

Optical Wireless Scattering Channel Estimation for Photon-Counting and Photomultiplier Tube Receivers

Chen Gong, Xiaoke Zhang, Zhengyuan Xu, and Lajos Hanzo

Abstract

Channel estimation is conceived for optical wireless scattering channels associated with Laser Diode transmitters and photon-counting/photomultiplier tube receivers. The proposed channel estimation approach consists of two stages, namely of the estimation of the channel tap second-order moments followed by the estimation of the channel taps based on the estimate of second-order moments. In the first stage, we provide the general framework of the moment estimation complemented by the conception of an estimation approach based on a sparse pilot structure, as well as by the analysis of the estimation error. We also propose the sparse pilot design and the associated low-complexity channel estimation, and prove the optimality of the proposed channel estimation. In the second stage, we conceive the channel tap estimation based on the eigenvalue decomposition of the matrix of estimated second-order moments, and analyze the associated performance. It is shown that as the length of the pilot sequence tends to infinity, the probability of having an estimation distortion above a threshold can be arbitrarily small. Simulation results show that the proposed sparse pilot sequence can lead to a smaller estimation error than the pilot design using random 0-1 bits.

Index Terms

Optical wireless scattering communication, Laser diode, channel estimation.

This work was supported by National Key Basic Research Program of China (Grant No. 2013CB329201), and Shenzhen Peacock Plan (No. 1108170036003286), and the Fundamental Research Funds for the Central Universities.

Chen Gong, Xiaoke Zhang, and Zhengyuan Xu are with Key Laboratory of Wireless-Optical Communications, Chinese Academy of Sciences, School of Information Science and Technology, University of Science and Technology of China, Hefei, Anhui 230026, China. Z. Xu is also with Shenzhen Graduate School, Tsinghua University, Shenzhen, China. Email addresses: zxiaoke@mail.ustc.edu.cn, {cgong821, xuzu}@ustc.edu.cn.

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K.. Email address: lh@ecs.soton.ac.uk.

I. INTRODUCTIONS

Optical wireless communication relies on a substantial license-free transmission bandwidth, whilst avoiding electromagnetic radiation. However, owing to its predominantly line-of-sight (LOS) propagation, the transmissions may be blocked by an obstacle between the transmitter and the receiver. Hence it is beneficial to exploit the scattered non-line-of-sight (NLOS) components [1], [2], where the transmission and the reception directions are not required to be perfectly aligned. Hence, the direct-link of NLOS optical communication has been extensively studied for example in [3], [4], [5], [6]. The relevant applications span from short-range atmospheric ultraviolet communication to sensing. In case of having a weak NLOS optical path, the received signal cannot be detected by the conventional waveform detector. The photon-level energy detector, such as the photon-counting receiver or a photomultiplier tube (PMT) receiver, needs to be employed. Explicitly, the photon-counting receiver counts the number of photons received; and the PMT receiver converts the received photons to electronic signals, while applying a certain amplification factor.

Existing contributions on single-input single-output optical wireless scattering channels predominantly focus on the channel capacity [7], [8], [9], as well as on the associated base-band digital signal processing [10], [11] and coded modulation aspects [12]. In a nutshell, those contributions address either the transmission capacity limits or the practical schemes that are capable of approaching those limits. Recently, a range of advanced transmission protocols have also been investigated, including the protocols designed for relay channels [13], [14], [15], [16], [17]. The channels' correlation and optimal linear receivers designed for single-input multi-output channels have been studied in [18] and [19], [20], respectively. These schemes are capable of significantly enhancing the achievable communication performance in the scenario of weak-link optical wireless scattering based communication. Similarly to the radio-frequency (RF) communications, the receiver side will suffer poor performance without estimating the channel state information before the signal detection/estimation. [Note that for the indoor visible light communication, the channel estimation approaches have been proposed in \[21\], \[22\], \[23\]. The received signal for the indoor visible light communication is characterized by continuous waveforms, while in this work the received signal is characterized by discrete photoelectrons.](#)

Let us consider the scenarios, where the scattering-induced time-domain dispersion is longer than the symbol duration. [Such channel model is typically adopted for the optical wireless communication in the ultra-violet spectrum, where the length of the NLOS pulse width](#)

broadening is larger than the symbol duration. Then, inter-symbol interference is imposed on the received signal. In this treatise, we consider a laser-diode (LD) based transmitter. The channel dispersion may bring phase change to the coherent optical signal of the LD. Since the energy detector such as the photon-counter or PMT cannot detect the phase of the received signal, the receiver only relies on the signal energy for channel estimation. Such channel estimation problem cannot be solved by utilizing the well-established techniques of RF communication, such as the linear minimum mean square error based steepest descent schemes, which rely on having a coherent receiver structure to retrieve the phase of the received signal. Note that the time-dispersion modeling has been addressed in [24] for the binary channel inputs and the associated capacity, and later in [25] for the channel estimation and symbol detection. These works focus on the non-coherent optical source, such as the light emitting diode, where the channel dispersion only brings the signal intensity change instead of the phase change. Thus the received signal model in this work is fundamentally different from that considered in [24], [25].

Against this background, our new contribution is that we conceive a novel channel estimation scheme for the energy-based receivers, where the phase change of the coherent optical signal due to the channel dispersion is considered. The channel estimation is carried out using a two-stage procedure. Explicitly, first the second-order moments of the channel taps are estimated, before estimating the channel taps themselves. The main contributions are outlined as follows.

- For the first stage, we propose a general framework applicable to both photon-counting and PMT receivers. We also propose the sparse pilot design and the associated low-complexity channel estimation, and prove the optimality of the proposed channel estimation.
- For the second stage, we propose a channel tap estimator based on the eigenvalue decomposition of the matrix of the estimated second-order moments obtained from the first stage.
- We analyze the estimation error of both stages. It is shown that as the length of the pilot sequence tends to infinity, the probability of having a tap estimation error above any positive threshold can be rendered arbitrarily small.

Simulations are also conducted both for a random 0-1 bit pilot sequence and for the proposed sparse pilot sequence. We will demonstrate that the sparse pilot sequence advocated leads to a smaller estimation error compared with the pilot sequence using random 0-1 bits, whilst imposing a lower computational complexity.

The remainder of this paper is organized as follows. In Section II, we provide our channel

model, formulate the channel estimation problem and the channel estimation ambiguity. In Section III, we set up the framework for the estimation of the second-order moments of the channel taps. In Section IV, we propose sparse pilots and estimate the moments of the channel taps. In Section V, we detail the estimation of channel taps, while our simulation results are provided in Section VI. Finally, Section VII concludes this paper.

II. SYSTEM MODEL

A. Signal Model for Energy-Type Receiver

Let us consider an NLOS optical wireless scattering communication system relying on a laser transmitter. Assume that the time-domain dispersion is longer than the symbol duration. In such a scenario, the inter-symbol interference (ISI) based channel model may be adopted.

The received signal consists of two components, the component of the dispersed signal and the component of the background radiation. Let $x(t)$ be the transmitted OOK symbol. The component of the dispersed signal, denoted as $y_{disp}(t)$, is given by

$$y_{disp}(t) = \sqrt{P} \sum_{n=0}^L h_n x(t - nT), \quad (1)$$

where T is the symbol duration; P is the transmission power; and $\{h_n\}_{n=0}^L$ represents the channel impulse response (CIR) taps. Note that since the channel dispersion may change the phase of the coherent optical signal, the channel taps h_n can be complex numbers. Such a scenario is fundamentally different from the previous study on the dispersion of non-coherent optical signal, where the channel taps are real numbers [24], [25].

We consider the energy-type detector, such as the photon-counting receiver and the PMT receiver. The energy of the received signal is the superposition of that of the dispersed signal $y_{disp}(t)$ and that of the background radiation. The energy of the dispersed signal within the symbol duration of $[mT, (m+1)T]$, denoted as E_m , is given by

$$E_m = \int_{mT}^{(m+1)T} |y_{disp}^2(t)| dt = P \int_{mT}^{(m+1)T} \left| \sum_{l=0}^L h_l x(t - lT) \right|^2 dt. \quad (2)$$

Assuming that the transmitted signal $x(t)$ is generated by modulating the data symbols x_n using a rectangular pulse $g(t)$, we have

$$x(t) = \sum_{n \geq 0} x_n g(t - nT), \quad (3)$$

where the pulse waveform is given by $g(t) = 1$ for $0 \leq t \leq T$ and $g(t) = 0$ otherwise. Then, according to (2) and (3), the signal energy E_m can be expressed as follows,

$$E_m = TP \left| \sum_{n=0}^L h_n x_{m-n} \right|^2. \quad (4)$$

Based on the energy E_m of the dispersed signal component within the symbol slot m [c.f. (4)], we characterize the statistical properties of the received signal both of the photon-counting receiver and of the PMT receiver. Let Λ_b denote the counting rate of the background radiation, such that the mean of the signal photons is given by $\lambda_b = T\Lambda_b$. Assume that the number of background radiation photons received obeys a Poisson distribution, which is statistically independent of the number of dispersed signal components.

1) *Photon-Counting Receiver*: The received signal is represented by the number of photons, which obeys a Poisson distribution. The mean of the Poisson distribution for the dispersed signal component, denoted as λ_m , is given by,

$$\lambda_m = \frac{E_m}{E_p}, \quad (5)$$

where E_p is the energy per photon determined by the Laser wavelength. The number of received photons N_m is the sum of the numbers for the dispersed signal component and the background radiation component, which satisfies the following distribution,

$$\mathbb{P}(N_m = n) = \frac{(\lambda_m + \lambda_b)^n}{n!} e^{-\lambda_m - \lambda_b}. \quad (6)$$

2) *Photomultiplier Tube Receiver*: The PMT receiver transforms the discrete number of received photons to the following analog electronic signals,

$$z_m = N_m A e + v_m, \quad (7)$$

where A is the PMT's amplification factor, e is the single electron charge, while v_m is the zero-mean Gaussian noise stimulated by receiving N_m photons. Explicitly, the additive Gaussian noise v_m is stimulated by the detected photons and the thermal noise, which satisfies the following distribution,

$$v_m \sim \mathcal{N}(0, N_m \sigma^2 + \sigma_0^2), \quad (8)$$

where σ^2 denotes the variance of the zero-mean shot noise stimulated by a single photon, while σ_0^2 denotes the thermal noise variance. The shot noise variance σ^2 and the thermal noise variance σ_0^2 are given by,

$$\sigma^2 = (\xi A e)^2, \quad \sigma_0^2 = \frac{2k_e T^o}{R_L} T, \quad (9)$$

where ξ is the PMT spreading factor; k_e , T^o , and R_L are the Boltzmann constant, the receiver temperature and the load resistance, respectively.

Let $G(z; \mu, \sigma^2)$ be the probability density function (PDF) of the Gaussian variable z with a mean of μ and a variance of σ^2 , which is given by

$$G(z; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z - \mu)^2}{2\sigma^2}\right). \quad (10)$$

Recall that the number of received photons N_m satisfies a Poisson distribution with a mean of $(\lambda_m + \lambda_b)$. Then for the PMT receiver, the electronic output signal satisfies the Gaussian mixture distribution, which is the Gaussian distribution $\mathcal{N}(0, N_m\sigma^2 + \sigma_0^2)$ modulated by the Poisson distributed number of the received photons. More specifically, the Gaussian mixture distribution for z_m is given as follows,

$$p(z_m) = \sum_{n=0}^{+\infty} \frac{(\lambda_m + \lambda_b)^n}{n!} e^{-\lambda_m - \lambda_b} G(z_m; nAe, n\sigma^2 + \sigma_0^2). \quad (11)$$

□

We aim for estimating the CIR taps $\{h_l\}_{l=0}^L$ of the scattering channel, based on the signals received from the length- M on-off keying (OOK) pilots $\{x_m\}_{m=1}^M$ within M symbol durations, i.e. on the pilot symbols $x_m \in \{0, 1\}$ for $1 \leq m \leq M$. Note that the phase information contained in the complex-valued channel taps $\{h_l\}_{l=0}^L$ cannot be resolved if the transmitter only sends OOK impulses, Explicitly, the complex values of the CIR taps cannot be inferred based on the amplitudes $\{|h_l|\}_{l=0}^L$.

