Sparse Bayesian RVM Regression Based Channel Estimation for IM/DD OFDM-VLC Systems with Reduced Training Overhead

Chen Chen, Wen-De Zhong, and Lifan Zhao
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798
Email: chen0884@e.ntu.edu.sg, {ewdzhong, lfzhao}@ntu.edu.sg

***Abstract*—We propose a novel channel estimation technique for intensity modulation/direct detection (IM/DD) based orthogonal frequency division multiplexing visible light communication (OFDM-VLC) systems, utilizing sparse Bayesian dual-variate relevance vector machine (RVM) regression. By exploiting sparse Bayesian framework, dual-variate RVM regression can provide accurate estimation of the real and imaginary parts of the complex channel response, and therefore the channel response can be estimated to perform channel compensation. Simulation results show that a 200 Mb/s OFDM-VLC system using sparse Bayesian RVM regression based channel estimation with only one complex training symbol (TS) achieves nearly the same bit error rate (BER) performance as the system using conventional time domain averaging (TDA) based channel estimation with a total of 20 complex TSs, indicating a significant reduction of training overhead. Moreover, by employing a fast marginal likelihood maximization method, the sparse Bayesian RVM regression based channel estimation can be computational efficient for practical application in high-speed OFDM-VLC systems.**

***Index Terms*—Orthogonal frequency division multiplexing; visible light communication; channel estimation; sparse Bayesian; relevance vector machine**

I. Introduction

White light-emitting diodes (LEDs) enabled visible light communications (VLCs) have attracted ever-increasing interest in recent years, owing to its advantages such as license-free spectrum, cost-effective front-ends, high security, and strong immunity to electromagnetic interference [1]–[3]. VLC using LEDs is very promising for many practical applications such as indoor wireless communications, localization or positioning, networking, and sensing [4], [5]. Although VLC is considered as a promising technology for future indoor communications, the development of high-speed VLC is very challenging as the modulation bandwidth of white LEDs is inherently small. The 3-dB modulation bandwidth of a commercially available white LED is typically a few MHz [6]. In order to increase the data rate of bandwidth-limited VLC systems, many techniques have been reported so far, including pre- or post-frequency domain equalization [7], [8], multiple-input multiple-output (MIMO) transmission schemes [9], [10], spectral-efficient modulation formats such as orthogonal frequency division multiplexing (OFDM) utilizing high-order quadrature amplitude modulation (QAM) constellations [11]–[13].

The channel response of an OFDM based VLC system using intensity modulation/direct detection (IM/DD) largely depends on the low-pass nature of the LED transmitter [7], [14]. In order to correctly recover the transmitted data, the response of the channel should be accurately estimated and hence compensated at the receiver side. Conventional time domain averaging (TDA) technique employing multiple training symbols (TSs) has been widely adopted for channel estimation in OFDM-VLC systems [11], [15]. Recently, many new channel estimation techniques have also been reported in the literature. In [16], a post-processing channel estimation technique was proposed which could eliminate the noise outside the maximum channel delay. In [17], the authors introduced an adaptive channel estimation technique which is robust to the changes in channel distribution and SNR range. In [18], three different channel estimation techniques were investigated and compared, including intra-symbol frequency-domain averaging (ISFA), minimum mean squared error (MMSE), and weighted inter-frame averaging (WIFA). Nevertheless, multiple TSs are generally required for accurate estimation of the channel response, which inevitably reduces the overall achievable data rate of bandwidth-limited OFDM-VLC systems due to the large training overhead.

In this paper, for the first time, we propose a novel channel estimation technique based on sparse Bayesian dual-variate relevance vector machine (RVM) regression for bandwidth-limited OFDM-VLC systems. As a machine learning method, RVM could perform accurate predictions in a probabilistic manner with only a limited number of training symbols (TSs) [19]–[21]. By exploiting dual-variate RVM regression, the real (Re) and imaginary (Im) parts of the complex channel response can be accurately estimated. Compared with conventional TDA based channel estimation, the required training overhead can be significantly reduced by utilizing the proposed sparse Bayesian dual-variate RVM regression based channel estimation.

The rest of the paper is organized as follows. Section II describes the model of an indoor IM/DD based OFDM-VLC system using sparse Bayesian RVM regression based channel estimation. The simulation setup and results are presented in Section III. Finally, Section IV gives the conclusion.

