

Signal Demodulation Using a Radial Basis Function Neural Network (RBFNN) in a Silicon Photomultiplier-Based Visible Light Communication System

Cuiwei He , *Member, IEEE*, and Steve Collins , *Member, IEEE*

Abstract—A silicon photomultiplier (SiPM) contains an array of microcells that can each detect individual photons. Consequently, it can arguably result in the most sensitive receiver in visible light communication (VLC). However, each microcell needs a period of several nanoseconds to recover after detecting a photon. This creates a non-linear response and introduces a unique form of inter-symbol interference. In this paper, we first show that this interference splits each element of the received signal constellation into multiple clusters. This observation motivates the investigation into the use of a Radial Basis Function Neural Network (RBFNN) to deal with the impact of the nonlinearity. Both the training procedures and the performance of the RBFNN are explained and discussed in detail. The influence of the number of the RBFNN centers, the widths of the centers, the constellation size and the period of the transmitted signal samples on the system performance are investigated. In addition, two different RBFNN-based data demodulation methods are introduced. The simulation results suggest that the new RBFNN-aided receivers reduce the negative impacts of the SiPM nonlinearity and can result in lower bit error rates (BERs) for a wide range of irradiances on the SiPM.

Index Terms—SiPM, SPAD, photon-counting, nonlinearity, machine learning, radial basis function neural network.

I. INTRODUCTION

DUE to the exponentially growing number of Internet of Things (IoT) devices, radio frequency (RF) wireless communications, including WiFi, is increasingly limited by the available bandwidth. One of the most efficient solutions to this problem is to add wireless communication capacity by using a different part of the electromagnetic spectrum, for example visible light. The result is a trending technology known as visible light communications (VLC) [1], [2]. Most VLC systems use intensity modulation/direct detection (IM/DD) in which the transmitted data is modulated onto the intensity of the light

emitted by the transmitter and detected using photodetectors in the receiver [3]. In VLC, the performance of a link then depends on the sensitivity of the optical receiver [4]. The most sensitive possible receiver is the one that can accurately count the number of photons incident on the receiver within a short period of time. One way in which a photon-counting receiver can be created is by biasing an avalanche photodiode (APD) above its breakdown voltage and placing it in series with a quenching device. This type of optical sensor is known as a single-photon avalanche diode (SPAD). Although a SPAD is very sensitive it needs a short period to recover after a single photon has been detected. During this period the SPAD can't detect any other photons [5], [6] and so this period is usually known as either the dead time or the recovery period of the SPAD. Since a minimum number of photons per bit are required to achieve a target bit error rate (BER) when photons are counted [5], the SPAD recovery period limits the data rates that can be supported [7].

Fortunately, the impact the SPAD recovery time can be reduced by using an array of SPADs so that when some SPADs are inactive other active SPADs can detect photons from the transmitter. One type of SPAD array, known as silicon photomultiplier (SiPM), is now commercially available from companies including Hamamatsu and onsemi. In a SiPM, a single SPAD is referred to as a microcell and all microcells share a common output. Recently, the use of SiPMs in VLC has been demonstrated with Gbits/s transmission data rates from several research groups [7]–[11]. These promising results have been obtained despite the fact that the recovery time of each microcell means that SiPMs have a non-linear response that can create a unique form of signal distortion when the sampling period of the transmitted signal is close to or less than the SiPM recovery period [9].

To compensate for the SiPM nonlinearity, a number of signal post-equalization and pre-equalization methods have been suggested. In [7], [12], when on-off keying (OOK) or pulse amplitude modulation (PAM) is used as the modulation method, decision feedback equalizers were used to simultaneously mitigate the signal distortion caused by the SiPM nonlinearity and any inter-symbol interference (ISI). In [13], the performance of a SiPM based orthogonal frequency-division multiplexing (OFDM) system was studied and the specific frequency response

Manuscript received 22 March 2022; revised 27 May 2022; accepted 16 June 2022. Date of publication 20 June 2022; date of current version 1 July 2022. (Corresponding author: Cuiwei He.)

Cuiwei He is with the School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1211, Japan (e-mail: cuiweihe@jaist.ac.jp).

Steve Collins is with the Department of Engineering Science, University of Oxford, Oxford OX1 2JD, U.K. (e-mail: steve.collins@eng.ox.ac.uk).

This article has supplementary downloadable material available at <https://doi.org/10.1109/JPHOT.2022.3184672>, provided by the authors.

Digital Object Identifier 10.1109/JPHOT.2022.3184672

due to the SiPM recovery period was analyzed. In addition, it was shown that, when the channel response at the receiver is pre-estimated, a group of single-tap equalizers can be implemented in the frequency domain to reduce the impact of the non-linearity. Furthermore, in [14], a time-domain based pre-equalization method which is specially designed to reduce the negative impact of the SiPM nonlinearity on OFDM signals was shown to give promising results. In a more recent study [11], this method was adapted to create a post-equalizer.

Although the signal distortion caused by the SiPM nonlinearity can be mitigated using classic digital signal processing methods, very accurate channel information and precise SiPM parameters need to be either pre-estimated or measured. This creates extra challenges and the complexity of the equalizers increases at higher data rates. Alternatively, data-driven solutions using machine learning algorithms can be considered. In recent years, the use of neural network based machine learning algorithms in digital communications has become a trending research topic [15] and its applications have also been considered in various types of VLC scenarios [16]–[19]. In more recent studies, artificial neural networks (ANN) or multilayer perceptrons (MLPs) have been shown to be a promising candidate for replacing the conventional demodulation techniques in SPAD-based optical wireless systems [20]. However, the main challenge is that the accuracy of the ANN-aided receiver depends on the complexity of the network and a complex ANN structure with many layers requires more training data and time. Moreover, in most cases, the trained ANN is used as a ‘black box’ and many of the trained parameters cannot be interpreted. Consequently, this approach does not provide any insights into how the performance of neural networks can be further improved or how they would perform in unusual circumstances.

In this paper, the SiPM non-linearity is shown to create a unique form of distortion when OFDM is used and the period of individual transmitted signal samples is less than the SiPM’s recovery time. In particular, the non-linearity splits each quadrature amplitude modulation (QAM) constellation point into several clusters. To emphasize this effect, the possible transmitted constellation points will be referred to as constellation elements in the rest of the paper. To deal with this new type of distortion, a new demodulation method is proposed and investigated. This new method is based upon a simple three-layer Radial Basis Function Neural Network (RBFNN). This RBFNN is formed from a group of Radial Basis Functions (RBFs) [21] and the parameters of the RBFs are obtained from training data. Each element of the QAM constellation is then associated with multiple RBFs. The network is trained to estimate the probability and each symbol has created the input to the RBFNN. This output probability, or soft information, can then be used to demodulate the input signal.

When a RBFNN is used in a SiPM based VLC system, its performance depends on a range of factors including the number of the RBFs per constellation element, the parameters of each RBF, the constellation size, the irradiance falling on the SiPM and the duration of individual transmitted signal samples. All these factors are analyzed in detail in this paper. Moreover, the performance of three different receivers are compared. These

receivers are a conventional receiver, a RBFNN receiver which is used to take hard decisions based upon the output with the maximum value and a RBFNN receiver that rejecting decisions when all the output probabilities are below a decision threshold. The BER results obtained using these methods show that RBFNN can reduce the impact of the SiPM nonlinearity over a wide range of irradiances falling on the SiPM.

The rest of the paper is structured as follows. Section II introduces the statistical model considered in the simulations of the photon counting process. The SiPM nonlinearity and its impact on received signals are described in Sections III. This is followed in Section IV by a description of RBFNNs and the details of the network training process in Section V. The performance of the trained RBFNN with different parameters is discussed in Section VI. Section VII describes a new RBFNN aided SiPM OFDM system and the BER results obtained with this system discussed in Section VIII. Finally, Section IX concludes the paper and discusses future work.

II. SIMULATION MODEL

In this section, the statistical model used to simulate the photon counting process is explained. In an IM/DD based VLC system, the transmitted information is carried on the instantaneous optical power of the light and therefore can be decoded based on the number of photons arrived at the optical sensor during individual signal sampling periods. The number of photons arrived at the SiPM can be accurately modelled using a Poisson distribution [7], [22]. When the influence of the photon detection efficiency (PDE), α_{PDE} , is considered, during the detection of the k th signal sample, the distribution of the effective number of photons arrived at a single SiPM microcell, v_k , has a probability mass function (PMF) given by

$$P(v_k = j) = \frac{\kappa_k^j e^{-\kappa_k}}{j!} \quad (1)$$

where κ_k is the average effective number of arriving photons, ! is the factorial function. If the SiPM has N_{cells} microcells and the energy of a single photon is E_p , κ_k is related to the instantaneous optical power received by the SiPM, $r_{\text{opt}}(t)$, by

$$\kappa_k = \frac{\alpha_{\text{PDE}}}{N_{\text{cells}} E_p} \int_{t_k}^{t_k + T_s} r_{\text{opt}}(t) dt \quad (2)$$

where $\int_{t_k}^{t_k + T_s} r_{\text{opt}}(t) dt$ is the received energy by the SiPM during the transmission of the k th sample and T_s is the transmission period of individual signal samples.

