Compact Optimal Pilot Design for Channel Estimation in MIMO VLC Systems

Ashok D. R. and A. Chockalingam
Department of ECE, Indian Institute of Science, Bangalore 560012

***Abstract*—In visible light communication (VLC) systems, like in radio frequency (RF) wireless communication systems, the channel has to be estimated to aid transceiver operations such as data detection and precoding. The LEDs which serve as VLC transmitters have to maintain a desired average intensity as they are simultaneously used for both data transmission as well as lighting, and each LED also has a maximum power constraint. In this paper, we are concerned with channel estimation in VLC systems under these constraints. Specifically, we consider extensions to combinational codes for optimal pilot transmission assuming that the receiver employs a zero-forcing decoder. We construct optimal codes which need much fewer pilot channel uses for channel estimation compared to that needed by simple concatenation of combinational codes. We propose a recursive algorithm to construct such codes that are short in length and meet the optimality constraints. The bit error performance achieved using the estimates of the channel obtained using the proposed codes are shown to be quite close to that with perfect channel knowledge.**

Keywords – *Visible light communication, channel estimation, MIMO-VLC, compact pilot design, combinational code.*

I. INTRODUCTION

Visible light communication (VLC) systems are proving to be attractive for wireless communications in indoor and vehicular environments [1],[2]. In VLC systems, light emitting diodes (LED) and photo diodes (PD) serve as optical wireless transmitters and receivers, respectively. VLC transceivers are simple and cost effective compared to RF communication transceivers. Simultaneous lighting and data transmission capability, security in closed-room applications, and free visible light spectrum are some key advantages of VLC. In VLC, information is conveyed through modulation of the optical intensity radiated by the LED(s). Consequently, the information signals that intensity modulate the LEDs are real and non-negative. Communication using multiple LEDs and multiple PDs in multiple-input multiple-output (MIMO) configuration is an attractive means to achieve high data rates [3],[4]. In indoor applications where LEDs have to support both lighting as well as communication needs, the LEDs have to maintain a desired average intensity, and each LED also must satisfy a maximum intensity constraint.

Like in RF wireless communication systems, estimates of the channel coefficients are needed in VLC systems as well for the purposes of implementing transceiver functions such as data detection at the receiver and precoder at the transmitter. In MIMO VLC systems, estimating the columns of the channel matrix by illuminating one LED at a time could render meeting the average intensity requirement without exceeding the maximum intensity constraint per LED difficult when more transmit LEDs are involved. This opens up the problem of finding optimal channel estimation schemes that take into account the aforementioned practical intensity constraints on the transmitter LEDs.

Channel estimation in VLC systems has been considered in the literature [6]-[10]. Channel estimation schemes for optical OFDM systems such as DCO-OFDM [11] and ACO-OFDM [12] and their performance are extensively studied [6]-[9]. The other work related to channel estimation in VLC include the problem of finding the scheme that minimizes the noise variance under practical considerations such as maximum power and average power constraints [10]. While average power constraints arise from the need to ensure a desired illumination of lighting, maximum power constraint arises from the LED's maximum power rating. The sufficient conditions under which a pilot code is optimal for the above problem has been reported in [10]. An example of a code which satisfies these conditions, namely combinational code, has also been reported.

Though the combinational code is optimal, the number of pilot channel uses required increases significantly as the in number of LEDs is increased. Whereas, optimal codes with fewer channel uses for pilot transmission are desired for increasing number of LEDs. Construction of such codes forms the key focus of this paper. Specifically, we construct optimal codes with short lengths for VLC channel estimation with average and maximum power constraints. The pilot matrices are chosen to have certain structure, which we call as concatenated circulant structure. The condition under which these pilot matrices are optimal when zero-forcing (ZF) detection is used is shown. Also, a recursive algorithm which generates these optimal pilot matrices for given number of LEDs is presented. It is shown that the optimal codes generated by the recursive algorithm are defined for much smaller number of channel uses compared to that of codes obtained by simple concatenation of combinational codes. We also evaluated the bit error rate (BER) performance of spatial multiplexing (SMP) in a MIMO VLC system when channel matrices estimated using the constructed codes are used in data detection. Results show that the BER performance achieved using the estimated channel is quite close (within about 1 dB) to that achieved with perfect channel knowledge.

The rest of this paper is organized as follows. The indoor MIMO VLC system model and the VLC channel estimation problem are introduced in Sec. II. The conditions for optimal pilot design and combinational code are presented in Sec. III. The proposed codes with concatenated circulant structure are presented in Sec. IV. Results and discussions are presented in Sec. V. Conclusions are presented in Sec. VI.

II. SYSTEM MODEL AND CHANNEL ESTIMATION PROBLEM

A. *Indoor MIMO VLC system model*

Consider an indoor MIMO VLC system with N_t LEDs (transmitter) and N_r PDs (receiver), where the LEDs are assumed to have Lambertian radiation pattern [5]. In a given

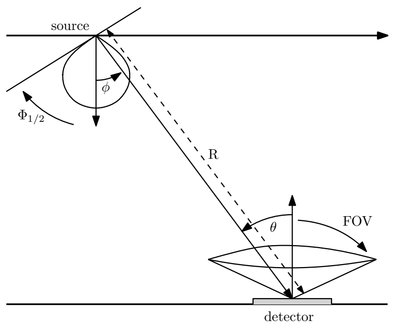

Fig. 1. Schematic showing various angles which determine LOS path gain.

signaling interval, an LED is either OFF or emits light with some intensity. Let $\mathbf{x} = [x_1 \ x_2 \cdots x_{N_t}]^T$ be the $N_t \times 1$ transmit vector, where x_i is the light intensity emitted by the ith LED, which depends on the modulation scheme used. The MIMO VLC channel matrix $\mathbf{H}$ is of the order $N_r \times N_t$ and its $(i,j)th$ element h_{ij} is the path gain from the jth LED to the ith PD, $j = 1, \cdots, N_t$, and $i = 1, \cdots, N_r$. The expression for the LOS path gain h_{ij} is given by [5]

$$h_{ij} = \frac{n+1}{2\pi} \cos^n \phi_{ij} \, \cos\theta_{ij} \frac{A}{R_{ij}^2} \mathrm{rect}\Big(\frac{\theta_{ij}}{FOV}\Big), \qquad (1)$$

where ϕ_{ij} is the angle of emergence from the jth source (LED) with respect to the normal at the source, n is the mode number of the radiating lobe which is given by $n = \frac{-\ln(2)}{\ln \cos \Phi_{\frac{1}{2}}}$, $\Phi_{\frac{1}{2}}$ is the half-power semiangle of the LED, θ_{ij} is the angle of incidence at the ith PD, A is the area of the PD, R_{ij} is the distance between the jth LED and the ith PD, FOV is the field-of-view of the PD, and rect$(x) = 1$, if $|x| \leq 1$, and rect$(x) = 0$, if $|x| > 1$. See Fig. 1 for the definition of various angles in the model. Assuming the electrical-to-optical conversion factor to be unity, the $N_r \times 1$ received signal vector at the receiver in the electrical domain is given by