For the turbulence channel, the receiver periodically estimates the channel taps, and adjusts the post-equalization accordingly. The receiver can also feedback the channel; estimate to the transmitter, such that the transmitter can perform pre-equalization and other pre-distortion processing.

B. Channel Estimation Ambiguity

There are two types of channel estimation ambiguities, because the phase rotated versions and complex conjugates of the transmitted signal cannot be distinguished purely based on the received signal energy. The above arguments are summarized in the following result. The proof hinges on the energy-receiving nature [c.f. (4)] of the receiver in a straightforward manner, and thus it is omitted here.

Proposition 1: For real-valued channel taps $\{h_l\}_{l=0}^L$, the following two types of ambiguities cannot be distinguished by the energy reception receivers considered:

- 1) rotational ambiguity: $\{h_l e^{j\theta}\}_{l=0}^L$, for any rotation angle $0 \leq \theta < 2\pi$;
- 2) complex conjugate ambiguity: $\{h_l^* e^{j\theta}\}_{l=0}^L$, for any rotation angle $0 \leq \theta < 2\pi$, where the superscript $*$ denotes complex conjugate.

□

It will be shown in Section V that the channel estimation only suffers from the rotational ambiguity and the complex conjugate ambiguity, as stated in Proposition 1. Explicitly, the channel estimation approach proposed in Section V does not impose any additional estimation ambiguities.

C. Overview of the Proposed Estimation Approach

Before delving into further details, we first provide an overview of the proposed channel estimation approach. Based on (2), the received signal power $P_m \triangleq \frac{E_m}{T}$ can be written as follows,

$$\begin{aligned}
 P_m &= P \left| \sum_{n=0}^L h_n x_{m-n} \right|^2 = P \sum_{n=0}^L h_n x_{m-n} \sum_{n=0}^L h_n^* x_{m-n} \\
 &= P \sum_{n=0}^L x_{m-n}^2 \text{Re}(h_n h_n^*) + P \sum_{0 \leq n_1 < n_2 \leq L} 2x_{m-n_1} x_{m-n_2} \text{Re}(h_{n_1} h_{n_2}^*), \quad (12)
 \end{aligned}$$

where $\text{Re}(\bullet)$ denotes the real part of a complex number.

According to (12), let \mathbf{h} denote the vector of the second-order moments of the CIR taps, and \mathbf{x}_m denote the vector of the related pilot symbols, given as follows,

$$\begin{aligned}
 \mathbf{h} &\triangleq \left[\text{Re}(h_0 h_0^*), \text{Re}(h_1 h_1^*), \dots, \text{Re}(h_L h_L^*), \text{Re}(h_0 h_1^*), \dots, \text{Re}(h_0 h_L^*), \text{Re}(h_1 h_2^*), \dots, \text{Re}(h_1 h_L^*), \right. \\
 &\quad \left. \text{Re}(h_2 h_3^*), \dots, \text{Re}(h_2 h_L^*), \dots, \text{Re}(h_{L-1} h_L^*) \right]^T, \\
 \mathbf{x}_m &\triangleq \left[x_m^2, x_{m-1}^2, \dots, x_{m-L}^2, 2x_m x_{m-1}, \dots, 2x_m x_{m-L}, 2x_{m-1} x_{m-2}, \dots, 2x_{m-1} x_{m-L}, 2x_{m-2} x_{m-3}, \dots, \right. \\
 &\quad \left. 2x_{m-2} x_{m-L}, \dots, 2x_{m-L+1} x_{m-L} \right]^T. \quad (13)
 \end{aligned}$$

The signal power P_m received within the symbol duration m [c.f. (12)] can be expressed as $P_m = P \mathbf{x}_m^T \mathbf{h}$; and the mean λ_m of the Poisson distributed number of signal photons in slot m [c.f. (5)] is given by

$$\lambda_m = \frac{E_m}{E_p} = \frac{T}{E_p} P_m \triangleq \alpha P_m = \alpha P \mathbf{x}_m^T \mathbf{h}, \quad L \leq m \leq M. \quad (14)$$

The channel estimation can be performed in two steps. In the first step, we estimate \mathbf{h} according to (12) by treating all elements of \mathbf{h} as independent variables; and subsequently in the second

step we estimate $\{h_n\}_{n=0}^L$ based on the estimate of \mathbf{h} , which has been obtained from the first step. The two steps will be addressed in detail in the following sections.

Remark: A preliminary channel estimation technique for the model under consideration has been investigated in [26], where the two-stage channel estimation framework has been proposed. In the current work, we propose a more rigorous channel tap estimation method for the second stage instead of the heuristic one in [26], prove the convergence of the proposed channel estimation framework, and extend the channel estimation to the PMT receiver.

III. ESTIMATION OF SECOND-ORDER MOMENTS \mathbf{h}

Let us now formulate a framework for the estimation of the second-order moments \mathbf{h} for both photon-counting and PMT receivers, based on the received signals $\{N_m\}_{m=L}^M$ and $\{z_m\}_{m=L}^M$, respectively. We also provide the analysis of the estimation error.

A. Estimation for the Photon-Counting Receiver

Recall that according to (14), the number of photons N_m received in slot m , $L \leq m \leq M$, satisfies a Poisson distribution with a mean of $\alpha P \mathbf{x}_m^T \mathbf{h} + \lambda_b$. Let $\mathbf{X} \triangleq [\mathbf{x}_L, \mathbf{x}_{L+1}, \dots, \mathbf{x}_M]^T$, $\mathbf{N} \triangleq [N_L, N_{L+1}, \dots, N_M]^T$, and $\mathbf{1} \triangleq [1, 1, \dots, 1]^T$ of length $(M - L + 1)$. We have the following result conceiving the statistics of the first- and second-order moments of the number of received photons \mathbf{N} .

Lemma 1: We have the following result on the number of received photons \mathbf{N} ,

$$\mathbb{E}[\mathbf{N}] = \alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}, \quad (15)$$

$$\mathbb{E}[\mathbf{N} \mathbf{N}^T] = (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})^T + \text{diag}(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}), \quad (16)$$

where $\text{diag}(\bullet)$ denotes the diagonal matrix consisting of the elements of the vector.

Proof: Note that, for $L \leq m \leq N$, we have $\mathbb{E}[N_m] = \alpha P \mathbf{x}_m^T \mathbf{h} + \lambda_b$. Thus, (15) follows from concatenating N_m as a column vector; and (16) follows from the fact that $\mathbb{E}[N_{m_1} N_{m_2}] = \mathbb{E}[N_{m_1}] \mathbb{E}[N_{m_2}]$ and $\mathbb{E}[N_m^2] = (\mathbb{E}[N_m])^2 + \mathbb{E}[N_m]$. ■

Based on Lemma 1, we define the distortion $D^{PC}(\mathbf{h})$ as the norm-2 distortion between a realization of the number \mathbf{N} of received photons and its expectation, given as follows,

$$D^{PC}(\mathbf{h}) \triangleq \left\| \alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1} - \mathbf{N} \right\|^2. \quad (17)$$

An estimate of \mathbf{h} , denoted as $\hat{\mathbf{h}}$, can be obtained as the one that minimizes the distortion $D^{PC}(\mathbf{h})$, formulated as $\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} D^{PC}(\mathbf{h})$. The following result provides the estimate $\hat{\mathbf{h}}$ according to the above arguments. The proof follows from setting $\frac{\partial D^{PC}(\mathbf{h})}{\partial \mathbf{h}} = \mathbf{0}$ to obtain the \mathbf{h} that minimizes $D^{PC}(\mathbf{h})$.

Theorem 1: The estimate $\hat{\mathbf{h}}$ that minimizes $D^{PC}(\mathbf{h})$ is given by,

$$\hat{\mathbf{h}} = \frac{1}{\alpha P} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{N} - \lambda_b \mathbf{1}). \quad (18)$$

□

In the following we analyze the estimate $\hat{\mathbf{h}}$. It can be proved that the estimate $\hat{\mathbf{h}}$ is unbiased, i.e., we have $\mathbb{E}[\hat{\mathbf{h}}] = \mathbf{h}$. We also determine the covariance matrix $\mathbb{E}\left[\left(\hat{\mathbf{h}} - \mathbf{h}\right)\left(\hat{\mathbf{h}} - \mathbf{h}\right)^T\right]$ of the estimate $\hat{\mathbf{h}}$, whose trace is the expectation of the estimation distortion $\mathbb{E}\left[\left\|\hat{\mathbf{h}} - \mathbf{h}\right\|^2\right]$.

Theorem 2: The estimate $\hat{\mathbf{h}}$ is unbiased, i.e., we have $\mathbb{E}[\hat{\mathbf{h}}] = \mathbf{h}$; and its co-variance matrix is given by,

$$\begin{aligned} \mathbb{E}\left[\left(\hat{\mathbf{h}} - \mathbf{h}\right)\left(\hat{\mathbf{h}} - \mathbf{h}\right)^T\right] &= \frac{1}{\alpha P} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{diag}(\mathbf{X}\mathbf{h}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &\quad + \frac{\lambda_b}{\alpha^2 P^2} (\mathbf{X}^T \mathbf{X})^{-1}. \end{aligned} \quad (19)$$

Thus, letting $Tr(\cdot)$ denote the trace of a matrix, we have the following

$$\begin{aligned} \mathbb{E}\left[\left\|\hat{\mathbf{h}} - \mathbf{h}\right\|^2\right] &= \frac{1}{\alpha P} Tr\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{diag}(\mathbf{X}\mathbf{h}) \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right) \\ &\quad + \frac{\lambda_b}{\alpha^2 P^2} Tr\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right). \end{aligned} \quad (20)$$

Proof: According to (15) in Lemma 1, we have

$$\mathbb{E}[\hat{\mathbf{h}}] = \frac{1}{\alpha P} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbb{E}[\mathbf{N}] - \lambda_b \mathbf{1}) = \frac{1}{\alpha P} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \alpha P \mathbf{X} \mathbf{h} = \mathbf{h}. \quad (21)$$

Thus the estimate $\hat{\mathbf{h}}$ is unbiased. Moreover, according to (18), we have

$$\hat{\mathbf{h}} - \mathbf{h} = \frac{1}{\alpha P} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{N} - (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})). \quad (22)$$

Then, we arrive that

$$\begin{aligned} \left(\hat{\mathbf{h}} - \mathbf{h}\right)\left(\hat{\mathbf{h}} - \mathbf{h}\right)^T &= \frac{1}{\alpha^2 P^2} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{N} - (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})) \\ &\quad \left(\mathbf{N} - (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})\right)^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}. \end{aligned} \quad (23)$$

Since \mathbf{N} is Poissonian with a mean of $\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}$, according to (15) we have,

$$\mathbb{E}\left[\left(\mathbf{N} - (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})\right)\left(\mathbf{N} - (\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1})\right)^T\right] = \alpha P \text{diag}(\mathbf{X}\mathbf{h}) + \lambda_b \mathbf{I}. \quad (24)$$

Based on (23) and (24), we can prove (19); and equation (20) follows from the fact that

$$\|\hat{\mathbf{h}} - \mathbf{h}\|^2 = \text{Tr}\left(\left(\hat{\mathbf{h}} - \mathbf{h}\right)\left(\hat{\mathbf{h}} - \mathbf{h}\right)^T\right). \quad (25)$$

■

From (19), it is seen that the estimation distortion $\mathbb{E}\left[\|\mathbf{h} - \hat{\mathbf{h}}\|^2\right]$ contains two terms, the traces of $\frac{1}{\alpha}\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{diag}\left(\mathbf{X} \mathbf{h}\right) \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}$ and $\frac{\lambda_b}{\alpha^2}\left(\mathbf{X}^T \mathbf{X}\right)^{-1}$. The former is due to the Poisson distributed signal component, which is unique for a Poissonian channel; and the latter represents the channel estimation distortion imposed by the additive noise, which is also part of the channel estimation distortion of RF communication. The pilot sequence needs to be designed to minimize the combination of distortion in the two terms.

