Piecewise Companding Transform Assisted Optical-OFDM Systems for Indoor Visible Light Communications

Hongming Zhang, *Student Member, IEEE*, Lie-Liang Yang, *Fellow, IEEE* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—In visible light communications (VLCs) relying on intensity-modulation and direct detection (IM/DD), the conversion from electrical signals to optical signals and the limited dynamic range of the light-emitting diodes (LEDs) constitute the fundamental impediments in the way of high-integrity communications, especially when orthogonal frequency-division multiplexing (OFDM) is employed. In IM/DD VLCs, only real-valued positive signals are used for signal transmission. However, the Fourier transform of OFDM systems is operated in the complex-domain. In order to meet the requirements of the IM/DD VLCs, the complex-to-real conversion is achieved at the cost of reducing the bandwidth efficiency. Moreover, OFDM signals experience a high peak-to-average power ratio (PAPR), hence typically clipping is used for confining the positive-valued signals within the LED's dynamic range. However, hard clipping leads to the loss of orthogonality for O-OFDM signals, generating inter-carrier interference (ICI). As a result, the performance of the clipping-based O-OFDM (CO-OFDM) systems may be severely degraded. In this paper, the concept of piecewise companding transform (CT) is introduced into the O-OFDM system advocated, forming the CTO-OFDM arrangement. We first investigate the general principles and design criteria of the piecewise CTO-OFDM. Based on our studies, three types of piecewise companders, namely the constant probability sub-distribution function (CPsDF), linear PsDF (LPsDF) and the nonlinear PsDF (NLPsDF) based CT are designed. Furthermore, we investigate the nonlinear effect of hard clipping and of our CT on O-OFDM systems in the context of different scenarios by both analytical and simulation techniques. Our investigations show that the CTO-OFDM constitutes a promising signalling scheme conceived for VLCs, which exhibits a high bandwidth efficiency, high flexibility, high reliability, as well as a high data-rate, despite experiencing nonlinear distortions.**

H. Zhang, L.-L. Yang and Lajos Hanzo are with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ, UK. (E-mail: zh1g11 lly lh@ecs.soton.ac.uk).

The financial support of the EPSRC projects EP/Noo4558/1 and EP/L018659/1, as well as of the European Research Council's Advanced Fellow Grant under the Beam-Me-Up project and of the Royal Society's Wolfson Research Merit Award is gratefully acknowledged.

NOTATIONS

$(\cdot)^T$ and $(\cdot)^*$	the transpose and the conjugate, respectively
$(\cdot)^H$ and $(\cdot)^{-1}$	the conjugate transpose and the inverse, respectively
$\mathbb{E}[\cdot]$	the expectation operator
$\mathbb{Z}_+^N$	a set of real integers containing {1,2,...,N}
$\mathbb{C}^{M\times L}$	the set of $(M \times L)$-element matrices in the complex field
$\boldsymbol{x}_{\mathrm{T}}$ and $\boldsymbol{x}_{\mathrm{F}}$	time-domain (TD) and frequency-domain (FD) vectors in boldface lower-case, respectively
$x_{\mathrm{T}}(l)$ and $x_{\mathrm{F}}(m)$	lth TD symbol and mth FD symbol, respectively
X and x	TD random variables and their realizations in upper-case and lower-case letters, respectively
X and x	FD random variables and their realizations in upper-case and lower-case letters, respectively
$diag\{\boldsymbol{x}\}$	the diagonal matrix with elements in $\boldsymbol{x}$ on its diagonal

I. Introduction

Visible light communication (VLC) has been recognized as a compelling technique of alleviating teletraffic congestion in the near future. The main advantage of VLCs is that a wide unlicensed visible light spectrum can be used for data transmission without inflicting interference upon the existing radio frequency (RF) communications [1]. Furthermore, since partitioning is impervious to visible light, a unity spatial reuse factor can be attained by VLCs operating in different rooms. Owing to their energy efficiency, longevity as well as reliability, light-emitting diodes (LEDs) are expected to be the dominant light sources, which can also be used for data transmission in VLCs. For this reason, the installation cost of VLCs is rather low, especially when power line communication is used as the backhaul technology for VLCs [2]. There are many applications for VLCs, such as RF/ VLC cooperation in the cellular network, device-to-device communications, underwater communications, as well as some others as described in [3, 4].

However, there are still a lot of open challenges concerning the implementation of high-performance VLC systems. Due to its low-cost and simplicity, a noncoherent scheme, which is known as intensity modulation and direct detection (IM/ DD), is commonly used in VLCs. In IM/DD VLC systems, only the real and positive signals can be used for modulating the intensity of optical sources, where the intensity envelope is detected by the optical receiver. For high data rate transmission of IM/DD signals over non-line-of-sight (NLOS) channels, although the multipath induced fading may be mitigated by exploiting the fact that the dimensions of optical detectors are usually much larger than the wavelength of visible light, the inter-symbol interference (ISI) caused by multipath propagation may still degrade the overall system performance. Thus, orthogonal frequency-division multiplexing (OFDM), which is capable of combating the multipath effects in radio communications [5, 6], has also become an attractive candidate for VLCs [7–9].

In the classic OFDM schemes, the inverse discrete Fourier transform (IDFT) and the DFT are applied for the implementation of the discrete multicarrier modulation and demodulation, respectively. Owing to the employment of the fast Fourier transform (FFT) and cyclic prefix, the DFT-based OFDM scheme is capable of mitigating ISI with the aid of a low-complexity implementation [10], when communicating over dispersive frequency-selective fading channels. However, the classic OFDM scheme cannot be directly employed for operation in dispersive IM/DD optical systems, since the time-domain (TD) symbols are complex-valued owing to modulation by FFT-based techniques. By contrast, only unipolar signals can be transmitted in the IM/DD systems [11]. In order to render the classic OFDM scheme suitable for the IM/DD systems, a range of optical OFDM (O-OFDM) solutions have been proposed [12–14]. In these approaches, the Hermitian symmetry is imposed on the frequency-domain (FD) subcarriers, in order to obtain real-valued baseband signals. Then, the real-valued baseband signals are converted to unipolar signals, for example, by adding a DC bias [12] (DCO-OFDM), by asymmetric clipping [13] (ACO-OFDM), as well as by joint asymmetric clipping and DC bias [14] (ADO-OFDM), etc. However, the major drawback of these approaches is their low bandwidth efficiency imposed by the Hermitian symmetry.

Nevertheless, due to the simplicity of implementation, the clipping-based optical OFDM (CO-OFDM) solutions, which include the DCO-OFDM and ACO-OFDM systems, have attracted considerable research attention in VLCs. As the investigations in [13–15] show, the ACO-OFDM has a higher optical power efficiency than the DCO-OFDM, when a low or moderate constellation size is used [13]. For a large constellation size, the optical power efficiency of the DCO-OFDM is higher than that of ACO-OFDM [15]. The reason for the above effects is that for a small or moderate-sized constellation, an insufficient DC bias degrades the system performance. By contrast, for a large constellation, the inefficient usage of bandwidth becomes the dominant limitation of the achievable performance [14, 15]. Furthermore, as shown in [16] that for a fixed symbol rate, the computational complexity of DCO-OFDM is lower than that of ACO-OFDM. However, we should emphasize that the above-mentioned observations are obtained based on the assumption that the upper clipping bound is infinite. In practice, however, the illumination level is limited for eye safety [12, 17]. Meanwhile, the electrical power consumption should be controlled by considering the efficiency of both the digital-to-analog conversion (DAC) and the power amplifier [18]. Thus, the dynamic range of O-OFDM systems is limited [19]. Furthermore, due to the anti-symmetric property [13], clipping the negative parts of ACO-OFDM signals does not introduce any loss of information, whilst clipping of the high positive peaks of ACO-OFDM signals results in nonlinear distortion, which degrades the system performance.

Another fundamental issue in OFDM is the high peak-to-average power ratio (PAPR) problem, making OFDM signals sensitive to nonlinear devices. Nonlinear distortion leads to both in-band distortion and out-of-band radiation, both of which degrade the system performance. In the literature, numerous PAPR reduction techniques have been proposed, see [20–22] and the references therein. Among these techniques, a simple yet effective method is the employment of a companding transform (CT) [23–31], which reduces the PAPR by applying a compander (a "compressor" and "expander" pair) to the time-domain signals, before their transmissions. A compander compresses the high signal peaks, while expanding the low-magnitude signals, thereby striking an attractive tradeoff between PAPR reduction and the bit error ratio (BER) performance achieved. Specifically, in [23], the so-called μ-law compander originally used in speech coding has been introduced for PAPR reduction. Later in [25], the design criteria of the CT has been derived for striking a tradeoff between the PAPR reduction and BER performance attained. Substantial research attention has been dedicated to transforming the amplitude distribution of the original OFDM signals to particular distributions, such as the uniform distribution [26], the trapezoidal distribution [27] and so forth. Note that in these techniques, the average power has not been changed by the CT. Most recently, the design of efficient

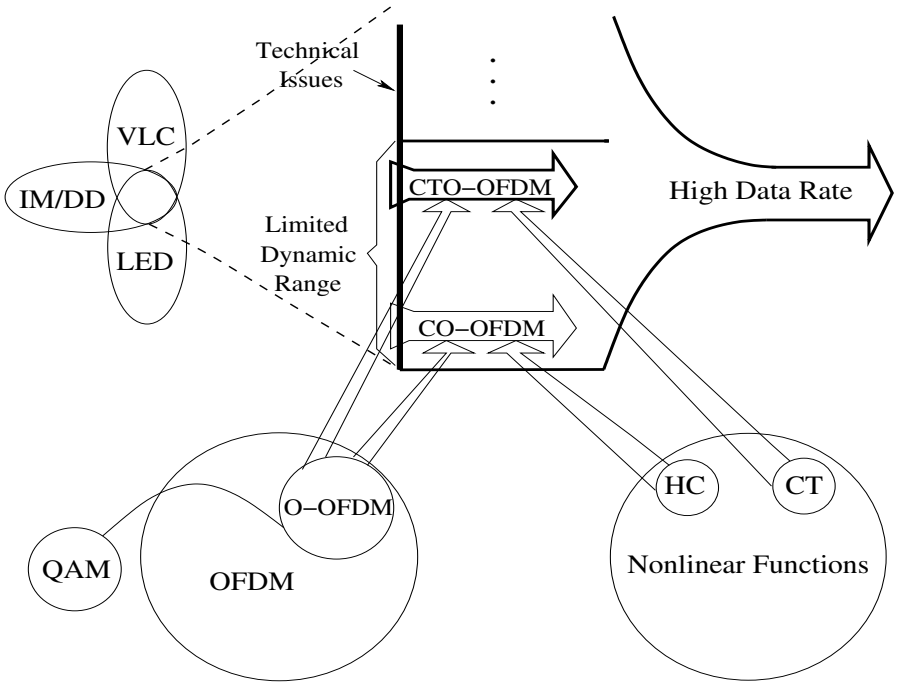


Fig. 1: Illustration of the relationships between optical OFDM (O-OFDM), clipping-based O-OFDM (CO-OFDM relying on hard clipping (HC)), as well as companding transform (CT) assisted O-OFDM (CTO-OFDM).

companders based on piecewise modification of the signal's amplitude distribution has been considered [28–31]. In these so-called piecewise companders, the amplitude distribution of original signals is only partially changed, whilst ensuring that both the PAPR reduction and BER performance can be improved.

Against the above background, in this paper, a symmetric piecewise CT is conceived for O-OFDM systems. The underlying concept is illustrated in Fig. 1. The general principles and design criteria of the piecewise CT schemes are studied in the context of VLC. Accordingly, three types of piecewise companding schemes, namely the constant probability sub-distribution function (CPsDF), the linear PsDF (LPsDF) and the nonlinear PsDF (NLPsDF) based CT are designed for the O-OFDM systems. In our CT-based optical OFDM (CTO-OFDM) systems, the CT is invoked for mitigating the clipping of high-peak TD signals. Thus, the system's performance can be improved, especially, when high-order QAM is used, which is sensitive to nonlinear distortions. In this paper, we investigate both the hard-clipping and the CT induced nonlinear effects imposed on O-OFDM systems relying on both analytical and on simulation approaches. Our studies show that the both the CTO-OFDM and DCO-OFDM systems are capable of attaining a higher bandwidth efficiency, than the ACO-OFDM system. Moreover, we demonstrate that the CTO-OFDM system outperforms the DCO-OFDM system both in terms of its reliability and flexibility, when supporting high-rate transmission experiencing nonlinear distortions.

The rest of the paper is organized as follows. In Section II, we describe the system model. The operating principle, design criteria and implementation of the symmetric piecewise CT are detailed in Section III. In Section IV, the nonlinear effects are analyzed. In Section V, our performance results are studied. Finally, we offer our conclusions in Section VI.