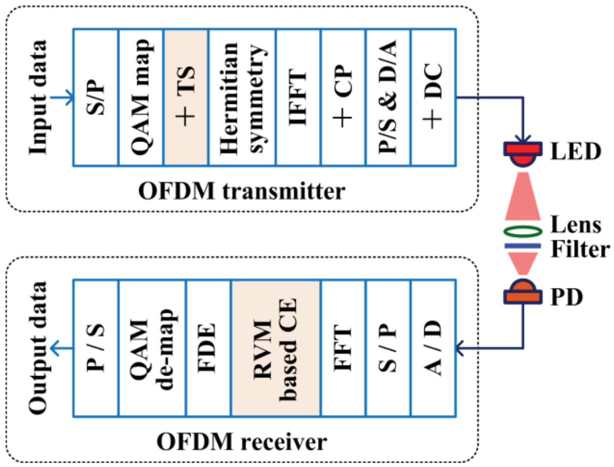

Fig. 1. Block diagram of an IM/DD OFDM-VLC system using sparse Bayesian RVM regression based channel estimation.

Notation: $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{-1}$ denote the transpose and the inverse operators, respectively. diag$(\cdot)$ stands for a diagonal matrix whose diagonal elements are the values given inside the round brackets. Non-boldface italic letters, lowercase boldface letters and capital boldface letters represent scalars, vectors and matrixes, respectively.

II. IM/DD OFDM-VLC Using Sparse Bayesian RVM Regression Based Channel Estimation

The model of an indoor IM/DD based OFDM-VLC system is described and the principle of sparse Bayesian RVM regression is also introduced. Moreover, the proposed channel estimation technique using dual-variate RVM regression is discussed.

A. System model

Fig. 1 shows the block diagram of an IM/DD OFDM-VLC system, where the serial input data are first converted to parallel data via serial-to-parallel (S/P) conversion. The parallel data are then mapped to quadrature amplitude modulation (QAM) symbols and training symbols (TSs) are also added. To obtain an LED-compatible real-valued signal, Hermitian symmetry is imposed before performing the inverse fast Fourier transform (IFFT). After cyclic prefix (CP) insertion, the resultant parallel and digital signal is converted to a serial and analog signal via parallel-to-serial (P/S) conversion and digital-to-analog (D/A) conversion. A DC bias current is further added so as to obtain a unipolar signal for modulating the intensity of an LED. The LED light modulated with data emits into the free space for simultaneous illumination and communication in the indoor environment. Usually, there are two types of light components received by the photodiode (PD): one is the line-of-sight (LOS) component, and the other is the diffuse components due to the reflections from the surfaces in the room. It has been shown that the strongest diffuse component is at least 7-dB lower than the weakest LOS component in electrical power [9]. Therefore, it is reasonable to only consider the LOS component in the system model. The LOS irradiance of an LED can be modeled as a generalized Lambertian pattern [12] and thus the LOS optical channel gain is calculated by (1).

$$h = \frac{(m+1)A}{2\pi d^2}\cos^m(\varphi) G_f G_l \cos(\phi)\,. \tag{1}$$

In (1), $m = -\ln 2/\ln(\cos\Phi_{1/2})$ is the Lambertian emission order and $\Phi_{1/2}$ is the semi-angle at half power of the LED, A is the active area of the PD, d is the distance between LED and PD, φ is the angle of irradiance, ϕ is the angle of incidence, G_f is the gain of the optical filter, and G_l is the gain of the optical lens. It should be noted that h becomes zero when the incident angle ϕ is outside the field-of-view (FOV) of the optical lens. After transmission over the free-space channel, the light is detected by a PD and the DC term is removed. The resultant electrical signal is given by

$$y(t) = RP_0 h\xi x(t) + n(t), \tag{2}$$

where R is the responsivity of the PD, P_0 is the output optical power of the LED without modulation, h is the channel gain, ξ is the modulation index, $x(t)$ is the transmitted OFDM signal with normalized electrical power, and $n(t)$ is the additive white Gaussian noise (AWGN) consisting of both shot and thermal noises. In the OFDM receiver, the received electrical signal is first converted to a digital and parallel signal via analog-to-digital (A/D) conversion and S/P conversion. After FFT, the proposed sparse Bayesian RVM regression enabled channel estimation is then executed and frequency domain equalization (FDE) is further performed by utilizing the estimated channel response. Hence, the output data can be obtained through QAM de-mapping and P/S conversion.

As per (2), the signal-to-noise ratio (SNR) of the received OFDM signal is expressed by

$$SNR = \frac{(RP_0 h\xi)^2}{\sigma^2_{shot} + \sigma^2_{thermal}}. \tag{3}$$

In (3), σ^2_{shot} and $\sigma^2_{thermal}$ are the variances of the shot and thermal noises, respectively, which are defined as [12]

$$\begin{cases} \sigma^2_{shot} = 2q(RP_r + I_{bg}I_2)B_n \\ \sigma^2_{thermal} = 8\pi k T_K \eta A B^2 \left(\dfrac{I_2}{G} + \dfrac{2\pi\Gamma}{g_m}\eta A I_3 B_n\right) \end{cases}, \tag{4}$$

where $P_r = hP_0$ is the average received optical power, I_{bg} is the background current, B_n is the equivalent noise bandwidth. The other parameters in (4) can be found in [1], [12].