B. Estimation for PMT Receiver

The estimation of $\hat{\mathbf{h}}$ for the PMT receiver is similar to that for the photon-counting receiver. Let $\mathbf{z} \triangleq [z_L, z_{L+1}, \dots, z_M]^T$. We then have the following result on the statistics of the first- and second-order moments of \mathbf{z} .

Lemma 2: The statistics of the first- and second-order moments of \mathbf{z} are given as follows,

$$\mathbb{E}\left[\mathbf{z}\right] = \left(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}\right) A e, \quad (26)$$

$$\begin{aligned} \mathbb{E}\left[\mathbf{z} \mathbf{z}^T\right] &= A^2 e^2 \left(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}\right) \left(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1}\right)^T + \alpha P \left(A^2 e^2 + \sigma^2\right) \text{diag}\left(\mathbf{X} \mathbf{h}\right) \\ &\quad + \left(\lambda_b A^2 e^2 + \lambda_b \sigma^2 + \sigma_0^2\right) \mathbf{I}. \end{aligned} \quad (27)$$

Proof: Note that z_m satisfied the Gaussian mixture distribution specified in (11) for $L \leq m \leq N$. We then have

$$\begin{aligned} \mathbb{E}\left[z_m\right] &= \sum_{n=0}^{+\infty} \frac{(\lambda_m + \lambda_b)^n}{n!} e^{-\lambda_m - \lambda_b} n A e \\ &= (\lambda_m + \lambda_b) A e = \left(\alpha P \mathbf{x}_m^T \mathbf{h} + \lambda_b\right) A e. \end{aligned} \quad (28)$$

Then, (26) follows from concatenating $\mathbb{E}\left[z_m\right]$ as a column vector.

Note that the elements z_{m_1} and z_{m_2} are statistically independent from each other for $m_1 \neq m_2$. Thus we have that $\mathbb{E}\left[z_{m_1} z_{m_2}\right] = \mathbb{E}\left[z_{m_1}\right] \mathbb{E}\left[z_{m_2}\right]$, from which the non-diagonal elements of $\mathbb{E}\left[\mathbf{z} \mathbf{z}^T\right]$ follow. The diagonal elements of $\mathbb{E}\left[\mathbf{z} \mathbf{z}^T\right]$ follow from the fact that for a mixed Gaussian distributed random variable z_m , we have

$$\begin{aligned} \mathbb{E}\left[z_m^2\right] &= \left(\alpha P \mathbf{x}_m^T \mathbf{h} + \lambda_b\right)^2 A^2 e^2 + \alpha P \left(A^2 e^2 + \sigma^2\right) \mathbf{x}_m^T \mathbf{h} \\ &\quad + \left(\lambda_b A^2 e^2 + \lambda_b \sigma^2 + \sigma_0^2\right). \end{aligned} \quad (29)$$

Compared with the first-order and second-order moments of the received signals for the photon-counting receiver [c.f. (15) - (16)], it is seen that first-order and second-order moments for the PMT receiver contains both the contribution from the background radiation and the additive shot and thermal noise [c.f. (26) - (27)]. We need to pursue the optimal estimate of \mathbf{h} for the PMT receiver based on its own signal characteristics. Similar to the photon-counting receiver, we want to minimize the following channel estimation distortion metric for the PMT receiver

$$D^{PMT}(\mathbf{h}) = \left\| \left(\alpha P \mathbf{X} \mathbf{h} + \lambda_b \mathbf{1} \right) A e - \mathbf{z} \right\|^2. \quad (30)$$

Similar to Theorem 1, we have the following result on the estimate $\hat{\mathbf{h}}$ that minimizes the distortion $D^{PMT}(\mathbf{h})$. The proof is similar to that of Theorem 1, and thus omitted here.

Theorem 3: The estimate $\hat{\mathbf{h}}$ that minimizes $D^{PMT}(\mathbf{h})$ [c.f. (30)] is given as follows,

$$\hat{\mathbf{h}} = \frac{1}{\alpha P} \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \left(\frac{\mathbf{z}}{A e} - \lambda_b \mathbf{1} \right). \quad (31)$$

□

It can also be shown that, similar to the photon-counting receiver [c.f. (18)], the estimate $\hat{\mathbf{h}}$ of the PMT receiver is unbiased too. Moreover, the estimation distortion of (30) consists of two terms, namely the one from the Poissonian signal characteristics and the other from the additive Poisson noise. The proof is similar to that of Theorem 2, and thus it is omitted here.

Theorem 4: The estimate $\hat{\mathbf{h}}$ is unbiased, i.e. $\mathbb{E}[\hat{\mathbf{h}}] = \mathbf{h}$; and its co-variance matrix is given by

$$\mathbb{E} \left[\left(\hat{\mathbf{h}} - \mathbf{h} \right) \left(\hat{\mathbf{h}} - \mathbf{h} \right)^T \right] = \frac{A^2 e^2 + \sigma^2}{\alpha P A^2 e^2} \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \text{diag} \left(\mathbf{X} \mathbf{h} \right) \mathbf{X} \left(\mathbf{X}^T \mathbf{X} \right)^{-1} + \frac{\lambda_b A^2 e^2 + \lambda_b \sigma^2 + \sigma_0^2}{\alpha^2 P^2 A^2 e^2} \left(\mathbf{X}^T \mathbf{X} \right)^{-1}. \quad (32)$$

Thus, letting $Tr(\cdot)$ denote the trace of a matrix, we have

$$\mathbb{E} \left[\left\| \hat{\mathbf{h}} - \mathbf{h} \right\|^2 \right] = \frac{A^2 e^2 + \sigma^2}{\alpha P A^2 e^2} Tr \left(\left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \text{diag} \left(\mathbf{X} \mathbf{h} \right) \mathbf{X} \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \right) + \frac{\lambda_b A^2 e^2 + \lambda_b \sigma^2 + \sigma_0^2}{\alpha^2 P^2 A^2 e^2} Tr \left(\left(\mathbf{X}^T \mathbf{X} \right)^{-1} \right), \quad (33)$$

which is a counterpart to equation (20) representing the the photon-counting receiver.

□

C. Discussions

It is seen from Theorems 1 and 3 that the computation of the estimate $\hat{\mathbf{h}}$ involves the evaluation of the matrix $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ and its multiplication with a column vector. While the former can be done offline, the latter involves a computational complexity on the order of $\frac{L(L+1)}{2} \times (M-L+1)$. Moreover, as we will demonstrate in Section VI on the numerical results, the pilots $\{x_m\}_{m=1}^M$ have to be designed for reducing the estimation distortion.

In order to reduce the computational complexity, in the following section we conceive a sparse pilot sequence structure, for ensuring that the complexity of the multiplication with the matrix $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ can be significantly reduced. Moreover, we will show in Section VI that such a pilot structure can lead to a reduced estimation distortion compared to certain realizations of random 0-1 pilot sequences.

IV. SPARSE PILOT DESIGN FOR EFFICIENT ESTIMATION OF SECOND-ORDER MOMENTS \mathbf{h}

We conceive a sparse pilot sequence structure, as well as the associated channel estimation approach, which allows us to construct an efficient estimation scheme for the second-order moments of \mathbf{h} . Moreover, it can be shown that the proposed estimation scheme also conforms to the optimality criterion of (18) and (31) in our general framework.

A. Estimation Using Sparse Pilots

Note that the non-sparsity of the matrix $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ stems from the cross-terms $x_{m-l_1} x_{m-l_2}$, $l_1 \neq l_2$, involved in the transmission power $P_m = P |\sum_{l=0}^L h_l x_{m-l}|^2$ [c.f. (12)]. If there are no such cross-terms, P_m only reflects the norm of the taps $\{h_l\}_{l=0}^L$, from which the exact values of the taps cannot be inferred. Thus there should be at least one such cross-term in the analytical expression of P_m in (12), in order to calculate the exact value of the taps $\{h_l\}_{l=0}^L$. In the following we will design pilot sequences, which have at most one cross-term involved in the calculation of the power P_m .

More specifically, the design conceived guarantees that the convolution $P_m = P |\sum_{l=0}^L h_l x_{m-l}|^2$ involves one or two non-zero pilots taps x_{m-l} , such that $|h_n|^2$ and $\text{Re}(h_{l_1} h_{l_2}^*)$ can be estimated in an efficient manner. Note that

$$P_m = P \left| \sum_{l=0}^L h_l x_{m-l} \right|^2 = \begin{cases} P |h_l|^2, & \text{if only } x_{m-l} = 1; \\ P |h_{l_1} + h_{l_2}|^2, & \text{if only } x_{m-l_1} = x_{m-l_2} = 1; \end{cases} \quad (34)$$

where

$$|h_{l_1} + h_{l_2}|^2 = |h_{l_1}|^2 + |h_{l_2}|^2 + 2\text{Re}(h_{l_1}h_{l_2}^*). \quad (35)$$

From (34), \mathbf{h} can be estimated using a successive approach. We first obtain estimates of $|h_l|^2$ for $0 \leq l \leq L$ via the power P_m that involves only one non-zero x_{m-l} , and then obtain the estimates of $\text{Re}(h_{l_1}h_{l_2}^*)$ for $0 \leq l_1 < l_2 \leq L$ based on (35) and the estimates of $|h_l|^2$.

1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0
1	0	0	0	×	×	×	×	×	×	×	×	×	×	×	×	×
.....																

TABLE I
PILOT SYMBOL UNIT FOR $L = 7$

1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
1	0	0	0	1	0	0	0	0	×	×	×	×	×	×	×	×	×
.....																	

TABLE II
PILOT SYMBOL UNIT FOR $L = 8$

Note that, for $0 \leq l_1 < l_2 \leq L$, in a pilot pattern that covers the cross-term $\text{Re}(h_{l_1}h_{l_2}^*)$, there are $(l_2 - l_1 - 1)$ zeros between two consecutive non-zero pilots. Based on the above intuition, one unit of our sparse pilot sequence contains a sequence consisting of $(L + 1)$ ones, where the l^{th} one is followed by t_l zeros, given by the following time instants

$$\begin{aligned} t_1 &= L, \quad t_2 = 0, \quad t_3 = L - 1, \quad t_4 = 1, \\ t_5 &= L - 2, \quad t_6 = 2, \quad t_7 = L - 3, \quad t_8 = 3, \dots \end{aligned} \quad (36)$$

Note that the length of such a sequence, denoted as L_P , is given as follows,

$$L_P = \frac{(L + 1)(L + 2)}{2}. \quad (37)$$

This pilot unit is then repeated several times, where $\lfloor \frac{L}{2} \rfloor$ zeros are added before the first unit, to complete the entire pilot sequence for the estimation of \mathbf{h} . Examples for such sequences for $L = 7$ and $L = 8$ are shown in Tables I and II, respectively. The pilot units are formed by concatenating the line of each table.

The following result shows that based on this specific pilot sequence structure, all combinations of $\{|h_l|^2\}_{l=0}^L$ and $\{|h_{l_1} + h_{l_2}|^2\}_{l_1, l_2=0}^L$ can be covered.

Theorem 5: All combinations of $\{|h_l|^2\}_{l=0}^L$ and $\{|h_{l_1} + h_{l_2}|^2\}_{l_1, l_2=0}^L$ can be covered by the pilot sequence characterized by (36). Moreover, each one is covered only once.

