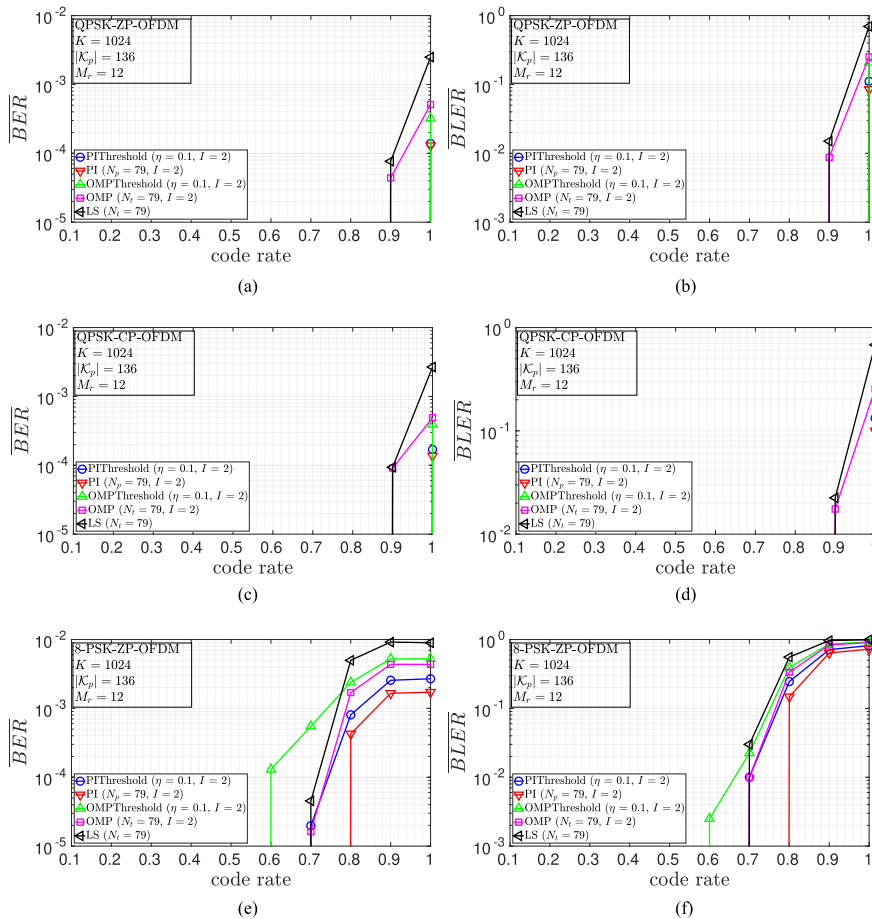

Fig. 13. Average BER versus the rate of the LDPC code (left). Average BLER as a function of rate of the LDPC code (right). The results in (a) and (b) reflect all 104 transmissions with 1024 carriers during MACE'10. Using code rates as high as 1, the PI channel estimation enables excellent performance with BER $= 1.5 \times 10^{-4}$ and BLER $= 1 \times 10^{-1}$ for CP-OFDM blocks. Using code rates as high as 0.8, the PI algorithm can also achieve BER and BLER as low as 4.5×10^{-4} and 1.5×10^{-1} for OFDM blocks whose carriers are 8-PSK modulated.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TADAYON AND STOJANOVIC: ITERATIVE SPARSE CHANNEL ESTIMATION AND SPATIAL CORRELATION LEARNING   15
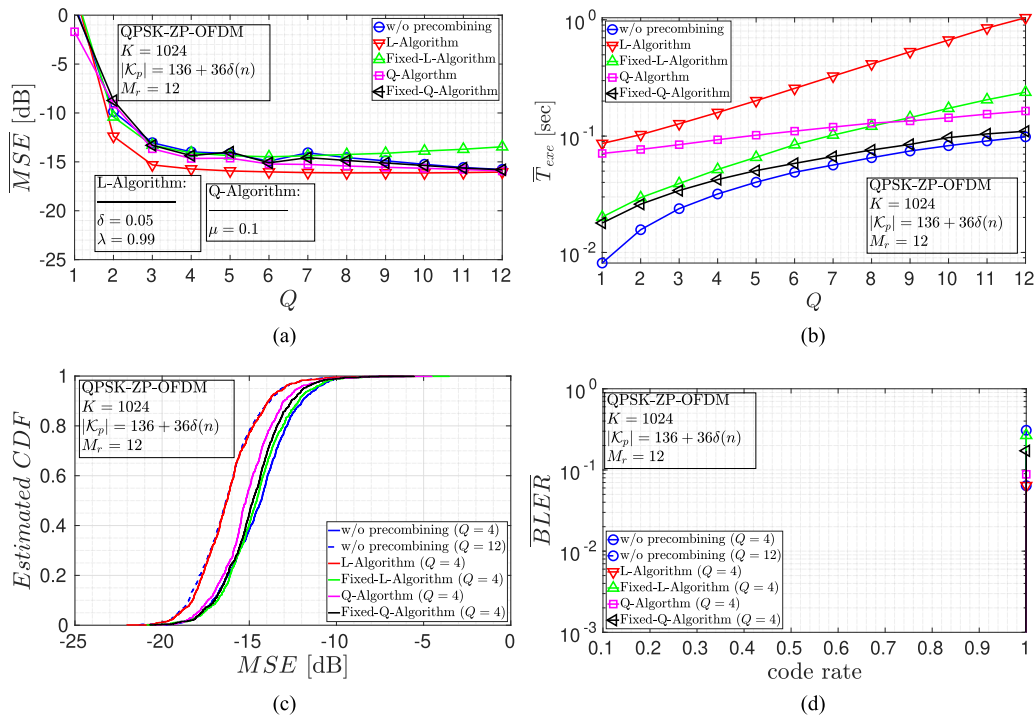

Fig. 14. (a) Average MSE as a function of the number of precombiner outputs Q for coherent receivers with and without the precombining schemes. (b) Average execution time as a function of the number of precombiner outputs Q. (c) Estimated CDF of the MSE for the adaptive precombining algorithms. (d) BLER as a function of rate of the LDPC code. The conventional receivers use Q maximally spaced receivers to detect the data symbols. The forgetting factor of the L-Algorithm is $\lambda = 0.99$ and the initial variance σ^2 is set to 0.05 in all of the cases. The step size of the Q-Algorithm is $\mu = 0.1$ in all of the cases. The L/Q Algorithms use the first 36 carriers of the first block as pilots for training, and thereafter switches to the decision-directed mode for the entire frame.

BER values that cannot be measured with the existing data. Using code rates as high as 0.7, the threshold-based PI algorithm can achieve BER and BLER as low as 2×10^{-5} and 10^{-2} for OFDM blocks whose carriers are 8-PSK modulated. Therefore, the PI algorithm enables higher bit rate than LS/OMP for the same BER.

