

# Interference Cancellation for Layered Asymmetrically Clipped Optical OFDM With Application to Optical Receiver Design

Thomas Q. Wang<sup>ib</sup>, Hang Li<sup>ib</sup>, and Xiaojing Huang<sup>ib</sup>, *Senior Member, IEEE*

**Abstract**—In this paper, we study a novel two-stage receiver to demodulate layered asymmetrically clipped optical orthogonal frequency division multiplexing for intensity modulation direct detection based visible light communications. Designed for avoiding the error propagation of the conventional receiver, the first stage of the receiver is a soft interference cancellation (SIC) module which evaluates the minimum mean square error (MMSE) estimates of the signals in different layers. For this stage, we derive the exact formula of the MMSE estimator, and compare the achieved mean square error and bit error rate (BER) with those of the existing simplified SIC receiver. We show that the estimation error in a layer has negligible impact on the design of estimators in the subsequent layers. Using the outputs of the SIC module, the second stage performs noise clipping to suppress the additive noise. For this stage, we present two schemes, the SIC-based iterative noise clipping (SIC-INC) and the SIC-based direct noise clipping (SIC-DNC). The simulation results show that SIC-INC can achieve BERs similar to those of the SIC-based diversity combining receiver with optimum combining coefficients. It is also shown that SIC-DNC outperforms the existing advanced receivers by up to 0.8 dB at the BER of  $10^{-4}$ .

**Index Terms**—Layered ACO-OFDM, minimum mean square error estimation, noise clipping receiver, soft interference cancellation, visible light communications.

## I. INTRODUCTION

OPTICAL orthogonal frequency division multiplexing (OFDM) is increasingly being considered for use in intensity modulation direct detection (IM/DD) based visible light communications (VLC) because of its robustness to inter-symbol interference and high spectral efficiency [1]. Unlike conventional radio frequency (RF) communications, optical OFDM signals used in IM/DD must be real and non-negative. This necessitates a series of signal processing techniques which lead to various optical OFDM schemes with different energy and spectral efficiencies. For all the representative schemes, Hermitian symmetry is employed at the inputs of inverse fast Fourier transform (IFFT) to create real signals in time domain. The non-negativity of the signals is then ensured typically using one of the two approaches: 1) the addition of a large direct

current (DC) bias and 2) asymmetrical signal clipping. The former approach produces the DC biased optical OFDM (DCO-OFDM) [2] which achieves high spectral efficiency through loading half of its available subcarriers with independent information symbols. However, due to the use of a large DC bias, the energy efficiency of DCO-OFDM is low. The clipping based approach, on the contrary, leads to a range of unipolar OFDM schemes with high energy efficiency. These include asymmetrically clipped optical OFDM (ACO-OFDM) [3], unipolar OFDM [4] and PAM-DMT [5]. Because of the restrictions in the processing techniques, all of these schemes inevitably experience inherent loss in spectral efficiency.

Recent years have witnessed the extensive research on advanced optical OFDM schemes that simultaneously provide high energy and spectral efficiencies [6]–[12]. These include asymmetrically clipped DC biased optical OFDM (ADO-OFDM) [6], hybrid ACO-OFDM [7], layered/enhanced ACO-OFDM [8], [9], enhanced unipolar OFDM (eU-OFDM) [10], enhanced PAM-DMT [11] and spectral and energy efficient OFDM [12]. In all these schemes, to achieve high spectral efficiency, multi-layer/depth structures are employed with the conventional schemes [3]–[5] included in their first layer/depth and the extra information symbols loaded onto the subsequent layers/depths.<sup>1</sup> In each layer/depth, Hermitian symmetry and asymmetrical clipping are applied to ensure the signal is real and non-negative with high energy efficiency.<sup>2</sup> Thus, improved spectral and energy efficiencies are achieved when the signals of different layers/depths are superimposed to form the intensity signal. In this paper, we focus on the layered ACO-OFDM which exploits the periodic and anti-periodic properties of IFFT to form multiple subcarrier groups (SCGs), each corresponding to a layer [13]. Then the gap of spectral efficiency between DCO-OFDM and ACO-OFDM can be closed when all the available SCGs are employed [8], [14].

The asymmetrical clipping applied in layered ACO-OFDM results in signal-dependent clipping noise which, after being created in a layer, falls only onto the subsequent layers as inter-layer interference (ILI) [13]. Therefore, the decoding of layered ACO-OFDM signals must be proceeded from low layer to high layer, where for each layer, the ILI is recovered and canceled. As far as we know, the methods available for the signal recovery fall into

Manuscript received May 10, 2017; revised September 29, 2017 and January 4, 2018; accepted February 4, 2018. Date of current version March 2, 2018. This work was supported under an Australian Research Councils Discovery funding scheme (DP 160101693). (Corresponding author: Thomas Q. Wang.)

The authors are with Global Big Data Technologies Centre, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: qian.wang@uts.edu.au; hang.li@uts.edu.au; xiaojing.huang@uts.edu.au).

<sup>1</sup>As described in [6]–[8], [11], it is convenient to define layers/depths for ADO-OFDM, hybrid ACO-OFDM, layered ACO-OFDM and enhanced PAM-DMT in frequency domain, whereas for eU-OFDM, in time domain [10].

<sup>2</sup>The only exception is ADO-OFDM which applies asymmetrical clipping on the ACO-OFDM component and a large DC bias on the DCO-OFDM component.

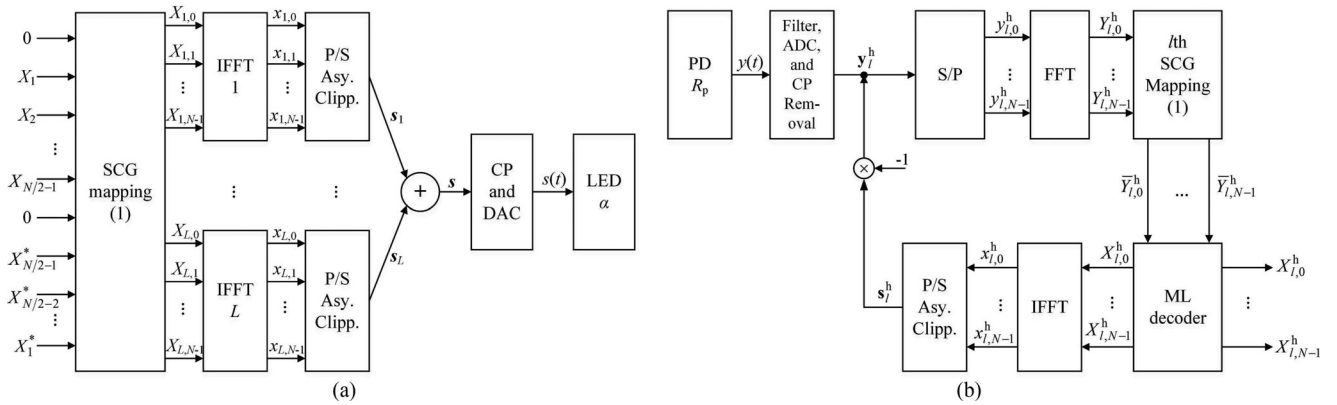


Fig. 1. Layered ACO-OFDM, (a) transmitter, and (b) conventional receiver.

two categories: the hard-decision based ILI cancellation [8] and the minimum mean square error (MMSE)-estimation based soft interference cancellation (SIC) [15]. Aiming to recover the exact transmitted signal, the hard-decision based method directly decodes the information symbols in each layer and uses the decoded symbols to reconstruct the ILI, assuming no decoding errors. This method, however, is sensitive to error propagation which occurs when symbols in a layer are decoded incorrectly. Alternatively, the SIC can be used in a receiver for layered ACO-OFDM [15]. Unlike its hard-decision based counterpart, in the ILI cancellation stage, the SIC module estimates other than decodes the information symbols in each layer, and uses the estimates to reconstruct the ILI for cancellation. The limitation of the SIC stems from the estimation errors of the MMSE which cause residue interference (RI) in the subsequent layers after ILI cancellation [15]. Similar with error propagation, the RI results in degraded performance in the high layers.

The research on the SIC based receiver is at a very early stage where there are lots of open problems to be addressed. These include 1) the robustness of the SIC to the inaccurate knowledge of noise power, and 2) the processing of the SIC outputs. In this paper, we will study these problems. The study on the robustness stems from the estimator developed in [15] where the RI after ILI cancellation is ignored. Thus, the resulting estimator is designed based on the inaccurate knowledge of noise power. To address this problem, we derive the exact formula of the MMSE estimators in the SIC module, taking into account the RI and additive noise, and compare the achieved results with those given by [15]. We show that although the RI has an impact on the mean square error (MSE) in low and medium signal-to-noise ratios (SNRs), the ignorance of it in the estimator design leads to no noticeable degradation in either MSE or BER. This confirms that the simplified estimator developed in [15] is effective. The study on the processing of SIC outputs is motivated by the advanced receivers recently proposed in [15]–[17]. In general, the improved reception of ACO-OFDM signals falls into two categories: diversity combining [18]–[20] and noise clipping [21], [22], both leading to improvements up to 3 dB. For both categories, a similar decision-making process must be performed to decide the sign of the samples (for diversity combining) or to decide the pure additive noise samples (for noise cancellation)<sup>3</sup>. As the decisions are made in accordance with the samples in

time domain which are Gaussian distributed, taking small values with high probability, incorrect decisions inevitably occurs constantly in the presence of additive noise, leading to performance degradation [21]. The SIC module provides a range of outputs which can be used to improve the performance of layered ACO-OFDM receiver. These include the estimated samples, the recovered (but yet estimated) samples and the recovered clipping noise samples. In this paper, we propose two new schemes for noise clipping receivers, one being SIC based iterative noise clipping (SIC-INC) and the other, SIC based direct noise clipping (SIC-DNC), both making use of the high-precision estimates of SIC module. For SIC-INC, the estimates of the ACO-OFDM samples and the decisions regarding pure noise samples are iteratively updated assisted by the recovered clipping noise. We show that without having to determine the combining coefficients, this scheme can achieve similar performance with that of the SIC based diversity combining (SIC-DC) receiver which employ the optimum coefficients [15]. For SIC-DNC, all the three outputs are used with the recovered samples added together to form the recovered ACO-OFDM samples in each layer and the estimates in the corresponding layer indicating the pure noise samples. It is shown that this scheme can significantly reduce the noise power, leading to BERs outperforming the existing advanced receivers by up to 0.8 dB at the BER of  $10^{-4}$ . The rest of this paper is organized as follows. The layered ACO-OFDM and its conventional hard decision based reception are described in Section II. In Section III, the SIC module is developed, with the exact MMSE estimators derived. In Section IV, we describe the two new noise clipping receivers. The simulation results are presented in Section V. The paper is concluded in Section VI.

## II. LAYERED ACO-OFDM

### A. Multi-Layer Optical OFDM

Fig. 1(a) shows an  $L$ -layer ACO-OFDM transmitter which converts independent symbols,  $\mathbf{X} = [0, X_1, \dots, X_{N/2-1}, 0, X_{N/2-1}^*, \dots, X_1^*]^T$  drawn from an  $M$ -QAM constellation, to a real nonnegative intensity signal,  $s(t)$ . Here, we denote the size of the IFFT by  $N$ , and the vector/matrix transpose and complex conjugate by the superscripts  $[\cdot]^T$  and  $[\cdot]^*$ , respectively. The transmitter consists of  $L$  single layer ACO-OFDM modulators (layers), each modulating the symbols onto the subcarriers included in their corresponding SCGs. The inputs to the  $l$ th,  $l \leq L$ ,

<sup>3</sup>In fact, similar process is also used in the improved receivers for unipolar OFDM [23], [24] and hybrid ACO-OFDM [25].

modulator (layer), denoted by  $\mathbf{X}_l = [X_{l,0}, X_{l,1}, \dots, X_{l,N-1}]^T$ , is related to the input symbol,  $\mathbf{X}$ , by

$$X_{l,n} = \begin{cases} X_n, & \text{if } n \in \mathcal{G}_l \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where  $\mathcal{G}_l$  denotes the  $l$ th SCG which includes those subcarriers with indexes dividable by  $2^{l-1}$  but not dividable by  $2^l$  [13]. Therefore, it can be seen that the conventional ACO-OFDM corresponds to the first layer modulator which loads symbols only onto the odd subcarriers.