Proof: Let us first consider the combination $\{|h_{l_1} + h_{l_2}|^2\}_{l_1, l_2=0}^L$ for $l_1 < l_2$. For $l_2 - l_1 = t > 0$, it can be proved that $|h_{l_1} + h_{l_2}|^2$ can be covered by two consecutive ones with t zeros between them. For $t_k = t$, note that $\{t_{k-1}, t_{k+1}\} = \{L - 1 - t, L - t\}$, which guarantees that all combinations $\{|h_{l_1} + h_{l_2}|^2\}_{l_2 - l_1 = t}$ can be covered by $\sum_{n=0}^L h_n x_{m-n}$ based on the following pattern

$$\left[\underbrace{00\dots0}_{L-1-t} 1 \underbrace{00\dots0}_t 1 \underbrace{00\dots0}_{L-t} \right], \text{ or } \left[\underbrace{00\dots0}_{L-t} 1 \underbrace{00\dots0}_t 1 \underbrace{00\dots0}_{L-t-1} \right], \quad (38)$$

which exists in the pilot sequence.

Next consider the representation $\{|h_l|^2\}_{l=0}^L$. It is seen that in the pattern under consideration we have $t_{L+1} = \lfloor \frac{L}{2} \rfloor$, thus $\{|h_l|^2\}_{l=L-\lfloor \frac{L}{2} \rfloor}^L$ can be covered by the following pilot pattern

$$\left[\underbrace{00\dots0}_{\lfloor \frac{L}{2} \rfloor} 1 \underbrace{00\dots0}_L \right]. \quad (39)$$

For $0 \leq l \leq L - \lfloor \frac{L}{2} \rfloor - 1$, $|h_l|^2$ can be covered by convolving the following pattern with the pilot sequence,

$$\left[\underbrace{00\dots0}_{L-l} 1 \underbrace{00\dots0}_l \right]. \quad (40)$$

Note the length of each unit of the pilot sequence is $\frac{(L+1)(L+2)}{2}$ and the number of combinations $\{|h_l|^2\}_{l=0}^L$ and $\{|h_{l_1} + h_{l_2}|^2\}_{l_1, l_2=0}^L$ is also $\frac{(L+1)(L+2)}{2}$. Thus each combination can only be covered once by the pilot sequence. ■

We then outline the successive estimation approach for the second-order moments \mathbf{h} . Let \mathcal{S}_i and \mathcal{S}_{ij} denote the set of m , $L \leq m \leq M$, corresponding to the received signal power P_m that involves only one non-zero channel tap $|h_i|^2$ and two non-zero channel taps $|h_i + h_j|^2$, respectively. Let $\hat{\text{Re}}(h_i h_i^*)$, $0 \leq i \leq L$, and $\hat{\text{Re}}(h_i h_j^*)$, $0 \leq i < j \leq L$, denote the estimate of $|h_i|^2$ and $\text{Re}(h_i h_j^*)$, respectively. Based on the expectation $\mathbb{E}[\mathbf{N}]$, we first obtain an estimate $\hat{\text{Re}}(h_i h_i^*)$ for $0 \leq i \leq L$, and then obtain the estimate $\hat{\text{Re}}(h_i h_j^*)$ for $0 \leq i < j \leq L$ based on the estimate $\{\hat{\text{Re}}(h_i h_i^*)\}_{i=0}^L$.

- 1) **Photon-counting Receiver:** Given the number of photons $\{N_m\}_{m=L}^M$ received over the slots $L \leq m \leq M$, $\hat{\text{Re}}(h_i h_i^*)$ and $\hat{\text{Re}}(h_i h_j^*)$ can be estimated in a successive manner as

follows,

$$\begin{aligned}\hat{\text{Re}}(h_i h_i^*) &= \frac{\sum_{m \in \mathcal{S}_i} N_m - \lambda_b |\mathcal{S}_i|}{\alpha P |\mathcal{S}_i|}, \quad 0 \leq i \leq L; \\ \hat{\text{Re}}(h_i h_j^*) &= \frac{\sum_{m \in \mathcal{S}_{ij}} N_m - \lambda_b |\mathcal{S}_{ij}|}{2\alpha P |\mathcal{S}_{ij}|} - \frac{\hat{\text{Re}}(h_i h_i^*) + \hat{\text{Re}}(h_j h_j^*)}{2}, \quad 0 \leq i < j \leq L.\end{aligned}\quad (41)$$

- 2) **PMT Receiver:** Given the number of photons $\{z_m\}_{m=L}^M$ received over the slots $L \leq m \leq M$, $\hat{\text{Re}}(h_i h_i^*)$ and $\hat{\text{Re}}(h_i h_j^*)$ can be estimated in a successive manner as follows,

$$\begin{aligned}\hat{\text{Re}}(h_i h_i^*) &= \frac{\sum_{m \in \mathcal{S}_i} z_m - \lambda_b |\mathcal{S}_i| A e}{\alpha P |\mathcal{S}_i| A e}, \quad 0 \leq i \leq L; \\ \hat{\text{Re}}(h_i h_j^*) &= \frac{\sum_{m \in \mathcal{S}_{ij}} z_m - \lambda_b |\mathcal{S}_{ij}| A e}{2\alpha P |\mathcal{S}_{ij}| A e} - \frac{\hat{\text{Re}}(h_i h_i^*) + \hat{\text{Re}}(h_j h_j^*)}{2}, \quad 0 \leq i < j \leq L.\end{aligned}\quad (42)$$

B. Optimality of the Sparse Pilot Sequence

The successive estimation approaches (41) and (42) serve as low-complexity solutions to the estimation of the second-order moments \mathbf{h} . It can be proved that this successive estimation approach conforms to the optimality criterion (18) and (31).

We specify the structure of the pilot matrix \mathbf{X} for the particular type of the pilot sequence specified by (34). Let \mathbf{G} be a $\frac{L(L+1)}{2} \times (L+1)$ matrix, where in each row two elements are one and all other elements are zero. It characterizes the terms $|h_{l_1}|^2$ and $|h_{l_2}|^2$ involved in the computation of $|h_{l_1} + h_{l_2}|^2$. The entire pilot structure can be characterized by the following matrix,

$$\mathbf{Q} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \text{---} & \text{---} \\ \mathbf{G} & 2\mathbf{I} \end{pmatrix}, \quad (43)$$

where the upper portion and lower portion characterize the pilots for the terms $|h_l|^2$ and $|h_{l_1} + h_{l_2}|^2 = |h_{l_1}|^2 + |h_{l_2}|^2 + 2\text{Re}(h_{l_1} h_{l_2}^*)$, respectively. According to Theorem 5, the output of each unit of the pilot sequence shown in (36) is given by $\mathbf{Q}\mathbf{h}$. Assuming that the unit of the pilot sequence is repeated r times for constructing the entire pilot sequence, we have

$$\mathbf{X} = \left[\mathbf{Q}^T \quad \mathbf{Q}^T \quad \dots \quad \mathbf{Q}^T \right]^T, \quad (44)$$

in conjunction with r component matrices \mathbf{Q}^T .

According to (44), we have

$$\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T = \frac{1}{r} \left[\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} \mathbf{Q}^T \left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} \mathbf{Q}^T \dots \left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} \mathbf{Q}^T \right], \quad (45)$$

with r repetitions of the component matrices $\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} \mathbf{Q}^T$. According to (43), we have

$$\mathbf{Q}^T \mathbf{Q} = \begin{pmatrix} \mathbf{I} + \mathbf{G}^T \mathbf{G} & 2\mathbf{G}^T \\ 2\mathbf{G} & 4\mathbf{I} \end{pmatrix}. \quad (46)$$

In order to further analyze the inverse matrix $\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1}$, we outline the following result on the matrix inversion, which has been provided in [27].

Proposition 2: If both matrices \mathbf{D} and $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ are nonsingular, we have (47) on the matrix inversion.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \left(\mathbf{A} - \mathbf{C}\mathbf{D}^{-1}\mathbf{B}\right)^{-1} & -\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} \mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} \mathbf{B}\mathbf{D}^{-1} \end{pmatrix}. \quad (47)$$

□

Upon substituting $\mathbf{A} = \mathbf{I} + \mathbf{G}^T \mathbf{G}$, $\mathbf{B} = \mathbf{C} = 2\mathbf{G}^T$, and $\mathbf{D} = 4\mathbf{I}$ into (47), we have the following result on the inverse matrix $\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1}$.

Lemma 3: According to (43), we have the following,

$$\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} = \begin{pmatrix} \mathbf{I} & -\frac{1}{2}\mathbf{G}^T \\ -\frac{1}{2}\mathbf{G} & \frac{1}{4}\mathbf{I} + \frac{1}{4}\mathbf{G}\mathbf{G}^T \end{pmatrix}, \text{ and } \left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} \mathbf{Q}^T = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\frac{1}{2}\mathbf{G} & \frac{1}{2}\mathbf{I} \end{pmatrix}. \quad (48)$$

□

According to (48), we can prove the conformance of the proposed low-complexity solution (41) and (42), to the optimality criterion (18) and (31) for the general estimation framework, respectively. The equivalence is summarized in the following result. The proof is straightforward and thus omitted here.

Theorem 6: The estimates (41) and (42) are equivalent to the solution provided in (18) and (31), respectively, for photon-counting and PMT receivers.

□

Finally we analyze the variance of the estimate $\hat{\mathbf{h}}$ generated with the sparse pilot structure based on Theorems 2 and 4. For the term arriving from the additive Poisson noise, we have

$$\text{Tr}\left(\mathbf{X}^T \mathbf{X}\right)^{-1} = \frac{1}{r} \text{Tr}\left(\mathbf{Q}^T \mathbf{Q}\right)^{-1} = \frac{(L+1)(3L+4)}{4r}; \quad (49)$$

and for the term engendered by the Poisson derived signal variance, we have

$$\begin{aligned} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{diag}(\mathbf{X} \mathbf{h}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} &= \frac{1}{r} (\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T \text{diag}(\mathbf{Q} \mathbf{h}) \mathbf{Q} (\mathbf{Q}^T \mathbf{Q})^{-1} \\ &= \frac{1}{r} \left(2 \sum_{l=0}^L |h_l|^2 + \frac{1}{2} \sum_{l_1=1}^{L+1} \sum_{l_2=l_1+1}^{L+1} \text{Re}(h_{l_1} h_{l_2}^*) \right). \end{aligned} \quad (50)$$

We can then obtain the variance of the estimate \mathbf{h} based on (49) and (50), for both the photon-counting receiver and the PMT receiver. The estimation variances are consistent with those obtained directly from (41) and (42).

V. ESTIMATION OF THE CHANNEL TAPS $\{h_n\}_{n=0}^L$

Let the channel taps be represented by $\mathbf{h}_c \triangleq [h_0, h_1, h_2, \dots, h_L]^T$. Assume that we have obtained an estimate of \mathbf{h} , given by

$$\begin{aligned} \hat{\mathbf{h}} \triangleq & \left[\hat{\text{Re}}(h_0 h_0^*), \hat{\text{Re}}(h_1 h_1^*), \dots, \hat{\text{Re}}(h_L h_L^*), \hat{\text{Re}}(h_0 h_1^*), \dots, \hat{\text{Re}}(h_0 h_L^*), \hat{\text{Re}}(h_1 h_2^*), \dots, \hat{\text{Re}}(h_1 h_L^*), \right. \\ & \left. \hat{\text{Re}}(h_2 h_3^*), \dots, \hat{\text{Re}}(h_2 h_L^*), \dots, \hat{\text{Re}}(h_{L-1} h_L^*) \right]^T, \end{aligned} \quad (51)$$

where $\hat{\text{Re}}(h_i h_j^*)$ denotes the estimate of $\text{Re}(h_i h_j^*)$. Let us now define $\mathbf{R} \triangleq \text{Re}(\mathbf{h}_c \mathbf{h}_c^H)$, where \mathbf{h}_c^H denotes the complex conjugate of the vector \mathbf{h}_c and $\text{Re}(\bullet)$ represents the real part of a complex matrix. An estimate of \mathbf{R} , denoted as $\hat{\mathbf{R}}$, can be obtained from (51) as follows,

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{\text{Re}}(h_0 h_0^*) & \hat{\text{Re}}(h_0 h_1^*) & \dots & \hat{\text{Re}}(h_0 h_L^*) \\ \hat{\text{Re}}(h_0 h_1^*) & \hat{\text{Re}}(h_1 h_1^*) & \dots & \hat{\text{Re}}(h_1 h_L^*) \\ \dots & \dots & \dots & \dots \\ \hat{\text{Re}}(h_0 h_L^*) & \hat{\text{Re}}(h_1 h_L^*) & \dots & \hat{\text{Re}}(h_L h_L^*) \end{pmatrix}. \quad (52)$$

We have to obtain an estimate of \mathbf{h}_c , denoted as $\hat{\mathbf{h}}_c$, based on the estimate $\hat{\mathbf{R}}$, as elaborated on in this section. We first investigate several properties of \mathbf{R} , then propose an algorithm for estimating \mathbf{h}_c , and finally analyze the performance of the proposed algorithm.

