PAPR Analysis and Mitigation Algorithms for Beamforming MIMO OFDM Systems

Ying-Che Hung and Shang-Ho (Lawrence) Tsai, *Senior Member, IEEE*

***Abstract*—Beamforming (or precoding) techniques have been widely adopted in modern MIMO OFDM systems. Using beamforming can significantly improve the receive SNR of OFDM systems. However, the combination of transmit signals after beamforming deteriorates the peak-to-average power ratio (PAPR), which has long been considered a major issue of OFDM systems. High PAPR not only complicates the design of the power amplifier, but also increases power consumption. In this paper, we theoretically analyze the PAPR performance of MIMO OFDM systems that adopt either one of the two popular beamforming schemes, *i.e.* MRT (maximum ratio transmission) and EGT (equal gain transmission). The analysis considers different numbers of channel taps after sampling. The results may provide important reference for practical designs when evaluating the required power amplifiers and power consumption. Moreover, the theoretical results show that MRT OFDM systems generally perform much worse than EGT OFDM systems in terms of PAPR. Furthermore, motivated from the derived results, PAPR reduction algorithms are proposed for both MRT OFDM and EGT OFDM systems. It is worth to mention that for MRT OFDM systems, the proposed algorithm can improve both PAPR and bit error rate; for EGT OFDM systems, the proposed algorithm improves PAPR while it only slightly degrades bit error rate.**

***Index Terms*—MIMO OFDM, beamforming, precoding, peak-to-average power ratio, PAPR, low power, maximum ratio transmission, MRT, equal gain transmission, EGT, extreme value theory, beta distribution.**

I. Introduction

MULTIPLE-INPUT Multiple-output orthogonal frequency division multiplexing (MIMO OFDM) is widely used in current and next generation broadband wireless communications, because it can provide high data rate transmission over multipath fading channels [1],[2]. Among the MIMO techniques, beamforming (or precoding) has been widely adopted in communication standards, *e.g.,* LTE, WiMAX and Wi-Fi applications, because it can achieve full diversity, which results in a reliable transmission. It is known that OFDM systems suffer from high peak-to-average power ratio (PAPR). High PAPR leads to high effort in designing the power amplifier (PA) in order to maintain a wide linear region for preventing signal clipping, which therefore increases not only hardware complexity but also power consumption. The PAPR issue is worse when beamforming is applied in OFDM systems, because the dynamic range of the signals increases after beamforming [3].

Many methods have been proposed for reducing the PAPR including deliberate clipping, companding, probabilistic methods, and coding, see *e.g.,* [4]-[11]. These methods may more or less distort signals and decrease the data rate. For example, the most straightforward PAPR reduction method is via clipping peak signals before passing them to the PA [5]. However, clipping signals induces in-band and out-of-band distortion and requires additional signal processing techniques to reconstruct the received signals. Another category of methods to reduce the PAPR is through probabilistic schemes such as partial transmit sequence (PTS) [6],[7], selected mapping (SLM) [8] and sign adjustment [9],[10]. The objective of probabilistic methods is to reduce the probability that peak power exceeds a certain PAPR threshold. These methods demand that the transmitter sends side information to the receiver. Consequently, the data rate decreases due to the side information, and the transmitted signals cannot be correctly reconstructed if the transmitted side information is polluted. Moreover, although there has been extensive research for PAPR on OFDM systems, to the best of the authors knowledge, few studies have been conducted in analyzing the PAPR for beamforming MIMO OFDM systems and developing corresponding PAPR reduction methods [11], which are especially important in practice, since most modern communication systems adopt beamforming MIMO OFDM techniques and Green communications is a worldwide trend to save power consumption. The discussion above motivates us to explore how PAPR increases if beamforming is adopted in OFDM systems, and then propose the corresponding PAPR mitigation methods.

In this paper, we analyze the PAPR performance for single-user MIMO OFDM systems adopting either one of the two most commonly used beamforming schemes, *i.e.,* maximum ratio transmission (MRT) [12] and equal gain transmission (EGT) [13]-[15]. MRT is the optimal beamforming scheme in terms of receive SNR, but the PA design for MRT is more complicated than EGT. It has been shown that the maximum SNR loss between MRT and EGT is only 1.05 dB under Rayleigh fading channels (see [16] and [17]). Here we use the Extreme Value Theory [18],[19] and order statistics [20] to derive the CCDF (complementary cumulative distribution function) of PAPR for EGT and MRT OFDM systems, and make some insightful observations. First, we found that EGT OFDM systems have constant power property for different OFDM blocks and different RF transmit branches. Thus the PAPR characteristic can be approximately obtained by simultaneously considering M_t unprecoded OFDM systems,

Manuscript received March 11, 2013; revised September 18, 2013; accepted January 29, 2014. The associate editor coordinating the review of this paper and approving it for publication was L. Liu.

The authors are with Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan (e-mail: yingjhe.ece95g@g2.nctu.edu.tw, shanghot@alumni.usc.edu).

This work is supported by the National Science Council (NSC), Taiwan, R.O.C. Cooperative Agreement No. 102-2221-E-009-017-MY3.

Digital Object Identifier 10.1109/TWC.2014.031914.130347

where M_t is the number of transmit antennas. In other words, the PAPR for EGT OFDM systems is the same as that of an unprecoded OFDM systems simultaneously transmitting M_t data streams. On the other hand, the PAPR performance of MRT OFDM systems is deeply affected by the number L of channel taps after sampling. MRT OFDM systems have the worst PAPR performance when $L = 1$, *i.e.,* flat fading, and achieve comparable PAPR as EGT OFDM systems when $L = N$, where N is the number of subcarriers. Since the case $L = N$ rarely occurs in practice, MRT OFDM systems in general have much worse PAPR performance than EGT OFDM systems; taking $L = 1$ for instance, the performance gap can be up to 5.3 dB when the CCDF is at 10^{-3}. Note that the maximum SNR loss of 1.05 dB between MRT and EGT derived in [16] and [17] can be easily extended to MRT OFDM and EGT OFDM systems. Hence, without proper PAPR reduction algorithms, designers may be in a dilemma to determine which performance index to pursuit, receive SNR or PAPR? Therefore, based on the derived results, we further propose PAPR reduction methods for both MRT OFDM and EGT OFDM systems. It is worth to emphasize that, unlike conventional PAPR reduction methods, there is no need to send side information from the transmitter to the receiver in the proposed algorithms. In addition, for MRT OFDM systems, the proposed algorithm not only can reduce the PAPR but also can improve the bit error rate performance. This satisfying result is obtained thanks to the motivation from the derived results. The proposed algorithm attempts to adjust the power at some subcarriers after beamforming as closely as possible. Since the subcarrier power is equalized in a certain level, both the PAPR and the bit error rate performance are improved simultaneously. For EGT OFDM systems, the proposed algorithm can reduce the PAPR, while at the same time it only slightly degrades bit error rate performance. Finally, simulation results are provided to show the accuracy of the derived theoretical PAPR results and the performance improvement achieved when using the proposed PAPR mitigation algorithms.

The rest of this paper is organized as follows: The system model and the problem formulation are given in Section II. We analyze the PAPR distribution for EGT and MRT OFDM systems in Section III. The proposed PAPR reduction algorithms are introduced in Section IV. Simulation results and conclusions are provided in Section V and VI, respectively.

Notations. All vectors are in lowercase boldface and matrices are in uppercase boldface. $(\cdot)^T$ and $(\cdot)^\dagger$ denote the transpose and the conjugate transpose of a vector. $\mathbb{E}[\cdot]$ and var$[\cdot]$ denote expectation and variance respectively. $(\cdot)^*$ denotes the complex conjugate of $(\cdot)$. $||\cdot||_2$ is the vector 2 norm. $Beta(\alpha,\beta)$ and $B(\alpha,\beta)$ represent the beta distribution and beta function with parameters α and β, respectively. log represents the natural log operation.

II. System Model and Problem Formulation

The block diagram of a MIMO OFDM system with transmit beamforming and receive combining, using M_t transmit antennas, M_r receive antennas and N subcarriers, is shown in Fig. 1. Assuming that the number L of channel taps after sampling is shorter than the length of the cyclic prefix (CP), the channel gain of the kth subcarrier after the discrete Fourier transform (DFT) can be represented by a M_r-by-M_t channel matrix, denoted by $\mathbf{H}_k$. At the transmitter, the kth symbol x_k is multiplied by a beamforming vector $\mathbf{g}_k = \left[g_k^{(0)}\ g_k^{(1)}\ \cdots\ g_k^{(M_t-1)}\right]^T$, where $g_k^{(i)}$ is the ith element of the beamforming vector at the kth subcarrier. At the receiver, the combined signal of the kth subcarrier after multiplying by the combining vector $\mathbf{z}_k = \left[z_k^{(0)}\ z_k^{(1)}\ \cdots\ z_k^{(M_r-1)}\right]^T$ can be expressed as

$$r_k = \mathbf{z}_k^\dagger \mathbf{H}_k \mathbf{g}_k x_k + \mathbf{z}_k^\dagger \mathbf{n}_k, \quad 0 \le k \le N-1, \tag{1}$$

where $\mathbf{n}_k \in \mathbb{C}^{M_r \times 1}$ is the noise vector after the DFT whose entries are independent and identically distributed (i.i.d.) complex Gaussian random variable with zero mean and variance N_0.

In the beamforming OFDM system, $\mathbf{g}_k$ and $\mathbf{z}_k$ are designed according to $\mathbf{H}_k$ to maximize the effective channel gain γ_k. That is,

$$\gamma_k = \max_{\{\mathbf{g}_k, \mathbf{z}_k\}} |\mathbf{z}_k^\dagger \mathbf{H}_k \mathbf{g}_k|^2. \tag{2}$$

In this paper we analyze the PAPR performance for MRT and EGT, the two most commonly used beamforming schemes. The complex baseband signal at the ith transmit branch after inverse-DFT (IDFT) can be expressed as

$$s^{(i)}(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k g_k^{(i)} e^{j\omega_k t}, \tag{3}$$

where $\omega_k = \frac{2\pi}{NT_c}(k - \frac{N-1}{2})$, and $T_c \in (0,\infty)$ is the subcarrier interval. The interval of an OFDM block is defined as $T = NT_c$. $\{x_k,\ k = 0, \ldots N-1\}$ is the modulated OFDM symbol with independent real part x_k^R and imaginary part x_k^I, which satisfy $\mathbb{E}[x_k^R] = \mathbb{E}[x_k^I] = 0$, and $\mathbb{E}[(x_k^R)^2] = \mathbb{E}[(x_k^I)^2] = \varepsilon_s/2$. Without loss of generality, ε_s is assumed to be 1. For description convenience, we define two types of average power: one is block average power P_{av} and the other is long-term average power $\overline{P}_{av}$. Referring to (3), the block average power is the average power of an OFDM block at a specific transmit branch, *i.e.,* $P_{av} = \frac{1}{T}\int_0^T |s^{(i)}(t)|^2 dt$; while the long-term average power is the expectation of the block average power, *i.e.,* $\overline{P}_{av} = \mathbb{E}[P_{av}]$. The baseband PAPR of the beamforming OFDM system is defined as

$$\max_{0 \le i \le M_t-1} \{\text{PAPR}^{(i)}\} = \max_{0 \le i \le M_t-1} \left\{ \max_{0 \le t \le T} \frac{|s^{(i)}(t)|^2}{\overline{P}_{av}} \right\}. \tag{4}$$

High PAPR is a main disadvantage of OFDM systems. The design effort of power amplifiers (PA) increases as PAPR increases. Moreover, high PAPR also leads to high power consumption because more complicated PA needs to be used. In beamforming OFDM systems, the PAPR becomes higher than the OFDM systems without beamforming. As a result, using beamforming in OFDM systems complicates the PA design and increases power consumption. In the following sections we analyze the PAPR performance of both MRT OFDM and EGT OFDM systems, and then propose corresponding PAPR reduction algorithms.

