# Sparse Channel Estimation and Equalization for OFDM-Based Underwater Cooperative Systems With Amplify-and-Forward Relaying

Erdal Panayirci, *Life Fellow, IEEE*, Habib Senol, *Member, IEEE*, Murat Uysal, *Senior Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—This paper is concerned with a challenging problem of channel estimation and equalization for amplify-and-forward cooperative relay based orthogonal frequency division multiplexing (OFDM) systems in sparse underwater acoustic (UWA) channels. The sparseness of the channel impulse response and prior information for the non-Gaussian channel gains, modeled by an exact continuous Gaussian mixture (CGM), are exploited to improve the performance of the channel estimation algorithm. The resulting novel algorithm initially estimates the overall sparse complex-valued channel taps from the source to the destination as well as their locations using the matching pursuit (MP) approach. The effective time-domain non-Gaussian noise is approximated well as a Gaussian noise in the frequency-domain, where the estimation takes place. An efficient and low complexity algorithm is developed based on a combination of the MP and the maximum a posteriori probability (MAP) based space-alternating generalized expectation-maximization technique, to improve the estimates of the channel taps and their locations in an iterative manner. Computer simulations show that the UWA channel is estimated very effectively and the proposed algorithm exhibits excellent symbol error rate and channel estimation performance.**

***Index Terms*—AF relaying, continuous Gaussian mixture, matching pursuit, OFDM, SAGE, underwater acoustic channel estimation.**

## I. Introduction

DEMANDS for underwater communication systems are increasing due to the on-going expansion of human activities in underwater environments such as environmental monitoring, pollution control and tracking, underwater exploration, scientific data collection, maritime archaeology, offshore oil field exploration, port security and tactical surveillance [1]–[4]. Wire-line systems, particularly fiber optical systems, can be deployed to provide real time communication in underwater applications. However, their high cost and operational disadvantages due to the lack of flexibility become restrictive for most practical cases. This triggers the growing demand for underwater wireless links.

Radio, optical, or acoustic waves can be used for wireless transmission underwater. The transmission ranges of radio and optical underwater systems are limited to short distances. With relatively favorable propagation characteristics of acoustic waves, acoustic systems achieve longer transmission ranges (on the order of kilometers) underwater and are widely deployed in practice. This technology however suffers from a very small available bandwidth and data rates are limited to a few tens of kilobits/sec (kb/s) [1]. With the emerging bandwidth-hungry underwater applications and the concept of "Underwater Internet of Things" [5], demanding requirements are further imposed on underwater acoustic (UWA) systems. To address these new challenges, innovative physical layer designs such as multiple-input multiple-output (MIMO) communication and orthogonal frequency division multiplexing (OFDM) have been used in UWA systems (see e.g., [6]–[10]) to exploit spatial and multipath diversities.

Another promising research direction to take advantage from diversity benefits in UWA systems is relay-assisted (cooperative) communication. Originally introduced in the context of terrestrial radio systems [11], cooperative communication has also been applied to UWA systems [12]–[17]. In [12], Choudhuri and Mitra have derived capacity bounds for UWA relay channels and investigated optimum power allocation. Through the comparisons with the rates achievable via direct transmission and two hop communication, it is concluded that cooperative relaying increases the rates significantly. In [13], Carbonelli *et al.* have considered a multi-hop system with decode-and-forward (DF) relaying and demonstrated that it outperforms the direct transmission since path loss degradation is much better addressed with the use of relays. In [14], Zhang *et al.* have considered a similar scenario in [13] and investigated the performance of multi-hop underwater systems taking into account practical constraints such as the frequency dependent signal attenuation, inter-hop interference, half-duplex constraint, and large propagation delay. In [15], Vajapeyam *et al.* have investigated a distributed space-time block coding scheme with amplify-and-forward (AF) relaying and reported

Manuscript received April 01, 2015; revised July 20, 2015; accepted August 22, 2015. Date of publication September 10, 2015; date of current version December 09, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yue Rong. This research was supported by the Turkish Scientific and Research Council (TUBITAK) under Grant 110E092 and in part by the U.S. National Science Foundation under Grants CCF-1420575, CNS-1456793 and ECCS-1343210. This paper was presented in part at the IEEE Global Comminations Conference, Anaheim, CA, USA, December 2012.

E. Panayirci is with the Department of Electrical and Electronics Engineering Kadir Has University, 34083, Istanbul, Turkey (e-mail: eepanay@khas.edu.tr).

H. Senol is with the Department of Computer Engineering, Kadir Has University, 34083, Istanbul, Turkey (e-mail: hsenol@khas.edu.tr).

M. Uysal is with the Department of Electrical and Electronics Engineering, Özyegin University, 34794, Istanbul, Turkey (e-mail: murat.uysal@ozyegin.edu.tr).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Digital Object Identifier 10.1109/TSP.2015.2477807

diversity gain improvements over direct transmission via simulations and experimental results. In [16], Rahmati and Duman have considered a cooperative underwater OFDM system and addressed asynchronism problems. They have proposed a delay diversity structure at the destination and, through the analysis of pairwise error probability, they have demonstrated that the proposed system achieves full spatial diversity. In [17], Jalil and Ghrayeb have proposed a distributed coding scheme that aims at achieving a good diversity-multiplexing trade-off in UWA cooperative networks. They have considered a network with multiple source nodes, multiple relay nodes, and a single destination and derived the end-to-end error performance of this scheme.

A common underlying assumption in most current works on cooperative UWA systems is the availability of perfect channel estimation at the receiver side with few exceptions [18]–[20]. In a practical coherent system, the fading channel coefficients should be estimated during the training period and then used in the detection process. The channel estimation problem in a cooperative system with DF relaying involves the individual estimation of source-to-relay and relay-to-destination channels. This estimation problem is essentially the same as in point-to-point links and the existing results on channel estimation for point-to-point UWA links (see e.g., [21]–[24] and the references therein) can be used for this purpose. On the other hand, for a cooperative system with AF relaying, the channel estimation problem involves the estimation of a cascaded channel consisting of source-to-relay and relay-to-destination links. Channel estimation for AF relaying has been studied earlier in the context of RF communications, see e.g., [25]–[28], but those results are not directly applicable to UWA systems since the underlying channels are much different. Particularly, the impulse response of the UWA channel is often sparse as the multipath arrivals should become resolvable. Furthermore, the effective noise entering the system between the source and the destination through the relay is correlated and non-Gaussian. The cascaded channel structure with the combination of sparsity and correlated non-Gaussian noise creates a challenging estimation problem for UWA AF relaying systems, which is the motivation for our work.

In this paper, we present a pilot assisted channel estimation technique for relay networks that employ the AF transmission scheme in UWA channels. Our main contribution in this work is two-fold. First, we exploit the sparse structure of the UWA channel impulse response to improve the performance of the channel estimation algorithm, due to the reduced number of taps to be estimated. The resulting algorithm initially estimates the overall sparse channel taps from the source to the destination as well as their locations using the matching pursuit (MP) approach [29]. The overall time-domain non-Gaussian noise, affecting the system from source to destination, is approximated well in the frequency domain by a Gaussian noise where the channel estimation is employed. We also model the prior probability density functions (pdfs) of the overall cascaded complex channel gains, from source-to-relay and relay-to-destination, as *continuous Gaussian mixtures* (CGMs) and show that an exponential type of mixing pdf admits this representation exactly. Second, based on the CGM model, we develop an efficient and low complexity novel channel estimation algorithm by combining the MP and the space-alternating generalized expectation-maximization (SAGE) techniques, called the CGM-MP-SAGE algorithm which relies on the concept of the *admissible hidden data*, to improve the estimates of the channel taps and their locations as well as the channel prior distribution parameters in an iterative manner. We demonstrate that by suitably choosing the admissible hidden data on which the SAGE algorithm relies, a subset of parameters is updated for analytical tractability and the remaining parameters for faster convergence [30], [31].

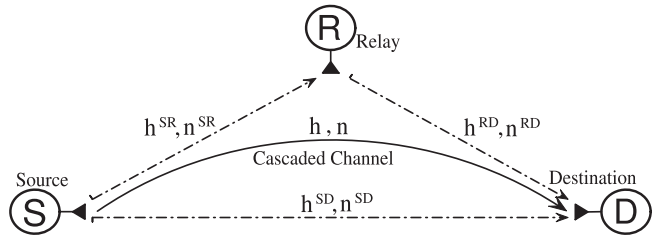


Fig. 1. Underwater cooperative system with amplify-and-forward relaying.

The remainder of the paper is organized as follows. Section II presents a system model for an OFDM-based underwater cooperative wireless communication system and descries the main parameters of the UWA channel. Section III proposes the new sparse channel estimation and equalization algorithms and Section IV provides the basic initialization techniques for the iterative algorithm proposed. Section V includes a computational complexity analysis and provides performance results based on simulations. Finally, Section VI contains concluding remarks.

## II. System Model

We consider an OFDM-based cooperative wireless communication scenario in which the source node $S$ transmits information to the destination node $D$ with the assistance of a relay node $R$ each of which is equipped with a single pair of transmit and receive antennas. The cooperation is based on the receive diversity (RD) protocol [32] with single-relay AF relaying with half-duplex nodes. In our work, we assume that the relay node does not perform channel estimation in order to keep its complexity as low as possible. As shown in Fig. 1, in the broadcasting phase, the source node transmits to the destination and the relay nodes. In the relaying phase, the relay node forwards a scaled noisy version of the signals received from the source. The UWA channel between each node pair is characterized by multipath propagation, typically with a few significant paths, resulting in a sparse multipath channel model [33]. This type of channel can be represented by a *parametric channel model*, consisting of a limited number of distinct paths parametrized by the path delays and path gains. The parametric channel model effectively reduces the dimension of the signal estimation problem, and the corresponding channel estimation can achieve better performance than that of non-parametric channel model-based estimators.

Time-varying UWA channel impulse responses (CIRs) for $S \rightarrow R$, $R \rightarrow D$ and $S \rightarrow D$ links are sparse and characterized by

$$h^{SR}(t,\tau) = \sum_{p=1}^{L_{SR}} A_p^{SR}(t)\delta(\tau - \tau_p^{SR}(t)) \quad (1)$$

$$h^{RD}(t,\tau) = \sum_{p=1}^{L_{SD}} A_p^{RD}(t)\delta(\tau - \tau_p^{RD}(t)) \quad (2)$$

$$h^{SD}(t,\tau) = \sum_{p=1}^{L_{SD}} A_p^{SD}(t)\delta(\tau - \tau_p^{SD}(t)), \quad (3)$$

where, for $\times \in \{SR, RD, SD\}$, $L_\times$, $A_p^\times(t)$ and $\tau_p^\times(t)$ denote the number of non-zero paths, the real channel path amplitudes and the time-varying path delays, respectively. We assume that the path gains on each link remain constant over one OFDM symbol transmission and vary independently from symbol to symbol. That is, $A_p^\times(t) \approx A_p^\times, p = 1, 2, \ldots, L_\times$.

The continuously time-varying delays $\tau_p^\times(t)$ are caused by motion of the transmitter/receiver as well as scattering of the moving sea surface or reflection due to sound speed variations [33]. For the duration of an OFDM symbol, the time variations of the path delays can be approximated well by a Doppler rate as $\tau_p^\times(t) = \tau_p^\times - \gamma_p^\times t$. We assume that the path delays, $\tau_p^\times$ are constant over an OFDM symbol duration and all paths have a similar Doppler scaling factor, that is $\gamma_p^\times \equiv \gamma^\times$. Note that, in general the Doppler scaling factor can be different for each path [34]. However, it was stated in [35] that as long as the dominant Doppler shift is caused by direct transmitter/receiver motion, this assumption can be justified. On the other hand, in [43], the problem with the path-specific Doppler scaling distortions in OFDM receiver designs were investigated and new receiver structures proposed to cope with these kinds of distortions. Also an efficient solution has been proposed in [44] for the challenges imposed by the different Doppler scaling factors by computing an average Doppler scaling factor. Consequently, we can in this case operate our channel estimation and equalization algorithm with this single (average) Doppler value as if it represents the fixed dominant Doppler scalar factor.

Taking these assumptions and approximations into account, the time-varying continuous-time multipath UWA channel impulse response models above are simplified to

$$h^\times(t,\tau) = \sum_{p=1}^{L_\times} A_p^\times\ \delta\left(\tau - \left(\tau_p^\times - \gamma^\times t\right)\right). \quad (4)$$

The baseband equivalent channel impulse response of $h^\times(t,\tau)$ in (4) can be determined as follows [36]:

$$h_b^\times(t,\tau) = \sum_{p=1}^{L_\times} \alpha_p^\times\ e^{j2\pi f_c \gamma^\times t}\delta\left(\tau - \left(\tau_p^\times - \gamma^\times t\right)\right), \quad (5)$$

where $\alpha_p^\times \triangleq A_p^\times e^{-j2\pi f_c \tau_p^\times}$. Note that the model in (4) deals only with real channel path amplitudes, $A_p^\times$, obtained from a ray tracing technique. However, there are many diffuse multipath components diffracted or scattered by the rough sea and bottom surface. Consequently, assuming high carrier frequency, multipath components will have random phases uniformly distributed over $[0, 2\pi)$, and by the central limit theorem, we can assume that the channel coefficients $\alpha_p^\times$ in (5) are uncorrelated, zero-mean complex Gaussian random variables each having variance $\sigma_p^2$. The average power of the channel is normalized to unity, that is, $\sum_{p=1}^{L} \sigma_p^2 = 1$. Depending on the sea conditions, each channel gain $|\alpha_p^\times|$ can be assumed to have a different distribution. When the receiver is in shallow water and close to the transmitter, diffuse random multipath contributions are negligible and the channel tap gains may be assumed to obey the Rician distribution. On the other hand, with increasing distance between transmitter and receiver, large sea dynamics prevent direct path contributions and mostly the diffuse multipaths dominate, resulting in the channel gains obeying a Rayleigh distribution. In our work we assume *a priori* that the channel path gains obey a Rayleigh distribution.