The IFFT in the  $l$ th layer generates a sequence of samples,  $\mathbf{x}_l = [x_{l,0}, x_{l,1}, \dots, x_{l,N-1}]^T$ , in time domain, given by

$$x_{l,k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_{l,n} e^{j \frac{2\pi}{N} kn}, \quad \text{for } 0 \leq k \leq N-1. \quad (2)$$

These samples are clipped at zero to produce the output of the  $l$ th modulator,  $\mathbf{s}_l = [s_{l,0}, s_{l,1}, \dots, s_{l,N-1}]^T$ , where [26]

$$s_{l,k} = \frac{1}{2}x_{l,k} + \frac{1}{2}|x_{l,k}| = \begin{cases} x_{l,k}, & \text{if } x_{l,k} \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Denote the FFT of  $\mathbf{s}_l$  by  $\mathbf{S}_l$ . Then (3) indicates that  $\mathbf{S}_l$  consists of two components: the information symbols with half amplitude loaded onto the subcarriers included in the  $l$ th SCG and the clipping noise given by half of the absolute value of  $\mathbf{x}_l$ . As the time sequence,  $\mathbf{x}_l$ , consists of  $2^{l-1}$  identical sub-sequences, each being anti-periodic [13], the clipping noise,  $|\mathbf{x}_l|/2$ , is  $2^l$ -folded periodic in time domain. This indicates that after being converted using FFT, the clipping noise only falls onto the subcarriers allocated to higher layers and the DC, leading to the ILI.

The outputs of the clipping devices are summed up to form the samples of the layered ACO-OFDM signal,  $\mathbf{s} = [s_0, \dots, s_{N-1}]^T$ , given by

$$s_k = \sum_{l=1}^L s_{l,k}, \quad \text{for } 0 \leq k \leq N-1. \quad (4)$$

Then the cyclic prefix is appended followed by the digital-to-analog conversion.

The electrical power of layered ACO-OFDM signal,  $P_{\text{elec}}$ , can be expressed as [27]

$$P_{\text{elec}} = E(s_k^2) = \frac{1}{2} \sum_{l=1}^L \sigma_l^2 + \frac{1}{\pi} \sum_{l_1, l_2, l_1 \neq l_2}^L \sigma_{l_1} \sigma_{l_2}, \quad (5)$$

where  $\sigma_l^2$  is the electrical power of  $x_{l,k}$ , given by  $E(x_{l,k}^2)$ . Denote the bit rate of a single layer ACO-OFDM by  $r_1$  bits per second. Then the rate of an  $L$ -layer signal employing identical constellation can be expressed as [8]

$$r_L = \left(1 + \frac{1}{2} + \dots + \frac{1}{2^{L-1}}\right) r_1 = 2 \left(1 - \left(\frac{1}{2}\right)^L\right) r_1, \quad (6)$$

leading to the average electrical energy per bit,  $E_{b,\text{elec}}$ , given by  $E_{b,\text{elec}} = P_{\text{elec}}/r_L$  [6].

## B. Conventional Receiver

Fig. 1(b) shows the conventional hard-decision based receiver for layered ACO-OFDM in the context of additive white

Gaussian noise (AWGN) channel<sup>4</sup> [8]. After the optical-to-electrical conversion at the photodetector (PD), the received signal,  $y(t)$ , can be expressed as

$$y(t) = \alpha R_p s(t) + w(t), \quad (7)$$

where  $\alpha$  and  $R_p$  denote the electrical-to-optical conversion coefficient at the LED and the PD's responsivity, respectively, and  $w(t)$ , the AWGN at the receiver, with single-sided power density,  $N_0$ . Without loss of generality, we assume  $\alpha R_p = 1$  [28]. Thus the received signal can be rewritten as

$$y(t) = s(t) + w(t), \quad (8)$$

which, after the analog-to-digital conversion and the removal of CP, produces the received samples,  $\mathbf{y} = [y_0, \dots, y_{N-1}]^T$ , given by

$$\mathbf{y} = \mathbf{s} + \mathbf{w}, \quad (9)$$

where  $\mathbf{w} = [w_0, \dots, w_{N-1}]^T$  denotes the Gaussian noise with power of  $\sigma_w^2$ . Assuming the anti-aliasing filter used in the receiver has a bandwidth of  $B$  which equals that of the transmitted signal, the noise power,  $\sigma_w^2$ , is given by  $N_0 B$ .

The conventional receiver performs a recursive procedure of hard-decoding and interference cancellation on the samples,  $\mathbf{y}$ . As the ILI generated in a layer only falls onto the subsequent layers, the demodulation must be performed from the first layer, where the information symbols are free of interference. Assuming no decoding errors, the samples in the first layer,  $\mathbf{s}_1^h$ , can be reconstructed using the decoded symbols. It is then cancelled from  $\mathbf{y}$  to form the modified samples,  $\mathbf{y}_2^h$ , in which the symbols in the second layer are ideally free of interference.<sup>5</sup> The whole procedure ends when the symbols in all the layers are decoded.

Fig. 1(b) shows the decoding of the symbols in the  $l$ th layer followed by the ILI cancellation. As shown in the figure, the samples,  $\mathbf{y}_l^h$ , are first converted to frequency domain using FFT. Then the outputs,  $\mathbf{Y}_l^h = [Y_{l,0}^h, \dots, Y_{l,N-1}^h]^T$ , are sent to the SCG mapping device which forces the values loaded on the subcarriers not included in the  $l$ th SCG to zeros to produce  $\bar{\mathbf{Y}}_l^h = [\bar{Y}_{l,0}^h, \dots, \bar{Y}_{l,N-1}^h]^T$ , where

$$\bar{Y}_{l,n}^h = \begin{cases} Y_{l,n}^h, & \text{if } n \in \mathcal{G}_l \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

The vector,  $\bar{\mathbf{Y}}_l^h$ , is input to the ML decoder which makes hard decisions on the values loaded in the  $l$ th layer in accordance with [8]

$$X_{l,n}^h = \arg \min_{X \in \mathcal{X}} |2X - \bar{Y}_{l,n}^h|^2, \quad \text{for } n \in \mathcal{G}_l, \quad (11)$$

where  $\mathcal{X}$  denotes the set including all the possible values of the employed constellation. For the subcarriers not included in the  $l$ th SCG, zeros are returned by the ML decoder. Assuming there are no decoding errors, the samples in the  $l$ th layer,  $\mathbf{s}_l^h$ , is reconstructed using the decoded symbols and the forced zeros. They are then subtracted from  $\mathbf{y}_l^h$  to produce the modified samples,

<sup>4</sup>We focus on the AWGN channel in this paper. The study can be extended to frequency selective channels by sequentially inserting an  $N$ -point FFT, a bank of single-tap equalizers and an  $N$ -point IFFT in between the Filter, ADC and CP removal module and the demodulation loop in Fig. 1(b) and 2.

<sup>5</sup>From the perspective of demodulation,  $\mathbf{y}$  is equivalent to  $\mathbf{y}_1^h$  from which the symbols in the first layer are decoded/estimated.

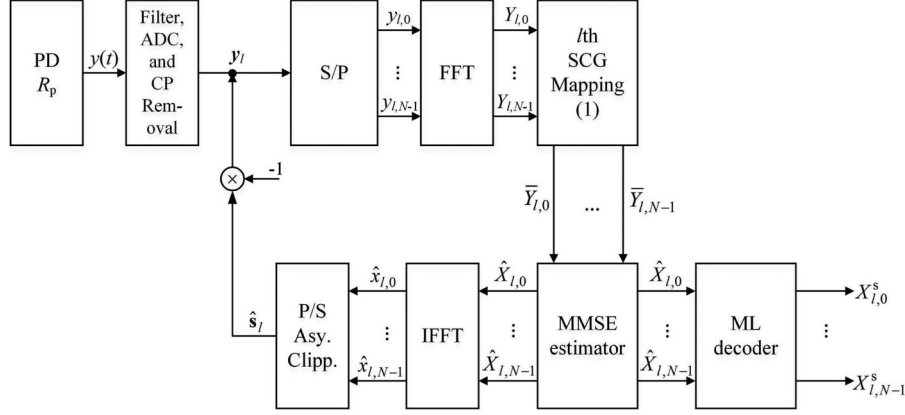


Fig. 2. SIC based layered ACO-OFDM receiver.

TABLE I  
SIC BASED RECEIVER FOR LAYERED ACO-OFDM

<b>Input:</b> Received sample, $\mathbf{y}$ , and noise power, $\sigma_w^2$
<b>Output:</b> Decoded symbol, $X_{l,n}^s$
1. <b>Initialization:</b> $\sigma_{Z_0}^2 = \sigma_w^2$ and $\chi_0^2 = 0$ .
2. <b>for</b> $l = 1 : L$ <b>do</b>
3.     Update the noise power, $\sigma_{Z_l}^2 = \sigma_{Z_{l-1}}^2 + \chi_{l-1}^2/4$ .
4.     Evaluate the estimates, $\hat{X}_{l,n}$ , $n \in \mathcal{G}_l$ using (26) and (27).
5.     Decode the estimates to have $X_{l,n}^s$ using ML.
6.     Calculate the power of $\hat{\mathbf{x}}_l$ , $\hat{\sigma}_l^2$ , using (28)-(31).
7.     Calculate the power of the RI, $\chi_l^2$ , using (21)-(25).
8.     Calculate the estimates, $\hat{\mathbf{s}}_l$ , using the estimated symbols.
9.     Remove the estimates, $\hat{\mathbf{s}}_l$ , from the received sample.
10. <b>end for</b>

$\mathbf{y}_{l+1}^h$ , which are used to demodulate the symbols in the  $(l+1)$ th layer.

### III. SOFT INTERFERENCE CANCELLATION

In this section, we describe the SIC based receiver (see Fig. 2) and derive the exact expression for the MMSE estimator. The demodulation process of an SIC based receiver is summarized in Table I. As shown in the table and figure, a recursive procedure is performed for the demodulation. However, unlike the hard-decision based receiver, an MMSE estimator, other than an ML decoder, is employed to evaluate the estimates of the symbols,  $\hat{X}_{l,n}$ ,  $n \in \mathcal{G}_l$ , which are used for 1) the reconstruction of the ACO-OFDM sample,  $\hat{\mathbf{s}}_l$ , and 2) ML decoding. The use of the MMSE estimator requires that the exact probability density function (PDF) of the interference-plus-noise must be available. This will be studied in Sections III-A and III-B, with the solutions presented in steps 3, 6, and 7 of the procedure. Using the exact PDF, we derive the estimator in Section III-C.

#### A. Residue Interference

The use of an MMSE estimator inevitably induces estimation error which results in the RI in the modified samples. Fig. 2 illustrates the generation of RI when the estimates of the samples in the  $l$ th layer,  $\hat{\mathbf{s}}_l$ , is cancelled to form the modified sample,  $\mathbf{y}_{l+1}$ . As shown in Fig. 2, for the  $l$ th layer, the estimates,  $\hat{X}_{l,n}$ ,  $n \in \mathcal{G}_l$ , and their real values,  $X_{l,n}$ , are typically different, with estimation error (in frequency domain) given by  $\Delta_{\mathbf{x}_l} = X_{l,n} - \hat{X}_{l,n}$ . This means that the resulting estimated samples,  $\hat{\mathbf{s}}_l$ , given by

$$\hat{\mathbf{s}}_l = \frac{1}{2}\hat{\mathbf{x}}_l + \frac{1}{2}|\hat{\mathbf{x}}_l|, \quad (12)$$

are also different from their real values,  $\mathbf{s}_l = (\mathbf{x}_l + |\mathbf{x}_l|)/2$ . Therefore, when  $\hat{\mathbf{s}}_l$  is removed from  $\mathbf{y}_l$  to form  $\mathbf{y}_{l+1}$ , the signal component of the  $l$ th layer is not completely cancelled. First, there are estimation errors (in time domain),  $(\mathbf{x}_l - \hat{\mathbf{x}}_l)/2$ , on the subcarriers included in the  $l$ th SCG in  $\mathbf{y}_{l+1}$ . However, they are orthogonal to the information symbols loaded in the  $(l+1)$ th SCG and thus can be completely removed using the  $(l+1)$ th SCG mapping device. Therefore, the estimation errors,  $(\mathbf{x}_l - \hat{\mathbf{x}}_l)/2$ , have no impact on the demodulation in the  $(l+1)$ th layer. Second, the estimation errors,  $(|\mathbf{x}_l| - |\hat{\mathbf{x}}_l|)/2$ , fall onto the subcarriers included in the  $(l+1)$ th,  $\dots$ ,  $L$ th SCGs. Hence, they cannot be removed using the SCG mapping device, resulting in the RI.