B. Sparse Bayesian RVM regression

Based on the probabilistic sparse Bayesian RVM regression model [19], for a given training set of real-valued input-target pairs $\{\mathbf{x}_n, z_n\}_{n=1}^{N}$, the target samples $\{z_n\}_{n=1}^{N}$ can be predicted by using the following linear regression model:

$$z_n = \sum_{i=1}^{M} w_i \phi_i(\mathbf{x}_n) + \epsilon_n = \mathbf{w}^{\mathrm{T}}\boldsymbol{\phi}(\mathbf{x}_n) + \epsilon_n. \tag{5}$$

In (5), $\mathbf{w} = (w_1\ w_2\ \cdots\ w_M)^{\mathrm{T}}$ is the weight vector of length M, $\phi_i(\mathbf{x}_n)$ $(i = 1, 2, \cdots, M)$ is the basis function generated by input $\mathbf{x}_n$, $\boldsymbol{\phi}(\mathbf{x}_n) = (\phi_1(\mathbf{x}_n)\ \phi_2(\mathbf{x}_n)\ \cdots\ \phi_M(\mathbf{x}_n)\ )^{\mathrm{T}}$ is the corresponding basis vector, and $\boldsymbol{\epsilon} = (\epsilon_1\ \epsilon_2\ \cdots\ \epsilon_M)^{\mathrm{T}}$ is the additive error vector. The error sample ϵ_n $(n = 1, 2, \cdots, N)$ is assumed to be independently distributed Gaussian with zero mean and a variance of σ^2, i.e. $p(\boldsymbol{\epsilon}) = \prod_{n=1}^{N} N(\epsilon_n|0,\sigma^2)$. Therefore, the likelihood function of the target vector $\mathbf{z}$ is expressed by

$$p(\mathbf{z}|\mathbf{w},\sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{\|\mathbf{z}\text{-}\boldsymbol{\Phi}(\mathbf{x})\mathbf{w}\|^2}{2\sigma^2}\right\}, \quad (6)$$

where $\boldsymbol{\Phi}(\mathbf{x}) = [\boldsymbol{\phi}(\mathbf{x}_1)\ \boldsymbol{\phi}(\mathbf{x}_2)\ \cdots\ \boldsymbol{\phi}(\mathbf{x}_N)]^{\mathrm{T}}$ is the $N \times (N+1)$ design matrix and $\boldsymbol{\phi}(\mathbf{x}_n)$ is the biased basis vector which is given by

$$\boldsymbol{\phi}(\mathbf{x}_n) = [1\ K(\mathbf{x}_n,\mathbf{x}_1)\ K(\mathbf{x}_n,\mathbf{x}_2)\ \cdots\ K(\mathbf{x}_n,\mathbf{x}_N)]^{\mathrm{T}}. \quad (7)$$

Therefore, the length of $\boldsymbol{\phi}(\mathbf{x}_n)$ is $M = N+1$ and $K(\mathbf{x},\mathbf{x}_n)$ is the kernel function. One of the widely used kernels is the Gaussian kernel which is defined as

$$K(\mathbf{x}_m,\mathbf{x}_n) = \exp\left(-\lambda^{-2}\|\mathbf{x}_m - \mathbf{x}_n\|^2\right), \quad (8)$$

where λ is known as the width parameter.

From the Bayesian perspective, we constrain the parameters by defining a zero-mean Gaussian prior distribution over them which takes the form

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^{N} N(w_i|0,\alpha_i^{-1}), \quad (9)$$

where $\boldsymbol{\alpha} = (\alpha_0\ \alpha_1\ \alpha_2\ \cdots\ \alpha_N)^{\mathrm{T}}$ is a vector of $N+1$ independent hyperparameters and each one is used to individually control the strength of the prior over its associated parameter [19]. Based on the likelihood and the prior, the posterior distribution can be obtained via the Bayes Rule

$$p(\mathbf{w}|\mathbf{z},\boldsymbol{\alpha},\sigma^2) = \frac{p(\mathbf{z}|\mathbf{w},\sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})}{p(\mathbf{z}|\boldsymbol{\alpha},\sigma^2)}, \quad (10)$$

which is Gaussian distributed $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where the mean and the covariance matrix are given by,

$$\boldsymbol{\Sigma} = \left(\sigma^{-2}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} + \mathrm{diag}(\boldsymbol{\alpha})\right)^{-1}, \quad (11)$$

$$\boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{z}. \quad (12)$$

Since it is generally intractable to obtain the full posterior of those parameters, a type-II maximum likelihood procedure can be used to find a most-probable point estimate $\boldsymbol{\alpha}_{\mathrm{MP}}$. Therefore, sparse Bayesian learning is then formulated as the local maximization with respect to $\boldsymbol{\alpha}$ of the marginal likelihood [19] and the logarithm of the marginal likelihood is given by (13).