II. System Model

A. Description of the Transmitter

Fig. 2 illustrates a generalized O-OFDM scheme invoked in IM/DD optical communication systems, which will be discussed below. Let us assume that an $L_{\rm b}$-length sequence of independent and identically distributed (i.i.d.) bits is transmitted, which is first mapped to $N_{\rm d}$ data symbols, according to a Q-ary QAM constellation relying on the alphabet $\mathcal{A} = \{a(1), \ldots, a(Q)\}$. Hence, we have the relationship of $L_{\rm b} = N_{\rm d}\log_2(Q)$. Let us denote the $N_{\rm d}$ data symbols by $\boldsymbol{x}_{\rm d} = [x_{\rm d}(0), \ldots, x_{\rm d}(N_{\rm d}-1)]^T$, where we have $x_{\rm d}(i) \in \mathcal{A}$ and $\mathbb{E}[|x_{\rm d}(i)|^2] = 1$. As shown in Fig. 2, the data symbols in $\boldsymbol{x}_{\rm d}$ are arranged to form M FD symbols denoted as $\boldsymbol{x}_{\rm F} = [x_{\rm F}(0), \ldots, x_{\rm F}(M-1)]^T$, where the elements satisfy the Hermitian symmetry formulated as $x_{\rm F}(0) = x_{\rm F}(M/2) = 0$ and $x_{\rm F}(M-m) = [x_{\rm F}(m)]^*$ for any $m \neq 0$. Furthermore, for both the DCO-OF DM and CTO-OFDM systems, we have the mappings of $x_{\rm F}(m) = x_{\rm d}(m-1)$ for $m = 1, 2, \ldots, (M-2)/2$, yielding $N_{\rm d} = (M-2)/2$. By contrast, for the ACO-OFDM system, the data symbols in $\boldsymbol{x}_{\rm d}$ are successively mapped to the odd-indexed subcarriers of the first half of $\boldsymbol{x}_{\rm F}$, while all the even indexed subcarriers of $\boldsymbol{x}_{\rm F}$ are set to zero, which gives $N_{\rm d} = M/4$. From the above description, we infer that the bandwidth efficiency of both the DCO-OFDM and CTO-OFDM systems is higher than that of the ACO-OFDM system.

As shown in Fig. 2, the FD symbols are entered into an M-point IFFT, yielding the TD symbols denoted as $\boldsymbol{x}_{\rm T} = [x_{\rm T}(0), \ldots, x_{\rm T}(M-1)]^T \in \mathbb{R}^{M\times 1}$, where the lth TD symbol

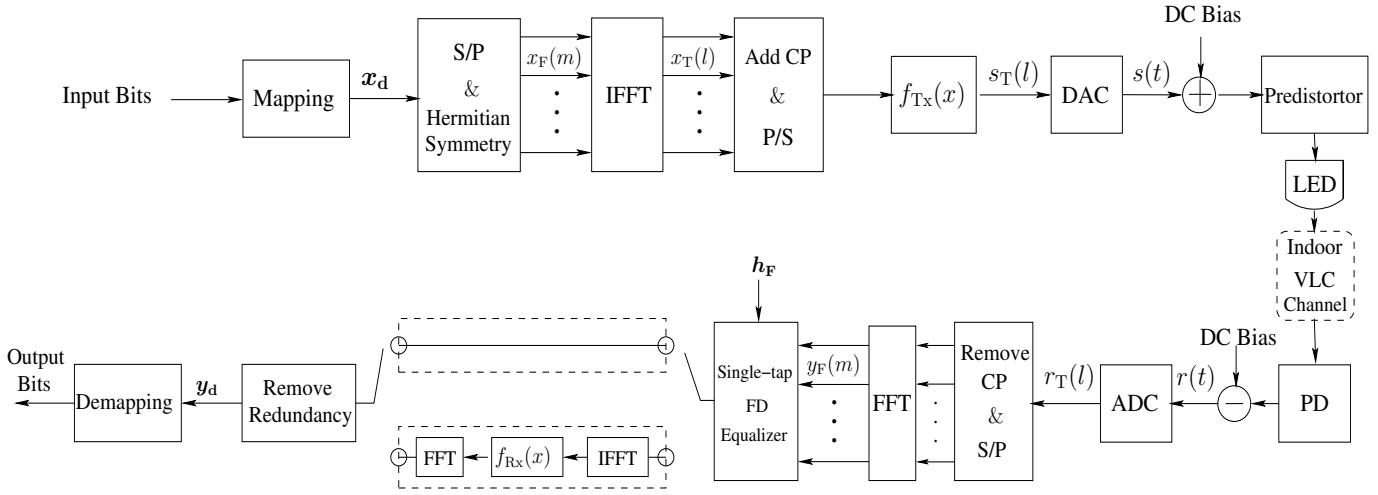


Fig. 2: Illustration of the O-OFDM systems.

can be expressed as

$$x_{\mathrm{T}}(l) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} x_{\mathrm{F}}(m) \exp\left(j2\pi \frac{m}{M} l\right) \tag{1}$$

for $l = 0, 1, \ldots M-1$. It is widely recognized that when M is sufficiently large, the TD symbols in $\boldsymbol{x}_{\mathrm{T}}$ can be approximated by Gaussian random variables with zero mean and a variance of $\sigma_x^2 = \mathbb{E}[|x_{\mathrm{T}}(l)|^2]$. As shown in Fig. 2, after adding a sufficiently long cyclic prefix (CP) and following parallel-to-serial conversion, the TD symbols in (1) are subjected to the specifically designed function $\{f_{\mathrm{Tx}}(x) : \mathbb{R} \to \mathbb{R}\}$. Explicitly, at the transmitter side, $f_{\mathrm{Tx}}(x)$ is used for guaranteeing that the information-bearing signals after adding a DC bias become real and positive. Let $s_{\mathrm{T}}(l) = f_{\mathrm{Tx}}[x_{\mathrm{T}}(l)]$ represent the output of the transmitter processor shown in Fig. 2, when $x_{\mathrm{T}}(l)$ expressed in (1) is the corresponding input. Then, as seen in Fig. 2, the outputs of the transmitter processor are fed into a DAC, where a DC bias associated with a level of A_{DC} is added in order to finally output the electric current signal denoted as $(s(t) + A_{\mathrm{DC}}) \in \mathbb{R}_+$. It should be noted that the inherent nonlinearity of LEDs often results in nonlinear distortion of the transmitted signals. In order to compensate for the nonlinearity of the LEDs, the electric current signal $s(t)$ can be pre-distorted [32]. An alternative technique of compensating for the nonlinearity is to employ a post-distorter [33] at the receiver. Finally, as seen in Fig. 2, the pre-distorted signals are input to the LED to modulate the intensity of the LED's light.

B. VLC Channels

Let us assume that the LEDs have the optical conversion factor of ρ_{LED} in Watt-per-ampere (W/A) and a photodetector (PD) with the responsivity of ρ_{PD} in ampere-per-Watt (A/W). Without loss of generality, let us assume that $\rho_{\mathrm{PD}}\rho_{\mathrm{LED}} = 1$ in order to simplify our analysis. Moreover, we assume that perfect pre-distortion or post-distortion is implemented at the transmitter or receiver, respectively. Under these assumptions, the IM/DD system can be assumed to be operated in the linear dynamic range, as shown in Fig. 3. Typically, in the IM/DD systems, information is conveyed by the optical intensity in response to the input electrical current, while the receiver outputs the electrical current proportional to the received optical intensity. In this paper, we assume that the optical intensity signals propagate over an indoor multipath VLC channel. In general, the variation of indoor VLC channels versus time is very slow. Hence, in our studies, we assume that the indoor VLC channels are time-invariant. Consequently, as seen in Fig. 2, after removing the DC bias, the electrical photocurrent signal obtained at the receiver can be expressed as

$$\begin{aligned} r(t) &= \int_{-\infty}^{\infty} \rho_{\mathrm{PD}}\rho_{\mathrm{LED}} h(t) s(t-\tau) \mathrm{d}\tau + n(t) \\ &= \int_{-\infty}^{\infty} h(t) s(t-\tau) \mathrm{d}\tau + n(t), \end{aligned} \tag{2}$$

where $h(t)$ denotes the VLC's channel impulse response and $n(t)$ is the noise. In particular, Barry's VLC channel model of [34, 35] is introduced in our study, which is formulated as

$$h(t) = \sum_{i=0}^{\infty} h_{\mathrm{B}}(i) \delta[t - \tau_{\mathrm{B}}(i)], \tag{3}$$

where $h_{\mathrm{B}}(i)$ and $\tau_{\mathrm{B}}(i)$ are the gain and delay of the ith bounce, respectively. Note that the gain and delay of the current bounce are dependent on that of the previous bounces. Let the frequency response of the VLC channel be denoted as $h(f)$, which can be obtained by taking the Fourier transform of (3). In practice, since the diameter of photodetectors is usually much larger than the wavelength of visible light, multipath propagation in VLCs does not impose fading on the optical signals. However, when signals carrying high-rate information are transmitted over such a channel, the time dispersion usually results in ISI. In (2), the noise $n(t)$ contains the effects of ambient light shot noise and thermal noise, which can commonly be modelled as mutually independent white Gaussian noise [34]. Therefore, we have $n(t) \sim \mathcal{N}(0, \sigma_{\mathrm{n}}^2)$.

C. Electrical Receiver

As seen in Fig. 2, the electrical current is low-pass filtered and analog-to-digital converted. Let us assume that perfect synchronization is achieved at the receiver. Then, the lth received TD sample can be expressed as

$$r_{\mathrm{T}}(l) = \sum_{i=0}^{L_{\mathrm{h}}-1} h_{\mathrm{B}}(i) s_{\mathrm{T}}(l-i) + n_{\mathrm{T}}(l), \tag{4}$$

where L_{h} denotes the maximum number of resolvable paths. Generally, a receiver processor is employed as a nonlinear mapping function of the received symbols. However, for the multipath propagation scenario of Section II-B, such a receiver may lead to excessive ICI in the FD, whilst additionally amplifying the noise. As a result, the system performance may in fact become degraded. In order to alleviate the nonlinearity-induced performance degradation, two different schemes are considered in this paper. To elaborate a little further, observe in Fig. 2 that, after removing the CP and carrying out the FFT, the received FD symbol can be formulated as

$$y_{\mathrm{F}}(m) = \frac{1}{\sqrt{M}} \sum_{l=0}^{M-1} r_{\mathrm{T}}(l) \exp\left(-j2\pi \frac{l}{M} m\right) \tag{5}$$

for $m = 0, 1, \ldots, M-1$. Then, the received FD symbol of (5) is fed into a single-tap FD equalizer, yielding $\widetilde{y}_{\mathrm{F}}(m) = y_{\mathrm{F}}(m)/h_{\mathrm{F}}(m)$, where without loss of generality h_{F} represents the FD channel transfer factor (FDCHTF) of the mth subchannel, which can be obtained from Section II-B. Here, we stipulate the idealized simplifying assumption that perfect channel estimation is achieved at the receiver. Next, either a low-complexity or a high-complexity receiver scheme is considered.

The low-complexity scheme is shown in the upper dashed box of Fig. 2, where the equalized symbols are directly demapped into the received bits by the conventional QAM demodulator, which results in an adequate performance.

However, when the nonlinearity induced performance degradation becomes severe at a high data rate, the higher-complexity scheme shown in the lower dashed box in Fig. 2 is selected. For the ease of understanding, let us temporarily employ vectored expressions. Firstly, without any ambiguity, the equalized symbols denoted as $\widetilde{\boldsymbol{y}}_{\mathrm{F}} = [\widetilde{y}_{\mathrm{F}}(0), \widetilde{y}_{\mathrm{F}}(1), \ldots, \widetilde{y}_{\mathrm{F}}(M-1)]^T$ can be expressed as

$$\widetilde{\boldsymbol{y}}_{\mathrm{F}} = \boldsymbol{\mathcal{F}}_M f_{\mathrm{Tx}}(\boldsymbol{x}_{\mathrm{T}}) + \widetilde{\boldsymbol{n}}_{\mathrm{F}}, \tag{6}$$

where $\widetilde{\boldsymbol{n}}_{\mathrm{F}}$ represents the filtered noise. Then, as seen in Fig. 2, the equalized symbols are entered into an M-point IFFT, yielding

$$\begin{aligned} \widetilde{\boldsymbol{y}}_{\mathrm{T}} &= \boldsymbol{\mathcal{F}}_M^H \widetilde{\boldsymbol{y}}_{\mathrm{F}} \\ &= f_{\mathrm{Tx}}(\boldsymbol{x}_{\mathrm{T}}) + \widetilde{\boldsymbol{n}}_{\mathrm{T}}, \end{aligned} \tag{7}$$

where, by definition, we have $\widetilde{\boldsymbol{n}}_{\mathrm{T}} = \boldsymbol{\mathcal{F}}_M^H \widetilde{\boldsymbol{n}}_{\mathrm{F}}$. Without loss of generality, we assume that a perfect expander is employed, i.e. we have $f_{\mathrm{Rx}}(x) = f_{\mathrm{Tx}}^{-1}(x)$. Hence, upon invoking an expansion by $f_{\mathrm{Rx}}(\widetilde{\boldsymbol{y}}_{\mathrm{T}})$, the resultant symbols are subjected to an M-point FFT, yielding

$$\begin{aligned} \boldsymbol{y}_{\mathrm{F}}^{'} &= \boldsymbol{\mathcal{F}}_M f_{\mathrm{Rx}}(\widetilde{\boldsymbol{y}}_{\mathrm{T}}) \\ &= \boldsymbol{\mathcal{F}}_M \boldsymbol{x}_{\mathrm{T}} + \boldsymbol{\mathcal{F}}_M f_{\mathrm{Rx}}(\widetilde{\boldsymbol{n}}_{\mathrm{T}}) \\ &= \boldsymbol{x}_{\mathrm{F}} + \boldsymbol{n}_{\mathrm{F}}^{'}, \end{aligned} \tag{8}$$

where, for convenience, we define $\boldsymbol{n}_{\mathrm{F}}^{'} = \boldsymbol{\mathcal{F}}_M f_{\mathrm{Rx}}(\widetilde{\boldsymbol{n}}_{\mathrm{T}})$. Finally, each symbol in $\boldsymbol{y}_{\mathrm{F}}^{'}$ is de-mapped by the conventional QAM demodulator, yielding the received bits. Specifically, when the transmitter processor is a CT function, the system's performance can be substantially improved by designing an effective CT function, which is detailed in the next section.