As estimation errors exist in all the layers, the RI accumulates as the demodulation proceeds. We examine the observation (recovered samples in frequency domain),  $\bar{\mathbf{Y}}_l = [\bar{Y}_{l,0}, \dots, \bar{Y}_{l,N-1}]^T$ , at the inputs of the estimator for the demodulation of the  $l$ th layer. As shown in Fig. 2, the modified samples,  $\mathbf{y}_l$ , are the difference between  $\mathbf{y}_{l-1}$  and the estimated samples,  $\hat{\mathbf{s}}_{l-1}$ . Thus using (3), (4) and (9), they can be expressed as

$$\begin{aligned} \mathbf{y}_l &= \mathbf{y}_{l-1} - \hat{\mathbf{s}}_{l-1} \\ &= \frac{1}{2}(\mathbf{x}_1 - \hat{\mathbf{x}}_1) + \dots + \frac{1}{2}(\mathbf{x}_{l-1} - \hat{\mathbf{x}}_{l-1}) \\ &\quad + \frac{1}{2}(|\mathbf{x}_1| - |\hat{\mathbf{x}}_1|) + \dots + \frac{1}{2}(|\mathbf{x}_{l-1}| - |\hat{\mathbf{x}}_{l-1}|) \\ &\quad + \frac{1}{2}\mathbf{x}_l + \dots + \frac{1}{2}\mathbf{x}_L + \frac{1}{2}|\mathbf{x}_l| + \dots + \frac{1}{2}|\mathbf{x}_L| + \mathbf{w}, \quad (13) \end{aligned}$$

leading to the FFT of  $\mathbf{y}_l$ ,  $\mathbf{Y}_l = [Y_{l,0}, \dots, Y_{l,N-1}]^T$ , given by

$$\begin{aligned} \mathbf{Y}_l &= \mathbf{Y}_{l-1} - \hat{\mathbf{S}}_{l-1} \\ &= \frac{1}{2} (\mathbf{X}_1 - \hat{\mathbf{X}}_1) + \dots + \frac{1}{2} (\mathbf{X}_{l-1} - \hat{\mathbf{X}}_{l-1}) \\ &\quad + \frac{1}{2} (\mathbf{C}_1 - \hat{\mathbf{C}}_1) + \dots + \frac{1}{2} (\mathbf{C}_{l-1} - \hat{\mathbf{C}}_{l-1}) \\ &\quad + \frac{1}{2} \mathbf{X}_l + \dots + \frac{1}{2} \mathbf{X}_L + \frac{1}{2} \mathbf{C}_l + \dots + \frac{1}{2} \mathbf{C}_L + \mathbf{W}, \end{aligned} \quad (14)$$

where  $\mathbf{C}_i = [C_{i,0}, \dots, C_{i,N-1}]^T$ ,  $\hat{\mathbf{C}}_i = [\hat{C}_{i,0}, \dots, \hat{C}_{i,N-1}]^T$ ,  $1 \leq i \leq l-1$  and  $\mathbf{W} = [W_0, \dots, W_{N-1}]^T$  denote the FFTs of  $|\mathbf{x}_i|$ ,  $|\hat{\mathbf{x}}_i|$  and  $\mathbf{w}$ , respectively, and  $\hat{\mathbf{X}}_i = [\hat{X}_{i,0}, \dots, \hat{X}_{i,N-1}]^T$ , represents the MMSE estimate of  $\mathbf{X}_i$  with zeros filled on the subcarriers not included in the  $i$ th SCG.

The SCG mapping device forces the values on the subcarriers not included in the  $l$ th SCG to zeros. This results in the observation on the  $n$ th subcarrier given by

$$\bar{Y}_{l,n} = \begin{cases} \frac{1}{2} X_{l,n} + W_n + \frac{1}{2} \sum_{i=1}^{l-1} (C_{i,n} - \hat{C}_{i,n}), & \text{if } n \in \mathcal{G}_l \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where when  $n \in \mathcal{G}_l$ , the first term denotes the symbols to be estimated, and the second and third, the additive Gaussian noise and the accumulated RI, respectively. As the RI,  $(C_{i,n} - \hat{C}_{i,n})/2$ ,  $1 \leq i \leq l-1$ , created in the  $i$ th layer is the  $n$ th element of  $(\mathbf{C}_i - \hat{\mathbf{C}}_i)/2$ , its statistical characteristic can be derived by analyzing  $(|\mathbf{x}_i| - |\hat{\mathbf{x}}_i|)$  which is the IFFT of  $(\mathbf{C}_i - \hat{\mathbf{C}}_i)$ .

### B. Distribution of Residue Interference

As shown in Fig. 2, the estimated samples,  $\hat{\mathbf{x}}_i$ , are the IFFT of  $\hat{\mathbf{X}}_i$ , i.e.,

$$\hat{x}_{i,k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{X}_{i,n} e^{j \frac{2\pi}{N} kn}, \quad \text{for } 0 \leq k \leq N-1. \quad (16)$$

Thus, it is Gaussian distributed with zero mean, assuming that the number of subcarriers included in the  $i$ th SCG is sufficiently large. As zeros are forced in  $\hat{\mathbf{X}}_i$ , for the subcarriers not included in the  $i$ th SCG, similar with  $\mathbf{x}_i$ ,  $\hat{\mathbf{x}}_i$  consists of  $2^{i-1}$  identical subsequences, each being anti-periodic. This indicates that the difference,  $(|\mathbf{x}_i| - |\hat{\mathbf{x}}_i|)$ , is  $2^i$ -folded periodical, falling only onto the subcarriers included in the  $(i+1)$ th,  $\dots$ ,  $L$ th SCGs.

The samples in a period of  $|\mathbf{x}_i|$  are independent and identically distributed (IID), and thus so are those of  $|\hat{\mathbf{x}}_i|$  [36]. Therefore, when the number of IID samples,  $N/2^i$ , is sufficiently large, using the central limit theorem, the FFT of the difference,  $(\mathbf{C}_i - \hat{\mathbf{C}}_i)$ , will have Gaussian distributed values on the subcarriers included in the  $(i+1)$ th,  $\dots$ ,  $L$ th SCGs, leading to the RI,  $(C_{i,n} - \hat{C}_{i,n})/2$ ,  $1 \leq i \leq l-1$ , Gaussian distributed. Denoting the expected value and variance of  $(C_{i,n} - \hat{C}_{i,n})$ ,  $1 \leq i \leq l-1$ , by  $\mu_i$  and  $\chi_i^2$ , respectively, the accumulated RI,  $\sum_{i=1}^{l-1} (C_{i,n} - \hat{C}_{i,n})/2$ , is complex Gaussian distributed,

given by

$$\frac{1}{2} \sum_{i=1}^{l-1} (C_{i,n} - \hat{C}_{i,n}) \sim \mathcal{CN} \left( \frac{1}{2} \sum_{i=1}^{l-1} \mu_i, \frac{1}{4} \sum_{i=1}^{l-1} \chi_i^2 \right). \quad (17)$$

Therefore, the observation given by (14) can be rewritten as

$$\bar{Y}_{l,n} = \frac{1}{2} X_{l,n} + Z_{l,n}, \quad \text{for } n \in \mathcal{G}_l, \quad (18)$$

where  $Z_{l,n}$  denotes the interference-plus-noise following Gaussian distribution, i.e.,

$$Z_{l,n} \sim \mathcal{CN} \left( \frac{1}{2} \sum_{i=1}^{l-1} \mu_i, \sigma_w^2 + \frac{1}{4} \sum_{i=1}^{l-1} \chi_i^2 \right). \quad (19)$$

Denote the electrical power of  $\hat{\mathbf{x}}_i$  by  $\hat{\sigma}_i^2$ , i.e.,  $\hat{\sigma}_i^2 = E(\hat{x}_{i,k}^2)$ ,  $0 \leq k \leq N-1$ . Then, the term,  $(|\mathbf{x}_i| - |\hat{\mathbf{x}}_i|)$ , is the difference of two folded Gaussian distributed variables, with the expected value and variance given by

$$E(|x_{i,k}| - |\hat{x}_{i,k}|) = \frac{2}{\sqrt{2\pi}} (\sigma_i - \hat{\sigma}_i) \quad (20)$$

and

$$\begin{aligned} \text{Var}(|x_{i,k}| - |\hat{x}_{i,k}|) &= \left(1 - \frac{2}{\pi}\right) \sigma_i^2 + \left(1 - \frac{2}{\pi}\right) \hat{\sigma}_i^2 \\ &\quad + \frac{4}{\pi} \sigma_i \hat{\sigma}_i - 2E(|x_{i,k} \hat{x}_{i,k}|), \end{aligned} \quad (21)$$

respectively, where [33]

$$E(|x_{i,k} \hat{x}_{i,k}|) = \frac{2}{\pi} \sigma_i \hat{\sigma}_i \left( \rho \arcsin \rho + (1 - \rho^2)^{1/2} \right), \quad (22)$$

and  $\rho$  denotes the correlation coefficient between  $x_{i,k}$  and  $\hat{x}_{i,k}$ , given by

$$\rho = \frac{E(x_{i,k} \hat{x}_{i,k})}{\sigma_i \hat{\sigma}_i} = \frac{\hat{\sigma}_i}{\sigma_i}, \quad (23)$$

where the second equation holds for symmetrical constellations in which the in-phase and quadrature components have identical probability mass functions. See the typical examples of these constellations which can be not only square but also non-square in [34], and Appendix A for the derivations of (20), (21) and (23).

When the difference,  $(|\mathbf{x}_i| - |\hat{\mathbf{x}}_i|)$ , is converted to  $(\mathbf{C}_i - \hat{\mathbf{C}}_i)$  using FFT, the power carried by the expected value will be allocated to the DC only, whereas the variance, equally to the subcarriers included in all the subsequent layers and the DC. The use of the unitary FFT, (2), means that [37]

$$N \text{Var}(|x_{i,k}| - |\hat{x}_{i,k}|) = (N/2^i) \chi_i^2, \quad (24)$$

where the left-hand-side (LHS) of (24) represents the overall power input to the FFT and the right-hand-side (RHS), that of the output. The value,  $N/2^i$ , on the RHS denotes the number of subcarriers with non-zero values loaded. Therefore, the expected value,  $\mu_i$ , and variance,  $\chi_i^2$  of  $(C_{i,n} - \hat{C}_{i,n})$  can be expressed as

$$\mu_i = 0, \quad \chi_i^2 = 2^i \text{Var}(|x_{i,k}| - |\hat{x}_{i,k}|). \quad (25)$$

### C. MMSE Estimator for SIC

The MMSE estimate of  $X_{l,n}$ ,  $n \in \mathcal{G}_l$ , with an observation,  $\bar{Y}_{l,n}$ , is given by the expectation of  $X_{l,n}$  conditioned on  $\bar{Y}_{l,n}$ , i.e.,  $\hat{X}_{l,n} = E(X_{l,n} | \bar{Y}_{l,n})$  [32]. For an  $M$ -QAM symbol perturbed by AWGN, its estimate can be expressed as  $\hat{X}_{l,n} = \hat{a}_{l,n} + j\hat{b}_{l,n}$ , where

$$\hat{a}_{l,n} = \frac{\sum_{a \in \mathcal{X}_R} a \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(\Re(\bar{Y}_{l,n}) - \frac{1}{2}a\right)^2\right]}{\sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(\Re(\bar{Y}_{l,n}) - \frac{1}{2}a\right)^2\right]} \quad (26)$$

and

$$\hat{b}_{l,n} = \frac{\sum_{b \in \mathcal{X}_I} b \Pr[\Im(X_{l,n}) = b] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(\Im(\bar{Y}_{l,n}) - \frac{1}{2}b\right)^2\right]}{\sum_{b \in \mathcal{X}_I} \Pr[\Im(X_{l,n}) = b] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(\Im(\bar{Y}_{l,n}) - \frac{1}{2}b\right)^2\right]} \quad (27)$$

with  $\sigma_{Z_l}^2 = \sigma_w^2 + \sum_{i=1}^{l-1} \chi_i^2/4$ ,  $l \geq 2$ , denoting the effective noise power in the  $l$ th layer, where  $\sigma_{Z_1}^2 = \sigma_w^2$  (see step 3 of Table I).  $\Pr(\cdot)$  denotes the probability of a specific incidence included in the brackets.  $\Re(\cdot)$ ,  $\Im(\cdot)$ ,  $\mathcal{X}_R$  and  $\mathcal{X}_I$  represent the real and imaginary parts of a complex number, and the sets that include all the possible values for the real and imaginary parts of the employed  $M$ -QAM constellation, respectively. Because of the symmetry,  $\mathcal{X}_R$  and  $\mathcal{X}_I$  are identical. See Appendix B for the derivation of (26) and (27).