In this paper, to accurately model the system, we simulate the arriving time of individual photons. When the number of photons arriving at the SiPM within a time period is a Poisson variable, the time interval between two adjacent arriving photons follows an exponential distribution and its probability density function (PDF) is given by

$$p(t) = \lambda e^{-\lambda t} \quad (3)$$

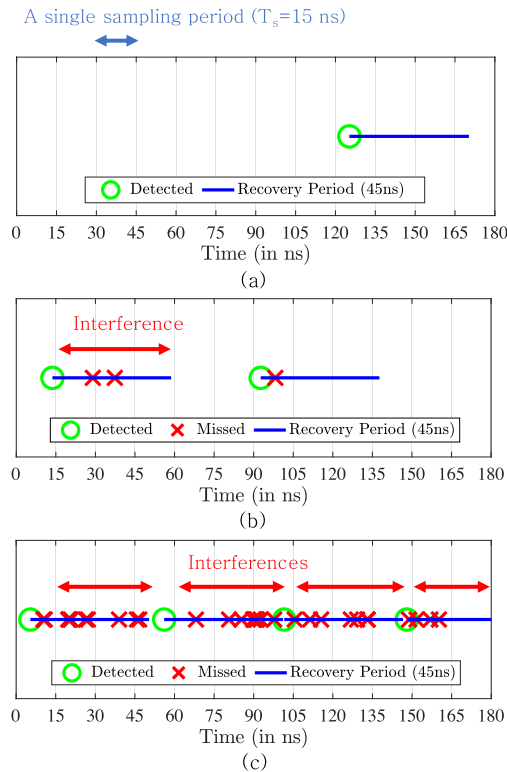


Fig. 1. The simulation trails of the photon counting process of an individual SiPM microcell when the transmission period of a signal sample, T_s , is 15 ns and different irradiance levels are considered (a) 1 mW/m^2 , (b) 10 mW/m^2 , (c) 100 mW/m^2 . In this figure, a green circle indicates the time of a detected photon which is followed by a blue line indicating its associated recovery period. The red crosses indicate the missed photons.

where λ is the rate parameter and it is related to κ_k and T_s by

$$\lambda = \frac{\kappa_k}{T_s} \quad (4)$$

As explained in the introduction, due to the recovery period of the microcell, not all arriving photons can be detected. Fig. 1 shows several simulation trails of the photon counting process of a single SiPM microcell under different irradiance levels. The details of the simulation method which includes the influences of the microcell recovery period are described in [13], [14]. From Fig. 1, we can see that the overall number of photons arriving at a SiPM microcell increases when the irradiance level is changed from 1 mW/m^2 to 100 mW/m^2 . At the same time, due to the microcell's recovery period, the number of missed photons also increases as the irradiance increases. Moreover, it can be seen that, when the sampling/counting period is 15 ns, which is much less than the microcell's recovery period of 45 ns, the recovery period spans multiple signal sampling periods which causes a unique form of interference [13], [14] and the associated signal distortion is discussed in the next section.

In this paper, we analyze the performance of the system by considering the received optical power or irradiance level rather than the transmitted optical power. In this case, the conclusions obtained in the following sections are independent of the optical channel gain as well as the properties of the optical transmitter. In a practical transmission system, a certain received irradiance

TABLE I
SiPM PARAMETERS [23]

| Parameters | Value |
|---------------------------------------|-------|
| SiPM Name | 30035 |
| Area (mm^2) | 9.42 |
| No. of Microcells, N_{cells} | 5676 |
| Microcell diameter (μm) | 35 |
| Fill Factor (%) | 75 |
| Recovery Period (ns) | 45 |
| PDE, α_{PDE} (405 nm) | 38% |

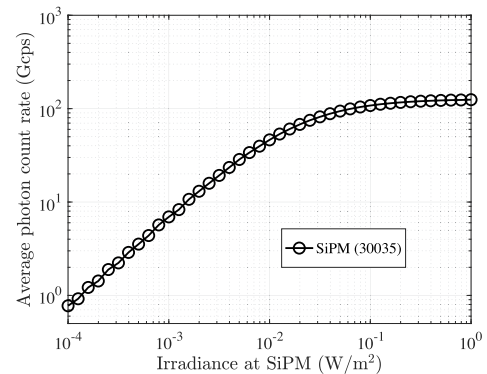


Fig. 2. The simulated average photon counting rate as a function of the irradiance of 405 nm light falling on a SiPM 30035.

level is related to a combination of the power of the transmitter and/or the distance between the transmitter and the receiver and/or the incident angle of the light.

III. SiPM NONLINEARITY AND SIGNAL DISTORTION

As shown in Section II, due to the microcell's recovery period, the number of detected photons is not always proportional to the irradiance falling on the SiPM. Moreover, when the period of the transmitted signal samples is less than the recovery period of the SiPM microcells, the microcell's recovery period introduces interference between signal samples which results in a unique form of signal distortion. To explain the motivation for using a RBFNN in a SiPM based VLC system, the SiPM nonlinearity and its associated signal distortion are discussed in this section. In this paper, a SiPM 30035 from onsemi is considered whose key parameters are listed in Table I. Fig. 2 shows the simulated average photon counting rate as a function of the irradiance falling on a SiPM. In particular, Fig. 2 shows that, when the irradiance of 405 nm light is up to 10 mW/m^2 , there is a linear relationship between irradiance and the photon counting rate. However, when the irradiance is above 10 mW/m^2 , the microcells recovery time means that some photons incident on the SiPM cannot be counted and this relationship becomes non-linear. Finally, the count rate saturates when the irradiance is above 100 mW/m^2 .

By simulating the SiPM OFDM system used in [13], Fig. 3 shows the received signal constellations when a transmitted signal sampling period of 15 ns is considered as an example

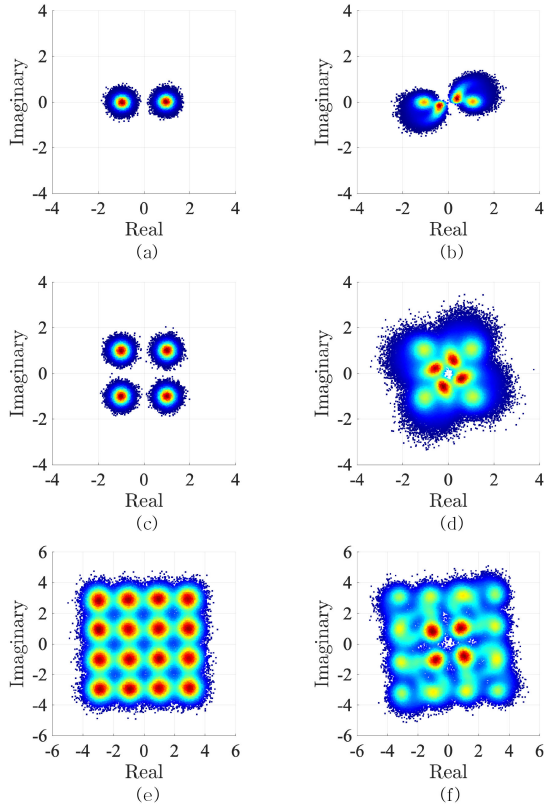


Fig. 3. Several examples of the received signal constellations from all OFDM subcarriers when the counting period, T_s , is 15 ns, (a) BPSK and the received irradiance level is 1 mW/m², (b) BPSK and the received irradiance level is 40 mW/m², (c) 4-QAM and the received irradiance level is 1 mW/m², (d) 4-QAM and the received irradiance level is 40 mW/m², (e) 16-QAM and the received irradiance level is 1 mW/m², (f) 16-QAM and the received irradiance level is 10 mW/m².

which is much less than the microcell recovery period of 45 ns. Each of the constellation diagrams in Fig. 3 is obtained from 1000 OFDM symbols, with each OFDM symbol contains 256 subcarriers. Three different constellation sizes including BPSK, 4-QAM, 16-QAM are considered. Fig. 3(a), (c), and (e) shows the constellations when the irradiance level is low enough for the SiPM to be linear. It can be seen that under these conditions noise will determine the BER. However, at a high irradiance the SiPM is non-linear and Fig. 3(b), (d), and (f) show that the constellations are very distorted. More importantly, these results show that the non-linearity can split the constellations points into multiple clusters. In most communication systems, demodulation of the received signal includes a procedure to identify the QAM constellation which corresponds to a particular input. Using a RBFNN is an efficient way of classifying data into different categories with each category associated with multiple data clusters [21], they therefore seem particularly well suited to demodulating QAM signal that have been distorted by the SiPM non-linearity.

IV. RBFNN FOR DATA DEMODULATION

In this section, the structure of the RBFNN that is investigated is described. As shown in Fig. 4, the RBFNN is formed from

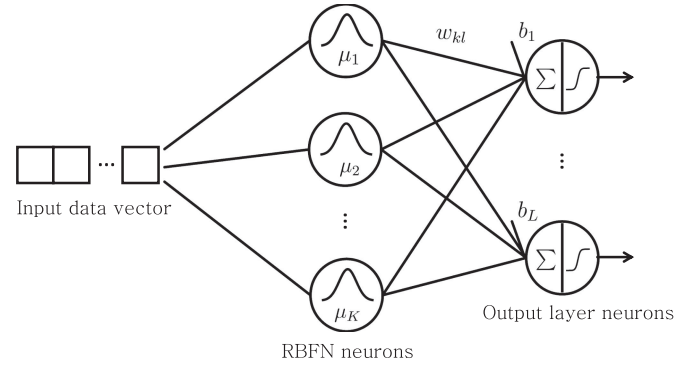


Fig. 4. A schematic diagram of a RBFNN used to demodulate the output of a SiPM VLC receiver.

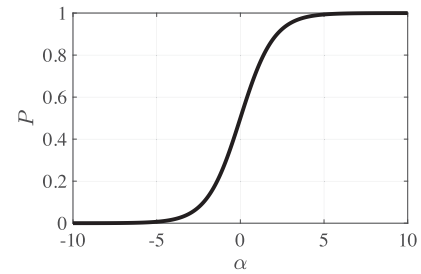


Fig. 5. A schematic diagram of the logistic function using in the output layer of the RBFNN.

three layers [21]. The input layer contains individual input data values and the hidden layer contains a group of RBFs which are referred to as RBF neurons in this paper. Finally, the output layer contains one neuron for each element in the QAM constellation and each of these output neurons contains a sum function followed by a logistic function. After the RBFNN has been trained, a received complex constellation value is first normalized and then supplied into all RBF neurons. In the rest of the paper, a received constellation value is denoted by $Y = Y_R + iY_I$ in which Y_R is the real part and Y_I is the imaginary part. In this paper, the RBFs are Gaussian and the output value of the k th RBF neuron is therefore calculated using

$$\varphi_k = e^{-\beta_k \|Y - \mu_k\|^2} \quad (5)$$

where μ_k is mean position or center of the k th Gaussian function or neuron. β_k then determines the variance of the Gaussian function and hence determines the width of the k th RBF neuron. Next, the output values from all RBF neurons are weighted and then summed using

$$\alpha_l = \sum_{k=1}^K w_{kl} \varphi_k + b_l \quad (6)$$

where w_{kl} is the weighting coefficient between the k th RBF neuron and the l th output neuron and b_l is the bias value for the l th output neuron. Each, α_l is then the input to a logistic function

$$P_l = \frac{1}{1 + \exp(-\alpha_l)}, \quad (7)$$

such as the one shown in Fig. 5. As shown in this figure the value of P_l is constrained to be between 0 and 1 and it is trained to estimate the probability that $Y = Y_R + iY_I$ belongs to the l th output neuron and hence the l th element of the constellation [24], [25]. The estimated probabilities from all output layer neurons can then be used to associate an element of the constellation with the input. The simplest way that this can be done is to make a hard decision by selecting the output with the maximum value using

$$c = \arg \max_l P_l. \quad (8)$$

Finally, c is converted back to binary data sequence using a look up table (LUT) which is based on Gray coding.

