A Tone Reservation-Based Optical OFDM System for Short-Range IM/DD Transmission

Liang Chen, Brian Krongold, and Jamie Evans

***Abstract*—This paper investigates the use of tone reservation for optical orthogonal frequency-division multiplexing (OFDM) transmission in a short-range, intensity-modulated, direct-detected (IM/DD) optical channel. Due to the unipolarity of the IM/DD channel, any asymmetric clipping of the bipolar OFDM signal and associated nonlinear distortion will result in an unfavorable dc power and significant performance penalty. Using a detailed analysis of the clipping distortion, we propose a new framework to enable IM/DD optical OFDM transmission, where certain subchannels are reserved to produce a negative peak cancellation (NPC) signal via a frequency-domain digital filter. This NPC signal guarantees a unipolar OFDM signal, hence eliminating any additional biasing and/or clipping. It also produces no in-band or out-of-and distortions and requires no transmission of any side information. Beneficial tradeoffs are identified that enable an efficient use of both the available power and spectrum.**

***Index Terms*—Direct detection, intensity modulation, optical orthogonal frequency-division multiplexing (OFDM), tone reservation.**

I. Introduction

With its high alignment tolerance and immunity to harsh environments, multimode fiber (MMF) has became increasingly attractive for high-speed, short-range applications, such as data centers, local area networks (LAN), storage networks, and supercomputers [1], [2]. For such cost-sensitive scenarios, the popular state-of-the-art coherent solutions proposed for long-haul transmission, such as coherent optical orthogonal frequency-division multiplexing (OFDM) or differential quadrature phase-shift keying, may not be suitable due to their high cost and/or complexity. An attractive alternative is an intensity modulated, direct-detected (IM/DD) optical OFDM system, such as dc-biased OFDM (DC-OFDM) in [2]–[4], which seems to provide a lower installation and maintenance cost, a wider compatibility to legacy IM/DD optical transceivers and links, as well as a better scalability to different channel conditions and data rates all at the same time. Although IM/DD optical OFDM systems can only compensate for linear impairments, the achievable rates in MMF links are likely to be limited by a combination of optical power and modal dispersion, and nonlinear effects should be insignificant.

Manuscript received July 27, 2011; revised September 25, 2011; accepted October 03, 2011. Date of publication October 24, 2011; date of current version December 23, 2011. This work was supported by National ICT, Australia (NICTA). NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

The authors are with National ICT Australia, Victoria Research Laboratory, Department of Electrical and Electronic Engineering, School of Engineering, University of Melbourne, Parkville, Vic. 3010, Australia (e-mail: lianchen@unimelb.edu.au; bsk@unimelb.edu.au; jse@unimelb.edu.au).

Digital Object Identifier 10.1109/JLT.2011.2173297

The modulated signal in IM/DD optical OFDM is the output of the inverse fast Fourier transform (IFFT) and is bipolar in general. Thus, it cannot be directly applied to conventional IM/DD systems that are designed for unipolar transmission, such as ON–OFF keying and pulse position modulation. A popular solution is to use biasing followed by clipping (BAC) to simply transform the bipolar IFFT output into a unipolar signal [2]–[8]. This method requires minimal modifications to existing IM/DD systems, and hence is very attractive for its compactness and cost effectiveness.

OFDM signals can be very sensitive to nonlinear distortion due to the instantaneous envelope of the its waveform generally following a Gaussian distribution with a large peak-to-average power ratio. The transmission of such a signal is largely limited by the linearity constraints in the modulation/demodulation processes, and by any nonlinear effects created on the unipolar IM/DD optical link. In order to avoid frequent clipping, a large bias power is usually needed, which unfortunately does not contribute to the receiver signal-to-noise ratio (SNR) at all. In fact, from a system design point of view and with a fixed power budget, a large biasing power can significantly penalize the overall performance [8].

Recent research has, hence, been focusing on reducing this bias via both distortion-introducing [8], [9] and distortion-less [10]–[12] techniques. In particular, in asymmetrically clipped optical OFDM (ACO-OFDM) systems [10], a tradeoff between power and spectral efficiency has been proposed, where by using only half of the available subchannels, a bipolar OFDM signal can be directly clipped with no additional biasing. As the clipping distortion is confined within certain frequencies, information can be recovered without any penalty. In general, however, these bias-reducing methods come with the cost of a lower spectral efficiency and/or a strong out-of-band clipping distortion.

In this paper, we propose a new tone-reservation-based optical OFDM (TR-OFDM) system that enables the unipolar transmission without any biasing and clipping. We start in Section II with a summary of the time- and frequency-domain representations of the clipping distortion. Using this theoretical analysis, a proposed TR-OFDM system model is introduced in Section III, where some of the subchannels/tones are reserved to produce a negative peak cancellation (NPC) signal. A simple filter-based estimator is developed in Section IV that enables distortion-less transmission without using any additional spectrum or bias, nor requiring any side-information at the receiver. An analysis of complexity and optimality is reported in Section V. We, then, show in Section VI that for a 40 Gb/s MMF link, a potential optical signal-to-noise ratio (OSNR) gain of 2.5 dB can be achieved for $\mathrm{BER} = 10^{-3}$.

II. Representation of the Clipping Distortion

Various types of IM/DD optical OFDM systems have been reported in the literature [2]–[8], [10], all with the same objective of using only simple intensity modulators and square-law detectors as the optical transceiver. Despite the detailed methodology differences, some basic concepts are shared among them. At the transmitter, incoming high-speed data are first split into a large number of low-speed datasets by a serial-to-parallel (S/P) converter before being encoded into quadrature amplitude modulation (QAM) symbols on N equally spaced subchannels. The complex symbols are then transformed into a time-domain signal via an IFFT, and the nth sample of the output is given by

$$x_n = \frac{1}{N}\sum_{m=0}^{N-1}\left(X_m^I + jX_m^Q\right)\exp\left(j\frac{2\pi}{N}nm\right) \tag{1}$$

where $X_m = X_m^I + jX_m^Q$ is the complex symbol modulated on the mth subchannel.

In order to focus on the effect of clipping, we look at the case where a real signal is produced. Therefore, as in the case of digital multitone modulation [13], we enforce conjugate symmetry by setting $X_m^I = X_{N-m}^I$ and $X_m^Q = -X_{N-m}^Q$ for all $0 < m < N/2$. For the case of complex output, an intermediate (IM) frequency can be inserted to generate real waveforms via RF I/Q modulation [14].

Since the sequence x_n is bipolar in general, without an additional electronic bias, received symbols can be deeply buried under clipping distortion [10], as phase information is not preserved during intensity modulation. In order to make x_n unipolar, a BAC process is normally employed. The unipolar time-domain samples $\tilde{x}_n$ that represents the distorted output signal sequence are

$$\tilde{x}_n = \begin{cases} 0, & x_n < -\gamma \\ x_n + \gamma, & x_n \geq -\gamma \end{cases} \tag{2}$$

where the biasing voltage is denoted by γ.

Many previous works have assumed a *sufficient* and variable dc offset is added to eliminate all possible clipping [5], [6]. Other works have used a fixed, yet large dc bias, to ensure an infrequent clipping from a statistical point of view [2]–[4], [7]. Recently, by using a linear minimum mean square error (LMMSE) approximation, the clipping distortion was effectively transformed into a deterministic attenuation and an uncorrelated noise. Consequently, an adaptive optimal biasing point was identified in our early work [8]. This method is not focused on eliminating all clipping distortion, but rather on providing "just enough" bias so that the clipping distortion is no longer the dominant noise source. Since the dc bias affects only the zeroth (dc) subchannel, which is already reserved along with the $(N/2)$th tone to produce conjugate symmetric signals, data transmission is enabled on all available subchannels.