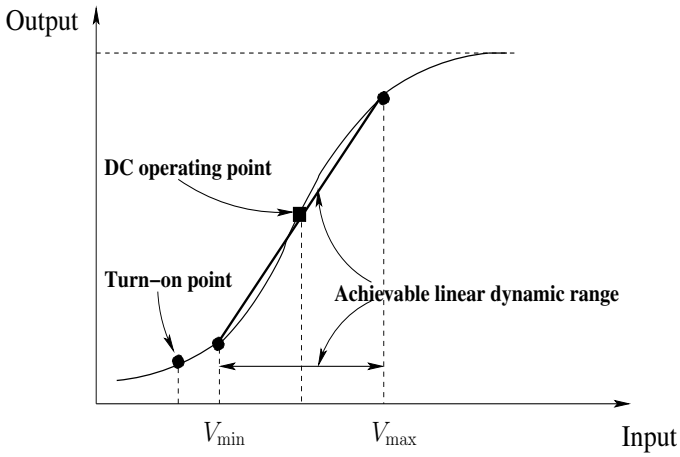


Fig. 3: Illustration of the linear dynamic range of a LED, which can be achieved by either employing a pre-distortor at the transmitter or using a post-distortor at the receiver.

III. Companding Transform of DCO-OFDM Signals

In the conventional CO-OFDM systems, the transmitter processor seen in Fig. 2 simply applies hard clipping. Explicitly, in DCO-OFDM systems, hard clipping tends to lead to a certain level of information loss, which degrades the attainable system performance. By contrast, in ACO-OFDM systems, due to the anti-symmetry property [13], clipping the negative part does not lead to any loss of information. Nonetheless, as shown in Section II-A, the anti-symmetric property is achieved at the cost of reduced bandwidth efficiency. In order to achieve a higher bandwidth efficiency at a reduced information loss, and furthermore, to guarantee that the DC-biased information-bearing signals are real and positive valued, in this paper, we introduce a particular type of CTs, namely the symmetric piecewise CT, as the transmitter processor. Before providing further details, let us first review the hard clipping operation of conventional CO-OFDM systems.

Let us model the input and output of the transmitter processor as random processes of

$$\left\{ X : X_l \triangleq x_{\mathrm{T}}(l) \in \mathbb{R} \right\} \tag{9a}$$

$$\left\{ S : S_l \triangleq s_{\mathrm{T}}(l) \in \mathbb{R} \right\}, \tag{9b}$$

respectively. Moreover, we have $S_l = f_{\mathrm{Tx}}(X_l)$ with respect to $s_{\mathrm{T}}(l) = f_{\mathrm{Tx}}(x_{\mathrm{T}}(l))$. Furthermore, according to Section II-A,

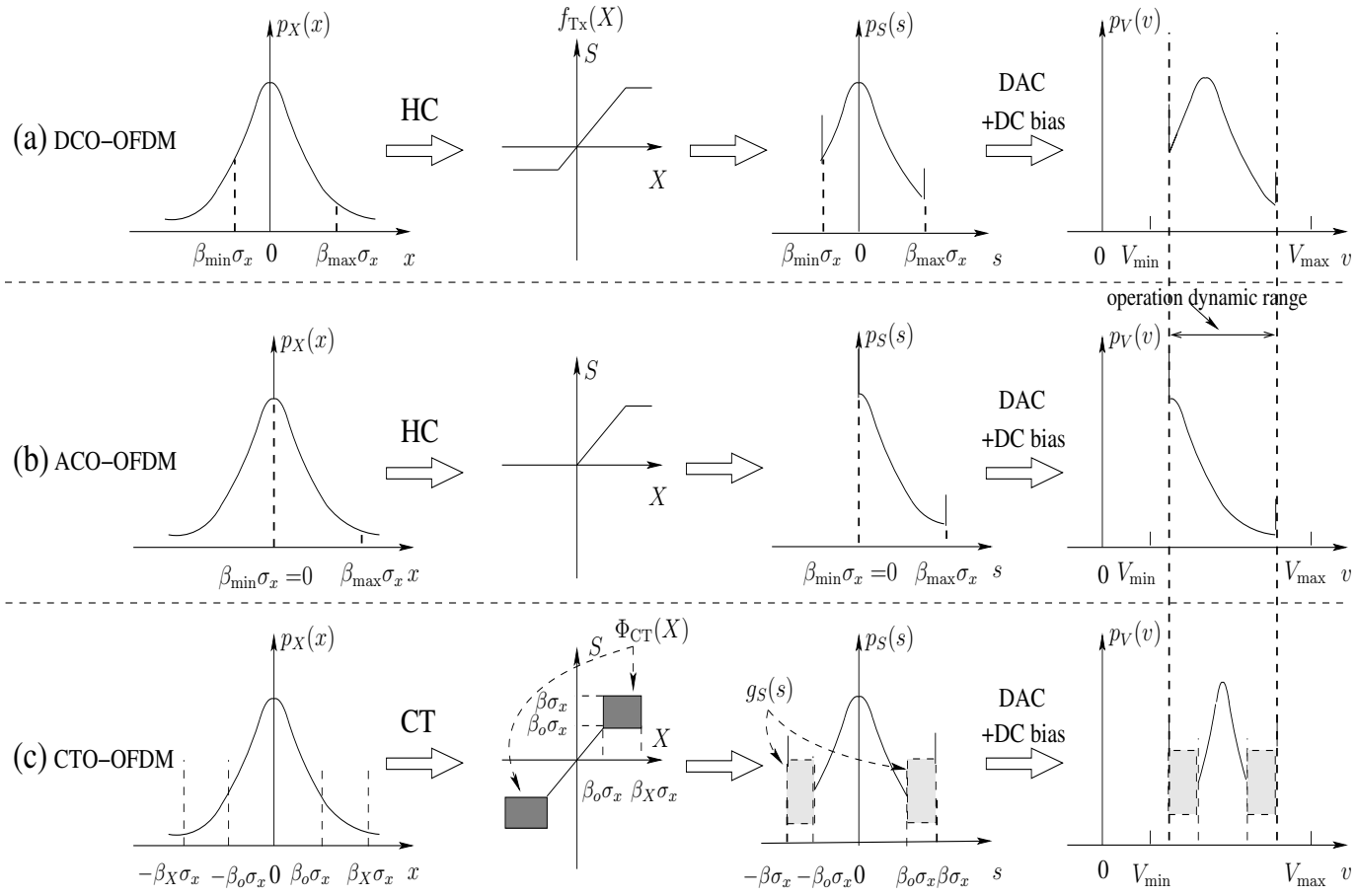


Fig. 4: Illustration of the nonlinear transformation that may be used by the transmitter processor.

the probability density function (PDF) and the cumulative distribution function (CDF) of the amplitudes of TD symbols can be approximated as

$$p_X(x) = \frac{Q_1\left(0, \frac{x}{\sigma_x}\right)}{\sigma_x\sqrt{2\pi}} \tag{10}$$

and as

$$P_X(x) = 1 - Q\left(\frac{x}{\sigma_x}\right), \tag{11}$$

respectively. In (10), we define the notation of $Q_1(0,x) \triangleq \exp(-x^2/2)$, which is actually a special case of the first order Marcum Q-function.

A. Review of Hard Clipping

In the conventional CO-OFDM systems, hard clipping is applied at the transmitter, implying that we have $f_{\text{Tx}}(X) = f_{\text{HC}}(X)$, where the hard clipping function can be expressed as

$$S = f_{\text{HC}}(X) = \begin{cases} \beta_{\min}\sigma_x, & \text{if } X \in \mathcal{X}_1^{\text{HC}} \\ x, & \text{if } X \in \mathcal{X}_2^{\text{HC}} \\ \beta_{\max}\sigma_x, & \text{if } X \in \mathcal{X}_3^{\text{HC}}, \end{cases} \tag{12}$$

where $\mathcal{X}_1^{\text{HC}} = (-\infty, \beta_{\min}\sigma_x]$, $\mathcal{X}_2^{\text{HC}} = (\beta_{\min}\sigma_x, \beta_{\max}\sigma_x)$ and $\mathcal{X}_3^{\text{HC}} = [\beta_{\max}\sigma_x, +\infty)$ are three disjoint regions of amplitudes satisfying $\bigcup_i \mathcal{X}_i^{\text{HC}} = \mathbb{R}$. The effect of (12) is shown in Fig. 4(a) and Fig. 4(b), where $\beta_{\min}\sigma_x$ and $\beta_{\max}\sigma_x$ denote the minimum and maximum clipping level, respectively. It should be noted that both the values of $\beta_{\min}\sigma_x$ and $\beta_{\max}\sigma_x$ are dependent on the achievable linear operational region of the LED, as depicted in Fig. 3. As shown in Fig. 4(b), the minimum clipping level associated with the ACO-OFDM system is given by $\beta_{\min}\sigma_x = 0$, i.e., all the negative components are clipped. By contrast, for the DCO-OFDM system shown in Fig. 4(a), only some of the high-peak negative components are clipped.

Based on (11) and (12), we can show that the PDF of the clipped symbol can be formulated as

$$p_S(s) = \begin{cases} P_X(\beta_{\min}\sigma_x) = Q(-\beta_{\min}), & \text{if } s = \beta_{\min}\sigma_x \\ \frac{Q_1\left(0, \frac{s}{\sigma_x}\right)}{\sigma_x\sqrt{2\pi}}, & \text{if } s \in \mathcal{S}_2^{\text{HC}} \\ 1 - P_X(\beta_{\max}\sigma_x) = Q(\beta_{\max}), & \text{if } s = \beta_{\max}\sigma_x \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where $\mathcal{S}_2^{\text{HC}} = \mathcal{X}_2^{\text{HC}}$ is the region of unchanged amplitudes. Here, the relationship between the PDF of the input amplitude and the PDF of the corresponding output amplitude of HC is shown. As seen in both Fig. 4(a) and Fig. 4(b), the hard clipping process usually leads to certain power loss, which is undesirable.

B. Principles of Symmetry Piecewise Companding Transform

As shown in Fig. 4(c), instead of using hard clipping, a specifically designed compander may be employed so that the biased signals can be operated in the desired region of the transmitter LED. In this case, we can let $f_{\text{Tx}}(x) = f_{\text{CT}}(x)$, where $f_{\text{CT}}(x)$ denotes the CT function used at the transmitter. The basic concept of a symmetric piecewise compander is illustrated in Fig. 4(c). Below, we develop the underlying principles in order to gain insights into the design of the symmetric piecewise compander.