In some cases, $Y = Y_R + iY_I$ might be a rare outlier which is far from all the RBF centers. In this situation, a more accurate method is to set a threshold value. Then if the maximum value of P_l is below this threshold, the hard decision is rejected and the associated bit is retransmitted. Although this approach can reduce the transmission data rate, it will lead to lower BERs. The results obtained using both of these methods are therefore included in this paper.

V. THE TRAINING OF A RBFNN

The training process of the RBFNN is divided into two stages. The first stage is to determine the locations of the centers of the RBF neurons, $\boldsymbol{\mu} = [\mu_1, \dots, \mu_k, \dots, \mu_K]$ and their associated width parameters, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_k, \dots, \beta_K]$. The second stage is to obtain the optimal weighting coefficients, w_{kl} , and the bias coefficients, b_l , between the RBF neurons and the output neurons.

A. *K*-Means Clustering Algorithm to Obtain RBF Centers

In this paper, unsupervised k -means clustering [26] was used to determine the locations of the centers of the RBF neurons, $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_K]$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_k, \dots, \beta_K]$ based on a set of training input data and associated constellation points, $\mathbf{Y} = [Y_1, Y_2, \dots, Y_m, \dots, Y_M]$. In the k -means clustering algorithm, when the number of RBF neurons, K , is decided, the values of $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_K]$ are first initialized randomly. Then for each constellation point in $\mathbf{Y} = [Y_1, Y_2, \dots, Y_m, \dots, Y_M]$, its Euclidean distances to all RBF neuron centers, $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_K]$, are calculated and then this constellation point is assigned to the closest RBF center. After this assignment process is implemented for all elements of \mathbf{Y} , \mathbf{Y} is divided into different groups.¹ The next step is for the center of each RBF neuron, μ_k , to be updated to be the average value of the data points which are assigned to this neuron. The above steps are implemented iteratively to update all RBF centers until the positions of the centers no longer change. Then, the standard deviation of the k th data group is calculated using

$$\sigma_k = \sum_{Y_{D,k} \in \mathbf{Y}_{D,k}} \frac{\|Y_{D,k} - \mu_k\|}{|\mathbf{Y}_{D,k}|} \quad (9)$$

¹Note that, in the case, if no data point is assigned to a certain RBF center, this center is removed and not considered in the following updating steps. Therefore, the final number of RBF centers is equal to or less than K .

TABLE II
K-MEANS ALGORITHM TO FIND RBF CENTERS

| | |
|----------------|--|
| Input: | $\mathbf{Y} = [Y_1, Y_2, \dots, Y_m, \dots, Y_M], \gamma$ |
| Output: | $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_K], \boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_k, \dots, \beta_K]$ |

- 1: Randomly generate initial values of all elements in $\boldsymbol{\mu}$
- 2: Set the number of iterations, I
- 3: $i \leftarrow 1$
- 4: **while** $i < I$ **do**
- 5: **for** each $Y_m \in \mathbf{Y} = [Y_1, Y_2, \dots, Y_m, \dots, Y_M]$ **do**
- 6: $d_m = \arg \min_k \|Y_m - \mu_k\|$
- 7: **end for**
- 8: **for** $m = 1$ To M **do**
- 9: **for** $k = 1$ To K **do**
- 10: **if** $d_m = k$ **then**
- 11: $\mathbf{Y}_{D,k} = [Y_{D,k}, Y_m]$
- 12: **else if** $d_m \neq k$ **then**
- 13: $\mathbf{Y}_{D,k} = \mathbf{Y}_{D,k}$
- 14: **end if**
- 15: **end for**
- 16: **end for**
- 17: **for** $k = 1$ To K **do**
- 18: $\mu_k \leftarrow \sum_{Y_{D,k} \in \mathbf{Y}_{D,k}} \frac{Y_{D,k}}{|\mathbf{Y}_{D,k}|}$
- 19: **end for**
- 20: $i \leftarrow i + 1$
- 21: **end while**
- 22: **for** $k = 1$ To K **do**
- 23: $\sigma_k = \sum_{Y_{D,k} \in \mathbf{Y}_{D,k}} \frac{\|Y_{D,k} - \mu_k\|}{|\mathbf{Y}_{D,k}|}$
- 24: $\beta_k = \frac{\gamma}{2\sigma_k^2}$
- 25: **end for**

where $\mathbf{Y}_{D,k}$ is a data vector which contains all the constellation values assigned to the k th RBF center and $|\mathbf{Y}_{D,k}|$ denotes the number of elements within $\mathbf{Y}_{D,k}$. In this way, the constellation values assigned to different RBF centers are determined after the training and consequently the values of σ_k are fixed. However, the performance of RBFNN depends upon the values of σ_k which are not necessarily optimized. A parameter, γ , has therefore been used to make these parameters adjustable, in particular,

$$\beta_k = \frac{\gamma}{2\sigma_k^2}. \quad (10)$$

In this paper, the width of the k th RBF neuron is defined as

$$\sigma_{\text{width},k} = \sqrt{\frac{1}{2\beta_k}}. \quad (11)$$

In this case, decreasing the value of γ increases the width of the RBF neurons. Note that when γ is one, $\sigma_{\text{width},k} = \sigma_k$ is the original obtained standard deviation of the k th data group. The simplified pseudocode of the considered k -means algorithm is summarized in Table II.

B. Gradient Descent Algorithm to Obtain Weighting Coefficients

As shown in Fig. 4, each RBF neuron is connected to all neurons in the output layer. Using the training data, the weighting coefficients, w_{kl} , and the bias coefficients, b_l , are updated based on the gradient descent principle [27].

To obtain the optimal values of w_{kl} and b_l , an error function is first defined and then minimized. In the RBFNN, the indices of output neurons are associated with an element of a constellation. Consequently, the desired output probability of the neuron associated with the correct element of the constellation is 1. In contrast, the target output probabilities of all the other output neurons are 0 s. In this paper, an error function [25], [28] for the l th output neuron is defined as

$$e_l = \begin{cases} -\frac{1}{M} \sum_{m=1}^M \log(P_{l,m}), & \text{if } c_{l,m} = 1 \\ -\frac{1}{M} \sum_{m=1}^M \log(1 - P_{l,m}), & \text{if } c_{l,m} = 0 \end{cases} \quad (12)$$

where $P_{l,m}$ is obtained value of the l th output neuron based on the m th training data point using (5)–(7). $c_{l,m} \in \{0, 1\}$ is the desired output and it is obtained based on the transmitted binary data. Using (12), when the desired value of the l th output neuron is one, ($c_{l,m} = 1$), and the predicted probability is one, (e.g. $P_{l,m} = 1$), the error coefficient, e_l , is zero. Similarly, when the desired output value is zero, (e.g. $c_{l,m} = 0$) and the predicted probability is zero, (e.g. $P_{l,m} = 0$), the error coefficient, e_l , is also zero. In contrast, when the difference between $P_{l,m}$ and $c_{l,m}$ is large, the value of e_l becomes high. To simplify the further analysis, (12) is rewritten as

$$\begin{aligned} e_l &= \frac{1}{M} \sum_{m=1}^M -(c_{l,m} \log P_{l,m} + (1 - c_{l,m}) \log(1 - P_{l,m})) \\ &= \frac{1}{M} \sum_{m=1}^M e_{l,m}. \end{aligned} \quad (13)$$

Next, to obtain the values of w_{kl} and b_l which can minimize e_l efficiently a gradient descent based coefficient updating approach is employed. The gradient of w_{kl} and b_l are determined based on partial derivatives. The partial derivative of e_l with respect to w_{kl} is calculated using

$$\frac{\partial e_l}{\partial w_{kl}} = \frac{1}{M} \sum_{m=1}^M \frac{\partial e_{l,m}}{\partial P_{l,m}} \frac{\partial P_{l,m}}{\partial \alpha_{l,m}} \frac{\partial \alpha_{l,m}}{\partial w_{kl}}, \quad (14)$$

where

$$\begin{aligned} \frac{\partial e_{l,m}}{\partial P_{l,m}} &= -\frac{\partial (c_{l,m} \log P_{l,m} + (1 - c_{l,m}) \log(1 - P_{l,m}))}{\partial P_{l,m}} \\ &= -\left(\frac{c_{l,m}}{P_{l,m}} - \frac{1 - c_{l,m}}{1 - P_{l,m}} \right), \end{aligned} \quad (15)$$

$$\frac{\partial P_{l,m}}{\partial \alpha_{l,m}} = P_{l,m}(1 - P_{l,m}), \quad (16)$$

TABLE III
THE PSEUDO-CODE OF THE UPDATING PROCESS OF b , w_l AND w_R

Input: $\mathbf{Y} = [Y_1, \dots, Y_m, \dots, Y_M]$, $c_l = [c_{l,1}, \dots, c_{l,m}, \dots, c_{l,M}]$ ($c \in \{0, 1\}$), $\boldsymbol{\mu} = [\mu_1, \dots, \mu_k, \dots, \mu_K]$, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_k, \dots, \beta_K]$