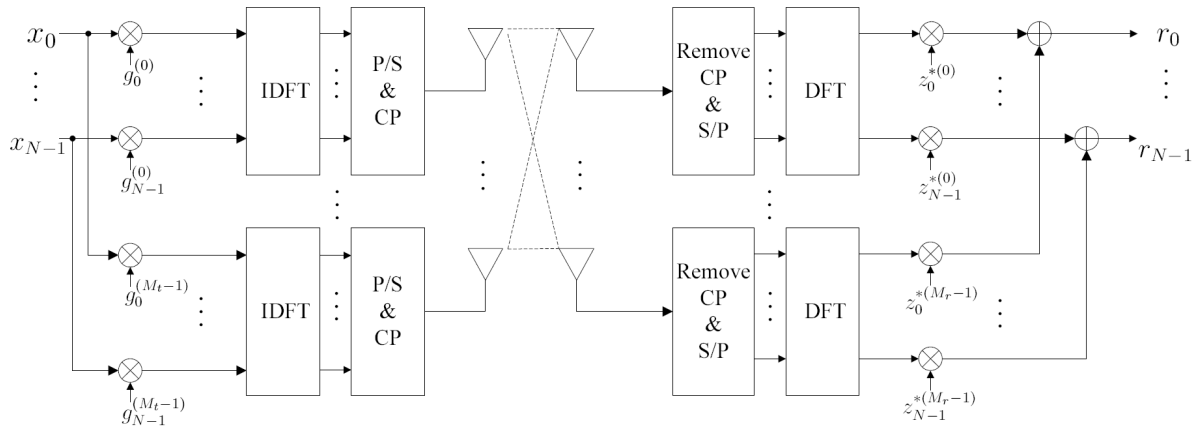

Fig. 1. A MIMO OFDM system with transmit beamforming and receive combining.

III. Analysis of Peak-to-Average Power Ratio

The distribution of the PAPR bears statistical characteristics in OFDM systems and is often expressed in terms of the complementary cumulative distribution function (CCDF). Recently, several analytical approaches to obtain the statistical distribution of the PAPR in OFDM systems have been proposed, such as empirical approximation and level-crossing rate approximation of the peak distribution [21]-[23]. In this paper, we analyze the PAPR by using the Extreme Value Theory [18], which can provide a simple, accurate, and rigorous expression of PAPR performance for beamforming OFDM systems. We briefly introduce the Extreme Value Theory below.

Extreme Value Theory: Let $\zeta(t)$ and $\eta(t)$ $(t > 0)$ be two independent stationary Gaussian processes with zero mean, unit variance and identical covariance function

$$r(t) = \text{Cov}(\zeta(s), \zeta(s+t)) = \text{Cov}(\eta(s), \eta(s+t)),$$

where $r(t)$ can be expressed by the expansion

$$r(t) = 1 - \lambda \frac{t^2}{2} + o(t^2) \ \text{ as } t \to 0. \tag{5}$$

Suppose that $\zeta(t)$ and $\eta(t)$ have continuous differentiable sample paths, with $\text{var}(\zeta'(t)) = \text{var}(\eta'(t)) = \lambda = -r''(0)$. Then $\chi^2(t) = \zeta^2(t) + \eta^2(t)$ is a stationary $\chi^2(2)$-process with continuously differentiable sample paths. Suppose further that

$$r(t)\log(t) \to 0, \ \text{ as } t \to \infty. \tag{6}$$

Then we have

$$\Pr\left\{\max_{0\le t\le T} \chi^2(t) \le d\right\} \to e^{-\tau}, \tag{7}$$

if

$$T\mu(d) = T\sqrt{\frac{\lambda d}{2\pi}} e^{-\frac{d}{2}} \to \tau, \ \ T \to \infty. \tag{8}$$

It is easy to show, based on (7) and (8), that as $T \to \infty$,

$$\Pr\left\{a_T\left(\max_{0\le t\le T} \chi^2(t) - b_T\right) \le x\right\} \to \exp\left(-e^{-x}\right), \tag{9}$$

where $a_T = 1/2$, and $b_T = 2\log T + \log\log T + \log(\lambda/\pi)$.

Applying the Extreme Value Theory, we first derive the PAPR at a specific transmit branch for both MRT and EGT OFDM systems, and then extend the results to the case where all transmit branches are considered. We assume that the transmit branches are uncorrelated†.

A. PAPR analysis for EGT OFDM systems

EGT is a supoptimal beamforming scheme compared to MRT. In MISO channels, an EGT vector that can most closely achieve the performance of the MRT is $\mathbf{g}_k = e^{-j\angle \mathbf{h}_k^*}/\sqrt{M_t}$, where $\angle \mathbf{h}_k^*$ is the phase of $\mathbf{h}_k^*$. For MIMO systems, however, there is no closed-form solution available for such EGT vector. Thus the authors in [15] and [17] have proposed the EGT solutions for MIMO channels.

Now let us first consider the PAPR at a specific transmit branch for EGT OFDM systems, and then extend the results to all transmit branches. Since the elements of $\mathbf{g}_k$ have constant amplitude $1/\sqrt{M_t}$, the power of the signal $s^{(i)}(t)$ in (3) is $1/M_t$ and the block average power P_{av} is thus $1/M_t$.

Lemma 1 *Consider EGT OFDM systems with constant-magnitude modulation, the block average power P_{av} is constant for all transmit branches and equal to the long-term average power $\overline{P}_{av}$.*

Proof: Please see Appendix A. ■

It is worth to point out that even if $\mathbf{x}$ does not have constant magnitude, $\frac{1}{N}\mathbf{x}^\dagger\mathbf{x}$ tends to a constant when N is sufficiently large, due to the Law of Large Numbers (LLN). Because P_{av} and $\overline{P}_{av}$ are constant, the PAPR analysis for EGT OFDM systems is much simpler than that for MRT OFDM systems as we will see later.

Lemma 2 *Consider the complex baseband signal $X_N(t) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} A_k e^{j\omega_k t}$, where $\omega_k = \frac{2\pi}{NT_c}\left(k - \frac{N-1}{2}\right)$, and $\{A_k,\ k = 0, \ldots, N-1\}$ is an i.i.d. complex random sequence with bounded real part A_k^R and imaginary part A_k^I; in addition, $\mathbb{E}\left[A_k^R\right] = \mathbb{E}\left[A_k^I\right] = 0$, $\mathbb{E}\left[(A_k^R)^2\right] = \mathbb{E}\left[(A_k^I)^2\right] = \sigma^2$, and $\mathbb{E}\left[A_k^R A_k^I\right] = 0$. Then for any closed and finite interval $T \subseteq R$, as $N \to \infty$, $X_N(t)$ converges in distribution to a zero-mean stationary complex Gaussian random process $X(t)$*

†In fact, the transmit branches are correlated because the beamforming vector is from the singular vector, and the elements of the singular vector have correlation. Nevertheless, the theoretical results obtained by assuming uncorrelation among transmit antennas are still quite consistent with the simulation results in most of the interested PAPR regime.

whose real and imaginary parts are independent and each part has autocorrelation function $r(\tau) = \sigma^2 sinc\left(\frac{\tau}{T_c}\right)$.

Proof: Please see [24] for details. ■

Lemma 3 *The CCDF of the PAPR at the ith transmit branch for EGT OFDM systems can be approximated as*

$$\Pr\left\{PAPR^{(i)} > \xi\right\} \approx 1 - \exp\left\{-e^{-\xi} N \sqrt{\frac{\pi}{3}\log N}\right\}. \quad (10)$$

Proof: Please see Appendix B. ■

Note that the result in (10) is the same with that for unprecoded OFDM systems (see [21]). That is, the PAPR obtained by only considering a specific transmit branch of EGT OFDM systems is the same as that for unprecoded OFDM systems. This is not surprising because the average power (denominator of the PAPR term) is approximately constant for sufficiently large N as mentioned in Lemma 1; also, the signal power of each subcarrier (numerator of the PAPR term) is only scaled down by $1/M_t$. These lead to the same result as unprecoded OFDM systems. Moreover, the PAPR of EGT MISO OFDM has the same characteristic as that of EGT MIMO OFDM because the magnitude of both generated beamforming vectors is equal to $1/\sqrt{M_t}$ and therefore contribute the same analytical results. It is worth to point out that although the theoretical results obtained by applying the Extreme Value Theory use the condition that $N \to \infty$, the derived results match the simulation results very well when $N \geq 256$.

Theorem 1 *Considering all transmit branches, the CCDF of the PAPR for EGT OFDM systems can be approximated as*

$$\Pr\left\{PAPR_{EGT} > \xi\right\} \approx 1 - \left\{\exp\left\{-e^{-\xi} N \sqrt{\frac{\pi}{3}\log N}\right\}\right\}^{M_t}. \quad (11)$$

Proof: Considering all transmit branches, the PAPR can be expressed as

$$\Pr\left\{\text{PAPR}_{\text{EGT}} > \xi\right\} = 1 - \Pr\left\{\max_{0 \leq i \leq M_t - 1} \text{PAPR}^{(i)} < \xi\right\}. \quad (12)$$

Since the transmit branches from different antennas are assumed to be mutually independent, utilizing the results of order statistics yields (11). ■

B. PAPR analysis for MRT OFDM systems

MRT is the optimal beamforming scheme, and can achieve 1.05 dB better receive SNR than EGT [17]. Unlike the aforementioned desirable property of EGT in Lemma 1, MRT has different average power for different transmit branches and different OFDM blocks because $(\mathbf{G}^{(i)})^{\dagger}\mathbf{G}^{(i)} \neq \frac{1}{M_t}\mathbf{I}$. Hence MRT does not have constant block average power and the long-term average is needed to identify the operation region of power amplifier. The power variation of MRT unavoidably increases the PAPR. As a result, it complicates the design of the PA and increases power consumption. Moreover, the PAPR analysis for MRT OFDM systems is more complicated than that for EGT OFDM systems because *1)* P_{av} is no longer a constant, and *2)* different numbers of channel taps after sampling L lead to different MRT vectors and consequently different PAPR values.

To obtain the PAPR of MRT OFDM systems, again, we first focus on a specific transmit branch, and then extend the results to all transmit branches. The MRT beamforming vector can be obtained from the right singular vector corresponding to the maximum singular value in MIMO/MISO channels. The PAPR statistics for MIMO and MISO channels should be the same, because their right singular vectors are both with the conditional Haar distribution [25]. Since the statistical characteristic of the term $P_{av} = \frac{1}{N}\sum_{k=0}^{N-1} |g_k^{(i)}|^2$ is used to derive the PAPR, the resulting PAPR statistics for MRT MISO OFDM and MRT MIMO OFDM are the same. Therefore we derive the PAPR results of MRT OFDM systems via MISO channels below.