Firstly, let us divide the TD amplitudes of O-OFDM symbols into five regions denoted as $\mathcal{X}_{1,-}^{\text{CT}} = (-\infty, -\beta_X\sigma_x]$, $\mathcal{X}_{3,-}^{\text{CT}} = (-\beta_X\sigma_x, -\beta_o\sigma_x]$, $\mathcal{X}_2^{\text{CT}} = (-\beta_o\sigma_x, \beta_o\sigma_x)$, $\mathcal{X}_{3,+}^{\text{CT}} = [\beta_o\sigma_x, \beta_X\sigma_x)$, $\mathcal{X}_{1,+}^{\text{CT}} = [\beta_X\sigma_x, +\infty)$, which can be further classified based on their magnitude into three types denoted as $\mathcal{X}_1^{\text{CT}} = \mathcal{X}_{1,-}^{\text{CT}} \bigcup \mathcal{X}_{1,+}^{\text{CT}}$, $\mathcal{X}_2^{\text{CT}}$ and $\mathcal{X}_3^{\text{CT}} = \mathcal{X}_{3,-}^{\text{CT}} \bigcup \mathcal{X}_{3,+}^{\text{CT}}$. The first region $\mathcal{X}_1^{\text{CT}}$ contains the high-peak components, which are hard clipped in our design. By contrast, in the second region $\mathcal{X}_2^{\text{CT}}$, the corresponding magnitudes are small enough so that they remain unchanged. Finally, the amplitudes within the third region $\mathcal{X}_3^{\text{CT}}$, which have moderate absolute values, are companded, as shown in Fig. 4(c). To simplify the implementation, we arrange for the CT function be symmetric about the origin. Hence, based on the above considerations, we can express the piecewise CT function as

$$S = f_{\text{CT}}(X) = \begin{cases} \text{sgn}(X)\beta\sigma_x, & \text{if } X \in \mathcal{X}_1^{\text{CT}} \\ X, & \text{if } X \in \mathcal{X}_2^{\text{CT}} \\ \text{sgn}(X)\Phi_{\text{CT}}(|X|), & \text{if } X \in \mathcal{X}_3^{\text{CT}}, \end{cases} \tag{14}$$

where the signum function is given by

$$\text{sgn}(X) = \begin{cases} -1, & \text{if } X < 0 \\ 0, & \text{if } X = 0 \\ 1, & \text{if } X > 0. \end{cases} \tag{15}$$

In (14), the sub-function denoted as $\text{sgn}(X)\Phi_{\text{CT}}(|X|)$ is applied for transforming its input amplitude $X \in \mathcal{X}_3^{\text{CT}}$ to the output amplitude in S, as shown by the shaded boxes in Fig. 4(c). Since there is an infinite variety of implementations for $\Phi_{\text{CT}}(|X|)$, we have to set up some conditions for narrowing the range of choices for the companding functions. At this moment, we should note that in order to strike an attractive tradeoff between the PAPR reduction and the bit error performance, the companded signals should be mapped smoothly to the desirable range of the LEDs. Bearing this in mind, some necessary conditions may be stated as

- **C.1** $f_{\text{CT}}(x)$ is symmetric across the origin;
- **C.2** $f_{\text{CT}}(x)$ has to be a bijective function over $X \in \mathcal{X}_2^{\text{CT}} \bigcup \mathcal{X}_3^{\text{CT}}$;
- **C.3** $\text{sgn}(X)\Phi_{\text{CT}}(|X|)$ has to be a strictly monotonically increasing function over $X \in \mathcal{X}_3^{\text{CT}}$;
- **C.4** $\Phi_{\text{CT}}(|X|)$ can be expressed in closed-form;
- **C.5** The PDF of companded signals denoted as $p_S(s)$ has to satisfy the usual probability condition of $\int_{-\infty}^{+\infty} p_S(s)\text{d}s = 1$.

Based on these conditions, we can now continue by showing the relationship between the companding function and the PDF of the corresponding output amplitudes as follows.

Based on **C.1**, it may be readily shown that the PDF of the corresponding output amplitudes is symmetric about the y-axis, as seen in the third sub-figure of Fig. 4(c). Hence, the PDF of the companded signals' amplitude in S given in (14) can be formulated as

$$p_S(s) = \begin{cases} Q(\beta_X), & \text{if } s = \beta\sigma_x, -\beta\sigma_x \\ \frac{Q_1(0, \frac{s}{\sigma_x})}{\sigma_x\sqrt{2\pi}}, & \text{if } s \in \mathcal{S}_2^{\text{CT}} \\ g_S(|s|), & \text{if } s \in \mathcal{S}_3^{\text{CT}} \\ 0, & \text{otherwise,} \end{cases} \tag{16}$$

where due to **C.2**, two output amplitude regions are given by $\mathcal{S}_2^{\text{CT}} = \mathcal{X}_2^{\text{CT}} = (-\beta_o\sigma_x, \beta_o\sigma_x)$ and $\mathcal{S}_3^{\text{CT}} = \mathcal{S}_{3,-}^{\text{CT}} \bigcup \mathcal{S}_{3,+}^{\text{CT}}$, with $\mathcal{S}_{3,-}^{\text{CT}} = (-\beta\sigma_x, -\beta_o\sigma_x]$ and $\mathcal{S}_{3,+}^{\text{CT}} = [\beta_o\sigma_x, \beta\sigma_x)$ corresponding to $\mathcal{X}_{3,-}^{\text{CT}}$ and $\mathcal{X}_{3,+}^{\text{CT}}$, respectively. Here, as shown in Fig. 4(c), $\beta_o\sigma_x$ represents the break-points of the compander, while $\beta_X\sigma_x$ and $\beta\sigma_x$ are the cutoff levels associated with the input X and the output S of the compander, respectively. In (16), $g_S(|s|)$, which is symmetric about the y-axis, represents the probability sub-distribution function (PsDF) of the amplitudes within $\mathcal{S}_3^{\text{CT}}$, as represented by the shaded boxes in Fig. 4(c).

Bearing the symmetry of $g_S(s)$ in mind, we can simplify our forthcoming discussions by considering the CT of only the positive amplitudes. In this case, we have $S = \text{sgn}(X)\Phi_{\text{CT}}(|X|) = \Phi_{\text{CT}}(X)$ and $g_S(|s|) = g_S(s)$ for the case of $X \in \mathcal{X}_{3,+}^{\text{CT}}$ and $S \in \mathcal{S}_{3,+}^{\text{CT}}$. To meet **C.5**, it is desirable to design a compander so that for each random variable $X_+ \in \mathcal{X}_{3,+}^{\text{CT}}$, we have

$$\int_{\beta_o\sigma_x}^{X_+} p_X(u_x)\text{d}u_x = \int_{\beta_o\sigma_x}^{\Phi_{\text{CT}}(X_+)} g_S(u_s)\text{d}u_s, \tag{17}$$

where we have $\Phi_{\text{CT}}(X_+) \geq \beta_o\sigma_x$ for $\Phi_{\text{CT}}(X_+) \in \mathcal{S}_{3,+}^{\text{CT}}$ due to the condition of **C.3**. Firstly, the left-hand-side (LHS) of (17) can be calculated with the aid of (11) as

$$\begin{aligned} \int_{\beta_o\sigma_x}^{X_+} p_X(u_x)\text{d}u_x =& P_X(X_+) - P_X(\beta_o\sigma_x) \\ =& Q(\beta_o) - Q\left(\frac{X_+}{\sigma_x}\right). \end{aligned} \tag{18}$$

Before proceeding with the right-hand-side (RHS) of (17), let us denote $G_S(s) = \int_{-\infty}^{s} g_S(u)\text{d}u$ as the CDF of $g_S(s)$. Moreover, due to **C.4**, we restrict our attention to specific PsDF having closed-form expressions for both $G_S(s)$ and for its inverse function $G_S^{-1}(s)$. Then, the RHS of (17) can be formulated as

$$\int_{\beta_o\sigma_x}^{\Phi_{\text{CT}}(X_+)} g_S(u_s)\text{d}u_s = G_S\left[\Phi_{\text{CT}}(X_+)\right] - G_S(\beta_o\sigma_x). \tag{19}$$

Upon substituting (18) and (19) into (17), we arrive at

$$Q(\beta_o) - Q\left(\frac{X_+}{\sigma_x}\right) = G_S\left[\Phi_{\text{CT}}(X_+)\right] - G_S(\beta_o\sigma_x). \tag{20}$$

$$S = f_{\text{CT}}(X) = \begin{cases} \text{sgn}(X)\beta\sigma_x, & \text{if } X \in \mathcal{X}_1^{\text{CT}} \\ X, & \text{if } X \in \mathcal{X}_2^{\text{CT}} \\ \text{sgn}(X)G_S^{-1}\left\{Q(\beta_o) - Q\left(\frac{|X|}{\sigma_x}\right) + G_S(\beta_o\sigma_x)\right\}, & \text{if } X \in \mathcal{X}_3^{\text{CT}} \end{cases} \tag{22}$$

With the aid of (20), the companding sub-function can now be expressed as

$$\Phi_{\text{CT}}(X_+) = G_S^{-1}\left\{Q(\beta_o) - Q\left(\frac{X_+}{\sigma_x}\right) + G_S(\beta_o\sigma_x)\right\} \tag{21}$$

for each input random variable of $X_+ \in \mathcal{X}_{3,+}^{\text{CT}}$. Based on the symmetry, the general companding function for the amplitude of the TD symbols can be finally formulated in (22), as shown at the top of this page.

So far, we have demonstrated the relationship between the companding function $S = f_{\text{CT}}(X)$ and the PDF $p_S(s)$ of the corresponding output amplitude. In particular, for a given PsDF $g_S(|s|)$, the companding sub-function $\text{sgn}(X)\Phi_{\text{CT}}(|X|)$ can be obtained with the aid of (21), which implies that we can design a compander in the probability domain. Below, we explore the characteristics of the PsDF $g_S(|s|)$ with the goal of identifying the conditions that can be used for the design of a compander.

C. Restrictions on the Design of the PsDF

As shown in (14), the companding sub-function $\text{sgn}(X)\Phi_{\text{CT}}(|X|)$ determines the final design of the symmetric piecewise compander. Furthermore, as shown in (22), the companding sub-function $\text{sgn}(X)\Phi_{\text{CT}}(|X|)$ can be obtained by specifying the PsDF $g_S(|s|)$ in the probability domain. Below, we detail some further restrictions imposed on the design of the PsDF.

Among the five necessary conditions of **C.1**-**C.5**, we can readily show that **C.2** and **C.5** can be used for quantifying both the PsDF and the companding sub-function, giving

$$G_S(\beta\sigma_x) - G_S(\beta_o\sigma_x) = Q(\beta_o) - Q(\beta_X). \tag{23}$$

We point out furthermore that since $\pm\beta_o\sigma_x$ are unique partitioning points (i.e. we have $f_{\text{CT}}(\pm\beta_o\sigma_x) = \pm\beta_o\sigma_x$), as a result of **C.2** and **C.5**, the PDF $p_S(s)$ should also be continuious at the break points of $\pm\beta_o\sigma_x$, yielding

$$g_S(|\pm\beta_o\sigma_x|) = p_X(\pm\beta_o\sigma_x) = \frac{Q_1(0,\beta_o)}{\sigma_x\sqrt{2\pi}}. \tag{24}$$

Next, let us consider the moments of S, which are important metrics directly affecting the system's performance. Firstly, due to the symmetry of the PDF $p_S(s)$, it can be readily shown that the first moment or the mean of S is given by

$$\mu_S \triangleq \mathbb{E}[S] = \int_{-\infty}^{+\infty} s p_S(s)\mathrm{d}s = 0. \tag{25}$$

Secondly, the second moment of S can be derived as

$$\begin{aligned}\mathbb{E}[S^2] =& \int_{-\infty}^{+\infty} s^2 p_S(s)\mathrm{d}s \\ =& 2(\beta\sigma_x)^2 Q(\beta_X) + \int_{\mathcal{S}_2^{\text{CT}}} s^2 \frac{Q_1\left(0, \frac{s}{\sigma_x}\right)}{\sigma_x\sqrt{2\pi}}\mathrm{d}s \\ &+ \int_{\mathcal{S}_3^{\text{CT}}} s^2 g_S(|s|)\mathrm{d}s \\ =& 2\beta^2\sigma_x^2 Q(\beta_X) + \sigma_x^2 - 2\sigma_x^2 Q(\beta_o) \\ &- \beta_o\sigma_x^2\sqrt{\frac{2}{\pi}}Q_1(0,\beta_o) + 2\int_{\mathcal{S}_{3,+}^{\text{CT}}} s^2 g_S(|s|)\mathrm{d}s, \end{aligned} \tag{26}$$

which is actually the average electrical power level of the companded signals. Intuitively, the lower the power-loss of the companded signals, the better the system performance becomes. Hence, let us define an average power-loss function as

$$\begin{aligned}\mathcal{L}_{\text{Po}} \triangleq& \mathbb{E}[X^2] - \mathbb{E}[S^2] \\ =& \sigma_x^2 - 2\beta^2\sigma_x^2 Q(\beta_X) - \sigma_x^2 + 2\sigma_x^2 Q(\beta_o) \\ &+ \beta_o\sigma_x^2\sqrt{\frac{2}{\pi}}Q_1(0,\beta_o) - 2\int_{\mathcal{S}_{3,+}^{\text{CT}}} s^2 g_S(|s|)\mathrm{d}s \\ =& 2\sigma_x^2 Q(\beta_o) - 2\beta^2\sigma_x^2 Q(\beta_X) + \beta_o\sigma_x^2\sqrt{\frac{2}{\pi}}Q_1(0,\beta_o) \\ &- 2\int_{\mathcal{S}_{3,+}^{\text{CT}}} s^2 g_S(|s|)\mathrm{d}s, \end{aligned} \tag{27}$$

where $\mathcal{L}_{\text{Po}} < 0$ means that the average power of signals is amplified after the CT, while $\mathcal{L}_{\text{Po}} > 0$ indicates that there is a certain power loss due to the CT. Generally, it is preferred to design a compander, which is capable of minimizing the average power-loss after the CT.