$$\mathbf{y} = a\mathbf{H}\mathbf{x} + \mathbf{n}, \qquad (2)$$

where a is the responsivity of the PD (in Amperes/Watt) and $\mathbf{n} = [n_1 \ n_2 \cdots n_{N_r}]^T$ is the noise vector. The electrical noise variables n_is in $\mathbf{n}$ are modeled as i.i.d. real AWGN with zero mean and variance σ^2. The SNR at a PD in the electrical domain is defined as $\frac{(aP_r)^2}{\sigma^2}$, where P_r is the total received optical power and σ^2 is the noise power at a PD. The total power received at the ith PD is given by $\mathbb{E}\{(\mathbf{H}_i\mathbf{x})^2\}$. Therefore, the average received optical power is given by $\mathbb{E}\{\|\mathbf{H}\mathbf{x}\|^2\} = \frac{1}{N_r}\sum_{i=1}^{N_r} \mathbb{E}\{(\mathbf{H}_i\mathbf{x})^2\}$, where $\mathbf{H}_i$ is the ith row of $\mathbf{H}$, $\|.\|$ is the Euclidean norm operator, $\mathbb{E}\{.\}$ is the expectation operator, and the expectation is w.r.t. the signal vector $\mathbf{x}$. Hence, the average SNR at the receiver in the electrical domain is given by $\overline{\gamma} = \frac{a^2}{\sigma^2 N_r}\sum_{i=1}^{N_r} \mathbb{E}\{(\mathbf{H}_i\mathbf{x})^2\}$, and the corresponding E_b/N_0 is given by $E_b/N_0 = \frac{\overline{\gamma}}{\eta}$, where η is the rate of the modulation scheme used, in bits per channel use (bpcu).

B. Channel estimation problem in VLC

In this subsection, we formulate the VLC channel estimation problem by considering a MISO system for simplicity of exposition though it extends to MIMO VLC systems in a straightforward manner. Consider a MISO VLC system consisting of N_t LEDs and one PD. The LOS path gains from the LEDs to the PD are to be estimated in the pilot transmission phase. Let h_j denote the LOS path gain from jth LED to the PD, N_{cu} ($N_{cu} \geq N_t$) denote the number of pilot channel uses used to estimate these path gains, and a_{jm} denote the intensity that is transmitted by jth LED in the mth channel use. It is assumed that h_js remains constant during the estimation phase (this is a reasonably valid assumption as indoor VLC channels are slow fading in nature). Without loss of generality, a responsitivity of $a = 1$ Amp/Watt is assumed. The observed variable y_m received by the PD in the mth channel use based on the system model considered in Sec. II-A is given by

$$y_m = \sum_{j=1}^{N_t} h_j a_{jm} + n_m, \quad m = 1, 2, \cdots N_{cu}. \qquad (3)$$

Grouping all observables over N_{cu} channel uses (3) can be written in vector form as

$$\mathbf{y}_p = \mathbf{h}\mathbf{A} + \mathbf{n}_p, \qquad (4)$$

where $\mathbf{y}_p = [y_1 \, y_2 \cdots y_{N_{cu}}]$, $\mathbf{h} = [h_1 \, h_2 \cdots h_{N_t}]$, $\mathbf{A} = [a_{jm}]_{N_t \times N_{cu}}$, and $\mathbf{n}_p = [n_1 \, n_2 \cdots n_{N_{cu}}]$. Here, the matrix $\mathbf{A}$ is the pilot matrix. In RF communication systems, typically $N_{cu} = N_t$ with pilot matrix $\mathbf{A}$ to be the identity matrix is used. Such a scheme in VLC will be inadequate when maximum and average power constraints are taken into consideration. Assume each LED supports a maximum intensity of μ lumens and the average illumination required from each LED is Φ lumens. Then the entries of the pilot matrix $\mathbf{A}$ should be chosen such that

$$0 \leq a_{jm} \leq \mu, \quad \forall j, m, \qquad (5)$$

$$\frac{1}{N_{cu}} \sum_{m=1}^{N_{cu}} a_{jm} = \Phi, \qquad \forall j. \qquad (6)$$

Now, assuming that a ZF decoder is used at the receiver, the estimate of $\mathbf{h}$ is given by

$$\begin{aligned} \hat{\mathbf{h}} &= \mathbf{y}_p\mathbf{W} \\ &= \mathbf{h} + \mathbf{n}_p\mathbf{W} \end{aligned} \qquad (7)$$

where $\mathbf{W} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$. For a fixed matrix $\mathbf{A}$, the mean square error (MSE) of the estimate in (7) is obtained as

$$\begin{aligned} \mathbb{E}\{\| \mathbf{h} - \hat{\mathbf{h}} \|^2\} &= \mathbb{E}\{\| \mathbf{n}_p\mathbf{W} \|^2\} \\ &= \mathbb{E}\{(\mathbf{n}_p\mathbf{W})(\mathbf{n}_p\mathbf{W})^T\} \\ &= \mathrm{Tr}\left(\mathbb{E}\{(\mathbf{n}_p\mathbf{W})^T(\mathbf{n}_p\mathbf{W})\}\right) \\ &= \mathrm{Tr}\left(\mathbf{W}^T\mathbb{E}\{\mathbf{n}_p{}^T\mathbf{n}_p\}\mathbf{W}\right) \\ &= \sigma^2\,\mathrm{Tr}\left(\mathbf{W}^T\mathbf{W}\right) \\ &= \sigma^2\,\mathrm{Tr}\left(\left((\mathbf{A}\mathbf{A}^T)^{-1}\right)^T\mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\right) \\ &= \sigma^2\,\mathrm{Tr}\left(\left((\mathbf{A}\mathbf{A}^T)^{-1}\right)^T\right) \\ &= \sigma^2\,\mathrm{Tr}\left((\mathbf{A}\mathbf{A}^T)^{-1}\right). \end{aligned} \qquad (8)$$

From (8) it is clear that minimizing the MSE amounts to minimizing the metric $\mathrm{Tr}\left(\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\right)$. Hence, the optimal pilot matrix is obtained as

$$\mathbf{A}_{opt} = \underset{\mathbf{A}}{\mathrm{argmin}}\ \mathrm{Tr}\left(\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\right), \text{ given constraints (5), (6).} \tag{9}$$

Note that the optimal pilot matrix in (9) is independent of the statistics of $\mathbf{h}$ (because of the ZF operation), which also implies that the problem is independent of the location of the receiver PD with respect to transmitter LEDs. Also, if there are multiple PDs at the receiver, the optimal matrix is the same as that in (9) for each PD. Therefore the optimal pilot matrix is given by (9) for MIMO VLC configuration as well.

III. Conditions for the optimal pilot design and Combinational code

Though the problem (9) is non-convex, the pilot matrix $\mathbf{A}$ is optimal solution to the problem if it satisfies the conditions given in Theorem 1 [10].

Theorem 1. *If $\mathbf{A} = [a_{jm}]_{N_t \times N_{cu}}$ constrained by* (5) *and* (6) *satisfies the following conditions:*

C1. *For all j and $j' \neq j$, both $\sum_m a_{jm}^2$ and $\sum_m (a_{jm})(a_{j'm})$ are constants,*

C2. $\sum_j a_{jm} = \Phi N_t$ *for all m,*

C3. *For all j and m, $a_{jm} = 0$ or μ,*

then $\mathbf{A}$ is an optimal solution to (9).