The complex baseband signal after the IDFT at the ith transmit branch is given by (3). Since the amplitude of the MRT vector is different for different OFDM blocks, the block average power of a specific transmit branch changes and its distribution is related to the power of MRT element, *i.e.* $|g_k^{(i)}|^2$. Moreover, the value of L affects the correlation of $g_k^{(i)}$ between different subcarriers. Therefore, we first consider two extreme scenarios, *i.e.,* $L = 1$ and $L = N$. More specifically, we show that the correlation coefficient of $g_k^{(i)}$ between different subcarriers is equal to 1 as $L = 1$; while that approaches to 0 when $L = N \to \infty$. For $L = 1$ and $L = N$, the Extreme Value Theorem can be applied to obtain the PAPR of MRT OFDM systems. For other values of L that are not equal to 1 or N, it is complicated to analyze the corresponding correlation coefficient of $g_k^{(i)}$ between different subcarriers. To overcome this, some assumptions and approximations are needed for obtaining insightful results for the PAPR of MRT OFDM systems. Let us discuss these cases separately as follows:

1) $L = 1$ (flat fading channel): For $L = 1$ and the same transmit branch index i, $|g_k^{(i)}|^2$ for different subcarriers k is equal to a constant c. Moreover, for description convenience, let us temporarily ignore the index i of the transmit branch. Hence $|g_k|^2 = c$ for $0 \leq k \leq N-1$. From Lemma 2, we know that the real part $s^R(t)$ and the imaginary part $s^I(t)$ of $s(t)$ are independent Gaussian processes with autocorrealtion function

$$\begin{aligned} r(\tau) = \mathbb{E}\left[s^R(t)s^R(t+\tau)\right] &= \mathbb{E}\left[s^I(t)s^I(t+\tau)\right] \\ &= \frac{c}{2}\text{sinc}(\frac{\pi}{T_c}). \end{aligned}$$

It is clear that the block average power $P_{av} = c$ and therefore $\sqrt{2}s^R(t)/\sqrt{P_{av}}$ and $\sqrt{2}s^I(t)/\sqrt{P_{av}}$ also satisfy conditions (5) and (6) with $\lambda = \frac{1}{3}(\frac{\pi}{T_c})^2$. Since P_{av} varies for different OFDM blocks, the PAPR characteristic is related to the distribution of P_{av}.

From the Extreme Value Theory, the CDF of the baseband PAPR for a specific transmit branch can be approximated to

$$\begin{aligned} &\Pr\left\{\max_{0 \leq t \leq T} \frac{1}{\overline{P}_{av}}\left[(s^R(t))^2 + (s^I(t))^2\right] \leq \xi\right\} \\ &= \int_C \Pr\left\{\max_{0 \leq t \leq T} \chi^2(t) \leq \frac{2\overline{P}_{av}\xi}{c}\middle| c\right\} f_C(c)dc \\ &\to \int_C \exp\left(-e^{-y}\right) f_C(c)dc, \end{aligned} \quad (13)$$

where $y = a_T(\frac{2\xi\overline{P}_{av}}{c} - b_T)$. Let $q = \frac{\overline{P}_{av}}{c}$. Since $T = NT_c$, the CCDF of the PAPR can be approximated to

$$\Pr\left\{\text{PAPR}^{(i)} > \xi\right\} \approx 1 - \int_Q \exp\left\{-e^{-q\xi}N\sqrt{\frac{\pi}{3}\log N}\right\} f_Q(q)dq. \quad (14)$$

From (14), the PAPR of MRT OFDM systems can be theoretically evaluated if we know the probability density function (PDF) of q, which is given by the following lemma.

Lemma 4 *Under Rayleigh fading channels, the power of the MRT element $|g_k^{(i)}|^2$ follows a beta distribution with parameter $(\alpha, \beta) = (1, M_t - 1)$,* i.e., *$|g_k^{(i)}|^2 \sim Beta(1, M_t - 1)$.*

Proof: $|g_k^{(i)}|^2 = \frac{|h_k^{(i)}|^2}{||\mathbf{h}_k||_2^2}$. Since $h_k^{(i)}$ is the DFT of the multipath channel, $h_k^{(i)} \sim \mathcal{CN}(0, \frac{L}{2N})$ and $|h_k^{(i)}|^2 \sim \Gamma(1, \frac{L}{2N})$. Assume that $|h_k^{(i)}|^2$ and $|h_k^{(j)}|^2$ are independent for $0 \le i, j \le M_t - 1,\ i \ne j$. Then the distribution of $|g_k^{(i)}|^2$ can be obtained as follows [19]

$$\frac{|h_k^{(i)}|^2}{||\mathbf{h}_k||_2^2} = \frac{|h_k^{(i)}|^2}{|h_k^{(0)}|^2 + \cdots + |h_k^{(M_t-1)}|^2} = \frac{\Gamma(1, \frac{L}{2N})}{\Gamma(1, \frac{L}{2N}) + \Gamma(M_t - 1, \frac{L}{2N})} \sim Beta(1, M_t - 1).$$

■

Using Lemma 4, the PDF of $q = \frac{\overline{P}_{av}}{c}$ can be and expressed as [19]

$$f_Q(q) = \frac{1}{\mathrm{B}(1, M_t - 1)q^2\overline{P}_{av}}\left(1 - \frac{1}{q\overline{P}_{av}}\right)^{M_t - 2}, \quad (15)$$

where $\frac{\overline{P}_{av}}{M_t} \le q \le \infty$. From the above discussion, we have the following theorem:

Theorem 2 *For MRT OFDM systems in flat fading channels, when $N \to \infty$, the CCDF of PAPR can be approximated by*

$$\Pr\{PAPR_{MRT} > \xi\} \approx 1 - \left(\Pr\left\{PAPR^{(i)} < \xi\right\}\right)^{M_t}, \quad (16)$$

where $\Pr\left\{PAPR^{(i)} < \xi\right\}$ can be obtained by (14) and (15).

2) $L = N$: If L equals to the number N of subcarriers and $N \to \infty$, the equivalent channel coefficient at the ith transmit branch $\left\{h_k^{(i)},\ 0 \le k \le N - 1\right\}$ is an i.i.d. complex Gaussian random variable due to the LLN. Therefore $\left\{|g_k^{(i)}|^2,\ 0 \le k \le N - 1\right\}$ are i.i.d. $Beta(1, M_t - 1)$. Note that generally L is much smaller than N in practice. However, discussing this case helps gain more insight into the PAPR relationship between MRT and EGT.

Lemma 5 *If the number of subcarriers $N \to \infty$, the block average power P_{av} is $1/M_t$ and is equal to the long-term average power $\overline{P}_{av}$.*

Proof: Since $|g_k^{(i)}|^2$ is i.i.d. $Beta(1, M_t - 1)$, the block average power at the ith transmit branch is given by

$$P_{av} = \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}|g_k^{(i)}|^2 = \mathbb{E}[|g_k^{(i)}|^2] = \frac{1}{M_t}.$$

Since the block average power of different OFDM blocks is equal to $1/M_t$, the long-term average power $\overline{P}_{av}$ is equal to $1/M_t$ as well. ■

Lemma 6 *Consider the complex baseband signal $X_N(t) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}A_k e^{j\omega_k t}$, where $\omega_k = \frac{2\pi}{NT_c}\left(k - \frac{N-1}{2}\right)$, and $\{A_k,\ k = 0, \ldots, N-1\}$ is an i.i.d. complex random sequence with bounded real part A_k^R and imaginary part A_k^I; in addition, $\mathbb{E}\left[A_k^R\right] = \mathbb{E}\left[A_k^I\right] = 0$, $\mathbb{E}\left[(A_k^R)^2\right] = \mathbb{E}\left[(A_k^I)^2\right] = \epsilon_k$, and $\mathbb{E}\left[A_k^R A_k^I\right] = 0$. Then for $N \to \infty$ and any closed and finite interval $T \subseteq R$, $X_N(t)$ converges in distribution to a zero-mean stationary complex Gaussian random process $X(t)$ whose real part $X^R(t)$ and imaginary part $X^I(t)$ satisfy*

$$\mathbb{E}\left[X^R(t)X^R(t+\tau)\right] = \mathbb{E}\left[X^I(t)X^I(t+\tau)\right] = \frac{1}{N}\sum_{k=0}^{N-1}\epsilon_k\cos(\omega_k\tau)$$

$$\mathbb{E}\left[X^R(t)X^I(t+\tau)\right] = \frac{1}{N}\sum_{k=0}^{N-1}\epsilon_k\sin(\omega_k\tau). \quad (17)$$

Proof: Please see [22] and [24]. ■

From Lemma 6, the real and imaginary parts of the output baseband signal converge to a Gaussian process with correlation functions

$$\mathbb{E}\left[s^R(t)s^R(t+\tau)\right] = \mathbb{E}\left[s^I(t)s^I(t+\tau)\right] = \frac{1}{2N}\sum_{k=0}^{N-1}|g_k^{(i)}|^2\cos\omega_k\tau, \quad (18)$$

and

$$\mathbb{E}\left[s^R(t)s^I(t+\tau)\right] = \frac{1}{2N}\sum_{k=0}^{N-1}|g_k^{(i)}|^2\sin\omega_k\tau. \quad (19)$$

From (19), $s^R(t)$ and $s^I(t)$ are independent at each time index t ($\tau = 0$). Therefore $\sqrt{2}s^R(t)/\sqrt{P_{av}}$ and $\sqrt{2}s^I(t)/\sqrt{P_{av}}$ satisfy (5) and (6) with

$$\lambda = \left(\frac{2\pi}{NT_c}\right)^2\frac{\sum_{k=0}^{N-1}|g_k^{(i)}|^2\left(k - \frac{N-1}{2}\right)^2}{\sum_{k=0}^{N-1}|g_k^{(i)}|^2},$$

and the Extreme Value Theory can be applied to approximate the PAPR value.

From Lemma 5, the CDF of PAPR at the ith transmit branch $\Pr\left\{\text{PAPR}^{(i)} \le \xi\right\}$ can be expressed as

$$\Pr\left\{\max_{0\le t\le T}\frac{1}{\overline{P}_{av}}\left[(s^R(t))^2 + (s^I(t))^2\right] \le \xi\right\} = \Pr\left\{\max_{0\le t\le T}\chi^2(t) \le 2\xi\right\} \to \exp\left\{-e^{-y}\right\}, \text{ as } T \to \infty, \quad (20)$$

where $y = a_T(2\xi - b_T) = \xi - \log T - (1/2)\log\log T - (1/2)\log(\lambda/\pi)$. Since $T = NT_c$, the CCDF of baseband PAPR has the following approximation:

$$\Pr\left\{\text{PAPR}^{(i)} > \xi\right\} \approx 1 - \exp\left\{-e^{-\xi}N\sqrt{\frac{\lambda}{\pi}\log N}\right\}, \quad (21)$$

with

$$\lambda = \left(\frac{2\pi}{N}\right)^2 \frac{\sum_{k=0}^{N-1} |g_k^{(i)}|^2 \left(k - \frac{N-1}{2}\right)^2}{\sum_{k=0}^{N-1} |g_k^{(i)}|^2}. \quad (22)$$

From (21), to evaluate the PAPR, λ needs to be calculated for every OFDM block and this increases computational complexity. To reduce the computational complexity, we prove that as $N \to \infty$ and sufficiently large M_t, λ converges to a constant no matter what the value $|g_k^{(i)}|^2$ is at each OFDM block. Before completing the proof, let us introduce the following lemma.