In an effort to improve the power efficiency and eliminate the use of additional bias, i.e., $\gamma = 0$, a tradeoff has been studied in ACO-OFDM [10]. By loading only odd subchannels, the IFFT output sequence x_n becomes antiperiodic. For any sample in the first half of the time-domain signal, there exists a corresponding sample in the second half with the same magnitude but opposite sign, i.e., $x_n = -x_{n+(N/2)}$. Thus, even with "0-bias clipping,"[1] there is no information loss in transmission, and symbols can be recovered without any distortion penalty. This benefit comes at the cost of abandoning all the even subchannels and thereby halving the spectral efficiency.

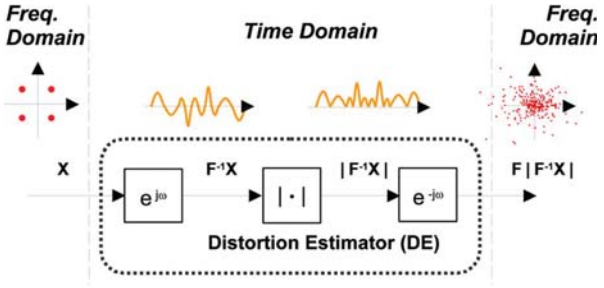

Fig. 1. Details of the DE with its input and output constellation diagrams and time-domain signals in between.

Notice, however, that with 0-bias clipping, i.e., $\gamma = 0$ in (2), the output distorted sequence $\tilde{x}_n$ has a relatively simply form, $\tilde{x}_n = 0.5(x_n + |x_n|)$. The distorted symbol after clipping becomes

$$\begin{aligned}\tilde{X}_m &= \frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{2}\left(x_n + |x_n|\right)\exp\left(-j\frac{2\pi}{N}mn\right) \\ &= \frac{1}{2}(X_m + D_m)\end{aligned} \tag{3}$$

where the clipping distortion D_m is equal to

$$\begin{aligned}D_m &= \frac{1}{N}\sum_{n=0}^{N-1}|x_n|\exp\left(-j\frac{2\pi}{N}mn\right) \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\left|\sum_{m=0}^{N-1}X_m\exp\left(j\frac{2\pi}{N}mn\right)\right| \\ &\quad\times\exp\left(-j\frac{2\pi}{N}mn\right).\end{aligned} \tag{4}$$

Although it exhibits a noise-like spectrum, D_m can actually be completely characterized by a simple distortion estimator (DE) that consists of a pair of Fourier transforms and an absolute value operation in the time domain, as shown in Fig. 1. By stacking D_m's and X_m's into vectors, the input–output relationship in (4) can be simply rewritten as

$$\mathbf{D} = \mathbf{F}|\mathbf{F}^{-1}\mathbf{X}| \tag{5}$$

where $\mathbf{F}$ is an $N \times N$ FFT matrix with its (i,k)th element defined as

$$[\mathbf{F}]_{ik} = \exp\left(-j2\pi\frac{ik}{N}\right) \tag{6}$$

and $\mathbf{F}^{-1}$ is an $N \times N$ IFFT matrix with its (i,k)th element defined as

$$[\mathbf{F}^{-1}]_{ik} = \frac{1}{N}\exp\left(j2\pi\frac{ik}{N}\right) \tag{7}$$

so that $\mathbf{F}\mathbf{F}^{-1} = \mathbf{I}$.

[1]With 0-bias clipping, all the negative data points of the IFFT output are simply replaced by zeros to form a unipolar signal.

In the frequency domain, the power spectral density (PSD) of the distortion signal also follows an interesting pattern. This is observed by denoting subchannels that are divisible by 2^k, but not 2^{k+1} as G_k. The N available subchannels can be divided into $\log_2 N$ sets as follows: [2]

$$\begin{aligned} G_0 &= \{1,\ 3,\ 5,\ 7,\ 9,\ 11,\ 13, \ldots, N-1\} \\ G_1 &= \{2,\ 6,\ 10,\ 14,\ 18, \ldots, N-2\} \\ G_2 &= \{4,\ 12,\ 20,\ 28, \ldots, N-4\} \\ &\vdots \\ G_k &= \{2^k,\ 3\cdot 2^k, \ldots, N-2^k\} \\ &\vdots \\ G_{\log_2 N-1} &= \left\{0, \frac{N}{2}\right\}. \end{aligned}$$

Consequently, the time-domain sample sequence x_n can be divided into $\log_2 N$ terms as well

$$x_n = x_n^0 + x_n^1 + x_n^2 + \cdots + x_n^{\log_2 N-1} \tag{8}$$

where each of these summands

$$x_n^k = \frac{1}{N}\sum_{u\in G_k} X_u \exp\left(j\frac{2\pi}{N}un\right) \tag{9}$$

is the contribution from the subchannels in G_k to the $\{x_n\}$ sequence.

According to this definition, it is easy to show that for all $n \in \{0,1,2,\ldots,(N)/(2^{k+1})-1\}$ and $\delta \in \{0,1,\ldots,2^{k+1}-1\}$

$$\begin{aligned} x^k_{n+\delta\frac{N}{2^{k+1}}} &= \frac{1}{N}\sum_{u\in G_k} X_u \exp\left[j\frac{2\pi}{N}u(n+\delta\frac{N}{2^{k+1}})\right] \\ &= \frac{1}{N}\sum_{u\in G_k} X_u \exp\left(j\frac{2\pi}{N}un\right)\exp\left(j\delta\pi\frac{u}{2^k}\right) \\ &= \frac{1}{N}\sum_{u\in G_k} X_u \exp\left(j\frac{2\pi}{N}un\right)\exp(j\delta\pi). \end{aligned} \tag{10}$$

Each of these x_n^k sequence are then both periodic as well as antiperiodic, i.e.,

$$x^k_{n+\delta\frac{N}{2^{k+1}}} = \begin{cases} -x_n^k, & \delta \text{ odd} \\ x_n^k, & \delta \text{ even} \end{cases} \tag{11}$$

with two simple example sequences shown in Fig. 2.

Furthermore, since the exponential terms of the FFT and IFFT operation in (6) and (7) are both circularly symmetric, when such a periodic signal x_n^k is 0-bias clipped, the resulting frequency-domain clipping distortion becomes

[2]Note that for simplicity, the zeroth subchannel is grouped with the $(N/2)$th subchannel together in $G_{\log_2 N-1}$, as these two must be reserved in the case of a real IFFT output.

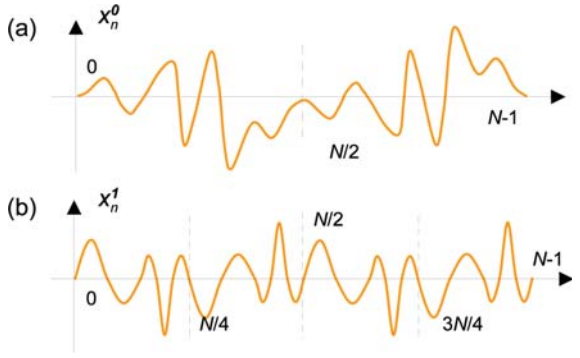

Fig. 2. Examples of time-domain waveforms when limited numbers of subchannels are loaded. (a) Antiperiodic waveform when G_0 is loaded. (b) Periodic, as well as antiperiodic, waveform when G_1 is loaded.