A. Combinational code

For N_t LEDs, each LED having maximum power of μ and average power Φ, combinational code is the pilot transmission scheme in which the columns of pilot matrix are chosen to be all possible permutation of $N_t\Theta$ μ's and $N_t(1-\Theta)$ 0's, where $\Theta = \frac{\Phi}{\mu}$. The number of channel uses the combinational code requires is given by $N_{cu} = \binom{N_t}{N_t\Theta}$. The combinational code and its time repetition satisfy all the conditions in Theorem 1. Hence they are optimal solutions to the problem (9) [10]. If a matrix $\mathbf{A}'$ satisfies the condition C1, C2, and C3 of Theorem 1, then it is easy to verify that any matrix obtained by interchanging the columns of $\mathbf{A}'$ also satisfies conditions C1, C2, and C3. This means a code obtained by interchanging columns of the optimal code is also optimal.

Example 1. *If $\mu = 1$, $N_t = 4$, and $\Phi = \frac{1}{2}$, then $N_{cu} = \binom{4}{2} = 6$ and the corresponding combinational code is*

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}. \tag{10}$$

IV. Proposed codes with concatenated circulant structure

A. Concatenated circulant structure

A pilot matrix $\mathbf{A}$ is said to have concatenated circulant structure if it is formed by a concatenation of matrices $\mathbf{A}_1, \mathbf{A}_2, \cdots \mathbf{A}_L$, i.e, $\mathbf{A} = [\mathbf{A}_1\,\mathbf{A}_2 \cdots \mathbf{A}_L]$, where L is a arbitrary positive finite integer and the lth matrix $\mathbf{A}_l$ is described by $N_t \times 1$ vector $\mathbf{v}_l$ and parameter β_l, $l = 1, 2, \cdots L$ as

$$\mathbf{A}_l = \left[\mathbf{v}_l^0\ \mathbf{v}_l^1 \cdots \mathbf{v}_l^{\beta_l - 1}\right]_{N_t \times \beta_l}, \tag{11}$$

where $\mathbf{v}_l^\rho$ denotes vector $\mathbf{v}_l$ circularly shifted by ρ in downward direction. For each l, the parameter β_l should be such that $\mathbf{v}_l^{\beta_l} = \mathbf{v}_l^0$. The number of columns in $\mathbf{A}$ will be the sum of the number of columns of $\mathbf{A}_l$, $l = 1, 2, \cdots N_t$. So, we have

$$N_{cu} = \sum_{l=1}^{L} \beta_l. \tag{12}$$

Now consider a vector $\mathbf{v}$ of size $N_t \times 1$. Let $\alpha_{\mathbf{v}}$ denote the minimum positive number such that $\mathbf{v}^{\alpha_{\mathbf{v}}} = \mathbf{v}^0$. Circularly shifting a vector by $\rho_1 + \rho_2$ is same as first circularly shifting by ρ_1 and then by ρ_2. Using this fact, it can be shown that for any positive integer ρ, $\mathbf{v}^\rho = \mathbf{v}^{\rho \bmod \alpha_{\mathbf{v}}}$. Let the matrix $\mathbf{B}_{\mathbf{v}}$ corresponding to vector $\mathbf{v}$ be

$$\mathbf{B}_{\mathbf{v}} = \left[\mathbf{v}^0\ \mathbf{v}^1 \cdots \mathbf{v}^{\alpha_{\mathbf{v}} - 1}\right]_{N_t \times \alpha_{\mathbf{v}}}. \tag{13}$$

If a vector of size N_t is circularly shifted by N_t all the elements in the vector get back to its initial position and we get back th same vector, i.e $\mathbf{v}^{N_t} = \mathbf{v}^0$. Hence N_t is multiple of $\alpha_{\mathbf{v}}$. Let this multiplying factor be $T_{\mathbf{v}}$, i.e., $N_t = T_{\mathbf{v}}\alpha_{\mathbf{v}}$. The repeated concatenation of matrix $\mathbf{B}_{\mathbf{v}}$ by $T_{\mathbf{v}}$ times gives us a circulant matrix with first column as $\mathbf{v}^0$ or simply $\mathbf{v}$. We denote this circulant matrix as $\mathbf{C}_{\mathbf{v}}$. The matrix $\mathbf{C}_{\mathbf{v}}^T$ will also be circulant and will have $\mathbf{v}$ as the first row. Using the fact that the eigen vectors of a circulant matrix are the columns of DFT matrix and the eigen values are DFT of the first row of the circulant matrix considered, and by applying eigen value decomposition on $\mathbf{C}_{\mathbf{v}}^T$, we have

$$\mathbf{C}_{\mathbf{v}}^T = \mathbf{F}_{N_t} \mathrm{diag}\{V(0), V(1), \cdots V(N_t - 1)\} \mathbf{F}_{N_t}^H, \tag{14}$$

where $\mathbf{F}_{N_t}$ denotes the $N_t \times N_t$ normalized DFT matrix, $V(k)$ is the $(k+1)$th entry of the DFT of the vector $\mathbf{v}$, $k = 0, 1 \cdots N_t - 1$, and $\mathrm{diag}\{p_1, p_2 \cdots p_{N_t}\}$ denotes the diagonal matrix with $p_1, p_2 \cdots p_{N_t}$ as the diagonal entries. Considering the entries of vector $\mathbf{v}$ to be real, $\mathbf{C}_{\mathbf{v}}^T = \mathbf{C}_{\mathbf{v}}^H$. Hence we will have

$$\mathbf{C}_{\mathbf{v}} = \mathbf{F}_{N_t} \mathrm{diag}\{V^*(0), V^*(1), \cdots V^*(N_t - 1)\} \mathbf{F}_{N_t}^H, \tag{15}$$

$$\mathbf{C}_{\mathbf{v}}\mathbf{C}_{\mathbf{v}}^T = \mathbf{F}_{N_t} \mathrm{diag}\{|V(0)|^2, |V(1)|^2, \cdots |V(N_t - 1)|^2\} \mathbf{F}_{N_t}^H, \tag{16}$$

where $V^*(k)$ is conjugate of $V(k)$ and $|V(k|^2$ is square of the magnitude of $V(k)$. Hence the eigen values of $\mathbf{C}_{\mathbf{v}}\mathbf{C}_{\mathbf{v}}^T$ are the square of the magnitudes of the DFT values of vector $\mathbf{v}$. Since $\mathbf{C}_{\mathbf{v}}$ is $T_{\mathbf{v}}$ times repeated concatenation of $\mathbf{B}_{\mathbf{v}}$, we have

$$\mathbf{B}_{\mathbf{v}}\mathbf{B}_{\mathbf{v}}^T = \left(\frac{1}{T_{\mathbf{v}}}\right) \mathbf{C}_{\mathbf{v}}\mathbf{C}_{\mathbf{v}}^T. \tag{17}$$

Hence the eigen values of $\mathbf{B}_{\mathbf{v}}\mathbf{B}_{\mathbf{v}}T$ are $|V(k)|^2/T_{\mathbf{v}}$, $k = 0, 1, \cdots N_t - 1$. Now, since $\mathbf{v}_l^{\beta_l} = \mathbf{v}_l^0$ for each l, $\beta_l = t_l\alpha_{\mathbf{v}_l}$, which means the matrix $\mathbf{A}_l$ is t_l times repeated concatenation of matrix $\mathbf{B}_{\mathbf{v}_l}$. Hence we have

$$\mathbf{A}_l\mathbf{A}_l^T = t_l\mathbf{B}_{\mathbf{v}_l}\mathbf{B}_{\mathbf{v}_l}^T = \left(\frac{t_l}{T_{\mathbf{v}_l}}\right)\mathbf{C}_{\mathbf{v}_l}\mathbf{C}_{\mathbf{v}_l}^T = \left(\frac{\beta_l/\alpha_{\mathbf{v}_l}}{N_t/\alpha_{\mathbf{v}_l}}\right)\mathbf{C}_{\mathbf{v}_l}\mathbf{C}_{\mathbf{v}_l}^T = \left(\frac{\beta_l}{N_t}\right)\mathbf{C}_{\mathbf{v}_l}\mathbf{C}_{\mathbf{v}_l}^T. \quad (18)$$