In summary, the PsDF $g_S(s)$ has to be designed to meet both (23) and (24), as well as to minimize the average power loss after the CT. Below, we consider the implementation of the symmetric piecewise compander.

D. Implementation of the Symmetric Piecewise Compander

Here we consider three types of functions as the implementation options for the PsDF, which we refer to as the constant probability sub-distribution function (CPsDF), the linear probability sub-distribution function (LPsDF) as well as the nonlinear probability sub-distribution function (NLPsDF). Again, we only detail the PsDF of positive amplitudes to simplify the presentation. We should note that in the following

analysis, the value of β is fixed since it is restricted by the operation dynamic range of the LED, as shown in Fig. 4.

1) CPsDF: In this scheme, the sub-function is chosen as $g_S(s) = \alpha_1$ for $s \in \mathcal{S}_{3,+}^{\mathrm{CT}}$, where α_1 is a constant. Correspondingly, we can readily see that $G_S(s) = \alpha_1 s + \alpha_2$ and $G_S^{-1}(s) = (s - \alpha_2)/\alpha_1$, where α_2 is also a constant. Hence, with the aid of (22), $\Phi_{\mathrm{CT}}(|X|)$ for $X \in \mathcal{X}_3^{\mathrm{CT}}$ can be formulated as

$$\Phi_{\mathrm{CT}}(|X|) = \frac{1}{\alpha_1}\left[-Q\left(\frac{|X|}{\sigma_x}\right) + Q(\beta_o)\right] + \beta_o\sigma_x, \tag{28}$$

where the constant α_1 can be obtained for meeting (23) and (24) as

$$\alpha_1 = \frac{Q(\beta_o) - Q(\beta_X)}{\sigma_x(\beta - \beta_o)} = \frac{Q_1(0, \beta_o)}{\sigma_x\sqrt{2\pi}}. \tag{29}$$

In particular, based on (29), β_X can be expressed in terms of β_o as

$$\beta_X = Q^{-1}\left\{Q(\beta_o) - \frac{\beta - \beta_o}{\sqrt{2\pi}}Q_1(0, \beta_o)\right\}, \tag{30}$$

where the value of β_o should be chosen to minimize the average power-loss of companded signals, under the constraint of

$$Q(\beta_o) - \frac{\beta - \beta_o}{\sqrt{2\pi}}Q_1(0, \beta_o) \geq 0 \tag{31}$$

for a given value of β. As proved in the Appendix, for a given $0 < \beta \leq \sqrt{2/\pi}$, the value of β_o is bounded by $0 \leq \beta_o \leq \beta$, while, for a given $\beta > \sqrt{2/\pi}$, the value of β_o is bounded by $\beta_{o,\min} \leq \beta_o < \beta$, where $\beta_{o,\min}$ is the lower bound of β_o still meeting the condition of (31).

Then, upon substituting (30) and $g_S(s) = \alpha_1$ into (27), the average power-loss function can be obtained as

$$\begin{aligned}\mathcal{L}_{\mathrm{Po}} =& 2\sigma_x^2(1 - \beta^2)Q(\beta_o) \\ &+ \sqrt{\frac{2}{\pi}}\sigma_x^2 Q_1(0, \beta_o)\left[\frac{1}{3}\beta_o^3 + \beta_o(1 - \beta^2) + \frac{2}{3}\beta^3\right].\end{aligned} \tag{32}$$

In order to obtain the optimal value of β_o for minimizing $\mathcal{L}_{\mathrm{Po}}$, let us take the first derivative of $\mathcal{L}_{\mathrm{Po}}$ with respect to β_o, yielding

$$\begin{aligned}\frac{\mathrm{d}\mathcal{L}_{\mathrm{Po}}}{\mathrm{d}\beta_o} =& -\frac{1}{3}\sqrt{\frac{2}{\pi}}\sigma_x^2 Q_1(0, \beta_o)\beta_o(\beta_o + 2\beta)(\beta_o - \beta)^2 \\ \leq& 0\end{aligned} \tag{33}$$

for any $0 \leq \beta_o < \beta$. Eq.(33) implies that the average-power loss function $\mathcal{L}_{\mathrm{Po}}$ is a non-increasing function. Hence, the optimal solution for $\min_{\beta_o} |\mathcal{L}_{\mathrm{Po}}|$ is $\beta_o = \beta$. In this case, the CT becomes HC, which goes against the goal of applying CT. Let the minimum average power-loss be expressed as

$$\mathcal{L}_{\mathrm{Po,min}} = \sigma_x^2\left[2(1 - \beta^2)Q(\beta) + \sqrt{\frac{2}{\pi}}\beta Q_1(0, \beta)\right]. \tag{34}$$

Then, a sub-optimal solution is to find $\beta_o < \beta$ so that the difference between the resultant average power-loss and the minimum average power-loss $\mathcal{L}_{\mathrm{Po,min}}$ is within a tolerable margin of ε_{C}, i.e., the value of β_o is chosen to satisfy

$$\mathcal{L}_{\mathrm{Po}} - \mathcal{L}_{\mathrm{Po,min}} \leq \varepsilon_{\mathrm{C}}, \tag{35}$$

where based on a certain value of ε_{C}, let us denote the corresponding parameters as $\beta_o^{\varepsilon_{\mathrm{C}}}$ and $\beta_X^{\varepsilon_{\mathrm{C}}}$. Finally, based on the values of $\beta_o^{\varepsilon_{\mathrm{C}}}$, $\beta_X^{\varepsilon_{\mathrm{C}}}$ and α_1, the desired CPsDF based compander can be obtained with the aid of substituting (28) into (22).

2) LPsDF: In the context of this scheme, the PsDF is chosen to be a linear function expressed as $g_S(s) = \alpha_1 s + \alpha_2$ for $s \in \mathcal{S}_{3,+}^{\mathrm{CT}}$, where both α_1 and α_2 are constants. Correspondingly, we can find that $G_S(s) = \alpha_1[s + (\alpha_2/\alpha_1)]^2/2 + \alpha_3 - \alpha_2^2/(2\alpha_1)$ and $G_S^{-1}(s) = [\sqrt{2\alpha_1 s + \alpha_2^2 - 2\alpha_1\alpha_3} - \alpha_2]/\alpha_1$, where α_3 is a constant. Then, with the aid of (22), $\Phi_{\mathrm{CT}}(|X|)$ for $X \in \mathcal{X}_3^{\mathrm{CT}}$ can be formulated as

$$\begin{aligned}\Phi_{\mathrm{CT}}(|X|) =& -\sqrt{\frac{-2Q\left(\frac{|X|}{\sigma_x}\right) + 2Q(\beta_o)}{\alpha_1} + \left(\beta_o\sigma_x + \frac{\alpha_2}{\alpha_1}\right)^2} \\ &- \frac{\alpha_2}{\alpha_1},\end{aligned} \tag{36}$$

where the constants α_1 and α_2 required to satisfy (23) and (24) can be found to be

$$\alpha_1 = \frac{2}{\sigma_x^2}\left[\frac{Q(\beta_o) - Q(\beta_X)}{(\beta - \beta_o)^2} - \frac{Q_1(0, \beta_o)}{\sqrt{2\pi}(\beta - \beta_o)}\right] \tag{37}$$

and

$$\alpha_2 = \frac{1}{\sigma_x}\left[\frac{(\beta + \beta_o)Q_1(0, \beta_o)}{\sqrt{2\pi}(\beta - \beta_o)} - \frac{2\beta_o\left[Q(\beta_o) - Q(\beta_X)\right]}{(\beta - \beta_o)^2}\right], \tag{38}$$

respectively. Note that as a result of (17), the LPsDF $g_S(s) = \alpha_1 s + \alpha_2$ is restricted to be a non-increasing function in the region of $\mathcal{S}_{3,+}^{\mathrm{CT}}$. Furthermore, the general condition of $g_S(s) = \alpha_1 s + \alpha_2 \geq 0$ always has to be satisfied. Thus, the values of β_o and β_X should be chosen to satisfy

$$\frac{Q_1(0, \beta_o)}{2\sqrt{2\pi}}(\beta - \beta_o) < Q(\beta_o) - Q(\beta_X) < \frac{Q_1(0, \beta_o)}{\sqrt{2\pi}}(\beta - \beta_o). \tag{39}$$

Next, based on (27), the average power-loss function of the LPsDF-based CT can be expressed as

$$\begin{aligned}\mathcal{L}_{\mathrm{Po}} =& 2\sigma_x^2\Big\{Q(\beta_X)[(\beta_o + \beta)^2 + \beta^2] \\ &- Q(\beta_o)[(\beta_o + \beta)^2 + 2\beta^2 - 1] \\ &+ \frac{Q_1(0, \beta_o)}{\sqrt{2\pi}}[2\beta_o^3 + \beta_o^2\beta + \beta_o\beta^2 + \beta_o + 2\beta^3]\Big\}.\end{aligned} \tag{40}$$

Since both the values of β_o and β_X are unknown, it is not easy to derive a closed-form expression for the minimum average power-loss in this case. Nevertheless, for the sake of comparison, we can simply let β_X be equal to the parameter obtained by the CPsDF-based CT, i.e. we have $\beta_X = \beta_X^{\varepsilon_{\mathrm{C}}}$. Then, the value of β_o can be obtained by solving the optimization problem of

$$\beta_o^{\mathrm{L}} = \min_{\beta_o} \mathcal{L}_{\mathrm{Po}} \quad \text{s.t.} \quad (39) \tag{41}$$

for given values of β and $\beta_X = \beta_X^{\varepsilon_C}$. Finally, based on the parameters β_o^{L}, $\beta_X^{\varepsilon_C}$ and β, the desired LPsDF based compander can be obtained with the aid of substituting (36) into (22).

3) NLPsDF: In this scenario, we consider a specific type of nonlinear functions, which is the square-root based function expressed as $g_S(s) = [\alpha_1(s+\alpha_2)]^{-\frac{1}{2}} + \alpha_3$ for $s \in \mathcal{S}_{3,+}^{\mathrm{CT}}$, where α_1, α_2 and α_3 are constants. Note that, other types of NLPsDF can be similarly analyzed. However, we leave other types of NLPsDF for our future research. Similarly, we can find $G_S(s) = 2\sqrt{(s+\alpha_2)/\alpha_1} + \alpha_3 s + \alpha_4$ and $G_S^{-1}(s) = \{\sqrt{[s-\alpha_4+\alpha_3\alpha_2+1/(\alpha_1\alpha_3)]/\alpha_3} - 1/(\alpha_3\sqrt{\alpha_1})\}^2 - \alpha_2$, where α_4 is a constant. Therefore, with the aid of (22), $\Phi_{\mathrm{CT}}(|X|)$ for $X \in \mathcal{X}_3^{\mathrm{CT}}$ can be formulated as

$$\Phi_{\mathrm{CT}}(|X|) = \Bigg\{ \Bigg[\frac{-Q\left(\frac{|X|}{\sigma_x}\right) + Q(\beta_o)}{\alpha_3} + \left(\sqrt{\beta_o\sigma_x + \alpha_2} + \frac{1}{\alpha_3\sqrt{\alpha_1}}\right)^2 \Bigg]^{\frac{1}{2}} - \frac{1}{\alpha_3\sqrt{\alpha_1}} \Bigg\}^2 - \alpha_2. \tag{42}$$

From (42), it can be readily shown that the compander is determined by the parameters α_1, α_2 and α_3. When either β_o or β_X is unknown, the problem of designing the NLPsDF based CT becomes undetermined. Hence, for the sake of both implementation and comparison, the values of β_o and β_X are assumed to be fixed in the case of NLPsDF. Specifically, we let $\beta_X = \beta_X^{\varepsilon_C}$. Thus, to meet (23) and (24), α_1 and α_3 can be formulated in terms of α_2, while the value of α_2 is chosen by minimizing the average power-loss function $\mathcal{L}_{\mathrm{Po}}$ given in (27). Finally, upon substituting the values of α_1, α_2 and α_3 into (42), we can obtain the desired NLPsDF based compander.

Finally, the CPsDF, LPsDF and NLPsDF based expanders can be simply obtained by calculating the related inverse functions of (14), respectively.

IV. Performance Analysis

Generally, the nonlinear distortion of TD OFDM signals results in loss of orthogonality, yielding ICI, which may degrade the attainable system performance. In this section, we analyze the nonlinear effect of both CO-OFDM and of our proposed CTO-OFDM systems. For notational brevity, the definitions given in (9a) and (9b) are also used in the following analysis.

