Sherman-Morrison Formula Aided Adaptive Channel Estimation for Underwater Visible Light Communication With Fractionally-Sampled OFDM

Junyu Chen, Lei Zhao, *Member, IEEE*, Ming Jiang, *Senior Member, IEEE*, and Zhiqiang Wu, *Senior Member, IEEE*

***Abstract*—In this paper, we investigate the channel estimation (CE) problem in an underwater visible light communication (UVLC) system invoking fractionally-sampled optical orthogonal frequency division multiplexing (FS-OOFDM). In practical UVLC scenarios, the communication links inevitably suffer from many stochastic channel effects including multi-path dispersion, scattering, turbulence, etc., and/or from the mobility of the transceiver, therefore resulting in a time-varying, location-dependent non-stationary propagation environment. Naturally, compared with the indoor visible light communication (VLC) scenario with a typical assumption being the time-flat channel models, it becomes a notable challenge for designing a low-complexity adaptive CE in the much more complicated UVLC scenarios. To solve this problem, we derive a class of Bayesian CE algorithms referred to as the Sherman-Morrison formula (SMF) based CE (SMF-CE), by exploiting the property of rank-one structure of the second-order channel statistics in the delay domain. Furthermore, an adaptive version of SMF-CE (ASMF-CE) can be obtained through updating the imperfect *a priori* knowledge of the channel and the noise's statistics. Simulation results demonstrate the superior performances of the proposed algorithms in comparison to existing methods, while maintaining a reduced computational complexity in comparison to the conventional linear minimum mean square error (LMMSE) scheme.**

***Index Terms*—Channel estimation (CE), fractional sampling (FS), optical orthogonal frequency division multiplexing (OOFDM), Sherman-Morrison formula (SMF), underwater visible light communication (UVLC).**

I. Introduction

With the growing demand for marine and oceanographic applications, wireless technologies play an increasingly important role and attract more interest from both the academia and the industry [1]. Compared with wired communication technologies such as fiber optics, wireless techniques are more cost-effective, offering a higher flexibility and more operational advantages especially in marine scenarios [1], [2]. Currently, acoustic communication is the most widely used technology targeting long-distance underwater information transmission [3]. However, it becomes restrictive for some practical applications, where a higher data rate is required, particularly in a short-to-medium coverage. In this case, underwater visible light communication (UVLC) attracts increasing attention, due to its capability of providing wider bandwidth, lower latency, better security and most importantly, much higher data rates [3]–[5], especially when combined with the optical orthogonal frequency division multiplexing (OOFDM) technologies [6]–[8].

Despite the promising benefits, many technical challenges still remain open for UVLC, for example channel estimation (CE) which is one of the essential function to enable broadband UVLC transmissions. So far, the studies on UVLC CE are very limited and how to acquire sufficiently accurate channel information is an open issue. As the most straightforward idea, one may propose to simply reuse existing CE methods developed for indoor visible light communication (VLC) or even for radio frequency (RF) systems in the UVLC scenario. Although such a direct migration may logically work, the actual achievable performance is far from optimal or could even degrade significantly, due to the specific disadvantageous effects of UVLC channels such as absorption, scattering and turbulence, which severely constrain UVLC systems' attainable link performance and coverage [9]–[11].

Compared with the indoor VLC channels, where temporal dispersion is often introduced by multi-path reflection, the multiple scattering produced by the rich types of underwater scatterers becomes the major source for the significant temporal dispersion in UVLC channels. Such a key difference implies that the existing CE schemes tailored for indoor VLC systems, for example the adaptive statistical Bayesian minimum mean square error CE (AS-BMMSE-CE) [12], are likely not optimal for UVLC systems. Inspired by the diagonalization approach [13], AS-BMMSE-CE is derived under the semi-orthogonality condition, which relies on a unique pilot pattern that is hard to be extended to multiple-input multiple-output (MIMO) scenarios. Moreover, the authors of [5], [9] point out that the non-turbulent channel impulse response (NT-CIR) in diffusive UVLC channel

Manuscript received November 3, 2019; revised March 11, 2020; accepted April 8, 2020. Date of publication April 20, 2020; date of current version May 13, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefano Tomasin. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1802300, in part by the General Project of National Natural Science Foundation of China under Grant 61771499, in part by the Guangdong Provincial Key Research and Development Project under Grant 2018B010114001, and in part by NSF Grant 1748494 and NSF Center for Surveillance Research. *(Corresponding author: Ming Jiang.)*

Junyu Chen, Lei Zhao, and Ming Jiang are with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510275, China (e-mail: chenjy256@mail2.sysu.edu.cn; zhaolei27@mail.sysu.edu.cn; jiangm7@mail.sysu.edu.cn).

Zhiqiang Wu is with the Department of Electrical Engineering, Wright State University, Dayton, OH 45435-0001 USA (e-mail: zhiqiang.wu@wright.edu).

Digital Object Identifier 10.1109/TSP.2020.2988355

links usually contains dense and energy-concentrated non-line-of-sight (NLOS) taps, due to interactions between photons and underwater suspended particles. As a result, the channel impulse response (CIR) of a diffusive UVLC channel is no longer sparse in the delay domain. Thus, some conventional CE methods, such as the compressive sensing (CS) based CEs [14], [15] or the sparse detection aided methods [16], may not be directly applicable to UVLC scenarios. Although Ma *et al.* [17] investigated the possibility of employing a CS-aided CE in UVLC systems, their channel model only assumes the sparsity of the reflective paths, without an accurate description of the scattering and turbulence effects of UVLC channels, which however may invalidate the sparsity assumption concerned.

In addition, note that the time-varying characteristics of UVLC channels mainly depend on two factors. The first aspect is the optical turbulence defined as the random variation of the refractive index [4], [5], which does not occur under indoor VLC channels. The second cause is the mobility effect of the user equipment (UE). Different from RF wireless channels, it is difficult to exactly model the time-varying characteristics in mobile UVLC scenarios, as the statistics of the channel state information (CSI) strongly depend on the specific moving path or location of the UE [18]. Due to such tricky issues, it is also practically difficult to acquire either the first-order statistics, namely the mean of the UVLC CIR, or the second-order statistics, namely the covariance and autocorrelation matrices of the UVLC CIR. Having said that, even if *a priori* channel information were made available, the relative movement between a UVLC transmitter and a UVLC receiver inevitably leads to spatial misalignment, thus naturally invalidates such information.

The problems mentioned above together make the design of an optimal or even a proper CE method for a UVLC system become more challenging than for its RF or indoor VLC counterparts. Naturally, it is highly desirable to design an efficient CE scheme by exploiting the specific properties of UVLC channels, such that the performance can be optimized for the underwater environment.

Against this background, in this paper we extend our preliminary work of [19] with the following main contributions:

1) Despite the increased path loss and delay spread, the rich diffusive links of UVLC channels can potentially alleviate the issue of light source tracking and alignment. Inspired by this property, we apply the time-domain (TD) fractionally-sampled based OOFDM (FS-OOFDM) to UVLC, which has not been investigated for optical underwater systems before. The new system is capable of reaping the multi-path diversity gain [20] offered by the diffusive UVLC channel links, hence effectively improving the achievable system performance.
2) A class of Bayesian estimators referred to as the Sherman-Morrison formula (SMF) based CE (SMF-CE) is derived utilizing the rank-one structure of the UVLC channel's second-order statistics. Compared with the optimal linear minimum mean square error CE (LMMSE-CE), the SMF-CEs avoid the matrix inversion that incurs an online cubic complexity. Particularly, the optimal SMF CE (OSMF-CE) can achieve the identical performance as the LMMSE-CE at a reduced-order complexity.

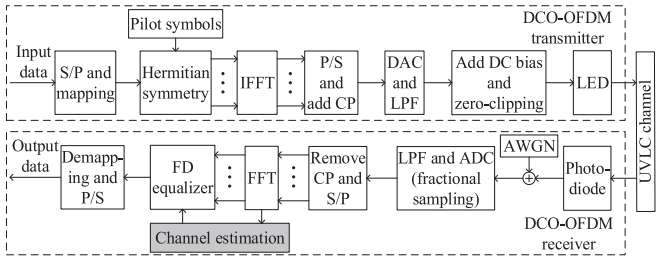


Fig. 1. The block diagram of the exemplified UVLC system.

3) Last but not least, we further propose an adaptive version of the SMF CE (ASMF-CE), where for the first time, specific approaches are designed for handling the imperfect *a priori* knowledge of the estimated statistics of both the UVLC channel and the noise. To preserve the rank-one structure of the estimated second-order statistics, we compute the largest eigenvalue and the corresponding eigenvector by taking advantage of a simple power iteration algorithm. Moreover, we devise a new common framework for the first-order statistics estimation, exploiting the channel correlations in both time and delay domains.

The organization of this paper is as follows. The system model is discussed in Section II. Then, the proposed class of SMF-CEs, subsuming its optimal and sub-optimal versions, is introduced in Section III. Furthermore, the adaptive versions of SMF-CEs are designed in Section IV, followed by the computational complexity analysis in Section V. Simulation results are provided in Section VI. Finally, our conclusions are given in Section VII.

Notations: Uppercase and lowercase bold variables denote matrices and vectors, respectively; $*$ and $\otimes$ denote the convolution and the Kronecker product operations, respectively; $\mathrm{vec}(\cdot), \mathrm{tr}(\cdot), (\cdot)^T, (\cdot)^H$ refer to the column-stacking vectorization, trace, transpose and Hermitian transpose operations, respectively; $(\cdot)^*$ represents the conjugate of $(\cdot)$; $\mathrm{E}\{\cdot\}$ is the expectation operation; $\lfloor\cdot\rfloor$ denotes the largest integer less than or equal to its argument; $\mathbf{I}_{\mathcal{L}}$ refers to an $\mathcal{L}\times\mathcal{L}$ identity matrix; $\mathrm{diag}(\mathbf{a})$ denotes the diagonal matrix with its diagonal entries being the elements of vector $\mathbf{a}$; $\|\cdot\|_2$ and $\|\cdot\|_\infty$ indicates the Euclidean norm and the maximum norm, respectively; $\Re(z)$ defines the real part of a complex number z. Furthermore, $[\mathbf{a}]_i$ denotes the i-th entry of vector $\mathbf{a}$, and $[\mathbf{A}]_{i,j}$ is the (i,j)-th entry of matrix $\mathbf{A}$. Finally, assuming that $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ with its associated eigenvalue λ, we refer to $(\lambda,\mathbf{x})$ as an eigenpair of $\mathbf{A}$.

II. System Model

In this work, we use light emitted diode (LED) transmitters and photodiode (PD) receivers, which operate under diffusive channel links that commonly exist in UVLC systems with relatively large divergence of source, large aperture, and wide field-of-view (FOV). Without loss of generality, we consider a UVLC system based on direct-current-biased optical OFDM (DCO-OFDM) [8], as depicted in Fig. 1, though the proposed CE schemes can be readily applicable to other types of OOFDM systems. Furthermore, a low UE mobility is assumed.

A. UVLC Channel Model

The typical characteristics of the UVLC channel model include absorption, scattering and turbulence effects. The absorption and scattering effects refer to the phenomenons, where photons suffer from energy loss, and interact with the molecules/atoms of the transmission medium, respectively [3]. On the other hand, the optical turbulence effect refers to the phenomenon commonly caused by the variation of the refraction index along the propagation path due to fluctuations in the temperature, salinity and density of the water [1].

In recent studies [5], [10], [21], a Monte Carlo (MC) approach has been utilized to investigate the absorption and scattering effects, which were verified through water-tank experiments [22]. Based on the numerical MC simulations [21], we may obtain the NT-CIR at the UE's n-th position by

$$c'_n(\tau) = \sum_{l=0}^{L_c-1} \alpha_{n,l}\delta(\tau - \tau_{n,l}), \tag{1}$$

where $\delta(x)$ is the unit impulse function, and L_c denotes the number of channel taps described by the position-dependent amplitude $\alpha_{n,l}$ and the delay $\tau_{n,l}$, where $l = 0, \ldots, L_c - 1$. More explicitly, the NT-CIR model of (1) may be viewed as the histogram of the received intensity versus propagation delay.

Furthermore, note that the parameters $L_c, \alpha_{n,l}, \tau_{n,l}$ in (1) characterize only the absorption and scattering effects, without reflecting the turbulence effect. Since the coherence time is typically at the order of 10^{-5}–10^{-2} seconds, which is much longer than the UVLC channel delay spread [5], we may assume that the same turbulence-induced fading coefficient, denoted by $\rho_n(t)$, is applied to all taps of a $c'_n(\tau)$ given in (1).

Following the above discussions, the overall CIR at the UE's n-th position may be formulated as

$$c_n(t, \tau) = \rho_n(t)c'_n(\tau), \tag{2}$$

where $\rho_n(t)$ follows the generalized Gamma (GG) distribution, which was experimentally validated for a wide range of scintillation index values spanning from weak to strong turbulence [23]. The GG distribution is defined as

$$f(\rho_n(t); a, d, p) = \frac{p}{a^d \Gamma\left(\frac{d}{p}\right)} \left[\rho_n(t)\right]^{d-1} e^{-\left[\frac{\rho_n(t)}{a}\right]^p}, \tag{3}$$

where the distribution parameters a, d and p should satisfy the condition [23]

$$\mathrm{E}\{\rho_n(t)\} = a\Gamma\left(\frac{d}{p} + \frac{1}{p}\right)\Gamma\left(\frac{d}{p}\right) = 1. \tag{4}$$

In this case, the scintillation index becomes [23]

$$\sigma_I^2 = \frac{\mathrm{E}\{\rho_n^2(t)\} - \mathrm{E}^2\{\rho_n(t)\}}{\mathrm{E}^2\{\rho_n(t)\}} = \frac{\Gamma(\frac{d}{p})\Gamma(\frac{d+2}{p})}{\Gamma^2(\frac{d+1}{p})} - 1. \tag{5}$$

To address the impact from realistic UVLC environment, the values of these distribution parameters for the GG distribution suggested by [23] were considered in our work.

Thus, we may define the *equivalent* CIR as $h_n(t, \tau) = c_n(t, \tau) * p(\tau)$, where $p(\tau)$ represents coefficients related to the pulse shaping and matched filter operations. In the sequel, CIR refers to the *equivalent* CIR unless otherwise stated. Based on the above definitions, $h_n(t, \tau)$ can be alternatively formulated by

$$h_n(t, \tau) = \rho_n(t)\mu_{h,n}(\tau), \tag{6}$$

where $\mu_{h,n}(\tau) = c'_n(\tau) * p(\tau)$.