$$\begin{aligned} D_m^k &= \frac{1}{N}\sum_{\delta=0}^{2^{k+1}-1}\sum_{n=0}^{\frac{N}{2^{k+1}}-1}\left|x^k_{n+\delta\frac{N}{2^{k+1}}}\right| \\ &\quad\times\exp\left[-j\frac{2\pi}{N}\left(n+\delta\frac{N}{2^{k+1}}\right)m\right] \\ &= \frac{1}{N}\sum_{n=0}^{\frac{N}{2^{k+1}}-1}\sum_{\delta=0}^{2^{k+1}-1}\left|x_n^k\right| \\ &\quad\times\exp\left(-j\frac{2\pi nm}{N}\right)\exp\left(-j\frac{2m\pi\delta}{2^{k+1}}\right) \\ &= \frac{1}{N}\sum_{n=0}^{\frac{N}{2^{k+1}}-1}\left|x_n^k\right|\exp\left(-j\frac{2\pi nm}{N}\right)\cdot\Gamma(m) \end{aligned} \tag{12}$$

where $\Gamma(m)$ is a vector summation defined as

$$\Gamma(m)\overset{\triangle}{=}\sum_{\delta=0}^{2^{k+1}-1}\exp\left(-j\frac{2m\pi\delta}{2^{k+1}}\right). \tag{13}$$

Notice that when m is not divisible by 2^{k+1} (i.e., $m \in G_u$ for all $u \le k$), the exponential terms in (13) are also circularly symmetric. On the other hand, when $m \in G_v$, where $v > k$, all exponential terms become equal to 1. Hence, we can write

$$\Gamma(m) = \begin{cases} 0, & m \in \bigcup_{i=0}^{k} G_i \\ 2^{k+1}, & m \in \bigcup_{i=k+1}^{\log_2 N-1} G_i. \end{cases} \tag{14}$$

In other words, since $|x_n^k|$ is a periodic signal with at most $(N/2^{k+1})$ independent time-domain samples, its Fourier transform D_m^k will have at most $(N)/(2^{k+1})$ nonzero frequency terms as well, as the cardinality

$$\left|\bigcup_{i=k+1}^{\log_2 N-1} G_i\right| = \sum_{i=k+1}^{\log_2 N-1}\frac{N}{2^{i+1}} = \frac{N}{2^{k+1}}. \tag{15}$$

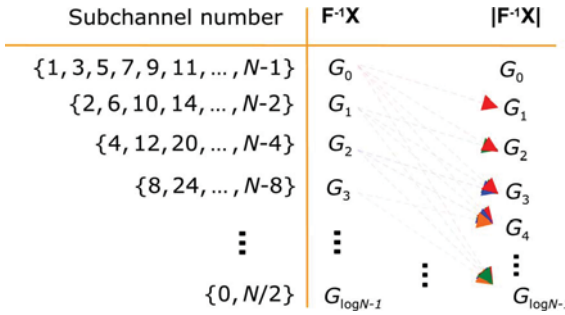

Fig. 3. Frequency-domain PSD transition diagram of the clipping distortion.

Intuitively speaking, as shown in Fig. 3, this means that when the G_k subchannels are loaded, the 0-bias clipping distortion falls only onto the $G_{k+1}, \ldots, G_{\log_2 N-1}$ subchannels. Therefore, when bit-loading is limited to only a single subchannel set G_k, the 0-bias clipped signal can easily be recovered as $D_m^k = 0$ for all $m \in G_k$. ACO-OFDM is therefore a special case where only G_0 (odd) subchannels are used.

On the other hand, the distortion power on each subchannel has only a limited, and not necessarily equal, number of possible sources. The distortion on the G_k subchannels depends only on whether data symbols have been loaded onto subchannel sets $G_0, G_1, \ldots, G_{k-1}$ and can be fully characterized by the time-domain DE described in Fig. 1. For example, evaluating $D_2, D_6, D_{10}, \ldots$ requires the knowledge of $X_1, X_3, X_5, \ldots$ only, and will remain exactly the same even if the other subchannels are also loaded. Similarly, D_m for $m \in G_2$ depends on the symbols loaded onto the G_0 and G_1 subchannels, i.e., it depends on the values of $3N/4$ elements in the $\mathbf{X}$ vector. We further note that G_0 subchannels are not affected by any distortion.[3]

The objective of this paper is to generalize the approaches mentioned previously, where there are only two choices of reserved subchannels, i.e., either 2 tones or $(N/2)$ tones. In the next section, we show that TR-OFDM provides a more flexible tradeoff between spectral and power efficiencies. In the frequency domain, the effect of clipping distortion is limited within certain subchannels that are carefully chosen. In the time domain, the additional bias γ is completely eliminated without introducing any in-band or out-of-band distortion.

III. Principle of Tone Reservation for IM/DD Optical OFDM Systems

The proposed TR-OFDM system diagram is shown in Fig. 4. At the transmitter, by stacking X_m's and x_n's into vectors, we denote the $N \times 1$ QAM symbol vector as $\mathbf{X} = [X_0 \cdots X_{N-1}]^T$ and its corresponding $N \times 1$ time-domain signal vector as $\mathbf{x} = [x_0 \cdots x_{N-1}]^T$. The modulation process in (1) can simply be expressed by the following matrix operation:

$$\mathbf{x} = \mathbf{F}^{-1}\mathbf{X}. \tag{16}$$

We again start with the case where a real signal is produced. With the proposed TR-OFDM systems, in order to reduce the cost of additional biasing, some subchannels are reserved, i.e., no data symbols will be loaded onto the corresponding elements of $\mathbf{X}$. Instead, an $N \times 1$ NPC symbol vector $\mathbf{C} = [C_0 \cdots C_{N-1}]^T$ is evaluated by the decision feedback distortion estimator (DFDE) and combined with $\mathbf{X}$ before being fed to the input of the IFFT. The reserved tones set is denoted by $\mathfrak{R}$, while the data-bearing subchannels then form the set $\mathfrak{D}$. The receiver has no side information about the details of the NPC vector, and to avoid any ambiguity, the set of reserved subchannels and the set of data-bearing subchannels must be disjoint, i.e., $\mathfrak{D} \cap \mathfrak{R} = \emptyset$.

[3]This interesting property means that with multiple 0-bias clippers, all clipping distortion can be completely characterized and consequently, iteratively compensated for. An example a system with a successive decoder at the receiver can be found in our early work in [11].

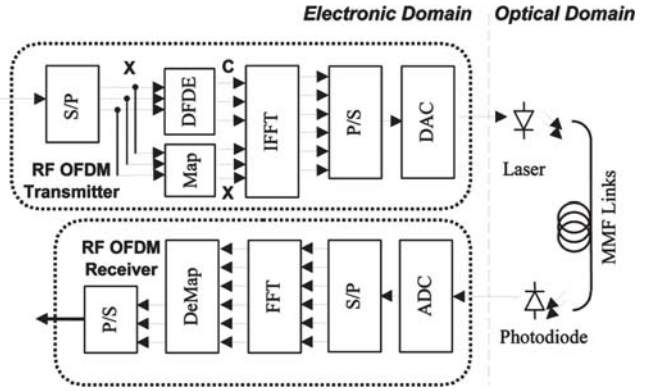

Fig. 4. Conceptual diagram of the TR-OFDM system: DFDE: decision feedback DE. For more details of DFDE, see Fig. 6.

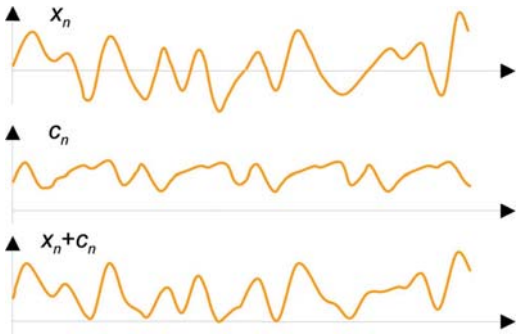

Fig. 5. Example of the time-domain waveforms for the data-bearing signal $\mathbf{x}$, the peak cancellation signal $\mathbf{c}$ and the unipolar drive signal $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{c}$.