So the eigen values of $\mathbf{A}_l\mathbf{A}_l^T$ are $\left(\frac{\beta_l}{N_t}\right)|V_l(k)|^2$, $k = 0, 1\cdots N_t-1$ and the eigen vectors are the columns of DFT matrix, i.e., $\left[1\ \exp\left(-j\frac{2\pi k}{N_t}\right)\cdots\exp\left(-j\frac{2\pi(N_t-1)k}{N_t}\right)\right], k = 0, 1, \cdots N_t - 1$. The matrix $\mathbf{A}$ being concatenation of $\mathbf{A}_1, \mathbf{A}_2, \cdots \mathbf{A}_L$, we have

$$\mathbf{A}\mathbf{A}^T = \sum_{l=1}^{L}\mathbf{A}_l\mathbf{A}_l^T. \quad (19)$$

Since the eigen vectors of $\mathbf{A}_l\mathbf{A}_l^T$ are same for all l, the eigen values of $\mathbf{A}\mathbf{A}^T$ are given by summing the eigen values of $\mathbf{A}_l\mathbf{A}_l^T$. So $\mathbf{A}\mathbf{A}^T$ will have $\sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right)|V_l(k)|^2$, $k = 0, 1\cdots N_t - 1$ as the eigen values and columns of DFT matrix as the eigen vectors.

Any code having concatenated circulant structure can be described completely by the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_L$ and corresponding parameters $\beta_1, \beta_2, \cdots \beta_L$ respectively. We denote code with these vectors and parameters as $\mathbf{circ}\left(\{\mathbf{v}_1, \beta_1\}, \{\mathbf{v}_2, \beta_2\}, \cdots \{\mathbf{v}_L, \beta_L\}\right)$.

B. Criteria for code to be optimal

Consider $\mathbb{A}_{\mu,\Theta}^{N_t}$ to be the set of all codes having concatenated circulant structure, where all the representative vectors $\mathbf{v}_l \in \mathbb{P}_{\mu,\Theta}^{N_t}$, $l = 1, 2\cdots L$, where $\mathbb{P}_{\mu,\Theta}^{N_t}$ is the set of all possible permutation of $N_t\Theta$ μ's and $N_t(1-\Theta)$ 0's. Any pilot matrix $\mathbf{A} \in \mathbb{A}_{\mu,\Theta}^{N_t}$ satisfies the condition C2 and C3 of Theorem 1 and also meets the constraints (5) and (6). This is justified by the following.

1) The columns of matrix $\mathbf{A}$ are the representative vectors and their circular shifts. Since the entries of these vectors are either μ or 0, all the entries of $\mathbf{A}$ are also either μ or 0. So constraint (5) and condition C3 of Theorem 1 are satisfied.
2) Each column of $\mathbf{A}$ has $N_t\Theta$ entries to be μ's and remaining entries to be 0's. So the sum of the entries in each column is $N_t\Theta\mu = \Phi N_t$. Hence condition C2 of Theorem 1 is satisfied.
3) For each l, the fraction of entries being μ in the representative vector $\mathbf{v}_l$ is Θ. As the matrix $\mathbf{C}_{\mathbf{v}_l}$ is circulant, each row will have the same elements as that of the columns, and as the columns are shifted versions of $\mathbf{v}_l$, the fraction of elements being μ in each row of $\mathbf{C}_{\mathbf{v}_l}$ will be Θ. Since matrices $\mathbf{C}_{\mathbf{v}_l}$ and $\mathbf{A}_l$ are repeated concatenated version of the same matrix $\mathbf{B}_{\mathbf{v}_l}$, the fraction of entries being μ in each row will be the same for both. So the fraction of entries being μ in each row of $\mathbf{A}_l$ will be Θ for each value of l. Hence the fraction of entries being μ in matrix $\mathbf{A}$, which is concatenated version of $\mathbf{A}_1, \mathbf{A}_2\cdots\mathbf{A}_L$ will also be Θ. Since the entries other than μ are 0, the average value of each row will be $\Theta\mu = \Phi$. So $\mathbf{A}$ also satisfies the constraint (6).

We have seen that any code from the set $\mathbb{A}_{\mu,\Theta}^{N_t}$ meets the constraints (5) and (6) and also satisfies conditions C1 and C2 of Theorem 1. If the code also satisfies C3, then it is optimal. To know under what condition the code from the set $\mathbb{A}_{\mu,\Theta}^{N_t}$ satisfy C3, we will make use of the following Lemma.

Lemma 1. *For a matrix $\mathbf{A} = [a_{jm}]_{N_t\times N_{cu}}$, given that the eigen vectors of $\mathbf{A}\mathbf{A}^T$ are columns of DFT matrix, for all j and $m \neq j$, both $\sum_m a_{jm}^2$ and $\sum_m(a_{jm})(a_{j'm})$ are constants if and only if the eigen values except the one which corresponds to the eigen vector $[1\ 1\cdots 1]^T$ are equal.*

Proof: Let $\sum_m a_{jm}^2 = \psi$ and $\sum_m(a_{jm})(a_{j'm}) = \omega$, for all j and $m \neq j$. This means that the matrix $\mathbf{A}\mathbf{A}^T$ is

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} \psi & \omega & \cdots & \omega \\ \omega & \psi & \cdots & \omega \\ \vdots & \vdots & \ddots & \vdots \\ \omega & \omega & \cdots & \psi \end{bmatrix}, \quad (20)$$

or simply $\mathbf{A}\mathbf{A}^T$ is a circulant matrix with the first row as $\mathbf{z} = [\psi\ \omega\cdots\omega]$. So the eigen values of $\mathbf{A}\mathbf{A}^T$ will be DFT of $\mathbf{z}$. Let DFT of $\mathbf{z}$ be denoted by Z, then

$$\begin{aligned} Z(k) &= \sum_{n=0}^{N_t-1}\mathbf{z}(n)\exp\left(-j\frac{2\pi nk}{N_t}\right), \ k = 0, 1, \cdots N_t - 1 \\ &= \psi + \omega\sum_{n=1}^{N_t-1}\exp\left(-j\frac{2\pi nk}{N_t}\right) \\ &= \psi - \omega + \omega\sum_{n=0}^{N_t-1}\exp\left(-j\frac{2\pi nk}{N_t}\right). \end{aligned} \quad (21)$$

Now, using the relation

$$\sum_{n=0}^{N-1}\exp\left(-j\frac{2\pi nk}{N}\right) = \begin{cases} N, & k = pN, \ p \in \mathbb{Z} \\ 0, & k \neq pN, \ p \in \mathbb{Z}, \end{cases} \quad (22)$$

we can write $Z(k)$ in (21) as

$$Z(k) = \begin{cases} \psi - \omega + \omega(0), & k = 1, \cdots N_t - 1 \\ \psi - \omega + \omega N_t, & k = 0, \end{cases} \quad (23)$$

$$Z(k) = \begin{cases} \psi - \omega, & k = 1, \cdots N_t - 1 \\ \psi + \omega(N_t - 1), & k = 0. \end{cases} \quad (24)$$

From (24), we can say that DFT of $[\psi\ \omega\cdots\omega]$ is $[\psi + (N_t-1)\,\omega\quad \psi-\omega\cdots\psi-\omega]$. Therefore, the eigen value corresponding to $[1\ 1\cdots 1]^T$ is $\psi + (N_t - 1)\,\omega$ and all other eigen values are equal to $\psi - \omega$. This completes the proof of forward implication.