B. FS-OOFDM Model

We consider a FS-OOFDM model [20] with an oversampling factor M and the sampling period $T_M = \frac{T_s}{M}$, where T_s is the sampling period of the conventional OOFDM model. A cyclic prefix is used to ensure the inter-symbol interference (ISI) is eliminated.

At the transmitter, each subcarrier is modulated by a complex pilot or data symbol $x_{n,k}$, where n and k denote the indices of the OOFDM symbol and the subcarrier, respectively. We assume that a comb-type pilot pattern is used in the system, where K_p pilot symbols having the unity amplitude at certain subcarrier indices $\mathcal{P} = \{k_i | 0 \le i \le K_p - 1\}$ are multiplexed with data symbols. The corresponding transmitted frequency-domain (FD) signal vector is denoted by $\mathbf{x}_n = [x_{n,0}, \ldots, x_{n,K-1}]^T \in \mathbb{C}^{K\times 1}$, where each element symbol is transmitted at one of the total K subcarriers in an OOFDM symbol. Note that in OOFDM systems [8], $x_{n,k}$ should fulfil the Hermitian symmetry (HS) constraint with respect to the $(\frac{K}{2})$-th subcarrier, namely $x_{n,k} = x^*_{n,K-k}$ $(1 \le k \le \frac{K}{2} - 1)$, where $x_{n,0} = x_{n,\frac{K}{2}} = 0$.

As pointed out by [24], slowly varying UVLC channels do not follow the complex Gaussian distribution. Nonetheless, they can still be partially described by their first- and second-order statistics. Since the coherence time of a UVLC channel is typically much longer than a normal OOFDM symbol duration, we assume that the CIR, which is impacted by both the mobility-induced and the turbulence-induced fading, is time invariant within one OOFDM symbol period.

Similar to [12], we use the OOFDM symbol index n to represent the index of the UE position, implying that the channel changes as the UE moves from one position to another. The FD signal received at the n-th position, $\mathbf{Y}_n = [\mathbf{y}_{n,0}, \ldots, \mathbf{y}_{n,M-1}] \in \mathbb{C}^{K\times M}$, can be modeled as

$$\mathbf{Y}_n = \mathbf{S}_n\mathbf{H}_n + \mathbf{V}_n, \tag{7}$$

where $\mathbf{H}_n = [\mathbf{H}_{n,0}, \ldots, \mathbf{H}_{n,M-1}] \in \mathbb{C}^{K\times M}$ is the FD channel matrix, $\mathbf{V}_n = [\mathbf{v}_{n,0}, \ldots, \mathbf{v}_{n,M-1}] \in \mathbb{C}^{K\times M}$ is the FD noise matrix, and $\mathbf{S}_n = \mathrm{diag}([s_{n,0}, \ldots, s_{n,K-1}]) \in \mathbb{C}^{K\times K}$ denotes the transmitted signals after the clipping operation. If a large direct-current (DC) bias B_{DC} of for example 13 dB is used, the clipping noise can be negligible [6], leading to

$$s_{n,k} = \begin{cases} \sqrt{K}B_{\mathrm{DC}}, & k = 0 \\ x_{n,k}, & 1 \le k \le K-1 \end{cases}. \tag{8}$$

More specifically, The channel transfer function (CTF) $\mathbf{H}_{n,m}$ can be viewed as the Fourier transform of $\mathbf{h}_{n,m} = [h_{n,m,0}, \ldots, h_{n,m,L-1}]^T \in \mathbb{R}^{L\times 1}_{\ge 0}$, where $h_{n,m,l} = h_n(t, lT_s + mT_M)$ denotes the l-th channel tap with the m-th oversampling offset at the n-th position, while L is the number of CIR taps.

Then we obtain

$$\mathbf{H}_n = \mathbf{F}_L \left[\mathbf{h}_{n,0}, \ldots, \mathbf{h}_{n,M-1}\right], \tag{9}$$

where $[\mathbf{F}_L]_{k,l} = e^{\frac{-j2\pi kl}{K}}, 0 \le k \le K-1, 0 \le l \le L-1$. Furthermore, $\mathbf{v}_{n,m} = [v_{n,m,0}, \ldots, v_{n,m,K-1}]^T \in \mathbb{C}^{K\times 1}$ is the complex additive white Gaussian noise (AWGN) complying with $\mathbf{v}_{n,m} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \sigma^2\mathbf{I}_K)$, where σ^2 is the FD noise power.

Substituting (9) into (7) and then vectorizing the FD received matrix, we have

$$\mathbf{y}_n = \widetilde{\mathbf{S}}_n \widetilde{\mathbf{F}}_L \mathbf{h}_n + \mathbf{v}_n, \tag{10}$$

where $\mathbf{y}_n = \mathrm{vec}\,(\mathbf{Y}_n) \in \mathbb{C}^{MK\times 1}$, $\widetilde{\mathbf{S}}_n = \mathbf{I}_M \otimes \mathbf{S}_n \in \mathbb{C}^{MK\times MK}$, $\widetilde{\mathbf{F}}_L = \mathbf{I}_M \otimes \mathbf{F}_L \in \mathbb{C}^{MK\times ML}$, $\mathbf{h}_n = \mathrm{vec}([\mathbf{h}_{n,0}, \ldots, \mathbf{h}_{n,M-1}]) \in \mathbb{R}_{\ge 0}^{ML\times 1}$, and $\mathbf{v}_n = \mathrm{vec}(\mathbf{V}_n) \in \mathbb{C}^{MK\times 1}$. The covariance matrix of $\mathbf{v}_n$ is denoted by $\mathbf{C}_\mathbf{v} = \mathrm{E}\{\mathbf{v}_n\mathbf{v}_n^H\}$, which has the form of $\mathbf{C}_\mathbf{v} = \sigma^2\bar{\mathbf{C}}_\mathbf{v}$ [20, (7)], and the $(m_1K + k_1, m_2K + k_2)$-th element of $\bar{\mathbf{C}}_\mathbf{v}$ is given by

$$[\bar{\mathbf{C}}_\mathbf{v}]_{m_1K+k_1, m_2K+k_2} = \frac{1}{K}\sum_{n_1=0}^{K-1}\sum_{n_2=0}^{K-1} p(\tau) e^{j\frac{2\pi}{K}(k_1n_2 - k_2n_1)}, \tag{11}$$

where $\{m_1, m_2\} = \{0, \ldots, M-1\}$, $\{k_1, k_2\} = \{0, \ldots, K-1\}$, and $\tau = (n_2 - n_1)T_s + (m_2 - m_1)T_M$. Moreover, we define the FD signal-to-noise ratio (SNR) as $\frac{P_t}{\sigma^2}$, where $P_t = \mathrm{E}\{|s_{n,k}|^2\}$.

Finally, a subcarrier-by-subcarrier maximum ratio combining (SS-MRC) approach is employed to combine the multi-path diversity, so as to acquire a gain similarly available in an equivalent single-input-multiple-output (SIMO) system. More details of this approach can be found in [20].

III. Sherman-Morrison Formula Based Channel Estimation

In this section, we propose a class of Bayesian channel estimators based on the SMF, exploiting the rank-one structure of the second-order channel statistics.

A. Overview of LMMSE-CE

For notation convenience, we use a superscript $(\cdot)^{(p)}$ to indicate the pilot symbols. Define $\mathbf{y}_n^{(p)} = \mathbf{\Pi}\mathbf{y}_n \in \mathbb{C}^{MK_p\times 1}$ as the received signal vector at pilot subcarriers, where $\mathbf{\Pi}$ is an $MK_p \times MK$ pilot index matrix with its (i,j)-th element being

$$[\mathbf{\Pi}]_{i,j} = \begin{cases} 1, & j = k_i + K\left\lfloor \frac{i}{K_p}\right\rfloor \\ 0, & \text{otherwise} \end{cases}, \tag{12}$$

where $0 \le i \le MK_p - 1, 0 \le j \le MK - 1$. With the aid of the least squares (LS) based CE, we obtain the channel estimates at pilot subcarriers as follows

$$\hat{\mathbf{H}}_n^{(p)} = (\widetilde{\mathbf{S}}_n^{(p)})^{-1}\mathbf{y}_n^{(p)} = \mathbf{\Phi}\mathbf{h}_n + \mathbf{z}_n, \tag{13}$$

where $\hat{\mathbf{H}}_n^{(p)}$ is the LS estimate of $\mathbf{H}_n^{(p)}$, $\mathbf{\Phi} = \mathbf{\Pi}\widetilde{\mathbf{F}}_L \in \mathbb{C}^{MK_p\times ML}$, $\mathbf{z}_n = (\widetilde{\mathbf{S}}_n^{(p)})^{-1}\mathbf{\Pi}\mathbf{v}_n \in \mathbb{C}^{MK_p\times 1}$ and $\widetilde{\mathbf{S}}_n^{(p)} = \mathbf{\Pi}\widetilde{\mathbf{S}}_n\mathbf{\Pi}^T \in \mathbb{C}^{MK_p\times MK_p}$. As pilot symbols have the unity amplitude, $\mathbf{z}_n$ is zero mean with the covariance matrix $\mathbf{C}_\mathbf{z} = \sigma^2\bar{\mathbf{C}}_\mathbf{z}$, where $\bar{\mathbf{C}}_\mathbf{z} = \mathbf{\Pi}\bar{\mathbf{C}}_\mathbf{v}\mathbf{\Pi}^T \in \mathbb{C}^{MK_p\times MK_p}$. Given the observation $\hat{\mathbf{H}}_n^{(p)}$, the objective of our work is to obtain a high-accuracy estimate of $\mathbf{h}_n$, denoted by $\hat{\mathbf{h}}_n$.

Based on (13), the classic LMMSE estimator has been proposed and widely used, which only depends on the first two moments of the channel PDF. With the mean vector $\boldsymbol{\mu}_{\mathbf{h},n} = [\boldsymbol{\mu}_{\mathbf{h},n,0}^T, \ldots, \boldsymbol{\mu}_{\mathbf{h},n,M-1}^T]^T \in \mathbb{C}^{ML\times 1}$, where $\boldsymbol{\mu}_{\mathbf{h},n,m} = [\mu_{h,n}(mT_M), \ldots, \mu_{h,n}(mT_M + (L-1)T_s)]^T \in \mathbb{C}^{L\times 1}$, and the covariance matrix $\mathbf{C}_{\mathbf{h},n} = \mathrm{E}\{\mathbf{\Delta}_{\mathbf{h},n}\mathbf{\Delta}_{\mathbf{h},n}^H\} = \mathrm{E}\{(\mathbf{h}_n - \boldsymbol{\mu}_{\mathbf{h},n})(\mathbf{h}_n - \boldsymbol{\mu}_{\mathbf{h},n})^H\} \in \mathbb{C}^{ML\times ML}$ being known *a priori*, the LMMSE estimator of the CIR is [25, Page 391]

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{LMMSE},n} &= \boldsymbol{\mu}_{\mathbf{h},n} + \mathbf{C}_{\mathbf{h},n}\mathbf{\Phi}^H(\mathbf{\Phi}\mathbf{C}_{\mathbf{h},n}\mathbf{\Phi}^H + \mathbf{C}_\mathbf{z})^{-1}\mathbf{\Delta}_{\mathbf{H},n} \\ &= \boldsymbol{\mu}_{\mathbf{h},n} + (\mathbf{C}_{\mathbf{h},n}^{-1} + \mathbf{\Phi}^H\mathbf{C}_\mathbf{z}^{-1}\mathbf{\Phi})^{-1}\mathbf{\Phi}^H\mathbf{C}_\mathbf{z}^{-1}\mathbf{\Delta}_{\mathbf{H},n},\end{aligned} \tag{14}$$

where $\mathbf{\Delta}_{\mathbf{H},n} = \hat{\mathbf{H}}_n^{(p)} - \Phi\boldsymbol{\mu}_{\mathbf{h},n}$, and we define $\boldsymbol{\upsilon}_n = \mathbf{\Phi}\mathbf{C}_{\mathbf{h},n}\mathbf{\Phi}^H + \mathbf{C}_\mathbf{z}$ for notation convenience later. Note that the second equality in (14) only holds for invertible $\mathbf{C}_{\mathbf{h},n}$. Furthermore, given the estimation error $\boldsymbol{\epsilon}_n = \mathbf{h}_n - \hat{\mathbf{h}}_{\mathrm{LMMSE},n}$, its covariance matrix is given by $\mathbf{C}_{\boldsymbol{\epsilon},n} = \mathrm{E}\{\boldsymbol{\epsilon}_n\boldsymbol{\epsilon}_n^H\} = \mathbf{C}_{\mathbf{h},n} - \mathbf{C}_{\mathbf{h},n}\mathbf{\Phi}^H\boldsymbol{\upsilon}_n^{-1}\mathbf{\Phi}\mathbf{C}_{\mathbf{h},n}$. The corresponding mean square error (MSE) is $J_{\mathrm{LMMSE}} = \mathrm{tr}\{\mathbf{C}_{\boldsymbol{\epsilon},n}\}$.

Remark 1: Note that the LMMSE estimator of (14) is not the optimal Bayesian estimator for UVLC channels, where $\mathbf{h}_n$ does not follow the Gaussian distribution. However, it is the optimal linear estimator in the MMSE sense and thus is still considered as a useful reference in practice. Having said that, LMMSE-CE requires the calculation of the $MK_p \times MK_p$ term $\boldsymbol{\upsilon}_n^{-1}$, which involves matrix inversion and thus imposes a cubic-order computational complexity, especially if the matrix has a high dimension. This motivates us to find a new solution for reducing the associated complexity, which would be extremely beneficial for applications in the highly resource-constrained UVLC environment.