Output: w_{kl} , b_l

- 1: **for** $k = 1$ To K **do**
- 2: $\boldsymbol{\varphi}_k = [e^{-\beta_k \|Y_1 - \mu_k\|^2}, \dots, e^{-\beta_k \|Y_m - \mu_k\|^2}, \dots, e^{-\beta_k \|Y_M - \mu_k\|^2}] = [\varphi_{k,1}, \dots, \varphi_{k,m}, \dots, \varphi_{k,M}]$
- 3: **end for**
- 4: Randomly generate initial w_{kl}
- 5: Randomly generate initial b_l
- 6: Set η , $i = 1$
- 7: **while** $i < I_{\text{iteration}}$ **do**
- 8: **for** $l = 1$ To L **do**
- 9: $P_l = \frac{1}{1 + \exp(-(\sum_{k=1}^K w_{kl} \varphi_k + b_l))}$
- 10: $w_{kl} \leftarrow w_{kl} - \frac{\eta}{M} \sum_{m=1}^M (P_{l,m} - c_{l,m}) \varphi_{k,m}$
- 11: $b_l \leftarrow b_l - \frac{\eta}{K} \sum_{m=1}^M (P_{l,m} - c_{l,m})$
- 12: **end for**
- 13: $i \leftarrow i + 1$
- 14: **end while**

and

$$\frac{\partial \alpha_m}{\partial w_{kl}} = \varphi_{k,m}. \quad (17)$$

Substituting (15)–(17) into (14) gives

$$\frac{\partial e_l}{\partial w_{kl}} = \frac{1}{M} \sum_{m=1}^M (P_{l,m} - c_{l,m}) \varphi_{k,m}. \quad (18)$$

Using the same approach the gradient of b_l is

$$\begin{aligned} \frac{\partial e_l}{\partial b_l} &= \frac{1}{M} \sum_{m=1}^M \frac{\partial e_{l,m}}{\partial P_{l,m}} \frac{\partial P_{l,m}}{\partial \alpha_{l,m}} \frac{\partial \alpha_{l,m}}{\partial b_l} \\ &= \frac{1}{M} \sum_{m=1}^M (P_{l,m} - c_{l,m}). \end{aligned} \quad (19)$$

Equations (18) and (19) indicate when the values of w_{kl} and b_l should be increased or decreased for minimizing the defined error, e_l . Furthermore, the optimal coefficients of w_{kl} and b_l are obtained using an iterative method. In each iteration, w_{kl} and b_l are updated using

$$w_{kl} \leftarrow w_{kl} - \eta \frac{\partial e_l}{\partial w_{kl}} = w_{kl} - \frac{\eta}{M} \sum_{m=1}^M (P_{l,m} - c_{l,m}) \varphi_{k,m}, \quad (20)$$

and

$$b_l \leftarrow b_l - \eta \frac{\partial e_l}{\partial b_l} = b_l - \frac{\eta}{K} \sum_{m=1}^M (P_{l,m} - c_{l,m}). \quad (21)$$

where η is the learning rate which is fixed at 0.5 in the following analysis. The overall procedures of this gradient descent algorithm are summarized in the Table III.

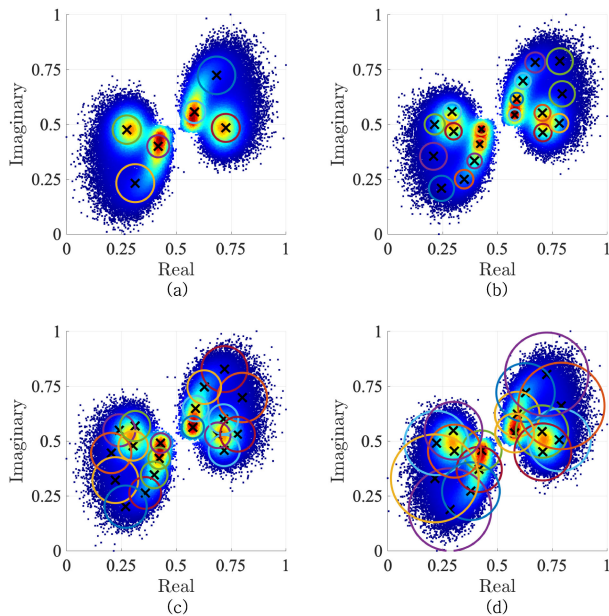


Fig. 6. The normalized BPSK signal constellations from all OFDM subcarriers used for the network training and the positions of the RBF neurons when the irradiance level is 40 mW/m^2 , (a) $K/L = 3$ and $\gamma = 1$, (b) $K/L = 10$ and $\gamma = 1$, (c) $K/L = 10$ and $\gamma = 0.5$, (d) $K/L = 10$ and $\gamma = 0.1$. The crosses in this figure are the centers of the RBF neurons, μ_k . The radii of the circles are the widths of the RBF neurons, $\sigma_{\text{width},k}$.

VI. RBFNN PERFORMANCE

The performance of the RBFNN depends on a range of parameters including the size of the constellation, the number of RBF neurons and the variance associated with each RBF neuron. In this section, the influences of these parameters are investigated. At the start of this investigation it was anticipated that larger RBFNNs will be required to deal with larger constellations and so the parameter, K/L , has been used where K is the number of RBF neurons and L is the size of the constellation. The data used for the network training is obtained from the statistical model described in Section II.

A. RBFNN Outputs for BPSK

In this section, we first focus on the results when BPSK is used with OFDM. Fig. 6 shows the results when the k -mean clustering algorithm described in Section V-A is used to locate the centers and widths of the RBF neurons. In the training process, 1000 OFDM symbols with each symbol containing 256 subcarriers were used. Fig. 6(a) shows the case when $K/L = 3$ and $\gamma = 1$. First, it can be seen that in the areas with high training data densities the widths of the RBF neurons are relatively small compared to the RBF neurons in the areas with low data density levels. Also, a comparison of Fig. 6(a) and (b) shows that, when the number of RBF neuron is increased the widths of the neurons decreases. Fig. 6(c) and (d) then show the impact of increasing the widths of the RBF neurons by using two smaller values of γ .

Next, the performance of the RBFNN containing the RBF neurons shown in Fig. 6(a) is investigated in detail. Firstly, the weighting coefficients as well as the bias coefficients were

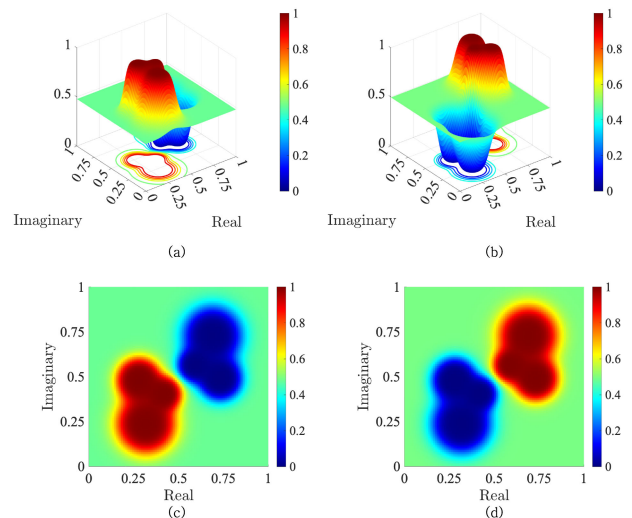


Fig. 7. (a) the 1st output values of the trained RBFNN (side view), (b) the 2nd output values of the trained RBFNN (side view), (c) the 1st output values of the trained RBFNN (top view), (d) the 2nd output values of the trained RBFNN (top view) when the irradiance level is 40 mW/m^2 , $K/L = 3$, $\gamma = 1$.

obtained using the gradient descent algorithm described in Section V-B. After the network was trained, all possible inputs within a normalized complex space were presented to the network. The results obtained for the two output neurons for all these possible inputs are shown in Fig. 7. Fig. 7(a) and (c) show that the outputs of the first neuron are very high where the input data is associated with a transmitted 0. In contrast, its output is very low for the positions associated input data arising from a transmitted 1. In contrast, the output of the second neuron, as shown in Fig. 7(b) and (d), is close to one when the input data is similar to training data arising from a transmitted 1 while they are low when the inputs are similar to the training data arising from a transmitted 0. Consequently, this RBFNN can be used to distinguish between the two BPSK symbols and classify them into two categories. In this case, most of the transmitted data can be decoded successfully. However, the results also show that when an input falls into a region from which training data was absent both output values are close to 0.5. This means that if an outlier input falls onto these locations, it can cause a detection error. However, this can be avoided by only making a classification decision if the maximum output value is larger than a threshold. The consequences of adopting this approach are described in Section IV.

Then, we analyze how the outputs are affected by changing the widths as well as the numbers of the RBF neurons. Fig. 8 shows the output values of the first neuron when the four cases of RBF neurons shown in Fig. 6 are considered. First, when a larger number of RBF neurons are considered, we can see that the distribution of the output values have a more complex pattern in Fig. 8(b) compared to Fig. 8(a) and consequently Fig. 8(b) can better represent the distribution of the constellations. Then, when the widths of the RBF neurons are increased by reducing the value of γ to 0.5 in Fig. 8(c) and 0.1 in Fig. 8(d), the area in which the two output neurons are similar, that is the area in

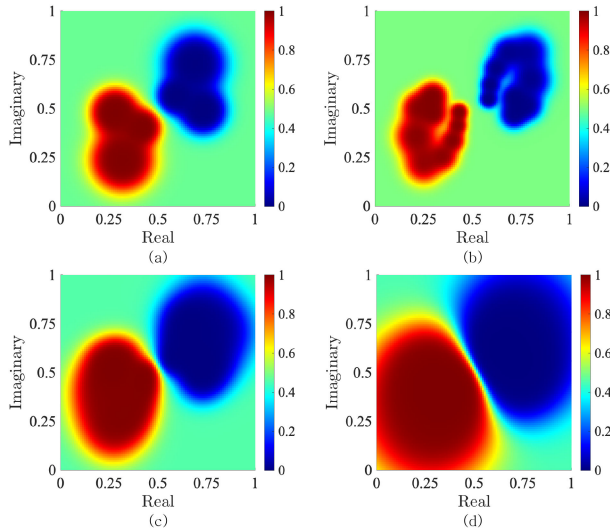


Fig. 8. The 1st output of the trained RBFNN for all possible constellation positions when the irradiance level is 40 mW/m^2 (a) $K/L = 3$ and $\gamma = 1$, (b) $K/L = 10$ and $\gamma = 1$, (c) $K/L = 10$ and $\gamma = 0.5$, (d) $K/L = 10$ and $\gamma = 0.1$.