With the help of the additional NPC symbol vector and its corresponding time-domain peak cancellation signal $\mathbf{c}$, the adjusted IFFT output $\tilde{\mathbf{x}}$ then becomes

$$\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{c} = \mathbf{F}^{-1}(\mathbf{X} + \mathbf{C}). \tag{17}$$

The objective is to minimize the additional bias that is required to make the $\tilde{x}_n$ sequence nonnegative. This is equivalent to maximizing the minimum value among all negative elements in $\tilde{\mathbf{x}}$. This is written as the following constrained optimization problem:

$$\min_{\mathbf{C} \in \mathbb{R}^N} \| |\mathbf{F}^{-1}(\mathbf{X} + \mathbf{C})| - \mathbf{F}^{-1}(\mathbf{X} + \mathbf{C}) \|_\infty \tag{18}$$
$$\text{subject to the nonsharing condition :}$$
$$\forall m \in \mathfrak{D}, C_m = 0.$$

where $\| \cdot \|_\infty$ is the L_∞ norm (i.e., maximum) defined on the finite-dimensional coordinate space.

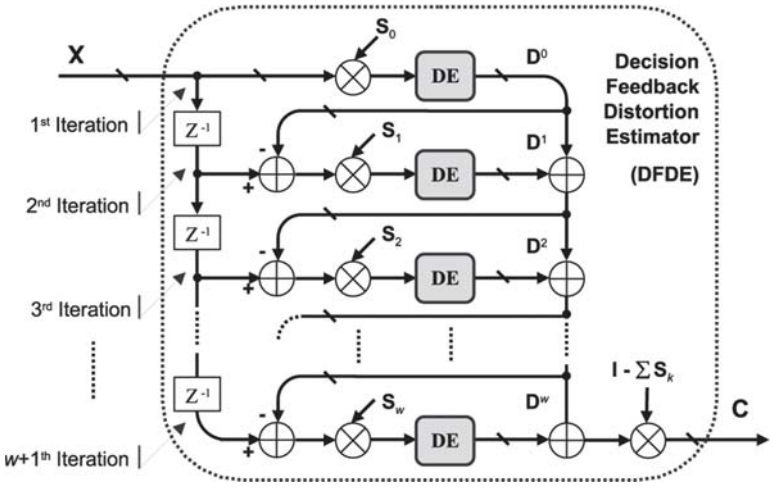

Fig. 6. Details of the DFDE. Z^{-1}: step delay; DE: distortion estimator as in Fig. 1.

In the optimal case(s), $\tilde{x}_n$ can be made unipolar as shown in Fig. 5. Consequently, unlike in previous approaches such as ACO-OFDM or DC-OFDM, no biasing and/or clipping is required. Nevertheless, under this framework, ACO-OFDM and DC-OFDM can be viewed as two special cases where the number of reserved subchannels is equal to 2 and $N/2$, respectively.

A cyclic prefix can be added so that both intersymbol interference and interchannel interference (ICI) due to delay spread/modal dispersion can be eliminated. Training symbols are inserted periodically for both time/frequency synchronization and estimating the coefficients for the 1-tap subchannel equalizers at the receiver. A digital-to-analog converter (DAC) then directly converts this sequence into analog driving waveforms which are free from both in-band and out-of-band distortion. In general, we note that although the analog signal may still dip below 0 due to DAC and filtering, the effect is negligible [15]. This unipolar waveform is then used to drive a linear optical modulator. For simplicity, we assume the use of an ideal intensity modulator, where the instantaneous output optical power is a replica of the corresponding electrical drive signal.

On the receiver side, after the photodiode, we assume appropriate demodulation can be performed. As all the clipping distortion has been fully compensated for at the transmitter and no transmission of side information on the NPC signal is required, the TR-OFDM receiver has a fairly simply structure, which has no major differences compared to a typical RF-OFDM receiver.

IV. Decision Feedback Distortion Estimator

The principle of tone reservation was introduced in [16] and [17], where a certain set of subchannels (known to both the transmitter and the receiver) is reserved to offset any large peaks from the data-bearing subchannels. In our scenario, these subchannels shall be equally spread among N candidates and are denoted again by the set $\mathfrak{R}$, while the data-bearing subchannels then form the set $\mathfrak{D}$. We further limit the choice of these two sets to be

$$\begin{aligned}\mathfrak{D} &= G_0 \cup G_1 \cup G_2 \cdots \cup G_w \\ \mathfrak{R} &= G_{w+1} \cup G_{w+2} \cdots \cup G_{\log_2 N-1}.\end{aligned} \tag{19}$$

As we shall see later in this section, with this setup, the clipping distortion and useful information can be easily separated.

Reserving subchannels obviously sacrifices some capacity, but our goal is to keep this overhead small while compensating for as many negative peaks as possible to minimize the need for additional electronic bias or extra bandwidth. Furthermore, a proposed method should also be scalable for all w from $0, 1, \ldots, \log_2 N - 2$, where even as small as two reserved subchannels are sufficient for distortion compensation. Existing approaches like DC-OFDM and ACO-OFDM then become the special cases, where w is equal to its minimum or maximum values.[4]

Determining an optimal $\mathbf{C}$ can be considered to be a digital filter design problem in the transform domain with a constrained set of taps (the reserved set $\mathfrak{R}$). In other words, design a signal in the frequency domain that optimizes some criteria in the time domain.

Reflecting on the time- and frequency-domain analysis of the clipping distortion in Section II, we notice that for x_n^k samples generated by loading symbols onto only the G_k subchannels, there naturally exists a simple peak cancellation signal $|x_n^k|$ that can be easily evaluated. Furthermore, as shown in Fig. 3, while x_n^k has nonzero PSD components on the G_k subchannels and $|x_n^k|$ has nonzero PSD components only on the $G_{k+1}, G_{k+2}, \ldots, G_{\log_2 N-1}$ subchannels. Hence, when bit-loading is limited to a single subchannel group G_k, the optimal solution to this filter design problem is precisely the DE introduced in Section II. Mathematically, if we denote the $N \times N$ mask matrix

$$\mathbf{S}_k = \begin{pmatrix} S_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & S_N \end{pmatrix} \tag{20}$$

where its trace elements $S_i = 1$ for all $i \in G_k$, and 0 otherwise. $\mathbf{X}^k \triangleq \mathbf{S}_k\mathbf{X}$ then contains nonzero values on the G_k elements only.

[4]It is worth to note that DC-OFDM utilizes only the zeroth tone in $G_{\log N-1}$, hence is not the best choice in general, while ACO-OFDM, as we later show, provides an optimal result for the case of $w = 0$.

The modified IFFT output is then

$$\begin{aligned}\tilde{\mathbf{x}}^k &= \mathbf{F}^{-1}(\mathbf{X}^k + \mathbf{F}|\mathbf{F}^{-1}\mathbf{X}^k|) \\ &= \mathbf{x}^k + |\mathbf{x}^k| \succcurlyeq \mathbf{0} \end{aligned} \tag{21}$$

where the sign "$\succcurlyeq$" is used to denote the element-wise larger or equal between two vectors. Using this formulation, we again see that ACO-OFDM [10] is just a special case where only the G_0 subchannels are loaded.

In a more general case, when w groups of subchannels are loaded, the peak cancellation signal can be evaluated in an iterative fashion based on the PSD transition diagram in Fig. 3. One optimal solution to this problem that guarantees a nonnegative IFFT output is shown in Fig. 6. The proposed DFDE filter has $w+1$ tapped delay lines in total, each producing a signal $\mathbf{D}^k$ that serves as a peak cancellation symbol for the G_k subchannels. Following (12), $\mathbf{D}^k$ has only $(N)/(2^{k+1})$ nonzero elements; hence, signals in all previous iterations, i.e., $\mathbf{X}^u, \forall u < k$, will not be affected.