B. The Optimal SMF-CE

In the sequel, we drop the time index t in $\rho_n(t)$ for notation convenience. Considering the fact that each tap of the CIR experiences the uniform turbulence fading in UVLC channels [5], we obtain

$$\begin{aligned}\mathbf{C}_{\mathbf{h},n} &= \mathrm{E}\{\mathbf{h}_n\mathbf{h}_n^H\} - \mathrm{E}\{\mathbf{h}_n\}\mathrm{E}\{\mathbf{h}_n\}^H \\ &= \mathrm{E}\{(\rho_n\boldsymbol{\mu}_{\mathbf{h},n})(\rho_n\boldsymbol{\mu}_{\mathbf{h},n})^H\} - \boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H \\ &= \frac{\mathrm{E}\{\rho_n^2\} - \mathrm{E}^2\{\rho_n\}}{\mathrm{E}^2\{\rho_n\}}\boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H,\end{aligned} \tag{15}$$

where we have exploited (4), and $\mathbf{h}_n = \rho_n\boldsymbol{\mu}_{\mathbf{h},n}$, which is the vectorized version of (6). Note that this is the UVLC channel's key property that differs from the indoor VLC channel. Then, inserting (5) into (15), we have

$$\mathbf{C}_{\mathbf{h},n} = \sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H. \tag{16}$$

According to (16), it can be noted that $\mathbf{C}_{\mathbf{h},n}$ depends on both the scintillation index and the first-order statistics of the UVLC

CIR. Moreover, another important observation from (16) is that the channel covariance matrix $\mathbf{C}_{\mathbf{h},n}$ is not of full column rank or, more precisely, it is of rank 1 due to the fact that it is the outer product of two nonzero vectors. Such a specific property implies that conventional CE schemes based on Bayesian-type estimators may not be directly applicable to the UVLC channels. For example, in this case, the diagonality and invertibility assumptions on $\mathbf{C}_{\mathbf{h},n}$ stated in [12] for indoor VLC channels no longer holds.

Remark 2: Due to the rank-one property of the UVLC channel covariance matrix, we cannot directly estimate the CIR with the aid of the second equality in (14). This inspires us of developing a new CE method particularly tailored for the UVLC channel, to exploit the rank-one structure of its channel covariance matrix, such that the complexity of the estimator may be reduced.

By inserting $\mathbf{C}_{\mathbf{h},n}$ of (16) into $\boldsymbol{\mho}_n$ defined in (14), we have

$$\boldsymbol{\mho}_n{}^{-1} = [\sigma_I^2(\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n})(\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n})^H + \mathbf{C}_{\mathbf{z}}]^{-1}, \tag{17}$$

which may be viewed as the inverse of a rank-one modification of $\mathbf{C}_{\mathbf{z}}$. If the FD noise power σ^2 is *a priori* known at the receiver, the inverse of $\boldsymbol{\mho}_n$ may be calculated by the following Lemma:

Lemma 1: (Sherman-Morrison formula [26]) If $\mathbf{A}$ is an $\mathcal{L}\times\mathcal{L}$ nonsingular square matrix, and if $\mathbf{u}$ and $\mathbf{v}$ are $\mathcal{L}\times 1$ vectors satisfying $1+\mathbf{v}^T\mathbf{A}\mathbf{u}\neq 0$, then the sum $\mathbf{A}+\mathbf{u}\mathbf{v}^T$ is nonsingular, and it holds that

$$(\mathbf{A}+\mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}. \tag{18}$$

Then through (18), we can simplify LMMSE-CE of (14) to the low-complexity OSMF-CE. Specifically, let us set $\mathbf{A}=\mathbf{C}_{\mathbf{z}}$, $\mathbf{u}=\sigma_I^2\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n}$, $\mathbf{v}=\boldsymbol{\Phi}^*\boldsymbol{\mu}_{\mathbf{h},n}^*$ and $\mathcal{L}=MK_p$. Then, applying Lemma 1 to (17) yields

$$\begin{aligned}\boldsymbol{\mho}_n{}^{-1} &= \mathbf{C}_{\mathbf{z}}^{-1} - \frac{\sigma_I^2\mathbf{C}_{\mathbf{z}}^{-1}\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H\boldsymbol{\Phi}^H\mathbf{C}_{\mathbf{z}}^{-1}}{1+\sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}^H\boldsymbol{\Phi}^H\mathbf{C}_{\mathbf{z}}^{-1}\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n}}\\ &= \frac{1}{\sigma^2}\left(\bar{\mathbf{C}}_{\mathbf{z}}^{-1} - \frac{\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\boldsymbol{\Phi}\}}\right)\\ &= \frac{1}{\sigma^2}\left(\bar{\mathbf{C}}_{\mathbf{z}}^{-1} - \frac{\mathbf{M}_1^H\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}\right),\end{aligned} \tag{19}$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are defined as

$$\begin{aligned}\mathbf{M}_1 &= \boldsymbol{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\in\mathbb{C}^{ML\times MK_p}\\ \mathbf{M}_2 &= \mathbf{M}_1\boldsymbol{\Phi} = \boldsymbol{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\boldsymbol{\Phi}\in\mathbb{C}^{ML\times ML},\end{aligned} \tag{20}$$

with $\mathbf{M}_2$ being a Hermitian positive definite matrix.

Next, by substituting (19) into the first equality in (14), we can obtain the proposed OSMF-CE by

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{OSMF},n} &= \boldsymbol{\mu}_{\mathbf{h},n} + \frac{1}{\sigma^2}\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\boldsymbol{\Delta}_{\mathbf{H},n}\\ &\quad - \frac{1}{\sigma^2}\frac{\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H\mathbf{M}_1^H\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}\boldsymbol{\Delta}_{\mathbf{H},n}\\ &= \boldsymbol{\mu}_{\mathbf{h},n} + \frac{1}{\sigma^2}\left(\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1 - \boldsymbol{\Theta}_n\right)\boldsymbol{\Delta}_{\mathbf{H},n},\end{aligned} \tag{21}$$

where we define $\boldsymbol{\Theta}_n = \frac{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}$. Using (16), we can expand $\boldsymbol{\Theta}_n$ as follows

$$\boldsymbol{\Theta}_n = \frac{\sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}(\sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}^H\mathbf{M}_2\boldsymbol{\mu}_{\mathbf{h},n})\boldsymbol{\mu}_{\mathbf{h},n}^H\mathbf{M}_1}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}. \tag{22}$$

Furthermore, observe that

$$\sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}^H\mathbf{M}_2\boldsymbol{\mu}_{\mathbf{h},n} = \mathrm{tr}\{\sigma_I^2\boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H\mathbf{M}_2\} = \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}. \tag{23}$$

Substituting (23) into (22) leads to the following equality

$$\boldsymbol{\Theta}_n = \frac{\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1. \tag{24}$$

Then, inserting (24) into (21), the OSMF-CE becomes

$$\hat{\mathbf{h}}_{\mathrm{OSMF},n} = \boldsymbol{\mu}_{\mathbf{h},n} + \frac{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}\boldsymbol{\Delta}_{\mathbf{H},n}, \tag{25}$$

which yields the MSE of OSMF-CE as

$$\begin{aligned}J_{\mathrm{OSMF}} &= \mathrm{tr}\{(\mathbf{h}_n-\hat{\mathbf{h}}_{\mathrm{OSMF},n})(\mathbf{h}_n-\hat{\mathbf{h}}_{\mathrm{OSMF},n})^H\}\\ &= \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\} - 2\frac{\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\mathbf{C}_{\mathbf{Hh},n}\}}{\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}\\ &\quad + \frac{\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\mathbf{C}_{\mathbf{H},n}\mathbf{M}_1^H\mathbf{C}_{\mathbf{h},n}\}}{(\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\})^2},\end{aligned} \tag{26}$$

where $\mathbf{C}_{\mathbf{Hh},n}=\mathrm{E}\{\boldsymbol{\Delta}_{\mathbf{H},n}\boldsymbol{\Delta}_{\mathbf{h},n}^H\}$ and $\mathbf{C}_{\mathbf{H},n}=\mathrm{E}\{\boldsymbol{\Delta}_{\mathbf{H},n}\boldsymbol{\Delta}_{\mathbf{H},n}^H\}$. Utilizing $\boldsymbol{\Delta}_{\mathbf{H},n}=\mathbf{z}_n+\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathbf{h},n}$, we expand $\mathbf{C}_{\mathbf{Hh},n}$ and $\mathbf{C}_{\mathbf{H},n}$ to

$$\begin{aligned}\mathbf{C}_{\mathbf{Hh},n} &= \mathrm{E}\left\{(\mathbf{z}_n+\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathbf{h},n})\boldsymbol{\Delta}_{\mathbf{h},n}^H\right\}\\ &= \mathrm{E}\left\{\mathbf{z}_n\boldsymbol{\Delta}_{\mathbf{h},n}^H\right\} + \boldsymbol{\Phi}\mathrm{E}\left\{\boldsymbol{\Delta}_{\mathbf{h},n}\boldsymbol{\Delta}_{\mathbf{h},n}^H\right\}\\ &= \mathbf{C}_{\mathbf{zh},n} + \boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n},\\ \mathbf{C}_{\mathbf{H},n} &= \mathrm{E}\left\{(\mathbf{z}_n+\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathbf{h},n})(\mathbf{z}_n+\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathbf{h},n})^H\right\}\\ &= \mathrm{E}\left\{\mathbf{z}_n\mathbf{z}_n^H\right\} + \mathrm{E}\left\{\mathbf{z}_n\boldsymbol{\Delta}_{\mathbf{h},n}^H\right\}\boldsymbol{\Phi}^H + \boldsymbol{\Phi}\mathrm{E}\left\{\boldsymbol{\Delta}_{\mathbf{h},n}\mathbf{z}_n^H\right\}\\ &\quad + \boldsymbol{\Phi}\mathrm{E}\left\{\boldsymbol{\Delta}_{\mathbf{h},n}\boldsymbol{\Delta}_{\mathbf{h},n}^H\right\}\boldsymbol{\Phi}^H\\ &= \mathbf{C}_{\mathbf{z}} + \mathbf{C}_{\mathbf{zh},n}\boldsymbol{\Phi}^H + \boldsymbol{\Phi}\mathbf{C}_{\mathbf{hz},n} + \boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H,\end{aligned} \tag{27}$$

where $\mathbf{C}_{\mathbf{zh},n}=\mathrm{E}\{\mathbf{z}_n\boldsymbol{\Delta}_{\mathbf{h},n}^H\}$ and $\mathbf{C}_{\mathbf{hz},n}=\mathbf{C}_{\mathbf{zh},n}^H$. Noting that $\mathbf{z}_n$ and $\boldsymbol{\Delta}_{\mathbf{h},n}$ are uncorrelated, we may assume $\mathbf{C}_{\mathbf{zh},n}=\mathbf{C}_{\mathbf{hz},n}=\mathbf{0}$. Thus, (27) can be simplified to

$$\begin{aligned}\mathbf{C}_{\mathbf{Hh},n} &= \boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n},\\ \mathbf{C}_{\mathbf{H},n} &= \mathbf{C}_{\mathbf{z}} + \boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n}\boldsymbol{\Phi}^H.\end{aligned} \tag{28}$$

By exploiting (20) and (28), we observe that

$$\begin{aligned}\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\mathbf{C}_{\mathbf{Hh},n}\} &= \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\boldsymbol{\Phi}\mathbf{C}_{\mathbf{h},n}\}\\ &= \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\mathbf{C}_{\mathbf{h},n}\}\\ &= \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}\cdot\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\},\\ \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\mathbf{C}_{\mathbf{H},n}\mathbf{M}_1^H\mathbf{C}_{\mathbf{h},n}\} &= \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_1\mathbf{C}_{\mathbf{z}}\mathbf{M}_1^H\mathbf{C}_{\mathbf{h},n}\}\\ &\quad + \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\mathbf{C}_{\mathbf{h},n}\}\\ &= (\sigma^2+\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\})\\ &\quad\cdot\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}\cdot\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\}.\end{aligned} \tag{29}$$

Incorporating (29) into (26) yields

$$J_{\mathrm{OSMF}} = \frac{\sigma^2 \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\}}{\sigma^2 + \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}. \tag{30}$$

From another perspective, (30) can also be considered as a function of σ_I^2 with the aid of (16), yielding

$$J_{\mathrm{OSMF}} = \frac{c_{n,1}\sigma^2\sigma_I^2}{\sigma^2 + c_{n,2}\sigma_I^2}, \tag{31}$$

where both $c_{n,1} = \boldsymbol{\mu}_{\mathbf{h},n}^H \boldsymbol{\mu}_{\mathbf{h},n}$ and $c_{n,2} = \boldsymbol{\mu}_{\mathbf{h},n}^H \mathbf{M}_2 \boldsymbol{\mu}_{\mathbf{h},n}$ are constants independent of the channel fading coefficient ρ_n and the noise variance σ^2.

As a brief summary, we point out that OSMF-CE of (25) shares the same optimality with LMMSE-CE of (14) among the family of linear CEs, since the former is derived from the form of the latter by only exploiting the rank-one property of the covariance matrix, which does not result in any loss of optimality. In other words, OSMF-CE can be viewed as a special version of LMMSE-CE customized for UVLC channels. However, while the LMMSE-CE has a high computational complexity due to the online matrix inversion operations, the proposed OSMF-CE only requires matrix/vector multiplications that result in a much lower complexity. This advantage is highly beneficial from the implementation perspective, especially in the challenging underwater environment where UVLC systems operate.

C. The Sub-optimal SMF-CE

In practice, only imperfect estimates of the mean vector $\boldsymbol{\mu}_{\mathbf{h},n}$ and the covariance matrix $\mathbf{C}_{\mathbf{h},n}$ are available. Such imperfect conditions greatly constrain the achievable performance of the Bayesian CEs, which are sensitive to the mismatch in the statistical knowledge [13]. Such an issue motivates us to adapt the OSMF-CE for striking a tradeoff between MSE performance and the availability of the *a priori* channel knowledge.