Regarding the proof of converse, let the eigen value corresponding to $[1\ 1\cdots 1]^T$ be γ and all other eigen values be equal to δ. Since it is given that eigen vectors of $\mathbf{A}\mathbf{A}^T$ are columns of DFT, $\mathbf{A}\mathbf{A}^T$ will be circulant and first row will be IDFT of eigen values. Now, by replacing γ and δ in place of ψ and ω in the above DFT-IDFT pair, IDFT of $[\gamma\ \delta\cdots\delta]$ will be $(1/N_t)[\gamma + (N_t-1)\,\delta\quad \gamma-\delta\cdots\gamma-\delta]$. So $\mathbf{A}\mathbf{A}^T$ is circulant with first row as $(1/N_t)[\gamma + (N_t-1)\,\delta\quad \gamma-\delta\cdots\gamma-\delta]$.

Algorithm 1 Obtaining combinational code in terms of representative vectors

Inputs: N_t, μ, Φ
Initialize: $\Theta = \frac{\Phi}{\mu}$; set $\mathbb{S} = \mathbb{P}^{N_t}_{\mu,\Theta}$; $l = 1$
while ($\mathbb{S} \neq \emptyset$) **do**
  choose a vector $\mathbf{v}_l$ randomly from the set $\mathbb{S}$
  $\beta_l = \alpha_{\mathbf{v}_l}$
  $\mathbb{S} = \mathbb{S} \backslash \{\mathbf{v}' | \mathbf{v}'$ is circular shift of $\mathbf{v}_l\}$
  $l = l + 1$
end while
$L = l - 1$
Output: $\mathbf{A} = \mathbf{circ}\left(\{\mathbf{v}_1, \beta_1\}, \{\mathbf{v}_2, \beta_2\}, \cdots \{\mathbf{v}_L, \beta_L\}\right)$

So $\sum_m a_{jm}^2 = \frac{\gamma + (N_t - 1)\,\delta}{N_t}$ and $\sum_m (a_{jm})(a_{j'm}) = \frac{\gamma - \delta}{N_t}$. Therefore, $\sum_m a_{jm}^2$ and $\sum_m (a_{jm})(a_{j'm})$ are constants.

For any code $\mathbf{A} \in \mathbb{A}^{N_t}_{\mu,\Theta}$, if condition C1 is satisfied, then it is optimal as it automatically satisfies other conditions. Since it has the concatenated circulant structure, the eigen values of $\mathbf{A}\mathbf{A}^T$ are $\sum_{l=1}^{L} \left(\frac{\beta_l}{N_t}\right) |V_l(k)|^2$, $k = 0, 1 \cdots N_t - 1$. Now, using Lemma 1, we can say that condition C1 is satisfied if all the eigen values except the one corresponding to eigen vector $[1\ 1 \cdots 1]^T$ are equal. Eigen value of $\mathbf{A}\mathbf{A}^T$ corresponding to $[1\ 1 \cdots 1]^T$ is $\sum_{l=1}^{L} \left(\frac{\beta_l}{N_t}\right) |V_l(0)|^2$. So $\mathbf{A}$ will be optimal if $\sum_{l=1}^{L} \left(\frac{\beta_l}{N_t}\right) |V_l(k)|^2$ is same for $k = 1, 2, \cdots N_t - 1$.

By definition, the vectors in the set $\mathbb{P}^{N_t}_{\mu,\Theta}$ are same as the columns of the combinational code. So, Algorithm 1 generates combinational code, where the columns are arranged such that it has concatenated circulant structure. This means that it can be represented in terms of the representative vectors and corresponding parameters. Using Algorithm 1, it can be verified that the combinational code in example 1 is same as the code $\mathbf{circ}(\{[1\ 1\ 0\ 0]^T, 4\}, \{[1\ 0\ 1\ 0]^T, 2\})$. All the combinational codes with columns arranged appropriately are examples of optimal codes among the code belonging to the set $\mathbb{A}^{N_t}_{\mu,\Theta}$. There can be some codes other than combinational codes which are also optimal in the set $\mathbb{A}^{N_t}_{\mu,\Theta}$. One such code is given in Example 2 below.

Example 2. Given $N_t = 10$, $\mu = 1$, $\Phi = \frac{1}{2}$, the code $\mathbf{A} = \mathbf{circ}(\{\mathbf{v}_1, 10\}, \{\mathbf{v}_2, 10\}, \{\mathbf{v}_3, 10\}, \{\mathbf{v}_4, 10\}, \{\mathbf{v}_5, 10\}, \{\mathbf{v}_6, 4\})$, where $\mathbf{v}_1 = [1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0]^T$, $\mathbf{v}_2 = [1\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 0]^T$, $\mathbf{v}_3 = [1\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0]^T$, $\mathbf{v}_4 = [1\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 0]^T$, $\mathbf{v}_5 = [1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0]^T$, $\mathbf{v}_6 = [1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0]^T$. The eigen values of $\mathbf{A}\mathbf{A}^T$ are [135 15 15 15 15 15 15 15 15 15]. Hence, using Lemma 1, we can say that code $\mathbf{A}$ is optimal. For this code $N_{cu} = 54$.

C. Recursive algorithm

The number of channel uses $N_{cu} = \binom{N_t}{N_t\Theta}$ used by combinational code increases rapidly with increase in N_t. For any even N_t, N_{cu} is maximum for the case of $\Theta = \frac{1}{2}$. So for this case, we propose Algorithm 2 to generate optimal codes for larger N_t, having much smaller N_{cu} compared to that in combinational code.

Algorithm 2 Algorithm to generate code $\mathbf{A}_{2N_t} \in \mathbb{A}^{2N_t}_{\mu,\frac{1}{2}}$ using the representative vectors and corresponding parameters of the code $\mathbf{A}_{N_t} \in \mathbb{A}^{N_t}_{\mu,\frac{1}{2}}$

Inputs: Representative vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_L \in \mathbb{P}^{N_t}_{\mu,\frac{1}{2}}$ and corresponding parameters $\beta_1, \beta_2, \cdots \beta_L$ of the code $\mathbf{A}_{N_t} \in \mathbb{A}^{N_t}_{\mu,\frac{1}{2}}$
The vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots \mathbf{u}_{2L+1} \in \mathbb{P}^{2N_t}_{\mu,\frac{1}{2}}$ and their corresponding parameters $\beta_1^{'}, \beta_2^{'} \cdots \beta_{2L+1}^{'}$ are obtained as follows (the vectors are indexed starting from 0)
for $l = 1, 2, \cdots L$
  for $m = 0, 1, \cdots N_t - 1$
    $\mathbf{u}_l(2m) = \mathbf{u}_l(2m+1) = \mathbf{v}_l(m)$
    $\mathbf{u}_{L+l}(2m) = \mathbf{v}_l(m)$
    $\mathbf{u}_{L+l}(2m+1) = \mu - \mathbf{v}_l(m)$
  end for
  ${\beta'}_l = 2\beta_l$
  ${\beta'}_{L+l} = 2\beta_l$
  $V_l = \mathrm{DFT}(\mathbf{v}_l)$
end for
$\mathbf{u}_{2L+1} = [\mu\ 0\ \mu\ 0 \cdots \mu\ 0]^T$
$\beta_{2L+1}^{'} = \frac{8}{N_t \mu^2} \sum_{l=1}^{L} \left(\frac{\beta_l}{N_t}\right) |V_l(\frac{N_t}{2})|^2$
Output:
$\mathbf{A}_{N_t} = \mathbf{circ}\left(\{\mathbf{u}_1, \beta_1^{'}\}, \{\mathbf{u}_2, \beta_2^{'}\}, \cdots \{\mathbf{u}_{2L+1}, \beta_{2L+1}^{'}\}\right)$