The algorithm starts by assuming only the G_0 subchannels are loaded. The corresponding peak cancellation signal $\mathbf{D}^0$ can easily be identified, which has non-zero values on all $G_1, G_2, \ldots, G_{\log N-1}$ subchannels. In the next iteration, the G_1 subchannels are now added into $\mathfrak{D}$ and can be loaded with data. Since this set is no longer available for NPC, the corresponding elements of $\mathbf{D}^0$ must be offset. Therefore, for each iteration, the peak cancellation signals from all previous groups should be compensated for before the input of the DE. $\mathbf{S}_k$ is then used to select the relevant elements for the kth iteration, where the input of the DE is always limited to a single subchannel group. The corresponding peak cancellation signal is then evaluated:

$\mathbf{D}^{-1} = \mathbf{0}, k = 0;$

while $k \leq w$ **do**

$\mathbf{Y}_k = \mathbf{X} - \sum_{i=0}^{k} \mathbf{D}^{i-1};$

$\mathbf{D}^k = \mathbf{F}|\mathbf{F}^{-1}\mathbf{S}_k\mathbf{Y}_k|;$

$k = k + 1;$

end while

Finally, after $w + 1$ iterations, all cancellation signals are summed up and relevant elements are picked to form the peak cancellation vector:

$$\mathbf{C} = \left(\mathbf{I} - \sum_{i=0}^{w} \mathbf{S}_i\right) \sum_{k=0}^{w} \mathbf{D}^k \tag{22}$$

where $\mathbf{I}$ is the identity matrix. Notice that with this setup, the nonsharing constraint is always guaranteed. Due to the unique PSD structure of the peak cancellation signal, we observe that $\forall k = 0, 1, 2, \ldots, w$

$$\mathbf{S}_i\mathbf{D}^k = \mathbf{0}, \quad \forall i \leq k. \tag{23}$$

TABLE I
COMPUTATIONAL COMPLEXITY INVOLVED IN EACH ITERATION

# of Iter.	$\mathbf{Y}_k$	$\mathbf{S}_k\mathbf{Y}_k$	$\mathbf{F}^{-1}$	$\lvert\cdot\rvert$	$\mathbf{F}$	$\sum \mathbf{D}^k$
1st	N	N	$\frac{N}{2}\log\frac{N}{2}$	$\frac{N}{2}$	$\frac{N}{2}\log\frac{N}{2}$	N
2nd	N	N	$\frac{N}{4}\log\frac{N}{4}$	$\frac{N}{4}$	$\frac{N}{4}\log\frac{N}{4}$	N
⋮	⋮	⋮	⋮	⋮	⋮	⋮
kth	N	N	$\frac{N}{2^k}\log\frac{N}{2^k}$	$\frac{N}{2^k}$	$\frac{N}{2^k}\log\frac{N}{2^k}$	N
⋮	⋮	⋮	⋮	⋮	⋮	⋮
log$N$-1th	N	N	2	2	2	N

Thus, after some simple manipulation, (22) can be re-organized into

$$\begin{aligned}\mathbf{C} &= \sum_{k=0}^{w}\left(\mathbf{I} - \sum_{i=0}^{w}\mathbf{S}_i\right)\mathbf{D}^k = \sum_{k=0}^{w}\mathbf{D}^k - \sum_{k=0}^{w-1}\sum_{i=k+1}^{w}\mathbf{S}_i\mathbf{D}^k \\ &= \sum_{k=0}^{w}\mathbf{D}^k - \sum_{k=1}^{w}\sum_{i=0}^{k-1}\mathbf{S}_k\mathbf{D}^i = \sum_{k=0}^{w}\left(\mathbf{D}^k - \sum_{i=0}^{k-1}\mathbf{S}_k\mathbf{D}^i\right)\end{aligned} \tag{24}$$

where the last equation is due to the fact that $\mathbf{D}^{-1} = \mathbf{0}$. The adjusted IFFT output in this case then becomes

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{F}^{-1}(\mathbf{X} + \mathbf{C}) \\ &= \mathbf{F}^{-1}\left[\sum_{k=0}^{w}\left(\mathbf{S}_k\mathbf{X} + \mathbf{D}^k - \sum_{i=0}^{k}\mathbf{S}_k\mathbf{D}^{i-1}\right)\right] \\ &= \mathbf{F}^{-1}\left[\sum_{k=0}^{w}\mathbf{S}_k\left(\mathbf{X} - \sum_{i=0}^{k}\mathbf{D}^{i-1}\right) + \mathbf{D}^k\right] \\ &= \mathbf{F}^{-1}\left[\sum_{k=0}^{w}\mathbf{S}_k\mathbf{Y}_k + \mathbf{D}^k\right] \\ &= \sum_{k=0}^{w}\left(\mathbf{F}^{-1}\mathbf{S}_k\mathbf{Y}_k + |\mathbf{F}^{-1}\mathbf{S}_k\mathbf{Y}_k|\right) \succcurlyeq \mathbf{0}.\end{aligned} \tag{25}$$

V. COMPLEXITY AND OPTIMALITY ANALYSIS

In terms of complexity, a worst case scenario for the proposed DFDE is summarized in Table I. In this case, only subchannels belonging to $G_{\log N-1}$ are reserved, i.e., the zeroth and $(N/2)$th subchannels. As a result, $\log N - 1$ iterations are required.

Throughout the evaluation, since we are only interested in the sum, $\sum_{i=0}^{k-1}\mathbf{D}^i$, rather than the individual values of $\{\mathbf{D}^k\}$, only N additions are required to calculate $\mathbf{Y}_k$. Furthermore, as $\mathbf{S}_k$ is actually a "sparse identity matrix," evaluating $\mathbf{S}_k\mathbf{Y}_k$ does not require any matrix multiplications, but only N comparisons. Therefore, the complexity of the DFDE filter depends on the IFFT and FFT operations inside the DE.

IFFT/FFT operations for the kth iteration can also be implemented with much fewer multiplications due to the input vector for the IFFT having only $N/2^k$ nonzero elements. Thus, after evaluating the first $N/2^k$ samples, the remaining samples can easily be generated using the periodic and antiperiodic properties reported in (11). For the FFT, since we are only interested

in the $N/2^k$ nonzero elements of $\mathbf{D}^k$, this calculation can be greatly simplified as well. Consequently, computational complexity for the kth iteration is then reduced from $\mathcal{O}(N \log N)$ to $\mathcal{O}((N/2^k)\log(N/2^k))$.

The overall computational complexity is then

$$\begin{aligned}
&\mathcal{O}\left(\sum_{k=1}^{\log N-1} \frac{N}{2^k}\log\frac{N}{2^k}\right)\\
&= \mathcal{O}\left(N\log N \sum_{k=1}^{\log N-1}\frac{1}{2^k} - N\sum_{k=1}^{\log N-1}\frac{k}{2^k}\right)\\
&= \mathcal{O}\left(N\log N - N\sum_{i=1}^{\log N-1}\sum_{k=i}^{\log N-1}\frac{1}{2^k}\right)\\
&= \mathcal{O}\left(N\log N - N\sum_{i=1}^{\log N-1}\frac{1}{2^{i-1}} - \frac{2}{N}\right)\\
&= \mathcal{O}(N\log N). \qquad (26)
\end{aligned}$$

Therefore, by using variable-size IFFTs/FFTs, the algorithm's total complexity can be reduced by a factor of $\log N$, from $\mathcal{O}(\log N \cdot N \log N)$ to $\mathcal{O}(N \log N)$.