In Fig. 14, we demonstrate the performance of the precombining schemes; namely, L/Q Algorithms and their fixed variants, where the precombining weights computed in the first block are used for the rest of the blocks within a frame, in terms of data detection MSE, average execution time, empirical CDF of MSE per OFDM block, and BLER. The performance of the receivers without precombining is also shown as a benchmark. The results shown in Fig. 14 reflect all 104 transmission of ZP-OFDM signals with $K = 1024$ carriers modulated by QPSK data symbols. The L/Q Algorithms use the first 36 carriers of the first block as pilots for training the precombining coefficients, and thereafter switches to the decision-directed mode for the entire frame, thus achieving the same bandwidth efficiency as the full-complexity receivers, which is mentioned in Table II, on average.

Fig. 14(a) illustrates the average data detection MSE of the precombining schemes; namely, L/Q Algorithms, which are used in coherent OFDM systems, as a function of the number of precombiner outputs Q varying from $Q = 1$ to $Q = M_r = 12$. The threshold-based PI with resolution $I = 4$ and threshold $\eta = 0.1$ is used to perform channel estimation on the Q precombined output channels. Clearly, there exists a form of saturation in performance with $Q = 4$ the average MSE reaches a value

that remains almost constant with further increase in Q. While an increase in Q from 1 to 4 improves the performance dramatically, changing the number of precombiner output channels from 4 to 12 results in a total fluctuation of the average MSE of less than 1 dB which is insignificant for the overall receiver performance at the given values of the average MSE. However, the corresponding change in complexity is considerable. As shown in Fig. 14(b), the complexity (measured in terms of average execution time) of the receiver equipped with fixed Q-Algorithm that combines 12 input channels into 4 output channels is 2.3 times less than the complexity of the conventional receiver.