Lemma 7 *$\frac{\sum_{k=0}^{N-1} |g_k^{(i)}|^2 \left(k-\frac{N-1}{2}\right)^2}{\sum_{k=0}^{N-1} \left(k-\frac{N-1}{2}\right)^2}$ converges in mean square to $\mathbb{E}\left[|g_k^{(i)}|^2\right]$ when $N \to \infty$ and M_t is sufficiently large.*

Proof: Please see Appendix C. ∎

Theorem 3 *For MRT OFDM systems with $L = N$, when $N \to \infty$ and M_t is sufficiently large, the CCDF of the PAPR is the same as that of EGT or unprecoded OFDM systems; that is, (21) tends to approach the same value as (12), and the CCDF of PAPR can be expressed by (11)*

Proof: When $N \to \infty$ and M_t is sufficiently large, λ in (22) can be rewritten as

$$\lambda = \left(\frac{2\pi}{N}\right)^2 \frac{\frac{\sum_{k=0}^{N-1} |g_k^{(i)}|^2 \left(k-\frac{N-1}{2}\right)^2}{\sum_{k=0}^{N-1} \left(k-\frac{N-1}{2}\right)^2}}{\frac{NP_{av}}{\sum_{k=0}^{N-1} \left(k-\frac{N-1}{2}\right)^2}}. \quad (23)$$

From Lemma 7, (23) can be approximated as

$$\lambda \approx \left(\frac{2\pi}{N}\right)^2 \frac{\frac{N^3}{12}\mathbb{E}\left[|g_k^{(i)}|^2\right]}{NP_{av}}. \quad (24)$$

According to Lemmas 4 and 5, λ in (24) is equal to $\pi^2/3$ and this completes the proof. ∎

Note that although we assume that M_t is sufficiently large to obtain the result in (24), later simulation results show that the approximation in (24) is satisfactorily accurate for $M_t \geq 4$.

3) Arbitrary L but $L \neq 1$ or $L \neq N$: For an arbitrary $L = L_0$, where $L_0 \neq 1$ or N, the PAPR is difficult to obtain as explained below. In this case, the MRT vectors for different subcarriers at a specific transmit branch are correlated. The derivation for $L = L_0$ is analogous to that for $L = N$, but we need to consider the distribution of $P_{av} = \frac{1}{N}\sum_{k=0}^{N-1} |g_k^{(i)}|^2$, where $|g_k^{(i)}|^2 \sim Beta(1, M_t - 1)$. The statistic distribution of P_{av} is difficult to obtain because now the $|g_k^{(i)}|^2$ are correlated between subcarriers. As shown in [19], obtaining the distribution of the summation of correlated random variables involves the joint probability density function (PDF) of $|g_k^{(i)}|^2$. However, it is hard to obtain the joint PDF because the correlation of $|g_k^{(i)}|^2$ varies according to L and is difficult to formulate. Therefore, obtaining the PDF of P_{av} directly is difficult and the resulting form may be cumbersome. Instead, we make the following conjecture for obtaining an approximate analytical result, which can help us gain more insight into how L_0 affects PAPR. Moreover, it is shown later that the analytical result matches the simulated result quite well.

Conjecture 1 *Let $|g_k^{(i)}|^2$ be the power of the MRT element at the kth subcarrier and the ith transmit branch and let $P_{av} = \frac{1}{N}\sum_{k=0}^{N-1} |g_k^{(i)}|^2$. Then P_{av} may be approximated by a beta distribution if $|g_k^{(i)}|^2$ are identically beta distributed; the parameters α and β of the beta distribution can be obtained via the mean and variance of P_{av},* i.e.,

$$\mathbb{E}[P_{av}] = \frac{\alpha}{\alpha+\beta},$$
$$var[P_{av}] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \quad (25)$$

Explanation: The motivation of this conjecture is from the corollary in [26]. The authors have shown that the summation of i.i.d. beta random variables is still beta distributed and (α, β) can be obtained by (25). Here we make this conjecture by utilizing the results to correlated but identical beta random variables. ∎

In the following we use Conjecture 1 to obtain the distribution of P_{av}. Since $|g_k^{(i)}|^2$ are i.i.d., $\mathbb{E}[P_{av}] = \mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1} |g_k^{(i)}|^2\right] = \mathbb{E}[|g_k^{(i)}|^2] = 1/M_t$. To compute var$[P_{av}]$, we need to know the correlation coefficient of $|g_k^{(i)}|^2$. However, the correlation coefficient of $|g_k^{(i)}|^2$ is too complicated to obtain. Fortunately observe that $|g_k^{(i)}|^2 = \frac{|h_k^{(i)}|^2}{|h_k^{(0)}|^2+\cdots+|h_k^{(M_t-1)}|^2}$. Hence the correlation coefficient of $|g_k^{(i)}|^2$ should strongly depend on $|h_k^{(i)}|^2$. Therefore we first calculate the correlation coefficient of $|h_k^{(i)}|^2$ and then find the relation between $|h_k^{(i)}|^2$ and $|g_k^{(i)}|^2$. The correlation coefficient of $|h_k^{(i)}|^2$ and $|h_{k+K}^{(i)}|^2$ can be shown to be

$$\rho_{|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2} = \frac{\mathbb{E}[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2 - \mathbb{E}[|h_k^{(i)}|^2]\mathbb{E}[|h_{k+K}^{(i)}|^2]}{\sigma_{|h_k^{(i)}|^2}\sigma_{|h_{k+K}^{(i)}|^2}}$$
$$= \frac{\left(2L_0 + L_0(L_0-1) + 2\sum_{n=1}^{L_0-1}(L_0-n)\cos\frac{2\pi}{N}nK\right) - L_0^2}{L_0^2}$$
$$= \frac{L_0 + 2\sum_{n=1}^{L_0-1}(L_0-n)\cos\frac{2\pi}{N}nK}{L_0^2}. \quad (26)$$

Lemma 8 *When M_t is sufficiently large, the correlation coefficient of $|g_k^{(i)}|^2$ may be approximated by $|h_k^{(i)}|^2$, i.e.,*

$$\rho_{|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2} \approx \rho_{|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2}.$$

Proof: Please see Appendix D. ∎

Fig. 2 shows the correlation coefficients $\rho_{|h_0^{(i)}|^2|h_K^{(i)}|^2}$ and $\rho_{|g_0^{(i)}|^2|g_K^{(i)}|^2}$ for $L_0 = 4$, where $\rho_{|h_0^{(i)}|^2|h_K^{(i)}|^2}$ is obtained by evaluating (26), and $\rho_{|g_0^{(i)}|^2|g_K^{(i)}|^2}$ is obtained by Monte Carlo simulation. Observe that when $M_t \geq 4$, $\rho_{|g_0^{(i)}|^2|g_K^{(i)}|^2}$ is very close to $\rho_{|h_0^{(i)}|^2|h_K^{(i)}|^2}$. This shows that (26) may be used to approximate $\rho_{|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2}$ when $M_t \geq 4$. According to Lemma 8, var$[P_{av}]$ can be obtained by (27), shown at the bottom of the next page. From (26) and (27), var$[P_{av}]$ can be analytically evaluated. Therefore the distribution of P_{av} is obtained from Conjecture 1 and the corresponding parameters α and β can be calculated by applying (25)-(27). For presentation convenience, let the random variable $P_{av} = s$. Fig. 3 shows the CDF of s for different numbers of transmit antennas for $L_0 = 4$. Observe that when $M_t \geq 4$ the

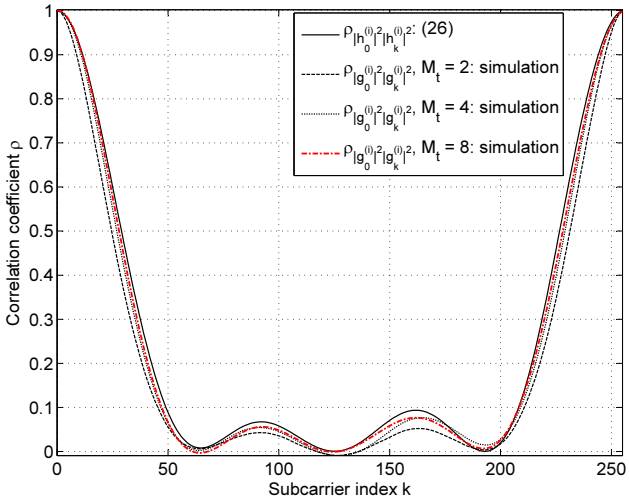

Fig. 2. Comparison of $\rho_{|h_0^{(i)}|^2|h_k^{(i)}|^2}$ and $\rho_{|g_0^{(i)}|^2|g_k^{(i)}|^2}$ for different numbers of transmit antennas.

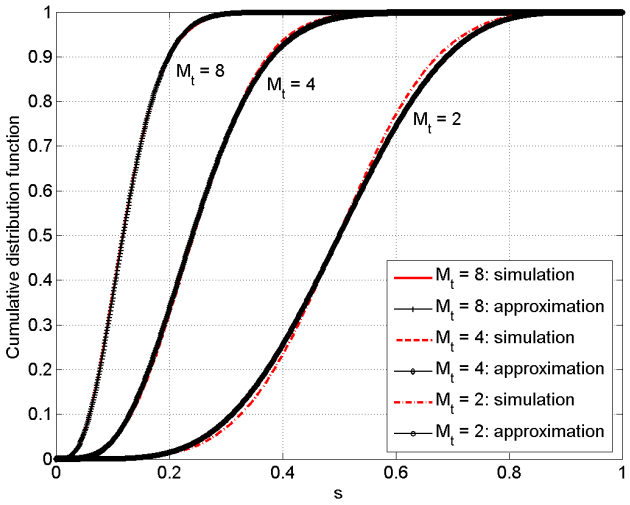

Fig. 3. The cumulative density function of P_{av}.

approximate results are consistent with the simulation results.

Therefore, when $N \to \infty$ and M_t is sufficiently large, the CCDF of the PAPR at the ith transmit branch can be obtained by the Extreme Value Theory and approximated to

$$\Pr\left\{\text{PAPR}^{(i)} > \xi\right\} \approx 1 - \int_0^1 \exp\left\{-e^{-\frac{\overline{P}_{av}\xi}{s}} N\sqrt{\frac{\pi}{3}\log N}\right\} f_S(s)ds, \quad (28)$$

where $f_S(s)$ is obtained using the following procedure:

- $\mathbb{E}[P_{av}] = \mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1}|g_k^{(i)}|^2\right] = \mathbb{E}[|g_k^{(i)}|^2] = 1/M_t$.
- Obtain $\text{var}[P_{av}]$ using (26), (27) and Lemma 8.
- Assume $f_S(s) \sim Beta(\alpha, \beta)$ and use (25) to obtain α and β.