As this estimator is designed to accommodate all possible values of w from 0 up to $\log N - 2$, a tradeoff between complexity and spectral efficiency should be considered in determining the numbers of iterations. From a practical system design point of view, it seems to be less cost effective to allow a large number of iterations, i.e., choosing a large w value. For each additional iteration, one more subchannel group is freed up for data transmission. However, the number of these extra subchannels gets smaller as the cardinality of the subchannel group G_k is $N/2^{k+1}$. Thus, the spectral efficiency increment tends to saturate, not to mention the accumulated processing delay and hardware cost.

In most cases, a DFDE designed for only a couple of iterations should be sufficient. For example, with only four iterations, distortion can be completely offset for the case where 93.75% of the subchannels are loaded with data. For these scenarios, using only one pair of fixed-size FFT/IFFT can provide a better tradeoff and the computational complexity is still $\mathcal{O}(N \log N)$.

Apart from the complexity issue, the other drawback of striving for a small number of reserved tones is that a higher average transmit power may be required. Therefore, for practical implementation of an efficient TR-OFDM system, the number of reserved subchannels/iterations should be carefully chosen so that both the electronic and optical power of the adjusted transmit signal $\tilde{x}_n$ is optimized for a desired system.

We denote σ_X^2 and σ_C^2 as the average transmit power for the data and NPC symbols, and σ_x^2 and σ_c^2 are the corresponding variances in the time domain, respectively. As only a limited number of subchannels will be loaded to generate $\mathbf{x}$ and $\mathbf{c}$. The relationship among the aforementioned quantities are then

$$\sigma_x^2 = \left(1 - \frac{1}{2^{w+1}}\right)\sigma_X^2, \quad \sigma_c^2 = \frac{1}{2^{w+1}}\sigma_C^2 \qquad (27)$$

where w is the index of the last subchannel group involved in data transmission as defined in (19). As for the mean time-domain values, they become $\mu_x = 0$ and $\mu_c = C_0/N$.

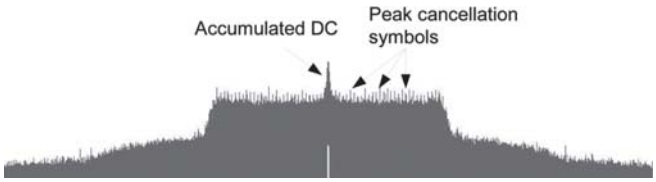

Fig. 7. Example of the optical spectrum of the TR-OFDM signal.

Fig. 7 shows an example of the optical spectrum for the proposed TR-OFDM system, where both in-band and out-of-band distortion is eliminated. Nevertheless, there is still an accumulated dc component on the zeroth subchannel, i.e., $C_0 > 0$. Apart from a large C_0, the other NPC symbols C_m, in general, also require a higher transmit power than the data-bearing symbols in $\mathbf{X}$, i.e., $\sigma_C^2 > \sigma_X^2$. Consequently, spikes can be observed across the frequency domain on the $\mathfrak{R}$ subchannels as shown in Fig. 7.

The average electrical power is proportional to $E[\tilde{x}^2]$, which is equal to

$$\begin{aligned}
\mathrm{E}\left[\tilde{x}^2\right] &= \mu_{\tilde{x}}^2 + \sigma_{\tilde{x}}^2 = \mu_c + \sigma_x^2 + \sigma_c^2\\
&= \mu_c^2 + \left(1 - \frac{1}{2^{w+1}}\right)\sigma_X^2 + \frac{\sigma_C^2}{2^{w+1}}. \qquad (28)
\end{aligned}$$

In general, μ_c, σ_X^2, and σ_C^2 are all monotonic functions of w. When more iterations are required, C_0 gradually builds up due to constructive contributions from all $\mathbf{D}_k$ signals on that tone (all of which are positive real numbers). Hence, μ_c tends to increase when w increases.

A large w also results in a decreased number of reserved subchannels where a large accumulated offset power must be compensated for. The variance of $\mathbf{Y}_k$, as well as σ_C^2, then tend to increase. On the other hand, with a fixed target rate, higher order constellations must be used when large number of subchannels are reserved to build up the filter. Thus, σ_X^2 decreases when more subchannels are freed up for data transmission with more iterations.

We observe that (28) is not a monotonic function of w. In the regions where w is small, the effect of σ_X^2 is dominating. For example, the first iteration itself can improve the spectral efficiency by 50%. In general, the improvement of one iteration equals the sum of all subsequent iterations could offer. As more subchannels become available for data, the electronic power starts to decrease as a result of a smaller average QAM constellation, as well as a higher flexibility for adaptive rate and power optimization [2], [3]. When w is already very large, the potential to improve the spectral efficiency quickly saturates. In this region, σ_X^2 is almost fixed, while $\sigma_C^2/2^{w+1}$ is negligible, and μ_c then becomes the dominating factor. In this case, with an increased w, μ_c^2 is monotonically increasing, leading to an unfavorably high transmit power.

Therefore, depending on the target bit error rate (BER), there exists an optimal w^* for which the system performance can be maximized. For practical implementation, based on this optimal choice, some subchannels are *intentionally reserved* from conveying data symbols in order to reduce the required electronic power.

Optical power, on the other hand, is proportional to $E[\tilde{x}_n] = \mu_c$ and is monotonically increasing with increasing w. If we denote a new zero-mean random signal as

$$s_n = \tilde{x}_n - \mu_c \tag{29}$$

μ_c is indeed the sufficient bias that is required to make s_n nonnegative. From this point of view, TR-OFDM is not only an extension of ACO-OFDM with an improved spectral efficiency, but is also an extension of DC-OFDM systems, where some subchannels are reserved to reduce the amount of sufficient bias and hence improve the power efficiency.

In general, for any bias ratio $\mu \leq \mu_c$, by utilizing the LMMSE approximation we proposed in [8], the effect of clipping can be transformed into a deterministic attenuation K and an uncorrelated noise z_n

$$\tilde{s}_n \approx K s_n + \mu + z_n \tag{30}$$

where $K = 1 - Q(\gamma)$ and $\gamma = \mu/\sigma_s$ is the normalized bias. The well-known Q-function is defined as the integral over normal pdf

$$Q(\nu) \triangleq \int_{\nu}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}} d\tau. \tag{31}$$

The variance of this additional clipping noise equals to [8]

$$\sigma_z^2 = \sigma_s^2 \left\{ (1+\gamma^2)[Q(\gamma) - Q^2(\gamma)] - \frac{1}{2\pi} e^{-\gamma^2} - \frac{\gamma}{\sqrt{2\pi}} e^{-\frac{\gamma^2}{2}} [1 - 2Q(\gamma)] \right\}. \tag{32}$$

Choosing a large dc bias reduces both the frequency and depth of clipping the negative peaks of s_n. But the resulting signal energy is significantly increased. Due to the circular symmetry of the "exp" components in receiver FFT, this unfavorable dc component does not contribute at all to the receiver SNR [8]. For any practical transmitter with a limited output optical power of σ_s^2, the biased signal must then be subjected to a power normalization, where the effective receiver SNR becomes [8]

$$\text{SNR} = \frac{\sigma_s^2/\gamma^2}{\sigma_z^2/\gamma^2 + \sigma_n^2} = \frac{\sigma_s^2}{\sigma_z^2 + \sigma_n^2 \cdot \gamma^2} \tag{33}$$

and σ_n^2 is the receiver noise figure.