The resulting MSE,  $E(|\Delta_{\mathbf{X}_l}|^2)$ , is given by

$$\begin{aligned} E(|\Delta_{\mathbf{X}_l}|^2) &= 2 \sum_{a \in \mathcal{X}_R} a^2 \Pr[\Re(X_{l,n}) = a] - \frac{2}{\sqrt{\pi\sigma_{Z_l}^2}} \\ &\times \int_{-\infty}^{\infty} \frac{\left(\sum_{a \in \mathcal{X}_R} a \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y - \frac{1}{2}a\right)^2\right]\right)^2}{\sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y - \frac{1}{2}a\right)^2\right]} dy. \end{aligned} \quad (28)$$

See Appendix C for the derivation of (28).

Using the orthogonal principle of MMSE estimation [32], the power of the estimates,  $E(|\hat{X}_{l,n}|^2)$ , is given by the difference between  $E(|X_{l,n}|^2)$  and  $E(|\Delta_{\mathbf{X}_l}|^2)$ , i.e.,

$$E(|\hat{X}_{l,n}|^2) = E(|X_{l,n}|^2) - E(|\Delta_{\mathbf{X}_l}|^2). \quad (29)$$

When unitary IFFT is applied to the estimates of the symbols in the  $l$ th layer, the following equation holds

$$\frac{N}{2^l} E(|\hat{X}_{l,n}|^2) = N \hat{\sigma}_l^2, \quad (30)$$

where the LHS of (30) denotes the overall power of the inputs and the RHS, that of the outputs. Thus, the power,  $\hat{\sigma}_l^2$ , is

given by

$$\hat{\sigma}_l^2 = \frac{1}{2^l} E(|\hat{X}_{l,n}|^2). \quad (31)$$

### IV. NOISE CLIPPING RECEIVER

The SIC based receiver provides two sequences which can be used to improve the reception of layered ACO-OFDM signals. These are 1) the estimated samples,  $\hat{\mathbf{x}}_l$ , given by (16), and 2) the recovered samples (in time domain),  $\bar{\mathbf{y}}_l = [\bar{y}_{l,0}, \dots, \bar{y}_{l,N-1}]^T$ ,  $1 \leq l \leq L$ , which are the IFFT of the inputs of the MMSE estimator, given by

$$\bar{y}_{l,k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \bar{Y}_{l,n} e^{j\frac{2\pi}{N}kn}. \quad (32)$$

In addition, a third sequence, the recovered clipping noise samples, can be derived using the received and estimated samples. In frequency domain, an estimate of the clipping noise created in the  $l$ th layer,  $\bar{\mathbf{C}}_l$ , can be expressed as [15]

$$\bar{\mathbf{C}}_l = 2\mathbf{Y} - \sum_{m=1}^L \hat{\mathbf{X}}_m - \sum_{m \neq l}^L \hat{\mathbf{C}}_m, \quad (33)$$

leading to the recovered clipping noise samples,  $\bar{\mathbf{c}}_l = [\bar{c}_{l,0}, \dots, \bar{c}_{l,N-1}]^T$ , given by the IFFT of  $\bar{\mathbf{C}}_l$ , where  $\mathbf{Y}$  denotes the FFT of the received sample,  $\mathbf{y}$ . Note that although representing an alternative of  $|\hat{\mathbf{x}}_l|$ ,  $\bar{\mathbf{c}}_l$  includes the additive noise on the subcarriers included in the  $(l+1)$ th,  $\dots$ ,  $L$ th SCGs, whereas  $|\hat{\mathbf{x}}_l|$ ,  $\hat{\mathbf{x}}_l$  and  $\bar{\mathbf{y}}_l$  depend only on the noise on those in the  $l$ th SCG and the accumulated RI. Therefore, the degree of correlation between the noise component of  $\bar{\mathbf{c}}_l$  and that of  $\hat{\mathbf{x}}_l$  ( $\bar{\mathbf{y}}_l$ ) is low. This enables the improved reception using these sequences.

In this section, we study two schemes for noise clipping receiver, both of which exploits the fact that an ACO-OFDM sequence,  $\mathbf{s}_l$ , always has a zero/non-negative sample and a non-negative/zero sample separated by  $N/2^l$  samples [13]. Unlike the existing works [16], [21]–[23], we use the estimated sample,  $\hat{\mathbf{x}}_l$ , to identify the pure noise samples in the constructed ACO-OFDM sequence and clip them to reduce the noise power. As the estimates,  $\hat{\mathbf{X}}_l$ , are the optimum estimates of  $\mathbf{X}_l$  in the sense of MSE, the use of the IFFT,  $\hat{\mathbf{x}}_l$ , will provide more accurate decisions than the recovered sample,  $\bar{\mathbf{y}}_l$ , will.

#### A. SIC Based Iterative Noise Clipping

The SIC-INC scheme uses the estimated samples,  $\hat{\mathbf{x}}_l$ , and the recovered clipping noise samples,  $\bar{\mathbf{c}}_l$ , to improve BERs. The idea behind iterative clipping is to successively update the two sequences to reduce their noise power. The iteration starts with  $\hat{\mathbf{x}}_l$  and  $\bar{\mathbf{c}}_l$  which are returned by the SIC module. First, for each layer, an estimate of the ACO-OFDM samples,  $\bar{\mathbf{s}}_l = [\bar{s}_{l,0}, \dots, \bar{s}_{l,N-1}]^T$ , are constructed as

$$\bar{\mathbf{s}}_l = (\hat{\mathbf{x}}_l + \bar{\mathbf{c}}_l)/2. \quad (34)$$

Then noise clipping is applied to the constructed signal, where the decisions on the pure additive noise samples are made in accordance with the signs of the estimated samples,  $\hat{\mathbf{x}}_l$ . Denoting the clipped samples by  $\bar{\mathbf{s}}_l^c = [\bar{s}_{l,0}^c, \dots, \bar{s}_{l,N-1}^c]^T$ , they are related with the constructed signal,  $\bar{\mathbf{s}}_l$ , by

$$\bar{s}_{l,k}^c = \frac{1}{2} \bar{s}_{l,k} (\text{sgn}(\hat{x}_{l,k}) + 1), \quad \text{for } 0 \leq k \leq N-1, \quad (35)$$

where  $\text{sgn}(\cdot)$  denotes the sign function returning 1 with a non-negative input and  $-1$ , otherwise. The clipped sample is then passed through an FFT followed by an SCG mapping device corresponding to the  $l$ th SCG and an IFFT to produce an updated estimate sample,  $\hat{\mathbf{x}}_l^u = [\hat{x}_{l,0}^u, \dots, \hat{x}_{l,N-1}^u]^T$ .

Next, we update the recovered clipping noise samples. Define two auxiliary variables,  $\tilde{\mathbf{c}}_l = |\hat{\mathbf{x}}_l^u|$  and  $\mathbf{x}_l^f = [x_{l,0}^f, \dots, x_{l,N-1}^f]^T$ , where  $\mathbf{x}_l^f$  denotes the clipping noise flipped in accordance with the sign of the updated estimate samples. Then  $\mathbf{x}_l^f$  is given by

$$\mathbf{x}_{l,k}^f = \text{sgn}(\hat{x}_{l,k}^u) \tilde{c}_{l,k}, \quad (36)$$

which is equivalent to the recovered component in diversity combining [18]. By summing up the two auxiliary variables, an alternative estimate of the ACO-OFDM samples,  $\tilde{\mathbf{s}}_l = [\tilde{s}_{l,0}, \dots, \tilde{s}_{l,N-1}]^T$ , can be constructed as

$$\tilde{s}_l = (\mathbf{x}_l^f + \tilde{\mathbf{c}}_l) / 2, \quad (37)$$

on which noise clipping is applied using the updated estimate samples,  $\hat{\mathbf{x}}_l^u$ , to produce the clipped samples,  $\tilde{\mathbf{s}}_l^c = [\tilde{s}_{l,0}^c, \dots, \tilde{s}_{l,N-1}^c]^T$ , given by

$$\tilde{s}_{l,k}^c = \frac{1}{2} \tilde{s}_{l,k} (\text{sgn}(\hat{x}_{l,k}^u) + 1), \quad \text{for } 0 \leq k \leq N - 1. \quad (38)$$

The clipped samples are then passed through an FFT followed by an SCG mapping device to force the values on the subcarriers included in the first  $l$  SCGs to zeros and an IFFT to have the updated clipping noise, given by  $\tilde{\mathbf{c}}_l^u = [\tilde{c}_{l,0}, \dots, \tilde{c}_{l,N-1}]^T$ .

The first iteration concludes with the updated sequences,  $\hat{\mathbf{x}}_l^u$  and  $\tilde{\mathbf{c}}_l^u$ , which substitute  $\hat{\mathbf{x}}_l$  and  $\tilde{\mathbf{c}}_l$ , respectively, to be used in the second iteration. The demodulation process for an  $I$ -iteration noise clipping receiver is summarized in Table II.

### B. SIC Based Direct Noise Clipping

The SIC-DNC scheme exploits the estimated samples,  $\hat{\mathbf{x}}_l$ , the recovered samples,  $\tilde{\mathbf{y}}_l$ , and the recovered clipping noise,  $\tilde{\mathbf{c}}_l$ , to improve BERs. First, for the  $l$ th layer, a sequence of ACO-OFDM samples,  $\tilde{\mathbf{s}}_l = [\tilde{s}_{l,0}, \dots, \tilde{s}_{l,N-1}]^T$ , is formed as

$$\tilde{s}_l = (\tilde{\mathbf{y}}_l + \tilde{\mathbf{c}}_l) / 2. \quad (39)$$

Then, noise clipping is performed to produce the clipped samples,  $\tilde{\mathbf{s}}_l^c = [\tilde{s}_{l,0}^c, \dots, \tilde{s}_{l,N-1}^c]^T$ , which are related to  $\tilde{\mathbf{s}}_l$  by

$$\tilde{s}_{l,k}^c = \frac{1}{2} \tilde{s}_{l,k} (\text{sgn}(\hat{x}_{l,k}) + 1), \quad \text{for } 0 \leq k \leq N - 1. \quad (40)$$

The clipped samples are then sent to a standard ACO-OFDM receiver to have the information symbols in this layer decoded. The demodulation process for an SIC-DNC receiver is summarized in Table III.

### C. Computational Complexity

In this section, the numbers of addition and multiplication that are additionally required by the proposed receivers are analyzed on top of the conventional layered ACO-OFDM [8]. We show that the overall complexity of the proposed receivers is still in the order of  $O(N \log_2 N)$ . In line with the evaluations for layered ACO-OFDM [8] and for FFT/IFFT algorithms [29], we assume the size of IFFT/FFT is sufficiently large. Compared with the conventional layered ACO-OFDM receiver, the demodulation

TABLE II  
SIC BASED ITERATIVE NOISE CLIPPING

<b>Input:</b> estimated samples, $\hat{\mathbf{x}}_l$ , recovered clipping noise, $\tilde{\mathbf{c}}_l$ , and number of iteration, $I$
<b>Output:</b> Decoded symbol, $X_{l,n}^{\text{inc}}$
1. <b>for</b> $l = 1 : L$ <b>do</b>
2. <b>for</b> $p = 1 : I$ <b>do</b>
3.         Construct an estimate of $\mathbf{s}_l$ , $\tilde{\mathbf{s}}_l$ , using (34)
4.         Noise clipping in accordance with the sign of $\hat{\mathbf{x}}_l$ , i.e. (35), to produce the clipped sample, $\tilde{\mathbf{s}}_l^c$ .
5.         Send $2\tilde{\mathbf{s}}_l^c$ to an FFT, followed by an SCG mapping device and an IFFT to produce an updated estimate, $\hat{\mathbf{x}}_l^u$ .
6.         Form the auxiliary sequences, $\tilde{\mathbf{c}}_l =  \hat{\mathbf{x}}_l^u $ and $\mathbf{x}_l^f$ .
7.         Construct an alternative estimate of $\mathbf{s}_l$ , $\tilde{\mathbf{s}}_l$ , using (37).
8.         Noise clipping in accordance with the sign of $\hat{\mathbf{x}}_l^u$ , i.e. (38), to produce the clipped sample, $\tilde{\mathbf{s}}_l^c$ .
9.         Send $2\tilde{\mathbf{s}}_l^c$ to an FFT followed by an SCG mapping device and an IFFT to produce an updated estimate, $\tilde{\mathbf{c}}_l^u$ .
10.        Update: $\hat{\mathbf{x}}_l^u \rightarrow \hat{\mathbf{x}}_l$ and $\tilde{\mathbf{c}}_l^u \rightarrow \tilde{\mathbf{c}}_l$ .
11. <b>end for</b>
12.     Demodulate $\hat{\mathbf{x}}_l$ using standard OFDM reception method.
13. <b>end for</b>

TABLE III  
SIC BASED DIRECT NOISE CLIPPING

<b>Input:</b> estimated samples, $\hat{\mathbf{x}}_l$ , recovered samples, $\tilde{\mathbf{y}}_l$ , and recovered clipping noise, $\tilde{\mathbf{c}}_l$ ,
<b>Output:</b> Decoded symbol, $X_{l,n}^{\text{dnc}}$
1. <b>for</b> $l = 1 : L$ <b>do</b>
2.     Construct an estimate of the ACO-OFDM samples in the $l$ th layer, $\tilde{\mathbf{s}}_l$ , using (39)
3.     Noise clipping in accordance with the sign of $\hat{\mathbf{x}}_l$ , i.e. (40), to produce the clipped sample, $\tilde{\mathbf{s}}_l^c$ .
4.     Send $2\tilde{\mathbf{s}}_l^c$ to an FFT followed by an SCG mapping device and ML decoders to retrieve the information symbol in the $l$ th layer.
5. <b>end for</b>

procedures for the exact SIC, SIC-INC and SIC-DNC shown in Tables I–III indicate that the increase of complexity in the receivers mainly comes from 1) the MMSE estimators (26) and (27), and 2) the noise clipping.