From the above discussion, we have the following theorem:

Theorem 4 *For MRT OFDM systems with arbitrary L, where $L \neq 1$ and N, when M_t is sufficiently large and $N \to \infty$, the CCDF of the PAPR can be approximated by*

$$\Pr\{PAPR_{MRT} > \xi\} \approx 1 - \left(\Pr\left\{PAPR^{(i)} < \xi\right\}\right)^{M_t}, \quad (29)$$

where $\Pr\left\{PAPR^{(i)} < \xi\right\}$ can be obtained by (28).

IV. Proposed Algorithms for PAPR Reduction

Based on the derived results in Section III, PAPR reduction algorithms for MRT and EGT OFDM systems are proposed in this section. It is worth to point out that unlike the conventional PAPR reduction methods, *e.g.,* partial transmit sequence (PTS) and selected mapping (SLM), where the side information of rotated phases are required to be sent from the transmitter to the receiver [6],[8], the proposed PAPR algorithms do not require sending side information because the PAPR is reduced by adjusting the beamforming vectors. Therefore, redundancy is not added in the transmitted data. Moreover, for MRT OFDM systems, the proposed PAPR reduction algorithm not only can reduce the PAPR but also can improve the bit error rate. We introduce the proposed PAPR reduction algorithms for MRT and EGT OFDM separately as follows:

A. Proposed PAPR reduction algorithm for MRT OFDM systems

From Thms. 2, 3 and 4, for MRT OFDM, the PAPR decreases as L increases. It is shown that increasing L would decrease the correlation of subcarriers. This result gives us an intuition that the PAPR may reduce if the correlation of subcarriers decreases. On the other hand, it is known that the bit error rate for MRT OFDM systems is dominated by those subcarriers with low effective channel gain. From [28] and [29], the overall BER is minimized when all subcarriers have equal BER. Moreover, the BER of individual subcarriers can be adjusted by modifying the corresponding magnitude of beamforming vector in MIMO OFDM systems. Therefore, the proposed algorithm attempts to increase the amplitude of those subcarriers with low effective channel gain, while at the same time decrease the amplitude of those subcarriers with large one at each OFDM block transmission. By performing the proposed algorithm, we noticed that the overall beamforming

$$\text{var}[P_{av}] = \text{var}\left[\frac{1}{N}\sum_{k=0}^{N-1}|g_k^{(i)}|^2\right] = \frac{1}{N^2}\left\{\sum_{k=0}^{N-1}\text{var}\left[|g_k^{(i)}|^2\right] + 2\sum_{k=0}^{N-1}\sum_{k'=0,k'\neq k}^{N-1}\sqrt{\text{var}[|g_k^{(i)}|^2]}\sqrt{\text{var}[|g_{k'}^{(i)}|^2]}\rho_{|g_k^{(i)}|^2|g_{k'}^{(i)}|^2}\right\}$$
$$\approx \frac{1}{N}\sigma^2_{|g_k^{(i)}|^2} + \frac{2}{N^2}\sigma_{|g_0^{(i)}|^2}\sigma_{|g_1^{(i)}|^2}\sum_{k=1}^{N-1}\left[(N-k)\rho_{|h_0^{(i)}|^2|h_k^{(i)}|^2}\right]. \quad (27)$$

gain of individual subcarriers is more or less 'equalized' This modification not only decreases the correlation of subcarriers and thus reduces the PAPR, but also increases the gain of the subcarriers with low effective channel gain and thus improves the bit error rate. That is, the proposed algorithm may more or less destroy the correlation among subcarriers and improve the PAPR performance. However, PAPR reduction via modifying the amplitude may increase the long-term average power and affect the operation region of the PA. To maintain a nearly unchanged long-term average power and yet mitigate the PAPR, we propose to scale down the subcarrier power corresponding to the M largest channel gains and scale up the subcarrier power corresponding to the M smallest channel gains, where the scaling down and scaling up are performed in pairs. That is, we first sort γ_k in descending order so that the maximum sorted gain $\gamma_{k_0} = \max_{0\le k\le N-1}\{\gamma_k\}$, and k_i represents the ith ordered subcarrier index. Then if the signal of subcarrier k_i is multiplied by a real constant δ, where $0 < \delta < 1$, subcarrier k_{N-i} should also be multiplied by $1/\delta$. By properly determining M and δ, and using the proposed algorithm, the PAPR can be considerably reduced. We also apply a greedy optimization approach to decide whether the power of the M paired subcarriers should be scaled or not. The proposed PAPR reduction algorithm is summarized in Algorithm 1.

Algorithm 1: Proposed PAPR reduction algorithm for MRT OFDM systems.

1: **Initialization:** Determine M and δ. Let Ψ be system PAPR for current OFDM block.
2: Sort γ_k in descending order and obtain ordered effective gains γ_{k_i}.
3: Let $\Lambda(k_i, k_{N-i})$ be the system PAPR obtained by multiplying $\mathbf{g}_{k_i}$ by δ, and $\mathbf{g}_{k_{N-i}}$ by $1/\delta$.
4: **for** $i = 0 : M-1$ **do**
5: **if** $\Lambda(k_i, N-k_i) > \Psi$ **then**
6: $\mathbf{g}_{k_i} = \delta \mathbf{g}_{k_i}, \quad \mathbf{g}_{k_{N-i}} = \mathbf{g}_{k_{N-i}}/\delta,$ $\Psi = \Lambda(k_i, N-k_i).$
7: **end if**
8: **end for**

B. Proposed PAPR reduction algorithm for EGT OFDM systems

From Thms. 1, the PAPR of one specific transmit branch for EGT OFDM systems is the same as that for unprecoded OFDM systems. Hence, the PAPR reduction algorithms for unprecoded OFDM systems can be applied to EGT OFDM systems. EGT already has constant power for different OFDM blocks and transmit branches; also it is known that changing subcarrier phases can reduce the PAPR of OFDM systems (see [10]). Therefore, we change the subcarrier phases corresponding to the largest effective channel gains, because the error rate performance is dominated by the subcarriers corresponding to the smallest effective channel gains. By doing this, we can effectively reduce the PAPR, yet the error rate performance is only slightly degraded. In the proposed algorithm, we first sort the effective channel gain, γ_k, in descending order so that the maximum sorted gain $\gamma_{k_0} = \max_{0\le k\le N-1}\{\gamma_k\}$, and k_i represents the ith ordered subcarrier index. The EGT vectors corresponding to the largest M gains are identified. Then we use a greedy algorithm to decide whether each individual of the M EGT vectors should keep its phases unchanged, or if it should be multiplied by a phase shift, either $e^{j\phi}$ or $e^{-j\phi}$, where $0 < \phi < \pi$, so that system PAPR be reduced. Therefore, by properly choosing M and ϕ, the PAPR can be effectively reduced without sending side information to the receiver, because we only modify the phase of beamforming vectors $\mathbf{g}_k$ not the transmit signals x_k. At the same time the error rate performance can still be maintained at a satisfactory level. The algorithm is summarized in Algorithm 2.

Algorithm 2: Proposed PAPR reduction algorithm for EGT OFDM system.

1: **Initialization:** Determine M and ϕ. Let Ψ be system PAPR for the current OFDM block.
2: Sort γ_k in descending order and obtain ordered effective gains γ_{k_i}.
3: Let $\Lambda(k_i, \phi)$ be the system PAPR obtained by multiplying $\mathbf{g}_{k_i}$ by $e^{j\phi}$.
4: **for** $i = 0 : M-1$ **do**
5: $\hat{\phi} = \arg \min_{\phi\in\{-\phi,0,\phi\}} \Lambda(k_i, \phi), \quad \Psi = \Lambda(k_i, \hat{\phi}).$
6: **end for**

V. Simulation results

In this section several examples are provided to demonstrate that the derived analytical results are verified by simulation results. Also, the simulation results show that the proposed PAPR reduction algorithms can effectively reduce the PAPR for both MRT and EGT OFDM systems. At the same time the performance is improved for MRT OFDM systems and is only slightly degraded for EGT OFDM systems. The simulations were conducted using the following settings. The number N of subcarriers is 256. The MIMO channel coefficients are assumed to be i.i.d. complex Gaussian distributed with zero mean and unit variance. The oversampling factor 4 is applied to coincide the PAPR distribution of continuous and discrete signals (see [27]). 4-QAM is used for the transmit symbol x. The notation mTnR corresponds to $M_t = m$ and $M_r = n$. PAPR_0 in the y-axis is used to represent ξ mentioned in the previous sessions.

Example 1. PAPR in EGT OFDM systems: The PAPR for EGT OFDM systems with $M_t = 4$ is shown in Fig. 4. The derived PAPR expression for EGT OFDM systems considering all transmit branches, *i.e.* Thm. 1, approximate the simulation results quite well. It is worth to point out that there exists a small gap between the theoretical and the simulation results for $L = 4$. The reason is explained below: For small L, the power of the transmit signals after multiplying beamforming vectors may be negatively associated (NA) [30]. Since we assume that the transmit signals are independent for all possible values of L, the theoretical results can be regarded as an upper bound, and the bound is tighter for large PAPR_0 than small one.

Example 2. PAPR for a specific transmit branch in beamforming MIMO OFDM systems: The PAPR considering

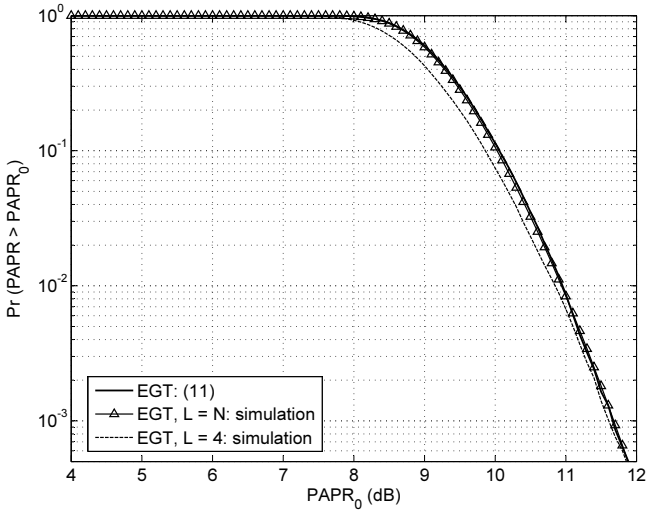

Fig. 4. PAPR for EGT OFDM systems.

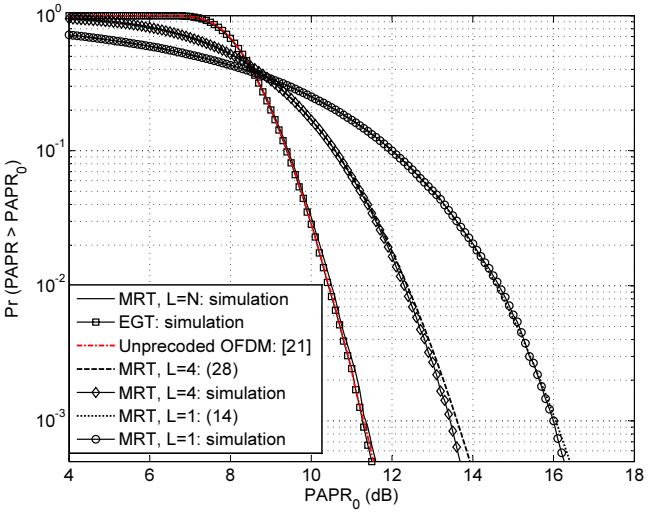

Fig. 6. PAPR considering only one transmit branch for different L in beamforming OFDM systems with $M_t = 8$.