Clearly, Fig. 14(c) and (d) illustrates that systems equipped with the L-Algorithm achieve the same level of reliability as the conventional receiver with full complexity. While the L-Algorithm delivers MSE below -14 for 90% of OFDM blocks, the Q-Algorithm delivers MSE below -13 dB for 90% of OFDM blocks. Using a code rate as high as 1, the L/Q-Algorithms in a 4-channel configuration achieve BLER as low as 6.5×10^{-2} and 9×10^{-2}, respectively.

VII. CONCLUSION

Within the framework of OFDM signal detection, we investigated a channel estimation method that operates on the transformed version of the input signal to identify the dominant propagation paths. Unlike the conventional methods, PI considers a continuum of delays and allows for increasing the delay resolution without undue penalty on complexity.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

16   IEEE JOURNAL OF OCEANIC ENGINEERING

We presented a comparative performance analysis using simulation and experimental signals recorded over a mobile acoustic channel. Simulation results, as well as experimental results obtained from all 210 transmissions spanning 3.5 h during the MACE'10 experiment, have clearly shown the effectiveness of the proposed PI algorithm. Our results show that the PI algorithm consistently outperforms the conventional LS and offers a performance comparable to that OMP, albeit with much lower computational complexity. Specifically, the PI algorithm is on average 8 times faster than the OMP algorithm. Such an advantage is of paramount importance for practical implementation of high data rate acoustic OFDM systems. In terms of performance, the PI algorithm delivered an average MSE below -14 dB for 93% of OFDM blocks, while the OMP and LS did so for 83% and 68% of OFDM blocks, respectively. We also showed that with the PI algorithm, we achieved higher bit rate than LS/OMP with the same BER.

We also presented two precombining schemes based on differential encoding for acoustic multichannel OFDM systems. Without requiring *a priori* knowledge about the spatial distribution of signals, the scheme extracts spatial coherence between receiving elements (if such exists) by linearly combining the M_r available channels into Q channels, so as to reduce the number of subsequent channel estimators. For further reduction in receiver complexity, we also considered the fixed variants of the presented schemes, where the precombining weights computed in the first block are used for the rest of the blocks within a frame. The algorithms learn spatial coherence of signal recursively, allowing the precombiner coefficients to change from one carrier to another, thus effectively achieving broadband processing.

Using the experimental signals, we demonstrated the performance of the proposed techniques. Our results show that precombining the 12 input channels into 4 for subsequent coherent processing yields the same performance in terms of average data detection MSE and BLER at the cost of very few additional pilots. We also showed that the 4-channel configuration delivers an average MSE below -14 dB for 90% of OFDM blocks and achieves BLER as low as 6.5×10^{-2} using code rates as high as 1.

REFERENCES

[1] M. Stojanovic, "Low complexity OFDM detector for underwater acoustic channels," in *Proc. IEEE OCEANS Conf.*, Boston, MA, USA, Sep. 2006, pp. 1–6.

[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] K. Tu, T. M. Duman, M. Stojanovic, and J. G. Proakis, "Multiple-resampling receiver design for OFDM over Doppler-distorted underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 2, pp. 333–346, Apr. 2013.

[4] Y. M. Aval and M. Stojanovic, "Differentially coherent multichannel detection of acoustic OFDM signals," *IEEE J. Ocean. Eng.*, vol. 40, no. 2, pp. 251–268, Apr. 2015.

[5] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1997, vol. 3, pp. 1845–1848.

[6] O. Edfors, M. Sandell, J. Van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[7] J. Van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE 45th Veh. Tech. Conf.*, Jul. 1995, vol. 2, pp. 815–819.

[8] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[9] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[10] H. Minn and V. K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation," *IEEE Trans. Broadcast.*, vol. 46, no. 4, pp. 240–248, Dec. 2000.

