

# Analysis of Clipping Noise and Tone-Reservation Algorithms for Peak Reduction in OFDM Systems

Luqing Wang, *Student Member, IEEE*, and Chintha Tellambura, *Senior Member, IEEE*

**Abstract**—Orthogonal frequency division multiplexing (OFDM) suffers from a high peak-to-average power ratio (PAR). Tone reservation is a popular PAR reduction technique that uses a set of reserved tones to design a peak-canceling signal. In a previous paper by Krongold and Jones, an active-set approach was developed to efficiently compute the peak-canceling signal. In this paper, we consider the use of clipping noise, which is generated when the OFDM signal is clipped at a predefined threshold, to design the peak-canceling signal. To this end, the clipping noise is analyzed as a series of parabolic pulses under tone-reservation constraints. The single-pulse case and the multiple-pulse case are treated. The analysis explains peak regrowth and the constancy of the clipping noise power spectrum over the whole OFDM band. Moreover, the clipping noise at the end of several clipping and filtering iterations is shown to be approximately proportional to that generated in the first iteration. The constant of proportionality is estimated via the level-crossing theory for high clipping thresholds. Using this analysis, a constant-scaling algorithm and an adaptive-scaling algorithm are proposed for tone reservation. These algorithms scale the filtered first-iteration clipping noise to compensate for peaks that are above the threshold. The simulation results show that the proposed algorithms achieve a larger peak reduction and lower complexity than the active-set algorithm.

**Index Terms**—Orthogonal frequency division multiplexing (OFDM), peak reduction, scaling, time and frequency domain analysis, tone reservation.

## I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM), although used in standards such as IEEE 802.11a/g, IEEE 802.16, high-performance radio LAN version 2, and digital video broadcasting [2], suffers from the high peak-to-average power ratio (PAR) [2]. A large PAR requires a linear high power amplifier (HPA), which is inefficient. Moreover, the combination of an insufficiently linear HPA range and a large PAR leads to in-band distortion and out-of-band radiation [2]. Various PAR reduction techniques have, therefore, been proposed, including clipping and filtering [3]–[6], tone reservation [1], [7], [8], multiple signal representation [9]–[11], and coding [12]–[14]. The clipping and filtering technique causes bit-error-rate (BER) degradation [15]–[17]. Although the degradation is small for high clipping thresholds, clipping noise cancellation techniques are required to lower the BER degradation due to low clipping thresholds [18]–[21].

Manuscript received July 31, 2006; revised June 11, 2007 and July 11, 2007. The review of this paper was coordinated by Prof. W. Su.

The authors are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: wlq@ece.ualberta.ca; chintha@ece.ualberta.ca).

Digital Object Identifier 10.1109/TVT.2007.907282

The tone-reservation technique exploits a small number of unused subcarriers (reserved tones) to generate a peak-canceling signal [7]. This peak-canceling signal does not distort data-bearing subcarriers. This technique not only eliminates the need for side information but also prevents the BER degradation, as occurs with other techniques. Tone reservation requires the efficient generation of the peak-canceling signal. One method is iterative clipping and filtering under tone-reservation constraints [4]. In each iteration, the time-domain OFDM signal is clipped to a predefined threshold and filtered to eliminate the clipping noise on the data tones and outside the OFDM band. In this case, the peak-canceling signal is simply the filtered clipping noise. Although time-domain filtering via a bandpass filter [3] or frequency-domain filtering by exploiting a fast Fourier transform (FFT)/inverse FFT (IFFT) is feasible, the latter has significantly lower complexity than the former [4]. A drawback of iterative clipping and filtering is peak regrowth. Depending on the number and positions of reserved tones, peak regrowth may be large, and the convergence rate may be slow after several iterations.

The active-set algorithm [1] for tone reservation, which obtains moderate PAR reductions, also uses a peak-canceling kernel. This algorithm maintains an active set that contains the OFDM signal peaks. All peaks in the active set have the same magnitude. In each iteration, the largest peak outside the active set is reduced by the peak-canceling kernel to the same magnitude of the samples in the active set. Since the convergence rate slows down after several iterations, a tradeoff exists between the complexity and the performance.