A. Properties of \mathbf{R}

Note that the complex channel taps \mathbf{h}_c can be decomposed as follows

$$\mathbf{h}_c = \mathbf{h}_{cr} + \mathbf{h}_{ci} \cdot \sqrt{-1}, \quad (53)$$

where \mathbf{h}_{cr} and \mathbf{h}_{ci} are the real and imaginary parts of \mathbf{h}_c , respectively. Then, the matrix \mathbf{R} can be expressed based on \mathbf{h}_{cr} and \mathbf{h}_{ci} as follows,

$$\mathbf{R} = \text{Re}(\mathbf{h}_c \mathbf{h}_c^H) = \mathbf{h}_{cr} \mathbf{h}_{cr}^T + \mathbf{h}_{ci} \mathbf{h}_{ci}^T \triangleq \mathbf{H} \mathbf{H}^T, \quad (54)$$

where $\mathbf{H} \triangleq [\mathbf{h}_{cr} \ \mathbf{h}_{ci}]$ is an $L \times 2$ matrix consisting of the real and imaginary parts of \mathbf{h}_c . It is readily seen from (54) that \mathbf{R} is a rank-two real symmetric (Hermitian) matrix, which we use for reconstructing \mathbf{H} .

We next investigate the eigenvalue decomposition of \mathbf{R} . Since \mathbf{R} is a Hermitian matrix, all its eigenvalues are real. Since \mathbf{R} is a rank-two real symmetric (Hermitian) matrix, the eigenvalue decomposition of \mathbf{R} can be written as follows,

$$\mathbf{R} = \kappa_1 \mathbf{v}_1 \mathbf{v}_1^T + \kappa_2 \mathbf{v}_2 \mathbf{v}_2^T = \tilde{\mathbf{V}} \tilde{\mathbf{V}}^T, \quad (55)$$

where $\tilde{\mathbf{V}} \triangleq [\sqrt{\kappa_1} \mathbf{v}_1, \sqrt{\kappa_2} \mathbf{v}_2]$. We now have to investigate the relationship between \mathbf{H} and $\tilde{\mathbf{V}}$.

If $\mathbf{h}_{ci} = \rho \mathbf{h}_{cr}$ for some real number ρ , then \mathbf{R} is a rank one matrix [c.f. (54)]. Then we have $\kappa_2 = 0$, and

$$\mathbf{v}_1 = \frac{\mathbf{h}_{cr}}{\|\mathbf{h}_{cr}\|}, \quad \text{and } \kappa_1 = (1 + \rho^2) \|\mathbf{h}_{cr}\|^2. \quad (56)$$

Otherwise, we can prove that $\kappa_1, \kappa_2 > 0$, and $\mathbf{v}_1, \mathbf{v}_2$ lie in the space spanned by \mathbf{h}_{ci} and \mathbf{h}_{cr} . More specifically, we have the following result.

Lemma 4: If \mathbf{h}_{ci} and \mathbf{h}_{cr} are linearly independent, we have $\kappa_1, \kappa_2 > 0$; and \mathbf{v}_1 and \mathbf{v}_2 lie in the space spanned by \mathbf{h}_{ci} and \mathbf{h}_{cr} .

Proof: Note that, since \mathbf{h}_{ci} and \mathbf{h}_{cr} are linearly independent, $\mathbf{R} = \mathbf{H} \mathbf{H}^T$ is a rank-two matrix. Then we have that the two eigenvalues satisfy $\kappa_1, \kappa_2 > 0$. Since \mathbf{v}_1 and \mathbf{v}_2 are orthogonal to each other, we have

$$\begin{aligned} \kappa_1 \mathbf{v}_1^T &= \mathbf{v}_1^T (\kappa_1 \mathbf{v}_1 \mathbf{v}_1^T + \kappa_2 \mathbf{v}_2 \mathbf{v}_2^T) = \mathbf{v}_1^T (\mathbf{h}_{cr} \mathbf{h}_{cr}^T + \mathbf{h}_{ci} \mathbf{h}_{ci}^T) \\ &= (\mathbf{v}_1^T \mathbf{h}_{cr}) \mathbf{h}_{cr}^T + (\mathbf{v}_1^T \mathbf{h}_{ci}) \mathbf{h}_{ci}^T; \end{aligned} \quad (57)$$

and thus \mathbf{v}_1 lies in the space spanned by \mathbf{h}_{ci} and \mathbf{h}_{cr} . Using similar arguments, we can prove that \mathbf{v}_2 lies in the space spanned by \mathbf{h}_{ci} and \mathbf{h}_{cr} . ■

According to Lemma 4, the eigenvectors satisfy $\tilde{\mathbf{V}} = \mathbf{H} \mathbf{T}$, where \mathbf{T} is the linear combinations matrix. Then, since both $\tilde{\mathbf{V}}$ and \mathbf{H} are real matrices, we have

$$\tilde{\mathbf{V}} = \mathbf{H} \mathbf{T} = \mathbf{H}^* \mathbf{T}^* = \mathbf{H} \mathbf{T}^*, \quad (58)$$

where \mathbf{H}^* and \mathbf{T}^* denote the element-wise complex conjugate of \mathbf{H} and \mathbf{T} , respectively. Again, since \mathbf{H} is of full column rank, we have that $\mathbf{T} = \mathbf{T}^*$, and thus \mathbf{T} is a real matrix.

According to (54), we have $\mathbf{R} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^T = \mathbf{H}\mathbf{T}\mathbf{T}^T\mathbf{H}^T = \mathbf{H}\mathbf{H}^T$. Since \mathbf{H} is of full column rank, we have $\mathbf{T}\mathbf{T}^T = \mathbf{I}$. Based on this we can further characterize the matrix \mathbf{T} . Letting

$$\mathbf{T} \triangleq \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}, \quad (59)$$

we have $t_{11}^2 + t_{12}^2 = t_{21}^2 + t_{22}^2 = 1$. Let $t_{11} = \cos \alpha$, $t_{12} = \sin \alpha$, $t_{21} = \sin \beta$, $t_{22} = \cos \beta$. Since $t_{11}t_{21} + t_{12}t_{22} = 0$, we have $\sin(\alpha + \beta) = 0$, i.e., the matrix \mathbf{T} can be written as follows,

$$\mathbf{T} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \triangleq \mathbf{R}^+(\alpha), \quad \text{or} \quad \mathbf{T} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} \triangleq \mathbf{R}^-(\alpha). \quad (60)$$

As $\mathbf{V} = \mathbf{H}\mathbf{T}$, we have the following relationship,

$$\mathbf{H} = \mathbf{V}\mathbf{R}^+(-\alpha), \quad \text{and} \quad \mathbf{H} = \mathbf{V}\mathbf{R}^-(-\alpha), \quad (61)$$

for any angle $0 \leq \alpha < 2\pi$. Note that $\mathbf{R}^+(\alpha)$ represents the rotation matrix in the two dimensional x - y plane; and $\mathbf{R}^-(\alpha)$ represents the rotation matrix followed by the mirror transform over the x -axis.

It may be obtained from (55) that a feasible complex channel tap solution \mathbf{h}_c is given by

$$\mathbf{h}_{c0} = \sqrt{\kappa_1}\mathbf{v}_1 + \sqrt{\kappa_2}\mathbf{v}_2 \cdot \sqrt{-1}, \quad (62)$$

which corresponds to the matrix $\tilde{\mathbf{V}}$. According to (61), it may be shown that *all feasible solutions* can be formulated based on the following expression,

$$\mathbf{h}_c = \mathbf{h}_{c0}e^{-j\alpha}, \quad \text{or} \quad \mathbf{h}_c = \mathbf{h}_{c0}^*e^{-j\alpha}, \quad \text{for any } 0 \leq \alpha < 2\pi. \quad (63)$$

Observe from Proposition 1 of Section II-B that the channel estimation may suffer from both a rotational ambiguity and a conjugate ambiguity, which serve as a ‘‘lower bound’’ on the channel estimation ambiguity. According to (63), it is seen that the proposed approach relying on first estimating \mathbf{R} and then estimating \mathbf{h}_c suffers from both of the above ambiguities, which shows that such an ‘‘lower bound’’ is tight. Therefore, the channel estimation problem under consideration suffers from both types of ambiguities, while the proposed estimation approach does not introduce any more ambiguities.

B. Estimation of the Taps $\{h_n\}_{n=0}^L$

We now propose an algorithm for estimating \mathbf{h}_c based on the eigenvalue decomposition of the estimate $\hat{\mathbf{R}}$. Since $\hat{\mathbf{R}}$ is a real symmetric matrix, all its eigenvalues and the corresponding eigenvectors are real. For $1 \leq j \leq L+1$, let $\hat{\kappa}_j$ be the j^{th} largest eigenvalue of $\hat{\mathbf{R}}$ and $\hat{\mathbf{v}}_j$ be the associated real eigenvector, obtained from the following equation

$$\left(\hat{\mathbf{R}} - \hat{\kappa}_j \mathbf{I}\right) \hat{\mathbf{v}}_j = \mathbf{0}. \quad (64)$$

The estimation of $\hat{\mathbf{h}}_c$ is based on the fact that the rank of \mathbf{R} is at most two and that the two non-zero eigenvalues of \mathbf{R} are both positive. More specifically, we find the largest two positive eigenvalues of $\hat{\mathbf{R}}$ and the corresponding eigenvectors. Then we obtain an estimate $\hat{\mathbf{h}}_c$ based on these eigenvalues and eigenvectors. An outage is reported if there is no positive eigenvalue.

The details of the proposed algorithm are outlined in Algorithm 1, as follows.

Algorithm 1 - Estimating \mathbf{h}_c based on the estimate $\hat{\mathbf{R}}$

- 1: Obtain the eigenvalues decomposition of $\hat{\mathbf{R}}$.
Let $\hat{\kappa}_1 \geq \hat{\kappa}_2 \geq \dots \geq \hat{\kappa}_{L+1}$ be the eigenvalues and the corresponding eigenvectors be $\{\hat{\mathbf{v}}_j\}_{j=1}^{L+1}$.
 - 2: If $\hat{\kappa}_1 > 0, \hat{\kappa}_2 > 0$, then let $\hat{\mathbf{h}}_c = \sqrt{\hat{\kappa}_1} \hat{\mathbf{v}}_1 + \sqrt{\hat{\kappa}_2} \hat{\mathbf{v}}_2 \cdot \sqrt{-1}$;
if $\hat{\kappa}_1 > 0, \hat{\kappa}_2 \leq 0$, then let $\hat{\mathbf{h}}_c = \sqrt{\hat{\kappa}_1} \hat{\mathbf{v}}_1$;
otherwise, output estimation outage.
 - 3: Output $\hat{\mathbf{h}}_c$ as an estimate of \mathbf{h}_c .
-

C. Performance Analysis

It is seen from Algorithm 1 that the estimate $\hat{\mathbf{h}}_c$ can be written as follows,

$$\hat{\mathbf{h}}_c = \sqrt{\hat{\kappa}_1} \hat{\mathbf{v}}_1 + \sqrt{(\hat{\kappa}_2)^+} \hat{\mathbf{v}}_2 \cdot \sqrt{-1}, \quad (65)$$

where $(\cdot)^+$ denotes the real part of a number. In the following we analyze the estimation distortion $\left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2$, assuming that the channel taps $\mathbf{h}_c = \mathbf{h}_{c0}$ obey (62). Note that the estimate $\hat{\mathbf{h}}_c$ is an approximation of \mathbf{h}_c under the perturbation of $\hat{\mathbf{R}}$ from its true value \mathbf{R} . The estimation distortion $\left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2$ is based on the perturbation of eigenvalues and eigenvectors of a Hermitian matrix.