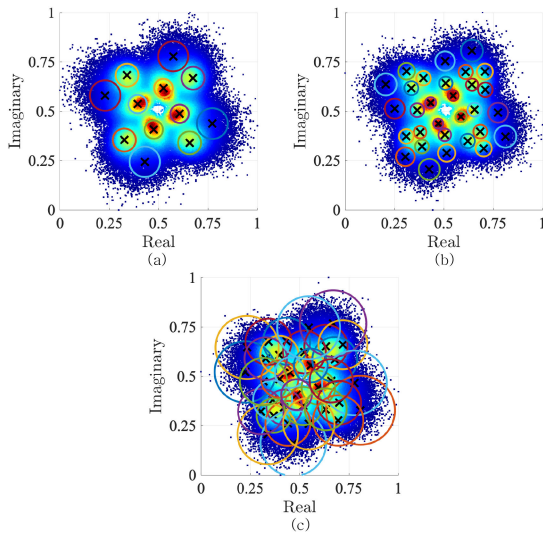


Fig. 9. The 4-QAM signal constellations from all OFDM subcarriers used for the RBFNN training when the irradiance level is 40 mW/m^2 , (a) $K/L = 3$, $\gamma = 1$, (b) $K/L = 10$, $\gamma = 1$, (c) $K/L = 10$, $\gamma = 0.1$. The crosses in this figure are the centers of the RBF neurons, μ_k . The radii of the circles are the widths of the RBF neurons, $\sigma_{\text{width},k}$.

which the outputs are indistinguishable, is reduced. In the cases of Fig. 8(c) and (d), the associated BERs are zeros.

B. RBFNN Outputs for 4-QAM and 16-QAM

In this section, the RBFNN outputs are discussed when larger constellation sizes, e.g. 4-QAM and 16-QAM, are used with OFDM. Fig. 9 shows the RBF neurons obtained with different values of K/L and γ when the 4-QAM constellations in Fig. 3(d) are used. The results in Fig. 9 show that using k -means clustering means that the regions in which there is a high density of input data are always covered by the RBF neurons. Similar to the cases of BPSK, for a given value of γ , the width of the RBF

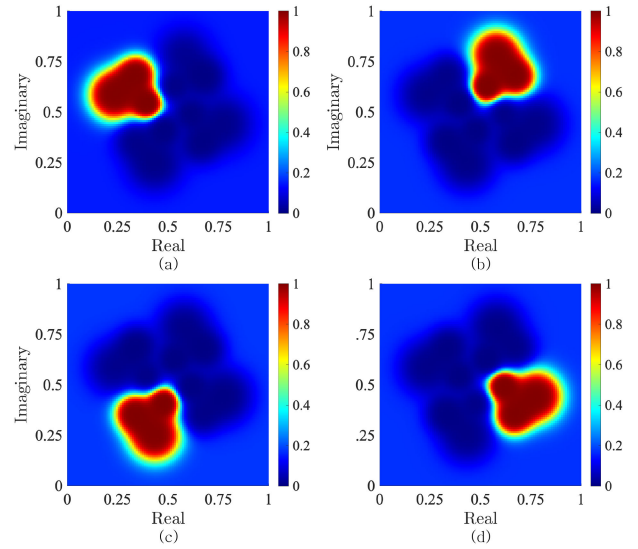


Fig. 10. The four outputs of the trained RBFNN for all possible input values in a normalized complex constellation space for 4-QAM when the irradiance level is 40 mW/m^2 , $K/L = 3$ and $\gamma = 1$.

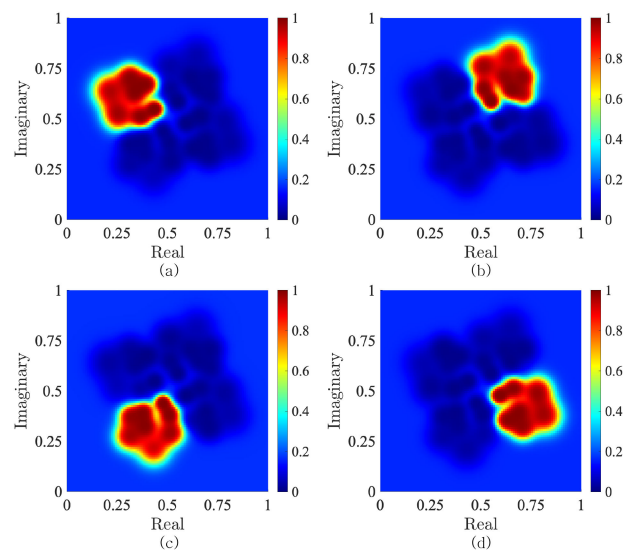


Fig. 11. The four outputs of the trained RBFNN for all possible input values in a normalized complex constellation space for 4-QAM when the irradiance level is 40 mW/m^2 , $K/L = 10$ and $\gamma = 1$.

neurons reduces when the number of RBFs increases. However, the width of the RBF neurons can be enlarged by using a smaller value of γ .

Fig. 10 shows the four output values of the RBFNN when the RBF neurons in Fig. 9(a) are used. The results in Fig. 10(a), (b), (c), and (d) show that output of each neuron is high when the input is similar to an input in the training data associated with its corresponding QAM element. At the same time, the values of the other three outputs are close to zero. Next, the number of RBF neurons is increased by using $K/L = 10$. The resulting RBF neurons are shown in Fig. 9(b) and the associated outputs of the network are shown in Fig. 11. In this case, a more complex distribution pattern of the output values is obtained. Also, similar

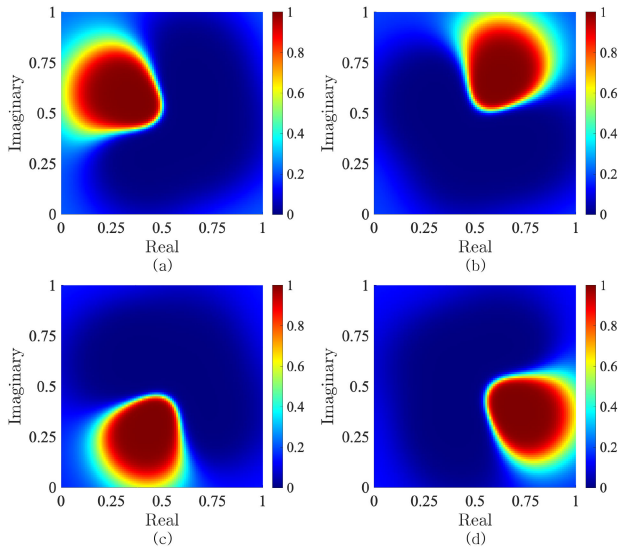


Fig. 12. The four outputs of the trained RBFNN for all possible input values in a normalized complex constellation space for 4-QAM when the irradiance level is 40 mW/m^2 , $K/L = 10$ and $\gamma = 0.1$.

to the case of BPSK, Fig. 12 shows that the undistinguishable area is reduced by using $\gamma = 0.1$ to increase the width of the RBF neurons. If 16-QAM is used then the RBFNN needs 16 outputs. In this case, Fig. 13 shows the outputs when $K/L = 10$ and $\gamma = 0.1$. Importantly, these result show that the overall space is divided into 16 areas and each area has a complex non-linear boundary which can potentially lead to more accurate results compared to the conventional approach, which relies upon partitioning the input space using multiple lines.

C. Results for Low Irradiances

Although the RBFNN has been suggested to handle the impact of SiPM non-linearity when the irradiance level is high, it must also performance well when the SiPMs response is linear. In the above section, the results show how a trained RBFNN can divide the constellation space into multiple areas based on the training data when the SiPM is non-linear. In this section, the results are also studied when the SiPM is linear and noise is the main cause of signal distortion. Figs. 14 and 15 show the outputs of the RBFNN when the received irradiance level is 1 mW/m^2 and hence the SiPM is in the linear region of Fig. 2. The horizontal and/or vertical lines in Figs. 14 and 15 indicate the decision boundaries when decisions/classifications are made based upon maximum likelihood (ML). Figs. 14 and 15 show that the trained RBFNN divides the complex constellation space into two areas for BPSK and four areas for 4-QAM. More importantly, the borders of these areas match the ML decision boundaries very well. Consequently, when the SiPM's response is linear, a receiver that uses a RBFNN is expected to achieve a very similar performance to one that uses ML.

VII. RBFNN AIDED SiPM OFDM SYSTEM

In this section, the RBFNN aided VLC transmission system is introduced. As shown in Fig. 16, at the transmitter,

a signal vector, $\mathbf{X} = [X_0, X_1, \dots, X_{N-1}]$, which contains N bipolar complex QAM data is input into an IFFT block. In VLC, since IM/DD is used, the transmitted signal needs to be both real and unipolar. To generate a real time-domain signal, $\mathbf{x} = [x_0, x_1, \dots, x_{N-1}]$, \mathbf{X} is constrained to have Hermitian symmetry. Next, a cyclic prefix (CP) and a DC bias are added and the negative part of the signal is clipped at a zero level so that the signal, $s_{\text{DCO}}(n)$, becomes unipolar. The optimal choice of the DC bias is discussed in [14] and in this paper the DC bias is fixed at 7 dB. Next, $s_{\text{DCO}}(n)$ is sent into a digital to analog converter (DAC) to obtain $s_{\text{DCO}}(t)$ which is used as the input to a 405 nm transmitter, a wavelength that was chosen because it is associated with a high SiPM photon detection efficiency. Finally, the emitted optical signal passes through an optical channel before arriving at the SiPM receiver. In the simulated transmission, the period of each transmitted signal sample is T_s and consequently the duration of one OFDM symbol is $(N + N_{\text{CP}})T_s$ and N_{CP} is the length of the CP. In the analysis, N is fixed at 256 and N_{CP} is fixed at 32. The data rate of DCO-OFDM can be calculated using

$$b = \frac{\left(\frac{N}{2} - 1\right) \log_2(M)}{(N + N_{\text{CP}})T_s} \quad (22)$$

where $\left(\frac{N}{2} - 1\right)$ is the number of subcarriers used for data-carrying in DCO-OFDM.