In this paper, we analyze the clipping noise as a series of parabolic pulses under tone-reservation constraints. We first consider the case that clipping noise consists of a single pulse and generalize our analysis to the case of multiple pulses. Our analysis explains peak regrowth and the constant clipping noise power spectrum over the whole OFDM band. We also establish the roughly proportional relationship between the clipping noise at the end of several clipping and filtering iterations and that generated in the first iteration. The constant of proportionality is estimated via the level-crossing theory [22], [23].

Using the clipping noise analysis, we propose a constant-scaling algorithm and an adaptive-scaling algorithm for tone reservation. These algorithms scale the filtered first-iteration clipping noise by a constant or adaptively calculated factor to compensate for peaks that are above the threshold. The simulation results show that our proposed algorithms achieve a larger PAR reduction and lower complexity than the active-set algorithm.

Compared to previous works [5], [7], [24], our main contributions are as follows.

- 1) Our analysis is focused on the complex OFDM signal. Compared to the baseband real OFDM signal [7], [24], which has a Gaussian distribution, the complex OFDM signal has a Rayleigh-distributed envelope and a complex phase, which make the theoretical analysis more difficult.
- 2) We exploit a new model where the basic clipping pulse is approximated as a parabolic magnitude function multiplied by a linear phase function. We derive the distribution of phase change (in Appendixes B and E) and prove that the phase change is small and can be omitted (in Appendix D).
- 3) We prove all conditions that are used in our analysis. Although these conditions are intuitive, proving them is nontrivial.
- 4) We extend the frequency spectrum analysis from the single-pulse case [5] to the multiple-pulse case. Our analysis explains peak regrowth and the constant clipping noise power spectrum over the whole OFDM band.
- 5) We propose two algorithms to find the near-optimal scaling factor. Compared to the active-set algorithm [1], our algorithms obtain a larger PAR reduction with less complexity.
- 6) We propose a fast method to calculate the PAR and to find the clipping noise. We give a necessary condition for the large peaks. Only the samples that satisfy this condition need to be calculated. Since the number of such samples is small, the complexity of calculating the PAR and finding the clipping noise is small. This method can also be used in other PAR reduction techniques, such as selective mapping (SLM) and partial transmit sequences (PTS).

This paper is organized as follows. Section II describes the OFDM system and the tone-reservation technique. Section III formulates the basic problems that are treated in this paper. Section IV analyzes iterative clipping and filtering. Several conditions that are necessary for our analysis are proved in the appendixes. The two new algorithms are proposed in Section V. In Section VI, simulation results are used to compare the proposed algorithms and the active-set algorithm. The conclusions are given in Section VII.

## II. TONE-RESERVATION TECHNIQUE

### A. OFDM System Basics

In OFDM systems,  $N$  data symbols  $X_k$ ,  $k = -(N/2), -(N/2) + 1, \dots, (N/2) - 1$  are modulated on a set of  $N$  orthogonal subcarriers. The time-domain signal  $x(t)$  is

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} X_k e^{j2\pi kt/T}, \quad 0 \leq t \leq T \quad (1)$$

where  $N$  data symbols  $X_k$  form an OFDM symbol  $\mathbf{X} = [X_{-(N/2)}, \dots, X_{(N/2)-1}]$ , and  $T$  is the OFDM symbol

period. Samples of  $x(t)$  are efficiently computed via an inverse discrete Fourier transform (IDFT)<sup>1</sup>, i.e.,

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} X_k e^{j2\pi \frac{nk}{JN}}, \quad n = 0, \dots, JN - 1 \quad (2)$$

where  $J$  is the oversampling factor. A cyclic prefix is appended to  $x_n$  to combat the intersymbol interference.  $x(t)$  is obtained from the cyclicly prefixed  $x_n$  via digital-to-analog conversion.