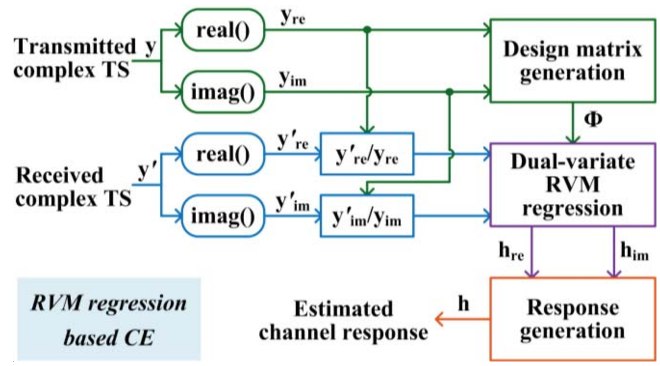

Fig. 2. Schematic diagram of the proposed sparse Bayesian dual-variate RVM regression based channel estimation.

$$\log(p(\mathbf{z}|\boldsymbol{\alpha},\sigma^2)) = \log\int_{-\infty}^{\infty} p(\mathbf{z}|\mathbf{w},\sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})\,d\mathbf{w},$$
$$= -(N\log 2\pi + \log|\mathbf{C}| + \mathbf{z}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{z})/2. \quad (13)$$

In (13), $\mathbf{C} = \sigma^2\mathbf{I} + \boldsymbol{\Phi}(\mathrm{diag}(\boldsymbol{\alpha}))^{-1}\boldsymbol{\Phi}^{\mathrm{T}}$ and $\mathbf{I}$ is an identity matrix. Hence, the estimate of $\boldsymbol{\mu}_{\mathrm{MP}}$ for the weights is given in (12) with $\boldsymbol{\alpha} = \boldsymbol{\alpha}_{\mathrm{MP}}$ and thus the final approximator of target $\mathbf{z}$ is given by $\tilde{\mathbf{z}} = \boldsymbol{\Phi}(\mathbf{x})\boldsymbol{\mu}_{\mathrm{MP}}$. It has been shown that the optimal values for most of the hyperparameters are infinite and hence $\boldsymbol{\mu}_{\mathrm{MP}}$ will consist of a few non-zero elements [19].

The original RVM regression model as proposed in [19] is initialized with all the M basis functions. The update rules for the hyperparameters depend on computing the posterior weight co-variance matrix as given in (11), which requires a matrix inverse operation with a complexity of $O(M^3)$. Therefore, the computational complexity of RVM regression could be rather high for practical applications. To reduce the computational complexity, a fast marginal likelihood maximization method was further proposed, where RVM regression is initialized with a single basis function, i.e. the bias [22]. Sequentially, the basis functions are iteratively added, updated or deleted to increase the marginal likelihood in a greedy manner. In this way, the computationally expensive matrix inversion operation can be appropriately avoided. Therefore, the new RVM regression model can achieve almost the same performance as the original one but with significantly reduced computational complexity. The implementation of such a fast RVM regression model using MATLAB is available in [23]. In this work, the fast RVM regression model is adopted and the computational complexity is analyzed in detail in Section III.

C. Channel estimation using dual-variate RVM regression

Since the sparse Bayesian RVM regression introduced above is a real-valued model, it cannot be directly applied to estimate the complex-valued channel of an OFDM-VLC system. In this subsection, we propose a sparse Bayesian dual-variate RVM regression based channel estimation technique for OFDM-VLC systems. Fig. 2 shows the schematic diagram of the proposed channel estimation technique. In order to estimate the channel response, one OFDM symbol consisting of totally N complex coefficients corresponding to N subcarriers is used for training.

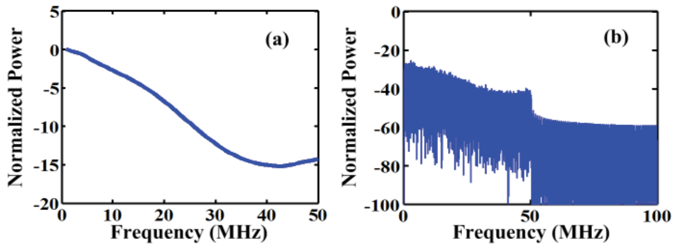

Fig. 3. (a) Channel response with a modulation bandwidth of 50 MHz used in the simulation (b) electrical spectrum of the received OFDM signal.