At the receiver, the light intensity is detected using a SiPM. The signal pulses generated from all SiPM microcells are added together via a common output. This output signal, $y(t)$, is input into an analog to digital converter (ADC) which has a sampling rate of $1/T_s$. The captured discrete signal sequence at the ADC output is then converted to the number of detected photons and used to create a vector of the number of photons detected during a period of NT_s , $\mathbf{y} = [y_0, y_1, \dots, y_{N-1}]$, which is sent into an FFT block to give $\mathbf{Y} = [Y_0, Y_1, \dots, Y_{N-1}]$. During the next step, each of the received signals is sent into a trained RBFNN to be classified. Finally, a look up table (LUT) converts the category ID into the received bits.

VIII. BER RESULTS

A. Influences of Irradiance, Sampling Period and Constellation Size

The BER results of the transmission system described in Section VII are analyzed in this section. Figs. 17 and 18 show the simulated BER results as a function of the irradiance falling on the SiPM when 4-QAM and 16-QAM are used. Although BPSK is good to be considered as an example to explain the principles of RBFNN, its associated BERs are usually very low and out of our interested range. Therefore, the BER results of BPSK are not discussed in this section. In the following discussion, for each of the three transmission sampling periods, three different methods have been used to determine the data that has been transmitted. First, it can be seen that, for all cases, the BER first decreases and then increases when the irradiance level is changed from 0.1 mW/m^2 to 100 mW/m^2 . This is because, when the irradiance level is low, the SiPM is linear and the performance is dominated

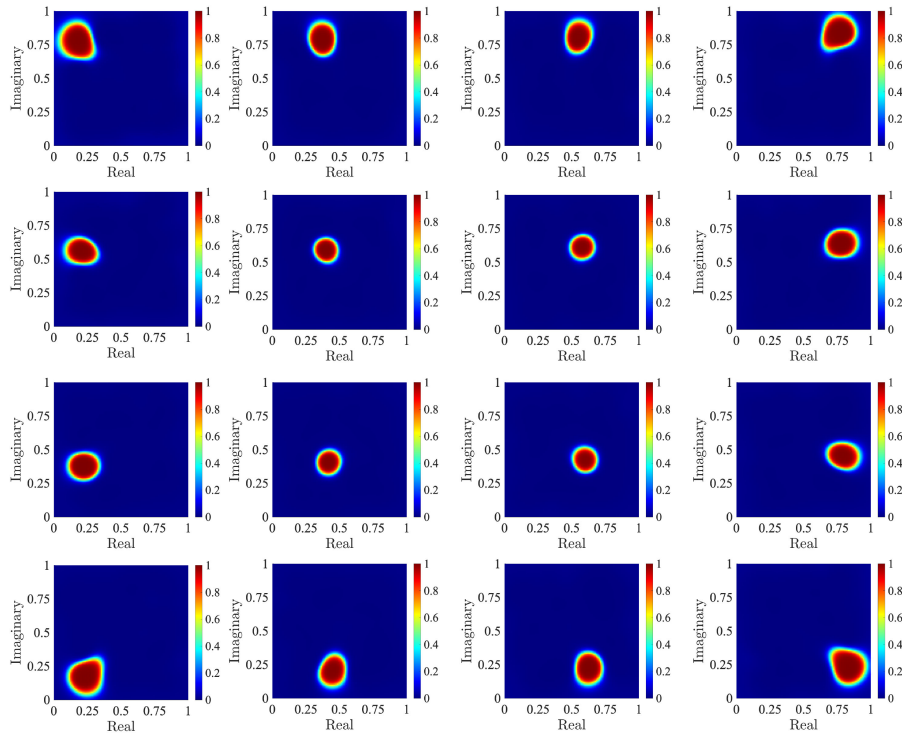


Fig. 13. The sixteen outputs of the trained RBFNN for all possible values in a normalized complex constellation space for 16-QAM when the irradiance level is 10 mW/m^2 , $K/L = 10$ and $\gamma = 0.1$.

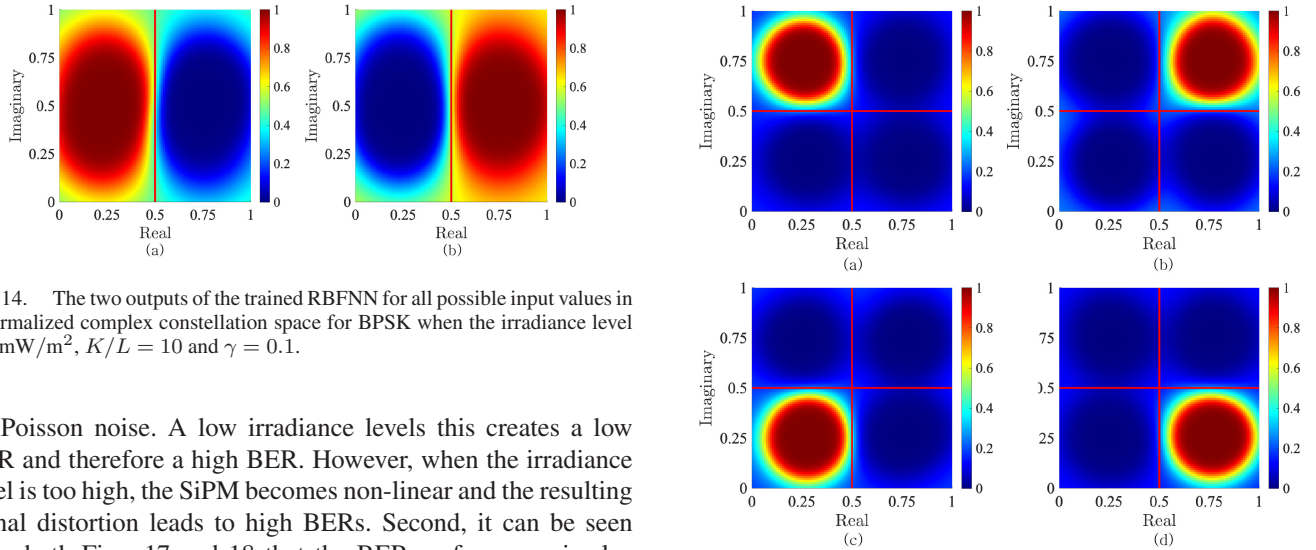


Fig. 14. The two outputs of the trained RBFNN for all possible input values in a normalized complex constellation space for BPSK when the irradiance level is 1 mW/m^2 , $K/L = 10$ and $\gamma = 0.1$.

by Poisson noise. A low irradiance levels this creates a low SNR and therefore a high BER. However, when the irradiance level is too high, the SiPM becomes non-linear and the resulting signal distortion leads to high BERs. Second, it can be seen from both Figs. 17 and 18 that the BER performance is also related to the sampling period of the transmitted signals and a shorter sampling period causes higher BERs. More importantly, as predicted in Section VI-C, in all cases the RBFNN receiver gives the same performance as the conventional receiver at low irradiances. More importantly, the results show that when the irradiance level is high, the receiver that incorporates a RBFNN can achieve much lower BERs than the conventional receiver without a RBFNN. This demonstrates that the RBFNN based receiver can significantly reduce the impact of the SiPM nonlinearity. Moreover, when decision rejection is used with a RBFNN receiver, the results show that the BERs are reduced even further in all cases. In these results, the decision rejection

Fig. 15. The four outputs of the trained RBFNN for all possible input values in a normalized complex constellation space for 4-QAM when the irradiance level is 1 mW/m^2 , $K/L = 10$ and $\gamma = 0.1$.

threshold is fixed at 0.7. The influences of the choice of the decision rejection threshold are discussed in Section VIII-D.

B. Low Error Irradiance Range (LEIR)

The results in Figs. 17 and 18 show that the BERs are only low for a range of irradiance levels. In this paper, in line with most of the VLC research work, the acceptable BER is considered to

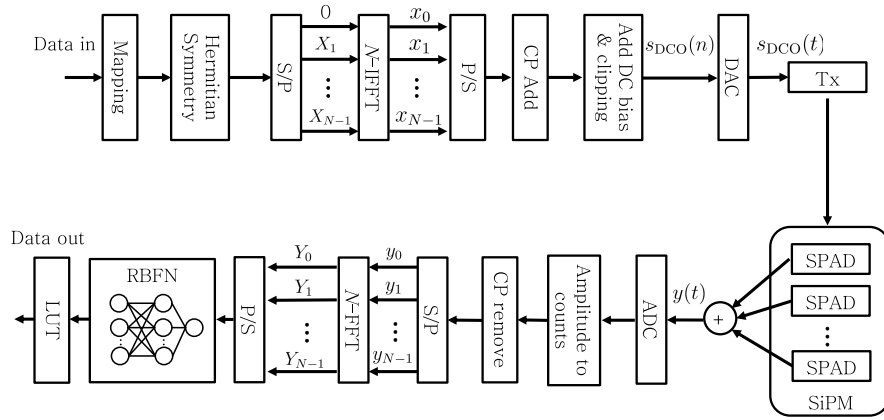


Fig. 16. The considered SiPM OFDM system with RBFNN-aided signal detection.

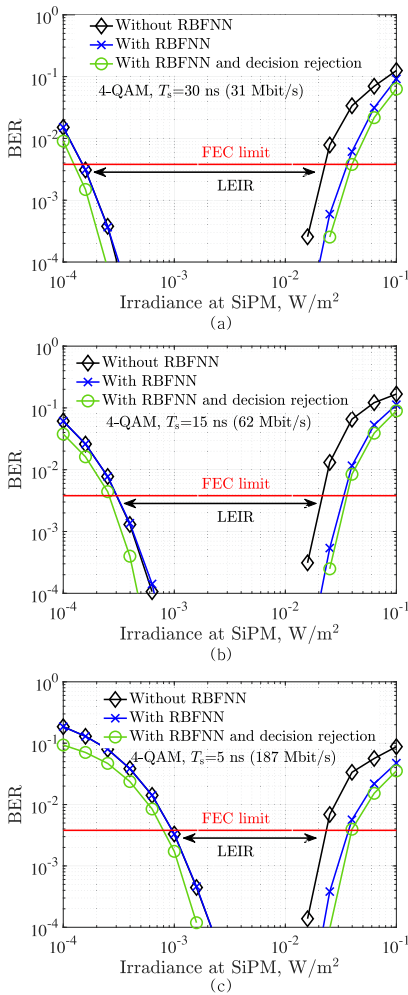


Fig. 17. The simulated BER plotted as a function of the irradiance level when $\gamma=0.1$ and $K/L=10$ for 4-QAM with different sampling periods of 30 ns, 15 ns and 5 ns.