At the source node, the OFDM based relay system with $N$ subcarriers employs actively $K$ subcarriers to transmit data symbols and nothing is transmitted on the remaining $N - K$ carriers. During any OFDM symbol, each active subcarrier is modulated by a data symbol $d_\varsigma[k]$, where $\varsigma$ and $k$ represents the OFDM symbol index and the subcarrier index, respectively. After taking an $N$-point inverse fast Fourier transform (IFFT) of the data sequence and adding a cyclic prefix (CP) of duration $T_{CP}$ before transmission, to avoid inter-symbol interference (ISI), the baseband-equivalent, continuous time-domain transmitted signal can be expressed as

$$s(t) = \frac{1}{\sqrt{N}} \sum_{\varsigma=0}^{M-1} \sum_{k=-K/2}^{K/2-1} d_\varsigma[k] \exp\left(\frac{j2\pi k}{T}(t - \varsigma T_{SYM} - T_{CP})\right) \otimes g_T(t), \quad 0 \le t \le T_{SYM}, \quad (6)$$

where $\otimes$ denotes linear convolution and $g_T(t)$ is the impulse response of the transmitter filter. $T_{SYM} = T + T_{CP}$ is the total OFDM symbol duration and $\Delta f = 1/T$ is the OFDM subcarrier spacing.

The received passband signals in the broadcasting phase (first time slot) at the relay and the destination nodes are given as

$$\widetilde{y}_1^R(t) = \widetilde{s}(t) \otimes h^{SR}(t,\tau) + \widetilde{n}^{SR}(t) = \sqrt{2}\Re e\left\{\left(\sum_{p=1}^{L_{SR}} \alpha_p^{SR} e^{j2\pi f_c \gamma^{SR} t} s\left((1+\gamma^{SR})t - \tau_p^{SR}\right)\right) e^{j2\pi f_c t}\right\} + \widetilde{n}^{SR}(t) \quad (7)$$

and

$$\widetilde{y}_1^D(t) = \widetilde{s}(t) \otimes h^{SD}(t,\tau) + \widetilde{n}^{SD}(t) = \sqrt{2}\Re e\left\{\left(\sum_{p=1}^{L_{SD}} \alpha_p^{SD} e^{j2\pi f_c \gamma^{SD} t} s\left((1+\gamma^{SD})t - \tau_p^{SD}\right)\right) e^{j2\pi f_c t}\right\} + \widetilde{n}^{SD}(t) \quad (8)$$

respectively. In the relaying phase (second time slot), the received passband signal at the destination node is shown in (9) at the bottom of the next page, where $\varpi = \left(\widetilde{y}_1^R(t)\right)_{rms}$ is the normalization factor and $y_1^R(\cdot)$ is the baseband equivalent of $\widetilde{y}_1^R(\cdot)$. $\widetilde{n}^{SR}(t), \widetilde{n}^{RD}(t)$ and $\widetilde{n}^{SD}(t)$ represent the additive white Gaussian noises (AWGNs) in the channels $S \rightarrow R$, $R \rightarrow D$ and

$S \to D$, respectively, each with one-sided power spectral density $N_0$. To ensure that the power budget is not violated, the relay node normalizes the receive signal $\widetilde{y}_1^R(t)$ by $\varpi$.

Substituting (7) into (9) and neglecting the second order products, (9) can be expressed as

$$\widetilde{y}_2(t) = \frac{\sqrt{2}}{\varpi}\Re\mathfrak{e}\left\{\Big(\sum_{p=1}^{L_{SR}}\sum_{q=1}^{L_{RD}} \alpha_p^{SR}\alpha_q^{RD} e^{j2\pi f_c(\gamma^{SR}+\gamma^{RD})t} \times s\left((1+\gamma^{SR}+\gamma^{SR})t-(\tau_p^{SR}+\tau_q^{RD})\right)\Big)e^{j2\pi f_c t}\right\}+\widetilde{n}_2(t) \tag{10}$$

$$= \frac{\sqrt{2}}{\varpi}\Re\mathfrak{e}\left\{\left(\sum_{\ell=1}^{L} h_\ell e^{j2\pi f_c\gamma t} s\left((1+\gamma)t-\tau_\ell\right)\right)e^{j2\pi f_c t}\right\}+\widetilde{n}_2(t), \tag{11}$$

where, $L = L_{SR} \times L_{RD}$,

$$h_\ell \triangleq \alpha_p^{SR}\alpha_q^{RD} \tag{12}$$

$$\tau_\ell \triangleq \tau_p^{SR} + \tau_q^{RD} \tag{13}$$

$$\gamma \triangleq \gamma^{SR} + \gamma^{RD} \tag{14}$$

and

$$\widetilde{n}_2(t) = \frac{1}{\varpi}\widetilde{n}^{SR}(t) \otimes h^{RD}(t,\tau) + \widetilde{n}^{RD}(t). \tag{15}$$

There is a one-to-one relationship between $\ell$ and $(p,q)$ in (12) and (13), and conversion between them can be easily obtained as follows:

$$\begin{aligned}&\text{Given } (p,q) \Rightarrow \ell = q + L_{RD}(p-1), \text{ and}\\ &\text{given} \qquad \ell \Rightarrow p = \left\lfloor \frac{\ell-1}{L_{RD}} \right\rfloor + 1, q = (\ell-1)\bmod(L_{RD})+1,\end{aligned} \tag{16}$$

where $\lfloor . \rfloor$ denotes the floor operator.

Note that in typical UWA communications systems, the receiver directly samples the passband signal due to a relatively low carrier frequency. Consequently, resampling, passband-to-baseband downshifting as well as Doppler shift estimation, compensation and channel estimation are often performed in the digital domain. As specified in [33] and [34], to mitigate the frequency dependent Doppler effect in the received signal (11), the following three operations are carried out on the sampled received *passband* signal to obtain the final discrete frequency-domain samples.

1) The main Doppler scaling effect is removed through resampling of the operation on the samples of the passband received signal taken at discrete times $t = mT_s$ with a sampling rate $1/T_s$ as $\widetilde{y}_2[m] = \widetilde{y}_2(mT_s)$, with a resampling factor $1+\widehat{\gamma}$, leading to the resampled signal $\widetilde{z}_2[m] = \widetilde{y}_2(mT_s/(1+\widehat{\gamma}))$. From (11) it follows that

$$\widetilde{z}_2[m] = \frac{\sqrt{2}}{\varpi}\Re\mathfrak{e}\left\{e^{j2\pi f_c mT_s}e^{j2\pi f_c bmT_s}\sum_{\ell=1}^{L} h_\ell s\left((1+b)mT_s-\tau_\ell\right)\right\} +\widetilde{n}_2[m], \tag{17}$$

where

$$b = \frac{\gamma-\widehat{\gamma}}{1+\widehat{\gamma}}. \tag{18}$$

The Doppler effect, $\gamma$, is divided into a non-zero valued Doppler rate, caused by relative movement of the transmitter, the relay and the receiver, and a Doppler spread, centered around zero caused by different travel paths and receiver angle shifts. As can be see from (17), the non-zero mean of $\gamma$ is removed by the resampling. Consequently, the new residual Doppler shift $b$ on each path is spread around zero between $[-\nu_{\max}, \nu_{\max}]$, after compensation by $\widehat{\gamma}$. $\nu_{\max}$ can be chosen based on the Doppler spread, with resolution $\Delta\nu = 2\nu_{\max}/N_\nu$. $N_\nu$ is the number of grid points on the Doppler spectrum.

2) The sampled version of the baseband-equivalent received signal $z_2(t)$ is obtained from (17) via a demodulation process. Demodulation is implemented digitally by passband-to-baseband down shifting followed by a lowpass filtering with the sampling rate $1/T_s = B$. At the sampling instants $t = nT_s + T_{CP}$, the baseband discrete-time signal can be expressed as follows:

$$z_2[n] = \frac{1}{\varpi}e^{j2\pi\nu t}\sum_{\ell=1}^{L} h_\ell s\left((1+b)t-\tau_\ell\right)\Big|_{t=nT_s+T_{CP}} + v_2[n], \tag{19}$$

where $\nu \triangleq f_c b$, represents the residual carrier frequency offset (CFO) and $v_2[n]$ is the baseband equivalent discrete-time noise signal of the resampled $\widetilde{n}_2[m]$ as in (17). Finally compensating the residual mean Doppler shift by its estimated value $\widehat{\nu}$ and substituting (6) in (19), for $t = nT_s + T_{CP}$, the baseband equivalent received discrete-time signal can be expressed as (see e.g., [33])

$$\begin{aligned} z[n] &= e^{-j2\pi\widehat{\nu}t} z_2[n]\\ z[n] &= e^{j2\pi(\nu-\widehat{\nu})t}\sum_{\ell=1}^{L} h_\ell \frac{1}{\varpi\sqrt{N}}\sum_{\varsigma=0}^{M-1}\sum_{q=-K/2}^{K/2-1} d_\varsigma[q]\\ &\quad\times \exp\left(\frac{j2\pi q}{T}(t-\varsigma T_{SYM}-T_{CP})\right)\exp\left(\frac{j2\pi q}{T}(bt-\tau_\ell)\right)\\ &\quad\otimes g_T\left((1+b)t-\tau_\ell\right) + v[n], \end{aligned} \tag{20}$$

where $v[n] = e^{-j2\pi\widehat{\nu}t}v_2[n]$.

$$\widetilde{y}_2(t) = \frac{1}{\varpi}\widetilde{y}_1^R(t) \otimes h^{RD}(t,\tau) + \widetilde{n}^{RD}(t) = \frac{\sqrt{2}}{\varpi}\Re\mathfrak{e}\left\{\left(\sum_{p=1}^{L_{RD}}\alpha_p^{RD}e^{j2\pi f_c\gamma^{RD}t}y_1^R\left((1+\gamma^{RD})t-\tau_p^{RD}\right)\right)e^{j2\pi f_c t}\right\} + \widetilde{n}^{RD}(t) \tag{9}$$

We assume that the cyclic prefix, $T_{CP}$, is greater than or equal to the maximum channel path delay to prevent the inter-carrier-interference (ICI) from entering the system, that $K$ active subcarriers are within the flat region of the frequency filters, and that the residual Doppler shift $\nu$ is fully compensated, that is $\nu = \widehat{\nu}$. Also, since the time varying-channel impulse response is assumed to be constant over the duration of one OFDM symbol, and that the number of channel paths and the path delays do not change during an OFDM symbol, it is sufficient to consider channel estimation only symbol by symbol. Therefore, omitting the OFDM symbol index for notational simplicity, the $n$th time sample within any OFDM symbol after the CP removal can be expressed as

$$z[n] = \sum_{\ell=1}^{L} h_\ell \frac{1}{\varpi\sqrt{N}} \sum_{q=-K/2}^{K/2-1} d[q] \exp\left(\frac{j2\pi q}{T}((1+b)nT_s + bT_{CP} - \tau_\ell)\right) + v[n], \quad (21)$$

where $T = NT_s$, and the $h_\ell$'s represent the overall discrete equivalent channel complex-valued path gains. They are non-Gaussian distributed random variables with zero-means and variances $\sigma_\ell^2 = \sigma_p^{2,SR} \times \sigma_q^{2,RL}$, *a priori*, since for each $\ell \Leftrightarrow (p,q)$, $h_\ell = \alpha_p^{SR} \times \alpha_q^{RD}$ and $\alpha_p^{SR}$ and $\alpha_q^{RD}$ are independent, complex Gaussian random variables with means zero and variances $\sigma_p^{2,SR}$ and $\sigma_q^{2,RD}$, $\left(\sum_p \sigma_p^{2,SR} = \sum_q \sigma_q^{2,RD} = 1\right)$, representing the complex channel gains on links $S \to R$ and $R \to D$, respectively.

Similarly, overall additive additive noise samples in the cascade channel on the $S \to R \to D$ link are denoted by $v[n], n = 0, 1, \ldots, N-1$. From (15) it follows that these are also non-Gaussian random variables with zero-means and variances $BN_0(1+\frac{1}{\varpi})$, since the additive noise signals $n^{SR}(.)$ and $n^{RD}(.)$ on links $S \to R$ and $R \to D$ are zero-mean, white complex Gaussian processes with one-sided spectral density $N_0$ and the complex channel multipath gains, $\alpha_q^{RD}$ and $\alpha_p^{SR}$, on these links are also independent Gaussian as specified in the paragraph above.