The PAR of OFDM may be defined as

$$\xi = \frac{\max_{t \in [0, T]} |x(t)|^2}{P_{av}} \quad (3)$$

where  $P_{av} = E\{|x(t)|^2\} = E\{|X_k|^2\}$  is the average power. The PAR may also be computed using the discrete samples  $x_n$  similar to (3) and is approximately equal to  $\xi$  when  $J \geq 4$  [25], [26].

### B. Tone-Reservation Technique

The tone-reservation technique [7] reserves  $N_r$  tones for a PAR reduction and uses the remaining  $(N - N_r)$  tones for data transmission. The tone-reservation ratio  $R = N_r/N$  is typically small. The peak-canceling signal  $c(t)$  is generated based on reserved tones, and the peak-reduced signal is given by

$$\hat{x}(t) = x(t) + c(t) = \frac{1}{\sqrt{N}} \sum_{k=-N/2}^{N/2-1} (X_k + C_k) e^{j2\pi kt/T} \quad (4)$$

where  $0 \leq t \leq T$ , and  $\mathbf{C} = [C_{-(N/2)}, \dots, C_{(N/2)-1}]$  is the set of peak-canceling tones.  $\hat{x}(t)$  is amplified by the HPA and transmitted to the receiver. Denote  $\mathcal{C}$  as the signal space of all possible  $\mathbf{C}$  vectors.

Let  $\mathcal{R} = \{i_0, \dots, i_{N_r-1}\}$  be the locations of the reserved tones, where  $-(N/2) \leq i_0 < i_1 < \dots < i_{N_r-1} \leq (N/2) - 1$ . Let the index set  $\mathcal{R}^c$  be the complement of  $\mathcal{R}$  in  $\mathcal{N} = \{-(N/2), \dots, (N/2) - 1\}$ . The constraint on  $c(t)$  is that  $\mathbf{C}$  must satisfy  $C_k \equiv 0$  for  $k \in \mathcal{R}^c$ . On the other hand,  $\mathbf{X}$  must satisfy  $X_k \equiv 0$  for  $k \in \mathcal{R}$ .  $\mathbf{X}$  and  $\mathbf{C}$  are not allowed to be nonzero on the same subcarriers, i.e.,

$$X_k + C_k = \begin{cases} X_k, & k \in \mathcal{R}^c \\ C_k, & k \in \mathcal{R}. \end{cases} \quad (5)$$

Clearly, this technique reduces the normalized system throughput to  $(1 - R)$ . For a frequency-selective fading channel (ignoring nonlinear amplification), demodulation is done at a per-tone basis. Thus, (5) allows the reserved tones to be readily discarded. The BER of data tones with this method is the same as that of the original OFDM system.<sup>2</sup>

<sup>1</sup>In this paper, we use the zero-inserting scheme to calculate  $x_n$ , i.e., the IDFT operation is applied to the extended vector  $\mathbf{X}_{ext} = [X_0, \dots, X_{(N/2)-1}, \underbrace{0, \dots, 0}_{(J-1)N \text{ zeros}}, X_{-(N/2)}, X_{-(N/2)+1}, \dots, X_{-1}]$ .

<sup>2</sup>The BER of the whole system is, however, slightly increased due to slightly increased average transmit power.

Note that, with tone reservation, the PAR is redefined as

$$\xi = \frac{\max |x(t) + c(t)|^2}{E \left\{ |x(t)|^2 \right\}}. \quad (6)$$

That is, the peak-canceling signal  $c(t)$  is excluded from the calculation of the average power to prevent a large  $c(t)$  [7]. Thus,  $\mathbf{C}$  must be chosen to minimize the maximum of the time-domain signal, i.e.,

$$\mathbf{C}^{(\text{opt})} = \arg \min_{\mathbf{C} \in \mathcal{C}} \max_{0 \leq t \leq T} \left| \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} (X_k + C_k) e^{j2\pi kt/T} \right|^2. \quad (7)$$

Equation (7) can be reformulated as a quadratically constrained quadratic program (QCQP) [7], i.e.,

$$\begin{aligned} & \min_{\mathbf{C} \in \mathcal{C}} E \\ & \text{subject to: } |x_n + \mathbf{q}_n \mathbf{C}|^2 \leq E \end{aligned}$$

for  $n = 0, 1, \dots, JN - 1$ , where  $\mathbf{q}_n$  is the  $n$ th row of the IDFT matrix. Although the optimum of a QCQP exists, the solution requires a high computational cost of  $O(N_r JN^2)$ . Suboptimal solutions are typically employed.