A. Generalized Analytical Model for O-OFDM Systems

Let us commence by introducing a generalized analytical model for the O-OFDM systems considered in this paper. Firstly, as mentioned in Section II-A, when M is sufficiently large, the TD symbols $x_{\mathrm{T}}(l)$ (i.e., X) approximately follow the Gaussian distribution with a mean of zero and a variance of σ_x^2. In this case, the Bussgang theorem [36] can be invoked for analyzing the effects of nonlinear distortion. Since the nonlinearity considered in this paper is memoryless, the nonlinearity of the transmitter processor expressed as $S = f_{\mathrm{Tx}}(X)$ can be modelled as [36]

$$S = \rho_{\mathrm{a}} X + W, \tag{43}$$

where ρ_{a} is the attenuation factor that can be obtained as

$$\begin{aligned}\rho_{\mathrm{a}} =& \frac{\mathbb{E}[SX]}{\mathbb{E}[X^2]} \\ =& \frac{1}{\sigma_x^2}\mathbb{E}\left[f_{\mathrm{Tx}}(X)X\right] \\ =& \frac{1}{\sigma_x^2}\int_{-\infty}^{\infty} x f_{\mathrm{Tx}}(x) p_X(x)\mathrm{d}x.\end{aligned} \tag{44}$$

In (43), W contains the distortion that is uncorrelated with X, i.e. we have $\mathbb{E}[WX] = 0$. The mean and variance of the distortion term W can be derived as

$$\begin{aligned}\mu_w \triangleq& \mathbb{E}[W] \\ =& \mathbb{E}[S - \rho_{\mathrm{a}}X] = \mathbb{E}[S] - \rho_{\mathrm{a}}\mathbb{E}[X] \\ =& \mathbb{E}[S] - 0 \\ =& \int_{-\infty}^{\infty} s p_S(s)\mathrm{d}s\end{aligned} \tag{45}$$

and

$$\begin{aligned}\sigma_w^2 \triangleq& \mathbb{E}\left[W^2\right] - (\mathbb{E}[W])^2 \\ =& \mathbb{E}[(S - \rho_{\mathrm{a}}X)^2] - \mu_w^2 \\ =& \mathbb{E}[S^2] - \rho_{\mathrm{a}}^2\mathbb{E}[X^2] - \mu_w^2 \\ =& \int_{-\infty}^{\infty} s^2 p_S(s)\mathrm{d}s - \rho_{\mathrm{a}}^2\sigma_x^2 - \mu_w^2,\end{aligned} \tag{46}$$

respectively.

Next, similarly to the TD modelling shown in (9), let the FD symbol of each subchannel shown in (4) and (5) be modelled as

$$\left\{\mathsf{X} : \mathsf{X}_m \triangleq x_{\mathrm{F}}(m) \in \mathbb{C}\right\} \tag{47a}$$

$$\left\{\mathsf{Y} : \mathsf{Y}_m \triangleq y_{\mathrm{F}}(m) \in \mathbb{C}\right\} \tag{47b}$$

$$\left\{\mathsf{N} : \mathsf{N}_m \triangleq n_{\mathrm{F}}(m) \in \mathbb{C}\right\} \tag{47c}$$

$$\left\{\mathsf{W} : \mathsf{W}_m \triangleq w_{\mathrm{F}}(m) \in \mathbb{C}\right\}, \tag{47d}$$

where $w_{\mathrm{F}}(m)$ represents the interference experienced by the mth subchannel, which is engendered by the distortion W given in (43). Then, based on (43), we can readily model the received FD symbol shown in (5) as

$$\mathsf{Y} = \rho_{\mathrm{a}} h_{\mathrm{F}} \mathsf{X} + h_{\mathrm{F}} \mathsf{W} + \mathsf{N}. \tag{48}$$

B. SINR Analysis

Based on (48), the received instantaneous signal-to-interference-plus-noise ratio (SINR) of the mth subchannel

can be expressed as

$$\gamma_{\mathrm{d}}(m) = \frac{\rho_{\mathrm{a}}^2 \left|h_{\mathrm{F}}(m)\right|^2 \mathbb{E}[|\mathsf{X}|^2]}{\left|h_{\mathrm{F}}(m)\right|^2 \mathbb{E}[|\mathsf{W}|^2] + \mathbb{E}[|\mathsf{N}|^2]} = \frac{\rho_{\mathrm{a}}^2 \sigma_x^2}{\sigma_w^2 + \mu_w^2 + \frac{\sigma_{\mathrm{n}}^2}{|h_{\mathrm{F}}(m)|^2}}. \tag{49}$$

According to Section II, only the subchannels indexed with $m = 1, \ldots, N_{\mathrm{d}}$ are used for data transmission. Hence, the nonlinear effect induced by the transmitter processor $f_{\mathrm{Tx}}(X)$ on the data symbols can be quantified in terms of the average SINR of

$$\bar{\gamma}_{\mathrm{d}} = \frac{1}{N_{\mathrm{d}}} \sum_{m=1}^{N_{\mathrm{d}}} \gamma_{\mathrm{d}}(m), \tag{50}$$

where, upon substituting (49) into (50), we can show that for the time-invariant VLC channels, the parameters that determine the nonlinear effects are ρ_{a}, σ_w^2 and μ_w^2. Furthermore, it can be readily shown that the higher $\bar{\gamma}_{\mathrm{d}}$, the lower the nonlinear effects experienced by the O-OFDM systems becomes. Hence the better the performance becomes. Below, the SINRs of the various O-OFDM systems considered in this paper are derived.

1) DCO-OFDM: First, for the DCO-OFDM system, let us rewrite (43) as

$$S = \rho_{\mathrm{a,DC}} X + W_{\mathrm{DC}}, \tag{51}$$

where, upon substituting (10) and (12) into (44), the attenuation factor $\rho_{\mathrm{a,DC}}$ is given by

$$\rho_{\mathrm{a,DC}} = Q(\beta_{\min}) - Q(\beta_{\max}). \tag{52}$$

Moreover, upon substituting (13) into (45) and (46), the mean and variance of the distortion term can be derived as

$$\mu_{w_{\mathrm{DC}}} = \sigma_x \left[\beta_{\min} Q(-\beta_{\min}) + \frac{Q_1(0,\beta_{\min})}{\sqrt{2\pi}} + \beta_{\max} Q(\beta_{\max}) - \frac{Q_1(0,\beta_{\max})}{\sqrt{2\pi}}\right] \tag{53}$$

and

$$\sigma_{w_{\mathrm{DC}}}^2 = \sigma_x^2 \left[\beta_{\min}^2 + (1-\beta_{\min}^2) Q(\beta_{\min}) + \beta_{\min} \frac{Q_1(0,\beta_{\min})}{\sqrt{2\pi}} - (1-\beta_{\max}^2) Q(\beta_{\max}) - \beta_{\max} \frac{Q_1(0,\beta_{\max})}{\sqrt{2\pi}}\right] - \rho_{\mathrm{a,DC}}^2 \sigma_x^2 - \mu_{w_{\mathrm{DC}}}^2, \tag{54}$$

respectively. Finally, upon substituting (52), (53) and (54) into (49), we obtain the SINR for the DCO-OFDM system.

2) ACO-OFDM: In the context of the ACO-OFDM system, we express (43) as

$$S = \rho_{\mathrm{a,AC}} X + W_{\mathrm{AC}}. \tag{55}$$

Similarly to the DCO-OFDM system, the attenuation factor $\rho_{\mathrm{a,AC}}$ can be derived as

$$\rho_{\mathrm{a,AC}} = \frac{1}{2} - Q(\beta_{\max}). \tag{56}$$

Correspondingly, the mean and variance of the distortion term W_{AC} can be expressed as

$$\mu_{w_{\mathrm{AC}}} = \sigma_x \left[\frac{1 - Q_1(0,\beta_{\max})}{\sqrt{2\pi}} + \beta_{\max} Q(\beta_{\max})\right] \tag{57}$$

and

$$\sigma_{w_{\mathrm{AC}}}^2 = \sigma_x^2 \left[\frac{1}{2} - (1-\beta_{\max}^2) Q(\beta_{\max}) - \frac{\beta_{\max} Q_1(0,\beta_{\max})}{\sqrt{2\pi}}\right] - \rho_{\mathrm{a,AC}}^2 \sigma_x^2 - \mu_{w_{\mathrm{AC}}}^2, \tag{58}$$

respectively. Finally, upon substituting (56), (57) and (58) into (49), we arrive at the SINR expression of the ACO-OFDM system.

3) CTO-OFDM: Here, we consider two specific cases for the CTO-OFDM systems.

In the first case, we assume that the expander is not invoked at the receiver, i.e. the scheme shown in the upper dashed box of Fig. 2 is used. In this case, both the CT and the HC components are analyzed as follows. Firstly, we can express (43) as

$$S = \rho_{\mathrm{a,CT}} X + W_{\mathrm{CT}}. \tag{59}$$

Upon substituting (10) and (14) into (44), the attenuation factor $\rho_{\mathrm{a,CT}}$ can be expressed as

$$\begin{aligned}\rho_{\mathrm{a,CT}} =& \frac{1}{\sigma_x^2} \int_{-\infty}^{\infty} x f_{\mathrm{CT}}(x) p_X(x) \mathrm{d}x \\ =& \frac{2}{\sigma_x^2} \left\{ \int_0^{\beta_o \sigma_x} x^2 p_X(x) \mathrm{d}x + \beta \sigma_x \int_{\beta_X \sigma_x}^{\infty} x p_X(x) \mathrm{d}x \right. \\ & \left. + \int_{\beta_o \sigma_x}^{\beta_X \sigma_x} x f_{\mathrm{CT}}(x) p_X(x) \mathrm{d}x \right\} \\ =& 1 - 2Q(\beta_o) + \sqrt{\frac{2}{\pi}} \left[\beta Q_1(0,\beta_X) - \beta_o Q_1(0,\beta_o)\right] \\ & + \frac{2}{\sigma_x^2} \mathbb{E}\left[SX | X \in \mathcal{X}_{3,+}^{\mathrm{CT}}\right], \end{aligned} \tag{60}$$

where the last term is defined as $\mathbb{E}\left[SX | X \in \mathcal{X}_{3,+}^{\mathrm{CT}}\right] \triangleq \int_{\beta_o \sigma_x}^{\beta_X \sigma_x} x f_{\mathrm{CT}}(x) p_X(x) \mathrm{d}x$, for which closed-form formulas can be derived, when the implementations of the CPsDF, LPsDF as well as NLPsDF based CTs given in Section III-D are assumed. However, for the sake of simplicity, in this paper, Monte Carlo simulations are applied for calculating the term $\mathbb{E}\left[SX | X \in \mathcal{X}_{3,+}^{\mathrm{CT}}\right]$ for each of the above-mentioned implementations. Next, the mean of the distortion term W_{CT} can be obtained by substituting (16) into (45), yielding

$$\mu_{w_{\mathrm{CT}}} = 0. \tag{61}$$

Furthermore, based on (27), we can express the variance of the distortion term W_{CT} in term of the average power-loss function as

$$\begin{aligned}\sigma_{w_{\mathrm{CT}}}^2 =& \mathbb{E}[S^2] - \rho_{\mathrm{a,CT}}^2 \sigma_x^2 - \mu_{w_{\mathrm{CT}}}^2 \\ =& \mathbb{E}[X^2] - \mathcal{L}_{\mathrm{Po}} - \rho_{\mathrm{a,CT}}^2 \sigma_x^2 \\ =& \sigma_x^2 - \mathcal{L}_{\mathrm{Po}} - \rho_{\mathrm{a,CT}}^2 \sigma_x^2. \end{aligned} \tag{62}$$

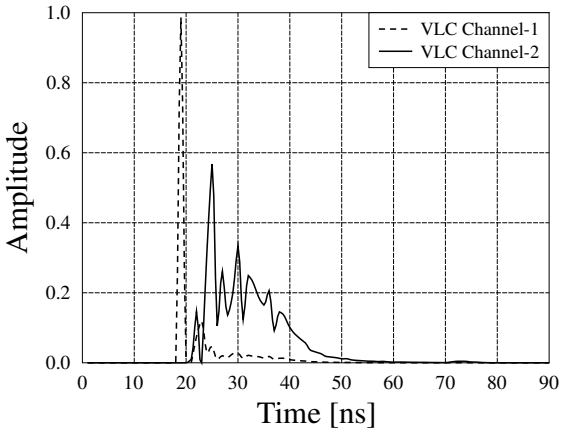


Fig. 5: Two indoor VLC channels generated according to [35] used in the simulations.

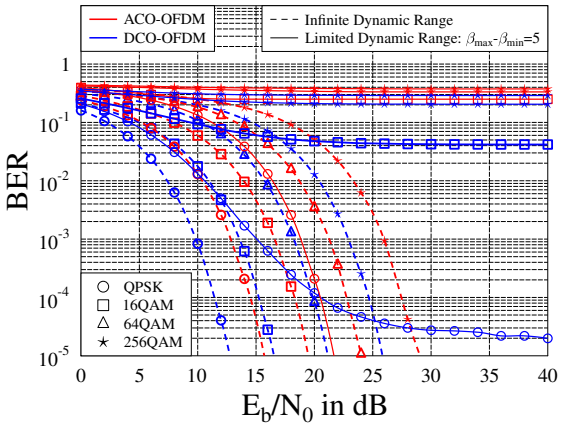


Fig. 6: Performance comparison between the DCO-OFDM and ACO-OFDM systems using QPSK, 16QAM and 256QAM, when communicating over AWGN channels.