Equation (33) is also not monotonic with respect to w. When γ is small, clipping noise σ_z^2 is the dominant term in the denominator. Increasing the biasing power in this region certainly results in an increased SNR. However, with excessive biasing, clipping noise is negligible and the effective SNR starts to decrease as the input signal contains a very large and unfavorable dc component. Hence, there exists an optimal biasing point, γ^*, for which the received SNR can be maximized. Applying the fast searching algorithm in [8], we observe that γ^* depends only on the SNR and when w is large, it is possible that $\gamma^* < \mu_c$. In other words, the performance can be improved by *deliberately clipping* (D-clipping) the TR-OFDM signal, as long as the clipping distortion is not the dominant noise source.

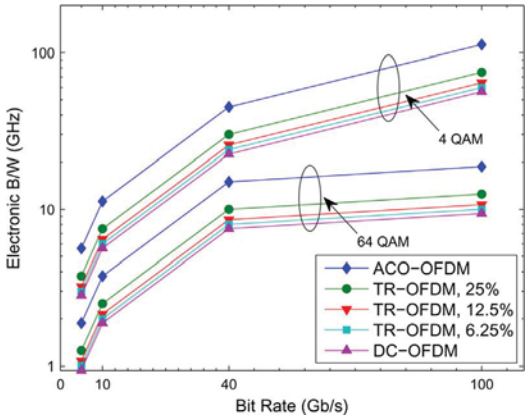

Fig. 8. Bandwidth requirements for different system with uniform bit-loading.

In general, we conclude that it is possible, but not always beneficial, to enable data transmission on all available subchannels in the frequency domain. It is also possible, but not always beneficial, to completely eliminate the clipping distortion in the time domain. As we show in the forthcoming simulation section, given a target BER, an optimal w^* can be easily determined via simulation.[5] For TR-OFDM systems with higher quality-of-service requirements and longer transmission distances, a larger w^* should be employed, where D-clipping is capable of offering a significant additional power savings.

VI. Simulation of a 40 Gb/s IM/DD Multimode Transmission

A detailed numerical simulation is carried out using VPI TransmissionMaker and MATLAB. The original 40 Gb/s data are first divided and mapped onto 1024 subchannels. If desired, the output unipolar OFDM signal after IFFT can be deliberately clipped to further improve the power efficiency. This real-baseband OFDM waveform is then interpolated to drive an external linear optical modulator. The DFB laser source is centered at 193.1 THz with a linewidth of 1 MHz and a 0 dB·m output power. The signal is center launched to a single span of 10 km, 50/125 μm graded-index MMF with n_1 =1.45, $\Delta = 0.01$, and a 0.2 dB/km loss. For channel estimation, we use four training symbols every 450 data symbols.

Fig. 8 shows the electronic bandwidth required for different systems to achieve a given set of rates with uniform bit-loading across all available subchannels. In all cases, TR-OFDM with different percentages of reserved tones achieves a significant spectral efficiency improvement compared to ACO-OFDM. It is slightly worse than DC-OFDM due to overhead and the associated delay.

To reduce this overhead, the delay/performance trade-off is compared in Fig. 9, where we examine the idea of intentionally reserving tones and looking for the optimal choice of w. In order to conduct a fair comparison, the maximum electronic bandwidth for all systems is set to be the same and adaptive bit-loading with variable QAM constellation is used to guarantee the target bit rate. Initially, when more data subchannels

[5]Theoretical characterizations of both μ_c and σ_c^2 are also possible, and therefore, w^* and γ^* can also be obtained analytically. This will be the subject of a future work.

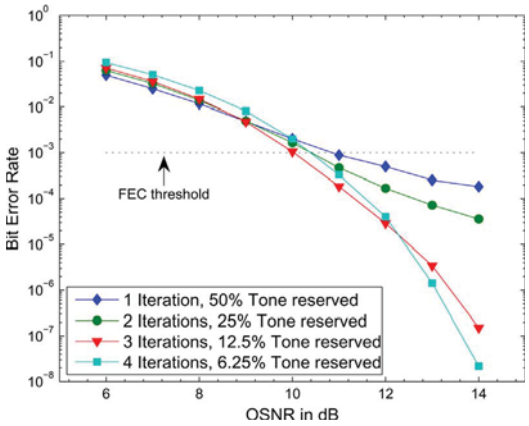

Fig. 9. BER versus OSNR for different numbers of reserved tones.

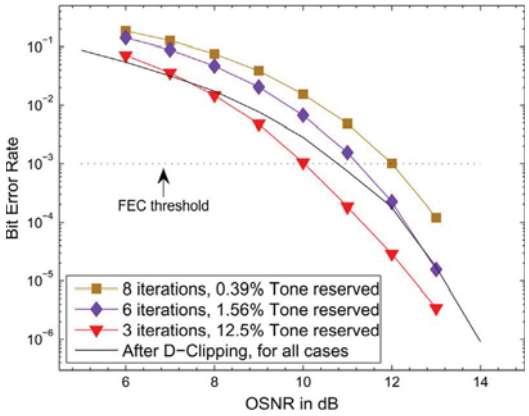

Fig. 10. BER improvement via deliberate clipping.

are available, the BER tends to decrease, but this is not always the case with more iterations. The peak-cancellation signal requires a higher average transmit energy, which leads to a reduced power efficiency and eventually a higher BER. We found that using three iterations seems to be the best choice for our scenario with a BER target of 10^{-3}. It is also worth noting that due to the nature of IM/DD optical OFDM, nonlinear impairment cannot be fully compensated and will accumulate along the transmission link. With larger QAM constellations and/or longer transmission distances, TR-OFDM become more sensitive to this nonlinearity. Hence, the tails of the BER curves for TR-OFDM with 1 or 2 iterations saturate in the high OSNR regions.

The effect of deliberately clipping is investigated in Fig. 10. With an increased w, the effect of an increased μ_c begins to dominate. γ^* depends only on the SNR; it gives the optimal amount of biasing to make clipping distortion negligible compared to receiver noise. As a result, a common bound is plotted where the dc term is fixed at γ^* for all TR-OFDM systems. The improvement by employing D-clipping ranges from significant, when μ_c is always larger than γ^*, to marginal or even negligible, when $\mu_c < \gamma^*$ in most cases. For any w value, the optimized BER then follows the minimum of this D-clipped bound and its original BER curve.

Finally, the power efficiencies across different systems are compared in Fig. 11. With a maximum electronic bandwidth

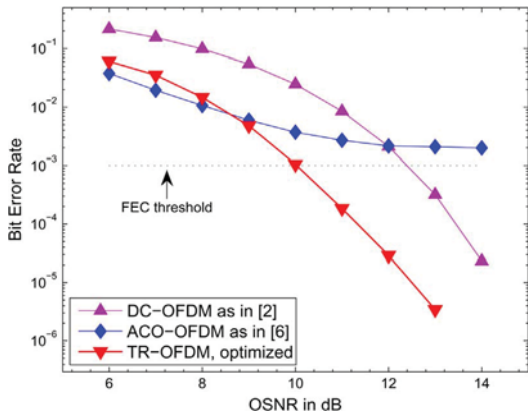

Fig. 11. BER versus OSNR for systems with different biasing strategies.

constraint, ACO-OFDM has no choice but to use larger QAM constellation, which exacerbates the residual ICI [18]. But for very low OSNRs, it outperforms the others due to the fact that it has the lowest dc power. DC-OFDM has the smallest average QAM size but requires more power due to the additional biasing. Compared to these two, TR-OFDM achieves a significant power efficiency improvement and gives more than 2.5 dB OSNR gain around the forward error correction (FEC) threshold. With an increased transmission distance, as discussed earlier, less number of subchannels should be reserved so that the average constellation size is smaller. Nevertheless, the proposed TR-OFDM scheme is still capable of providing a steady performance improvement. However, due to the increased computational complexity, simple DC-OFDM scheme may be preferred in some cases.