The following result shows that the distortions of eigenvalues and eigenvectors are bounded under a small perturbation of the matrix \mathbf{R} [28], [29].

Proposition 3: We have the following bound on the perturbation of the eigenvalues,

$$|\kappa_i - \hat{\kappa}_i| \leq \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (66)$$

Let us consider the angle between the eigenvectors \mathbf{v}_i and $\hat{\mathbf{v}}_i$, $1 \leq i \leq L + 1$, denoted as θ_i . Then $\sin \theta_i$ is also bounded as follows,

$$\left| \sin \theta_i \right| \leq \frac{1}{\min_{j \neq i} |\kappa_j - \hat{\kappa}_i|} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (67)$$

□

Note that, from the definition of \mathbf{h} [c.f. (13)] and \mathbf{R} [c.f. (52)], we have the following,

$$\left\| \mathbf{R} - \hat{\mathbf{R}} \right\| \leq \sqrt{2} \left\| \mathbf{h} - \hat{\mathbf{h}} \right\|. \quad (68)$$

Naturally, the distortion $\left\| \mathbf{R} - \hat{\mathbf{R}} \right\|$ tends to zero, as the distortion $\left\| \mathbf{h} - \hat{\mathbf{h}} \right\|$ approaches zero. Note that according to (49) - (50), for sparse pilot sequences consisting of r repeated pilot sequence units, the expected distortion obeys

$$\mathbb{E} \left[\left\| \mathbf{h} - \hat{\mathbf{h}} \right\|^2 \right] \leq \frac{C}{r}, \quad (69)$$

for some constant C , which approaches zero as the number of repeated units r approaches infinity.

Let us first analyze the eigenvalues of $\hat{\mathbf{R}}$. According to (66), we have

$$\mathbb{E} \left[|\kappa_i - \hat{\kappa}_i|^2 \right] \leq \mathbb{E} \left[\left\| \hat{\mathbf{R}} - \mathbf{R} \right\|^2 \right] \leq \mathbb{E} \left[2 \left\| \mathbf{h} - \hat{\mathbf{h}} \right\|^2 \right] \leq \frac{2C}{r}. \quad (70)$$

Then, according to Chebyshev's inequality, given any distortion threshold $\delta > 0$, we have

$$\mathbb{P} \left(|\kappa_i - \hat{\kappa}_i| > \delta \right) \leq \frac{2C}{r\delta^2}, \quad (71)$$

which approaches zero as the number of repeated units r approaches infinity.

Based on the above arguments, we analyze the estimation induced perturbation imposed on linearly dependent pair and linearly independent pair \mathbf{h}_{cr} and \mathbf{h}_{ci} . We construct a set of "typical scenarios" conceiving the eigenvalues of $\hat{\mathbf{R}}$, whose probability approaches one as the number of repeated units r approaches infinity, and then prove that the distortion $\left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2$ can be reduced arbitrarily small for those events.

1) *Analysis for the Linearly Dependent \mathbf{h}_{cr} and \mathbf{h}_{ci} Pair:* For linearly dependent \mathbf{h}_{cr} and \mathbf{h}_{ci} , we only have a single positive eigenvalue κ_1 . Let us define the following event,

$$\mathcal{E}_1 \triangleq \left\{ \hat{\kappa}_1 > \frac{\kappa_1}{2} \right\}. \quad (72)$$

It can be shown that the probability $\mathbb{P}(\mathcal{E}_1)$ of event \mathcal{E}_1 approaches one as the number of repeated units r approaches infinity. More specifically, since

$$\left\{ |\hat{\kappa}_1 - \kappa_1| \leq \frac{\kappa_1}{2} \right\} \subseteq \mathcal{E}_1, \quad (73)$$

we have the following lower bound on $\mathbb{P}(\mathcal{E}_1)$ based on (71),

$$\mathbb{P}(\mathcal{E}_1) \geq \mathbb{P}\left(\left\{|\hat{\kappa}_1 - \kappa_1| \leq \frac{\kappa_1}{2}\right\}\right) \geq 1 - \frac{8C}{r\kappa_1^2}. \quad (74)$$

Under the event \mathcal{E}_1 , we can further bound the estimation distortion of $\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|^2$ based on (66) and (67) for the perturbation of the eigenvalues and eigenvectors, respectively. The following results show that the estimation distortion can be reduced arbitrarily small with the probability arbitrarily close to one, given that the number of repeated pilot units r approaches infinity.

Theorem 7: For any $\delta > 0$, we have the following upper bound on the estimation distortion

$$\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|^2 \leq 6\delta + \frac{24\delta^2}{\kappa_1}, \quad (75)$$

with the probability of at least

$$P_1(\delta, r) \triangleq 1 - \frac{2C}{r\delta^2} - \frac{8C}{r\kappa_1^2}. \quad (76)$$

Proof: Please refer to Appendix VIII-A. ■

2) *Analysis for the Linearly Independent Pair \mathbf{h}_{cr} and \mathbf{h}_{ci} :* Note that in this scenario, the eigenvalues obey $\kappa_1 \geq \kappa_2 > 0$, and $\kappa_i = 0$ for $i \geq 3$. We consider the following the event,

$$\mathcal{E}_2 = \left\{\kappa_1 > \kappa_2\right\} \cap \left\{\hat{\kappa}_1 > \frac{\kappa_1 + \kappa_2}{2} > \hat{\kappa}_2 > \frac{\kappa_2}{2}\right\}. \quad (77)$$

Note that the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^T$ are also the eigenvalues of $\mathbf{H}^T\mathbf{H}$. Then, it is seen that, if the two eigenvalues of $\mathbf{H}^T\mathbf{H}$ are the same, i.e. $\kappa_1 = \kappa_2 \triangleq \kappa$, then we have $\mathbf{H}^T\mathbf{H} = \kappa\mathbf{I}$. In such a scenario, we have that $\|\mathbf{h}_{ci}\| = \|\mathbf{h}_{cr}\|$ and $\mathbf{h}_{ci}^T\mathbf{h}_{cr} = 0$. Note that the above scenario happens with probability zero. Since $\mathbb{P}\left(\left\{\kappa_1 > \kappa_2\right\}\right) = 1$, we have

$$\begin{aligned} \mathbb{P}(\mathcal{E}_2) &\geq 1 - \mathbb{P}\left(|\kappa_1 - \hat{\kappa}_1| > \delta_{12}\right) - \mathbb{P}\left(|\kappa_2 - \hat{\kappa}_2| > \delta_{23}\right) \\ &= 1 - \frac{2C}{r\delta_1^2} - \frac{2C}{r\delta_2^2}, \end{aligned} \quad (78)$$

where $\delta_1 \triangleq 0.5 \cdot (\kappa_1 - \kappa_2)$ and $\delta_2 \triangleq 0.5 \cdot \min\{\kappa_1 - \kappa_2, \kappa_2\}$.

We analyze the distortion $\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|^2$ in case that \mathcal{E}_2 is satisfied. Again, we can prove that the estimation distortion $\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|^2$ can be made arbitrarily small with a probability arbitrarily close to one, given that the number of repeated pilot units r approaches infinity. More specifically, we have the following result.

Theorem 8: For any $\delta > 0$, we have the following upper bound on the estimation distortion,

$$\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|^2 \leq 8\delta + \left(\frac{8\kappa_1}{\delta_1^2} + \frac{8\kappa_2}{\delta_2^2}\right)\delta^2, \quad (79)$$

which holds with a probability of at least

$$P_2(\delta, r) \triangleq 1 - \frac{2C}{r\delta^2} - \frac{2C}{r\delta_1^2} - \frac{2C}{r\delta_2^2}. \quad (80)$$

Proof: Please refer to Appendix VIII-B. ■

VI. SIMULATION RESULTS

A. Simulation Setup and Parameters

We adopt the following system parameters. The receiver temperature is $T^o = 300K$; the load resistance is $R_L = 5M\Omega$; the PMT spreading factor is $\xi = 0.10$; the photon noise rate is $\Lambda_b = 14500s^{-1}$; and the bit rate is $R_b = 2Mbps$ and thus the symbol duration is $T = 1/R_b = 0.5\mu s$. The Planck Constant is $h = 6.62606957 \times 10^{-34}m^2kg/s$; the Boltzmann Constant is $k_e = 1.3806505 \times 10^{-23}J \cdot K^{-1}$; and the single electron charge is $e = 1.602 \times 10^{-19}C$.

The CIR taps are generated according to stochastic physics. Consider the transmission optical signal with a wavelength 260nm. Accordingly, each photon has energy $E_p = 7.65 \times 10^{-19}J$; the Rayleigh scattering coefficient is given by $k_r = 0.266 \times 10^{-3}m^{-1}$; the Mie scattering coefficient is given by $k_m = 0.284 \times 10^{-3}m^{-1}$; and the absorption coefficient is given by $k_a = 0.802 \times 10^{-3}m^{-1}$. Set $\gamma = 0.017$, $f = 0.5$, and $g = 0.72$ in the atmosphere scattering model, and the receiver area $S_R = 1.77 \times 10^{-4}m^2$. Let the positions of the transmitter and the receiver be $(100, 0, 0)m$ and $(0, 0, 0)m$, respectively. Let the pointing angles of the transmitter and the receiver be both $\pi/3$; and the transmitter beamwidth after beam expansion and receiver field of view be both $\pi/6$. Let the azimuth angles of the transmitter and the receiver measured from the positive x -axis be π and 0, respectively. We have that the number of channel taps is $L + 1 = 4$, and $\left[|h_0|^2, |h_1|^2, |h_2|^2, |h_3|^2 \right] = \left[6.5, 3.7, 1.0, 0.2 \right] \times 10^{-11}$. Here we generate the phases of h_0 , h_1 , h_2 , and h_3 randomly within the interval $[0, 2\pi)$, to test the effectiveness of the proposed two-stage channel estimation.

Let us now characterize the performance of the proposed channel estimation approach, for both a 0-1 random pilot sequence and the sparse sequence, where for the former the probabilities of both the 0 and 1 OOK symbols are 0.5. Recall that \mathbf{h}_c and $\hat{\mathbf{h}}_c$ denote the original and estimated CIR taps, respectively. The normalized channel estimation error, denoted as E_{CE} , is given by

$$E_{CE} = \frac{\left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2}{\left\| \mathbf{h}_c \right\|^2}, \quad (81)$$

which is employed as our metric of evaluating the channel estimation performance. In this section, we compare both the complementary cumulative distribution function (CCDF) of E_{CE} ,

as well as the mean of E_{CE} , for both the 0-1 random and the sparse pilot sequences. A total number of 500,000 channel realizations are simulated for generating the CCDF and the mean value of E_{CE} . The same lengths are adopted for both types of pilots, which consists of four repeated units of the pilots. The length of the pilot sequences is given by $\frac{(L+1)(L+2)}{2} \times 4 + 1 = 41$.