Let us consider a general form of the CE as

$$\hat{\mathbf{h}}_n = \mathbf{W}\hat{\mathbf{H}}_n^{(p)}, \tag{32}$$

where $\mathbf{W}$ is to be optimized. By minimizing the MSE, we can get the optimal version of $\mathbf{W}$ as

$$\begin{aligned}\mathbf{W}_{\mathrm{O}} &= \arg\min_{\mathbf{W}} \mathrm{E}\{\|\mathbf{h}_n - \hat{\mathbf{h}}_n\|_2^2\} \\ &= \mathbf{R}_{\mathbf{h},n}\boldsymbol{\Phi}^H(\boldsymbol{\Phi}\mathbf{R}_{\mathbf{h},n}\boldsymbol{\Phi}^H + \mathbf{C}_{\mathbf{z}})^{-1},\end{aligned} \tag{33}$$

where $\mathbf{R}_{\mathbf{h},n}$ is the autocorrelation matrix of $\mathbf{h}_n$. Interestingly, based on (16), $\mathbf{R}_{\mathbf{h},n}$ is also a rank-one matrix

$$\mathbf{R}_{\mathbf{h},n} = \mathbf{C}_{\mathbf{h},n} + \boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H = (\sigma_I^2 + 1)\boldsymbol{\mu}_{\mathbf{h},n}\boldsymbol{\mu}_{\mathbf{h},n}^H. \tag{34}$$

If we let $\mathbf{A} = \mathbf{C}_{\mathbf{z}}$, $\mathbf{u} = (\sigma_I^2+1)\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{h},n}$, $\mathbf{v} = \boldsymbol{\Phi}^*\boldsymbol{\mu}_{\mathbf{h},n}^*$, and $\mathcal{L} = MK_p$, then we may apply Lemma 1 to the inverse term of (33), resulting in

$$\begin{aligned}\mathbf{W}_{\mathrm{O}} &= \frac{1}{\sigma^2}\mathbf{R}_{\mathbf{h},n}\boldsymbol{\Phi}^H\left(\bar{\mathbf{C}}_{\mathbf{z}}^{-1} - \frac{\mathbf{M}_1^H\mathbf{R}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2 + \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}\right) \\ &= \frac{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2 + \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}.\end{aligned} \tag{35}$$

Substituting (35) into (32) yields the sub-optimal Sherman-Morrison formula based CE (SSMF-CE)

$$\hat{\mathbf{h}}_{\mathrm{SSMF},n} = \mathbf{W}\hat{\mathbf{H}}_n^{(p)}|_{\mathbf{W}=\mathbf{W}_{\mathrm{O}}} = \frac{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_1}{\sigma^2 + \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}\hat{\mathbf{H}}_n^{(p)}. \tag{36}$$

Similar to the OSMF-CE, we can also develop the MSE of SSMF-CE

$$J_{\mathrm{SSMF}} = \frac{\sigma^2\mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\}}{\sigma^2 + \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}} = \frac{c_{n,1}\sigma^2(\sigma_I^2+1)}{\sigma^2 + c_{n,2}(\sigma_I^2+1)} > J_{\mathrm{OSMF}}, \tag{37}$$

where J_{OSMF} is given in (31). This implies that the OSMF-CE always performs better than the SSMF-CE from the MSE perspective. Furthermore, we have the following property of the SSMF-CE:

Theorem 1: Given the perfect *a priori* knowledge on CIR, the performance of the SSMF-CE is asymptotically equivalent to that of the OSMF-CE, when σ^2 approaches zero.

Proof: To compare the performance of the OSMF-CE and SSMF-CE, we define the ratio η as follows

$$\eta = \frac{J_{\mathrm{OSMF}}}{J_{\mathrm{SSMF}}} = \frac{\sigma_I^2}{\sigma_I^2+1}\cdot\frac{\sigma^2 + \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}{\sigma^2 + \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}}. \tag{38}$$

Due to the positive definite property of $\mathbf{M}_2$, both $\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}$ and $\mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}$ are positive numbers, which satisfy

$$\frac{\mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}{\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}} = \frac{\sigma_I^2+1}{\sigma_I^2} > 1, \tag{39}$$

where we have exploited (23) and (34). Furthermore, noting $\mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\} > \mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}$ from (39), it can be observed that the derivative of η with respect to σ^2 is always negative. Thus, η is a decreasing function of σ^2. Particularly, when $\sigma^2 \to 0$, we obtain

$$\lim_{\sigma^2\to 0}\eta = \frac{\sigma_I^2}{\sigma_I^2+1}\cdot\frac{\mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\}}{\mathrm{tr}\{\mathbf{C}_{\mathbf{h},n}\mathbf{M}_2\}} = 1, \tag{40}$$

implying that $\sigma^2 \to 0 \Rightarrow J_{\mathrm{OSMF}} = J_{\mathrm{SSMF}}$. Hence, we may conclude that when the SNR becomes sufficiently high, the performance of the SSMF-CE is asymptotically equivalent to that of the OSMF-CE. The proof completes. ■

Note that comparing with the need to acquire both $\boldsymbol{\mu}_{\mathbf{h},n}$ and $\mathbf{C}_{\mathbf{h},n}$ in (25) by OSMF-CE, the SSMF-CE scheme of (36) relaxes the requirement for *a priori* information to $\mathbf{R}_{\mathbf{h},n}$ only. Furthermore, similar to OSMF-CE, SSMF-CE does not involve the matrix inversion operation either.

IV. Adaptive SMF-Based Channel Estimators

Both the OSMF/SSMF-CEs proposed in Section III require a sufficiently accurate *a prior* knowledge of CIR. If a good link alignment can be maintained for a UVLC UE, it might be unnecessary to frequently update the *a prior* knowledge of CSI due to the almost constant average optical power at the UE. In practical scenarios, however, link misalignment often occurs or the UE can move to another position. In such cases, the dynamic change of the *a priori* knowledge of CIR in space and time, which unfortunately is generally not known at the UE side, should be considered in the CE design.

To tackle the above-mentioned issue, in the sequel we propose the adaptive versions of the OSMF/SSMF CEs.

A. Noise Variance Estimation

The noise variance is required by the OSMF/SSMF-CEs. Although many efforts have been dedicated to noise variance estimation for conventional OFDM systems [27], [28], the existing solutions were not designed for FS-OOFDM, thus ignoring some specific features that may be exploited for further optimization. Aiming to solve such issues, we propose a simple yet efficient noise variance estimator tailored for the FS-OOFDM system introduced in Section II-B.

In DCO-OFDM systems, both the 0-th and $(\frac{K}{2})$-th subcarriers are not used due to the HS constraint. When adding the DC bias to the original signal, as given by (8), it may overwhelm the FD noise superimposed on the 0-th subcarrier. Therefore, we may choose to exploit the received signal on the $(\frac{K}{2})$-th subcarrier for noise variance estimation. Thanks to the FS technique, we can effectively average the multiple signal receptions on the $(\frac{K}{2})$-th subcarrier as the initial estimate, denoted by $\bar{\sigma}_n^2$, corresponding to the n-th position

$$\bar{\sigma}_n^2 = \frac{1}{M}\sum_{m=0}^{M-1} |[\mathbf{y}_{n,m}]_{\frac{K}{2}}|^2, \tag{41}$$

Assuming that the noise variance remains constant during the UE's movement, the estimated noise variance $\hat{\sigma}_n^2$ at the n-th position can be refined with the following rank-one update

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\bar{\sigma}_i^2 = \frac{n-1}{n}\hat{\sigma}_{n-1}^2 + \frac{1}{n}\bar{\sigma}_n^2, \tag{42}$$

where $\hat{\sigma}_n^2$ is an unbiased estimator of the FD noise power σ^2.

B. Imperfect Second-Order Statistics Estimation

As the first step, we obtain the maximum likelihood (ML) estimate of the CIR [24], [25] without any *a priori* knowledge

$$\hat{\mathbf{h}}_{\mathrm{ML},n} = \mathbf{W}_{\mathrm{ML}}\hat{\mathbf{H}}_n^{(p)}, \tag{43}$$

where $\mathbf{W}_{\mathrm{ML}} = (\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{\Phi})^{-1}\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}$. Substituting $\hat{\mathbf{H}}_n^{(p)}$ from (13) into (43) yields

$$\hat{\mathbf{h}}_{\mathrm{ML},n} = \mathbf{W}_{\mathrm{ML}}\mathbf{\Phi}\mathbf{h}_n + \mathbf{W}_{\mathrm{ML}}\mathbf{z}_n = \mathbf{h}_n + \mathbf{e}_n, \tag{44}$$

where

$$\mathbf{e}_n = \mathbf{W}_{\mathrm{ML}}\mathbf{z}_n = \left(\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{z}_n \tag{45}$$

denotes the ML estimation (MLE) error, with its covariance matrix given by

$$\mathbf{C}_{\mathbf{e}} = \mathrm{E}\{\mathbf{e}_n\mathbf{e}_n^H\} = \sigma^2\left(\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{\Phi}\right)^{-1} = \sigma^2\mathbf{M}_2^{-1}. \tag{46}$$

Then, we may approximate $\mathbf{C}_{\mathbf{e}}$ by substituting $\hat{\sigma}_n^2$ in (42) for σ^2 in (46), yielding

$$\hat{\mathbf{C}}_{\mathbf{e},n} = \hat{\sigma}_n^2\mathbf{M}_2^{-1}. \tag{47}$$

Now we need to either obtain both $\boldsymbol{\mu}_{\mathbf{h},n}$ and $\mathbf{C}_{\mathbf{h},n}$ for the OSMF-CE of (25), or get $\mathbf{R}_{\mathbf{h},n}$ for the SSMF-CE of (36). For non-stationary channels, both $\mathbf{C}_{\mathbf{h},n}$ and $\mathbf{R}_{\mathbf{h},n}$ can be estimated by applying the exponential forgetting window approach [24]. More specifically, the initial estimate $\bar{\mathbf{C}}_{\mathbf{h},n}$ can be recursively updated by exploiting the following strategy with a forgetting factor β_{C} [29]

$$\bar{\mathbf{C}}_{\mathbf{h},n} = \beta_{\mathrm{C}}\left[\bar{\mathbf{C}}_{\mathbf{h},n-1} + (1-\beta_{\mathrm{C}})\bar{\mathbf{\Delta}}_{\mathbf{h},n}\bar{\mathbf{\Delta}}_{\mathbf{h},n}^H\right], \tag{48}$$

where $\bar{\mathbf{\Delta}}_{\mathbf{h},n} = \hat{\mathbf{h}}_{\mathrm{ML},n} - \bar{\boldsymbol{\mu}}_{\mathbf{h},n-1}$ denotes the difference between the ML estimate vector $\hat{\mathbf{h}}_{\mathrm{ML},n}$ measured at the n-th UE position and the exponentially weighted moving average vector $\bar{\boldsymbol{\mu}}_{\mathbf{h},n-1}$ associated with the $(n-1)$-th UE position. Note that $\bar{\boldsymbol{\mu}}_{\mathbf{h},n}$ is recursively updated by [29]

$$\bar{\boldsymbol{\mu}}_{\mathbf{h},n} = \bar{\boldsymbol{\mu}}_{\mathbf{h},n-1} + (1-\beta_{\mathrm{C}})\bar{\mathbf{\Delta}}_{\mathbf{h},n}. \tag{49}$$

Similarly, the estimate $\bar{\mathbf{R}}_{\mathbf{h},n}$ can be recursively updated by using the exponential forgetting window approach [24], [30]

$$\bar{\mathbf{R}}_{\mathbf{h},n} = \beta_{\mathrm{R}}\bar{\mathbf{R}}_{\mathbf{h},n-1} + (1-\beta_{\mathrm{R}})\hat{\mathbf{h}}_{\mathrm{ML},n}\hat{\mathbf{h}}_{\mathrm{ML},n}^H, \tag{50}$$

where β_{R} is a forgetting factor.

Note that the rank-one structures of $\mathbf{C}_{\mathbf{h},n}$ and $\mathbf{R}_{\mathbf{h},n}$ are the prerequisites for OSMF-CE and SSMF-CE, respectively. However, the rank of $\bar{\mathbf{C}}_{\mathbf{h},n}$ in (48) or the rank of $\bar{\mathbf{R}}_{\mathbf{h},n}$ in (50) might be larger than 1 due to the residual errors in the ML estimates denoted by (43). Under this circumstance, we need to monitor the dominant eigenpair of $\bar{\mathbf{C}}_{\mathbf{h},n}$ and $\bar{\mathbf{R}}_{\mathbf{h},n}$ to compute their rank-one approximation, where the dominant eigenpair refers to the pair of the largest eigenvalue and its corresponding eigenvector.

A simple yet effective method to compute the dominant eigenpair is the power iteration algorithm [31], which has a linear convergence rate. To accelerate the convergence, a shifting technique can be employed [32], where the value of the shift associated with the optimal convergence rate is calculated based on the average of the second-largest and the smallest eigenvalues. However, it is somewhat difficult to determine an appropriate shift value in practical scenarios [32]. Therefore, no shift is applied in our implementation of the power iteration algorithm outlined in Algorithm 1, where $I_{\max}$ and ϵ denote the maximum number of iterations and the termination threshold, respectively.

Algorithm 1: The Power Iteration Algorithm.

Input: $\mathbf{A}$, $I_{\max}$, ϵ
Initialization: $\lambda = 0$, $\mathbf{x} = [1, \ldots, 1]^T$
for $i = 1, \ldots, I_{\max}$ **do**
  $\lambda' = \lambda$, $\mathbf{x}' = \mathbf{x}$
  Generate the next vector $\mathbf{y} = \mathbf{A}\mathbf{x}'$
  Update the largest eigenvalue $\lambda = \|\mathbf{y}\|_\infty$
  Update the corresponding eigenvector $\mathbf{x} = \frac{\mathbf{y}}{\|\mathbf{y}\|_\infty}$
  if $|\lambda - \lambda'| < \epsilon$ **and** $\|\mathbf{x} - \mathbf{x}'\|_2 < \epsilon$ **then**
    break
  end if
end for
Output: λ, $\mathbf{x}$

Denote the dominant eigenpairs of $\bar{\mathbf{C}}_{\mathbf{h},n}$ and $\bar{\mathbf{R}}_{\mathbf{h},n}$ as $(\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}, \mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}})$ and $(\lambda_{\bar{\mathbf{R}}_{\mathbf{h},n}}, \mathbf{u}_{\bar{\mathbf{R}}_{\mathbf{h},n}})$ respectively. Based on the Eckart-Young Theorem [33], the best rank-one approximation

Algorithm 2: The ASSMF-CE Algorithm.