Proposition 1. *If $\mathbf{A}_{N_t} \in \mathbb{A}^{N_t}_{\mu,\frac{1}{2}}$ is optimal, then the code $\mathbf{A}_{2N_t} \in \mathbb{A}^{2N_t}_{\mu,\frac{1}{2}}$ generated by Algorithm 2 using $\mathbf{A}_{N_t}$ is also optimal.*

Proof: The proof is given in Appendix A.

Example 3. In this example, we present an optimal code generated by Algorithm 2 for $N_t = 8$, $\mu = 1$, $\Theta = \frac{1}{2}$ from $N_t = 48$, $\mu = 1$, $\Theta = \frac{1}{2}$. The code $\mathbf{A}_{N_t} = \mathbf{circ}(\{\mathbf{v}_1, 4\}, \{\mathbf{v}_2, 2\})$, where $\mathbf{v}_1 = [1\ 1\ 0\ 0]^T$ and $\mathbf{v}_2 = [1\ 0\ 1\ 0]^T$. The eigen values of $\mathbf{A}_{N_t}\mathbf{A}_{N_t}^T$ are $[6\ 2\ 2\ 2]$, making $\mathbf{A}_{N_t}$ optimal. Using Algorithm 2 we have, $\mathbf{u}_1 = [1\ 1\ 1\ 1\ 0\ 0\ 0\ 0]^T$, $\mathbf{u}_2 = [1\ 1\ 0\ 0\ 1\ 1\ 0\ 0]^T$, $\mathbf{u}_3 = [1\ 0\ 1\ 0\ 0\ 1\ 0\ 1]^T$, $\mathbf{u}_4 = [1\ 0\ 0\ 1\ 1\ 0\ 0\ 1]^T$, $\mathbf{u}_5 = [1\ 0\ 1\ 0\ 1\ 0\ 1\ 0]^T$, and $\mathbf{A}_{2N_t} = \mathbf{circ}(\{\mathbf{u}_1, 8\}, \{\mathbf{u}_2, 4\}, \{\mathbf{u}_3, 8\}, \{\mathbf{u}_4, 4\}, \{\mathbf{u}_5, 4\})$. The eigen values of $\mathbf{A}_{2N_t}\mathbf{A}_{2N_t}^T$ are [56 8 8 8 8 8 8 8], and so $\mathbf{A}_{2N_t}$ is also optimal.

V. Results and discussion

In this section, we present the results on the optimal codes generated, MSE performance, and BER performance of the proposed optimal codes and combinational codes. The number of channel uses N_{cu} for different N_t for combinational codes and optimal codes generated using Algorithm 2 are given in Table I. The codes that are compared with combinational codes for $N_t = 8, 16, 32$ are generated by repeatedly using Algorithm 2 starting with $N_t = 4$ combinational code as the input. Similarly, codes for $N_t = 12, 24$ are generated using Algorithm 2 starting with $N_t = 6$ combinational code, and the code for $N_t = 20$ is obtained from $N_t = 10$ scheme

N_t	N_{cu}	
	Optimal code generated using proposed recursive algorithm	Combinational code
8	28	70
12	88	924
16	120	12870
20	228	184756
24	368	2704156
32	496	601080390

TABLE I
COMPARISON OF OPTIMAL CODE GENERATED USING RECURSIVE ALGORITHM WITH COMBINATIONAL CODE IN TERMS OF N_{cu} FOR $\Theta = \frac{1}{2}$.

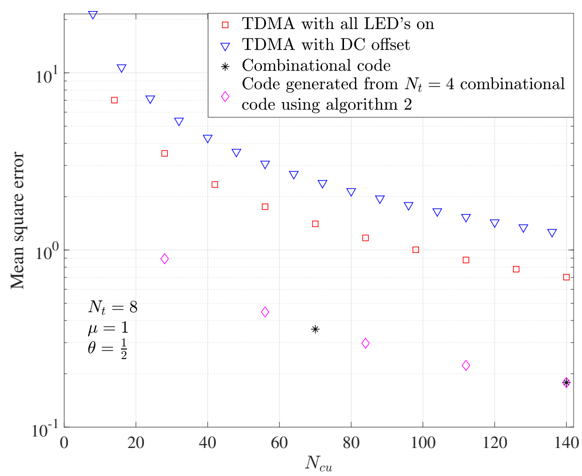

Fig. 2. Mean square error of various channel estimation schemes as a function of N_{cu}.

that is given in Example 2. From Table I, we see that the codes generated using the proposed recursive algorithm have much fewer channel uses compared to those of combinational codes, and that this advantage gets increasingly more for larger values of N_t.

Like in the case of combinational codes, time repetition of optimal codes generated using the recursive algorithm is also optimal. If we consider $N_t = 8$, N_{cu} for combinational code is 70 and its time repetition will be multiple of 70, whereas N_{cu} for the code obtained using recursive algorithm is 28 and its time repetition will be multiple of 28. The MSE for these schemes along with the those of other schemes which satisfy constraints (5) and (6) are shown in Fig. 2. TDMA with all LEDs ON is a scheme in which the pilot

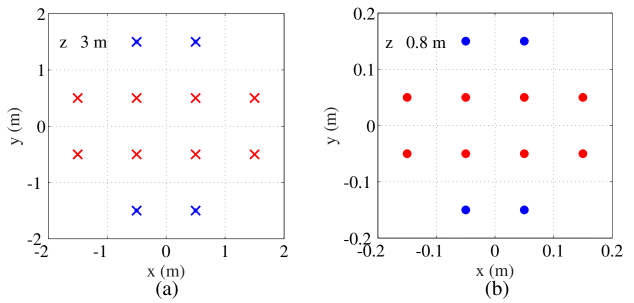

Fig. 3. (a) Location of LEDs and (b) location of PDs considered for BER simulations.