An $N$–point fast Fourier transform (FFT) is applied to transform the sequence $z[n]$ into the frequency domain. By the assumption that the $K$ subcarriers on which data is transmitted are within the flat region of the transmitter and receiver filters of unity gains, the $k$th subcarrier output of the FFT during one OFDM symbol can be represented by

$$\begin{aligned} Z[k] &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z[n] \exp\Big(-j\frac{2\pi nk}{N}\Big) \\ &= \sum_{\ell=1}^{L} h_\ell A[k,\ell] + V[k], k = -\frac{K}{2}, \ldots, \frac{K}{2}-1, \end{aligned} \quad (22)$$

where

$$A[k,\ell] = \sum_{q=-K/2}^{K/2-1} d[q] F_q(k,\ell)$$

$$F_q[k,\ell] = \frac{1}{\varpi N} \sum_{n=0}^{N-1} \exp\left(j\frac{2\pi q}{N}(bn + b\frac{T_{CP}}{T_s} - \widetilde{\tau}_\ell)\right) \exp\left(j\frac{2\pi n}{N}(q-k)\right), \quad (23)$$

where $\widetilde{\tau}_\ell = \tau_\ell / T_s$ is the normalized $\ell$th path delay. Frequency domain noise samples $V[k]$ in (22) are determined from $V[k] = (1/\sqrt{N}) \sum_{n=0}^{N-1} v[n] \exp(-j\frac{2\pi nk}{N})$. Note that although the time-domain noise samples, $v[n]$, are generated from a wide-sense stationary, non-Gaussian random process, as explained earlier, the complex frequency-domain noise samples $V[k]$ are approximately zero-mean Gaussian distributed due to the statistical independence of carriers (via the central-limit theorem). It can be easily shown from (15) that the frequency-domain noise samples $V[k]$ are uncorrelated and consequently independent since they are approximately Gaussian, with zero means and variances $\sigma_{V[k]}^2 \triangleq \sigma_V^2 = N_0 B(1 + \frac{1}{\varpi^2})$. Consequently, by inserting (23) into (22), the vector form of (22) can be expressed as

$$\mathbf{Z} = \mathbf{A}\mathbf{h} + \mathbf{V}, \quad (24)$$

where

$$\begin{aligned} \mathbf{Z} &= [Z[-K/2], Z[-K/2+1], \ldots, Z[K/2-1]]^T \in \mathcal{C}^K \\ \mathbf{h} &= [h_1, h_2, \ldots, h_L]^T \in \mathcal{C}^L \\ \mathbf{V} &= [V[-K/2], V[-K/2+1], \ldots, V[K/2-1]]^T \in \mathcal{C}^K \end{aligned}$$

and the $[k,\ell]$th element of $\mathbf{A} \in \mathcal{C}^{K \times L}$ is determined from (23).

In this work, we are mainly interested in estimation of a fast time-varying, cascade UWA channel in a cooperative system with amplify-and-forward relaying, based on the observation model (24). As explained earlier, the overall continuous-time channel impulse response is represented by a parametric model in which each distinct path is characterized by a few significant complex-valued path amplitudes, $\{h_\ell\}_{\ell=1}^L$, the normalized path delays $\{\widetilde{\tau}_\ell \triangleq \tau_\ell / T_s\}_{\ell=1}^L$ and the Doppler spread $b$, resulting in a sparse multipath channel model. Note that the overall cascade channel parameters $h_l, \widetilde{\tau}_\ell$ and $b$ are determined by means of the individual parameters of the channels from the source to relay and the relay to destination in (12), (13) and (18), respectively.

Conventional matching pursuit algorithms have been applied to sparse channel estimation [37]. However, in practice, the sparsity assumption does not always hold due to the non-integer normalized path delays in the equivalent discrete-time baseband representation of the channel. Therefore, such an estimated channel may differ substantially from the original channel. To improve the channel estimation performance, the analog-to-digital (A/D) conversion at the input of the OFDM receiver is implemented with an oversampling rate $R_s^{(\varrho)} = \varrho / T_s, \varrho = 1, 2, \ldots$, leading to finer delay resolution, where $\varrho$ is the oversampling factor, and $1/T_s$ is the baseband Nyquist sampling rate. Consequently, the real-valued normalized path delays $\widetilde{\tau}_\ell, \ell = 1, 2, \ldots, L$, can be discretized as $\eta_\ell = \lfloor \varrho \widetilde{\tau}_\ell \rfloor$ and take values from the set of the possible discrete path delays:

$$\eta_\ell \in \mathcal{T} = \{0, 1, 2, \ldots, N_\tau - 1\}, \quad (25)$$

where $N_\tau = \varrho L_{CP}$, $L_{CP} = T_{CP}/T_s$ and $T_{CP}$ is the duration of the CP. Note that $\varrho = 1$ is typically chosen in MP-based channel estimators.

Similarly, the real-valued Doppler spread $b$ can be discretized as $\beta = \lfloor (b + \nu_{\max}) / \Delta\nu \rfloor$ and takes values from the set of possible discrete Doppler rates:

$$\beta \in \mathcal{B} = \{0, 1, 2, \ldots, N_\nu - 1\}, \quad (26)$$

where $N_\nu = (2\nu_{\max})/\Delta\nu$.

Based on the associated discrete Doppler rates and the random channel tap positions, $\beta, \{\eta_\ell\}_{\ell=1}^{L}$, the received signal in (24) can be rewritten as

$$\mathbf{Z} = \sum_{\ell=1}^{L} \mathbf{a}_{r_\ell} h_\ell + \mathbf{V}, \tag{27}$$

where, $\mathbf{a}_{r_\ell}$ is the $r_\ell$th column vector of the finer resolution matrix $\mathbf{A}^{(\varrho)} \in \mathcal{C}^{K \times N_\tau N_\nu}$, called a *dictionary matrix*, whose columns correspond to the $(\beta, \eta_\ell)$th discrete multipath channel taps and Doppler rate positions, respectively. For a given $r_\ell \Leftrightarrow (\beta, \eta_\ell)$, the $k$th component of $\mathbf{a}_{r_\ell}$ is determined from (23) as

$$a_{r_\ell}[k] = A[k,\ell]\Big|_{(b=-\nu_{\max}+\beta\Delta\nu, \widetilde{\tau}_\ell=\eta_\ell/\varrho)}. \tag{28}$$

Note that for $\beta \in \{0, 1, \ldots N_\nu - 1\}$, and $\eta_\ell \in \{0, 1, \ldots N_\tau - 1\}$, the total number of columns of $\mathbf{A}^{(\varrho)}$ is $N_\tau N_\nu$ and they are labeled as $r_\ell = 1, 2, \ldots, N_\tau N_\nu$. The conversion between $r_\ell$ and $(\beta, \eta_\ell)$ can be easily obtained as follows:

$$\begin{aligned} r_\ell &= \beta + N_\nu \eta_\ell + 1, \\ \beta &= r_\ell \bmod(N_\nu) - 1 \text{ and, } \eta_\ell = \left\lfloor \frac{r_\ell}{N_\nu} \right\rfloor - 1. \end{aligned}$$

Since $L \ll N_\tau N_\nu$, the overall cascade channel can be assumed to be sparse and the channel estimation problem can be solved as a sparse signal recovery problem. The MP or orthogonal MP algorithms are popular sparse recovery methods and can be used for channel estimation problems. However, the overall discrete equivalent cascade channel impulse response in UWA cooperative communication systems may not represent a truly sparse channel impulse response even though the individual channels are perfectly sparse. Thus, the channel estimation performance of these algorithms based on compressed sensing may degrade substantially.

To improve the channel estimation performance, we propose a new technique that combines the MP and the *maximum a posteriori probability* (MAP) algorithms, using the CGM pdf as prior information for the channel gains, under the SAGE framework, which we call the CGM-MP-SAGE algorithm. We incorporate priors for the unknown complex valued channel gains, $h_\ell, \ell = 1, 2, \ldots, L$, in a Bayesian framework, which are defined in (12) as $h_\ell = \alpha_p^{SR} \times \alpha_q^{RD}$ where $\alpha_q^{SR} \sim \mathcal{CN}(0, \sigma_p^{SR})$ and $\alpha_q^{RD} \sim \mathcal{CN}(0, \sigma_q^{SR})$. Since the bivariate probability density functions of $\alpha_q^{SR}$ and $\alpha_q^{RD}$ have spherical symmetry, it can be clearly seen that the $h_\ell$'s are also spherically symmetric having non-Gaussian pdfs. Consequently, they can be represented by a CGM $f(h_\ell) = \int_0^\infty \phi(h_\ell \mid \sigma_\ell^2 \nu_\ell) p(\nu_\ell) d\nu_\ell$, where $\phi(x|\sigma^2)$ denotes the zero-mean Gaussian pdf with variance $\sigma^2$, and where $p(\nu_\ell)$ represents the mixing probability distribution [38].

In the following section, we demonstrate how the CGM-MP-SAGE algorithm estimates the sparse complex-valued channel gains, channel path delays and Doppler spread in an iterative way. We assume the CGM model, as mentioned above, for the *a priori* pdfs of $h_\ell, \ell = 1, 2, \ldots, L$. The initial channel gains as well as the path delays and the Doppler shift are estimated by an MP algorithm and they are updated within the SAGE iterations to improve their estimation performance. Computer simulations show that the improvement in the estimation performance resulting from this approach is very substantial.

## III. Sparse Multipath Channel Estimation With the CGM-MP-SAGE Algorithm

We now propose a new iterative algorithm, called the CGM-MP-SAGE algorithm, based on the SAGE and MP techniques for channel estimation employing the signal model given by (27). The SAGE algorithm, proposed by Fessler and Hero [39], is a twofold generalization of the so-called *expectation maximization* (EM) algorithm that provides updated estimates for a set $\boldsymbol{\Theta}$of unknown parameters. First, rather than updating all parameters simultaneously at iteration $(i)$, only a subset of $\boldsymbol{\Theta}_S$ indexed by $S = S[i]$ is updated while keeping the parameters in the complement set $\boldsymbol{\Theta}_{\overline{S}}$ fixed; and second, the concept of the *complete data* $\chi$ is extended to that of the so-called *admissible hidden data* $\chi_S$ to which the observed signal $\mathcal{R}$ is related by means of a possibly nondeterministic mapping. The convergence rate of the SAGE algorithm is usually higher than that of the EM algorithm, because the conditional Fisher information matrix for each set of parameters is likely smaller than that of the complete data, given for the entire space. At the $i$th iteration, the *expectation-step* (E-step) of the SAGE algorithm is defined

$$Q_S(\boldsymbol{\Theta}_S \mid \boldsymbol{\Theta}^{[i]}) = E\left\{ \log p\left(\chi_S, \boldsymbol{\Theta}_S \mid \boldsymbol{\Theta}_{\overline{S}}^{[i]}\right) \mid \mathcal{R}, \boldsymbol{\Theta}^{[i]} \right\}.$$

In the *maximization step* (M-step), only $\boldsymbol{\Theta}_S$ is updated, i.e.,

$$\begin{aligned} \boldsymbol{\Theta}_S^{[i+1]} &= \arg\max_{\boldsymbol{\Theta}_S} Q_S(\boldsymbol{\Theta}_S \mid \boldsymbol{\Theta}^{[i]}) \\ \boldsymbol{\Theta}_{\overline{S}}^{[i+1]} &= \boldsymbol{\Theta}_{\overline{S}}^{[i]}. \end{aligned}$$

On the other hand, the MP algorithm is an iterative procedure that can sequentially identify the dominant channel taps and estimate the associated tap coefficients by choosing the column $\mathbf{a}_{r_\ell}^{(\varrho)}$ of $\mathbf{A}^{(\varrho)}$ that best align with the residual vector until all the taps are identified. Our proposed CGM-MP-SAGE algorithm implements the MP algorithm at each SAGE iteration step by updating all the dominant channel taps and the associated tap coefficients sequentially. The details of the algorithm are presented below.

The unknown parameter set to be estimated in our problem is

$$\boldsymbol{\Phi} = \{\mathbf{h}, \mathbf{r}, \boldsymbol{\vartheta}\}, \tag{29}$$

where $\mathbf{h} = [h_1, h_2, \ldots, h_L]^T$, $\mathbf{r} = (\beta, \boldsymbol{\eta})$, $\boldsymbol{\eta} = [\eta_1, \eta_2, \ldots, \eta_L]^T$, and $\boldsymbol{\vartheta} = \{\sigma_1^2, \ldots, \sigma_L^2\}$. The first step in deriving the CGM-MP-SAGE algorithm for estimating $\boldsymbol{\Phi}$ based on the received vector $\mathbf{Z}$ in (27) is the specifications of complete data and admissible hidden data sets whose pdfs are characterized by the common parameter set $\boldsymbol{\Phi}$. To obtain a receiver architecture that iterates between soft-data and channel estimation in the CGM-MP-SAGE algorithm, we decompose $\boldsymbol{\Phi}$ into $L+1$ subsets, representing the parameters, $\mathbf{h}$, $\mathbf{r}$ and $\boldsymbol{\vartheta}$, as follows:

- The first $L$ subsets of $\boldsymbol{\Phi}$ are chosen as $\boldsymbol{\Phi}_\ell = \{h_\ell, r_\ell\}$, $\ell = 1, 2, \ldots, L$, representing $\mathbf{h}$ and $\mathbf{r} = (\beta, \boldsymbol{\eta})$. For each subset, we define $\bar{\boldsymbol{\Phi}}_\ell = \boldsymbol{\Phi} \setminus \boldsymbol{\Phi}_\ell = \{\bar{\mathbf{h}}_\ell, \bar{\mathbf{r}}_\ell, \boldsymbol{\vartheta}\}$, $\bar{\mathbf{h}}_\ell = \mathbf{h} \setminus h_\ell$, $\bar{\mathbf{r}}_\ell = \mathbf{r} \setminus r_\ell$, where $\setminus$ denotes the exclusion operator.

- The $(L+1)$st subset of $\boldsymbol{\Phi}$ is chosen as $\boldsymbol{\Phi}_{L+1} = \boldsymbol{\vartheta}$ and $\bar{\boldsymbol{\Phi}}_{L+1} \triangleq \boldsymbol{\Phi} \setminus \boldsymbol{\Phi}_{L+1} = \boldsymbol{\Phi} \setminus \boldsymbol{\vartheta} = \{\mathbf{h}, \mathbf{r}\}$.

At SAGE iteration $(i)$, only the parameters in one set are updated, while the other parameters are kept fixed, and this process is repeated until all parameters are updated. According to the above parameter subset definitions, each iteration of the SAGE algorithm for our problem has two steps:

1) $\boldsymbol{\Phi}_\ell, \ell = 1, 2, \ldots, L$, are updated with the CGM-MP-SAGE algorithm while $\boldsymbol{\Phi}_{L+1}$ is fixed.
2) $\boldsymbol{\Phi}_{L+1}$ is updated with the SAGE algorithm while $\boldsymbol{\Phi}_\ell, \ell = 1, 2, \ldots, L$, are fixed and this process is repeated until all parameters are updated.