### C. Optimization Techniques for Tone Reservation

The simplest optimization technique for tone reservation is iterative clipping and filtering [4]. In each iteration, this technique clips the OFDM signal to a predefined threshold  $A$ . The clipped signal is then filtered such that the clipping noise exists on reserved tones only. The convergence rate of this technique is slow. In the next section, we will analyze this technique and derive two new algorithms based on the analysis.

The active-set algorithm is proposed in [1] for tone reservation. This algorithm iteratively reduces the peaks under tone-reservation constraints. For a detailed description, see [1].

## III. PROBLEM FORMULATION

The filtered clipping noise must be scaled to form the peak-canceling signal  $c(t)$ . For clipping  $x(t)$  using a soft limiter (SL) [27], the clipped OFDM signal  $\tilde{x}(t)$  becomes

$$\tilde{x}(t) = \begin{cases} Ae^{j\theta(t)}, & |x(t)| > A \\ x(t), & |x(t)| \leq A \end{cases} \quad (8)$$

where  $A$  is the predefined threshold, and  $\theta(t)$  is the phase of  $x(t)$ . The clipping noise is given by

$$f(t) = x(t) - \tilde{x}(t). \quad (9)$$

The clipping noise  $f(t)$  consists of the segments of  $x(t)$ , where  $|x(t)|$  exceeds  $A$ . Unless  $A$  is small,  $f(t)$  is, thus, a series of

pulses, i.e.,

$$f(t) = \sum_{i=1}^{N_p} f_i(t)$$

where  $f_i(t)$  is the  $i$ th clipping pulse with pulse duration  $\tau_i$ , with its amplitude maximum at  $t_i$ , and  $N_p$  is the number of clipping pulses.

The filtered clipping noise  $\hat{f}(t)$  is obtained by passing  $f(t)$  through a filter whose passbands are on reserved tones. The peak-canceling signal is a scaled version of the filtered clipping noise, i.e.,  $c(t) = -\beta \hat{f}(t)$ , where  $\beta$  is the scaling factor to be optimized. One of our targets is to optimize  $\beta$  such that the PAR is minimized. Thus, the optimization problem is

$$\min_{\beta} \max_{0 \leq t \leq T} |x(t) - \beta \hat{f}(t)|^2. \quad (10)$$

We will first analyze clipping and filtering and reveal the mechanism of peak regrowth, i.e., the clipped peaks may grow and exceed  $A$  after filtering. The analysis of peak regrowth will facilitate the optimization of  $\beta$ . We will also give an analytical explanation of the flat spectrum of the clipping noise, which has been observed in [28] by simulation.

## IV. ANALYSIS OF CLIPPING AND FILTERING

### A. Time Domain Analysis of the Clipping Noise

Our analysis assumes that the real and imaginary parts of the input data symbol  $X_k$  are independent identically distributed (i.i.d.) random variables with zero mean and variance  $\sigma^2$ . We also assume that  $A$  and  $N$  are large,  $T$  is small, and the OFDM bandwidth  $W = N/T$  is a constant.