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

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

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

[14] A. Radosevic, R. Ahmed, T. M. Duman, J. G. Proakis, and M. Stojanovic, "Adaptive OFDM modulation for underwater acoustic communications: Design considerations and experimental results," *IEEE J. Ocean. Eng.*, vol. 39, no. 2, pp. 357–370, Apr. 2014.

[15] M. Stojanovic and S. Tadayon, "Estimation and tracking of time-varying channels in OFDM systems," in *Proc. 52nd Annu. Conf. Commun., Control, Comput.*, Sep. 2014, pp. 116–122.

[16] M. Stojanovic, J. Catipovic, and J. Proakis, "Reduced-complexity spatial and temporal processing of underwater acoustic communication signals," *J. Acoust. Soc. Amer.*, vol. 98, no. 2, pp. 961–972, 1995.

[17] A. Tadayon and M. Stojanovic, "Exploitation of spatial coherence for reducing the complexity of acoustic OFDM systems," in *Proc. 4th Underwater Commun. Netw. Conf.*, Aug. 2018, pp. 1–4.

[18] A. Tadayon and M. Stojanovic, "Low-complexity superresolution frequency offset estimation for high data rate acoustic OFDM systems," *IEEE J. Ocean. Eng.*, to be published, doi: 10.1109/JOE.2018.2869657.

[19] 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.

[20] A. Björck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA, USA: SIAM, 1996.

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

[22] G. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2012, p. 293.

[23] S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge, U.K.: Cambridge Univ. Press, 2018, p. 232.

[24] R. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Norwood, MA, USA: Artech House, 2000, pp. 112–115.

[25] M. Stojanovic, "An adaptive algorithm for differentially coherent detection in the presence of intersymbol interference," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1884–1890, Sep. 2005.

[26] M. B. Porter, The Bellhop Manual and User's Guide. (Preliminary Draft), 2011. Accessed: May 31, 2019. [Online]. Available: http://oalib.hlsresearch.com/Rays/HLS-2010-1.pdf

[27] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, Oct. 2013.

[28] L. Freitag and S. Singh, "Performance of micro-modem psk signaling with a mobile transmitter during the 2010 MACE experiment," in *Proc. IEEE OCEANS Conf.*, Genova, Italy, May 2015, pp. 1–7.

[29] P. Qarabaqi and M. Stojanovic, Acoustic Channel Simulator, 2014. Accessed: May 31, 2019. [Online]. Available: http://millitsa.coe.neu.edu/sites/millitsa.coe.neu.edu/files/acoustic_ channel_simulator_info.pdf

[30] W. Ryan and S. Lin, *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009, pp. 201–202.

[31] J. Hamkins, "Performance of low-density parity-check coded modulation," in *Proc. IEEE Aerosp. Conf.*, Mar. 2010, pp. 1–14.

[32] D. J. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

Amir Tadayon (S'07–GS'15) received the B.S. degree in electrical engineering from Tehran University, Tehran, Iran, in 2011 and the M.S. degree in electrical engineering from Northeastern University, Boston, MA, USA, in 2016, where he is currently working toward the Ph.D. degree.

His research interests include statistical signal processing and digital communications, and their applications to underwater acoustic systems.

Milica Stojanovic (SM'08–F'10) received the Dipl.Ing. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston, MA, USA, in 1991 and 1993, respectively.

She was a Principal Scientist with the Massachusetts Institute of Technology, and in 2008, joined Northeastern University, where she is currently a Professor of electrical and computer engineering. She is also a Guest Investigator with the Woods Hole Oceanographic Institution. Her research interests include digital communications theory, statistical signal processing, and wireless networks, and their applications to underwater acoustic systems.

Prof. Stojanovic is an Associate Editor for the IEEE JOURNAL OF OCEANIC ENGINEERING and a past Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. She also serves on the Advisory Board of IEEE COMMUNICATIONS LETTERS and the Editorial Board of *IEEE Signal Processing Magazine*, and Chairs of the IEEE Ocean Engineering Society's Technical Committee for Underwater Communication, Navigation, and Positioning. She is the recipient of the 2015 IEEE/OES Distinguished Technical Achievement Award, and is the 2008 IEEE/OES Distinguished Lecturer.