Input: $n_{\max}$, β_{R}
Initialization: $\bar{\mathbf{R}}_{\mathbf{h},0} = \mathbf{0}_{ML\times ML}$, $\hat{\sigma}_0^2 = 0$
Determine $\mathbf{M}_1$ and $\mathbf{M}_2$ by (20)
for $n = 1, \ldots, n_{\max}$ **do**
  Execute (13) to obtain the LS estimate $\hat{\mathbf{H}}_n^{(p)}$
  Calculate (43) to obtain $\hat{\mathbf{h}}_{\mathrm{ML},n}$
  Update $\hat{\sigma}_n^2$ according to (41) and (42)
  Calculate $\bar{\mathbf{R}}_{\mathbf{h},n}$ by (50)
  Execute Algorithm 1 to obtain $(\lambda_{\bar{\mathbf{R}}_{\mathbf{h},n}}, \mathbf{u}_{\bar{\mathbf{R}}_{\mathbf{h},n}})$ of $\bar{\mathbf{R}}_{\mathbf{h},n}$
  Calculate $\hat{\mathbf{R}}_{\mathbf{h},n}$ by (51)
  Calculate $\hat{\mathbf{h}}_{\mathrm{ASSMF},n}$ by (52)
  $\hat{\mathbf{H}}_n = \widetilde{\mathbf{F}}_L \hat{\mathbf{h}}_{\mathrm{ASSMF},n}$
end for
Output: $\hat{\mathbf{H}}_n, n = 1, \ldots, n_{\max}$

of $\bar{\mathbf{C}}_{\mathbf{h},n}$ and $\bar{\mathbf{R}}_{\mathbf{h},n}$ are given by

$$\hat{\mathbf{C}}_{\mathbf{h},n} = \lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}} \mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}} \mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H, \hat{\mathbf{R}}_{\mathbf{h},n} = \lambda_{\bar{\mathbf{R}}_{\mathbf{h},n}} \mathbf{u}_{\bar{\mathbf{R}}_{\mathbf{h},n}} \mathbf{u}_{\bar{\mathbf{R}}_{\mathbf{h},n}}^H. \quad (51)$$

Replacing σ^2, $\mathbf{R}_{\mathbf{h},n}$ in (36) with $\hat{\sigma}_n^2$ in (42) and $\hat{\mathbf{R}}_{\mathbf{h},n}$ in (51), it yields our proposed adaptive SSMF-CE (ASSMF-CE)

$$\hat{\mathbf{h}}_{\mathrm{ASSMF},n} = \frac{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_1}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\}} \hat{\mathbf{H}}_n^{(p)}. \quad (52)$$

After some derivations, we can formulate its MSE as

$$\begin{aligned} J_{\mathrm{ASSMF}} = \mathrm{tr}\{\mathbf{R}_{\mathbf{h},n}\} &- 2\frac{\mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\mathbf{R}_{\mathbf{h},n}\}}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\}} \\ &+ \frac{\sigma^2 \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\}\mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\}}{(\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\})^2} \\ &+ \frac{\mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\mathbf{R}_{\mathbf{h},n}\mathbf{M}_2\hat{\mathbf{R}}_{\mathbf{h},n}\}}{(\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{h},n}\mathbf{M}_2\})^2}. \end{aligned} \quad (53)$$

We summarize the ASSMF-CE in Algorithm 2, where $n_{\max}$ is the number of OOFDM symbols or, equivalently, the number of UE positions.

C. A Common Framework for Imperfect First-order Statistics Estimation

For the OSMF-CE, we need the *a priori* knowledge of $\boldsymbol{\mu}_{\mathbf{h},n}$, too. In this subsection, we establish a common framework to keep track of $\boldsymbol{\mu}_{\mathbf{h},n}$ based on recent ML estimates.

Assumption 1: Let $\rho_n, n = 1, \ldots, n_{\max}$ denote the turbulence-induced fading coefficients of the UVLC channel when the UE moves to the n-th position. Under the condition of weak turbulence, the temporal correlation model [34], [35] is utilized to describe the correlation of the consecutive ρ_n values at different time instants and different positions.

Under this assumption, a linear multiple-input multiple-output (MIMO) filter [36] of length Q may be applied to obtain the estimate of $\boldsymbol{\mu}_{\mathbf{h},n}$ by

$$\hat{\boldsymbol{\mu}}_{\mathbf{h},n} = \sum_{q=0}^{Q-1} \mathbf{W}_{n,q} \hat{\mathbf{h}}_{\mathrm{ML},n-q}, \quad (54)$$

where $\mathbf{W}_{n,q} \in \mathbb{C}^{ML\times ML} (0 \le q \le Q-1)$ denotes the q-th weighting matrix at the UE's n-th position.

Remark 3: Note that $\hat{\boldsymbol{\mu}}_{\mathbf{h},n}$ in the AS-BMMSE estimator proposed by [12] can also be formulated as (54), where Q is replaced by $\omega_{\max}$, which is the maximum statistic window size, and $\mathbf{W}_{n,q} (0 \le q \le \omega_{\max} - 1)$ are restricted to be diagonal with their entries being

$$[\mathbf{W}_{n,q}]_{l+1,l+1} = \begin{cases} \frac{1}{\omega_l^n}, & 0 \le q \le \omega_l^n \\ 0, & \omega_l^n < q \le \omega_{\max} - 1 \end{cases}, \quad (55)$$

where ω_l^n denotes the statistic window size for the l-th tap at the n-th position. It has to select the optimal size $\omega_{l,\mathrm{opt}}^n$ for each tap l by using the variable statistic window (VSW) approach [12]. However, the formulation of $\mathbf{W}_{n,q}$ in (55) considers the characteristics of indoor VLC channels only, and does not fully exploit the temporal and scattering correlations of UVLC channels. Thus, the AS-BMMSE-CE is not optimal for UVLC channels. To improve the CE accuracy, we propose that the diagonality constraint for $\mathbf{W}_{n,q}$ is removed and the prerequisite of (55) is eliminated from (54). Through this way, the temporal and scattering correlations of the UVLC CIR exhibiting from the recent ML estimates can be efficiently exploited, thus yielding a quality-improved $\hat{\boldsymbol{\mu}}_{\mathbf{h},n}$.

After substituting σ^2, $\mathbf{C}_{\mathbf{h},n}$ and $\boldsymbol{\mu}_{\mathbf{h},n}$ in (25) with $\hat{\sigma}_n^2$, $\hat{\mathbf{C}}_{\mathbf{h},n}$ and $\hat{\boldsymbol{\mu}}_{\mathbf{h},n}$, respectively, we obtain the proposed adaptive OSMF-CE (AOSMF-CE) as

$$\hat{\mathbf{h}}_{\mathrm{AOSMF},n} = \hat{\boldsymbol{\mu}}_{\mathbf{h},n} + \frac{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_1}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}} \hat{\boldsymbol{\Delta}}_{\mathbf{H},n}, \quad (56)$$

where $\hat{\boldsymbol{\Delta}}_{\mathbf{H},n} = \hat{\mathbf{H}}_n^{(p)} - \boldsymbol{\Phi}\hat{\boldsymbol{\mu}}_{\mathbf{h},n}$ with $\boldsymbol{\Phi}$ given by (13). For notation simplicity, we define

$$\mathbf{A}_n = \frac{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_1}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}}, \mathbf{B}_n = \mathbf{I}_{ML} - \mathbf{A}_n\boldsymbol{\Phi}. \quad (57)$$

Then the residual CE error term of (56), denoted by $\hat{\boldsymbol{\epsilon}}_n$, can be written as

$$\hat{\boldsymbol{\epsilon}}_n = \mathbf{h}_n - \hat{\mathbf{h}}_{\mathrm{AOSMF},n} = \mathbf{h}_n - \mathbf{A}_n\hat{\mathbf{H}}_n^{(p)} - \mathbf{B}_n\hat{\boldsymbol{\mu}}_{\mathbf{h},n}. \quad (58)$$

Utilizing (13), we develop (58) to

$$\begin{aligned} \hat{\boldsymbol{\epsilon}}_n &= \mathbf{h}_n - \mathbf{A}_n(\boldsymbol{\Phi}\mathbf{h}_n + \mathbf{z}_n) - \mathbf{B}_n\hat{\boldsymbol{\mu}}_{\mathbf{h},n} \\ &= (\mathbf{I}_{ML} - \mathbf{A}_n\boldsymbol{\Phi})\mathbf{h}_n - \mathbf{A}_n\mathbf{z}_n - \mathbf{B}_n\hat{\boldsymbol{\mu}}_{\mathbf{h},n} \\ &= \mathbf{B}_n\hat{\boldsymbol{\Delta}}_{\mathbf{h},n} - \mathbf{A}_n\mathbf{z}_n, \end{aligned} \quad (59)$$

where we have

$$\hat{\boldsymbol{\Delta}}_{\mathbf{h},n} = \mathbf{h}_n - \hat{\boldsymbol{\mu}}_{\mathbf{h},n}. \quad (60)$$

Hence, based on (59), the MSE of AOSMF-CE becomes

$$\begin{aligned} J_{\mathrm{AOSMF}} &= \mathrm{tr}\{\mathrm{E}\{\hat{\boldsymbol{\epsilon}}_n \hat{\boldsymbol{\epsilon}}_n^H\}\} \\ &= \mathrm{tr}\{\mathbf{B}_n \mathbf{C}_{\hat{\mathbf{h}},n} \mathbf{B}_n^H\} + \mathrm{tr}\{\mathbf{A}_n \mathbf{C}_{\mathbf{z}} \mathbf{A}_n^H\} \\ &\quad - \mathrm{tr}\{\mathbf{A}_n \mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n} \mathbf{B}_n^H + (\mathbf{A}_n \mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n} \mathbf{B}_n^H)^H\}, \end{aligned} \quad (61)$$

where we define

$$\mathbf{C}_{\hat{\mathbf{h}},n} = \mathrm{E}\{\hat{\boldsymbol{\Delta}}_{\mathbf{h},n} \hat{\boldsymbol{\Delta}}_{\mathbf{h},n}^H\}, \mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n} = \mathrm{E}\{\mathbf{z}_n \hat{\boldsymbol{\Delta}}_{\mathbf{h},n}^H\}. \quad (62)$$

Note that $\hat{\boldsymbol{\Delta}}_{\mathbf{h},n}$ in (62) can be obtained by first substituting (44) into (54) and then by inserting (54) into (60), yielding

$$\hat{\boldsymbol{\Delta}}_{\mathbf{h},n} = \mathbf{h}_n - \sum_{q=0}^{Q-1} \mathbf{W}_{n,q} \mathbf{h}_{n-q} - \sum_{q=0}^{Q-1} \mathbf{W}_{n,q} \mathbf{e}_{n-q}. \quad (63)$$

On the other hand, we define the channel cross-correlation matrix between the $(n-q_1)$-th and $(n-q_2)$-th position as follows

$$\mathbf{R}_{\mathbf{h},n,|q_1-q_2|} = \mathrm{E}\{\mathbf{h}_{n-q_1} \mathbf{h}_{n-q_2}^H\}, \{q_1, q_2\} = 0, \ldots, Q-1. \quad (64)$$

In particular, $\mathbf{R}_{\mathbf{h},n,0}$ is the same as $\mathbf{R}_{\mathbf{h},n}$.

By substituting (63) into (62) and utilizing the definition in (64), we can transform (61) to

$$\begin{aligned} J_{\mathrm{AOSMF}} &= \mathrm{tr}\left\{\mathbf{B}_n \mathbf{R}_{\mathbf{h},n} \mathbf{B}_n^H\right\} + \mathrm{tr}\left\{\mathbf{A}_n \mathbf{C}_{\mathbf{z}} \mathbf{A}_n^H\right\} \\ &\quad - 2 \sum_{q=0}^{Q-1} \Re(\mathrm{tr}\{\mathbf{B}_n \mathbf{W}_{n,q} (\mathbf{R}_{\mathbf{h},n,q} - \delta_q \boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}) \mathbf{B}_n^H\}) \\ &\quad + \sum_{q_1=0}^{Q-1} \sum_{q_2=0}^{Q-1} \mathrm{tr}\left\{\mathbf{B}_n \mathbf{W}_{n,q_1} \mathbf{R}_{\mathbf{h},n,|q_1-q_2|} \mathbf{W}_{n,q_2}^H \mathbf{B}_n^H\right\} \\ &\quad + \sum_{q_1=0}^{Q-1} \sum_{q_2=0}^{Q-1} \delta_{q_1-q_2} \mathrm{tr}\left\{\mathbf{B}_n \mathbf{W}_{n,q_1} \mathbf{C}_{\mathbf{e}} \mathbf{W}_{n,q_2}^H \mathbf{B}_n^H\right\}, \end{aligned} \quad (65)$$

where

$$\boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}} = \mathbf{B}_n^{-1} \mathbf{A}_n \boldsymbol{\Phi} \mathbf{C}_{\mathbf{e}}, \quad (66)$$

and $\mathbf{C}_{\mathbf{e}}$ is given by (46), while δ_x is the indicator function that takes the value of 1 if and only if $x = 0$, and takes the value of 0 otherwise. The derivation details can be found in Appendix A.

Note that (65) can be viewed as a function of $\mathbf{W}_{n,q}$. In the sequel, we intend to solve $\mathbf{W}_{n,q}$ by minimizing (65), yielding

$$\mathbf{W}_{n,q}^{\mathrm{opt}} = \arg\min_{\{\mathbf{W}_{n,q}\}} J_{\mathrm{AOSMF}}, 0 \le q \le Q-1, \quad (67)$$

where $\mathbf{W}_{n,q}^{\mathrm{opt}}$ denotes the optimal version of $\mathbf{W}_{n,q}$ and can be obtained through the following theorem.

Theorem 2: Given the covariance matrix $\hat{\mathbf{C}}_{\mathbf{h},n}$, the optimal weighting matrices $\{\mathbf{W}_{n,q}^{\mathrm{opt}}\}(0 \le q \le Q-1)$ should satisfy

$$\sum_{q'=0}^{Q-1} \mathbf{W}_{n,q'}^{\mathrm{opt}} (\mathbf{R}_{\mathbf{h},n,|q'-q|} + \delta_{q'-q} \mathbf{C}_{\mathbf{e}}) = \mathbf{R}_{\mathbf{h},n,q} - \delta_q \boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}, \quad (68)$$

where $\boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}} = \frac{\sigma^2}{\hat{\sigma}_n^2} \hat{\mathbf{C}}_{\mathbf{h},n}$ can be derived from (66).

Proof: See Appendix B. ■

TABLE I
COMPUTATIONAL COMPLEXITIES OF VARIOUS CES

CE schemes	Computational complexity
1D RWF [37]	$\mathcal{O}(K_f \widetilde{K})$
MLE [39]	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + \widetilde{K} \log K)$
LMMSE [25], [39]	$\mathcal{O}(\widetilde{K}_p^3 + \widetilde{L} \widetilde{K}_p^2 + \widetilde{L}^2 \widetilde{K}_p + \widetilde{K} \log K)$
RLS [37], [38]	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + K_t^2 \widetilde{L} + \widetilde{K} \log K)$
AS-BMMSE [12]	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + \omega_{\max} \widetilde{L} + \widetilde{K} \log K)$
OSMF	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + \widetilde{L}^2 + \widetilde{K} \log K)$
SSMF	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + \widetilde{L}^2 + \widetilde{K} \log K)$
AOSMF	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + I_{\max} \widetilde{L}^2 + Q^3 \widetilde{L}^3 + \widetilde{K} \log K)$
ASSMF	$\mathcal{O}(\widetilde{K}_p \widetilde{L} + I_{\max} \widetilde{L}^2 + \widetilde{K} \log K)$

Note that for $q \neq 0$, the indicator function δ_q defined below (65) becomes 0, reducing the right-hand side of (68) to $\mathbf{R}_{\mathbf{h},n,q}$, which is equivalent to the Wiener-Hopf structure [36]. For $q = 0$, the right-hand side of (68) simplifies to $\mathbf{R}_{\mathbf{h},n,0} - \boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}$.