The transmitted and received TSs are given by $\mathbf{y}=\mathbf{y}_{re}+j\mathbf{y}_{im}$ and $\mathbf{y}'=\mathbf{y}'_{re}+j\mathbf{y}'_{im}$, respectively. It is assumed that the real (Re) and imaginary (Im) parts of the transmitted TS $\mathbf{y}$ are the same, i.e. $\mathbf{y}_{re}=\mathbf{y}_{im}$. As per (8), the design matrix $\mathbf{\Phi}$ is generated by using $\mathbf{y}_{re}$ and $\mathbf{y}_{im}$. After separating the Re part and the Im part of the received TS $\mathbf{y}'$, the responses of the Re and Im parts of the complex channel are obtained by $\mathbf{h}_{re}=\mathbf{y}'_{re}/\mathbf{y}_{re}$ and $\mathbf{h}_{im}=\mathbf{y}'_{im}/\mathbf{y}_{im}$, respectively. Taking $\mathbf{h}_{re}$ and $\mathbf{h}_{im}$ as two inputs and utilizing the design matrix $\mathbf{\Phi}$, sparse Bayesian dual-variate RVM regression can be performed and the estimated responses of the Re and Im parts of the complex channel, i.e. $\mathbf{f}_{re}$ and $\mathbf{f}_{im}$, can be obtained.

Based on the transmitted and received TSs, i.e. $\mathbf{y}$ and $\mathbf{y}'$, the complex channel response is expressed by

$$\mathbf{h}=\frac{\mathbf{y}'}{\mathbf{y}}=\frac{\mathbf{y}'_{re}+j\mathbf{y}'_{im}}{\mathbf{y}_{re}+j\mathbf{y}_{im}}. \tag{14}$$

Using $\mathbf{y}_{re}=\mathbf{y}_{im}$, (14) can be rewritten as

$$\mathbf{h}=\frac{\mathbf{y}'_{re}+j\mathbf{y}'_{im}}{\mathbf{y}_{re}(1+j)}=\frac{\mathbf{y}'_{re}+\mathbf{y}'_{im}-j(\mathbf{y}'_{re}-\mathbf{y}'_{im})}{2\mathbf{y}_{re}}. \tag{15}$$

Therefore, we have

$$\mathbf{h}=\frac{1}{2}[\mathbf{h}_{re}+\mathbf{h}_{im}-j(\mathbf{h}_{re}-\mathbf{h}_{im})]. \tag{16}$$

Since $\mathbf{f}_{re}$ and $\mathbf{f}_{im}$ are respectively the estimates of $\mathbf{h}_{re}$ and $\mathbf{h}_{im}$, the estimated complex channel response $\mathbf{f}$ is obtained by

$$\mathbf{f}=\frac{1}{2}[\mathbf{f}_{re}+\mathbf{f}_{im}-j(\mathbf{f}_{re}-\mathbf{f}_{im})]. \tag{17}$$

Based on the above estimated complex channel response, FDE can be successfully performed.

III. Simulation Setup and Results

In this section, Monte Carlo simulations are performed to evaluate the performance of an IM/DD OFDM-VLC system using sparse Bayesian dual-variate RVM regression based channel estimation. In the simulation, a commercially available white LED (Luxeon Star) is considered. Fig. 3(a) shows the measured channel response of the LED after blue filtering with 50 MHz modulation bandwidth. As we can observe, the power attenuation is up to 15 dB. The OFDM signal has an IFFT/FFT size of 256 and a CP length of 8. The 1st subcarrier is corresponding to the DC term and totally 64 subcarriers (2nd ~ 65th, N=64) are used to carry data. The 66th ~ 128th subcarriers are left unmodulated for oversampling. The sampling rate is set at 200 MSa/s and 16QAM mapping is used. Hence, the overall bit rate of the OFDM-VLC system is $50\times\log_2 16$ = 200 Mb/s. Fig. 3(b) illustrates the electrical spectrum of the received OFDM signal and totally 500 OFDM symbols are transmitted and collected for BER calculation.

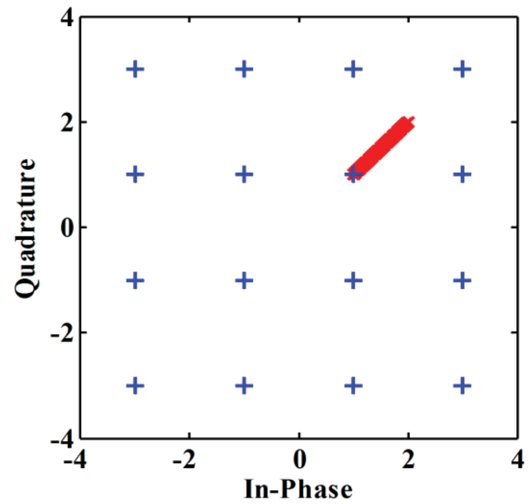

Fig. 4. Constellation diagram of the transmitted OFDM signal. Blue '+' signs show the constellation points of the 16QAM encoded data while red '×' signs give the constellation points of the complex training data of the TS.