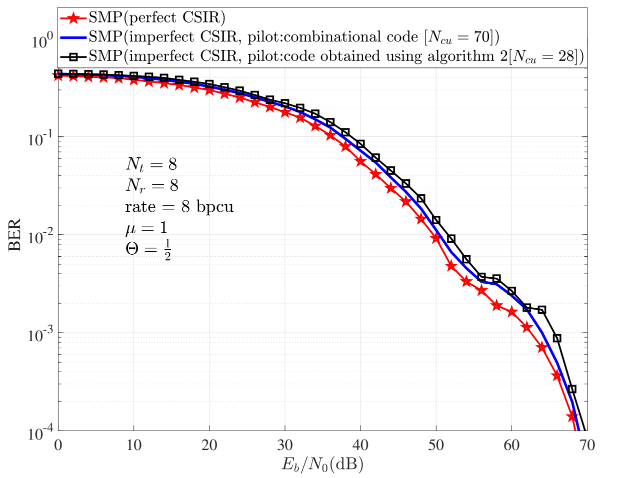

Fig. 4. Comparison of BER performance of SMP in 8×8 MIMO VLC system with estimated and perfect channel knowledge.

matrix is either concatenation of $\mu \mathbf{I}_{N_t}$ and $\mu \mathbf{1}_{N_t \times (N_t - 2)}$ or its time repetition, where $\mathbf{I}_N$ refers to $N \times N$ identity matrix and $\mathbf{1}_{N_1 \times N_2}$ refers to matrix of order $N_1 \times N_2$ where all the elements are 1. TDMA with DC offset is a scheme in which pilot matrix is either $\frac{\mu}{N_t - 1}\left(\frac{N_t}{2}\mathbf{I}_{N_t} + \frac{N_t - 2}{2}\mathbf{1}_{N_t \times N_t}\right)$ or its time repetition. From Fig. 2, we see that, being optimal, the combinational codes and the codes obtained using the proposed recursive algorithm achieve significantly better MSE performance compared to the TDMA based schemes.

Finally, in Figs. 4 and 5 we present the BER performance of SMP in 8×8 and 12×12 MIMO VLC systems, respectively, for the cases of perfect and estimated channel knowledge. Each LED uses ON/OFF keying where intensity 0 or μ is radiated based on information bits. The location of LEDs and PDs used for the simulation are shown in Fig. 3. For 12×12 system, all LEDs and PDs (marked in blue and red) are used. For 8×8 system, only those LEDs and PDs marked in red are used. From Figs. 4 and 5, we observe that the degradation of BER performance with estimated channel compared to that with perfect channel knowledge is quite small. For example, the SNR degradation is not more than 1 dB compared to the performance with perfect channel knowledge which demonstrates the effectiveness of the proposed optimal codes with small number of pilot channel uses.

VI. CONCLUSIONS

We considered the problem of channel estimation in VLC systems under maximum power and average power constraints. We chose the pilot matrices to have a structure, which we called as concatenated circulant structure. We gave the representation of these pilot matrices in terms of representative vectors and a parameter associated with each vector. We also presented the condition under which pilot matrices having this structure are optimal. We proposed a recursive algorithm which generates optimal code for $2N_t$ given the optimal code for N_t for the case of $\Theta = \frac{1}{2}$. The recursive algorithm gave the optimal codes having much fewer pilot channel uses compared to combinational codes. So the proposed codes can be practical even for large N_t, making it suitable for applications requiring high data rates. Simulation

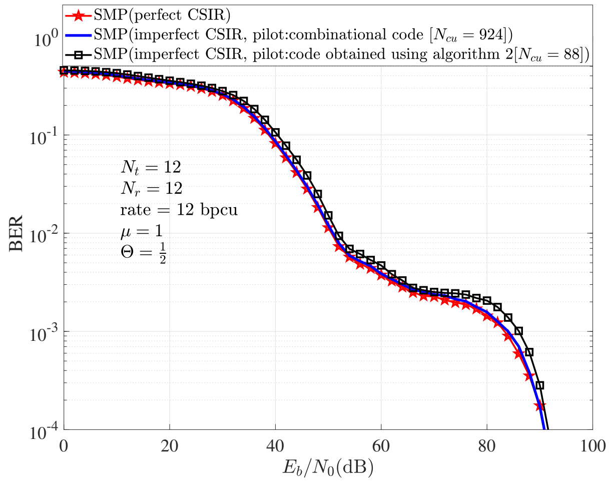

Fig. 5. Comparison of BER performance of SMP in 12×12 MIMO VLC system with estimated and perfect channel knowledge.

results showed that there is only about 1 dB SNR degradation in BER performance when estimated channel using the proposed codes is used (compared to the performance with perfect channel knowledge). Obtaining compact optimal codes for Θ other than $\frac{1}{2}$ is a topic for further research.

Appendix A: Proof of Proposition 1

Since $\mathbf{A}_{N_t}$ is optimal, from Lemma 1, the eigen values of $\mathbf{A}_{N_t}\mathbf{A}_{N_t}^T$ other than that corresponding to the eigen vector $[1\ 1 \cdots 1]^T$ are equal. Let these eigen values be equal to ω, i.e.,

$$\sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right)|V_l(k)|^2 = \omega, \ \text{for}\ k = 1, 2, \cdots N_t - 1. \tag{25}$$

Using Lemma 1, we can say that code $\mathbf{A}_{2N_t}$ will be optimal if all the eigen values of $\mathbf{A}_{2N_t}\mathbf{A}_{2N_t}^T$ are same except the one corresponding to eigen vector $[1\ 1 \cdots 1]^T$. So we will have to show that $\sum_{l=1}^{2L+1}\left(\frac{\beta_l'}{2N_t}\right)|U_l(k)|^2$ is same for $k = 1, 2, \cdots 2N_t - 1$, where V_l and U_l are DFT of the vectors $\mathbf{v}_l$ and $\mathbf{u}_l$, respectively. For any $\{l \in 1, 2, \cdots L\}$ and $\{k \in 0, 1, \cdots 2N_t - 1\}$, we have

$$\begin{aligned}
U_l(k) &= \sum_{m'=0}^{2N_t-1}\mathbf{u}_l(m')\exp\left(-j\frac{2\pi m'k}{2N_t}\right) \\
&= \sum_{m=0}^{N_t-1}\mathbf{u}_l(2m)\exp\left(-j\frac{2\pi(2m)k}{2N_t}\right) + \\
&\quad \sum_{m=0}^{N_t-1}\mathbf{u}_l(2m+1)\exp\left(-j\frac{2\pi(2m+1)k}{2N_t}\right) \\
&= \sum_{m=0}^{N_t-1}\mathbf{v}_l(m)\exp\left(-j\frac{2\pi mk}{N_t}\right) + \\
&\quad \sum_{m=0}^{N_t-1}\mathbf{v}_l(m)\exp\left(-j\frac{2\pi mk}{N_t}\right)\exp\left(-j\frac{\pi k}{N_t}\right) \\
&= V_l(k) + V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right),
\end{aligned} \tag{26}$$

which, using (22), can be written as

$$U_l(k) = \begin{cases} V_l(k) + V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right), & k \neq 0, N_t \\ 2V_l(0), & k = 0 \\ 0, & k = N_t, \end{cases} \tag{27}$$

and

$$\begin{aligned}
U_{L+l}(k) &= \sum_{m'=0}^{2N_t-1}\mathbf{u}_{L+l}(m')\exp\left(-j\frac{2\pi m'k}{2N_t}\right) \\
&= \sum_{m=0}^{N_t-1}\mathbf{u}_{L+l}(2m)\exp\left(-j\frac{2\pi(2m)k}{2N_t}\right) + \\
&\quad \sum_{m=0}^{N_t-1}\mathbf{u}_{L+l}(2m+1)\exp\left(-j\frac{2\pi(2m+1)k}{2N_t}\right) \\
&= \sum_{m=0}^{N_t-1}\mathbf{v}_l(m)\exp\left(-j\frac{2\pi mk}{N_t}\right) + \\
&\quad \sum_{m=0}^{N_t-1}(\mu - \mathbf{v}_l(m))\exp\left(-j\frac{2\pi mk}{N_t}\right)\exp\left(-j\frac{\pi k}{N_t}\right) \\
&= V_l(k) + \mu\sum_{m=0}^{N_t-1}\exp\left(-j\frac{2\pi mk}{N_t}\right)\exp\left(-j\frac{\pi k}{N_t}\right) \\
&\quad - V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right),
\end{aligned} \tag{28}$$

which again, using (22), can be written as

$$U_{L+l}(k) = \begin{cases} V_l(k) - V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right), & k \neq 0, N_t \\ N_t\mu, & k = 0 \\ 2V_l(N_t) - N_t\mu, & k = N_t. \end{cases} \tag{29}$$