In order to obtain a closed-form formulation of $\mathbf{W}_{n,q}^{\mathrm{opt}}(0 \le q \le Q-1)$ in (68), for notation convenience, let us first define the $ML \times QML$ matrix $\mathbf{W}_n^{\mathrm{opt}} = [\mathbf{W}_{n,0}^{\mathrm{opt}}, \ldots, \mathbf{W}_{n,Q-1}^{\mathrm{opt}}]$, the $ML \times QML$ matrix $\boldsymbol{\Upsilon}_{\mathbf{h},n} = [\mathbf{R}_{\mathbf{h},n,0} - \boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}, \mathbf{R}_{\mathbf{h},n,1}, \ldots, \mathbf{R}_{\mathbf{h},n,Q-1}]$, and the $QML \times QML$ block Toeplitz matrix $\boldsymbol{\Psi}_{\mathbf{h},n}$ with its (i,j)-th block element being $\mathbf{R}_{\mathbf{h},n,|i-j|}$. Then, we may rewrite (68) as

$$\mathbf{W}_n^{\mathrm{opt}} (\boldsymbol{\Psi}_{\mathbf{h},n} + \mathbf{I}_Q \otimes \mathbf{C}_{\mathbf{e}}) = \boldsymbol{\Upsilon}_{\mathbf{h},n}. \quad (69)$$

Therefore, the optimal weighting matrices $\mathbf{W}_n^{\mathrm{opt}}$ are given by

$$\mathbf{W}_n^{\mathrm{opt}} = [\mathbf{W}_{n,0}^{\mathrm{opt}}, \ldots, \mathbf{W}_{n,Q-1}^{\mathrm{opt}}] = \boldsymbol{\Upsilon}_{\mathbf{h},n} (\boldsymbol{\Psi}_{\mathbf{h},n} + \mathbf{I}_Q \otimes \mathbf{C}_{\mathbf{e}})^{-1}. \quad (70)$$

In practical scenarios, we may replace $\mathbf{C}_{\mathbf{e}}$ in (70) and σ^2 in $\boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}$ with $\hat{\mathbf{C}}_{\mathbf{e},n}$ in (47) and $\hat{\sigma}_n^2$ in (42), respectively. Thus $\boldsymbol{\Delta}_{\mathbf{R}_{\mathbf{h},n}}$ is approximated by $\hat{\mathbf{C}}_{\mathbf{h},n}$. In addition, $\mathbf{R}_{\mathbf{h},n,q}(0 \le q \le Q-1)$ in $\boldsymbol{\Psi}_{\mathbf{h},n}$ of (70) may be replaced by $\hat{\mathbf{R}}_{\mathbf{h},n,q}$, which can be recursively updated by [24], [30]

$$\hat{\mathbf{R}}_{\mathbf{h},n,q} = \beta_{\mathrm{R}} \hat{\mathbf{R}}_{\mathbf{h},n-1,q} + (1-\beta_{\mathrm{R}}) \hat{\mathbf{h}}_{\mathrm{ML},n} \hat{\mathbf{h}}_{\mathrm{ML},n-q}^H. \quad (71)$$

Then, substituting $\mathbf{W}_{n,q}^{\mathrm{opt}}$ of (70) for $\mathbf{W}_{n,q}$ of (54), we have

$$\hat{\boldsymbol{\mu}}_{\mathbf{h},n} = \sum_{q=0}^{Q-1} \mathbf{W}_{n,q}^{\mathrm{opt}} \hat{\mathbf{h}}_{\mathrm{ML},n-q}, \quad (72)$$

which can then be applied to (56) to derive the estimated CIR $\hat{\mathbf{h}}_{\mathrm{AOSMF},n}$. Furthermore, the minimum MSE of AOSMF-CE defined in (65) can be calculated by substituting (70) into (65)

$$\begin{aligned} J_{\mathrm{AOSMF}}^{\min} &= \mathrm{tr}\left\{\mathbf{B}_n \mathbf{R}_{\mathbf{h},n} \mathbf{B}_n^H\right\} + \mathrm{tr}\{\mathbf{A}_n \mathbf{C}_{\mathbf{z}} \mathbf{A}_n^H\} \\ &\quad - \mathrm{tr}\left\{\mathbf{B}_n \boldsymbol{\Upsilon}_{\mathbf{h},n} (\boldsymbol{\Psi}_{\mathbf{h},n} + \mathbf{I}_Q \otimes \mathbf{C}_{\mathbf{e}})^{-1} \boldsymbol{\Upsilon}_{\mathbf{h},n}^H \mathbf{B}_n^H\right\}. \end{aligned} \quad (73)$$

Finally, we outline the AOSMF-CE scheme in Algorithm 3.

V. COMPUTATIONAL COMPLEXITY ANALYSIS

In Table I, we compare the asymptotic computational complexities of various CEs normalized to one OOFDM symbol duration, where for notation convenience we define $\widetilde{K} = MK$,

Algorithm 3: The AOSMF-CE Algorithm.

Input: $n_{\max}$, Q, β_{C}, β_{R}

Initialization: $\hat{\mathbf{h}}_{\mathrm{ML},n} = \mathbf{0}_{ML\times 1}, -(Q-2) \le n \le 0$, $\bar{\boldsymbol{\mu}}_{\mathbf{h},0} = \mathbf{0}_{ML\times 1}$, $\hat{\mathbf{R}}_{\mathbf{h},0,q} = \mathbf{0}_{ML\times ML}$, $0 \le q \le Q-1$, $\bar{\mathbf{C}}_{\mathbf{h},0} = \mathbf{0}_{ML\times ML}$

Determine $\mathbf{M}_1, \mathbf{M}_2$ by (20) and calculate $\mathbf{M}_2^{-1}$

for $n = 1, \ldots, n_{\max}$ **do**

  Execute (13) to obtain the LS estimate $\hat{\mathbf{H}}_n^{(p)}$

  Calculate (43) to obtain $\hat{\mathbf{h}}_{\mathrm{ML},n}$

  Update $\hat{\sigma}_n^2$ according to (41) and (42)

  Calculate $\hat{\mathbf{C}}_{\mathbf{e},n}$ by (47)

  for $q = 0, \ldots, Q-1$ **do**

    Update $\hat{\mathbf{R}}_{\mathbf{h},n,q}$ by (71)

  end for

  Calculate $\bar{\mathbf{C}}_{\mathbf{h},n}$ by (48)

  Execute Algorithm 1 to obtain $(\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}, \mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}})$ of $\bar{\mathbf{C}}_{\mathbf{h},n}$

  Calculate $\hat{\mathbf{C}}_{\mathbf{h},n}$ by (51)

  Calculate $\mathbf{W}_n^{\mathrm{opt}}$ by (70)

  Calculate $\hat{\boldsymbol{\mu}}_{\mathbf{h},n}$ by (72)

  Calculate $\hat{\mathbf{h}}_{\mathrm{AOSMF},n}$ by (56)

  $\hat{\mathbf{H}}_n = \widetilde{\mathbf{F}}_L \hat{\mathbf{h}}_{\mathrm{AOSMF},n}$

end for

Output: $\hat{\mathbf{H}}_n, n = 1, \ldots, n_{\max}$

$\widetilde{K}_p = MK_p$ and $\widetilde{L} = ML$. Other notations include the FD filter order K_f for the one-dimensional (1D) robust Wiener filtering (RWF) estimator [37] and the TD filter order K_t for the recursive least squares (RLS) estimator [37], [38].

The complexities of the various CE problems are outlined in Table I in terms of number of floating-point operations (FLOPs), which are represented by the highest-order terms of the corresponding polynomial functions. Note that LMMSE-CE, OSMF-CE and SSMF-CE assume perfect knowledge of the channel statistics, while other Bayesian schemes have to pay additional complexities for retrieving the channel statistics. Note from Table I that LMMSE-CE imposes a cubic-order complexity with respect to $\widetilde{K}_p$ in each OFDM symbol duration, while OSMF/SSMF-CEs only require a linear complexity scaled by $\widetilde{L}$. Hence, OSMF/SSMF-CEs have the same asymptotic complexity as MLE and a lower complexity than LMMSE-CE under the typical condition of $\widetilde{K}_p > \widetilde{L}$.

Compared with other estimators, AOSMF/ASSMF-CEs have superior performances as to be revealed in Section VI, though resulting in a quadratic order of computational complexity with respect to $\widetilde{L}$, denoted by the term $I_{\max}\widetilde{L}^2$, due to the power iteration operations specified by Algorithm 1. Furthermore, additional computations are necessary to track and approximate the first-order statistics of AOSMF-CE, where a cubic order of complexity with respect to $\widetilde{L}$ dominates the overall complexity of post-processing. Thus, noting that AOSMF-CE and ASSMF-CE have comparable performances at medium-to-high SNRs, as shown in Section VI, we conclude that given a modest value of $I_{\max}$, ASSMF-CE is more cost-efficient in striking a good tradeoff between performance and computational complexity.

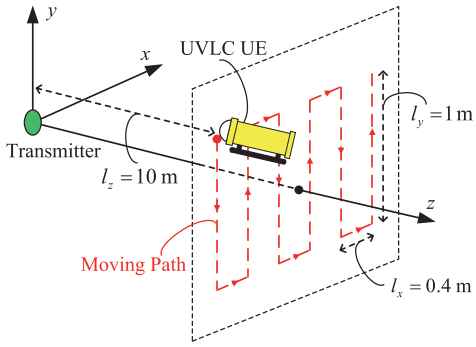


Fig. 2. Illustration of the moving route of UE.

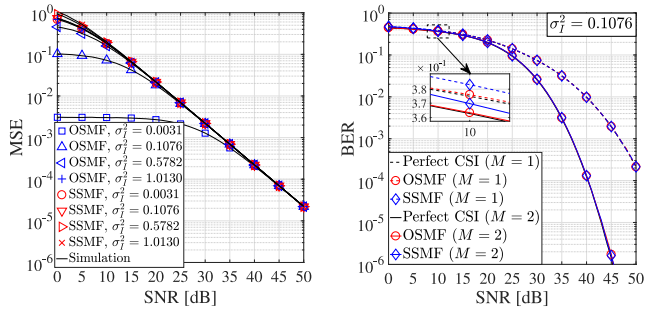
TABLE II
MAJOR PARAMETERS FOR THE UVLC CHANNEL MODEL AND THE SYSTEM CONFIGURATION

Entity	Parameter	Value
Channel	Transmitter-receiver distance, l_z	10 m
	Source light wavelength	532 nm
	Water type	Turbid harbor
	LED semi-half power angle	15°
	LED size	5 mm
	FOV	80°
	Aperture diameter	20 cm
	Photon weight threshold at the receiver	10^{-6}
	Scintillation index, σ_I^2	Various
System	Number of OOFDM subcarriers, K	512
	Cyclic prefix length, K_{cp}	32
	Number of pilot subcarriers, K_p	8
	Sampling period, T_s	2 ns
	Oversampling factor, M	2
	Pulse shaping filter span	4 samples
	Roll-off factor of pulse shaping filter	1
	DC bias	13 dB
	Signal constellation	16-QAM
	FD Equalizer	SS-MRC
UE	Velocity	2 m/s
	Spatial resolution of moving route	4 mm
	Start position	$(-1, 0.5, 10)$
	End position	$(1, 0.5, 10)$

VI. SIMULATION RESULTS

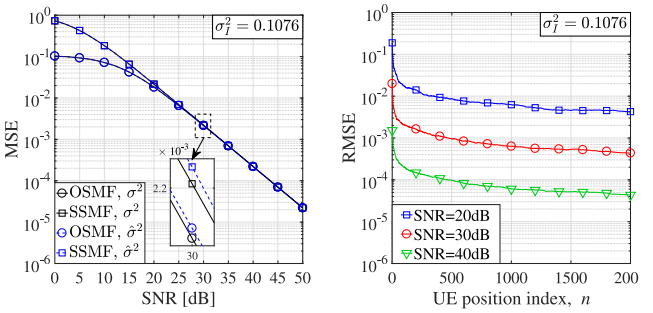
In this section, the simulation results of the proposed SMF-CE schemes are provided and compared with existing methods. The UVLC channel model was simulated based on the approach stated in Section II-A. As an example, we considered a UVLC scenario, where the transmitter was fixed at the origin of the virtual Cartesian coordinate system, while the UE slowly moved along a common zigzag route on the two-dimensional (2D) plane perpendicular to z-axis, as shown in Fig. 2. The major parameters used in the simulations are provided in Table II.

As the first test, we evaluate the theoretical MSEs of OSMF/SSMF-CEs under a wide range of σ_I^2 spanning from weak to strong turbulence, as specified by (30) and (37), respectively. For comparison of different CE schemes, we define the FD MSE metric as $\Gamma_n = \frac{\mathrm{E}\{\|\mathbf{H}_n - \hat{\mathbf{H}}_n\|^2\}}{MK} = \frac{\mathrm{E}\{\|\mathbf{h}_n - \hat{\mathbf{h}}_n\|^2\}}{M}$.



(a) MSE performance with different values of σ_I^2 (b) BER performance with different values of M

Fig. 3. The MSE and BER performances of OSMF/SSMF-CEs.



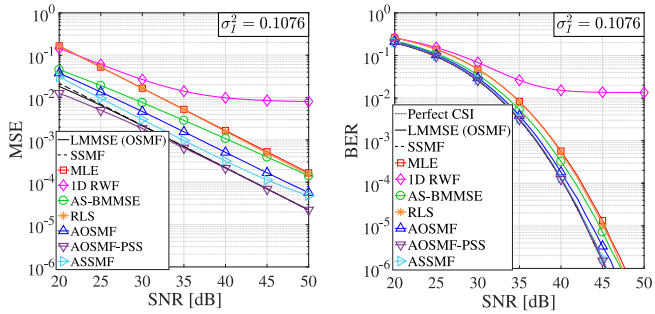
(a) MSE performance (b) RMSE performance

Fig. 4. The performances of the proposed noise variance estimator.