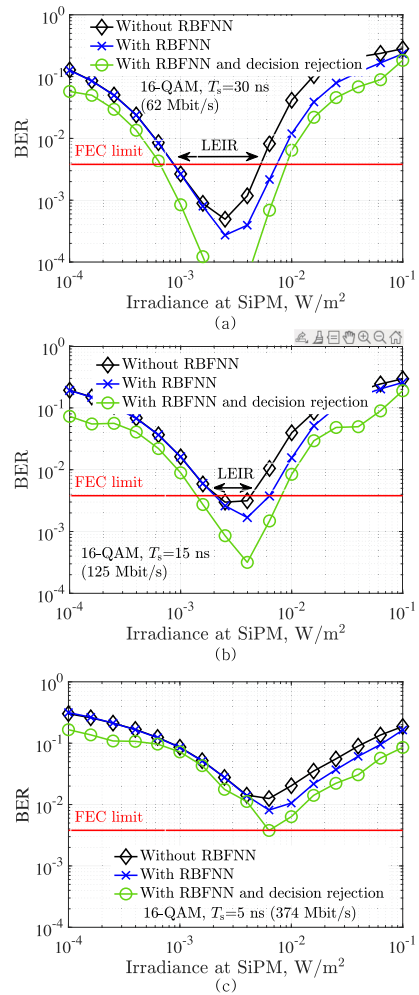


Fig. 18. The simulated BER plotted as a function of the irradiance level when $\gamma=0.1$ and $K/L=10$ for 16-QAM with different sampling periods of 30 ns, 15 ns and 5 ns.

be the forward error correction (FEC) limit which is typically fixed at 3.8×10^{-3} [29]. As shown in Figs. 17 and 18, the irradiance range in which the BER is below the FEC limit is called the low error irradiance range (LEIR) and it is considered

as a performance metric in this paper. A higher value of LEIR means that the SiPM works for a wider range of irradiance levels which is very desirable for the transmission system. In Figs. 17 and 18, the irradiance is considered in a log scale and

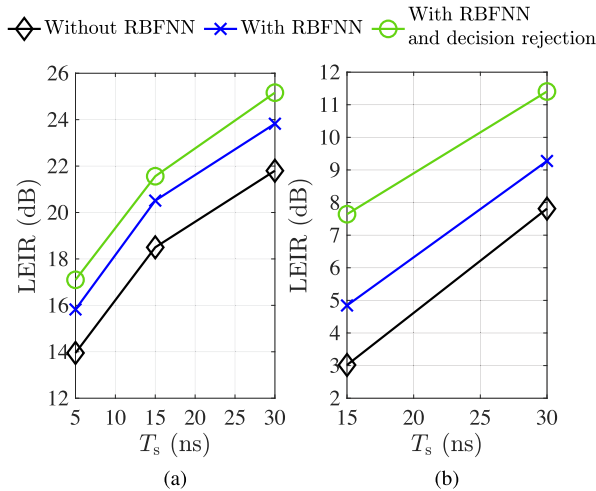


Fig. 19. The LEIR for different sampling periods, (a) 4-QAM, (b) 16-QAM.

between 10^{-4} W/m² and 10^{-1} W/m². Since the area of the SiPM is 9.42 mm² as listed in Table I, this irradiance range is equivalent to the received optical power between -60.25 dBm and -30.25 dBm. To give the values of the LEIR in a log scale, it is expressed in dB by using the higher received optical power level (in dBm) which is associated with a BER of 3.8×10^{-3} minus the lower received optical power level (in dBm) which is also associated with a BER of 3.8×10^{-3} . Fig. 19 shows the obtained LEIRs for different sampling periods. Note that in the case when 16-QAM and T_s is 5 ns, the obtained BERs without RBFNN and with RBFNN are all above the FEC limit and only the RBFNN with decision rejection can achieve a BER below the FEC limit, its LEIR results are therefore not shown in Fig. 19. First, by comparing Fig. 19(a) with Fig. 19(b), we can see that the use of 4-QAM results in much lower LEIRs than the case of 16-QAM. Second, it can be seen that, for both 4-QAM and 16-QAM, the LEIR increases when a larger value of T_s is considered. This is because the influences of the SiPM nonlinearity reduce when the value of T_s increases. Third, in the case of 4-QAM, we can see that the use of RBFNN can result in a 2 dB improvement compared to the case without RBFNN. Also, using RBFNN with decision rejection, a further improvement of 1 dB~1.5 dB can be achieved. In the case of 16-QAM as shown in Fig. 19(b), the use of RBFNN can result in an improvement of 1 dB~2 dB. Moreover, the use of decision rejection can lead to additional gains of 2 dB~2.5 dB and therefore the overall performance gain is up to 4.5 dB.

C. Influence of RBF Neuron Width

The results in Section VI suggest that the performance of the RBFNN will depend critically on the width of the RBF neurons. This parameter can be varied using the parameter γ , which is defined in a way that means that smaller values of γ are associated with wider RBF neurons. The changes in BER as γ is varied for two typical scenarios are shown in Fig. 20. Since the conventional detection method doesn't include any RBF neurons the BER obtained using this method is independent of γ . For

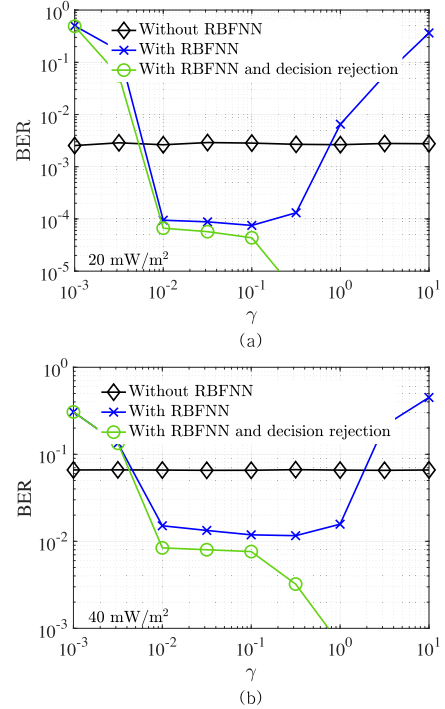


Fig. 20. The simulated BER plotted as a function of γ when $T_s = 15$ ns and two irradiance levels of 20 mW/m² and 40 mW/m² are considered for 4-QAM.

the case of the RBFNN based receiver, the BER first decreases and then increases when γ is changed from 0.001 to 10. This is because when γ is small, that is the RBF neurons are large, each RBF neuron covers areas that should be associated with multiple outputs and therefore the BERs are high. However, as shown in Section VI, when the widths of the RBF neurons are too small, a large area of the constellation space will generate similar outputs and the correct category can't be reliably determined for many different inputs. However, as shown in Fig. 20, this problem can be solved by using the enhanced RBFNN receiver with decision rejection. Moreover, if the widths of the RBF neurons are small, the distribution pattern of the constellations can be well reflected into the RBFNN outputs and this results in much lower BERs. Moreover, for a given irradiance level, since smaller RBF widths would cause more decisions to be rejected, the number of rejected decisions can be used to determine if the widths of the RBFs are too narrow.

D. Influences of Decision Rejection Threshold

In the above sections, the results suggest that the BER performance of the receiver can be improved by using decision rejection. Since the performance of this approach also depends on the choice of the decision rejection threshold, its influences are investigated in this section. Fig. 21 shows the simulated BER and the decision rejection ratio (e.g. a ratio between the number of transmitted bits which are rejected and the overall number of transmitted bits) plotted as a function of the decision rejection threshold. It can be seen that, when the decision rejection threshold is changed from 0 to 1, the BER decreases

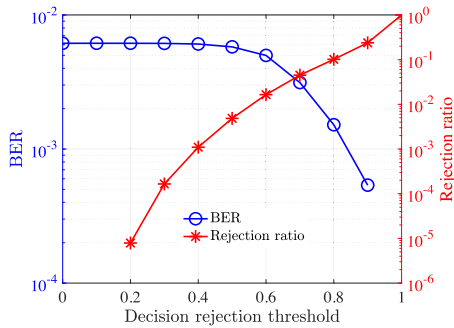


Fig. 21. The BER and the decision rejection ratio plotted as a function of the decision rejection threshold when $T_s = 15$ ns and the irradiance levels is 35 mW/m^2 for 4-QAM.

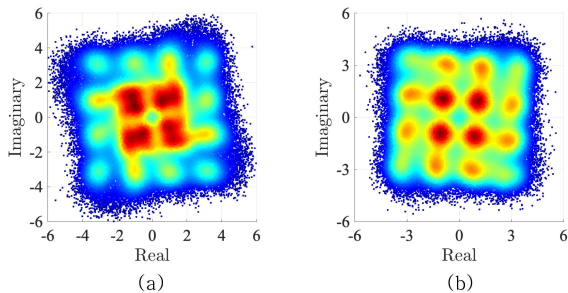


Fig. 22. The 16-QAM constellation when $T_s = 15$ ns and the irradiance is 20 mW/m^2 , (a) without channel equalization (b) with channel equalization.

and the rejection ratio increases. This is because a higher decision rejection threshold would result in more rejections and consequently a lower BER. It also shows that, when the decision rejection threshold is lower than 0.3, no decisions are rejected and consequently the BER is not affected. When the decision rejection threshold is greater than 0.7, the simulated BER is zero. In this case, further increasing the rejection threshold would no longer reduce the BER but only significantly affects the transmission data rate. In this paper, the decision rejection threshold is considered to be 0.7 so that the BER can be reduced and the transmission data rate is not crucially affected.