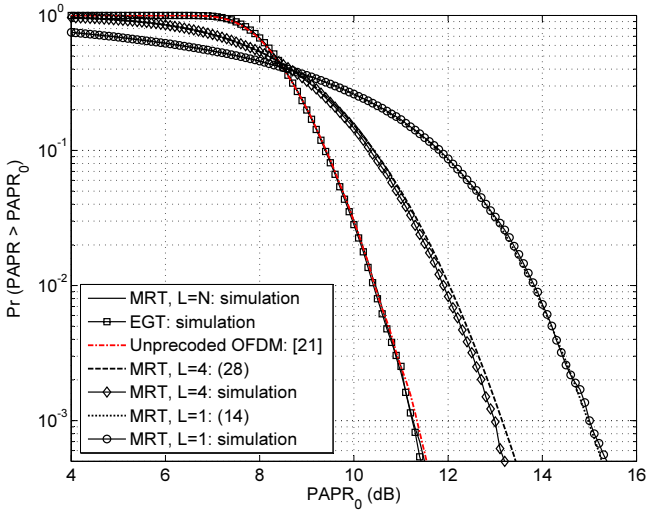

Fig. 5. PAPR considering only one transmit branch for different L in beamforming OFDM systems with $M_t = 4$.

only one transmit branch for different L in MISO OFDM systems with $M_t = 4$ and $M_t = 8$ are shown in Figs. 5 and 6, respectively. The two figures show that our derived results in (14), (21) and (28) are consistent with the simulation results. Note that we have mentioned, when $L = N \to \infty$, the PAPR of the MRT OFDM systems considering only one transmit branch is the same as that for unprecoded OFDM systems, and this is observed in the solid and dash-dot curves. Moreover, we mentioned in (10) that the PAPR of EGT OFDM systems considering only one specific transmit branch is the same as unprecoded OFDM systems, and this is also observed from the square curve.

Example 3. PAPR for MRT OFDM systems: The PAPR of MRT OFDM systems considering all transmit branches are shown in Fig. 7 for $M_t = 4$, and Fig. 8 for $M_t = 8$. Different values of $L = 1,\ 4$, and N are considered in this example. From the figures, we observe the derived results in Thms. 2, 3 and 4 approximate the simulation results very well. In addition, the gap between the theoretical and the simulation results for $L = 1$ is due to the negatively associated property for different transmit branches. Hence the derived result can be regarded as an upper bound. Moreover, the theoretical result for $L = 4$ is based on Thm. 4. Although we made some assumptions and approximations to derive the theorem, the theoretical result still matches the simulation result quite well (see dash and diamond curves). Note that without making the assumptions and approximations, the results in Thm. 4 would be cumbersome and may not be able to provide any insight into how L affects the PAPR. For example, we see from the figures that the CCDF of the PAPR in MRT OFDM systems decreases when L increases. This can be deduced from Lemma 8 and the derived result in (26), since as L increases, the correlation between the $|g_k^{(i)}|$ decreases and the CCDF of the PAPR approaches to Thm. 3. In addition, from Figs. 7 and 8, the performance gap between $L = 1$ and $L = N$ increases when M_t increases from 4 to 8. This is reasonable because as N_t increases, the elements of the beamforming vector also increases and this leads to a wide dynamic range of signals after beamforming.

Finally Fig. 7 also compares the PAPR of 4T2R MRT OFDM and 4T1R MRT OFDM systems. Since the MRT vectors for MISO and MIMO systems both follow a conditional Haar distribution, the PAPR characteristics of the two systems are identical.

Example 4. Proposed PAPR reduction algorithm for MRT OFDM systems: For $L = 4$ in a 4T2R MRT OFDM system, the PAPR applying Algorithm 1 for $\delta = 0.5$ and 0.8 is shown in Fig. 9. Here the number of subcarriers for greedy optimization is determined off-line to be $M = N/8$. The corresponding bit error rate (BER) is shown in Fig. 10. For $\delta = 0.8$, the BER as well as the PAPR is improved. More specifically, the PAPR improves around 0.5 dB at $\text{PAPR}_0 = 10^{-3}$ observed from Fig. 9, and the BER improves around 0.4 dB at BER$= 10^{-5}$ observed from Fig. 10. These observations show that the proposed algorithm not only can reduce the PAPR for MRT OFDM systems but also can improve the BER performance. It is worth to mention that if a slight degradation of the bit error

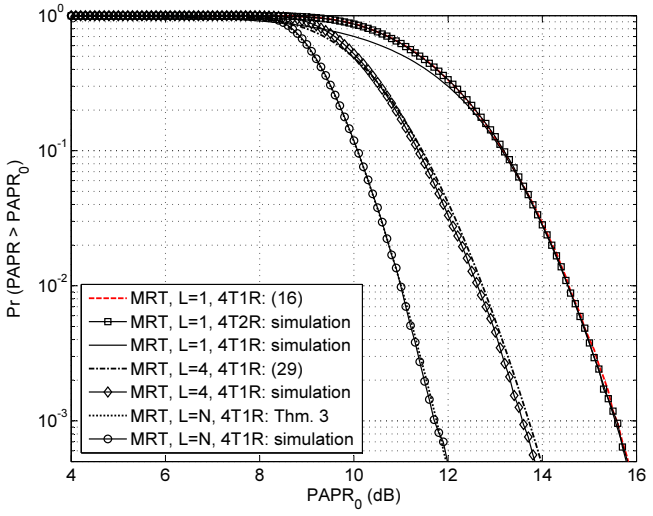

Fig. 7. PAPR for different L for MRT OFDM systems with $M_t = 4$ and $M_r = 1$ and 2.

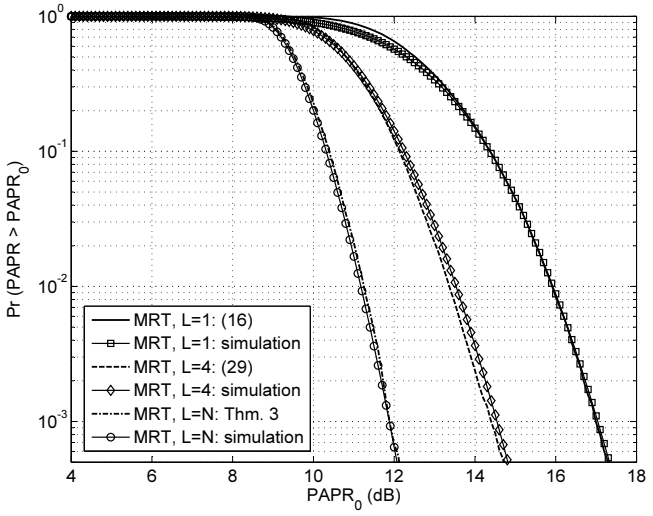

Fig. 8. PAPR for different L for MRT OFDM systems with $M_t = 8$ and $M_r = 1$.

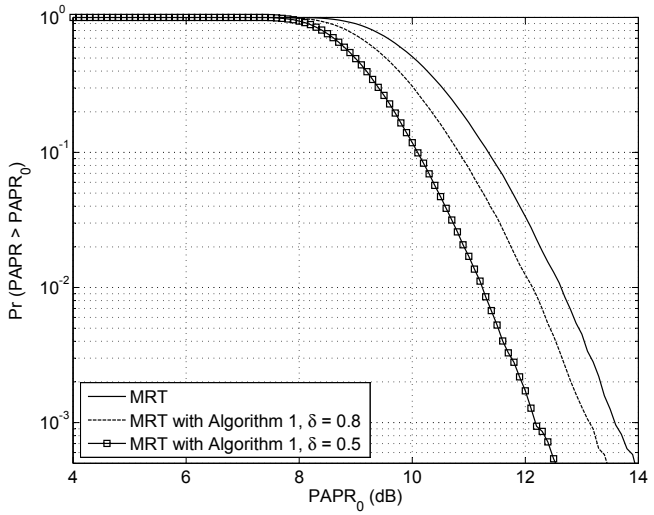

Fig. 9. Comparison of PAPR with and without the proposed algorithm for MRT OFDM systems.

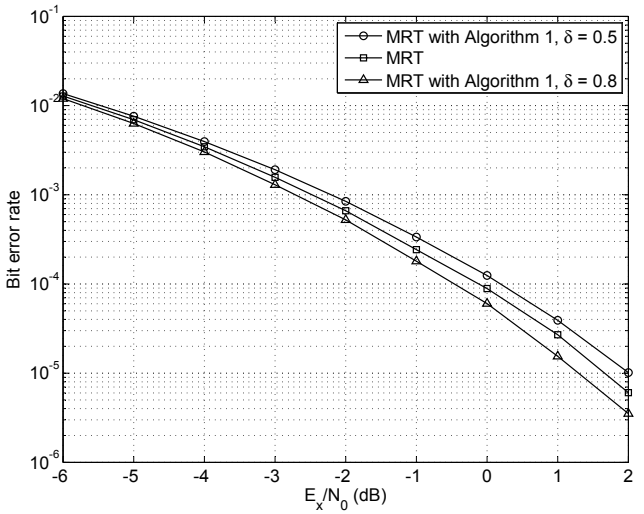

Fig. 10. Comparison of BER with and without the proposed algorithm for MRT OFDM systems.

rate performance, *i.e.,* 0.3 dB in this example, is allowable, the PAPR performance can be improved by up to 1.5 dB by letting $\delta = 0.5$.

Example 5. Proposed PAPR reduction algorithm for EGT OFDM systems: For $L = 4$ in a 4T2R EGT OFDM system, the PAPR applying Algorithm 2 with different shift angles ϕ is shown in Fig. 11. The EGT vectors are generated by the algorithm proposed in [15]. Here the number of subcarriers for greedy optimization is determined off-line to be $M = N/8$. The CCDF of PAPR can be reduced by around 1 and 1.5 dB when $\phi = \pi/8$ and $\phi = \pi/4$, respectively. As can be observed from Fig. 11, the PAPR has better CCDF distribution than those depicted in the figure if the set size and shift angle are enlarged. However, these modifications would also degrade the system error rate performance and increase the computational complexity, as observed from the error performance shown in Fig. 12. Therefore, M and ϕ should be carefully determined to keep error performance slightly degraded and the complexity low. Finally, from Fig. 12, the performance gap between MRT OFDM and EGT OFDM systems (without the proposed algorithms) is less than 0.8 dB, which is consistent with the results in [17].