Fig. 3(a) shows that the simulated results of both OSMF/SSMF-CEs match the theoretical results, which validates our analytical derivations. Furthermore, we observe that the performance gap between OSMF-CE and SSMF-CE becomes smaller as the SNR increases. At high SNR levels, the MSE curves of both OSMF/SSMF-CEs converge, which is consistent with Theorem 1 that SSMF-CE has an asymptotically equivalent performance as OSMF-CE. An interesting phenomenon is that the MSE performances of OSMF-CE are more sensitive to σ_I^2 than those of SSMF estimators at low SNR levels. This behaviour can be interpreted by computing the first-order derivatives of (30) and (37) with respect to σ_I^2, where it shows that the former is larger than the latter given the same σ_I^2.

Fig. 3(b) shows the simulated bit error rate (BER) performances of the proposed schemes. The curves with $M = 1$ correspond to the conventional OOFDM system, which can be turned into a FS-OOFDM system by setting $M > 1$. From Fig. 3(b), we note that the FS-OOFDM system offers a large SNR gain with perfect CSI for both OSMF/SSMF-CEs, thanks to its exploitation of the multi-path diversity gains offered by the diffusive UVLC channels. Furthermore, the BER performances of OSMF/SSMF-CEs are very close to the benchmark scheme with perfect CSI.

From Fig. 4(a), it can be seen that both OSMF/SSMF-CEs with estimated $\hat{\sigma}^2$ perform almost the same as their counterparts with exact σ^2. To investigate the instantaneous estimation error when the UE moves, in Fig. 4(b) we measured the root MSE (RMSE) performances of the proposed noise variance estimator



(a) MSE performance (b) BER performance

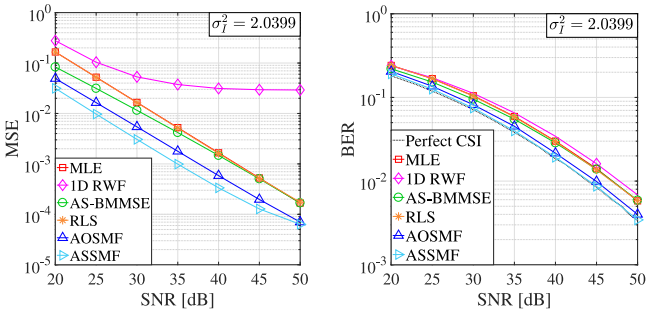
Fig. 5. The MSE and BER performances of various CEs.

at different positions along the entire route. It can be seen that the convergence rate approximately remains the same at various SNR levels. This property is indeed beneficial for Bayesian CEs, since the estimated noise variance, once obtained by the proposed method, may be applied upon a relatively long time or UE's moving distance regardless of the SNR level. In the remaining part of this section, we applied this noise variance estimator to all Bayesian CEs including AS-BMMSE [12], AOSMF and ASSMF CEs.

Fig. 5 compares the MSE and BER performances of the proposed CEs with their counterparts mentioned in Table I. The MSE and BER curves were plotted using the results averaged upon the entire UE moving route. Since perfect *a priori* knowledge of the channel is practically unavailable, LMMSE-CE and OSMF/SSMF-CEs, which are designed upon perfect CSI, only serve as theoretical lower bounds. In case of imperfect channel and noise statistics, AOSMF/ASSMF-CEs achieved better MSE and BER performances compared with other CE methods, as demonstrated by Fig. 5. In particular, compared with the conventional AS-BMMSE-CE scheme, AOSMF-CE and ASSMF-CE improved the system BER performance by about 1 dB and 2 dB at the BER of 10^{-6}, respectively.

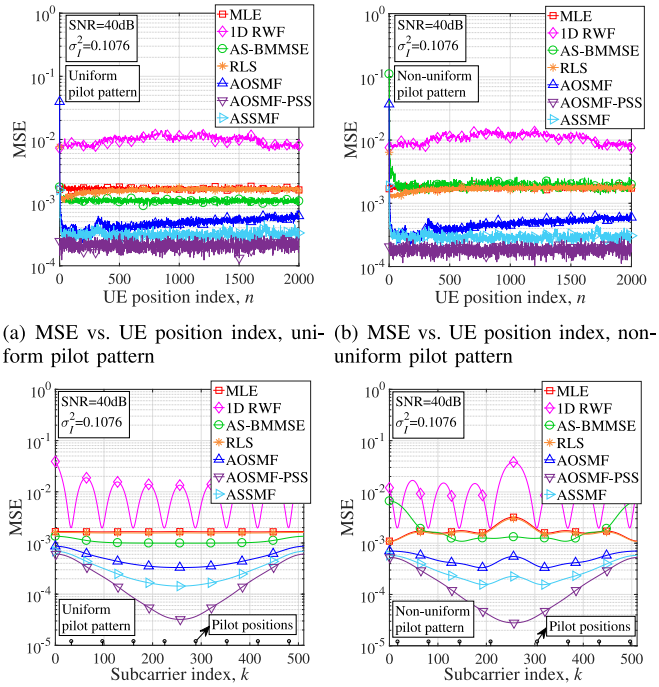
Interestingly, it can be noted that ASSMF-CE outperformed AOSMF-CE, despite that the latter is claimed to be the optimal linear CE method under the common framework discussed in Section IV-C. Such a phenomenon is mainly due to the fact that in AOSMF-CE, the estimation of the first-order channel statistics relies on the accuracy of the second-order channel statistics, as revealed by (70) and (72). To elaborate further, the second-order channel statistics utilized by AOSMF-CE are recursively obtained through (71), which may result in error propagation when its initial estimate is not accurate, hence inevitably degrading the quality of the estimated first-order channel statistics.

However, if the estimated channel cross-correlation matrix $\hat{\mathbf{R}}_{\mathbf{h},n,q}$ in (71) is replaced with its ideal version, yielding the reference scheme of AOSMF-CE with perfect second-order statistics (AOSMF-CE-PSS), then the estimated first-order channel statistics would become more reliable and the overall performance can be improved. This is evidenced by Fig. 5, where we can see that the AOSMF-CE-PSS curve is the closest to the lower MSE/BER bounds in medium-to-high SNRs. Note that the MSE performance gains of AOSMF-CE-PSS over LMMSE/OSMF-CE in the low-to-medium SNR region seen in



(a) MSE performance (b) BER performance

Fig. 6. The MSE and BER performances of various CEs for QPSK modulation and $\sigma_I^2 = 2.0399$.



(a) MSE vs. UE position index, uniform pilot pattern (b) MSE vs. UE position index, non-uniform pilot pattern

(c) MSE vs. subcarrier index, uniform pilot pattern (d) MSE vs. subcarrier index, non-uniform pilot pattern

Fig. 7. The MSE performances under different pilot patterns.

Fig. 5(a) was achieved, mainly due to the fact that the traditional Bayesian lower bound (TBLB) referring to (30) may be broken, if certain past channel information can be efficiently exploited. This observation was similar to that found in [12]. In any case, Fig. 5(b) indicates that both AOSMF/ASSMF-CEs provided a BER performance approaching the ideal benchmark scheme with perfect CSI.

Furthermore, we also compared the MSE and BER performances of the various CE schemes under a strong turbulence channel with $\sigma_I^2 = 2.0399$, as shown in Fig. 6. We can note that the MSE gains of AOSMF-CE and ASSMF-CE over AS-BMMSE-CE were about 4 dB and 6 dB, respectively, while the BER gains were about 2 dB and 3 dB, respectively. These results indicated that the proposed AOSMF/ASSMF-CEs were capable of attaining reasonable gains against their conventional counterparts under strong turbulence.

Remark 4: Comparing Fig. 5(b) and Fig. 6(b), we can see that a strong turbulence highly degraded the BER performance of the UVLC systems concerned. In practical scenarios, however, channel codes, transmitter beam-expander-and-collimator (BEC) [23] and receiver aperture averaging lens (AAL) [23] may be employed to mitigate the impact from strong turbulence.

Last but not least, in Fig. 7(a) and Fig. 7(c) we portray the MSE performances from the UE position and the subcarrier index perspectives, respectively, employing the unique uniform pilot pattern proposed for OOFDM-VLC systems invoking the HS operation [12]. We averaged the results of all subcarriers for every UE position in Fig. 7(a) and the results of all UE positions for every subcarrier in Fig. 7(c), respectively. From Fig. 7(a), we can see that the AOSMF-CE curve fluctuates significantly at the initial UE positions, where the quality of the estimated second-order channel statistics was not good. In contrast, AOSMF-CE-PSS and ASSMF-CE maintained a relatively stable performance along the UE's entire moving route. On the other hand, evaluating the systems from the FD in Fig. 7(c), where the vertical bars seen at the bottom of the figure indicate the subcarriers accommodating pilot symbols, it can be noticed that the proposed AOSMF/ASSMF-CEs achieved superior performances compared with other CEs.

Nevertheless, the unique uniform pilot pattern [12] may not be applicable in MIMO-UVLC or multi-cell UVLC systems, where the pilot indices have to be shifted from one LED or cell to another, in order to for example avoid inter-cell pilot collisions. Then, the shifted pilot patterns are no longer uniform and may result in performance degradation for conventional schemes, as revealed by Fig. 7(b) and Fig. 7(d), where an example non-uniform pilot pattern was employed. While the changed pilot pattern incurred a performance loss in the conventional CEs, however, it had little average impact on the proposed CEs, whose design does not depend on the pilot pattern used.

VII. Conclusion

In this paper, a specific form of Bayesian CEs, namely the family of SMF-CEs, were designed for the UVLC system equipped with FS-OOFDM. The SMF-CEs exploit the Sherman-Morrison formula to avoid the online matrix inversion involved in the conventional LMMSE-CE, thus greatly reducing the associated computational complexity. With the perfect channel and noise statistics, we derived the OSMF-CE which has the identical MSE performance as LMMSE-CE. Moreover, without the first-order statistical information, we proposed the SSMF-CE which asymptotically approaches OSMF-CE in terms of the MSE performance. We further developed the power iteration algorithm to make the rank-one approximation of imperfect covariance matrix of UVLC CIRs, which facilitated the design of a system framework for estimating the first-order statistics that are generally not available in practical systems. The proposed adaptive implementations of OSMF/SSMF-CEs, namely AOSMF/ASSMF-CEs, showed robustness to imperfect channel and noise statistics in comparison to a range of existing methods in the UVLC environment. Simulation results indicated that at the target BER of 10^{-6}, the AOSMF/ASSMF-CEs improved

the system BER performance by about 1 dB and 2 dB over the conventional AS-BMMSE-CE, respectively.

APPENDIX A
DERIVATION OF (65)

Utilizing (63) and (64), we may expand $\mathbf{C}_{\hat{\mathbf{h}},n}$ and $\mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n}$ in (62) to

$$\begin{aligned}\mathbf{C}_{\hat{\mathbf{h}},n} = \mathbf{R}_{\mathbf{h},n} &- \sum_{q=0}^{Q-1}(\mathbf{R}_{\mathbf{h},n,q}\mathbf{W}_{n,q}^H + \mathbf{W}_{n,q}\mathbf{R}_{\mathbf{h},n,q}) \\ &- \sum_{q=0}^{Q-1}(\mathrm{E}\{\mathbf{h}_n\mathbf{e}_{n-q}^H\}\mathbf{W}_{n,q}^H + \mathbf{W}_{n,q}\mathrm{E}\{\mathbf{e}_{n-q}\mathbf{h}_n^H\}) \\ &+ \sum_{q_1=0}^{Q-1}\sum_{q_2=0}^{Q-1}\mathbf{W}_{n,q_1}\mathbf{R}_{\mathbf{h},n,|q_1-q_2|}\mathbf{W}_{n,q_2}^H \\ &+ \sum_{q_1=0}^{Q-1}\sum_{q_2=0}^{Q-1}\mathbf{W}_{n,q_1}\mathrm{E}\{\mathbf{e}_{n-q_1}\mathbf{e}_{n-q_2}^H\}\mathbf{W}_{n,q_2}^H \\ &+ \sum_{q_1=0}^{Q-1}\sum_{q_2=0}^{Q-1}\mathbf{W}_{n,q_1}\mathrm{E}\{\mathbf{h}_{n-q_1}\mathbf{e}_{n-q_2}^H\}\mathbf{W}_{n,q_2}^H \\ &+ \sum_{q_1=0}^{Q-1}\sum_{q_2=0}^{Q-1}\mathbf{W}_{n,q_1}\mathrm{E}\{\mathbf{e}_{n-q_1}\mathbf{h}_{n-q_2}^H\}\mathbf{W}_{n,q_2}^H \end{aligned} \tag{74}$$

and

$$\begin{aligned}\mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n} = \mathrm{E}\{\mathbf{z}_n\mathbf{h}_n^H\} &- \sum_{q=0}^{Q-1}\mathrm{E}\{\mathbf{z}_n\mathbf{h}_{n-q}^H\}\mathbf{W}_{n,q}^H \\ &- \sum_{q=0}^{Q-1}\mathrm{E}\{\mathbf{z}_n\mathbf{e}_{n-q}^H\}\mathbf{W}_{n,q}^H, \end{aligned} \tag{75}$$

respectively. Since $\mathbf{z}_n$ and $\mathbf{h}_n$ are uncorrelated, we have $\mathrm{E}\{\mathbf{z}_n\mathbf{h}_n^H\} = \mathbf{0}_{MK_p\times ML}$. Recalling from (45) that $\mathbf{e}_n$ is a linear transform of $\mathbf{z}_n$, we also have $\mathrm{E}\{\mathbf{e}_n\mathbf{h}_n^H\} = \mathbf{0}_{ML\times ML}$. In addition, we assume that $\mathbf{z}_n$ is independent for different values of n, resulting in $\mathrm{E}\{\mathbf{e}_i\mathbf{e}_j^H\} = \mathbf{0}_{ML\times ML}(i \neq j)$ and $\mathrm{E}\{\mathbf{z}_i\mathbf{e}_j^H\} = \mathbf{0}_{ML\times ML}(i \neq j)$.