B. The Estimation Distortion for the Photon-counting Receiver

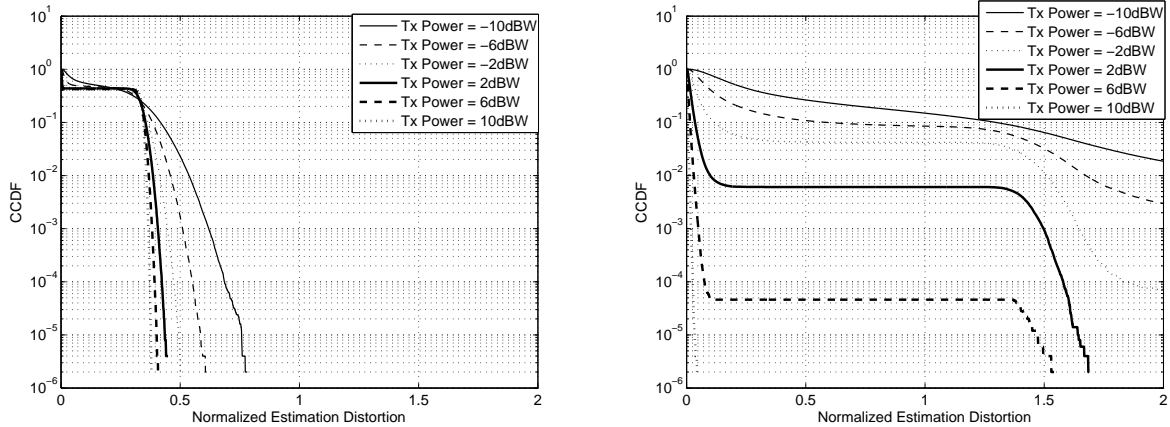
The CCDF of the normalized channel estimation error for the 0-1 random pilot sequence relying on the photon-counting receiver is shown in Figure 1, for two different pilot sequences and transmission power P varying from -10dBW to 10dBW . Observe for the random pilot sequence in Figure 1(a) that there is a probability of around 0.3 that the normalized distortion is about 0.3, resulting in a high “distortion floor”. By contrast, for the pilot sequence in Figure 1(b), such “distortion floor” is reduced. Hence Figure 1 demonstrates the importance of the pilot sequence design.

By contrast, the CCDF of the normalized channel estimation error E_{CE} recorded for the sparse pilot sequence and the photon-counting receiver is shown in Figure 2. Again, the “distortion floor” is reduced upon increasing the transmission power. We can contrast the average normalized estimation distortion of the random 0-1 pilot sequence and of the sparse sequence, by comparing Figure 1 to Figure 2. The distortion recorded in Figure 1(a) for the random 0-1 pilot sequence with a high distortion floor is above 0.1; while the sparse pilot sequence shows a slightly lower distortion in Figure 2.

Figure 3 compares the distortion of the proposed sparse pilot design with those of the 0-1 random pilots with high distortion floor and low distortion floor. It is seen that the distortion of the sparse pilot is close to that of the 0-1 random pilots with low distortion floor, and significantly lower than that of the 0-1 random pilots with high distortion floor.

C. The Estimation Distortion for the PMT Receiver

Next we study the CCDF of the normalized channel estimation error for the PMT receiver. Figure 4 shows the CCDF for both 0-1 random pilot sequence and the sparse sequence, for the transmission power ranging from -2dBW to 10dBW and for the PMT amplification factors of $A = 100, 200, 500, \text{ and } 1000$. As shown in both figures, the CCDFs plotted for the same transmission power are grouped, where the trend is that in each group the CCDFs decrease with the amplification factor. This is because a higher amplification factor can lead to a reduced



(a) The 0-1 random pilot sequence with high distortion floor. (b) The 0-1 random pilot sequence with low distortion floor.

Fig. 1. The CCDF of the normalized estimation distortion for the 0-1 random pilot sequence (high and low distortion floor) in conjunction with a photon-counting receiver.

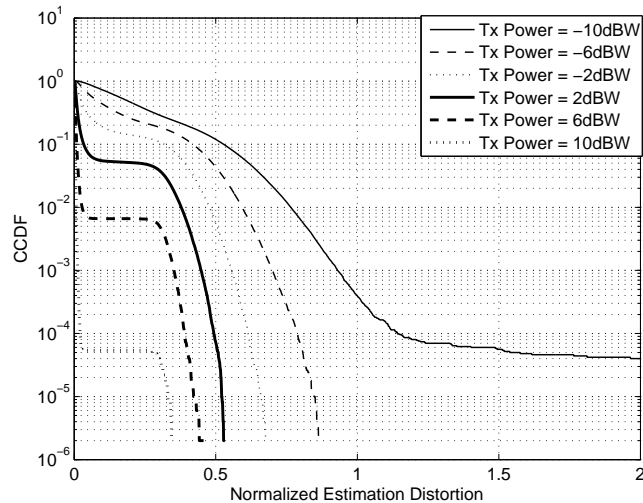


Fig. 2. The CCDF of the normalized estimation distortion using the sparse pilot sequence for a photon-counting receiver.

channel estimation distortion. It is seen in Figure 4 that sparse sequence leads to a lower distortion floor than the 0-1 random pilot sequence.

The average normalized estimation distortions recorded for both the 0-1 random pilot and the sparse pilot sequences are shown in Figure 5. It is seen that the normalized channel estimation error decreases with the amplification factor for both types of pilot sequences. Note that for the ultraviolet LD source, the transmission power can reach between 200mW and 400mW, which is between -7dBW and -4dBW . In such a power regime, the sparse pilot sequence is capable of

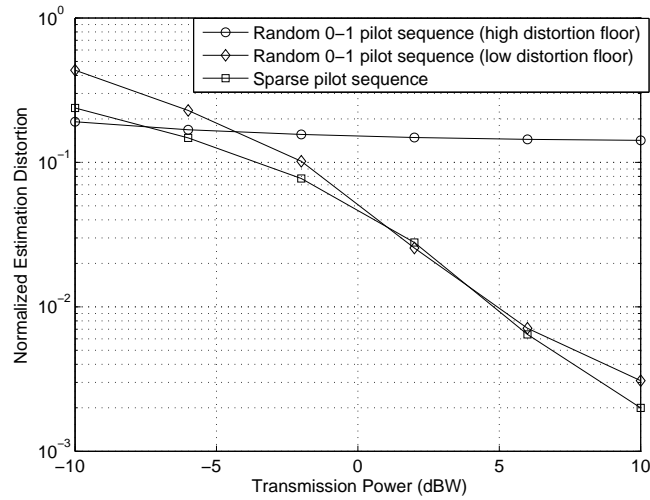
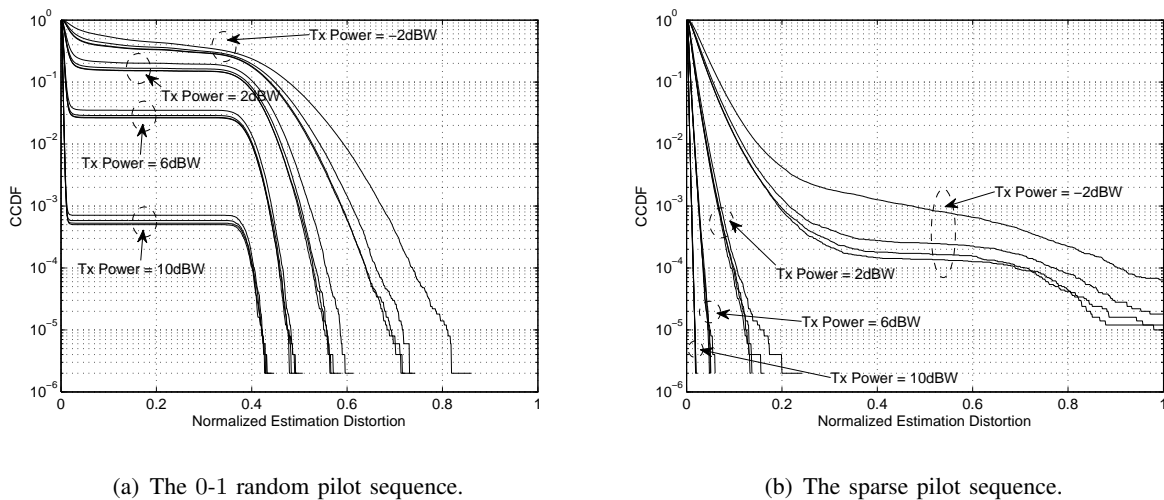


Fig. 3. The normalized estimation distortion for photon-counting receiver.

reducing the distortion compared with the 0-1 random pilot sequence.



(a) The 0-1 random pilot sequence.

(b) The sparse pilot sequence.

Fig. 4. The CCDF of the normalized estimation distortion with 0-1 random and sparse pilot sequence for PMT receiver.

For the sparse pilot, the mutual interference of “1”s can be well controlled by the pattern design, which contributes to the reduced distortion. For the randomly chosen pilot pattern, the number of “1”s and the number of “0”s are approximately equal, but for the sparse pilot design the number of “1”s is much smaller. This implies that the sparse pilot design can reduce the estimation distortion at the cost smaller average transmission power. In this work we assume the same peak transmission power for both the random pilots and the sparse pilot, which is the

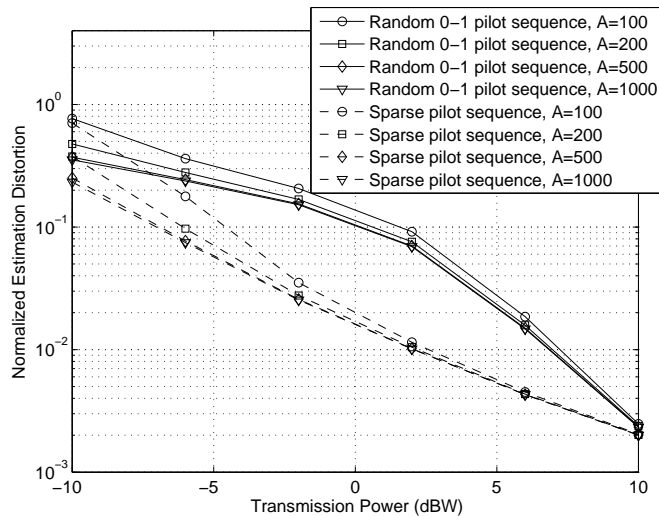


Fig. 5. The normalized estimation distortion for PMT receiver.

typical scenario of the external modulator for the UV laser.

VII. CONCLUSIONS

We have proposed a two-stage channel estimation framework for the energy-type receivers of optical wireless scattering communications. Based on the framework, we have also proposed a sparse structure for the pilot sequences as used, which guarantees a reduced computational complexity. We have also analyzed the performance of the proposed estimation approach for both stages. We prove that as the length of the pilot sequence approaches infinity, the probability of having an estimation distortion larger than any positive threshold may approach zero. Numerical results show that compared to the pilot sequence using 0-1 random bits, the proposed sparse structure is capable of reducing the estimation distortion.

VIII. APPENDIX

A. Proof of Theorem 7

The estimation distortion $\|\hat{\mathbf{h}}_c - \mathbf{h}_c\|$ may be expressed as

$$\begin{aligned}
 \|\hat{\mathbf{h}}_c - \mathbf{h}_c\| &= \left\| \left(\sqrt{\hat{\kappa}_1} - \sqrt{\kappa_1} \right) \hat{\mathbf{v}}_1 + \sqrt{\kappa_1} \left(\hat{\mathbf{v}}_1 - \mathbf{v}_1 \right) + \sqrt{\left(\hat{\kappa}_2 \right)^+} \hat{\mathbf{v}}_2 \cdot \sqrt{-1} \right\| \\
 &\leq \left| \sqrt{\hat{\kappa}_1} - \sqrt{\kappa_1} \right| + \sqrt{\kappa_1} \left\| \hat{\mathbf{v}}_1 - \mathbf{v}_1 \right\| + \left| \sqrt{\hat{\kappa}_2} - \sqrt{\kappa_2} \right| \\
 &\leq \sqrt{|\hat{\kappa}_1 - \kappa_1|} + \sqrt{|\hat{\kappa}_2 - \kappa_2|} + \sqrt{\kappa_1} \left\| \hat{\mathbf{v}}_1 - \mathbf{v}_1 \right\|. \tag{82}
 \end{aligned}$$

For the event \mathcal{E}_1 , if $\hat{\kappa}_1 > \frac{\kappa_1}{2}$, the angle between $\hat{\mathbf{v}}_1$ and \mathbf{v}_1 , denoted as θ_1 , satisfies

$$\left| \sin \theta_1 \right| \leq \frac{2}{\kappa_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (83)$$