VI. Conclusion

In this paper we investigate the PAPR performance and corresponding mitigation algorithms for beamforming OFDM systems. We analytically derive the PAPR distribution for EGT OFDM and MRT OFDM systems. The theoretical results show that generally MRT OFDM systems perform much worse than EGT OFDM systems in terms of PAPR. Therefore, although MRT is the optimal beamforming scheme, which can achieve 1.05 dB more receive SNR than EGT, when the cost of the PA and the better power consumption are of concern, EGT may be a preferred solution due to its superior PAPR performance in OFDM systems. Moreover, we propose PAPR reduction

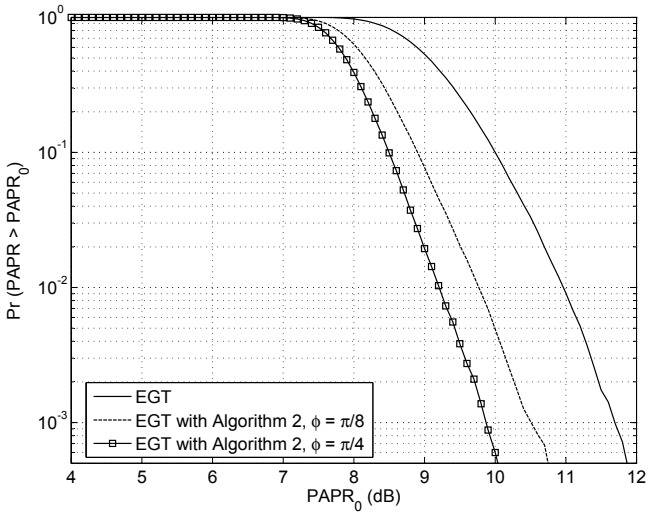

Fig. 11. Comparison of PAPR with and without the proposed algorithm for EGT OFDM systems.

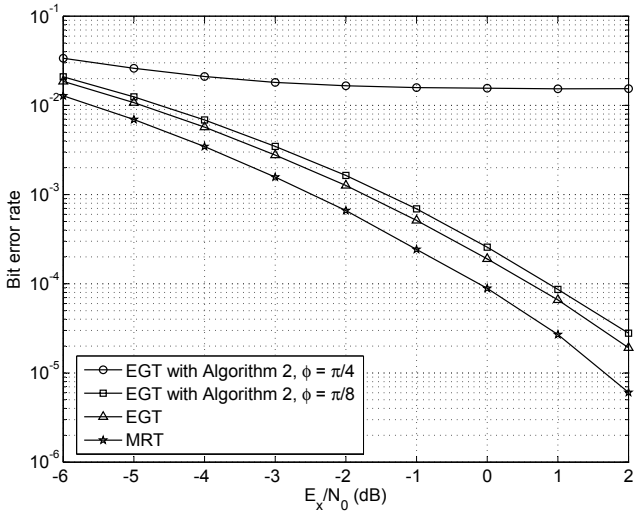

Fig. 12. Comparison of BER with and without the proposed algorithm for EGT OFDM systems.

algorithms for both MRT OFDM and EGT OFDM systems. The performance improvement using the proposed algorithm is more pronounced in MRT OFDM systems; that is, both the PAPR and the bit error rate performance can be improved simultaneously. If a more aggressive PAPR improvement is needed, by carefully determining the design parameters, the proposed algorithms can greatly improve PAPR performance, yet the bit error rate performance is slightly degraded. Simulation results show that the proposed PAPR mitigation algorithms indeed significantly boost the PAPR performance. Also, the analytical results for the PAPR match simulation results well. Consequently, the derived outcomes could be used to evaluate the required PA specification and the power consumption for beamforming OFDM systems in practical designs. An interesting extension of this work is to analyze the PAPR for multi-stream and SDMA systems, and this question is still open.

APPENDIX

A. Proof of Lemma 1

The discrete time-domain baseband signal at the ith branch $\mathbf{s}^{(i)}$ can be expressed as

$$\mathbf{s}^{(i)} = \mathbf{F}^{\dagger}\mathbf{G}^{(i)}\mathbf{x},$$

where $\mathbf{x} \in \mathbb{C}^{N\times 1}$ is the transmitted data vector, $\mathbf{F} \in \mathbb{C}^{N\times N}$ and $[\mathbf{F}]_{kn} = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}(k-\frac{N-1}{2})n}$, $\mathbf{G}^{(i)} = \text{diag}([g_0^{(i)} g_1^{(i)}, \ldots, g_{N-1}^{(i)}])$. The block average power P_{av} at the ith transmit branch can be expressed as

$$\begin{aligned} P_{av} = \frac{1}{N}(\mathbf{s}^{(i)})^{\dagger}\mathbf{s}^{(i)} &= \frac{1}{N}\mathbf{x}^{\dagger}(\mathbf{G}^{(i)})^{\dagger}\mathbf{F}\mathbf{F}^{\dagger}\mathbf{G}^{(i)}\mathbf{x} \\ &= \frac{1}{N}\mathbf{x}^{\dagger}(\mathbf{G}^{(i)})^{\dagger}\mathbf{G}^{(i)}\mathbf{x} = \frac{1}{NM_t}\mathbf{x}^{\dagger}\mathbf{x}, \end{aligned} \tag{30}$$

where we have used the fact that $(\mathbf{G}^{(i)})^{\dagger}\mathbf{G}_i = \frac{1}{M_t}\mathbf{I}$ for EGT beamforming. From (30), we see that if constant amplitude modulation is used for $\mathbf{x}$, *e.g.* BPSK or QPSK, $\mathbf{x}^{\dagger}\mathbf{x}$ is constant, and P_{av} is therefore constant and independent of the index i of transmit branch. Since P_{av} is constant, the long-term average power $\overline{P}_{av} = \mathbb{E}[P_{av}] = 1/M_t$. ■

B. Proof of Lemma 3

From Lemma 2, the complex signal $s^{(i)}(t)$ in (3) converges to zero-mean stationary complex Gaussian random process with real part $s^R(t)$ and imaginary part $s^I(t)$ such that

$$\begin{aligned} r(\tau) = \mathbb{E}\left[s^R(t)s^R(t+\tau)\right] &= \mathbb{E}\left[s^I(t)s^I(t+\tau)\right] \\ &= \frac{1}{2M_t}\text{sinc}\left(\frac{\tau}{T_c}\right) \end{aligned}$$

and

$$\mathbb{E}\left[s^R(t_1)s^I(t_2)\right] = 0, \ \forall\, t_1 \text{ and } t_2.$$

Therefore, $\sqrt{2}s^R(t)/\sqrt{P_{av}}$ and $\sqrt{2}s^I(t)/\sqrt{P_{av}}$ satisfy (5) and (6) with $\lambda = \frac{1}{3}(\frac{\pi}{T_c})^2$. By applying Lemma 1 and the Extreme Value Theory, the cumulative distribution function (CDF) of the PAPR at the ith transmit branch $\Pr\left\{\text{PAPR}^{(i)} \le \xi\right\}$ follows the asymptotic characteristic

$$\begin{aligned} &\Pr\left\{\max_{0\le t\le T}\frac{1}{\overline{P}_{av}}\left[(s^R(t))^2 + (s^I(t))^2\right] \le \xi\right\} \\ &= \Pr\left\{\max_{0\le t\le T}\chi^2(t) \le 2\xi\right\} \to \exp(-e^{-y}), \ \text{ as } T\to\infty, \end{aligned} \tag{31}$$

where $y = a_T(2\xi - b_T) = \xi - \log T - (1/2)\log\log T - (1/2)\log(\lambda/\pi)$.

Let the time interval of output baseband signal be sampled with T_c, *i.e.,* $T = NT_c$ and $\lambda = \frac{\pi^2}{3}$. Therefore the CCDF of the PAPR at the ith transmit branch can be approximated as that in (10), where the approximation is due to the Extreme Value Theory. ■

C. Proof of Lemma 7

Proving Lemma 7 is equivalent to proving that when M_t is sufficiently large, the following equality holds

$$\lim_{N\to\infty} \mathbb{E}\left[\left|u^{(i)} - \mathbb{E}\left[|g_k^{(i)}|^2\right]\right|^2\right] = 0, \tag{32}$$

where

$$u^{(i)} = \frac{\sum_{k=0}^{N-1} |g_k^{(i)}|^2 \left(k - \frac{N-1}{2}\right)^2}{\sum_{k=0}^{N-1} \left(k - \frac{N-1}{2}\right)^2}.$$

The left side term of (32) can be rewritten as

$$\lim_{N\to\infty} \mathbb{E}\left[\left(u^{(i)}\right)^2\right] - 2\mathbb{E}\left[|g_k^{(i)}|^2\right]\mathbb{E}\left[u^{(i)}\right] + \left(\mathbb{E}\left[|g_k^{(i)}|^2\right]\right)^2. \tag{33}$$

Since $\left\{|g_k^{(i)}|^2,\ 0 \le k \le N-1\right\}$ is i.i.d. when $L = N \to \infty$, $g_k^{(i)}$ is independent of k and we denote $\mathbb{E}[|g_k^{(i)}|^2] = \mathbb{E}[|g^{(i)}|^2]$ for simplicity. Therefore the second term of (33) can be written to $(2\mathbb{E}[|g^{(i)}|^2])^2$. We expand the first term of (33) and it yields (34), shown at the top of the next page. Since $|g_k^{(i)}|^2 \sim Beta(1, M_t - 1)$, $\text{var}[|g_k^{(i)}|^2] = \frac{M_t - 1}{M_t^2(M_t+1)} \to 0$ when M_t is sufficiently large. Thus we have $\mathbb{E}[(|g_k^{(i)}|^2)^2] = \left(\mathbb{E}[|g_k^{(i)}|^2]\right)^2$ and (34) can be rewritten as

$$\frac{\left(\mathbb{E}[|g^{(i)}|^2]\right)^2 \sum_{k=0}^{N-1}\left(k - \frac{N-1}{2}\right)^2 \sum_{k'=0}^{N-1}\left(k' - \frac{N-1}{2}\right)^2}{\left(\sum_{k=0}^{N-1}\left(k - \frac{N-1}{2}\right)^2\right)^2} = \left(\mathbb{E}[|g^{(i)}|^2]\right)^2. \tag{35}$$

Therefore, we have

$$\lim_{N\to\infty} \mathbb{E}\left[\left|u^{(i)} - \mathbb{E}\left[|g_k^{(i)}|^2\right]\right|^2\right] = \left(\mathbb{E}[|g^{(i)}|^2]\right)^2 - 2\left(\mathbb{E}[|g^{(i)}|^2]\right)^2 + \left(\mathbb{E}[|g^{(i)}|^2]\right)^2 = 0. \tag{36}$$

■

D. Proof of Lemma 8

The correlation coefficient of $|h_k^{(i)}|^2$ and $|g_k^{(i)}|^2$ can be respectively expressed as

$$\rho_{|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2} = \frac{\mathbb{E}[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2] - \mathbb{E}[|h_k^{(i)}|^2]\mathbb{E}[|h_{k+K}^{(i)}|^2]}{\sigma_{|h_k^{(i)}|^2}\sigma_{|h_{k+K}^{(i)}|^2}},$$
$$\rho_{|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2} = \frac{\mathbb{E}[|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2] - \mathbb{E}[|g_k^{(i)}|^2]\mathbb{E}[|g_{k+K}^{(i)}|^2]}{\sigma_{|g_k^{(i)}|^2}\sigma_{|g_{k+K}^{(i)}|^2}}. \tag{37}$$

Since $|h_k^{(i)}|^2 \sim \Gamma(1, L/2N)$ and $|g_k^{(i)}|^2 \sim Beta(1, M_t - 1)$, (37) can be rewritten as

$$\rho_{|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2} = \frac{\mathbb{E}[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2] - \left(\frac{L}{2N}\right)^2}{\left(\frac{L}{2N}\right)^2},$$
$$\rho_{|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2} = \frac{\mathbb{E}[|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2] - \left(\frac{1}{M_t}\right)^2}{\frac{M_t-1}{M_t^2(M_t+1)}}. \tag{38}$$