As discussed in Section II.C, only one complex TS is needed when using the proposed sparse Bayesian dual-variate RVM regression based channel estimation technique. Moreover, the Re and Im parts of the TS should be the same. Fig .4 depicts the constellation diagram of the transmitted OFDM signal. It can be found that the constellation points of the 16QAM encoded data are shown by the blue '+' signs while the red '×' signs give the constellation points of the complex training data in the TS. The Re and Im parts of the complex training data are the same, which are assumed to be uniformly distributed between 1 and 2. Using the OFDM signal described in Fig. 4, the design matrix is generated by the Gaussian kernel given by (8) and the width parameter is set to 1. Since the basis functions are iteratively added, updated or deleted to increase the marginal likelihood, we assume that the maximum number of iterations is 10. After obtaining the responses of the Re and Im parts of the complex channel by exploiting the TS, dual-variate RVM regression is performed to achieve the amplitude responses of the Re and Im parts of the complex channel. Figs. 5(a) and (b) illustrate the estimated amplitude responses of the Re and Im parts of the complex channel, respectively, with an SNR of 20 dB. As we can see, the amplitude responses are severely distorted before regression because of the additive noise while the amplitude responses become much smoother after regression. Moreover, there are only three relevance vectors (RVs) used in the RVM regression for both the Re and Im parts, indicating that only three out of totally 65 weights ($M=N+1=65$) are nonzero.

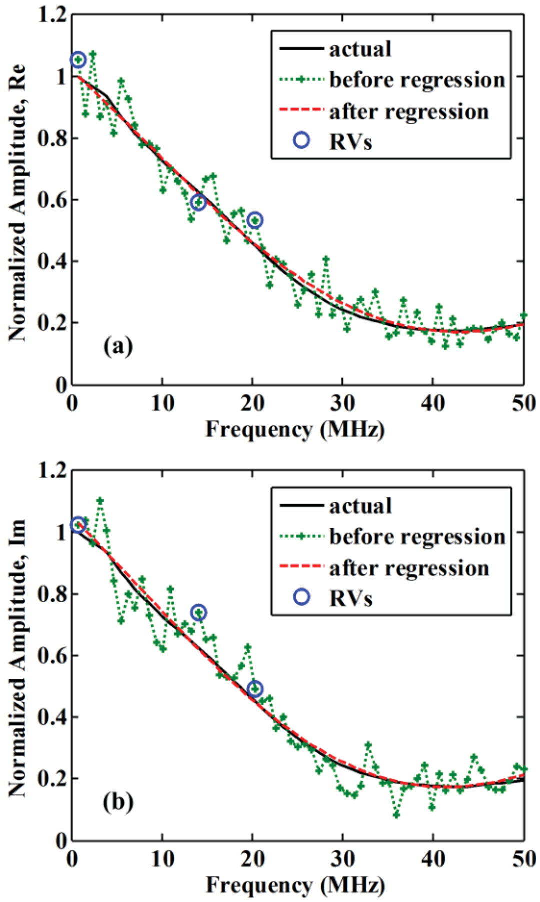

Fig. 5. Amplitude response of (a) real part and (b) imaginary part of the channel with an SNR of 20 dB. RVs: relevance vectors.

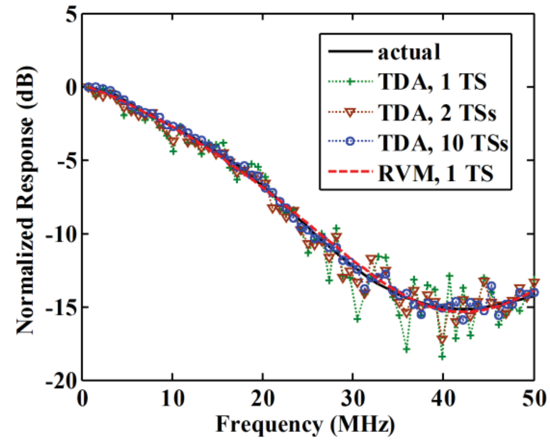

Fig. 6. Comparison of actual and estimated channel responses (SNR=20 dB).

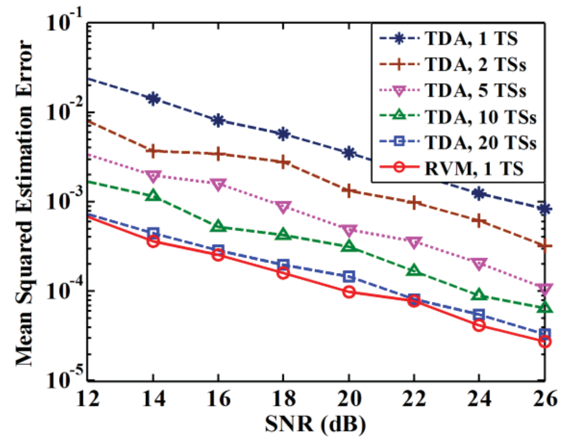

Fig. 7. Mean square estimation error versus SNR for 16QAM based OFDM-VLC system with different channel estimation techniques.