Without loss of generality, let $\theta_1 < \pi/2$, and thus

$$\left\| \hat{\mathbf{v}}_1 - \mathbf{v}_1 \right\| = 2 \left| \sin \frac{\theta}{2} \right| = \frac{2 \sin \theta}{2 \cos \frac{\theta}{2}} \leq \frac{2\sqrt{2}}{\kappa_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (84)$$

Then according to (82) and (84), we have

$$\left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\| \leq \sqrt{|\hat{\kappa}_1 - \kappa_1|} + \sqrt{|\hat{\kappa}_2 - \kappa_2|} + \frac{2\sqrt{2}}{\sqrt{\kappa_1}} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (85)$$

Based on the perturbation of eigenvalues [c.f. (66)], we have the following upper bound on the estimation distortion

$$\begin{aligned} \left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2 &\leq 3|\hat{\kappa}_1 - \kappa_1| + 3|\hat{\kappa}_2 - \kappa_2| + \frac{24}{\kappa_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|^2 \\ &\leq 6 \left\| \mathbf{R} - \hat{\mathbf{R}} \right\| + \frac{24}{\kappa_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|^2. \end{aligned} \quad (86)$$

Note that, according to (71), for any $\delta > 0$, we have

$$\mathbb{P} \left(\left\| \hat{\mathbf{R}} - \mathbf{R} \right\| > \delta \right) \leq \frac{2C}{r\delta^2}; \quad (87)$$

and we arrive that $\left\| \hat{\mathbf{R}} - \mathbf{R} \right\| < \delta$ with a probability of at least $1 - \frac{2C}{r\delta^2}$. Noting that (84) holds for the event \mathcal{E}_1 , we have that (75) holds under the event

$$\tilde{\mathcal{E}}_1 \triangleq \mathcal{E}_1 \cap \left\{ \left\| \hat{\mathbf{R}} - \mathbf{R} \right\| < \delta \right\}. \quad (88)$$

Note that the probability of the event $\tilde{\mathcal{E}}_1$ can be bounded as follows,

$$\mathbb{P} \left(\tilde{\mathcal{E}}_1 \right) \geq \mathbb{P} \left(\mathcal{E}_1 \right) - \mathbb{P} \left(\left\| \hat{\mathbf{R}} - \mathbf{R} \right\| > \delta \right), \quad (89)$$

which leads to (76).

B. Proof of Theorem 8

Note that the estimate distortion satisfies

$$\begin{aligned} \left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\| &= \left\| \sqrt{\hat{\kappa}_1} \hat{\mathbf{v}}_1 - \sqrt{\kappa_1} \mathbf{v}_1 + \sqrt{\hat{\kappa}_2} \hat{\mathbf{v}}_2 \cdot \sqrt{-1} - \sqrt{\kappa_2} \mathbf{v}_2 \cdot \sqrt{-1} \right\| \\ &\leq \left\| \left(\sqrt{\hat{\kappa}_1} - \sqrt{\kappa_1} \right) \hat{\mathbf{v}}_1 + \sqrt{\kappa_1} \left(\hat{\mathbf{v}}_1 - \mathbf{v}_1 \right) \right\| + \left\| \left(\sqrt{\hat{\kappa}_2} - \sqrt{\kappa_2} \right) \hat{\mathbf{v}}_2 + \sqrt{\kappa_2} \left(\hat{\mathbf{v}}_2 - \mathbf{v}_2 \right) \right\| \\ &\leq \sqrt{|\hat{\kappa}_1 - \kappa_1|} + \sqrt{\kappa_1} \left\| \hat{\mathbf{v}}_1 - \mathbf{v}_1 \right\| + \sqrt{|\hat{\kappa}_2 - \kappa_2|} + \sqrt{\kappa_2} \left\| \hat{\mathbf{v}}_2 - \mathbf{v}_2 \right\|. \end{aligned} \quad (90)$$

According to (67), the angle between \mathbf{v}_1 and $\hat{\mathbf{v}}_1$, denoted as θ_1 , as well as that between \mathbf{v}_2 and $\hat{\mathbf{v}}_2$, denoted as θ_2 , satisfy

$$\left| \sin \theta_1 \right| \leq \frac{1}{\delta_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|, \quad \left| \sin \theta_2 \right| \leq \frac{1}{\delta_2} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (91)$$

Similar to (84), we have the following upper bound on the distortion $\left\| \mathbf{v}_1 - \hat{\mathbf{v}}_1 \right\|$ and $\left\| \mathbf{v}_2 - \hat{\mathbf{v}}_2 \right\|$,

$$\left\| \hat{\mathbf{v}}_1 - \mathbf{v}_1 \right\| \leq \frac{\sqrt{2}}{\delta_1} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|, \quad \left\| \hat{\mathbf{v}}_2 - \mathbf{v}_2 \right\| \leq \frac{\sqrt{2}}{\delta_2} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|. \quad (92)$$

According to (66) and (92), we have

$$\begin{aligned} \left\| \hat{\mathbf{h}}_c - \mathbf{h}_c \right\|^2 &\leq 4|\hat{\kappa}_1 - \kappa_1| + 4|\hat{\kappa}_2 - \kappa_2| + \frac{8\kappa_1}{\delta_1^2} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|^2 + \frac{8\kappa_2}{\delta_2^2} \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|^2 \\ &\leq 8 \left\| \mathbf{R} - \hat{\mathbf{R}} \right\| + \left(\frac{8\kappa_1}{\delta_1^2} + \frac{8\kappa_2}{\delta_2^2} \right) \left\| \mathbf{R} - \hat{\mathbf{R}} \right\|^2. \end{aligned} \quad (93)$$

Noting that (93) holds under the event \mathcal{E}_2 , (79) holds under the event

$$\tilde{\mathcal{E}}_2 \triangleq \mathcal{E}_2 \cap \left\{ \left\| \hat{\mathbf{R}} - \mathbf{R} \right\| < \delta \right\}. \quad (94)$$

We have the following lower bound on the probability $\mathbb{P}(\tilde{\mathcal{E}}_2)$,

$$\mathbb{P}(\tilde{\mathcal{E}}_2) \geq \mathbb{P}(\mathcal{E}_2) - \mathbb{P}\left(\left\| \hat{\mathbf{R}} - \mathbf{R} \right\| > \delta\right), \quad (95)$$

which leads to (80).

REFERENCES

- [1] H. Ding, G. Chen, A. K. Majumdar, B. M. Sadler, and Z. Xu, "Modeling of non-line-of-sight ultraviolet scattering channels for communication," *IEEE Journal Select. Areas Commun.*, vol. 27, no. 9, pp. 1535–1544, Dec. 2009.
- [2] Z. Xu and B. M. Sadler, "Ultraviolet communications: potential and state-of-the-art," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 67–73, May 2008.
- [3] Z. Xu, H. Ding, B. M. Sadler, and G. Chen, "Analytical performance study of solar blind non-line-of-sight ultraviolet short-range communication links," *Optics Letters*, vol. 33, no. 16, pp. 1860–1862, Aug. 2008.
- [4] G. Chen, Z. Xu, H. Ding, and B. M. Sadler, "Path loss modeling and performance trade-off study for short-range non-line-of-sight ultraviolet communications," *Optics Express*, vol. 17, no. 5, pp. 3929–3940, Mar. 2009.
- [5] Q. He, Z. Xu, and B. M. Sadler, "Performance of non-line-of-sight LED based ultraviolet communication receivers," *Optics Express*, vol. 18, no. 12, pp. 12226–12238, May 2010.
- [6] Z. Xu, H. Ding, B. M. Sadler, and G. Chen, "Analytical performance study of solar blind Non-Line-of-Sight ultraviolet short-range communication links," *Optics Letters*, vol. 33, no. 16, pp. 1860–1862, Aug. 2008.
- [7] A. D. Wyner, "Capacity and error component for the direct detection photon channel - Part I-II," *IEEE Trans. Info. Theory*, vol. 34, no. 6, pp. 1449–1471, Nov. 1988.
- [8] M. R. Frey, "Information capacity of the Poisson channel," *IEEE Trans. Info. Theory*, vol. 37, no. 2, pp. 244–256, Mar. 1991.

- [9] A. Lapidoth and S. M. Moser, "On the capacity of the discrete-time Poisson channel," *IEEE Trans. Info. Theory*, vol. 55, no. 1, pp. 303–322, Jan. 2009.
- [10] A. R. Calderbank and C. N. Georghiades, "Synchronizable codes for the optical OPPM channel," *IEEE Trans. Info. Theory*, vol. 40, no. 4, pp. 1097–1107, Jul. 1994.
- [11] N. Chatzidiamantis, G. K. Karagiannidis, and M. Uysal, "Generalized maximum-likelihood sequence detection for photon-counting free space optical systems," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3381–3385, Dec. 2010.
- [12] T. T. Nguyen and L. Lampe, "Coded multipulse pulse-position modulation for free-space optical communications," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1036–1041, Apr. 2010.
- [13] C. Abou-Rjeily and S. Haddad, "Cooperative fso systems: performance analysis and optimal power allocation," *IEEE J. Lightw. Technol.*, vol. 29, no. 7, pp. 1058–1065, Apr. 2011.
- [14] C. Abou-Rjeily and A. Slim, "Cooperative diversity for free-space optical communications: transceiver design and performance analysis," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 658–663, Mar. 2011.
- [15] C. Abou-Rjeily, "Achievable diversity orders of decode-and-forward cooperative protocols over Gamma-Gamma fading FSO links," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3919–3930, Sept. 2013.
- [16] Q. He, Z. Xu, and B. M. Sadler, "Non-Line-of-Sight serial relayed link for optical wireless communications," in *Proc. 2010 IEEE Military Communications Conference, San Jose, CA, Oct. 31-Nov. 3, 2010*.
- [17] C. Gong and Z. Xu, "Non-line of sight optical wireless relaying with the photon counting receiver: A count-and-forward protocol," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 376–388, Jan. 2015.
- [18] B. Huang, C. Gong, and Z. Xu, "Correlation study for single-input multiple-output optical wireless scattering channels," in *Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Nov. 2014*.
- [19] C. Gong and Z. Xu, "Linear receivers for optical wireless scattering communication with multiple photon detectors," in *IEEE Globecom, Austin, TX, Dec. 2014*.
- [20] —, "LMMSE SIMO receiver for short-range non-line-of-sight scattering communication," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5338–5349, Oct. 2015.
- [21] K. Kumar, R. Kaushik, and R. C. Jain, "Blind channel estimation for indoor optical wireless communication systems," in *Proc. 2015 International Conference on Signal Processing and Communication (ICSC), India, Greater Noida, Mar. 16 - 18, 2015*.
- [22] S. K. Hashemi, Z. Ghassemlooy, L. Chao, and D. Benhaddou, "Channel estimation for indoor diffuse optical ofdm wireless communications," in *Proc. 5th International Conference on Broadband Communications, Networks and Systems, London, Sept. 8 - 11, 2008*.
- [23] D. Wu, Z. Wang, R. Wang, J. He, Q. Zuo, and H. Zhao, "Channel estimation for asymmetrically clipped optical orthogonal frequency division multiplexing optical wireless communications," *IET Commun.*, vol. 6, no. 5, pp. 532–540, May 2012.
- [24] M. A. El-Shimy and S. Hranilovic, "Binary-input non-line-of-sight solar-blind uv channels: Modeling, capacity and coding," *IEEE J. Opt. Commun. Netw.*, vol. 12, no. 4, pp. 1008–1017, Dec. 2012.
- [25] C. Gong and Z. Xu, "Channel estimation and signal detection for optical wireless scattering communication with inter-symbol interference," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5326–5337, Oct. 2015.
- [26] X. Zhang, C. Gong, and Z. Xu, "Estimation of NLOS optical wireless communication channels with laser transmitters," in *48th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Nov. 2014*.
- [27] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Princeton University Press, 2005.
- [28] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation," *SIAM J. Nume. Anal.*, no. 7, pp. 1–46, 1970.
- [29] B. N. Parlett, *The symmetric eigenvalue problem*. Prentice-Hall, Englewood Cliffs, NJ, 1980.