Using the above conditions, we may develop (74) to

$$\begin{aligned}\mathbf{C}_{\hat{\mathbf{h}},n} = \mathbf{R}_{\mathbf{h},n} &- \sum_{q=0}^{Q-1}\mathbf{R}_{\mathbf{h},n,q}\mathbf{W}_{n,q}^H - \sum_{q=0}^{Q-1}\mathbf{W}_{n,q}\mathbf{R}_{\mathbf{h},n,q} \\ &+ \sum_{q_1=0}^{Q-1}\sum_{q_2=0}^{Q-1}\mathbf{W}_{n,q_1}(\mathbf{R}_{\mathbf{h},n,|q_1-q_2|} + \delta_{q_1-q_2}\mathbf{C}_{\mathbf{e}})\mathbf{W}_{n,q_2}^H, \end{aligned} \tag{76}$$

and then simplify (75) to

$$\begin{aligned}\mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n} &= -\mathrm{E}\{\mathbf{z}_n\mathbf{e}_n^H\}\mathbf{W}_{n,0} \\ &= -\mathbf{C}_{\mathbf{z}}\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{\Phi}\left(\mathbf{\Phi}^H\bar{\mathbf{C}}_{\mathbf{z}}^{-1}\mathbf{\Phi}\right)^{-1}\mathbf{W}_{n,0} \\ &= -\mathbf{\Phi}\mathbf{C}_{\mathbf{e}}\mathbf{W}_{n,0}, \end{aligned} \tag{77}$$

where we also exploit (45) and (46). Thus, the term $\mathbf{A}_n\mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n}\mathbf{B}_n^H$ in (61) can be rewritten as

$$\mathbf{A}_n\mathbf{C}_{\mathbf{z}\hat{\mathbf{h}},n}\mathbf{B}_n^H = -\mathbf{B}_n\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}}\mathbf{W}_{n,0}\mathbf{B}_n^H, \tag{78}$$

where

$$\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}} = \mathbf{B}_n^{-1}\mathbf{A}_n\mathbf{\Phi}\mathbf{C}_{\mathbf{e}}. \tag{79}$$

Finally, by substituting (76) and (78) into (61), we arrive at (65).

APPENDIX B
PROOF OF THEOREM 2

Proof: Taking the partial derivative of (65) with respect to $\mathbf{W}_{n,q}^*$ for $0 \le q \le Q-1$, we have

$$\begin{aligned}\frac{\partial J_{\mathrm{AOSMF}}}{\partial \mathbf{W}_{n,q}^*} = &-\mathbf{B}_n^H\mathbf{B}_n\left(\mathbf{R}_{\mathbf{h},n,q} - \delta_q\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}}\right) \\ &+ \mathbf{B}_n^H\mathbf{B}_n\sum_{q'=0}^{Q-1}\mathbf{W}_{n,q'}(\mathbf{R}_{\mathbf{h},n,|q'-q|} + \delta_{q'-q}\mathbf{C}_{\mathbf{e}}). \end{aligned} \tag{80}$$

In order to simplify $\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}}$ in (80), we utilize $\mathbf{M}_2$ in (20) and $\hat{\mathbf{C}}_{\mathbf{h},n}$ in (51), and then rewrite $\mathbf{B}_n$ in (57) as

$$\mathbf{B}_n = \mathbf{I}_{ML} - \frac{\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H\mathbf{M}_2. \tag{81}$$

Next, we define $\mathbf{A} = \mathbf{I}_{ML}$, $\mathbf{u} = -\frac{\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}}{\hat{\sigma}_n^2+\mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}$, $\mathbf{v} = \mathbf{M}_2^*\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^*$, and $\mathcal{L} = ML$. Then, we apply Lemma 1 to (81), leading to

$$\begin{aligned}\mathbf{B}_n^{-1} &= \mathbf{I}_{ML} - \frac{-\frac{\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}}{\hat{\sigma}_n^2+\mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H\mathbf{M}_2}{1 - \frac{\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}}{\hat{\sigma}_n^2+\mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H\mathbf{M}_2\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}} \\ &= \mathbf{I}_{ML} + \frac{\lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H\mathbf{M}_2}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\} - \lambda_{\bar{\mathbf{C}}_{\mathbf{h},n}}\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}^H\mathbf{M}_2\mathbf{u}_{\bar{\mathbf{C}}_{\mathbf{h},n}}} \\ &= \mathbf{I}_{ML} + \frac{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2}{\hat{\sigma}_n^2}. \end{aligned} \tag{82}$$

Hence, we can exploit (20), (46), (57) and (82) to obtain

$$\begin{aligned}\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}} &= \left(\mathbf{I}_{ML} + \frac{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2}{\hat{\sigma}_n^2}\right)\cdot\frac{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_1\mathbf{\Phi}\sigma^2\mathbf{M}_2^{-1}}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}} \\ &= \frac{\sigma^2\hat{\mathbf{C}}_{\mathbf{h},n}}{\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}} + \frac{\sigma^2\mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\}\hat{\mathbf{C}}_{\mathbf{h},n}}{\hat{\sigma}_n^2(\hat{\sigma}_n^2 + \mathrm{tr}\{\hat{\mathbf{C}}_{\mathbf{h},n}\mathbf{M}_2\})} \\ &= \frac{\sigma^2}{\hat{\sigma}_n^2}\hat{\mathbf{C}}_{\mathbf{h},n}. \end{aligned} \tag{83}$$

Finally, by setting (80) to zero for $0 \le q \le Q-1$, we arrive at (68), where $\mathbf{\Delta}_{\mathbf{R}_{\mathbf{h},n}}$ is given by (83).

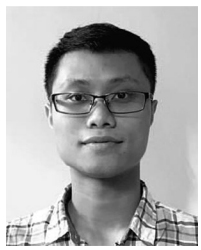
The proof of Theorem 2 completes. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. Julian Cheng from the University of British Columbia for his useful suggestions on writing in selected parts of this paper.

REFERENCES

[1] H. Kaushal and G. Kadcdoum, "Underwater optical wireless communication," *IEEE Access*, vol. 4, pp. 1518–1547, Apr. 2016.
[2] M. Elamassie, F. Miramirkhani, and M. Uysal, "Performance characterization of underwater visible light communication," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 543–552, Jan. 2019.
[3] Z. Zeng, S. Fu, H. Zhang, Y. Dong, and J. Cheng, "A survey of underwater optical wireless communications," *IEEE Commun. Surv. Tut.*, vol. 19, no. 1, pp. 204–238, Oct. 2017.
[4] M. V. Jamali, J. A. Salehi, and F. Akhoundi, "Performance studies of underwater wireless optical communication systems with spatial diversity: MIMO scheme," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1176–1192, Mar. 2017.
[5] M. V. Jamali, P. Nabavi, and J. A. Salehi, "MIMO underwater visible light communications: comprehensive channel study, performance analysis, and multiple-symbol detection," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8223–8237, Sep. 2018.
[6] J. Armstrong and B. J. C. Schmidt, "Comparison of asymmetrically clipped optical OFDM and DC-biased optical OFDM in AWGN," *IEEE Commun. Lett.*, vol. 12, no. 5, pp. 343–345, May 2008.
[7] J. Armstrong, "OFDM for optical communications," *J. Lightw. Technol.*, vol. 27, no. 3, pp. 189–204, 2009.
[8] S. D. Dissanayake and J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD systems," *J. Lightw. Technol.*, vol. 31, no. 7, pp. 1063–1072, Apr. 2013.
[9] S. Tang, Y. Dong, and X. Zhang, "Impulse response modeling for underwater wireless optical communication links," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 226–234, Jan. 2014.
[10] C. Gabriel, M.-A. Khalighi, S. Bourennane, P. Léon, and V. Rigaud, "Monte-Carlo-based channel characterization for underwater optical communication systems," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 1, pp. 1–12, Jan. 2013.
[11] O. Korotkova, N. Farwell, and E. Shchepakina, "Light scintillation in oceanic turbulence," *Waves Random Complex Media*, vol. 22, no. 2, pp. 260–266, Feb. 2012.
[12] X. Chen and M. Jiang, "Adaptive statistical Bayesian MMSE channel estimation for visible light communication," *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1287–1299, Nov. 2017.
[13] N. Shariati, E. Bjrnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.
[14] W. Ding, F. Yang, C. Pan, L. Dai, and J. Song, "Compressive sensing based channel estimation for OFDM systems under long delay channels," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 313–321, Jun. 2014.
[15] J. Meng, W. Yin, Y. Li, N. T. Nguyen, and Z. Han, "Compressive sensing based high-resolution channel estimation for OFDM system," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 15–25, Feb. 2012.
[16] X. Chen and M. Jiang, "Enhanced adaptive polar-linear interpolation aided channel estimation," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 693–696, Jun. 2019.
[17] X. Ma, F. Yang, S. Liu, and J. Song, "Channel estimation for wideband underwater visible light communication: A compressive sensing perspective," *Opt. Express*, vol. 26, no. 1, pp. 311–321, Jan. 2018.
[18] C. Gabriel, M. Khalighi, S. Bourennane, P. Léon, and V. Rigaud, "Misalignment considerations in point-to-point underwater wireless optical links," in *Proc. MTS/IEEE OCEANS*, Jun. 2013, pp. 1–5.
[19] J. Chen, L. Zhao, and M. Jiang, "Low complexity channel estimation for fractionally sampled underwater visible light communication," in *Proc. IEEE Int. Conf. Commun.*, Dublin, Ireland, Jun. 2020.
[20] C. Tepedelenlioglu and R. Challagulla, "Low-complexity multipath diversity through fractional sampling in OFDM," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3104–3116, Nov. 2004.
[21] W. C. Cox Jr., *Simulation, Modeling, and Design of Underwater Optical Communication Systems*. Raleigh, NC, USA: North Carolina State Univ., 2012.
[22] F. Hanson and S. Radic, "High bandwidth underwater optical communication," *Appl. Opt.*, vol. 47, no. 2, pp. 277–283, 2008.
[23] M. V. Jamali *et al.*, "Statistical studies of fading in underwater wireless optical channels in the presence of air bubble, temperature, and salinity random variations," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4706–4723, Oct. 2018.
[24] F. A. Dietrich and W. Utschick, "Pilot-assisted channel estimation based on second-order statistics," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1178–1193, Mar. 2005.
[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[26] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: SIAM, 2000.
[27] G. Ren, H. Zhang, and Y. Chang, "SNR estimation algorithm based on the preamble for OFDM systems in frequency selective channels," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2230–2234, Aug. 2009.
[28] V. Savaux, F. Bader, and Y. Louët, "A joint MMSE channel and noise variance estimation for OFDM/OQAM modulation," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4254–4266, Nov. 2015.
[29] T. Finch, "Incremental calculation of weighted mean and variance," *Univ. Cambridge*, vol. 4, no. 11-5, pp. 41–42, Feb. 2009.
[30] Y. Zheng, "A novel channel estimation and tracking method for wireless ofdm systems based on pilots and Kalman filtering," *IEEE Trans. Consum. Electron.*, vol. 49, no. 2, pp. 275–283, May 2003.
[31] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: John Hopkins Univ. Press, 2012.
[32] M. T. Heath, *Scientific Computing: An Introductory Survey*. Philadelphia, PA, USA: SIAM, 2002.
[33] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*. Philadelphia, PA, USA: SIAM, 1997, vol. 50.
[34] S. Tang, X. Zhang, and Y. Dong, "Temporal statistics of irradiance in moving turbulent ocean," in *Proc. MTS/IEEE OCEANS*, 2013, pp. 1–4.
[35] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*, vol. 152, Bellingham, WA, USA: SPIE press, 2005.
[36] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 593–602, Mar. 2005.
[37] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-user Communications, WLANs and Broadcasting*. Reading, MA, USA: Wiley, 2003.
[38] J. Zhang, H. Luo, and R. Jin, "Recursive MMSE channel estimation for MIMO-OFDM systems," in *Proc. 5th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2009, pp. 1–4.
[39] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.



Junyu Chen received the B.Eng. degree in communication engineering, in 2018, from the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China, where he is currently working towards the M.Eng. degree. His research interests include underwater visible light communication, deep learning, array signal processing, estimation theory, etc.

Lei Zhao (Member, IEEE) received the B.S. and M.S. degrees from Lanzhou University, Lanzhou, China, in 2008 and 2011, respectively, and the Ph.D. degree from the University of Nantes, Nantes, France, in 2014. From 2014 to 2017, he was with Microsystem and Terahertz Research Center, China Academy of Engineering Physics, Mianyang, China, where he was involved in terahertz radar and communication systems design. From 2017 to 2018, he was a System Engineer with Huawei Technologies, Co., Ltd., where he focused on 5G physical layer algorithm design. He is currently a Research Associate with Sun Yat-sen University, Guangzhou, China. His current research interests include visible light communications, space-air-ground communications, and 5G massive MIMO systems.

Ming Jiang (Senior Member, IEEE) received B.Eng. and M.Eng. degrees from South China University of Technology (SCUT), China, and Ph.D. degree from University of Southampton, UK, respectively, all in electronic engineering. Dr. Jiang has substantial international and industrial experience with Fortune 500 telecom companies. From 2006 to 2013, he had held key R&D or management positions at Samsung Electronics Research Institute (SERI), UK, Nortel Networks' R&D Centre, China, and telecom equipment maker New Postcom, China, where he actively participated in numerous collaborative projects across the EU, North America and Asia, contributing to algorithm and system research and standardization, as well as radio access and core network product designs. Since June 2013, he has been a Full Professor and Ph.D. Supervisor at Sun Yat-sen University, China, where he focuses on both fundamental research and technology transfer, and leads a number of national, provincial and industrial research projects. His research interest includes a wide range of areas in next-generation wireless mobile communications. He has coauthored or contributed to five Wiley books, 70+ papers, 80+ patents as well as 400+ 3GPP/IEEE standardization contributions. He received several Chinese council awards in 2011, including Innovative Leading Talents, Outstanding Experts, and Top Overseas Scholars.

Zhiqiang Wu (Senior Member, IEEE) received his B.S. from Beijing University of Posts and Telecommunications in 1993, M.S. from Peking University in 1996, and Ph.D. from Colorado State University in 2002, all in electrical engineering. He served as an Assistant Professor at Department of Electrical and computer Engineering, West Virginia University Institute of Technology, from 2003 to 2005. He joined Wright State University in 2005 where he currently serves as a Full Professor at the Department of Electrical Engineering. Dr. Wu's research has been supported by NSF, AFRL, ONR, AFOSR, and OFRN. He has also held visiting positions at Peking University, Harbin Engineering University, Guizhou Normal University and Tibet University.