Let  $x(t) = x_R(t) + jx_I(t) = r(t)e^{j\theta(t)}$ , where  $x_R(t)$ ,  $x_I(t)$ ,  $r(t) \geq 0$ , and  $\theta(t)$  are the real and imaginary parts, the magnitude, and the phase of  $x(t)$ , respectively. Based on the central limit theorem,  $x_R(t)$  and  $x_I(t)$  are i.i.d. Gaussian random processes with zero mean and variance<sup>3</sup>  $\sigma^2$ ,  $r(t)$  is a Rayleigh process,  $\theta(t)$  is uniformly distributed between  $[0, 2\pi)$ , and  $r(t)$  is independent to  $\theta(t)$ .

The clipping noise  $f(t)$  is the consequence of upward level crossing of  $r(t)$  at level  $A$ . The level-crossing rate (the expected number of crossings of level  $A$  per second) can be found as [23]

$$\lambda_A = \frac{\dot{\sigma}}{\sqrt{2\pi}} \frac{A}{\sigma^2} e^{-A^2/2\sigma^2} \quad (11)$$

where [22]

$$\dot{\sigma}^2 = E \{ \dot{x}_R^2(t) \} = E \{ \dot{x}_I^2(t) \} = \frac{1}{2\pi} \int \omega^2 S(\omega) d\omega$$

<sup>3</sup>If  $N_r = RN$  tones are reserved, the variance of  $x_R(t)$  and  $x_I(t)$  is then equal to  $\sigma^2 R$ .

and  $S(\omega)$  is the power spectral density (PSD) of  $x_R(t)$  or  $x_I(t)$ . When  $N$  is large,  $S(\omega)$  is (approximately) constant over a fixed frequency band  $[-W/2, W/2]$ . We have

$$\dot{\sigma}^2 = \frac{(\pi N)^2 \sigma^2}{3T^2} = \frac{\pi^2}{3} W^2 \sigma^2. \quad (12)$$

Substituting (12) into (11), the level-crossing rate is

$$\lambda_A = \sqrt{\frac{\pi}{6}} \frac{A}{\sigma} \frac{N}{T} e^{-A^2/2\sigma^2}. \quad (13)$$

For large  $A$ , each up-crossing of level  $A$  leads to a clipping pulse. Therefore, the average number of clipping pulses during one OFDM signal period can be calculated as

$$\bar{N}_p = E\{N_p\} = \lambda_A T = N \sqrt{\frac{\pi}{6}} \frac{A}{\sigma} e^{-A^2/2\sigma^2}. \quad (14)$$

The clipping pulse duration  $\tau$  is a Rayleigh random variable with the following probability density function (pdf) [29]:

$$p(\tau) = \frac{\pi\tau}{2\bar{\tau}^2} \exp\left(-\frac{\pi\tau^2}{4\bar{\tau}^2}\right) \quad (15)$$

where  $\bar{\tau}$  is the mean of  $\tau$ . Since  $\lambda_A \bar{\tau} = \Pr[r(t) > A]$ ,  $\bar{\tau}$  can be calculated as

$$\bar{\tau} = \frac{\Pr[r(t) > A]}{\lambda_A} = \frac{\sigma^2 \sqrt{2\pi}}{\dot{\sigma} A} = \sqrt{\frac{6}{\pi}} \frac{\sigma}{AW}. \quad (16)$$

Let us consider a clipping pulse  $f_i(t)$  that reaches its maximum magnitude at  $t_i$  and has a time duration  $\tau_i$ . That is,  $f_i(t) = (r(t) - A)e^{j\theta(t)}$  within its pulse duration and is zero elsewhere. Equations (15) and (16) imply that, most probably,  $\tau$  is small in practical OFDM systems. Thus,  $r(t)$  can be approximated as a parabolic function using its Taylor series expansion at  $t = t_i$ . Let  $\Delta t_i = t - t_i$ , and note that  $r(t_i) > A$ ,  $\dot{r}(t_i) = 0$ , and  $\ddot{r}(t_i) < 0$ . We have

$$r(t) = r(t_i + \Delta t_i) \approx r(t_i) + \frac{1}{2} \ddot{r}(t_i) \Delta t_i^2. \quad (17)$$