We now derive the CGM-MP-SAGE algorithm as follows.

### A. Estimation of Soft Data Symbols

Suppose now that we employ $P$ known data symbols (pilots) in each OFDM symbol. We assume they are are evenly inserted into the $K$ subcarriers. Let $\mathcal{P} = \{i_1, i_2, \ldots, i_P\}$ denote the set that contains the position indices of the $P$ pilots, denoted by $d_P(q), q \in \mathcal{P}$. During the computation of (23), in the iterative channel estimation algorithm as descried shortly, we use $P$ pilot symbols and replace the $K-P$ unknown data symbols by the soft symbols that are obtained from an equalizer output determined in the previous iteration. This enables us to obtain refined channel estimates as the soft symbols become more reliable from iteration to iteration. Initially, the unknown data symbols $d_D(q)$ are set to zero [33]. From (22) and (23), it can be easily shown that $\mathbf{Z}$ can be expressed as

$$\mathbf{Z} = \mathbf{G}\mathbf{d} + \mathbf{V}, \tag{30}$$

where the $[k,q]$th element of the matrix $\mathbf{G} \in \mathcal{C}^{K\times K}$ can be written as

$$G[k,q] = \sum_{\ell=1}^{L} h_\ell F_q[k,\ell] \tag{31}$$

and $F_q[k,\ell]$ is given by (23). The data vector $\mathbf{d} = [d_0, d_1, \ldots, d_{K-1}]^T$ in (30) is a superposition of the pilot vector, $\mathbf{d}_P$, and the vector of the unknown data symbols, $\mathbf{d}_D$, as $\mathbf{d} = \mathbf{d}_P + \mathbf{d}_D$. Note that the vectors $\mathbf{d}_P, \mathbf{d}_D \in \mathcal{C}^K$ contains nonzero components only at the pilot positions $\mathcal{P} = \{i_1, i_2, \ldots, i_P\}$ and at the data positions $\overline{\mathcal{P}}$, respectively. Consequently, (30) can be expressed as

$$\begin{aligned}\mathbf{Z}_P &\triangleq \mathbf{Z} - \mathbf{G}\mathbf{d}_P \\ &= \mathbf{G}\mathbf{d}_D + \mathbf{V}.\end{aligned} \tag{32}$$

At the $i$th SAGE iteration step, then, the soft data symbols $\widetilde{\mathbf{d}}_D$ are recovered at the output of a linear minimum-mean-square-error (MMSE) equalizer as

$$\widetilde{\mathbf{d}}_D^{(i)} = \mathbf{G}^{\dagger(i)} \left(\mathbf{G}^{(i)}\mathbf{G}^{\dagger(i)} + \gamma^{-1}\mathbf{I}_K\right)^{-1} \mathbf{Z}_P^{(i)}, \tag{33}$$

where $\gamma$ is the signal-to-noise ratio (SNR). The $[k,q]$th element of $\mathbf{G}^{(i)}$ is computed from (31) by replacing the channel estimates $\{h_\ell^{(i)}, (b^{(i)}, \widetilde{\tau}_\ell^{(i)})\}_{\ell=1}^L$, obtained at the previous iteration, and $\mathbf{Z}_P^{(i)}$ is calculated from (32) as $\mathbf{Z}_P^{(i)} = \mathbf{Z} - \mathbf{G}^{(i)}\mathbf{d}_P$. Note that although the MMSE equalizer outperforms other linear equalizers, the matrix inversion in (33) requires $\mathcal{O}(K^3)$ flops which represent significant burden when $K$ is large. However, as already documented in [33], the UWA channels produce nearly-banded channel matrices $\mathbf{G}$ and this property can be exploited to reduce the complexity substantially by means of the factorization of Hermitian banded matrices [40].

Finally, the matrix $\mathbf{A}^{(i)}$ in (24), formed by the column vectors $\{\mathbf{a}_{r_1^{(i)}}, \mathbf{a}_{r_2^{(i)}}, \ldots, \mathbf{a}_{r_L^{(i)}}\}$, that is necessary for the next estimation step, can be computed easily from the estimates $\mathbf{r}^{(i)} = [r_1^{(i)}, r_2^{(i)}, \ldots, r_L^{(i)}]^T$ and $\widetilde{\mathbf{d}}^{(i)} = \mathbf{d}_P + \widetilde{\mathbf{d}}_D^{(i)}$ from the previous iteration.

### B. Estimation of $\boldsymbol{\phi}_\ell = \{h_\ell, r_\ell \equiv (\beta, \eta_\ell)\}, \ell = 1, 2, \ldots, L$

First, we state the prior information about the channel coefficients $\mathbf{h} = [h_1, h_2, \ldots, h_L]^T$. From (12) and as shown earlier, for $\ell = 1, 2, \ldots, L$, the pdf of the complex-valued, cascade channel $h_\ell$ is distributed as a multiplication of two independent complex Gaussian random variables $\alpha_p^{SR} \sim \mathcal{CN}(0, \sigma_p^{2,SR})$ and $\alpha_q^{RD} \sim \mathcal{CN}(0, \sigma_p^{2,RD})$. Consequently, $h_\ell$ is non-Gaussian and, it can be shown that the pdf of $h_\ell = \alpha_p^{SR} \times \alpha_q^{RD}$ is

$$f(h_\ell) = \frac{2}{\pi\sigma_\ell^2} K_0\left(\frac{2 \mid h_\ell \mid}{\sigma_\ell}\right), \tag{34}$$

where $\sigma_\ell = \sigma_p^{SR}\sigma_q^{RD}$ and $K_0(\cdot)$ is the zeroth-order modified Bessel function of the second kind. The spherical symmetric non-Gaussian pdf of the complex channel gains, $\{h_\ell\}$, can be represented *a priori* by a CGM distribution

$$f(h_\ell) = \int_0^\infty \phi(h_\ell | \sigma_\ell^2 \nu_\ell) p(\nu_\ell) d\nu_\ell, \tag{35}$$

where $\phi(x|s^2) \triangleq (1/\pi s^2)\exp\left(-|x|^2/s^2\right)$ denotes a zero-mean, complex Gaussian pdf with variance $s^2$. The pdf $p(\nu_\ell)$ represents a mixing distribution on $[0, \infty)$. It is interesting to see that the non-Gaussian pdf of $f(h_\ell)$ in (34) admits this representation exactly with $\sigma_\ell^2 = \sigma_p^{2,SR}\sigma_q^{2,RD}$ for the mixing pdf $p(\cdot)$ having an exponential form as follows [41]:

$$p(\nu_\ell) = e^{-\nu_\ell}, \nu_\ell \geq 0. \tag{36}$$

Taking then into account the prior distribution of the channel coefficients specified above, a suitable approach for applying the CGM-MP-SAGE algorithm for estimation of $\boldsymbol{\Phi}_\ell$ is to decompose the $k$th sample of the receive signal in (27) into the sum

$$Z[k] = X_\ell[k] + \bar{X}_\ell[k], k = -\frac{K}{2}, \ldots, \frac{K}{2} - 1, \tag{37}$$

where

$$X_\ell[k] = a_{r_\ell}[k]h_\ell + V[k], \bar{X}_\ell[k] = \sum_{p=1, p\neq\ell}^{L} a_{r_p}[k]h_p \tag{38}$$

and $a_{r_\ell}[k]$ denotes the $k$th element of the $\mathbf{a}_{r_\ell}$. We define the *admissible hidden data* as $\chi_\ell = \{\mathbf{X}_\ell, \boldsymbol{\nu}\}$, where $\mathbf{X}_\ell = [X_\ell[-K/2], X_\ell[-K/2+1], \ldots, X_\ell[K/2-1]]^T$ and $\boldsymbol{\nu} = [\nu_1, \nu_2, \ldots, \nu_L]^T$.

To perform the *E-Step* of the CGM-MP-SAGE algorithm, the conditional expectation is taken over $\chi_\ell$ given the observation

$\mathbf{Z}$ and given that $\mathbf{\Phi}$ equals its estimate calculated at the $i$th iteration:

$$\begin{aligned} Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)}) &= E\left\{\log p\left(\chi_\ell, \mathbf{\Phi}_\ell|, \bar{\mathbf{\Phi}}_\ell^{(i)}\right)|\mathbf{Z}, \mathbf{\Phi}^{(i)}\right\} \\ &\sim E\left\{\log p\left(\mathbf{X}_\ell|h_\ell, r_\ell, \boldsymbol{\nu}, \bar{\mathbf{\Phi}}_\ell^{(i)}\right)|\mathbf{Z}, \bar{\mathbf{\Phi}}^{(i)}\right\} \\ &\quad + E\left\{\log p\left(h_\ell|\nu_\ell\right)|\mathbf{Z}, \bar{\mathbf{\Phi}}^{(i)}\right\} \end{aligned} \tag{39}$$

where, from (38)

$$\log p\left(\mathbf{X}_\ell|h_\ell, r_\ell, \boldsymbol{\nu}, \bar{\mathbf{\Phi}}_\ell^{(i)}\right) = -\frac{1}{\sigma_V^2}\sum_{k=-K/2}^{K/2-1}\left|X_\ell[k] - a_{r_\ell}[k]h_\ell\right|^2, \tag{40}$$

and, since under the condition that $\nu_\ell$ is given, $h_\ell$ is a zero-mean Gaussian random variable with variance equal to $\sigma_\ell^{2^{(i)}}\nu_\ell$, it follows from (35) that

$$\begin{aligned} \log p\left(h_\ell \mid \nu_\ell\right) &= \log\left(\phi(h_\ell \mid \sigma_\ell^{2^{(i)}}\nu_\ell)\right) \\ &\sim -\frac{|h_\ell|^2}{\sigma_\ell^{2^{(i)}}\nu_\ell}. \end{aligned} \tag{41}$$

Inserting (40) and (41) into (39) and taking expectations it follows that

$$\begin{aligned} Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)}) = \frac{1}{\sigma_V^2}\sum_{k=-K/2}^{K/2-1}\left(2\Re\left\{\widehat{X_\ell[k]}^{(i)}a_{r_\ell}^*[k]h_\ell^*\right\} - |a_{r_\ell}[k]|^2|h_\ell|^2\right) \\ - \frac{\widehat{\nu_\ell^{-1}}^{(i)}}{\sigma_\ell^{2^{(i)}}}|h_\ell|^2, \end{aligned} \tag{42}$$

where $\Re(\cdot)$ and $(\cdot)^*$ denote the real part and the conjugate operators, respectively and $\widehat{X_\ell[k]}^{(i)}$, $\widehat{\nu_\ell^{-1}}^{(i)}$ are defined as

$$\begin{aligned} \widehat{X_\ell[k]}^{(i)} &\triangleq E\left\{X_\ell[k] \mid Z[k], \mathbf{\Phi}^{(i)}\right\} \\ \widehat{\nu_\ell^{-1}}^{(i)} &\triangleq E\left\{\frac{1}{\nu_\ell} \mid \mathbf{Z}, \mathbf{\Phi}^{(i)}\right\}. \end{aligned} \tag{43}$$

Recalling (37), it follows that

$$\widehat{X_\ell[k]}^{(i)} = Z[k] - \sum_{p=1, p\neq\ell}^{L} a_{r_p^{(i)}}[k]h_p^{(i)}. \tag{44}$$

The vector form of (42) can be expressed as follow:

$$\begin{aligned} Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)}) = \frac{2}{\sigma_V^2}\Re\{\mathbf{a}_{r_\ell}^\dagger\widehat{\mathbf{X}_\ell}^{(i)}h_\ell^*\} \\ - \left(\frac{1}{\sigma_V^2}\mathbf{a}_{r_\ell}^\dagger\mathbf{a}_{r_\ell} + \frac{\widehat{\nu_\ell^{-1}}^{(i)}}{\sigma_\ell^{2^{(i)}}}\right)|h_\ell|^2, \end{aligned} \tag{45}$$

where from (44),

$$\widehat{\mathbf{X}_\ell}^{(i)} = \left[\widehat{X_\ell[-K/2]}^{(i)}, \ldots, \widehat{X_\ell[K/2-1]}^{(i)}\right]^T = \mathbf{Z} - \sum_{p=1, p\neq\ell}^{L}\mathbf{a}_{r_p^{(i)}}h_p^{(i)}.$$

In the *M-step* of the CGM-MP-SAGE algorithm, the estimates of $\mathbf{\Phi}_\ell = \{h_\ell, r_\ell\}$ are updated sequentially for $\ell = 1, 2\ldots, L$, at the $(i+1)$st iteration according to

$$\mathbf{\Phi}_\ell^{(i+1)} = \arg\max_{\mathbf{\Phi}_\ell} Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)}), \tag{46}$$

where $Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)})$ is given by (45). So, taking the derivative of $Q_\ell(\mathbf{\Phi}_\ell|\mathbf{\Phi}^{(i)})$ with respect to $h_\ell^*$ and equating it to zero, we find the final SAGE estimates of $(r_\ell, h_\ell)$ at the $(i+1)$st iteration as follows:

$$\begin{aligned} r_\ell^{(i+1)} &= \left(\beta^{(i+1)}, \eta_\ell^{(i+1)}\right) = \arg\max_r \frac{\left|\mathbf{a}_r^\dagger\widehat{\mathbf{X}_\ell}^{(i)}\right|^2}{\mathbf{a}_r^\dagger\mathbf{a}_r}, \\ h_\ell^{(i+1)} &= \frac{\frac{1}{\sigma_V^2}\mathbf{a}_{r_\ell^{(i+1)}}^\dagger\widehat{\mathbf{X}_\ell}^{(i)}}{\frac{1}{\sigma_V^2}\mathbf{a}_{r_\ell^{(i+1)}}^\dagger\mathbf{a}_{r_\ell^{(i+1)}} + \frac{\widehat{\nu_\ell^{-1}}^{(i)}}{\sigma_\ell^{2^{(i)}}}}, \end{aligned} \tag{47}$$

where $r = (\beta, \eta)$, $\beta \in \{0, 1, \ldots, N_\nu - 1\}$ and $\eta \in \{0, 1, \ldots, N_\tau - 1\}$ with $\eta \notin \{\eta_1^{(i+1)}, \eta_2^{(i+1)} \ldots, \eta_{\ell-1}^{(i+1)}\}$.