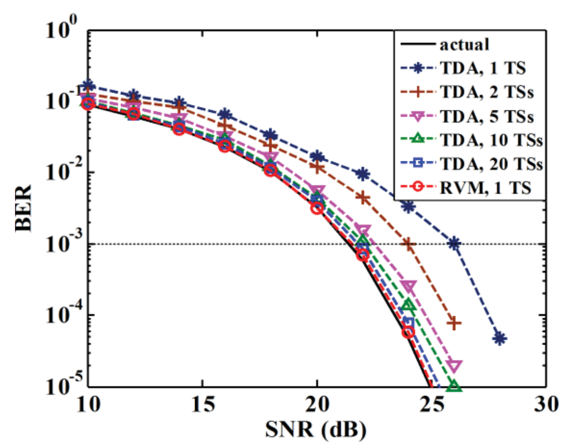

Fig. 8. BER versus SNR for 16QAM based OFDM-VLC system with different channel estimation techniques.

The estimated channel responses using conventional TDA based channel estimation and the proposed sparse Bayesian dual-variate RVM regression based channel estimation are compared in Fig. 6. It can be clearly observed that the estimated channel response is least accurate when TDA with one TS is used. With the increase of the number of the TSs, the estimated channel response employing TDA becomes more accurate. In contrast, the channel response can be accurately estimated by using sparse Bayesian dual-variate RVM regression technique with only one TS. Fig. 7 shows the mean square estimation error versus SNR for the 16QAM based OFDM-VLC system. It can be seen that the mean square estimation error is gradually reduced with the increase of SNR. Moreover, the mean square estimation error using conventional TDA is decreased when more TSs are utilized. It is also revealed that nearly the same mean square estimation error performance can be achieved for conventional TDA based channel estimation with 20 TSs and sparse Bayesian dual-variate RVM regression based channel estimation with only one TS.

The BER performance of the 16QAM based OFDM-VLC system is shown in Fig. 8. To achieve a target BER of 10^{-3}, the required SNR using the actual channel information is about 21.4 dB. When conventional TDA based channel estimation with one TS is used, the required SNR for achieving BER=10^{-3} is about 26 dB, resulting in an SNR penalty of 4.6 dB due to the inaccurate channel information. When more TSs are used for TDA based channel estimation, the SNR penalty is reduced. For example, when five TSs are used, the SNR penalty can be reduced to 1.1 dB. Moreover, when total 20 TSs are used, the SNR penalty becomes negligible (~ 0.4 dB). In contrast, the SNR penalty is only about 0.2 dB when using sparse Bayesian dual-variate RVM regression based channel estimation with only one TS. As a result, the required number of TSs is reduced from 20 to 1 for accurate channel estimation when employing the proposed channel estimation technique in comparison to conventional TDA, indicating a substantially reduced training overhead in bandwidth-limited OFDM-VLC systems.

Finally, we analyze the computational complexity of the proposed sparse Bayesian dual-variate RVM regression based channel estimation technique. For the original RVM regression model as proposed in [19], the computational complexity is related to both the iteration number N_{it} and the number of basis functions M. An inverse operation with a complexity of $O(M^3)$ is required to compute the posterior weight co-variance matrix so as to update the hyperparameters at each iteration. Therefore, the computational complexity of the original RVM regression is $O(N_{it} \times M^3)$ with a memory storage of $O(M^2)$. However, for the fast RVM regression which is adopted in this work [22], a fast marginal likelihood maximization method is used, where RVM regression is initialized with the bias only and the basis functions are iteratively added, updated or deleted to increase the marginal likelihood. In the worst case, a new basis function is added at each iteration since adding basis functions requires most of the computations. Hence, the worst-case computational complexity of the fast RVM regression is $O(N_{it} \times M^2)$ [22]. It should also be noted that the worst case scenarios are highly impossible to occur. For common cases, an approximation of the computational complexity is about $O(N_{it} \times N_{RV}^2)$ where N_{RV} is the number of RVs. In our numerical simulations, only three out of totally 65 basis vectors are used as RVs, suggesting a substantially reduced computational complexity.

IV. CONCLUSION

We have proposed a channel estimation technique based on sparse Bayesian dual-variate RVM regression in a 200 Mb/s OFDM-VLC system. The performance of the proposed channel estimation technique has been evaluated and compared with the widely used TDA technique by numerical simulations. The simulation results have shown that, to accurately estimate the channel response, totally 20 complex TSs are required when using conventional TDA based channel estimation, while only one complex TS is needed when the proposed sparse Bayesian RVM regression based channel estimation is employed. Hence, a significantly training overhead reduction can be achieved, resulting in an improved spectral efficiency in bandwidth-limited VLC systems. It has also been demonstrated that, by adopting the fast marginal likelihood maximization method, the sparse Bayesian RVM regression based channel estimation can be computational efficient for the practical application in high-speed OFDM-VLC systems.