The MMSE estimators, given by (26) and (27), show that only elementary functions, such as exponential function, are involved in the calculation. Therefore, the estimates can be efficiently evaluated using the existing well-developed algorithms, such as those in [30].

For each layer, SIC-DNC sums up the recovered samples and clipping noise, followed by noise clipping (steps 2 and

3 of Table III). This causes extra  $2N$  additions and multiplications on top of the exact SIC. Taking into account the FFT after noise clipping (step 4), we can see that for each layer, extra  $N \log_2 N + 2N$  additions and multiplications are required. Thus, when  $L$  layers are to be demodulated, the overall extra computational complexity of SIC-DNC is given by  $L(N \log_2 N + 2N)$ .

As shown in Table II, the extra computations performed in each layer by SIC-INC are caused by steps 3–9 and 12. The steps, 3–9, included in each iteration result in  $N$  (eq. (34)),  $N$  (eq. (35)),  $2N \log_2 N$  (FFT and IFFT),  $N$  (eq. (36)),  $N$  (eq. (37)),  $N$  (eq. (38)) and  $2N \log_2 N$  (FFT and IFFT) additions and multiplications, respectively. Thus, when  $I$  iterations are to be performed, the resulting computations are given by  $(5N + 4N \log_2 N)I$ . Taking into account the  $N \log_2 N$  additions and multiplications caused by step 12, the computations required for demodulating a layer are given by  $(5N + 4N \log_2 N)I + N \log_2 N$ . Therefore, the overall extra computational complexity of SIC-INC on top of exact SIC is given by  $((5N + 4N \log_2 N)I + N \log_2 N)L$ .

## V. SIMULATION RESULTS

In this section, we present the simulation results to show the performance of the SIC based receiver and the proposed noise clipping receivers. Comparisons with the existing receivers [8], [15]–[17] are also provided. We suppose 100,000 layered ACO-OFDM signals using  $N = 1024$  subcarriers are transmitted through an AWGN channel, with the transmitted power,  $P_{\text{elec}} = 1$ , and received signal given by (8). Perfect synchronization is assumed at the receiver which employs an anti-aliasing filter with a bandwidth of  $B = 2$  MHz, equalling that of the wanted signal. The received signal is then sampled with Nyquist sampling rate, producing the samples given by (9), which are used in the considered receivers. In order to implement the MMSE estimator, the noise power,  $\sigma_w^2$ , is assumed known to the receiver.

### A. Exact Versus Simplified SIC

First, we examine the robustness of the SIC based receiver to the inaccurate knowledge of noise power. As a counterpart of the derived exact SIC, the simplified SIC [15] is employed for comparison. For simplified SIC, the accumulated RI is ignored. Thus, the noise power is assumed to be a constant equalling  $\sigma_w^2$  in all the layers.

The comparisons between the two SIC receivers are made using four-layer ACO-OFDM signals which modulate 4-, 16-, 64-, and 256-QAM symbols onto the subcarriers. Fig. 3 shows the achieved MSEs of the four layers of the two receivers, where the derived MSEs, (28), are verified by the simulation results (marked with asterisks) and compared with the simulated ones achieved by the simplified receiver (marked with circles). From this figure, we can see that the MSEs take higher values in the higher layers for low to medium SNRs. This is caused by the accumulated RI which is superimposed onto the additive noise. Thus, the effective noise power increases as shown in step 3 of Table I, leading to higher MSE in the higher layers. The MSEs achieved by the two receivers are identical in the first layer. This is because given the demodulation process in Table I, the symbols in the first layer are free of RI, leading to the estimator given by (26) and (27) identical with

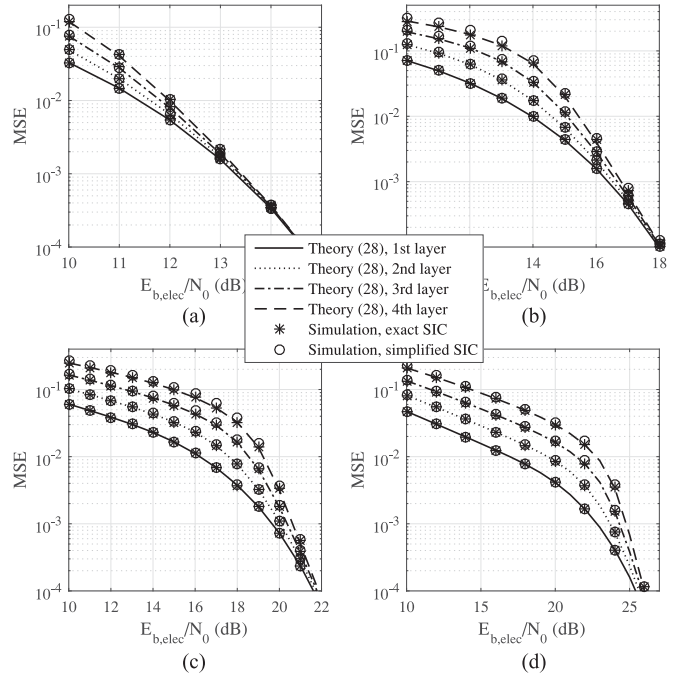


Fig. 3. Comparison of MSEs achieved by the exact and simplified SIC based receivers, (a) 4-QAM, (b) 16-QAM, (c) 64-QAM, and (d) 256-QAM.

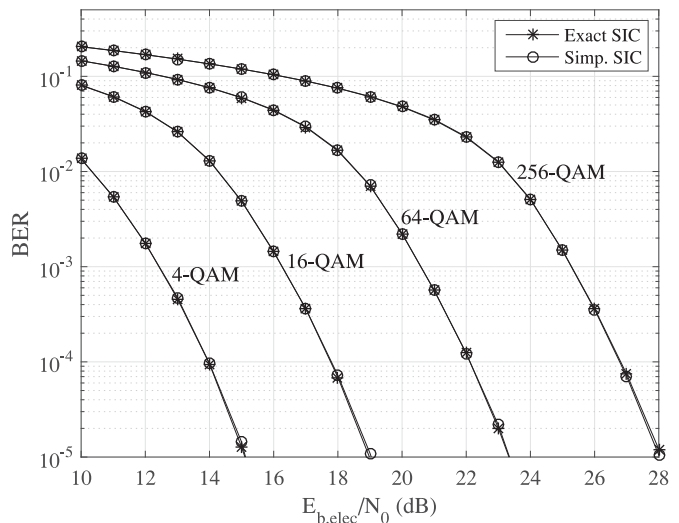


Fig. 4. Comparison of BERs achieved by the exact and simplified SIC receivers.

those of the simplified SIC. We can also see that the MSEs achieved by the two receivers deviate in the subsequent three layers, with the exact SIC slightly outperforming its simplified counterpart. This means that although the use of inaccurate knowledge of noise power causes performance loss, the loss is negligible in terms of MSEs. The similar MSEs of the two receivers indicate that the resulting BERs are identical. This is confirmed by Fig. 4, where BERs are plotted as a function of  $E_{b,\text{elec}}/N_0$  for all the considered constellations. Consequently, we can see that when its lower complexity is taken into account, simplified SIC is an effective approach to demodulating layered ACO-OFDM.



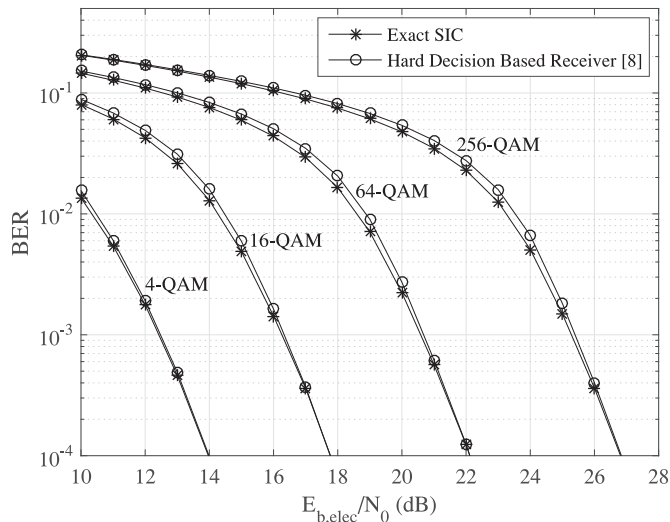


Fig. 5. Comparison of BERs achieved by the exact SIC receivers and the conventional hard decoding.

### B. Exact SIC Versus Hard-Decision Based Receiver

Next we compare the BER of SIC based receiver with those achieved by its hard-decision based counterpart [8]. The block diagrams of the two considered receivers are shown in Figs. 2 and 1(b), respectively. As we have discussed, the two receivers take completely different approaches to processing the information symbols. Therefore, their performance are limited by different facts. For the hard-decision based receiver, error propagation has been shown a major cause of performance degradation. This is reflected by the degrading BERs in the high layers [8]. For SIC based receivers, on the contrary, error propagations are absent. However, the RI superimposed onto the additive noise accumulates in the demodulation and thus contributes increasingly to the effective noise power. The resulting MSEs have been demonstrated in Fig. 3 where the performance is shown to degrade as the demodulation proceeds. Fig. 5 shows the BERs of the two receivers which receive four-layer ACO-OFDM signals. From this figure, we can see that although limited by different facts, they have similar performance with the SIC based receiver slightly outperforming at low and medium SNRs. This indicates that the SIC alone cannot provide significant improvement on the BERs over its hard-decision based counterpart.<sup>6</sup>

The BER of each layer is plotted in Fig. 6 for the two receivers which demodulate four-layer ACO-OFDM signals with 4-QAM symbols loaded onto the subcarriers. From this figure, we can see that they achieve identical performance in the first layer. This indicates that for the first layer, the use of MMSE estimators cannot provide improved BER over the conventional hard-decision based receiver. As the symbols loaded on the first layer take up  $256/(256 + 128 + 64 + 32) = 53.3\%$  of all the symbols, the BER of this layer dominates the overall BER of the four layers. Therefore, very close BERs are observed in Fig. 5 for the two receivers. The advantage of using the estimators is shown constantly in the subsequent layers, where slightly lower BERs are achieved by the exact SIC. For these layers, the SIC based

<sup>6</sup>There is a mistake in [15] for the performance simulation of the hard-decision based iterative noise clipping receiver (HD-INC). It has been corrected in this paper. See Fig. 8 and Table IV.

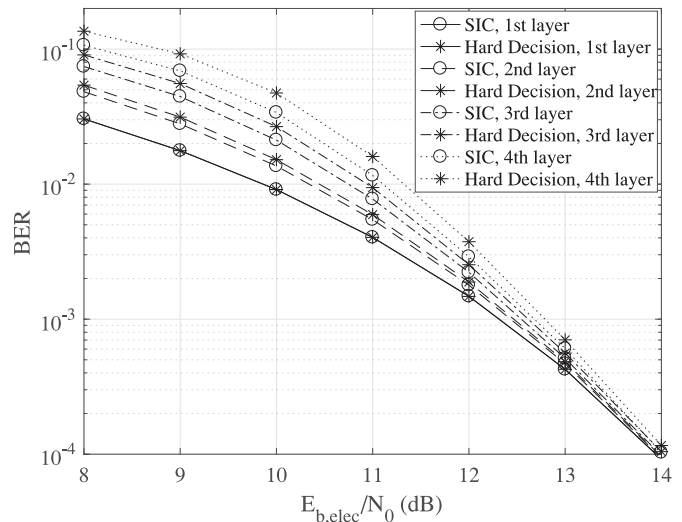


Fig. 6. BERs of each layer achieved by the exact SIC and the conventional hard-decision based receivers.

receiver cancels the ILI caused by previous layers using the estimates,  $\hat{X}_l$ , whereas its hard-decision based counterpart uses the decoded symbols to reconstruct the ILI. From this figure, it can be seen that although both RI and error propagation cause BER degradations in 2nd-4th layers, the degradation caused by RI is not as significant as that by error propagation. Therefore, from the perspective of interference cancellation, SIC slightly outperforms the hard-decision based receiver.