### C. Computation of $\widehat{\nu_\ell^{-1}}^{(i)}$

From (43) it follows that

$$\widehat{\nu_\ell^{-1}}^{(i)} = \int_0^\infty \frac{1}{\nu_\ell}p(\nu_\ell \mid \mathbf{Z}, \mathbf{h}^{(i)}, \mathbf{r}^{(i)}). \tag{48}$$

Note that

$$p\left(\nu_\ell \mid \mathbf{Z}, \mathbf{h}^{(i)}, \mathbf{r}^{(i)}\right) = \frac{\phi\left(h_\ell^{(i)} \mid \sigma_\ell^{2^{(i)}}\nu_\ell\right)p(\nu_\ell)}{f(h_\ell^{(i)})}, \tag{49}$$

where $p(\nu_\ell)$ and $f(h_\ell^{(i)})$ were given earlier by (36) and (34), respectively. Inserting (49) in (48) we have

$$\widehat{\nu_\ell^{-1}}^{(i)} = \frac{1}{\pi\sigma_\ell^{2^{(i)}}f(h_\ell^{(i)})}\int_0^\infty \frac{1}{\nu_\ell^2}\exp\left\{-\left(\frac{|h_\ell^{(i)}|^2/\sigma_\ell^{2^{(i)}}}{\nu_\ell} + \nu_\ell\right)\right\}d\nu_\ell. \tag{50}$$

The integral in (50) can be computed with the aid of a formula given in [41] as follows:

$$\widehat{\nu_\ell^{-1}}^{(i)} = \frac{\sigma_\ell^{(i)}}{|h_\ell^{(i)}|}\frac{K_1\left(2|h_\ell^{(i)}|/\sigma_\ell^{(i)}\right)}{K_0\left(2|h_\ell^{(i)}|/\sigma_\ell^{(i)}\right)}, \ell = 1, 2, \ldots, L. \tag{51}$$

### D. Estimation of $\boldsymbol{\phi}_{L+1} = \boldsymbol{\vartheta} = \{\sigma_1^2, \ldots, \sigma_L^2\}$

We define the admissible hidden data as $\chi_{L+1} = \{\mathbf{h}, \boldsymbol{\nu}\}$ to estimate the mixture parameters $\boldsymbol{\vartheta} = \{\sigma_1^2, \ldots, \sigma_L^2\}$. To perform the *E-Step* of the algorithm, the conditional expectation is taken over $\chi_{L+1}$ given the observation $\mathbf{Z}$ and given that $\mathbf{\Phi}$ equals its estimate calculated at the $i$th iteration:

$$\begin{aligned} Q_{L+1}(\mathbf{\Phi}_{L+1}|\mathbf{\Phi}^{(i)}) &= E\left\{\log p(\chi_{L+1}|\mathbf{\Phi}_{L+1}, \bar{\mathbf{\Phi}}_{L+1}^{(i)}) \mid \mathbf{Z}, \mathbf{\Phi}^{(i)}\right\} \\ &= E\left\{\log p(\mathbf{h}, \boldsymbol{\nu}|\boldsymbol{\vartheta}, \mathbf{h}^{(i)}, \mathbf{r}^{(i)}) \,|\, \mathbf{Z}, \mathbf{h}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\vartheta}^{(i)}\right\} \\ &\sim E\left\{\log p(\mathbf{h}|\boldsymbol{\vartheta}, \boldsymbol{\nu}, \mathbf{h}^{(i)}) \,|\, \mathbf{Z}, \mathbf{h}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\vartheta}^{(i)}\right\}. \end{aligned} \tag{52}$$

Note that the last expression in (52) follows since $\mathbf{r}$ and $\boldsymbol{\nu}$ are independent of $\boldsymbol{\vartheta} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2]^T$. Also, since the condition that $\boldsymbol{\nu}$ is given, the components of $\mathbf{h} = [h_1, h_2, \ldots, h_L]^T$ are zero-mean, complex Gaussian with variances $\sigma_\ell^2, \ell = 1, 2, \ldots, L$. Consequently, from (35) we have

$$\log p(\mathbf{h}|\boldsymbol{\vartheta}, \boldsymbol{\nu}, \mathbf{h}^{(i)}) = -\sum_{\ell=1}^{L}\left(\log\sigma_\ell^2 + \frac{|h_\ell^{(i)}|^2}{\nu_\ell\sigma_\ell^2}\right). \tag{53}$$

Inserting (53) in (52) and taking the expectation with respect to $\boldsymbol{\nu}$, we obtain

$$Q_{L+1}(\boldsymbol{\Phi}_{L+1}|\boldsymbol{\Phi}^{(i)}) = -\sum_{\ell=1}^{L}\left(\log\sigma_\ell^2 + \frac{\widehat{\nu_\ell^{-1}}^{(i)}|h_\ell^{(i)}|^2}{\sigma_\ell^2}\right), \tag{54}$$

where $\widehat{\nu_\ell^{-1}}^{(i)}$ is given in (51).

In the *M-step* of the algorithm, the estimates $\boldsymbol{\Phi}_{L+1} = \boldsymbol{\vartheta}$ are updated at the $(i+1)$st iteration according to the following constrained maximization problem:

$$\begin{aligned}\boldsymbol{\Phi}_{L+1}^{(i+1)} &= \arg\max_{\boldsymbol{\Phi}_{L+1}} Q_{L+1}(\boldsymbol{\Phi}_{L+1}|\boldsymbol{\Phi}^{(i)})\\ &\text{subject to :}\\ &\sum_{\ell=1}^{L}\sigma_\ell^2 = 1.\end{aligned} \tag{55}$$

It is straightforward to show that the solution of the above optimization problem, yielding the variance estimator of the CGM components, $h_\ell$, at the $(i+1)$th iteration is

$$(\sigma_\ell^2)^{(i+1)} = \frac{\widehat{\nu_\ell^{-1}}^{(i)}|h_\ell^{(i)}|^2}{\sum_{m=1}^{L}\widehat{\nu_m^{-1}}^{(i)}|h_m^{(i)}|^2}, \ell = 1,2,\ldots,L. \tag{56}$$

## IV. Initialization

### A. Initialization of $\{h_\ell^{(0)}, r_\ell^{(0)}\}$ With the Matching Pursuit Algorithm

The initialization of the SAGE algorithm is a critical issue since the algorithm may not converge if the initial values of the parameters to be estimated are not chosen properly. We apply the matching pursuit algorithm to determine the initial $\boldsymbol{\Phi}^{(0)} = \left\{h_\ell^{(0)}, r_\ell^{(0)} \equiv (\beta^{(0)}, \eta_\ell^{(0)}), \ell = 1,2,\ldots,L\right\}$ considering the observation model in (27). It is well known that as the MP algorithm can estimate the positions $r_\ell^{(0)}$ quite well, its estimation performance of the complex channel gains would not be sufficiently good. However, following the initialization step, the SAGE algorithm is able to compensate for this weakness by using those rough estimates, $h_\ell^{(0)}$, to enable the algorithm to converge to an optimal MAP solution. The details of the MP algorithm are given below as follows:

From (27) we compute the observation vector $\mathbf{Z}$ at the pilot subcarriers, $P = \{i_1, i_2, \ldots, i_P\}$, of the OFDM symbol which leads to the new reduced-dimensional observation model

$$\mathbf{Z}_P = \sum_{\ell=1}^{L}\widetilde{\mathbf{a}}_{r_\ell}h_\ell + \mathbf{V}_P, \tag{57}$$

where $\mathbf{Z}_P = [z(i_1), z(i_2), \ldots, z(i_P)]^T \in \mathcal{C}^P$, $\mathbf{V}_P = [v(i_1), v(i_2), \ldots, v(i_P)]^T \in \mathcal{C}^P$ and $\widetilde{\mathbf{a}}_{r_\ell} \in \mathcal{C}^P$ is the lower dimensional $r_\ell$th column vector of the dictionary matrix $\mathbf{A}_P^{(\varrho)} \in \mathcal{C}^{P\times N_\tau N_\nu}$. The $k$th component of $\widetilde{\mathbf{a}}_{r_\ell}$ is evaluated from (23) by replacing the index $k$ with $i_k$ for $k = 1,2,\ldots,P$, and setting the unknown data symbols to zero [33] as follows:

$$\begin{aligned}\widetilde{a}_{r_\ell}[k] &= A_P[k,\ell]\\ &= \frac{1}{\varpi N}\sum_{n=0}^{N-1}\sum_{p=1}^{P} d[i_p]\exp\left(j\frac{2\pi i_p}{N}(bn-\widetilde{\tau}_\ell)\right)\\ &\quad\times\exp\left(j\frac{2\pi n}{N}(i_p - i_k)\right)\Big|_{\left(b=-\nu_{\max}+\beta\Delta\nu, \widetilde{\tau}_\ell=\frac{\eta_\ell}{\varrho}\right)}.\end{aligned} \tag{58}$$

As a first step in the MP algorithm, the column in the dictionary matrix, $\mathbf{A}_P^{(\varrho)} = [\widetilde{\mathbf{a}}_0, \widetilde{\mathbf{a}}_1, \ldots, \widetilde{\mathbf{a}}_{N_\tau N_\nu - 1}]$, associated with the observation (57), that is best aligned with the residue vector $\boldsymbol{\rho}_0 = \mathbf{Z}_P$ is found and denoted by $\mathbf{a}_{r_1}$. Then the projection of $\boldsymbol{\rho}_0$ along this direction is removed from $\boldsymbol{\rho}_0$ and the residual $\boldsymbol{\rho}_1$ is obtained. The algorithm proceeds by sequentially choosing the column that is the best match until a termination criterion is met. At the $\ell$th iteration, the index of the vector from $\mathbf{A}_P^{(\varrho)}$ most closely aligned with the residual vector $\boldsymbol{\rho}_{\ell-1}$ is obtained as follows [33]:

$$\begin{aligned}r_\ell^{(0)} = \arg\max_j \frac{|\widetilde{\mathbf{a}}_j^\dagger\boldsymbol{\rho}_{\ell-1}|^2}{\widetilde{\mathbf{a}}_j^\dagger\widetilde{\mathbf{a}}_j}, j = 0,1,\ldots,N_\tau N_\nu - 1,\\ j \notin \{r_1^{(0)}, r_2^{(0)}, \ldots, r_{\ell-1}^{(0)}\},\end{aligned} \tag{59}$$

and a coarse channel path amplitude estimation at $r_\ell^{(0)}$ is

$$h_\ell^{(0)} = \frac{\widetilde{\mathbf{a}}_{r_\ell^{(0)}}^\dagger\boldsymbol{\rho}_{\ell-1}}{\widetilde{\mathbf{a}}_{r_\ell^{(0)}}^\dagger\widetilde{\mathbf{a}}_{r_\ell^{(0)}}}, \tag{60}$$

from which the new residual vector is computed as $\boldsymbol{\rho}_\ell = \boldsymbol{\rho}_{\ell-1} - h_\ell^{(0)}\mathbf{a}_{r_\ell^{(0)}}$. The iteration is repeated until a specified number of channel taps, $L$, have been selected or the residual becomes sufficiently small, i.e., $\|\boldsymbol{\rho}_\ell\| < \epsilon$.

### B. Initializations of the Gaussian-Mixture Parameters $\boldsymbol{\vartheta} = \{(\sigma_\ell^2)^{(0)}, \ell = 1,2,\ldots,L\}$

We use the reduced dimensional observation model given by (57) to estimate the Gaussian mixture parameters $\boldsymbol{\vartheta}$ of the complex-valued channel gains $h_\ell, \ell = 1,2.\ldots,L$. For notational simplicity, we re-write (57) in the following form:

$$\mathbf{z} = \boldsymbol{\Upsilon}\mathbf{h} + \mathbf{v}, \tag{61}$$

where $\mathbf{z} \equiv \mathbf{Z}_P, \boldsymbol{\Upsilon} \equiv [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_L], \mathbf{e}_\ell \equiv \widetilde{\mathbf{a}}_{r_\ell^{(0)}}$, $\mathbf{h} = [h_1, h_2, \ldots, h_L]^T$ and $\mathbf{v} \equiv \mathbf{V}_P$. We also use the notation $s_\ell \equiv (\sigma_\ell^2)^{(0)}, \beta \equiv 1/\sigma_V^2$ and $\alpha_\ell \equiv 1/(\nu_\ell s_\ell)$ in the following derivations. The maximum likelihood (ML) estimate of the variances $\mathbf{s} = [s_1, s_2, \ldots, s_L]^T$ is given by

$$\widehat{\mathbf{s}} = \arg\max_{\mathbf{s}} p(\mathbf{z}|\mathbf{s}). \tag{62}$$

In (62), $p(\mathbf{z}|\mathbf{s})$ can be evaluated by first averaging it over $\mathbf{h}$, conditioned on the mixing parameters $\boldsymbol{\nu} = [\nu_1, \nu_2, \ldots, \nu_L]^T$, and then taking the average over the exponential mixing distribution in (36) as

$$p(\mathbf{z}|\mathbf{s}) = \int_{\boldsymbol{\nu}}\left\{\int_{\mathbf{h}} p(\mathbf{z}|\mathbf{h},\boldsymbol{\alpha})p(\mathbf{h}|\boldsymbol{\alpha})d\mathbf{h}\right\}p(\boldsymbol{\nu})d\boldsymbol{\nu}. \tag{63}$$

From (61) and, since the components of $\mathbf{h}$ and $\boldsymbol{\nu}$ are independent, it follows that

$$p(\mathbf{z}|\mathbf{h},\boldsymbol{\alpha}) \backsim \exp\left\{-\beta\|\mathbf{z}-\boldsymbol{\Upsilon}\mathbf{h}\|^2\right\}, \tag{64}$$

$$p(\mathbf{h}|\boldsymbol{\alpha}) = \prod_{\ell=1}^{L}(\alpha_\ell/\pi)\exp\left\{-\alpha_\ell|h_\ell|^2\right\}, \tag{65}$$

$$p(\boldsymbol{\nu}) = \prod_{\ell=1}^{L}\exp\{-\nu_\ell\}. \tag{66}$$

Substituting (64) and (65) into (63), the inner integral of (63) with respect to $\mathbf{h}$ is computable and given by

$$\begin{aligned} I(\boldsymbol{\nu}) &\triangleq \int_{\mathbf{h}} p(\mathbf{z}|\mathbf{h},\boldsymbol{\alpha})p(\mathbf{h}|\boldsymbol{\alpha})d\mathbf{h} \\ &= \pi^{-K}\det\left(\boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Upsilon}^\dagger + (1/\beta)\mathbf{I}_K\right)^{-1} \\ &\times \exp\left\{\mathbf{z}^\dagger\left(\boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Upsilon}^\dagger + (1/\beta)\mathbf{I}_K\right)^{-1}\mathbf{z}\right\}, \end{aligned} \tag{67}$$

where $\mathbf{A} = \operatorname{diag}(\alpha_1,\alpha_2,\ldots,\alpha_L)$.