Notice that $\mathbb{E}[|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2] = \mathbb{E}\left[\frac{|h_k^{(i)}|^2}{|h_k^{(0)}|^2+\cdots+|h_k^{(M_t-1)}|^2} \cdot \frac{|h_{k+K}^{(i)}|^2}{|h_{k+K}^{(0)}|^2+\cdots+|h_{k+K}^{(M_t-1)}|^2}\right]$, and $|h_k^{(i)}|^2$ is correlated to itself and $|h_{k+K}^{(i)}|^2$. When M_t is sufficiently large, $|h_{k+K}^{(i)}|^2$ and $|h_k^{(0)}|^2 + \cdots + |h_k^{(M_t-1)}|^2$ are nearly uncorrelated because $|h_{k+K}^{(i)}|^2$ is only correlated with the term $|h_{k+K}^{(i)}|^2$ in the sum. Similarly, $(|h_k^{(0)}|^2 + \cdots + |h_k^{(M_t-1)}|^2)$ and $(|h_{k+K}^{(0)}|^2 + \cdots + |h_{k+K}^{(M_t-1)}|^2)$ are also nearly uncorrelated when M_t is sufficiently large, because there are M_t correlated terms, but the remaining $M_t(M_t-1)$ terms are uncorrelated. Therefore,

$$\begin{aligned}
&\mathbb{E}[|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2] \\
&= \mathbb{E}\left[\frac{|h_k^{(i)}|^2}{|h_k^{(0)}|^2 + \cdots + |h_k^{(M_t-1)}|^2} \cdot \frac{|h_{k+K}^{(i)}|^2}{|h_{k+K}^{(0)}|^2 + \cdots + |h_{k+K}^{(M_t-1)}|^2}\right] \\
&\approx \mathbb{E}\left[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2\right] \cdot \mathbb{E}\left[\frac{1}{|h_k^{(0)}|^2 + \cdots + |h_k^{(M_t-1)}|^2}\right] \\
&\quad \cdot \mathbb{E}\left[\frac{1}{|h_{k+K}^{(0)}|^2 + \cdots + |h_{k+K}^{(M_t-1)}|^2}\right] \\
&= \mathbb{E}\left[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2\right] \cdot \left(\frac{1}{M_t - 1}\right)^2 \left(\frac{2N}{L}\right)^2.
\end{aligned} \tag{39}$$

From (39), we can approximate $\rho_{|g_k^{(i)}|^2|g_{k+K}^{(i)}|^2}$ in (38) by

$$\frac{\mathbb{E}\left[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2\right] \cdot \left(\frac{1}{M_t-1}\right)^2 \left(\frac{2N}{L}\right)^2 - \left(\frac{1}{M_t}\right)^2}{\frac{M_t-1}{M_t^2(M_t+1)}} \approx \frac{\mathbb{E}\left[|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2\right] - \left(\frac{L}{2N}\right)^2}{\left(\frac{L}{2N}\right)^2} = \rho_{|h_k^{(i)}|^2|h_{k+K}^{(i)}|^2}. \tag{40}$$

■

ACKNOWLEDGMENT

The authors would like to thank all the anonymous reviewers for their constructive suggestions, which have significantly improved the quality of this work.

REFERENCES

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—a key to gigabit wireless," *Proc. IEEE,* vol. 92, no. 2, pp. 198–218, Feb. 2004.

[2] K. F. Lee and D. B. Williams, "A space-frequency transmitter diversity technique for OFDM systems," in *Proc. 2000 IEEE GLOBECOM*, pp. 1473–1477.

[3] S. Khademi, A.-J. van der Veen, and T. Svantesson, "Precoding technique for peak-to-average-power-ratio (PAPR) reduction in MIMO OFDM/A systems," in *Proc. 2012 IEEE ICASSP*, vol. 20, pp. 3005–3008.

[4] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun.,* vol. 12, no. 2, pp. 56–65, Apr. 2005.

[5] U.-K. Kwon, D. Kim, and G.-H. Im, "Amplitude clipping and iterative reconstruction of MIMO-OFDM signals with optimum equalization," *IEEE Trans. Wireless Commun.,* vol. 8, no. 1, pp. 268–277, Jan. 2009.

[6] S. H. Müller and J. B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electron. Lett.,* vol. 33, no. 5, pp. 368–369, Feb. 1997.

[7] L. Wang and J. Liu, "Cooperative PTS for PAPR reduction in MIMO-OFDM," *Electron. Lett.,* vol. 47, no. 5, pp. 351–352, Mar. 2011.

[8] R. W. Bäuml, R. F. H. Fischer, and J. B. Huber, "Reducting the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electron. Lett.,* vol. 32, no. 22, pp. 2056–2057, Oct. 1996.

$$\mathbb{E}\left[\left(u^{(i)}\right)^2\right] = \frac{\mathbb{E}\left[\left(\sum_{k=0}^{N-1}|g_k^{(i)}|^2\left(k-\frac{N-1}{2}\right)^2\right)\left(\sum_{k'=0}^{N-1}|g_{k'}^{(i)}|^2\left(k'-\frac{N-1}{2}\right)^2\right)\right]}{\left(\sum_{k=0}^{N-1}\left(k-\frac{N-1}{2}\right)^2\right)^2}$$
$$= \frac{\mathbb{E}\left[\sum_{k=0}^{N-1}\left(|g_k^{(i)}|^2\right)^2\left(k-\frac{N-1}{2}\right)^4 + \sum_{k=0}^{N-1}|g_k^{(i)}|^2\left(k-\frac{N-1}{2}\right)^2\sum_{k'=0,k'\neq k}^{N-1}|g_{k'}^{(i)}|^2\left(k-\frac{N-1}{2}\right)^2\right]}{\left(\sum_{k=0}^{N-1}\left(k-\frac{N-1}{2}\right)^2\right)^2}. \quad (34)$$

[9] M. Sharif and B. Hassibi, "Existence of codes with constant PMEPR and related design," *IEEE Trans. Signal Process.,* vol. 52, no. 10, pp. 2836–2846, Oct. 2004.

[10] M. Sharif, V. Tarokh, and B. Hassibi, "Peak power reduction of OFDM signals with sign adjustment," *IEEE Trans. Commun.,* vol. 57, no. 7, pp. 2160–2166, July 2009.

[11] J. Joung, E.-R. Jeong, and Y. H. Lee, "Beamforming and PAPR reduction for MISO-OFDM systems," in *Proc. 2007 IEEE ICASSP,* vol. 3, pp. 377–380.

[12] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Sel. Areas Commun.,* vol. 18, no. 11, pp. 2172–2178, Nov. 2000.

[13] D. J. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-out wireless systems," *IEEE Trans. Commun.,* vol. 51, no. 7, pp. 1102–1110, July 2003.

[14] C. R. Murthy and B. D. Rao, "Quantization methods for equal gain transmission with finite rate feedback," *IEEE Trans. Signal Process.,* vol. 55, no. 1, pp. 233–245, Jan. 2007.

[15] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Trans. Signal Process.,* vol. 55, no. 11, pp. 5395–5406, Nov. 2007.

[16] S.-H. Tsai, "Transmit equal gain precoder in Rayleigh fading channels," *IEEE Trans. Signal Process.,* vol. 57, no. 9, pp. 3717–3721, Sept. 2009.

[17] S.-H. Tsai, "Equal gain transmission with antenna selection in MIMO communications," *IEEE Trans. Wireless Commun.,* vol. 10, no. 5, pp. 1470–1479, May 2011.

[18] M. R. Leadbetter and H. Rootzen, "An extremal theory for stochasic processes," *Annals Probability,* vol. 16, pp. 431–478, Apr. 1988.

[19] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochasitic Process.* McGraw-Hill, 2002.

[20] H. A. David and H. N. Nagaraja, *Order Statistics,* 3rd ed. John Wiley & Sons, Inc., 2003.

[21] S. Wei, D. Goeckel, and P. Kelly, "A modern extreme value theory approach to calculating the distribution of the peak-to-average power ratio in OFDM Systems," in *Proc. 2002 IEEE ICC,* vol. 3, pp. 1686–1690.

[22] T. Jiang, M. Guizani, H.-H. Chen, W. Xiang, and Y. Wu, "Derivation of PAPR distribution for OFDM wireless systems based on extreme value theory," *IEEE Trans. Wireless Commun.,* vol. 7, no. 4, pp. 1298–1305, Apr. 2008.

[23] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.,* vol. 49, no. 2, pp. 282–289, Feb. 2001.

[24] S. Wei, D. Goeckel, and P. Kelly, "Convergence of the complex envelope of bandlimited OFDM signals," *IEEE Trans. Inf. Theory,* vol. 56, no. 10, pp. 4893–4904, Oct. 2010.

[25] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis,* 2nd ed. John Wiley & Sons, Inc., 1984.

[26] T. G. Pham and N. Turkkan, "Reliability of a standby system with beta-distributed component lives," *IEEE Trans. Reliability*, vol. 43, no. 1, pp. 71–75, Mar. 1994.

[27] C. Tellambura, "Computation of the continuous time PAR of an OFDM signal with BPSK subcarriers," *IEEE Commun. Lett.,* vol. 5, no. 5, pp. 185–187, May 2001.

[28] Y.-P. Lin, S.-M. Phoong, and P. P. Vaidyanathan, *Filter Bank Transceivers for OFDM and DMT Systems.* Cambridge University Press, 2011.

[29] Y.-P. Lin and S.-M. Phoong, "Statistical bit allocation and statistical precoding for correlated MIMO channels with decision feedback," *IEEE Signal Process. Lett.,* vol. 19, no. 11, pp. 761–764, Nov. 2012.

[30] T. Hu and J. Hu, "Comparison of order statistics between dependent and independent random variables," *Statist. Probab. Lett.,* vol. 37, pp. 1–6, Jan. 1998.

Ying-Che Hung was born in Taichung, Taiwan, in 1983. He received the B.S. degree in Power Mechanical Engineering from the National Tsing-Hua University, Taiwan, in 2006. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, National Chiao-Tung University, Taiwan. His research interests mainly include digital signal processing, multiple-input-multiple-output (MIMO) wireless communications, and compressive sensing.

Shang-Ho (Lawrence) Tsai (SM'12) was born in Kaohsiung, Taiwan, 1973. He received the Ph.D. degree in Electrical Engineering from the University of Southern California (USC), USA, in Aug. 2005. From June 1999 to July 2002, he was with the Silicon Integrated Systems Corp. (SiS), where he participated the VLSI design for DMT-ADSL systems. From Sep. 2005 to Jan. 2007, he was with the MediaTek Inc. (MTK) and participated the VLSI design for MIMO-OFDM systems. From Jun. 2013 to Dec. 2013, he was a visiting fellow in the department of Electrical Engineering at the Princeton University. Since Feb. 2007, he joined the Department of Electrical and Control Engineering (now Department of Electrical Engineering) at the National Chiao Tung University where he is now an associate professor. His research interests include signal processing for communications, statistical signal processing, and signal processing for VLSI designs. He was awarded a government scholarship for overseas study from the Ministry of Education, Taiwan, in 2002-2005.