TABLE I: Parameters for All Simulations

	Q	M	$N_{\rm d}$	$N_{\rm d}/M$	$L_{\rm cp}$
ACO-OFDM	$4, 16, 64, 256$	256	64	25%	5 or 13
DCO-OFDM			127	49.61%	
CTO-OFDM			127	49.61%	

Consequently, the SINR of this case can be obtained by substituting (60), (61) and (62) into (49).

In the second case, we assume that the CT components are perfectly restored without amplifying the noise at the receiver side. However, no processing is assumed by the HC components. It can be readily shown that this is actually equivalent to the case of DCO-OFDM. But, the detailed expressions are omitted here since they can be obtained from those of the DCO-OFDM by substituting $\beta_{\rm min}$ and $\beta_{\rm max}$ into $-\beta_X$ and β_X, respectively. Thus, when the HC of DCO-OFDM is assumed to be symmetric in conjunction with $\beta_{\rm min} = -\beta$ and $\beta_{\rm max} = \beta$, we can readily show that our CTO-OFDM is capable of outperforming DCO-OFDM, since we have $\beta_X > \beta$ implying a higher SINR for our CTO-OFDM than that of DCO-OFDM. However, it should be noted that this is achieved at the cost of an increased complexity.

V. Performance Results

In this section, simulation results are provided for investigating the performance of the CO-OFDM and CTO-OFDM systems. The system setup and the parameters used in our simulations are summarized in Table I. We consider the communication scenario, where the transmitter and receiver are placed in a room with dimensions of [length=4 m, width=4 m, height=8 m]. Correspondingly, the transmitter LED is placed at the position of [0.06 m, 0.06 m, 6 m], while the PD is placed at the position of [1.5 m, 2.5 m, 1 m]. The other parameters for the transmitter, receiver, as well as the reflectance of building material are assumed to be the same as those stated in [35]. Specifically, three bounces are considered for calculating the impulse response of the VLC channels expressed in (3). In our simulations, the pair of indoor VLC multipath channels shown in Fig. 5 are considered. As seen in Fig. 5, the second channel model is more dispersive than the first channel model. Furthermore, we consider O-OFDM systems using QPSK, 16QAM, 64QAM or 256QAM in conjunction with $M = 256$ subcarriers, among which $N_{\rm d}$ subcarriers are used for data transmission. According to the principles of the O-OFDM systems considered, we have $N_{\rm d} = (M-2)/2$ for both the DCO-OFDM and CTO-OFDM systems, while $N_{\rm d} = M/4$ for the ACO-OFDM system. Therefore, it can be shown that about $N_{\rm d}/M = 49.61\%$ of the subchannels are actively used for data transmission in both the DCO-OFDM and the CTO-OFDM systems. By contrast, only $N_{\rm d}/M = 25\%$ subchannels are employed by the ACO-OFDM system. In all the O-OFDM systems, the CP length is chosen to be $L_{\rm cp} = 5$ and 13 samples for the first and second VLC channel models, respectively. For the sake of comparison, the transmit power of baseband signals is normalized to unity, i.e. we have $\sigma_x^2 = 1$. Moreover, three dynamic ranges are considered, namely $2\beta\sigma_x^2 = (\beta_{\rm max} - \beta_{\rm min})\sigma_x^2 = 4$, 5 and 6.

In Fig. 6, the BER performance of the ACO-OFDM and DCO-OFDM systems is compared, when communicating over non-dispersive AWGN channels. In this figure, QPSK, 16QAM, 64QAM and 256QAM are considered. Correspondingly, a CP length of $L_{\rm cp} = 0$ is set for both the systems. For the dynamic range, in addition to $\beta_{\rm max} - \beta_{\rm min} = 5$, we also consider the ideal scenario of having an infinite dynamic range. Specifically, the dynamic range set for the DCO-OFDM system is $[\beta_{\rm min}\sigma_x^2, \beta_{\rm max}\sigma_x^2] = [-1, 4]$. From the results of Fig. 6, we infer the following observations. First, for a given CO-OFDM system operating with a limited dynamic range, as the constellation size Q increases, the BER becomes more and more sensitive to the nonlinear distortion imposed by the hard clipping. This trend is the same as that observed in the classic OFDM system in RF communications, experiencing nonlinear distortion, as shown in [5, 6]. Secondly, for the QPSK scheme, the BER performance of the ACO-OFDM

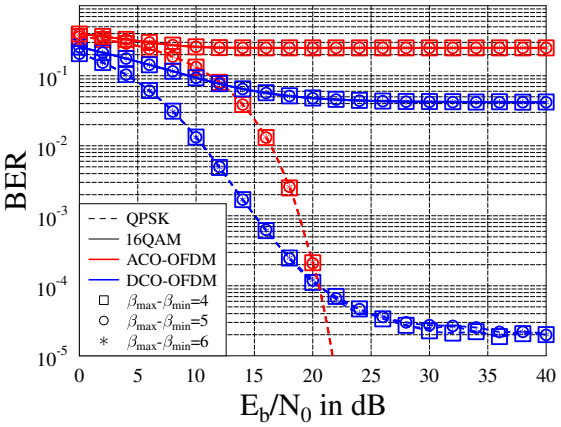
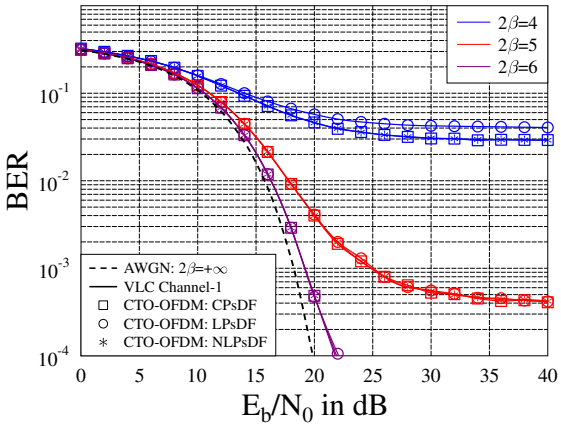
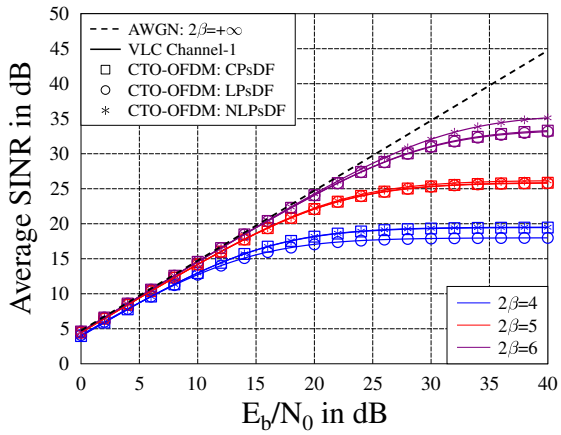


Fig. 7: BER performance of both the DCO-OFDM and ACO-OFDM systems operated with different dynamic ranges, when communicating over AWGN channels.



(a)



(b)

Fig. 8: Performance comparison of CPsDF, LPsDF and NLPsDF based CTO-OFDM systems using 64QAM, when communicating over VLC channel-1.

system is better than that of the DCO-OFDM system, when the SNR exceeds a certain level, such as 20.3 dB in this figure. Before this crossing point, DCO-OFDM outperforms ACO-OFDM. By contrast, for the relatively large constellations of 16QAM, 64QAM and 256QAM, as shown in Fig. 6, the DCO-OFDM system outperforms the ACO-OFDM system within the observation range considered. These observations can be explained with the aid of the analysis provided in Section IV-B. For a given dynamic range of $\beta_{\max} - \beta_{\min} = 5$, the attenuation factors obtained from (52) and (56) are $\rho_{\text{a,DC}} = 0.8413$ for the DCO-OFDM and $\rho_{\text{a,AC}} = 0.5$ for the ACO-OFDM, respectively. Thus, for a given transmitted energy per bit, the BER performance of the DCO-OFDM system associated with a lower power is better than that of the ACO-OFDM system, since 16QAM, 64QAM or 256QAM, are highly sensitive to the nonlinear distortion engendered by hard clipping. Finally, as seen in Fig. 6, an error floor occurs for the CO-OFDM systems associated with large constellation sizes for a limited dynamic range. This is because the nonlinear distortion generated by hard clipping usually results in ICI, as detailed in Section IV. Hence, when the SNR is sufficiently high, the SINR is dominated by the ICI, as seen in (49). As a result, the system exhibits an error floor.

In Fig. 7, we study the BER performance of the DCO-OFDM and ACO-OFDM systems for different dynamic ranges of $\beta_{\max} - \beta_{\min} = 4$, 5 and 6, when communicating over AWGN channels. Both QPSK and 16QAM are characterized. Again, due to the AWGN channels considered, a CP length of $L_{\text{cp}} = 0$ is assumed. Specifically, the dynamic ranges of the DCO-OFDM system are set to $[-1, 3]$, $[-1, 4]$ and $[-1, 5]$. Observe from Fig. 7 that there is no significant difference among the BER curves of a given CO-OFDM system for a given constellation scheme, when different dynamic ranges are used. This observation can be explained as follows. According to our analytical results given in Section IV-B, the achievable SINR of both CO-OFDM systems is mainly dependent on the lower clipping limit $\beta_{\min}$, as shown in (49). In other words, given a sufficiently wide dynamic range, the upper clipping limit $\beta_{\max}$ has little impact on the resultant parameters of ρ_{a}, μ_w and σ_w^2, as indicated by (52)-(54) and (56)-(58). In this case, when the hard clipping scheme is employed, the system performance can be hardly improved by further increasing the dynamic range by increasing the upper clipping limit $\beta_{\max}$.

In Fig. 8, we study the performance of CTO-OFDM systems employing CPsDF, LPsDF or NLPsDF based CTs, when communicating over multipath indoor VLC channels. For the sake of comparison, the VLC channel-1 associated with a CP length of $L_{\text{cp}} = 5$ and a 64QAM scheme are employed. It should be noted that the performance of other modulation schemes and channel scenarios will obey similar trends, as inferred from our theoretical analysis provided in Section IV-B. The dynamic ranges are set to $[-2, 2]$, $[-2.5, 2.5]$ and $[-3, 3]$. Furthermore, in order to investigate the inherent nonlinear distortion of each CT scheme, no expanders are employed in these simulations. Observe from the results of Fig. 8 (a) that

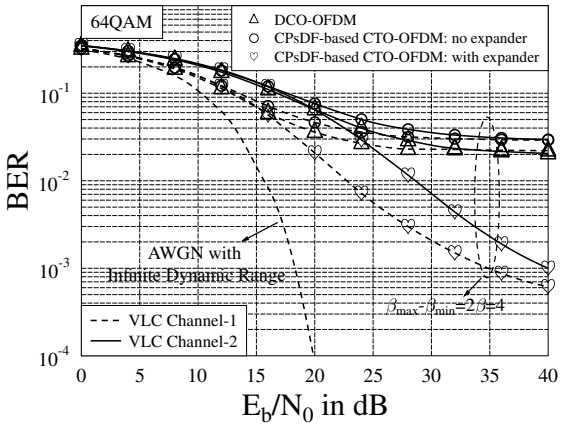


Fig. 9: Performance comparison between the DCO-OFDM and CTO-OFDM systems using 64QAM, when communicating over indoor VLC channels. The dynamic ranges of both systems are set to $[-2, 2]$.

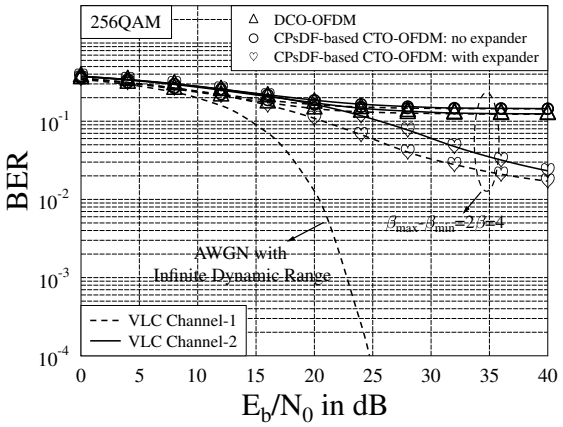


Fig. 11: Performance comparison between the DCO-OFDM and CTO-OFDM systems using 256QAM, when communicating over indoor VLC channels. The dynamic ranges of both systems are set to $[-2, 2]$.