Using the matrix inversion lemma, $\left(\boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Upsilon}^\dagger + (1/\beta)\mathbf{I}_K\right)^{-1}$ in (67) can be expressed as

$$\left(\boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Upsilon}^\dagger + (1/\beta)\mathbf{I}_K\right)^{-1} = \beta\mathbf{I}_K - \beta\boldsymbol{\Upsilon}\left(\boldsymbol{\Upsilon}^\dagger\boldsymbol{\Upsilon} + (1/\beta)\mathbf{A}\right)^{-1}\boldsymbol{\Upsilon}^\dagger. \tag{68}$$

Taking into account the fact that $\boldsymbol{\Upsilon}^\dagger\boldsymbol{\Upsilon}$ is a banded matrix, it can be approximated as

$$\boldsymbol{\Upsilon}^\dagger\boldsymbol{\Upsilon} = \operatorname{diag}\left(\|\mathbf{e}_1\|^2,\|\mathbf{e}_2\|^2\ldots,\|\mathbf{e}_L\|^2\right), \tag{69}$$

where $\mathbf{e}_\ell$ is the $\ell$th column vector of $\boldsymbol{\Upsilon}$. Substituting (68) into (67), after replacing (69) in (68) and after some algebra, (67) can be computed as

$$I(\boldsymbol{\nu}) = \left(\prod_{\ell=1}^{L}\left(\frac{|\mathbf{e}_\ell|^2}{\alpha_\ell}+\frac{1}{\beta}\right)\right)^{-1}\exp\left\{\beta\sum_{\ell=1}^{L}\left(\xi_\ell\lambda_\ell - \|\mathbf{z}\|^2\right)\right\}, \tag{70}$$

where $\xi_\ell \triangleq (\mathbf{z}^\dagger\mathbf{e}_\ell)^2$ and $\lambda_\ell \triangleq \left(\|\mathbf{e}_\ell\|^2 + (1/\beta)\alpha_\ell\right)^{-1}$. Substituting (70) into (63), discarding the terms independent of $\mathbf{s}$ and then performing the integration with respect to $\boldsymbol{\nu}$, after some algebra we have

$$p(\mathbf{z}|\mathbf{s}) \sim \prod_{\ell=1}^{L}\frac{\exp\left\{1/(s_\ell\beta\|\mathbf{e}_\ell\|^2)\right\}}{s_\ell\|\mathbf{e}_\ell\|^2}K_0\left(\frac{2(\xi_\ell/s_\ell)^{1/2}}{\|\mathbf{e}_\ell\|^2}\right). \tag{71}$$

Finally, the ML variance estimates $\mathbf{s}$ are found by solving the following constrained optimization problem:

$$\begin{aligned} &\widehat{\mathbf{s}} = \arg\max_{\mathbf{s}}\log p(\mathbf{z}|\mathbf{s}) \\ &\text{subject to} \\ &\sum_{\ell=1}^{L}s_\ell = 1, \text{ and } \quad s_\ell \geq 0 \quad \text{for} \quad \ell = 1,2,\ldots,L. \end{aligned} \tag{72}$$

It can be easily shown that each term in the product of (71) is a convex function of $\nu_\ell, \ell = 1,2,\ldots,L$. Consequently, (72) is a convex optimization problem whose solution can be obtained by a water filling algorithm [42]. The details of the derivations and the resulting optimal estimates $\{\widehat{s}_\ell\}_{\ell=1}^{L}$ for the mixture variances are given in the appendix.

The final CGM-MP-SAGE algorithm is summarized in Fig. 2.

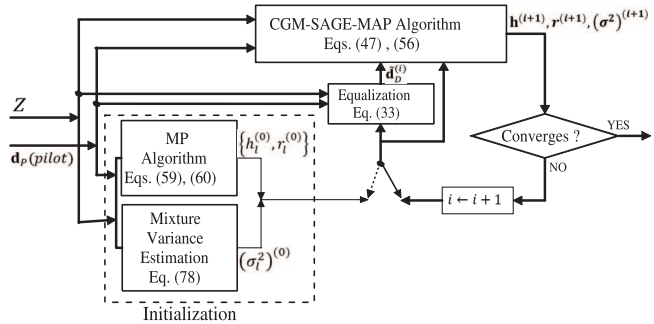


Fig. 2. Block diagram of the CGM-MP-SAGE channel estimation and equalization algorithm.

TABLE I
COMPUTATIONAL COMPLEXITY DETAILS

| INITIALIZATION | | | |
|---|---|---|---|
| Eq. No | Variable | CMs | CAs |
| (59), (60) | $r_\ell^{(0)}$, $h_\ell^{(0)}$ | $\approx(P+1)LN_\tau N_\nu$ | $\approx(P-1)LN_\tau N_\nu$ |
| (78) | $s_\ell = \sigma_\ell^{2\,(0)}$ | $\approx 31L$ | $\approx 8L$ |
| **CGM-SAGE-MAP ITERATION** | | | |
| (33) | $\widetilde{\mathbf{d}}_D^{(i)}$ | $\approx 3N^2(L+\Delta+2)+NL$ | $\approx 2N^2(L+\Delta+2)+N(L-1)$ |
| (51) | $\widehat{\nu_\ell^{-1}}^{(i)}$ | $\approx L$ | 0 |
| (47) | $r_\ell^{(i+1)}, h_\ell^{(i+1)}$ | $\approx NL(N_\tau N_\nu+3)$ | $\approx NL(N_\tau N_\nu+3)$ |
| (56) | $(\sigma_\ell^2)^{(i+1)}$ | $\approx 5L/2$ | $\approx L/2$ |

## V. COMPLEXITY ANALYSIS AND SIMULATION RESULTS

The computational complexity of the proposed algorithm is presented in Table I. Note that, the initial values, $r_\ell^{(0)}, h_\ell^{(0)}$ in (59) and (60), are obtained by the MP algorithm and require approximately $(P+1)LN_\tau N_\nu$ complex multiplications (CMs) and $(P-1)LN_\tau N_\nu$ complex additions (CAs) as given in Table I. Also, the initial values, $s_\ell = \sigma_\ell^{2\,(0)}$, are obtained by a water filling (WF) algorithm which requires $31L$ CMs and $8L$ CAs for each step of the WF algorithm.

We now present the computation complexity of the main CGM-MP-SAGE algorithm. In the computation of $\widetilde{\mathbf{d}}_D^{(i)}$ in (33), the channel matrix $\mathbf{G}^{(i)}$ and the Hermitian matrix $\left(\mathbf{G}^{(i)}\mathbf{G}^{(i)\,\dagger} + \gamma^{-1}\mathbf{I}_K\right)$ can be approximated as nearly banded whose total number of diagonal and subdiagonals is $\Delta$. This property results in a substantial reduction of the computational complexity. Namely, the computation of $\widetilde{\mathbf{d}}_D^{(i)}$ requires approximately $3N^2(L+\Delta+2)+NL$ CMs and $2N^2(L+\Delta+2)+N(L-1)$ CAs. Consequently, for updating of $\widehat{\nu_\ell^{-1}}^{(i)}, r_\ell^{(i+1)}, h_\ell^{(i+1)}$ and $(\sigma_\ell^2)^{(i+1)}$, we need approximately $NL(N_\tau N_\nu+3)+7L/2$ CMs and $NL(N_\tau N_\nu+3)+L/2$ CAs. As a result, it follows from Table I that the total computational complexity per iteration of the CGM-MP-SAGE channel estimation algorithm is approximately $3N^2(L+\Delta+2)+(N+P)LN_\tau N_\nu$ CMs and $2N^2(L+\Delta+2)+(N+P)LN_\tau N_\nu+N(L-1)$ CAs

TABLE II
CHANNEL AND SIMULATION PARAMETERS OF THE SPACE'08 EXPERIMENT

| | |
|---|---|
| carrier frequency ($f_c$) | 13 KHz |
| channel bandwidth ($BW$) | 2.442 KHz |
| number of subcarriers ($N$) | 256 |
| number of occupied subcarriers ($K$) | 224 |
| OFDM symbol duration ($T$) | 104.8 ms |
| Subcarrier spacing ($\Delta f := 1/T$) | 9.54 Hz |
| cyclic prefix duration ($T_{CP}$) | 12.4 ms |
| number of paths on each link ($L_\times$) | $3, 6, 9, 12$ |
| maximum Doppler rate ($b_{max}$) | $10^{-3}, 5\times10^{-3}, 10^{-2}, 5\times10^{-2}$ |
| Doppler spread resolution $\Delta\nu$ | $10^{-3}$ |
| modulation formats | BPSK, QPSK, 16QAM |
| number of CGM-MP-SAGE iterations | 5 |
| pilot spacing ($\Delta_p$) | 1, 2, 4, 8 |
| oversampling factor ($\varrho$) | 1, 2, 4, 8, 16 |

$\sim \mathcal{O}(N^2L)$. Consequently the complexity of the algorithm is on the order of $\mathcal{O}(NL)$ per OFDM subcarrier.

In this section, we now present computer simulation results to assess the performance of OFDM-based cooperative communication systems in UWA channels with the proposed channel estimation algorithm. The UWA OFDM specifications in the numerical simulation are chosen to match the settings used in the *SPACE'08 Experiment* as given in [33]. The experimental data of the *SPACE'08 Experiment*, which included stormy days with strong wind and wave activity leading to severe Doppler spread, was recorded off the coast of Martha's Vineyard, MA, from Oct. 14 to Nov. 1, 2008. Channel and simulation parameters, associated with the UWA channel, are summarized in Table II.

Channel delays on each link of the cascaded channel, $\{\tau_p^\times\}_{p=1}^L$, are assumed to be independent of each other and uniformly distributed within the interval $[0, T_{cp}]$. The amplitudes $\{A_p^\times\}_{p=1}^L$ are assumed to be Rayleigh distributed with the average power decreasing exponentially with delay, $E\{(A_p^\times)^2\} = \Omega_p = Ce^{-\tau_p^\times/T_{cp}}$, where $C$ is the power normalization constant such that $\sum_{p=1}^L \Omega_p = 1$. The Doppler rate $b$ is assumed to be uniformly distributed within $[-b_{\max}, b_{\max}]$, where $b_{\max}$ is chosen as in Table II.

We consider a comb-type pilot structure with equally spaced pilot subcarriers. We measure the performance of the system in terms of the frequency-domain mean square error (MSE) of our proposed channel estimator and the corresponding symbol error rate (SER). The initial estimates of the complex-valued channel gains and the corresponding positions used in the water filling algorithm to determine the initial values of the CGM parameters are obtained using the reduced complexity MP algorithm.

Figs. 3 and 4 show MSE and SER performance curves of the MP and CGM-MP-SAGE algorithms for binary phase shift-keying (BPSK), quadrature phase shift-keying (QPSK) and 16-ary quadrature amplitude modulation (16QAM) signaling formats. As seen from these curves, the CGM-MP-SAGE algorithm, having excellent channel estimation performance and symbol error rate, outperforms the MP estimator although the MP algorithm uses the linear MMSE equalizer to detect the data symbols. We conclude from these curves that the CGM-MP-SAGE algorithm exhibits superior performance in the estimation of channel tap positions, tap coefficients as well as the CGM parameters that model the *a priori* pdf of the tap coefficients of the cascaded channel.

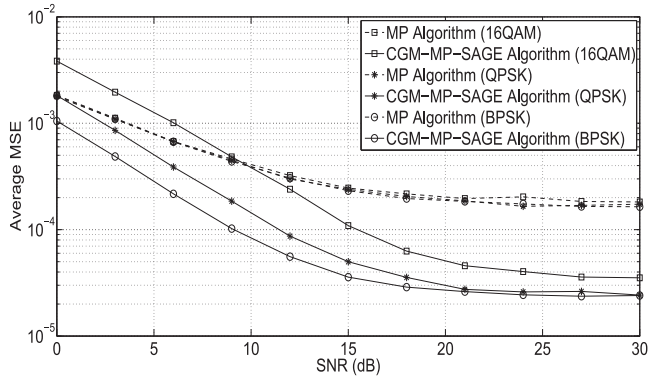


Fig. 3. MSE vs. SNR performance comparisons of the CGM-MP-SAGE and MP algorithms for different constellations: $\varrho = 8, b_{\max} = 5\times10^{-3}, \Delta_p = 4, L = 3$.