As $\mathbf{v}_l \in \mathbb{P}_{\mu,\frac{1}{2}}^{2N_t}$, $\frac{N_t}{2}$ entries are μ and remaining entries are 0. So we have

$$\begin{aligned}
V_l(N_t) &= \sum_{m=0}^{N_t-1}\mathbf{v}_l(m)\exp\left(-j\frac{2\pi m N_t}{N_t}\right) \\
&= \sum_{m=0}^{N_t-1}\mathbf{v}_l(m) = \frac{N_t}{2}\mu.
\end{aligned} \tag{30}$$

Substituting (30) in (29), we get

$$U_{L+l}(k) = \begin{cases} V_l(k) - V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right) & , k \neq 0, N_t \\ N_t\mu, & k = 0 \\ 0, & k = N_t. \end{cases} \tag{31}$$

In Algorithm 2, we have chosen vector $\mathbf{u}_{2L+1}$ to have even entries as μ and odd entries as 0. So we have

$$\begin{aligned}
U_{2L+1}(k) &= \sum_{m'=0}^{2N_t-1} \mathbf{u}_{2L+1}(m') \exp\left(-j\frac{2\pi m' k}{2N_t}\right) \\
&= \sum_{m=0}^{N_t-1} \mu \exp\left(-j\frac{2\pi(2m)k}{2N_t}\right) \\
&= \mu \sum_{m=0}^{N_t-1} \exp\left(-j\frac{2\pi mk}{N_t}\right) \\
&= \begin{cases} N_t\mu, & k = 0, N_t \\ 0, & k \neq 0, N_t. \end{cases} \qquad (32)
\end{aligned}$$

Let $\widetilde{V}_l(k) = V_l(k)\exp\left(-j\frac{\pi k}{N_t}\right)$. Now, we have

$$\begin{aligned}
&\sum_{l=1}^{2L+1} \left(\frac{\beta_l'}{2N_t}\right) |U_l(k)|^2 \\
&= \sum_{l=1}^{L} \left(\frac{\beta_l'}{2N_t}\right) |U_l(k)|^2 + \sum_{l=1}^{L} \left(\frac{\beta_{L+l}'}{2N_t}\right) |U_{L+l}(k)|^2 \\
&\quad + \left(\frac{\beta_{2L+1}'}{2N_t}\right) |U_{2L+1}(k)|^2 \\
&= \sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) \left|V_l(k) + \widetilde{V}_l(k)\right|^2 + \sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) \left|V_l(k) - \widetilde{V}_l(k)\right|^2 \qquad (33) \\
&\quad + 0\,,\ k \neq 0, N_t \\
&= \sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right)\left(\left|V_l(k)+\widetilde{V}_l(k)\right|^2 + \left|V_l(k)-\widetilde{V}_l(k)\right|^2\right),\ k \neq 0, N_t \\
&= \sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) 2\left(|V_l(k)|^2 + |\widetilde{V}_l(k)|^2\right),\ k \neq 0, N_t \\
&= \sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) 2\left(|V_l(k)|^2 + |V_l(k)|^2\right),\ k \neq 0, N_t \\
&= 4\sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) |V_l(k)|^2,\ k \neq 0, N_t \\
&= 4\,\omega,\ k \neq 0, N_t, \qquad (34)
\end{aligned}$$

where (34) is obtained by using (25). Now, for $k = N_t$, (33) is given by

$$\sum_{l=1}^{2L+1} \left(\frac{\beta_l'}{2N_t}\right) |U_l(N_t)|^2 = 0 + 0 + \left(\frac{\beta_{2L+1}'}{2N_t}\right) (N_t\mu)^2. \quad (35)$$

Substituting for β_{2L+1}' as given in Algorithm 2, (35) can be written as

$$\begin{aligned}
\sum_{l=1}^{2L+1} \left(\frac{\beta'_l}{2N_t}\right) |U_l(N_t)|^2 &= 4\sum_{l=1}^{L}\left(\frac{\beta_l}{N_t}\right) |V_l(N_t/2)|^2 \\
&= 4\,\omega. \qquad (36)
\end{aligned}$$

So $\sum_{l=1}^{2L+1} \left(\frac{\beta'_l}{2N_t}\right) |U_l(k)|^2$ is same and equal to $4\,\omega$ for $k = 1, 2, \cdots 2N_t - 1$. This means the eigen values of $\mathbf{A}_{2N_t}$ are equal except the one corresponding to eigen vector $[1\ 1 \cdots 1]^T$. Hence Proposition 1 is proved.

References

[1] H. Elgala, R. Mesleh, and H. Haas, "Indoor optical wireless communication: potential and state-of-the-art," *IEEE Commun. Mag.*, vol. 49, no. 9, pp. 56-62, Sep. 2011.

[2] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: a survey, potential and challenges," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 4, pp. 2047-2077, 4th quarter 2015.

[3] T. Fath and H. Haas, "Performance comparison of MIMO techniques for optical wireless communications in indoor environments," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 733-742, Feb. 2013.

[4] A. K. Gupta and A. Chockalingam, "Performance of MIMO modulation schemes with imaging receivers in visible light communication," *IEEE/OSA J. Lightwave Tech.*, vol. 36, no. 10, pp. 1912-1927, May 2018.

[5] J. Barry, J. Kahn, W. Krause, E. Lee, and D. Messerschmitt, "Simulation of multipath impulse response for indoor wireless optical channels," *IEEE J. Sel. Areas in Commun.*, vol. 11, no. 3, pp. 367-379, Apr. 1993.

[6] Y. S. Hussein, M. Y. Alias, and A. A. Abdulkafi, "On performance analysis of LS and MMSE for channel estimation in VLC systems," *Proc. IEEE CSPA'2016*, Mar. 2016.

[7] O. ayli, H. Doan, and E. Panayirci, "On channel estimation in DC Biased optical OFDM systems over VLC channels," *Proc. ATC'2016*, Oct. 2016.

[8] H. Doan, O. ayli, and Erdal Panayirci, "Pilot assisted channel estimation for asymmetrically clipped optical OFDM over visible light channels," *Proc. IEEE BlackSeaCom'2016*, Jun. 2016.

[9] 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, Mar. 2017.

[10] S-W. Ho, A. A. Saed, L. Lai, and C. W. Sung, "Coding and bounds for channel estimation in visible light communications and positioning," *IEEE J. Sel. Areas in Commun.*, vol. 36, no. 1, pp. 34-44, Jan. 2018 .

[11] O. Gonzlez, R. Prez-Jimnez, S. Rodriguez, J. Rabadn, and A. Ayala, "OFDM over indoor wireless optical channel," *Proc. IEE Optoelectron.*, vol. 152, no. 4, pp. 199-204, Aug. 2005.

[12] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electron. Lett.*, vol. 42, no. 6, pp. 370-372, Mar. 2006.