VII. Conclusion

Based on the concept of tone reservation, a nonlinear tapped-delay-line filter can be constructed to optimally offset all clipping distortion. OFDM signals can then be directly converted into unipolar driving currents without using any additional power or bandwidth. The proposed TR-OFDM system is not only an extension of ACO-OFDM with an improved spectral efficiency, but also is an extension of DC-OFDM, where some subchannels are reserved to reduce the amount of additional bias and hence improve the power efficiency. TR-OFDM eliminates the use of both biasing and clipping and hence offers both a more flexible system design and an improved bandwidth-power tradeoff. Not only simple to run in terms of complexity, we further notice that the proposed scheme requires only a couple of simple electronic components to perform the appropriate digital signal processing. Hence, the additional hardware cost is also marginal.

We notice that it is possible, but not always beneficial, to enable data transmission on all available subchannels, or to completely eliminate the clipping distortion. Hence, differing from previous approaches where the focus is solely on maximizing spectral utilization or minimizing additional bias, we show that the system performance is maximized by locating a best tradeoff point, where both the number of reserved subchannels and the size of the dc component in the resulting unipolar signal are carefully optimized.

REFERENCES

[1] D. Cunningham, "Multimode fiber data communication," in *Proc. Opt. Fiber Commun./Nat. Fiber Opt. Eng. Conf.*, 2008, pp. 1–31.

[2] H. Yang, S. Lee, E. Tangdiongga, C. Okonkwo, H. van den Boom, F. Breyer, S. Randel, and A. Koonen, "47.4 Gb/s transmission over 100 m graded-index plastic optical fiber based on rate-adaptive discrete multitone modulation," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 352–359, Feb. 2010.

[3] J. M. Tang, P. M. Lane, and K. A. Shore, "High-speed transmission of adaptively modulated optical OFDM signals over multimode fibers using directly modulated DFBs," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 429–441, Jan. 2006.

[4] X. Jin, R. Giddings, E. Hugues-Salas, and J. Tang, "End-to-end real-time demonstration of 128-QAM encoded optical OFDM transmission with a 5.25bit/s/Hz spectral efficiency over IMDD systems," in *Proc. Opt. Fiber Commun./Nat. Fiber Opt. Eng. Conf.*, 2010, pp. 1–3.

[5] J. Carruthers and J. Kahn, "Multi-subcarrier modulation for non-directed wireless infrared communication," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 538–546, Apr. 1996.

[6] S. Randel, F. Breyer, and S. C. J. Lee, "High-speed transmission over multimode optical fibers," in *Proc. Opt. Fiber Commun./Nat. Fiber Opt. Eng. Conf.*, 2008, pp. 1–3.

[7] O. Gonzalez, R. Perez-Jimemnez, S. Rodriguze, J. Rabadan, and A. Ayala, "OFDM over indoor wireless optical channel," *IEE Proc. Optoelectron.*, vol. 152, no. 4, pp. 199–204, 2005.

[8] L. Chen, B. Krongold, and J. Evans, "Performance evaluation of optical OFDM systems with nonlinear clipping distortion," in *Proc. Int. Conf. Commun.*, Jun. 2009, pp. 1–5.

[9] D. Chanda, A. Sesay, and B. Davies, "Performance of clipped OFDM signal in fiber," in *Proc. IEEE Can. Conf. Elect. Comput. Eng.*, May 2004, vol. 4, pp. 2401–2404.

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

[11] L. Chen, B. Krongold, and J. Evans, "Diversity combining for asymmetrically clipped optical OFDM in IM/DD channels," in *Proc. IEEE Int. Global Commun. Conf.*, Dec. 2009, pp. 1–6.

[12] L. Chen, B. Krongold, and J. Evans, "Successive decoding of antiperiodic OFDM signals in IM/DD optical channel," in *Proc. Int. Conf. Commun.*, May 2010, pp. 1–6.

[13] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Sel. Areas Commun.*, vol. 9, no. 6, pp. 895–908, Aug. 1991.

[14] A. J. Lowery, L. B. Du, and J. Armstrong, "Performance of optical OFDM in ultralong-haul WDM lightwave systems," *J. Lightw. Technol.*, vol. 25, no. 1, pp. 131–138, Jan. 2007.

[15] D. Barros and J. Kahn, "Optical modulator optimization for orthogonal frequency-division multiplexing," *J. Lightw. Technol.*, vol. 27, no. 13, pp. 2370–2378, Jul. 2009.

[16] A. Gatherer and M. Polley, "Controlling clipping probablity in DMT transmission," in *Proc. Asilomar*, 1997, pp. 578–584.

[17] J. Tellado, "Peak to average power reduction for multicarrier modulations," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2000.

[18] D. Barros and J. Kahn, "OFDM versus OOK with MLSD for IM/DD systems," in *Proc. Opt. Fiber Commun./Nat. Fiber Opt. Eng. Conf.*, 2010, pp. 1–3.

Liang Chen received the B.E. degree from Zhejiang University, Hangzhou, China, in 2003, and the M.S. degree from the University of Melbourne, Parkville, Australia, in 2006, where he is currently working toward the Ph.D. degree.

During the winter of 2010, he was with the National Institute of Informatics, Tokyo, Japan, as a Visiting Researcher. His current research interests include signal processing for both wireless and optical orthogonal frequency-division multiplexing systems.

Mr. Chen was the recipient of Best Paper Awards from the 2006 Australian Communications Theory Workshop and the 2006 European Wireless Conference, respectively. In 2011, he also received a Victoria Fellowship from the Victoria State Government of Australia

Brian Krongold received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, in 1995, 1997 and 2001, respectively.

He was as a Research Assistant at the Coordinated Science Laboratory from 1995 to 2001. From December 2001 to December 2004, he was a Research Fellow in the ARC Special Research Centre for Ultra-Broadband Information Networks in the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia. He received an ARC Postdoctoral Research Fellowship and held this from 2005 to 2008. Since January 2008, he has been a Senior Lecturer at the University of Melbourne. During the summer of 1994, he interned for Martin Marietta at the Oak Ridge National Laboratory, Oak Ridge, Tennessee. From January to August 1995, he consulted at Bell Laboratories in Middletown, NJ. During the summer of 1998, he was with the Electronics and Telecommunications Research Institute, Taejon, South Korea, under a National Science Foundation summer research grant. During the first half of 2011, he was on sabbatical at Alcatel-Lucent Bell Laboratories, Murray Hill, NJ. His current research interests include multicarrier communications systems, energy-efficient communications, and signal processing for optical communications.

Dr. Krongold received second prize in the Student Paper Contest at the 2001 Asilomar Conference on Signals, Systems, and Computers, and the Best Paper Award at the 2006 European Wireless Conference. His work on active constellation extension for orthogonal frequency-division multiplexing peak-to-average power ratio reduction is part of the DVB-T2 digital video broadcast standard.

Jamie Evans was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, Newcastle, in 1992 and 1993, respectively, and the M.S. and Ph.D. degrees from the University of Melbourne, Parkville, Australia, in 1996 and 1998, respectively, both in electrical engineering.

From March 1998 to June 1999, he was a Visiting Researcher in the Department of Electrical Engineering and Computer Science, University of California, Berkeley. He returned to Australia to take up a position as a Lecturer at the University of Sydney, Sydney, Australia, where he stayed until July 2001. Since then, he has been with the Department of Electrical and Electronic Engineering, University of Melbourne, where he is currently a Professor. His research interests include communications theory, information theory, and statistical signal processing with a focus on wireless communications networks.

Dr. Evans received the University Medal upon graduation from University of Newcastle. He also received the Chancellors Prize for excellence for his Ph.D. thesis.