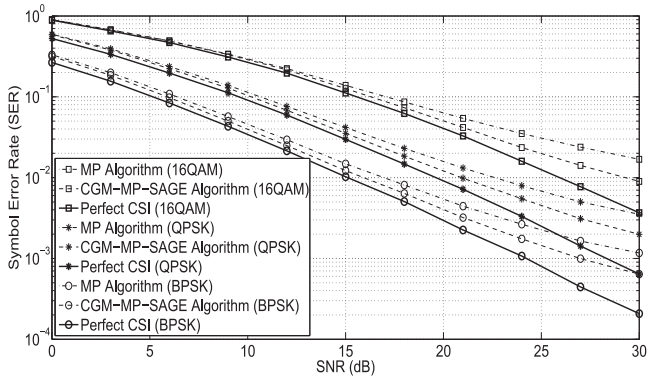


Fig. 4. SER vs. SNR performance comparisons of the CGM-MP-SAGE and MP algorithms for different constellations: $\varrho = 8, b_{\max} = 5\times10^{-3}, \Delta_p = 4, L = 3$.

In Figs. 5 and 6, the MSE and SER performance of the CGM-MP-SAGE algorithm is plotted as a function of SNR for oversampling factors of $\varrho = 1, 2, 4, 8, 16$. It can be seen from these figures that an oversampling factor of $\varrho = 8$ would be sufficient to obtain better sparse channel estimation performance of the proposed CGM-MP-SAGE algorithm. We also investigate the Doppler effect on the MSE and SER performance of the system. Figs. 7 and 8 show that the proposed CGM-MP-SAGE algorithm is quite robust to Doppler shifts up to $b_{\max} = 5\times10^{-3}$ which can be considered as severe Doppler effects. The effects of channel estimation on the average MSE performance are investigated as functions of SNR with different pilot spacings $(\Delta_p)$. The plots in Fig. 9 exhibit that the MSE performance gain becomes much less for high SNR levels beyond 20 dB despite denser pilot symbols. It can be seen from these curves that a pilot spacing of $\Delta_p = 4$ would be sufficient for maximum Doppler rates around $b_{\max} = 5\times10^{-3}$. In Fig. 10, we also examine the SER performance of the proposed CGM-MP-SAGE algorithm under tougher channel settings, i.e., $L = 3, 6, 9, 12$. It can be seen from Fig. 10 that the performance of the proposed algorithm yields quite convincing performance results even under these harsh channel conditions.

Finally we note that by examining all the SER-vs-SNR performance results presented in this section, it is clear that the channel estimation results in a decrease in the diversity order.

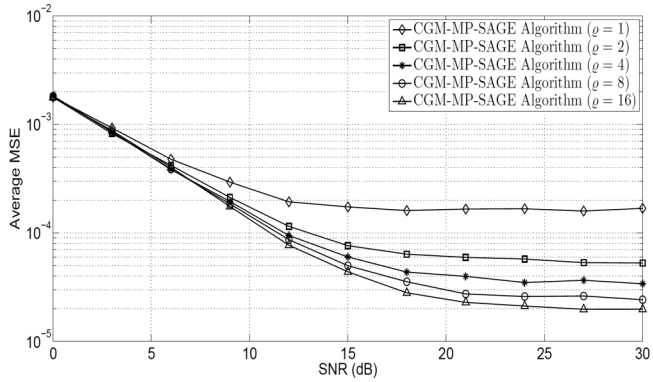


Fig. 5. MSE vs. SNR performance of the CGM-MP-SAGE algorithm for different resolution factors: $b_{\max} = 5 \times 10^{-3}$, $\Delta_p = 4$, $L = 3$, QPSK signaling.

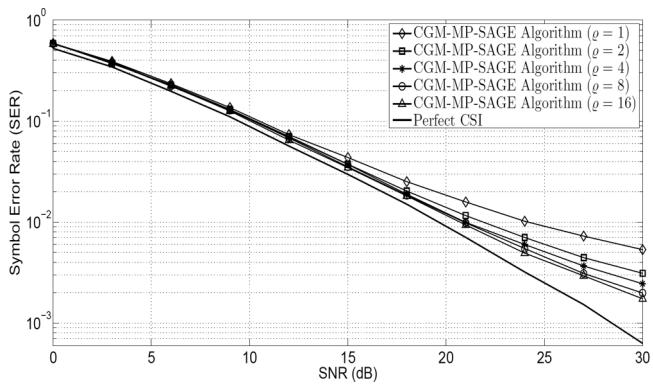


Fig. 6. SER vs. SNR performance of the CGM-MP-SAGE algorithm for different resolution factors: $b_{\max} = 5 \times 10^{-3}$, $\Delta_p = 4$, $L = 3$, QPSK signaling.

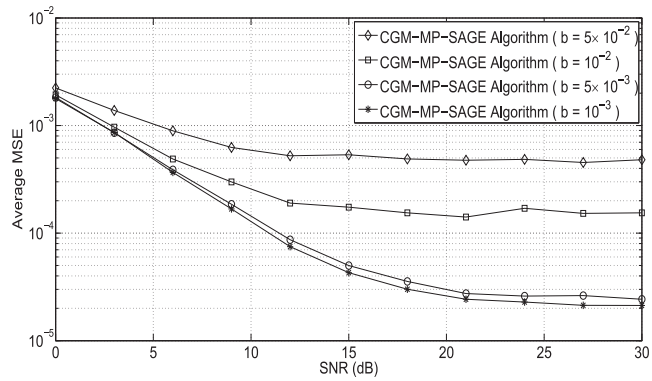


Fig. 7. MSE vs. SNR performance of the CGM-MP-SAGE algorithm for different Doppler rates: $\varrho = 8$, $\Delta_p = 4$, $L = 3$, QPSK signaling.

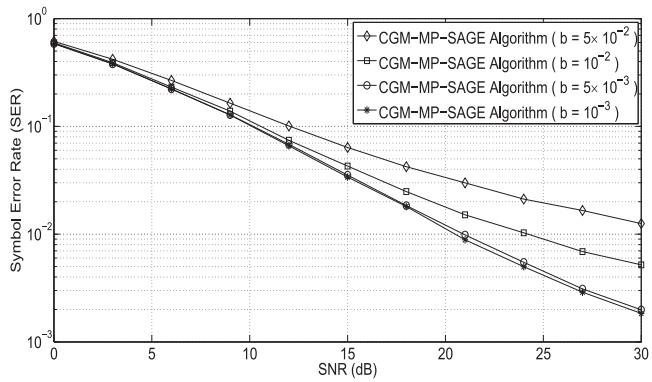


Fig. 8. SER vs. SNR performance of the CGM-MP-SAGE algorithm for different Doppler rates: $\varrho = 8$, $\Delta_p = 4$, $L = 3$, QPSK signaling.

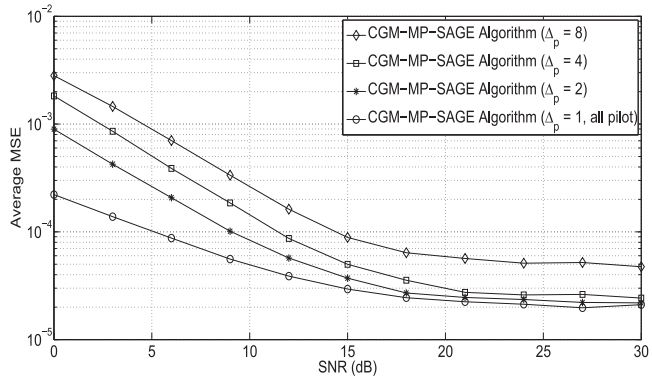


Fig. 9. MSE vs. SNR performance of the CGM-MP-SAGE algorithm for different pilot spacings: $\varrho = 8$, $b_{\max} = 5 \times 10^{-3}$, $L = 3$, QPSK signaling.

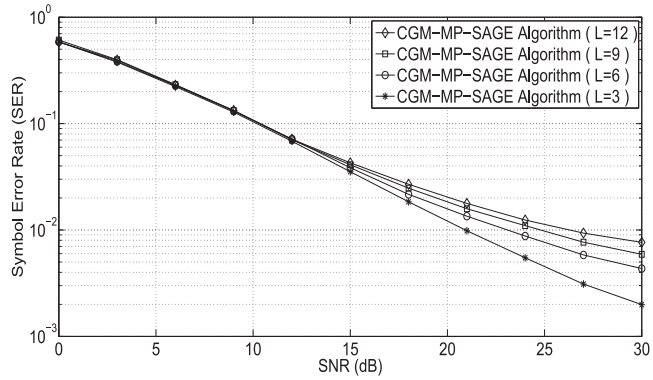


Fig. 10. SER vs. SNR performance of the CGM-MP-SAGE algorithm for different number of paths: $\varrho = 8$, $b_{\max} = 5 \times 10^{-3}$, $\Delta_p = 4$, QPSK signaling.

This is mainly due to the time-selective nature of the UWA channels. Especially, since the speed of sound in water is much lower than that of electromagnetic waves in air, the UWA channel becomes rapidly time-varying due to the severe motion-induced Doppler effect. In fading channels with very high mobilities, the time variation of the channel over an OFDM symbol period results in a loss of subchannel orthogonality which leads to inter-channel interference due to power leakage among OFDM subcarriers. Consequently, estimation of the UWA channel parameters as well as channel equalization become critical issues in the design of UWA systems operating under such conditions. In particular, the channel estimation performance degrades substantially especially in the higher SNR region forming an error floor and consequently causing loss of diversity in SER performance. The diversity loss depends on the ratio of the power allocated for the pilot sequences and that allocated for the information signal. In the computer simulations carried out in our paper, 25 percent of the total power has been allocated to the pilot sequences and the corresponding diversity loss can be seen from the SER vs SNR and MSE vs SNR curves in Figs. 3–9. If this ratio is increased by allocating more power to the pilot sequences, an improvement in SER and MSE performance would be expected.

## VI. Conclusions

In this work we have presented a novel pilot assisted cascade channel estimation and equalization algorithm for AF cooperative relay based OFDM systems operating in sparse and cascade connected UWA channels. In this algorithm, the initial values of the unknown parameters of the UWA channel such as sparse channel gains, channel delays and the variances of the CGM densities, representing the prior information for the overall cascaded channels gains, are estimated by the MP algorithm and a new variance estimation algorithm. The prior pdf of the overall cascaded complex channel gains, from source-to-relay and relay-to-destination, have been modeled as CGMs and it has been shown that an exponential type of mixing pdf admits this representation exactly. We have further developed an efficient and low complexity novel iterative channel estimation algorithm based on the SAGE technique which makes use of the soft data estimates obtained from the equalizer output at each iteration step.

The performance of the proposed algorithms has been assessed by detailed computer simulations on a typical UWA channel model widely used in the literature. Thus, the proposed approach and the resulting channel estimation algorithm seem to be very promising for this kind of challenging and difficult channel estimation scenario involving cascade connected channels with non-Gaussian priors. Finally, the computer simulations have shown that the UWA channel is estimated very effectively and the proposed algorithm has excellent symbol error rate and channel estimation performance, and is robust to the effects of Doppler mismatch.

## Appendix
## Derivation of Optimal Estimates for Mixture Variances

The convex optimization problem in (72) can be solved by the Lagrangian method. Consider the expression

$$J(\mathbf{s},\mu)=\sum_{\ell=1}^{L}\left\{\frac{1}{s_\ell\beta\|\mathbf{e}_\ell\|^2}-\log(s_\ell\|\mathbf{e}_\ell\|^2)+\log K_0\left(\frac{2(\xi_\ell/s_\ell)^{1/2}}{\|\mathbf{e}_\ell\|^2}\right)\right\}-\mu\left(\sum_{\ell=1}^{L}s_\ell-1\right), \quad (73)$$

where $\mu>0$ is the Lagrange multiplier. The Kuhn-Tucker conditions for the optimal solution are

$$\begin{aligned}\frac{\partial J}{\partial s_\ell}&=0,\ \text{if } s_\ell>0,\\ \frac{\partial J}{\partial s_\ell}&\le 0,\ \text{if } s_\ell=0.\end{aligned} \quad (74)$$

Taking the derivative of $J$ with respect to $s_\ell$ and equating to zero we have,

$$-\frac{1}{s_\ell^2\beta\|\mathbf{e}_\ell\|^2}-\frac{1}{s_\ell}+\frac{(\beta\sqrt{\xi_\ell})^{-1}}{s_\ell\sqrt{s_\ell}}\frac{K_1\left(\frac{2(\xi_\ell/s_\ell)^{1/2}}{\|\mathbf{e}_\ell\|^2}\right)}{K_0\left(\frac{2(\xi_\ell/s_\ell)^{1/2}}{\|\mathbf{e}_\ell\|^2}\right)}-\mu=0. \quad (75)$$

The ratio $K_1(\cdot)/K_0(\cdot)$ in (75) can be well approximated as

$$\frac{K_1\left(\frac{2}{\beta}(s_\ell\xi_\ell)^{-1/2}\right)}{K_0\left(\frac{2(\xi_\ell/s_\ell)^{1/2}}{\|\mathbf{e}_\ell\|^2}\right)}\approxeq 1+(a-1)\sqrt{s_\ell},\ \text{for } 0\le s_\ell\le 1, \quad (76)$$

where, $a\triangleq K_1\left(\frac{2}{\|\mathbf{e}_\ell\|^2}(\xi_\ell/s_\ell)^{1/2}\right)/K_0\left(\frac{2}{\|\mathbf{e}_\ell\|^2}(\xi_\ell/s_\ell)^{1/2}\right)$. Using this approximation and after some algebra (75) can be expressed as

$$s_\ell^3+Ps_\ell^2+Qs_\ell+R=0, \quad (77)$$

where, $P\triangleq -2B/\mu; Q\triangleq (B^2-2A\mu)/\mu^2; R\triangleq (2AB-C^2)/\mu^2$ with $A\triangleq -1/(\beta\|\mathbf{e}_\ell\|^2)$; $B\triangleq a/(\beta\sqrt{\xi_\ell})-1$ and $C\triangleq 1/(\beta\sqrt{\xi_\ell})$. It can be easily verified that the cubic equation in (77) has only one real root and this root can be obtained as follows:

$$\widehat{s}_\ell\equiv\widehat{s}_\ell(\mu)=\left[\left(-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\right)^{1/3}+\left(-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\right)^{1/3}+\frac{P}{3}\right]^+,\quad 1\le\ell\le L, \quad (78)$$

where $[x]^+\triangleq\max(0,x)$ and $p\triangleq(3Q-P)/3; q\triangleq(2P^3-9PQ+27R)/27$.