### C. BER of SIC Based Iterative Noise Clipping

Fig. 7 shows the BERs achieved by the SIC-INC scheme in which 4-QAM symbols are loaded onto triple-layer ACO-OFDM signals. Up to four iterations are configured in the receiver to avoid introducing unacceptably high complexity. From this figure, we can see that noise clipping is an effective approach to improving the BER. The scheme with one iteration provides a significant improvement of approximately 1 dB at the BER of  $10^{-4}$  compared with the output of the SIC alone (shown in Fig. 2). We can also see that the BER of the scheme improves with increasing number of iterations. In particular, the most significant improvement is seen when the number of iteration increases from one to two, whereas little improvement can be achieved when over three iterations are used. It can also be seen that as the number of iteration increases, the BERs approach to those of the SIC-DC receiver with optimum combining coefficient [15].

Hard-decision based iterative noise clipping receiver (HD-INC) is an effective noise clipping scheme for demodulating layered ACO-OFDM [16]. It distinguishes from SIC-INC in two major aspects. First, unlike SIC-INC in which the updated signal and clipping noise in a layer depend on those in the corresponding layer only, it updates the signal in a layer based on those in other layers. Second, HD-INC decodes the information symbols in each iteration, whereas for SIC-INC, hard decisions are made only after all the iterations are executed. Therefore, the two iterative receivers are fundamentally different. A comparison between them is presented in Fig. 8, where two-, three- and four-layer ACO-OFDM signals are demodulated by the two iterative receivers, each running five iterations. From Fig. 7

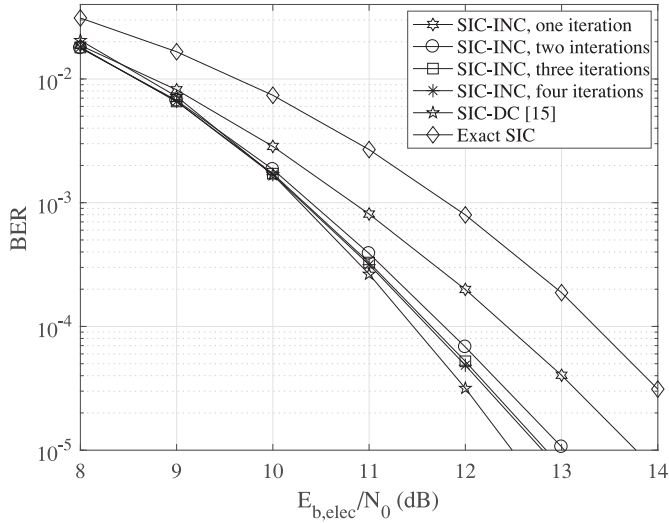


Fig. 7. Comparison of BERs between SIC-INC, SIC-DC, and the exact SIC.

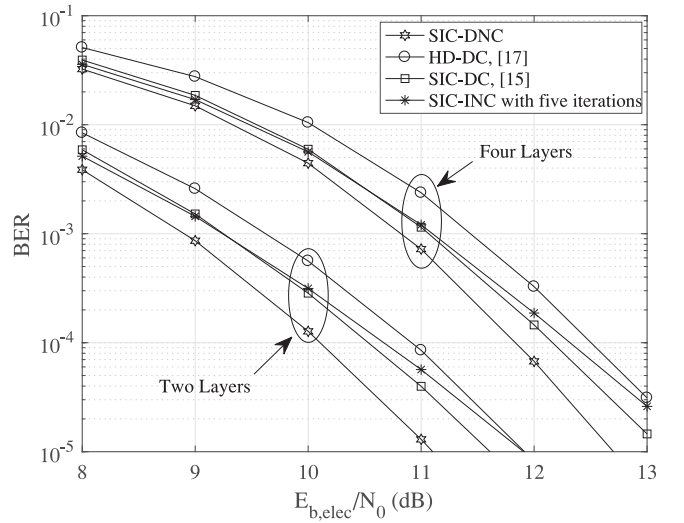


Fig. 9. Comparison of BERs of SIC-INC, SIC-DNC, SIC-DC and HD-DC.

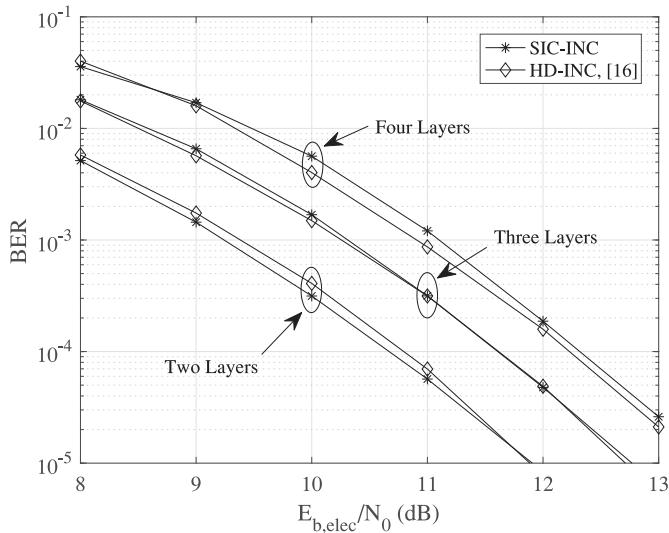


Fig. 8. Comparison between SIC-INC and HD-INC.

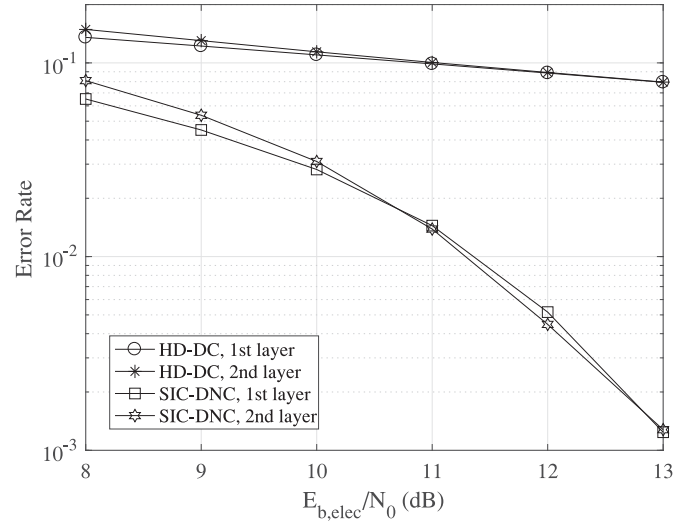


Fig. 10. Comparison of error rates between SIC-DNC and HD-DC.

and [16], it can be seen that five iterations are sufficient for both receivers to converge. From Fig. 8, we can see that they achieve similar BERs, with SIC-INC slightly outperforming for two-layer ACO-OFDM and the other for four layers.

#### D. BER of SIC Based Direct Noise Clipping

Fig. 9 shows the BERs achieved by the SIC-DNC scheme, where 4-QAM symbols are loaded onto two- and four-layer ACO-OFDM signals. BERs of the SIC-INC with five iterations, and the two diversity combining receivers [15], [17] are also plotted for comparison. The combining coefficient used for the SIC-DC receiver [15] is configured at 0.75 which is close to the optimum, whereas that for the hard decision based diversity combining receiver (HD-DC) [17], 0.5. From this figure, we can see that the SIC-DNC scheme achieves the lowest BERs at all the considered values for  $E_{b,elec}/N_0$ . It constantly outperforms the other three receiving schemes by up to 0.8 dB at the BER of  $10^{-4}$ . We attribute this to the distribution of the noise in the

recovered samples,  $\bar{s}_l$ , and the accurate identification of zero samples using the estimates  $\hat{x}_l$ .

First, consider SIC-INC and SIC-DNC, where the estimated samples,  $\hat{x}_l$ , are added to the recovered clipping noise samples,  $\bar{c}_l$ , in SIC-INC, whereas the recovered samples,  $\bar{y}_l$  are used in SIC-DNC. Although from the perspective of detection, the estimated and recovered samples,  $\hat{x}_l$  and  $\bar{y}_l$  lead to identical BERs because of the property of MMSE [32], the noise components in the two sets of samples have different power. Using the central limit theorem, the noise in both sets of samples and that in  $\bar{c}_l$  are all Gaussian distributed. Denote the noise power for the estimated samples, the recovered samples and the recovered clipping noise samples by  $P_1$ ,  $P_2$  and  $P_3$  respectively. Then we can see that  $P_1 < P_3$  and  $P_1 < P_2$ , due to the fact that  $\hat{x}_l$  are estimated samples and the other two, recovered yet estimated. The first inequality indicates that in SIC-INC, the removed noise samples will mostly come from the recovered clipping noise samples, whereas the samples,  $\hat{x}_l$ , which are used for decoding are only slightly corrected. The second inequality means that

TABLE IV  
COMPARISONS BETWEEN ADVANCED RECEIVERS

	Exact SIC	HD-DC [17]	SIC-INC	HD-INC [16]	SIC-DC [15]	SIC-DNC
Inputs	–	–	$\hat{\mathbf{x}}_l, \bar{\mathbf{c}}_l$	–	$\hat{\mathbf{x}}_l, \bar{\mathbf{c}}_l, \alpha_{\text{SIC-DC}}$	$\hat{\mathbf{x}}_l, \bar{\mathbf{y}}_l, \bar{\mathbf{c}}_l$
$E_{b,\text{elec}}/N_0$ (dB), 2-layer	12.53	10.91 (–1.62)	10.67 (–1.86)	10.80 (–1.73)	10.53 (–2)	10.10 (–2.43)
$E_{b,\text{elec}}/N_0$ (dB), 3-layer	13.34	11.82 (–1.52)	11.61 (–1.73)	11.61 (–1.73)	11.46 (–1.88)	11.10 (–2.24)
$E_{b,\text{elec}}/N_0$ (dB), 4-layer	13.97	12.50 (–1.47)	12.32 (–1.65)	12.23 (–1.74)	12.16 (–1.81)	11.83 (–2.14)

in SIC-DNC, the noise power,  $P_2$  is not as dominating as in SIC-INC. Therefore, more pure noise samples can be removed from the recovered samples,  $\bar{\mathbf{y}}_l$ . This leads to improved BERs for the SIC-DNC scheme.

Second, consider SIC-DNC and HD-DC [17]. We can see that the most important difference from the two schemes is the decision making processes that determine the positions of the non-negative samples. For HD-DC, this is implemented using the signs of the samples (IFFT of  $\bar{\mathbf{Y}}_l^h$  shown in Fig. 1(b)) which are perturbed by additive Gaussian noise, whereas for SIC-DNC, using the estimates,  $\hat{\mathbf{x}}_l$ . Fig. 10 shows the error rate of the process for double-layer ACO-OFDM, where the error rate is defined as the number of wrong decisions over the number of samples,  $N$ . From this figure, we can see that as expected, the SIC-DNC scheme leads to much lower error rate than HD-DC. This leads to the superior BERs of SIC-DNC.

#### E. Comparisons Between Advanced Receivers

In this section, we present comparisons between the advanced receivers including HD-DC [17], HD-INC with five iterations [16], SIC-DC with the optimum combining coefficient [15], SIC-INC with five iterations, and SIC-DNC. The comparisons are made in two aspects, 1) the inputs to the second stage of the SIC based receivers and 2) the required values for  $E_{b,\text{elec}}/N_0$  when a BER of  $10^{-4}$  is to be achieved for two-, three- and four-layer ACO-OFDM loaded by 4-QAM symbols. The results are shown in Table IV.