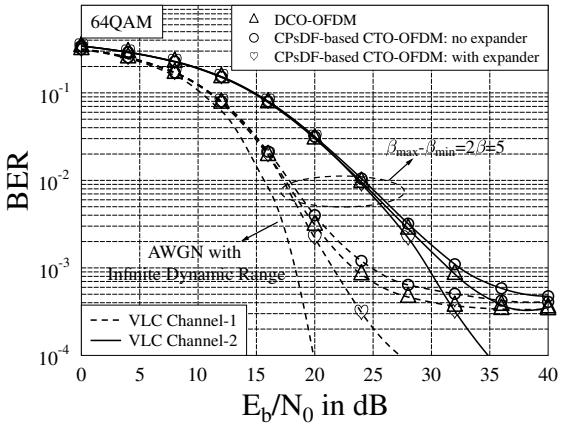


Fig. 10: Performance comparison between the DCO-OFDM and CTO-OFDM systems using 64QAM, when communicating over indoor VLC channels. The dynamic ranges of both systems are set to $[-2.5, 2.5]$.

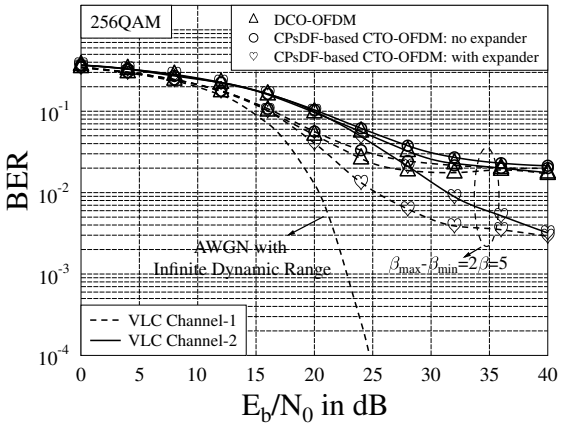


Fig. 12: Performance comparison between the DCO-OFDM and CTO-OFDM systems using 256QAM, when communicating over indoor VLC channels. The dynamic ranges of both systems are set to $[-2.5, 2.5]$.

the BER performance of the CPsDF-based CT is competitive in comparison to that of the NLPsDF-based CT, both of which slightly outperform the LPsDF-based CT. This observation can in fact be explained with the aid of the analytical SINR results of Fig. 8 (b). As seen in these figures, the average SINR of the CPsDF-based CT is almost the same as that of the NLPsDF-based CT. However, the LPsDF-based CT suffers from a higher interference than both the CPsDF and NLPsDF based CTs. Correspondingly, the BER performance of the LPsDF-based CT is worse than that of the CPsDF and NLPsDF based CTs.

In Fig. 9-12, we investigate the performance of the DCO-OFDM and CTO-OFDM system using 64QAM or 256QAM, when communicating over indoor VLC multipath channels. Specifically, the VLC channel-1 associated with a CP length of $L_{\text{cp}} = 5$ and the VLC channel-2 with a CP length of $L_{\text{cp}} = 13$ are employed. For the sake of comparison, the dynamic ranges of both systems are set to $[-2, 2]$ or $[-2.5, 2.5]$. As inferred from Fig. 8, the CPsDF-based CT is employed by the CTO-OFDM system, where the pair of receiver processing schemes shown in Fig. 2 are used. From the results of Fig. 9- 12 we infer the following observations.

Firstly, as observed from Fig. 9-12, the CTO-OFDM system using the expander always outperforms the DCO-OFDM system. This observation confirms our analysis of CTO-OFDM carried out in Section IV-B, where we have shown that if the CT components can be recovered at the receiver sider, our CTO-OFDM becomes capable of outperforming DCO-OFDM. Secondly, as seen in Fig. 9 -12, as the operational

range increases, the performance gain attained by CTO-OFDM becomes less attractive in the light of the corresponding receiver complexity. This implies that a tradeoff has to be struck between the performance gain and complexity cost for designing efficient CTO-OFDM systems. Finally, as inferred from Fig. 9 and Fig. 12, if DCO-OFDM fails to support reliable communications, CTO-OFDM may be activated. This becomes more evident, when error correction codes are employ. Therefore, the flexibility of our CTO-OFDM system is higher than that of the CO-OFDM system.

VI. Conclusions

In this paper, piecewise CT schemes have been conceived for O-OFDM systems, in order to constrain the transmitted signals within the limited dynamic range of LEDs. We have outlined the general principles and design criteria of piecewise CT schemes operated in different optical propagation scenarios. Three types of piecewise companders have been designed for the O-OFDM systems, which are the CPsDF, LPsDF and NLPsDF based CTs. Furthermore, the nonlinear effect of both the hard clipping and CT on the performance of the O-OFDM systems has been investigated in the context of different scenarios by both analysis and simulations. Our studies show that in comparison to the ACO-OFDM system, both the DCO-OFDM and CTO-OFDM systems are capable of attaining a higher bandwidth efficiency. Moreover, the CTO-OFDM system is shown to exhibit a high flexibility, as well as a capability to support high data-rate transmissions in the face of nonlinear effects. We also point out that a tradeoff has to be struck between the performance gain and the complexity cost for designing efficient CTO-OFDM systems, which will be part of our future work.

Appendix

Firstly, let us define the LHS of (31) as

$$f(\beta_o) \triangleq Q(\beta_o) - \frac{\beta - \beta_o}{\sqrt{2\pi}} Q_1(0, \beta_o), \tag{A.1}$$

which has the first derivative of

$$\frac{\mathrm{d}f(\beta_o)}{\mathrm{d}\beta_o} = \frac{\beta_o(\beta - \beta_o)}{\sqrt{2\pi}} Q_1(0, \beta_o) \geq 0, \tag{A.2}$$

showing that $f(\beta_o)$ is a non-decreasing function. Hence, for a given $\beta \leq \sqrt{2/\pi}$, we can readily show that

$$f(\beta_o) \geq f(0) = \frac{1}{2} - \frac{\beta}{\sqrt{2\pi}} \geq 0, \tag{A.3}$$

implying that for any $0 \leq \beta_{\mathrm{o}} < \beta$, Eq.(31) is always satisfied. On the other hand, for a given $\beta > \sqrt{2/\pi}$, there exists a lower bound of β_o to meet the condition of (31), i.e., we have

$$\begin{cases} f(\beta_o) \geq 0, & \text{if } \beta_{o,\min} \leq \beta_o < \beta \\ f(\beta_o) < 0, & \text{if } 0 \leq \beta_o < \beta_{o,\min}. \end{cases} \tag{A.4}$$

References

[1] L. Hanzo *et al.*, "Wireless myths, realities, and futures: From 3G/4G to optical and quantum wireless," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1853–1888, May 2012.

[2] T. Komine and M. Nakagawa, "Integrated system of white LED visible-light communication and power-line communication," *IEEE Transactions on Consumer Electronics,*, vol. 49, no. 1, pp. 71–79, Feb 2003.

[3] A. Jovicic, J. Li and T. Richardson, "Visible light communication: opportunities, challenges and the path to market," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 26–32, December 2013.

[4] H. Burchardt, N. Serafimovski, D. Tsonev, S. Videv and H. Haas, "VLC: Beyond point-to-point communication," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 98–105, July 2014.

[5] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-user Communications, WLANs and Broadcasting*. John Wiley & Sons, May 2003. [Online]. Available: http://eprints.soton.ac.uk/258228/

[6] Lie-Liang Yang, *Multicarrier Communications*. John Wiley & Sons, 2009.

[7] R. Zhang and L. Hanzo, "Multi-layer modulation for intensity-modulated direct-detection optical OFDM," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 5, no. 12, pp. 1402–1412, Dec 2013.

[8] J. Jiang, P. Zhang, R. Zhang, S. Chen, and L. Hanzo, "Aperture selection for ACO-OFDM in free-space optical turbulence channel," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6089–6100, Aug 2016.

[9] J. Jiang *et al.*, "Video streaming in the multiuser indoor visible light downlink," *IEEE Access*, vol. 3, pp. 2959–2986, 2015.

[10] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications," *Proceedings of the IEEE*, vol. 88, no. 5, pp. 611–640, May 2000.

[11] J. Armstrong, "OFDM for optical communications," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 189–204, Feb 2009.

[12] J. Carruthers and J. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 3, pp. 538–546, April 1996.

[13] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electronics Letters*, vol. 42, no. 6, pp. 370–372, March 2006.

[14] S. D. Dissanayake and J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD systems," *Journal of Lightwave Technology*, vol. 31, no. 7, pp. 1063–1072, April 2013.

[15] J. Armstrong and B. J. C. Schmidt, "Comparison of asymmetrically clipped optical OFDM and DC-biased optical OFDM in AWGN," *IEEE Communications Letters*, vol. 12, no. 5, pp. 343–345, May 2008.

[16] D. Barros, S. Wilson and J. Kahn, "Comparison of orthogonal frequency-division multiplexing and pulse-amplitude modulation in indoor optical wireless links," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 153–163, January 2012.

[17] S. Dimitrov, S. Sinanovic and H. Haas, "Clipping noise in OFDM-based optical wireless communication systems," *IEEE Transactions on Communications*, vol. 60, no. 4, pp. 1072–1081, April 2012.

[18] X. Ling, J. Wang, X. Liang, Z. Ding and C. Zhao, "Offset and power optimization for DCO-OFDM in visible light communication systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 349–363, Jan 2016.

[19] L. Chen, B. Krongold, and J. Evans, "Theoretical characterization of nonlinear clipping effects in IM/DD optical OFDM systems," *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2304–2312, August 2012.

[20] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, no. 2, 2005.

[21] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, June 2008.

[22] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Communications Surveys Tutorials*, vol. 15, no. 4, pp. 1567–1592, Fourth 2013.

[23] X. Wang, T. T. Tjhung and C. S. Ng, "Reduction of peak-to-average power ratio of OFDM system using a companding technique," *IEEE Transactions on Broadcasting*, vol. 45, no. 3, pp. 303–307, Sep 1999.

[24] X. Wang, T. T. Tjhung, C. S. Ng and A. A. Kassim, "On the SER analysis of A-law companded OFDM system," in *IEEE Global Telecommunications Conference, 2000. GLOBECOM '00*, vol. 2, 2000, pp. 756–760.

[25] X. Huang, J. Lu, J. Zheng, K. B. Letaief and J. Gu, "Companding transform for reduction in peak-to-average power ratio of OFDM signals," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2030–2039, Nov 2004.

[26] T. Jiang, Y. Yang and Y.-H. Song, "Exponential companding technique for PAPR reduction in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 51, no. 2, pp. 244–248, June 2005.

[27] J. Hou, J. H. Ge and J. Li, "Trapezoidal companding scheme for peak-to-average power ratio reduction of OFDM signals," *Electronics Letters*, vol. 45, no. 25, pp. 1349–1351, December 2009.

[28] J. Hou, J. Ge, D. Zhai and J. Li, "Peak-to-average power ratio reduction of OFDM signals with nonlinear companding scheme," *IEEE Transactions on Broadcasting*, vol. 56, no. 2, pp. 258–262, June 2010.

[29] Y. Wang, L. H. Wang, J. H. Ge and B. Ai, "An efficient nonlinear companding transform for reducing PAPR of OFDM signals," *IEEE Transactions on Broadcasting*, vol. 58, no. 4, pp. 677–684, Dec 2012.

[30] S. P. DelMarco, "General closed-form family of companders for PAPR reduction in OFDM signals using amplitude distribution modification," *IEEE Transactions on Broadcasting*, vol. 60, no. 1, pp. 102–109, March 2014.

[31] M. Hu, Y. Li, W. Wang and H. Zhang, "A piecewise linear companding transform for PAPR reduction of OFDM signals with companding distortion mitigation," *IEEE Transactions on Broadcasting*, vol. 60, no. 3, pp. 532–539, Sept 2014.

[32] H. Elgala, R. Mesleh and H. Haas, "Predistortion in optical wireless transmission using OFDM," in *Ninth International Conference on Hybrid Intelligent Systems, 2009. HIS '09.*, vol. 2, Aug 2009, pp. 184–189.

[33] H. Qian, S. J. Yao, S. Z. Cai and T. Zhou, "Adaptive postdistortion for nonlinear LEDs in visible light communications," *IEEE Photonics Journal*, vol. 6, no. 4, pp. 1–8, Aug 2014.

[34] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb 1997.

[35] K. Lee, H. Park and J. R. Barry, "Indoor channel characteristics for visible light communications," *IEEE Communications Letters*, vol. 15, no. 2, pp. 217–219, February 2011.

[36] J. Bussgang, "Cross correlation function of amplitude-distorted Gaussian signals," Research Laboratory of Electronics, Massachusetts Institute of Technology, Tech. Rep., 1952.

[37] S. Randel, F. Breyer, S. Lee and J. Walewski, "Advanced modulation schemes for short-range optical communications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1280–1289, Sept 2010.