ACKNOWLEDGMENT

This work was conducted within the Delta-NTU Corporate Lab for Cyber-Physical Systems with funding support from Delta Electronics Inc. and the National Research Foundation (NRF) Singapore under the Corp Lab@University Scheme.

REFERENCES

[1] Z. Ghaseemlooy, W. Popoola, and S. Rajbhandri, *Optical Wireless Communications, System and Channel Modelling With Matlab*. London, U.K.: CRC Press, 2012.

[2] H. Haas, "Visible light communication," in *Proc. OFC*, 2015, p. Tu2G.5.

[3] S. Arnon, Ed., *Visible Light Communication*. Cambridge, U.K.: Cambridge Univ. Press, Feb. 2015.

[4] D. Wu, Z. Ghassemlooy, W.-D. Zhong, M.-A. Khalighi, H. L. Minh, C. Chen, S. Zvanovec, and A. Boucouvalas, "Effect of optimal Lambertian order for cellular indoor optical wireless communication and positioning systems," *Opt. Eng.*, vol. 55, no. 6, Jun. 2016, Art. no. 066114.

[5] P. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: a survey, potential and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047–2077, 2015.

[6] S. Rajbhandari, et al., "High-speed integrated visible light communication system: device constraints and design considerations," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1750–1757, 2015.

[7] H. L. Minh, D. O'Brien, G. Faulkner, L. Zeng, K. Lee, D. Jung, Y. Oh, and E. Won, "100-Mb/s NRZ visible light communications using a postequalized white LED," *IEEE Photon. Technol. Lett.*, vol. 21, no. 15, pp. 1063–1065, Aug. 2009.

[8] C. Chen, W.-D. Zhong, and D. Wu, "Indoor OFDM visible light communications employing adaptive digital pre-frequency domain equalization", in *Proc. CLEO*, 2016, p. JTh2A.118.

[9] L. Zeng, D. O'Brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. Won, "High data rate multiple input multiple output (MIMO) optical wireless communications using white led lighting," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 9, pp. 1654–1662, 2009.

[10] C. Chen, W.-D. Zhong, D. Wu, and Z. Ghassemlooy, "Wide-FOV and high-gain imaging angle diversity receiver for indoor SDM-VLC systems," *IEEE Photon. Technol. Lett.*, vol. 28, no. 19, pp. 2078–2081, 2016.

[11] S. K. Hashemi, Z. Ghassemlooy, L. Chao, and D. Benhaddou, "Orthogonal frequency division multiplexing for indoor optical wireless communications using visible light LEDs," in *Proc. CNSDSP*, 2008, pp. 174–178.

[12] C. Chen, W.-D. Zhong, and D. Wu, "Communication coverage improvement of indoor SDM-VLC system using NHS-OFDM with a modified imaging receiver", in *Proc. IEEE ICC Workshops*, 2016, pp. 315–320.

[13] C. Chen, W.-D. Zhong, and D. Wu, "Non-Hermitian symmetry orthogonal frequency division multiplexing for multiple-input multiple-output visible light communications," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 36–44, Jan. 2017.

[14] D. Bykhovsky and S. Arnon, "Multiple access resource allocation in visible light communication systems," *J. Lightw. Technol.*, vol. 32, no. 8, pp. 1594–1600, Apr. 2014.

[15] H. Elgala, R. Mesleh, H. Haas, and B. Pricope, "OFDM visible light wireless communication based on white LEDs," in *Proc. VTC*, 2007, pp. 2185–2189.

[16] Y. Xiang, M. Zhang, X. Tang, J. Wu, and D. Han, "A post-processing channel estimation method for DCO-OFDM visible light communication," in *Proc. CNSDSP*, 2012, pp. 1–4.

[17] T. Zhang, S. Guo, H. Chen, F. Zhong, and C. Ma, "Enhancing the bit error rate of indoor visible light communication systems using adaptive channel estimation algorithm," *IET Commun.*, vol. 9, no. 4, pp. 501–507, Mar. 2015.

[18] B. Lin, X. Tang, Z. Ghassemlooy, S. Zhang, Y. Li, Y. Wu, and H. Li, "Efficient frequency-domain channel equalisation methods for OFDM visible light communications," *IET Commun.*, vol. 11, no. 1, pp. 25–29, Jan. 2017.

[19] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Jun. 2001.

[20] L. Zhao, L. Wang, G. Bi, L. Zhang, and H. Zhang, "Robust frequency-hopping spectrum estimation based on sparse Bayesian method," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 781–793, Feb. 2015.

[21] L. Zhao, L. Wang, L. Yang, Z. Abdelhak, and G. Bi, "The race to improve radar imagery: An overview of recent progress in statistical sparsity-based techniques," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 85–102, Nov. 2016.

[22] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proc. of 9th Int. Workshop on Artificial Intelligence and Statistics*, C. M. Bishop and B. J. Frey, Eds., 2003.

[23] Available: http://www.miketipping.com/sparsebayes.htm#software.