From a system reliability point of view, the inputs required by the second stage can typically reflect the robustness of an advanced receiver. This is because the receiver completely depends on the inputs to operate and it fails when any input becomes unavailable due to unknown failures. As a result, we can see that SIC-INC has the highest system reliability compared with SIC-DC and SIC-DNC. As shown in the table, SIC-INC can be implemented when only the estimated samples,  $\hat{\mathbf{x}}_l$ , and the recovered clipping noise,  $\bar{\mathbf{c}}_l$ , are available. This is in contrast with the other two SIC based schemes, SIC-DNC and SIC-DC, which demand more information from the SIC module. From the table it can be seen that SIC-DNC requires the recovered samples,  $\bar{\mathbf{y}}_l$ , in addition to  $\hat{\mathbf{x}}_l$  and  $\bar{\mathbf{c}}_l$ , and that SIC-DC needs the optimum combining coefficient,  $\alpha_{\text{SIC-DC}}$ , as an input. In the context of SIC, the determination of  $\alpha_{\text{SIC-DC}}$  is an open problem which can be solved only by using the inefficient exhaustive search<sup>7</sup> [15]. The use of SIC-INC and SIC-DNC can

effectively bypass this open problem. This is an advantage of SIC-INC and SIC-DNC over SIC-DC.

The values for  $E_{b,\text{elec}}/N_0$  are presented in the last three rows of the table where the numbers in decibel included in the brackets denote the improvement relative to the exact SIC. From Table IV, we can see that the best performance (the lowest value for  $E_{b,\text{elec}}/N_0$ ) is consistently achieved by SIC-DNC followed by SIC-DC with optimum combining coefficient. HD-DC, on the contrary, results in the highest required value for all the considered ACO-OFDM signals. The gap between SIC-DNC and HD-DC is up to 0.8 dB, depending on the number of layers. We can also see that the two iterative receivers, SIC-INC and HD-INC, have similar performance. They constantly outperform HD-DC, but are inferior to SIC-DC for all the considered ACO-OFDM signals. From this table, it can be seen that, except for HD-INC, the improvement of BER provided by the advanced receivers decreases with increasing layers used in the signal. For example, a gain of 2.14 dB can be achieved by SIC-DNC for four-layer ACO-OFDM, whereas these gains are given by 2.43 dB and 2.24 dB for double- and triple-layer signals, respectively. The improvement provided by HD-INC, on the contrary, is relatively stable regardless of the number of layers used in the signal. It can be seen that the gain provided by HD-INC is approximately 1.73 dB for all the three ACO-OFDM signals.

## VI. CONCLUSION

In this paper, we have studied the MMSE estimation based SIC for layered ACO-OFDM and its application to noise clipping receivers. A two-stage receiver paradigm is proposed, in which the SIC is applied in the first stage, and noise clipping, in the second stage. For the first stage, we derive the exact formula for the MMSE estimators, taking into account the RI accumulated in the process of demodulation. We show that the RI is Gaussian distributed with zero mean and a variance related to the MSE of the estimation error. The resulting MSEs are evaluated for a range of constellations and compared with the existing simplified SIC. It is demonstrated that the exact SIC provides little improvement on MSE and BER over the simplified SIC, confirming that the simplified SIC is effective. For the second stage, we present two schemes for noise clipping using the outputs of the first stage, namely SIC-INC and SIC-DNC. These schemes can effectively bypass the challenge of determining the optimum combining coefficient when the diversity combining is applied, while providing comparable performance. We show that the SIC-INC scheme can achieve similar BERs to those of the existing diversity combining receivers using optimum com-

<sup>7</sup>The simulation result not shown in the paper shows that  $\alpha_{\text{SIC-DC}}$  is a function of  $E_{b,\text{elec}}/N_0$  and the size of the constellation,  $M$ . Therefore, the optimum combining coefficient can be evaluated through simulation for each possible value for  $E_{b,\text{elec}}/N_0$  and  $M$  before data transmission. These coeffi-

cients are then stored in the memory, and used to configure the receiver according to  $E_{b,\text{elec}}/N_0$  and  $M$  during data transmission.

binning coefficients. It is also shown that although the exact SIC cannot provide superior BER to its hard-decision based counterpart, it produces more accurate decisions on the pure noise samples than the hard-decision based diversity combining receivers can do, leading to the BER achieved by the SIC-DNC outperforming the existing advanced receivers by up to 0.8 dB at the BER of  $10^{-4}$ .

#### APPENDIX A

##### DERIVATION OF THE EXPECTED VALUE AND VARIANCE

The random variables,  $x_{i,k}$  and  $\hat{x}_{i,k}$  are Gaussian distributed with zero mean, i.e.,  $x_{i,k} \sim \mathcal{N}(0, \sigma_i^2)$  and  $\hat{x}_{i,k} \sim \mathcal{N}(0, \hat{\sigma}_i^2)$ . Thus, the expected values of  $|x_{i,k}|$  and  $|\hat{x}_{i,k}|$  can be expressed as [35]

$$E(|x_{i,k}|) = \frac{2}{\sqrt{2\pi}}\sigma_i \quad (41)$$

and

$$E(|\hat{x}_{i,k}|) = \frac{2}{\sqrt{2\pi}}\hat{\sigma}_i, \quad (42)$$

respectively, leading to the expectation of  $(|x_{i,k}| - |\hat{x}_{i,k}|)$  given by (16).

The correlation coefficient,  $\rho$ , can be expressed as

$$\begin{aligned} \rho &= \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n_1 \in \mathcal{G}_i} \sum_{n_2 \in \mathcal{G}_i} E \left( X_{i,n_1} \hat{X}_{i,n_2}^* \right) e^{j \frac{2\pi}{N} k(n_1 - n_2)} \right\} \\ &\stackrel{(a)}{=} \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n \in \mathcal{G}_i} \left[ E \left( X_{i,n} \hat{X}_{i,n}^* \right) \right. \right. \\ &\quad \left. \left. + (-1)^k E \left( X_{i,n} \hat{X}_{i,(n+N/2)_N}^* \right) \right] \right\} \\ &\stackrel{(b)}{=} \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n \in \mathcal{G}_i} \left[ E \left( X_{i,n} \hat{X}_{i,n}^* \right) + (-1)^k E \left( X_{i,n} \hat{X}_{i,n} \right) \right] \right\} \\ &= \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n \in \mathcal{G}_i} \left[ E \left( X_{i,n} \hat{X}_{i,n}^* \right) + (-1)^k \right. \right. \\ &\quad \left. \left. \left( E \left( a_{i,n} \hat{a}_{i,n} - b_{i,n} \hat{b}_{i,n} \right) + j E \left( a_{i,n} \hat{b}_{i,n} + \hat{a}_{i,n} b_{i,n} \right) \right) \right] \right\} \\ &\stackrel{(c)}{=} \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n \in \mathcal{G}_i} E \left( X_{i,n} \hat{X}_{i,n}^* \right) \right\} \\ &\stackrel{(d)}{=} \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ \frac{1}{N} \sum_{n \in \mathcal{G}_i} E \left( |\hat{X}_{i,n}|^2 \right) \right\} \\ &\stackrel{(e)}{=} \frac{1}{\sigma_i \hat{\sigma}_i} \left\{ E \left( \hat{x}_{i,n}^2 \right) \right\} \\ &= \frac{\hat{\sigma}_i}{\sigma_i}, \quad (43) \end{aligned}$$

where  $(\cdot)_N$  denotes modulo  $N$ . Equation (a) holds because the symbols loaded are independent with zero mean and Hermitian

symmetry. Equation (b) holds because of the Hermitian symmetry. Equation (c) holds for the symmetrical constellations. Using the symmetry, the in-phase and quadrature components,  $a_{i,n}$  and  $b_{i,n}$ , and their MMSE estimates,  $\hat{a}_{i,n}$  and  $\hat{b}_{i,n}$  have the following relationships, 1)  $E(a_{i,n}) = E(b_{i,n}) = 0$ , 2)  $E(a_{i,n} \hat{a}_{i,n}) = E(b_{i,n} \hat{b}_{i,n})$ , 3)  $E(a_{i,n} b_{i,n}) = 0$ , 4)  $E(\hat{a}_{i,n} b_{i,n}) = E(a_{i,n} \hat{b}_{i,n}) = 0$ . Equation (d) holds due to the orthogonality principle of MMSE estimation, given by  $E\{(X_{i,n} - \hat{X}_{i,n})\hat{X}_{i,n}^*\} = 0$  [32], and equation (e), Parseval's theorem [31].

The variance,  $\text{Var}(|x_{i,k}| - |\hat{x}_{i,k}|)$ , is given by

$$\begin{aligned} &\text{Var}(|x_{i,k}| - |\hat{x}_{i,k}|) \\ &= E(|x_{i,k}| - |\hat{x}_{i,k}|)^2 - (E|x_{i,k}| - E|\hat{x}_{i,k}|)^2 \\ &= E(x_{i,k}^2) - 2E(|x_{i,k} \hat{x}_{i,k}|) + E(\hat{x}_{i,n}^2) - (E|x_{i,k}|)^2 \\ &\quad + 2E(|x_{i,k}|)E(|\hat{x}_{i,k}|) - (E|\hat{x}_{i,k}|)^2 \\ &= \left(1 - \frac{2}{\pi}\right)\sigma_i^2 + \left(1 - \frac{2}{\pi}\right)\hat{\sigma}_i^2 + \frac{4}{\pi}\sigma_i\hat{\sigma}_i - 2E(|x_{i,k} \hat{x}_{i,k}|). \quad (44) \end{aligned}$$

#### APPENDIX B

##### DERIVATION OF MMSE ESTIMATOR

We rewrite the information symbol, the interference-plus-noise and the observation at the MMSE estimator as

$$\begin{aligned} X_{l,n} &= \Re(X_{l,n}) + j\Im(X_{l,n}) \\ Z_{l,n} &= \Re(Z_{l,n}) + j\Im(Z_{l,n}) \\ \bar{Y}_{l,n} &= \Re(\bar{Y}_{l,n}) + j\Im(\bar{Y}_{l,n}), \quad (45) \end{aligned}$$

respectively. Then  $\Re(Z_{l,n})$  and  $\Im(Z_{l,n})$  are Gaussian random variables with zero-mean and identical variance, given by

$$E(\Re(Z_{l,n}))^2 = E(\Im(Z_{l,n}))^2 = \frac{1}{2}\sigma_{Z_l}^2. \quad (46)$$

We first study the real part. As the information symbols are drawn from  $M$ -QAM constellation, their real parts are pulse amplitude modulation (PAM) symbols with the PDF given by [32]

$$p_{\Re(X_{l,n})}(x_R) = \sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \delta(x_R - a) \quad (47)$$

where  $\delta(\cdot)$  denotes the Kronecker delta function. Therefore, because of the relationship,  $\Re(\bar{Y}_{l,n}) = \Re(X_{l,n})/2 + \Re(Z_{l,n})$  (see (18)), the PDF of  $\Re(\bar{Y}_{l,n})$  conditioned on the transmitted symbol can be expressed as

$$p_{\Re(\bar{Y}_{l,n})|\Re(X_{l,n})}(y_R|x_R) = \frac{1}{\sqrt{\pi\sigma_{Z_l}^2}} \exp\left[-\frac{1}{\sigma_{Z_l}^2}\left(y_R - \frac{1}{2}x_R\right)^2\right], \quad (48)$$

leading to the joint PDF of the transmitted symbol and the observation given by

$$\begin{aligned}
& p_{\Re(X_{l,n}), \Re(\bar{Y}_{l,n})}(x_R, y_R) \\
&= p_{\Re(\bar{Y}_{l,n}) | \Re(X_{l,n})}(y_R | x_R) p_{\Re(X_{l,n})}(x_R) \\
&= \frac{1}{\sqrt{\pi \sigma_{Z_l}^2}} \sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \\
&\quad \times \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}x_R\right)^2\right] \delta(x_R - a). \quad (49)
\end{aligned}$$

As the transmitted symbol is independent of the interference-plus-noise, the PDF of the real part of the observation is given by the convolution of their individual PDFs, i.e.,

$$\begin{aligned}
p_{\Re(\bar{Y}_{l,n})}(y_R) &= \frac{1}{\sqrt{\pi \sigma_{Z_l}^2}} \sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \\
&\quad \times \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}a\right)^2\right]. \quad (50)
\end{aligned}$$

Therefore, the PDF of the transmitted symbol conditioned on the observation can be expressed as

$$\begin{aligned}
& p_{\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})}(x_R | y_R) \\
&= \frac{p_{\Re(X_{l,n}), \Re(\bar{Y}_{l,n})}(x_R, y_R)}{p_{\Re(\bar{Y}_{l,n})}(y_R)} \\
&= \frac{\sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}x_R\right)^2\right] \delta(x_R - a)}{\sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}a\right)^2\right]}, \quad (51)
\end{aligned}$$

leading to the MMSE estimates, given by the conditional expected value, expressed as (26).

Using a similar approach as those presented in (47)–(51), the MMSE estimate of imaginary part can be derived as (27). In fact, because of the symmetry of the constellations, the estimator for the imaginary part can be derived by simply substitute  $a$ ,  $\mathcal{X}_R$  and  $\Re(\bar{Y}_{l,n})$  in (26) with  $b$ ,  $\mathcal{X}_I$  and  $\Im(\bar{Y}_{l,n})$ , respectively.