E. The Combination of Signal Pre-Equalization With RBFNN

Using a RBFNN can result in lower BERs, however, its performance will be degraded whenever there is a strong overlap between inputs that are associated with different elements of the constellation, especially when the constellation size becomes large (e.g. 16-QAM). Fig. 22(a) shows an example of the received 16-QAM constellations when the irradiance is 20 mW/m^2 . It can be seen that the overlapping between different constellation elements is strong. In this case, even the complex non-linear classification boundaries achieved using a group of RBFs cannot correctly classify some inputs. In [14], we showed that this overlapping effect can be reduced by pre-equalizing the transmitted signal. Fig. 22(b) shows the received signal constellations after implementing the signal equalization technique proposed in [14] when the irradiance is 20 mW/m^2 . We can see that the overlapping effect between different constellation elements is much reduced. However, in order to more correctly

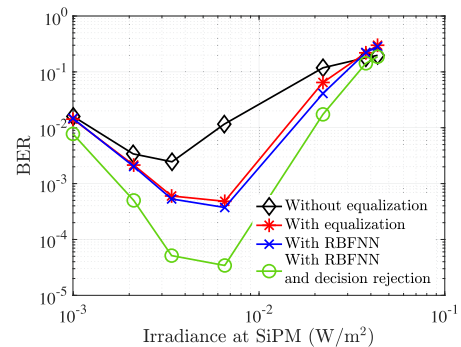


Fig. 23. The BER performance of the system when an additional signal pre-equalization step is used together with the proposed RBFNN receiver ($\gamma=0.1$, $K/L=10$, 16-QAM).

classify these constellation values into their right categories, non-linear classification boundaries are still required which means the RBFNN can be used together with signal equalization to further enhance the system performance.

Fig. 23 shows the simulated BER as a function of the irradiance level for four different signal demodulation methods. The first approach is to directly demodulate the received signal without signal equalization or a RBFNN. The second approach is to demodulate the signal after implementing signal pre-equalization [14]. The third approach is to combine signal pre-equalization with the proposed RBFNN. The fourth approach is to combine signal pre-equalization with the RBFNN and decision rejection. It shows that the BER can be reduced by pre-equalizing the transmitted signal. Then, by applying the proposed RBFNN onto the equalized signal, the BERs can be further reduced. In this case, the reduction in the BER is relatively minor. However, by combining the pre-equalization with the RBFNN with decision rejection, the BER is significantly reduced compared to other three cases.

IX. CONCLUSION

In this paper, a new RBFNN-based signal detection method has been introduced to deal with the impact of the nonlinear response of SiPMs. The motivation for using the new approach has been explained by showing the impact of the SiPM non-linearity on the received signal constellations. In particular, results were presented which showed that the SiPM non-linearity causes a unique form of distortion to the elements of a constellation. In particular each element of the constellation becoming associated with multiple clusters of input data and the clusters associated with each element of the constellation form an irregular region on the input space. A RBFNN is a particularly suitable way of dealing with these multiple clusters and irregular regions.

In this study, k -means algorithm was implemented to determine both the locations of the centers of the RBF neurons and the initial widths of these neurons. Results have been presented which showed that this method places the RBF neurons in regions of the input space associated with a high density of training data. Then, a gradient descent algorithm was applied to train the weighting coefficients between the RBF layer and

the output layer of the RBFNN. The result is a RBFNN that can divide the input space into areas associated with each element of the constellation. Furthermore, the borders of each area can have a complex, non-linear shape which can lead to a more accurate determination of the constellation element associated with each input. Consequently, a RBFNN was shown to significantly reduce the BER results obtained when a VLC channel that uses OFDM and a SiPM receiver is simulated. In addition, unlike other methods a RBFNN can detect when an input is so atypical that the outputs are unreliable and shouldn't therefore be used. In this paper, we defined a performance metric named LEIR to quantify the irradiance range in which the transmission error rate is below the FEC limit, 3.8×10^{-3} . We show that the use of the RBFNN can result in a 2 dB gain for 4-QAM and a gain of 1 dB~2 dB for 16-QAM. The use of decision rejection with RBFNN can lead to further gains of 1 dB for 4-QAM and 2.5 dB for 16-QAM. In this paper, we also investigate the choice of the decision rejection threshold used in the enhanced RBFNN. We show that increasing the decision rejection threshold can lead to a lower BER at the cost of reducing the transmission data rate. Our simulation results show that a good choice of the decision rejection threshold is around 0.7 so that the BER can be reduced significantly and the transmission data rate is not crucially affected. Furthermore, we also look into the combination of the conventional signal equalization techniques with the proposed RBFNN. We show that, although the use of signal equalization can reduce the influence of the SiPM nonlinearity, its performance can be further enhanced by using the proposed RBFNN demodulation methods.

REFERENCES

- [1] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: Opportunities, challenges and the path to market," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–32, Dec. 2013.
- [2] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1533–1544, Mar. 2016.
- [3] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
- [4] D. O'Brien, S. Rajbhandari, and H. Chun, "Transmitter and receiver technologies for optical wireless," *Philos. Trans. Roy. Soc. A*, vol. 378, no. 2169, 2020, Art. no. 20190182.
- [5] D. Chitnis and S. Collins, "A SPAD-based photon detecting system for optical communications," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 2028–2034, 2014.
- [6] E. Fisher, I. Underwood, and R. Henderson, "A reconfigurable 14-bit 60GPhoton/s single-photon receiver for visible light communications," in *Proc. Eur. Solid-State Circuits Conf.*, 2012, pp. 85–88.
- [7] L. Zhang, H. Chun, Z. Ahmed, G. Faulkner, D. O'Brien, and S. Collins, "The future prospects for SiPM-based receivers for visible light communications," *J. Lightw. Technol.*, vol. 37, no. 17, pp. 4367–4374, 2019.
- [8] W. Matthews, Z. Ahmed, W. Ali, and S. Collins, "A SiPM-based VLC receiver for 3.45 Gigabits/s communication using OOK modulation," in *Proc. IEEE Photon. Conf.*, 2020, pp. 1–2.
- [9] W. Matthews, Z. Ahmed, W. Ali, and S. Collins, "A 3.45 Gigabits/s SiPM-based OOK VLC receiver," *IEEE Photon. Technol. Lett.*, vol. 33, no. 10, pp. 487–490, May 2021.
- [10] S. Huang, C. Chen, R. Bian, H. Haas, and M. Safari, "5 Gbps optical wireless communication using commercial SPAD array receivers," *Opt. Lett.*, vol. 47, no. 9, pp. 2294–2297, 2022.
- [11] L. Zhang, R. Jiang, X. Tang, Z. Chen, J. Chen, and H. Wang, "A simplified post equalizer for mitigating the nonlinear distortion in SiPM based OFDM-VLC system," *IEEE Photon. J.*, vol. 14, no. 1, Feb. 2022, Art. no. 7301707.
- [12] W. Matthews, W. Ali, Z. Ahmed, G. Faulkner, and S. Collins, "Inter-symbol interference and silicon photomultiplier VLC receivers in ambient light," *IEEE Photon. Technol. Lett.*, vol. 33, no. 9, pp. 449–452, May 2021.
- [13] C. He, Z. Ahmed, and S. Collins, "Optical OFDM and SiPM receivers," in *Proc. IEEE Globecom Workshops*, 2020, pp. 1–6.
- [14] C. He, Z. Ahmed, and S. Collins, "Signal pre-equalization in a silicon photomultiplier-based optical OFDM system," *IEEE Access*, vol. 9, pp. 23344–23356, 2021.
- [15] Y. Liu, S. Bi, Z. Shi, and L. Hanzo, "When machine learning meets Big Data: A wireless communication perspective," *IEEE Veh. Technol. Mag.*, vol. 15, no. 1, pp. 63–72, Mar. 2020.
- [16] X. Wu and D. C. O'Brien, "A novel machine learning-based handover scheme for hybrid LiFi and WiFi networks," in *Proc. IEEE Globecom Workshops*, 2020, pp. 1–5.
- [17] C. Chen, L. Zeng, X. Zhong, S. Fu, M. Liu, and P. Du, "Deep learning-aided OFDM-based generalized optical quadrature spatial modulation," *IEEE Photon. J.*, vol. 14, no. 1, Feb. 2022, Art. no. 7302306.
- [18] F. Dong and D. O'Brien, "A neural network-based adaptive MIMO-VLC system," in *Proc. IEEE Photon. Conf.*, 2021, pp. 1–2.
- [19] Y. Zhao, P. Zou, Z. He, Z. Li, and N. Chi, "Low spatial complexity adaptive artificial neural network post-equalization algorithms in MIMO visible light communication systems," *Opt. Exp.*, vol. 29, no. 20, pp. 32728–32738, 2021.
- [20] R. Jiang, C. Sun, L. Zhang, X. Tang, H. Wang, and A. Zhang, "Deep learning aided signal detection for SPAD-based underwater optical wireless communications," *IEEE Access*, vol. 8, pp. 20363–20374, 2020.
- [21] D. S. Broomhead and D. Lowe, "Radial basis functions, multi-variable functional interpolation and adaptive networks," Royal Signals and Radar Establishment Malvern, U.K., Tech. Rep. 4148, 1988.
- [22] E. Sarbazi, M. Safari, and H. Haas, "The bit error performance and information transfer rate of SPAD array optical receivers," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5689–5705, Sep. 2020.
- [23] Silicon Photomultiplier Sensors, J-Series, Accessed: Jun. 24, 2022. [Online]. Available: <https://www.onsemi.com/products/sensors/photodetectors-sipm-spad/silicon-photomultipliers-sipm/j-series-sipm>
- [24] M. I. Jordan *et al.*, "Why the logistic function? a tutorial discussion on probabilities and neural networks," MIT, Computational cognitive science technical report, Tech. Rep. 9503, Aug. 1995.
- [25] M. D. Richard and R. P. Lippmann, "Neural network classifiers estimate bayesian a posteriori probabilities," *Neural Comput.*, vol. 3, no. 4, pp. 461–483, 1991.
- [26] D. J. MacKay, *Information Theory, Inference and Learning Algorithms*. New York, NY, USA: Cambridge Univ. Press, 2003.
- [27] S. Ruder, "An overview of gradient descent optimization algorithms," 2016, *arXiv:1609.04747*.
- [28] J. Cid-Sueiro, J. I. Arribas, S. Urbán-Munoz, and A. R. Figueiras-Vidal, "Cost functions to estimate a posteriori probabilities in multiclass problems," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 645–656, May 1999.
- [29] Forward error correction for high bit-rate DWDM submarine systems (G.975.1), 2004, Accessed: Jun. 24, 2022. [Online]. Available: <https://www.itu.int/rec/T-REC-G.975.1-200402-1/en>