Note that (78) is the optimal solution if the Lagrange multiplier $\mu$ satisfies the condition

$$\sum_{\ell=1}^{L}\widehat{s}_\ell(\mu)=1.$$

The inverse of the Lagrange multiplier can be regarded as a water level. Generally, the water level can be found by the binary search method [42].

## References


[1] L. Liu, S. Zhou, and J.-H. Cui, "Prospects and problems of wireless communication for underwater sensor networks," *Wireless Commun. Mobile Comput.*, pp. 977–994, 2008.

[2] R. Headrick, "Growth of underwater communication technology in the U.S. navy," *IEEE Commun. Mag.*, pp. 80–82, Jan. 2009.

[3] S. I. Al-Dharrab, M. Uysal, and T. M. Duman, "Cooperative underwater acoustic communications," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 146–153, Jul. 2013.

[4] K. Chen, M. Ma, E. Cheng, F. Yuan, and W. Su, "A survey on MAC protocols for underwater wireless sensor networks," *IEEE Commun. Surveys Tutor., Third Quarter*, vol. 16, no. 3, pp. 1433–1447, 2014.

[5] SUNRISE: Building the Internet of Underwater Things, Dec. 5, 2014 [Online]. Available: http://fp7-sunrise.eu/

[6] S. Roy, T. M. Duman, V. McDonald, and J. G. Proakis, "Highrate communication for underwater acoustic channels using multiple transmitters and space-time coding: Receiver structures and experimental results," *IEEE J. Ocean. Eng.*, vol. 32, no. 3, pp. 663–688, Jul. 2007.

[7] B. Li, J. Huang, S. Zhou, K. Ball, M. Stojanovic, L. Freitag, and P. Willett, "MIMO-OFDM for high-rate underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 34, no. 4, pp. 634–644, Oct. 2009.

[8] K. Tu, D. Fertonani, T. M. Duman, M. Stojanovic, J. G. Proakis, and P. Hursky, "Mitigation of intercarrier interference for OFDM over timevarying underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 36, no. 2, pp. 156–171, Apr. 2011.

[9] J. Huang, S. Zhou, and Z. Wang, "Performance results of two iterative receivers for distributed MIMO OFDM with large Doppler deviations," *IEEE J. Ocean. Eng.*, vol. 38, no. 2, pp. 347–357, Apr. 2013.

[10] L. Zhiqiang and T. C. Yang, "On overhead reduction in time-reversed OFDM underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 39, no. 4, pp. 788–800, Oct. 2014.

[11] *Cooperative Communications for Improved Wireless Network Transmission: Frameworks for Virtual Antenna Array Applications*, M. Uysal, Ed. Hershey, PA: IGI Global, 2009.
[12] C. Choudhuri and U. Mitra, "Capacity bounds and power allocation for underwater acoustic relay channels with ISI," in *Proc. ACM Int. Workshop on Underwater Networks*, Berkeley, CA, USA, 2009.
[13] C. Carbonelli, S. H. Chen, and U. Mitra, "Error propagation analysis for underwater cooperative multi-hop communications," *Elsevier J. Ad Hoc Net.*, vol. 7, no. 4, pp. 759–769, Jun. 2009.
[14] W. Zhang, M. Stojanovic, and U. Mitra, "Analysis of a linear multihop underwater acoustic network," *IEEE J. Ocean. Eng.*, vol. 35, no. 4, pp. 961–970, Oct. 2010.
[15] M. Vajapeyam, S. Vedantam, U. Mitra, J. C. Preisig, and M. Stojanovic, "Distributed space-time cooperative schemes for underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 33, no. 4, pp. 489–501, Oct. 2008.
[16] M. Rahmati and T. M. Duman, "Achieving delay diversity in asynchronous underwater acoustic (UWA) cooperative communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, 3, pp. 1367–1379, Feb. 2014.
[17] A. M. Jalil and A. Ghrayeb, "Distributed channel coding for underwater acoustic cooperative networks," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 848–856, Feb. 2014.
[18] N. Richard and U. Mitra, "Sparse channel estimation for cooperative underwater communications: A structured multichannel approach," in *Proc. IEEE Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 5300–5303.
[19] S. Choudhary and U. Mitra, "Sparse recovery from convolved output in underwater acoustic relay networks," in *Proc. Signal Inf. Process. Assoc. Ann. Summit Conf. (APSIPA ASC), Asia-Pacific*, Dec. 2012, pp. 1–8.
[20] H. Senol, E. Panayirci, M. Erdogan, and M. Uysal, "Channel estimation in underwater cooperative OFDM system with amplify-and-forward relaying," in *Proc. IEEE Global Commun. Conf.*, Anaheim, CA, USA, Dec. 2012.
[21] T. Kang and R. A. Iltis, "Iterative carrier frequency offset and channel estimation for underwater acoustic OFDM systems," *IEEE J. Sel. Area. Commun*, vol. 26, no. 9, pp. 1650–1661, Dec. 2008.
[22] P. C. Carrascosa and M. Stojanovic, "Adaptive channel estimation and data detection for underwater acoustic MIMO-OFDM systems," in *Proc. IEEE Conf., Acoust., Speech Signal Process.*, Mar. 2010, pp. 635–646.
[23] S. Kim, "Angle-domain frequency-selective sparse channel estimation for underwater MIMO-OFDM systems," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 685–687, May 2012.
[24] J. Ling, X. Tan, T. Yardibi, J. Li, M. L. Nordenvaad, H. He, and K. Zhao, "On Bayesian channel estimation and FFT-based symbol detection in MIMO underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 39, no. 1, pp. 59–73, Jan. 2014.
[25] H. Mheidat and M. Uysal, "Non-coherent and mismatched-coherent receivers for distributed STBCs with amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4060–4070, Nov. 2007.
[26] B. Gedik and M. Uysal, "Impact of imperfect channel estimation on the performance of amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1468–1479, Mar. 2009.
[27] H. Lanlan, W. Yik-Chung, M. Shaodan, N. Tung-Sang, and H. V. Poor, "Superimposed training-based channel estimation and data detection for OFDM amplify-and-forward cooperative systems under high mobility," *IEEE Trans. Signal Process*, vol. 60, no. 1, pp. 274–284, Jan. 2012.
[28] M. Seyfi, S. Muhaidat, and L. Jie, "Amplify-and-forward selection cooperation over Rayleigh fading channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 199–209, Jan. 2012.
[29] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, Mar. 2002.
[30] E. Panayirci, H. Senol, and H. V. Poor, "Joint channel estimation, equalization and data detection for OFDM systems in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4225–4238, Aug. 2010.
[31] H. Senol, E. Panayirci, and H. V. Poor, "Non-data-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4236–4253, Aug. 2012.
[32] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.
[33] 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 Process.*, vol. 58, no. 3, pp. 1708–1721, 2010.
[34] S. F. Mason, C. R. Berger, Z. Zhou, and P. Willett, "Detection, synchronization, and Doppler scale estimation with multicarrier waveforms in underwater acoustic communication," *IEEE J. Selec. Areas Commun.*, vol. 26, no. 9, pp. 1638–1649, 2008.
[35] B. Li, S. Zhou, M. Stojavonic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channles with nonuniform Doppler shift," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 198–209, Apr. 2008.
[36] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[37] S. F. Cotter and B. D. Rao, "Maximum-likelihood array processing in non-Gaussian mixtures," *IEEE Trans. Commun.*, vol. 40, no. 3, pp. 374–377, Mar. 2002.
[38] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-Gaussian mixtures," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3520–3535, Dec. 2000.
[39] J. A. Fessler and A. O. Hero, "Space-alternating generalized EM algorithm," *IEEE Trans. Signal Process.*, vol. 42, pp. 2664–2677, Oct. 1994.
[40] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Commun. Lett.*, vol. 9, no. 7, pp. 619–621, Jul. 2005.
[41] D. Teichroew, "The mixture of normal distributions with different variances," *Ann. Math. Statist.*, vol. 28, no. 2, pp. 510–512, 1957.
[42] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb. 2005.
[43] 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, Apr. 2013.
[44] S. Yerramalli and U. Mitra, "On optimal resampling for OFDM signaling in doubly-selective underwater acoustic channels," in *Proc. IEEE Oceans Conf.*, Quebec City, QC, Canada, Sep. 2008.

**Erdal Panayirci** (M'80–SM'91–F'03–LF'06) received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing, MI, USA.

Until 1998, he was with the Faculty of Electrical and Electronics Engineering, Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is Professor of Electrical Engineering and Head of the Electronics Engineering Department at Kadir Has University, Istanbul. His recent research interests include communication theory, synchronization, advanced signal processing techniques and their applications to wireless electrical, underwater and optical communications. He spent the academic year 2008–2009 with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, working on new channel estimation and equalization algorithms for high mobility WIMAX and LTE systems. He has published extensively in leading scientific journals and international conference and coauthored the book *Principles of Integrated Maritime Surveillance Systems* (Boston, MA: Kluwer Academic, 2000).

Dr. Panayirci has been the principal coordinator of a 6th and 7th Frame European project called Network of Excellent on Wireless Communications (NEWCOM) and WIMAGIC Strep project for two years, representing Kadir Has University. He was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the areas of Synchronizations and Equalizations during 1995–2000. He served as a Member of IEEE Fellow Committee during 2005–2008. He was the Technical Program Co-Chair of the IEEE International Conference on Communications (ICC2006) and the Technical Program Chair of the IEEE PIMRC, both held in Istanbul in 2006 and 2010, respectively. He is the Executive Vice Chairman of the upcoming IEEE Wireless Communications and Networking Conference (WCNC), Istanbul, in April 2014. Presently he is head of the Turkish Scientific Commission on Signals and Systems of International Union of Radio Science (URSI).

**Habib Senol** (S'04–M'07) received the B.S. and M.S. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, both in electronics engineering. He received the Ph.D. degree in electronics engineering from Isik University, Istanbul, in 2006.

He is currently on the faculty of Computer Engineering, Kadir Has University, Istanbul. He spent the academic year 2007–2008 at the Department of Electrical Engineering, Arizona State University, Tempe, USA, working on channel estimation and power optimization algorithms for wireless sensor networks. His recent research interests include communication theory, advanced signal processing techniques and their applications to wireless electrical, underwater acoustic, and optical communications.

**Murat Uysal** (M'02–SM'07) received the B.Sc. and the M.Sc. degree in electronics and communication engineering from Istanbul Technical University, Istanbul, Turkey, in 1995 and 1998, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, Texas, in 2001.

He is currently a Full Professor and Chair of the Department of Electrical and Electronics Engineering at Ozyegin University, Istanbul, Turkey. Prior to joining Ozyegin University, he was a tenured Associate Professor at the University of Waterloo, Canada, where he still holds an adjunct faculty position. His research interests are in the broad areas of communication theory and signal processing with a particular emphasis on the physical layer aspects of wireless communication systems in radio, acoustic, and optical frequency bands. He has authored approximately 200 journal and conference papers on these topics and received more than 4000 citations.

Prof. Uysal currently leads the EU COST Action OPTICWISE which is a European scientific network for interdisciplinary research activities in the area of optical wireless communications. He serves on the editorial boards of IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, Wiley *Wireless Communications and Mobile Computing (WCMC) Journal*, and Wiley *Transactions on Emerging Telecommunications Technologies (ETT)*. In the past, he served as an Editor for IEEE COMMUNICATIONS LETTERS, Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issues on Optical Wireless Communication (2009 and 2015), Chair of the Communication Theory Symposium of IEEE ICC 2007, Chair of the Communications and Networking Symposium of IEEE CCECE 2008, Chair of the Communication and Information Theory Symposium of IWCMC 2011, TPC Co-Chair of IEEE IWOW 2012, TPC Co-Chair of IEEE WCNC 2014, and General Chair of IEEE IWOW 2015. His distinctions include the Marsland Faculty Fellowship in 2004, NSERC Discovery Accelerator Supplement Award in 2008, University of Waterloo Engineering Research Excellence Award in 2010, Turkish Academy of Sciences Distinguished Young Scientist Award in 2011, and Ozyegin University Best Researcher Award in 2014 among others.

**H. Vincent Poor** (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, USA, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign, USA. Since 1990 he has been on the faculty at Princeton University, where he is the Michael Henry Strater University Professor of Electrical Engineering and Dean of the School of Engineering and Applied Science. His research interests are in the areas of stochastic analysis, statistical signal processing, and information theory, and their applications in wireless networks and related fields such as social networks and smart grid. Among his publications in these areas is the recent book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge, U.K.: Cambridge Univ. Press, 2014).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and a foreign member of Academia Europaea and the Royal Society. He is also a fellow of the American Academy of Arts and Sciences, the Royal Academy of Engineering (U.K.), and the Royal Society of Edinburgh. He received the Technical Achievement and Society Awards of the IEEE Signal Processing Society in 2007 and 2011, respectively. Recent recognition of his work includes the 2014 URSI Booker Gold Medal, the 2015 EURASIP Athanasios Papoulis Award, and honorary doctorates from several universities in Europe and Asia, including an honorary D.Sc. degree from Aalto University in 2014.