#### APPENDIX C DERIVATION OF MSE

Because of the symmetry of constellations, the MSE of the real part equals that of the imaginary part. Therefore, the overall MSE will be twice of that of the real part.

We first study the real part. The MSE conditioned on an observation is the variance of the conditional density,

$p_{\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})}(x_R | y_R)$  [31], i.e.,

$$\begin{aligned}
& \sigma_{\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})}^2 \\
&= E\left(\left(\Re(X_{l,n})\right)^2 | \Re(\bar{Y}_{l,n})\right) - \left(E\left(\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})\right)\right)^2 \\
&= \frac{\left(\sum_{a \in \mathcal{X}_R} a^2 \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}a\right)^2\right]\right)}{\left(\sum_{a \in \mathcal{X}_R} \Pr[\Re(X_{l,n}) = a] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_R - \frac{1}{2}a\right)^2\right]\right)} \\
&\quad - \hat{a}_{l,n}^2, \quad (52)
\end{aligned}$$

where the second equation is derived by using the PDF given by (51) and the MMSE estimator given by (26).

Using a similar approach, the conditional MSE for the imaginary part can be expressed as

$$\begin{aligned}
& \sigma_{\Im(X_{l,n}) | \Im(\bar{Y}_{l,n})}^2 \\
&= E\left(\left(\Im(X_{l,n})\right)^2 | \Im(\bar{Y}_{l,n})\right) - \left(E\left(\Im(X_{l,n}) | \Im(\bar{Y}_{l,n})\right)\right)^2 \\
&= \frac{\left(\sum_{b \in \mathcal{X}_I} b^2 \Pr[\Im(X_{l,n}) = b] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_I - \frac{1}{2}b\right)^2\right]\right)}{\left(\sum_{b \in \mathcal{X}_I} \Pr[\Im(X_{l,n}) = b] \exp\left[-\frac{1}{\sigma_{Z_l}^2} \left(y_I - \frac{1}{2}b\right)^2\right]\right)} \\
&\quad - \hat{b}_{l,n}^2. \quad (53)
\end{aligned}$$

Therefore, by averaging (52) and (53) with respect to  $\Re(\bar{Y}_{l,n})$  and  $\Im(\bar{Y}_{l,n})$ , respectively, the expected MSE can be expressed as

$$\begin{aligned}
E(|\Delta_{\mathbf{X}_I}|^2) &= \int_{-\infty}^{\infty} \sigma_{\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})}^2 p_{\Re(\bar{Y}_{l,n})}(y_R) dy_R \\
&\quad + \int_{-\infty}^{\infty} \sigma_{\Im(X_{l,n}) | \Im(\bar{Y}_{l,n})}^2 p_{\Im(\bar{Y}_{l,n})}(y_I) dy_I \\
&= 2 \int_{-\infty}^{\infty} \sigma_{\Re(X_{l,n}) | \Re(\bar{Y}_{l,n})}^2 p_{\Re(\bar{Y}_{l,n})}(y_R) dy_R, \quad (54)
\end{aligned}$$

which can be rewritten as (28).

#### REFERENCES

- [1] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi?," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1533–1544, Mar. 2016.
- [2] J. B. Carruthers and J. M. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 538–546, Apr. 1996.
- [3] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electron. Lett.*, vol. 42, pp. 370–372, Mar. 2006.
- [4] D. Tsonev, S. Sinanovic, and H. Haas, "Novel unipolar orthogonal frequency division multiplexing (U-OFDM) for optical wireless," in *Proc. IEEE Veh. Technol. Conf.*, Yokohama, Japan, 2012, pp. 1–5.
- [5] S. C. J. Lee, S. Randel, F. Breyer, and A. M. J. Koonen, "PAM-DMT for intensity-modulated and direct-detection optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 23, pp. 1749–1751, Dec. 2009.
- [6] S. D. Dissanayake and J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD Systems," *J. Lightw. Technol.*, vol. 31, no. 7, pp. 1063–1072, Apr. 2013.
- [7] B. Ranjha and M. Kavehrad, "Hybrid asymmetrically clipped OFDM-Based IM/DD optical wireless system," *J. Opt. Commun. Netw.*, vol. 6, no. 4, pp. 387–396, Apr. 2014.

- [8] Q. Wang, C. Qian, X. Guo, Z. Wang, D. G. Cunningham, and I. H. White, "Layered ACO-OFDM for intensity-modulated direct-detection optical wireless transmission," *Opt. Express*, vol. 23, pp. 12382–12393, May 2015.
- [9] M. S. Islam, D. Tsonev, and H. Haas, "On the superposition modulation for OFDM-based optical wireless communication," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Orlando, FL, USA, 2015, pp. 1022–1026.
- [10] D. Tsonev, S. Videv, and H. Haas, "Unlocking spectral efficiency in intensity modulation and direct detection systems," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1758–1770, Sep. 2015.
- [11] M. S. Islam and H. Haas, "Augmenting the spectral efficiency of enhanced PAM-DMT-based optical wireless communications," *Opt. Express*, vol. 24, pp. 11932–11949, May 2016.
- [12] H. Elgala and T. D. C. Little, "SEE-OFDM: Spectral and energy efficient OFDM for optical IM/DD systems," in *Proc. IEEE 25th Annu. Int. Symp. Personal, Indoor, Mobile Radio Commun.*, Washington, DC, USA, 2014, pp. 851–855.
- [13] L. Chen, B. Krongold, and J. Evans, "Successive decoding of anti-periodic OFDM signals in IM/DD optical channel," in *Proc. Int. Conf. Commun.*, Cape Town, South Africa, 2010, pp. 1–6.
- [14] A. J. Lowery, "Comparisons of spectrally-enhanced asymmetrically-clipped optical OFDM systems," *Opt. Express*, vol. 24, pp. 3950–3966, Feb. 2016.
- [15] T. Q. Wang, H. Li, and X. Huang, "Diversity combining for layered asymmetrically clipped optical OFDM using soft successive interference cancellation," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1309–1312, Jun. 2017.
- [16] Q. Wang, Z. Wang, X. Guo, and L. Dai, "Improved receiver design for layered ACO-OFDM in optical wireless communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 3, pp. 319–322, Feb. 2016.
- [17] M. M. A. Mohammed, C. He, and J. Armstrong, "Diversity combining in layered asymmetrically clipped optical OFDM," *J. Lightw. Technol.*, vol. 35, no. 11, pp. 2078–2085, Jun. 2017.
- [18] L. Chen, B. Krongold, and J. Evans, "Diversity combining for asymmetrically clipped optical OFDM in IM/DD channels," in *Proc. IEEE Globecom*, Honolulu, HI, USA, 2009, pp. 1–6.
- [19] M. Tahar, T. Q. Wang, M. F. G. Medina, O. Gonzalez, and J. Armstrong, "Experimental demonstration of diversity combining for asymmetrically clipped optical OFDM," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 906–909, May 2016.
- [20] J. Dang, Z. Zhang, and L. Wu, "Frequency-domain diversity combining receiver for ACO-OFDM system," *IEEE Photon. J.*, vol. 7, no. 6, Dec. 2015, Art. no. 7802510.
- [21] K. Asadzadeh, A. Dabbo, and S. Hranilovic, "Receiver design for asymmetrically clipped optical OFDM," in *Proc. Globecom Workshops 2011*, Houston, TX, USA, 2011, pp. 777–781.
- [22] Z. Chen, D. Tsonev, and H. Haas, "Improved receivers for asymmetrically-clipped optical OFDM," in *Proc. IEEE Veh. Technol. Conf.*, Seoul, South Korea, 2014, pp. 1–5.
- [23] N. Fernando, Y. Hong, and E. Viterbo, "Flip-OFDM for unipolar communication systems," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3726–3733, Dec. 2012.
- [24] N. Huang, J. -B. Wang, C. Pan, J. -Y. Wang, Y. Pan, and M. Chen, "Iterative receiver for Flip-OFDM in optical wireless communication," *IEEE Photon. Technol. Lett.*, vol. 27, no. 16, pp. 1729–1732, Aug. 2015.
- [25] Q. Wang, Z. Wang, and L. Dai, "Iterative receiver for hybrid asymmetrically clipped optical OFDM," *J. Lightw. Technol.*, vol. 32, no. 22, pp. 4471–4477, Nov. 2014.
- [26] D. Tsonev, S. Sinanovic, and H. Haas, "Complete modeling of nonlinear distortion in OFDM-based optical wireless communication," *J. Lightw. Technol.*, vol. 31, no. 18, pp. 3064–3076, Sep. 2013.
- [27] L. Chen, B. Krongold, and J. Evans, "Performance analysis for optical OFDM transmission in short-range IM/DD systems," *J. Lightw. Technol.*, vol. 30, no. 7, pp. 974–983, Apr. 2012.
- [28] S. Hranilovic, *Wireless Optical Communication Systems*. New York, NY, USA: Springer, 2004.
- [29] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [30] P. T. P. Tang, "Table-Lookup algorithms for elementary functions and their error analysis," in *Proc. 10th IEEE Symp. Comput. Arithmetic*, Grenoble, France, pp. 232–236, Jun. 1991.
- [31] A. V. Oppenheim and G. C. Verghese, *Class Notes for Introduction to Communication, Control and Signal Processing*. Cambridge, MA, USA: MIT Press, 2010.
- [32] A. H. Sayed, *Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2008.
- [33] S. R. H. Shojaie, M. Aminghafari, and A. Mohammadpour, "On the expected absolute value of a bivariate normal distribution," *J. Statist. Theory Appl.*, vol. 11, pp. 371–377, 2012.
- [34] D. Tsonev *et al.*, "A 3-Gb/s single-LED OFDM-Based wireless VLC link using a gallium nitride  $\mu$  LED," *IEEE Photon. Technol. Lett.*, vol. 26, no. 7, pp. 637–640, Apr. 2014.
- [35] S. Dimitrov, S. Sinanovic, and H. Haas, "Clipping noise in OFDM-Based optical wireless communication systems," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1072–1081, Apr. 2012.
- [36] T. Q. Wang and X. Huang, "Dimming compatible optical OFDM for high speed visible light communications," in *Proc. IEEE 85th Veh. Technol. Conf.*, Sydney, NSW, Australia, 2017.
- [37] X. Li, J. Vucic, V. Jungnickel, and J. Armstrong, "On the capacity of intensity-modulated direct-detection systems and the information rate of ACO-OFDM for indoor optical wireless applications," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 799–809, Mar. 2012.

**Thomas Q. Wang** received B.E. degree in electrical engineering from Dalian Jiaotong University, Dalian, China, in 2006, and the M.E. and Ph.D. degrees in communication and information systems from Dalian Maritime University, Dalian, China, in 2008 and 2011, respectively. From March 2012 to December 2015, he was a Research Fellow with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, VIC, Australia. He is currently a Research Fellow with the Global Big Data Technologies Centre, University of Technology Sydney, Sydney, NSW, Australia. His research interests include optical wireless communications and multiple-input multiple-output technology.

**Hang Li** received the B.Eng. and M.Eng. degrees from Beijing Jiaotong University, Beijing, China, in 2003 and 2006, respectively, and the Ph.D. degree from the University of Western Australia, Perth, WA, Australia, in 2014, all in electronic engineering. He is currently a Postdoctoral Research Fellow with the Global Big Data Technologies Centre, University of Technology Sydney. His research interests include wireless and optical communications signal processing, wireless cross-layer protocols design, and mmWave Massive MIMO.

**Xiaojing Huang** (M'99–SM'11) received the B.E., M.E., and Ph.D. degrees from Shanghai Jiao Tong University, Shanghai, China, in 1983, 1986, and 1989, respectively, all in electronic engineering. He is currently a Professor of Information and Communications Technology with the School of Electrical and Data Engineering and the Program Leader for mobile sensing and communications with the Global Big Data Technologies Centre, University of Technology Sydney, Australia. He had been a Principal Research Scientist with Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, and the Project Leader of CSIRO microwave and mm-wave backhaul projects since 2009, an Associate Professor at the University of Wollongong, Australia, since 2004, and a Principal Research Engineer with Motorola Australian Research Centre since 1998. With over 29 years of combined industrial, academic, and scientific research experience, he has published more than 270 book chapters, refereed journal and conference papers, and major commercial research reports. He has also filed 31 patents. He was a recipient of the CSIRO Chairman's Medal and the Australian Engineering Innovation Award in 2012 for exceptional research achievements in multigigabit wireless communications.