$r(t_i + \Delta t_i)$  is symmetric to  $t_i$ . Then,  $r(t_i - \tau_i/2) \approx r(t_i + \tau_i/2) \approx A$ , and

$$\tau_i \approx \sqrt{-\frac{8(r(t_i) - A)}{\ddot{r}(t_i)}}. \quad (18)$$

Letting  $b_i = -\ddot{r}(t_i)$ , we have

$$r(t_i + \Delta t) - A \approx -\frac{1}{2} b_i \Delta t_i^2 + \frac{1}{8} b_i \tau_i^2, \quad -\frac{\tau_i}{2} \leq \Delta t_i < \frac{\tau_i}{2}. \quad (19)$$

Now, we consider the phase  $\theta(t) = \theta(t_i + \Delta t_i)$ . The phase change within  $-(\tau_i/2) \leq \Delta t_i \leq (\tau_i/2)$  is generally small. Note that the phase of  $f_i(t_i + \Delta t_i)$  is determined by all the constituent frequency components of  $x(t)$ , where  $x(t)$  is a

band-limited signal. The phase change of the  $k$ th frequency component, from  $t = t_i$  to  $t = t_i + (\tau_i/2)$ , is

$$\Delta\theta_k = 2\pi k \frac{\tau_i}{2T}, \quad k = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, \frac{N}{2} - 1.$$

Substituting  $\tau_i$  by  $\bar{\tau}$ , we find

$$\Delta\theta_k = \frac{\sqrt{6\pi} k \sigma}{NA}, \quad k = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, \frac{N}{2} - 1.$$

The largest phase change, which happens on  $k = -(N/2)$ , does not depend on  $N$ . By letting  $\theta_k = \theta_{(N/2)}$  for all  $k$ , the phase variation of  $f_i(t)$  from  $t = t_i$  to  $t = t_i + (\tau_i/2)$  is upper-bounded by  $\sqrt{6\pi}\sigma/2A$ . Clearly, the upper bound is quite loose, and the actual phase variation of  $f_i(t)$  is much smaller than this bound since some negative and positive phase changes may cancel each other out. Nevertheless, since this upper bound is small when  $A$  is large, we can approximate  $\theta(t_i + \Delta t_i) = \arcsin(x_I(t_i + \Delta t_i)/r(t_i + \Delta t_i))$  by its Taylor series expansion at  $t = t_i$ , i.e.,

$$\theta(t_i + \Delta t_i) \approx \theta_i + \gamma_i \Delta t_i$$

where  $\theta_i = \theta(t_i)$ , and  $\gamma_i = \dot{x}_I(t_i)/|x_R(t_i)|$ . Thus, we have

$$\begin{aligned} f_i(t) &= f_i(t_i + \Delta t_i) = (r(t_i + \Delta t_i) - A) e^{j\theta(t_i + \Delta t_i)} \\ &\approx \left(-\frac{1}{2} b_i \Delta t_i^2 + \frac{1}{8} b_i \tau_i^2\right) e^{j(\theta_i + \gamma_i \Delta t_i)}, \quad -\frac{\tau_i}{2} \leq \Delta t_i < \frac{\tau_i}{2}. \end{aligned}$$

The phase term  $\gamma_i \Delta t_i$  is most probably small and can be omitted. Appendix B gives the conditional pdf and moments of  $\gamma_i$  given  $\dot{r}(t_i) = 0$  and  $r(t_i) \geq A$ . Using these results, we have

$$E\{|\gamma_i|\} E\{\tau_i\} = \frac{\sqrt{2\pi}\sigma}{A} \operatorname{erfc}\left(\frac{A}{\sqrt{2}\sigma}\right) e^{A^2/2\sigma^2} \quad (20)$$

where

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

In Appendix D, we show that  $\gamma_i$  and  $\tau_i$  are uncorrelated. Thus, (20) indicates how small  $|\gamma_i \tau_i|$  usually is. For example,  $E\{|\gamma_i|\} E\{\tau_i\} \approx 0.07\pi$  when  $A = 6$  dB compared to the average power and is  $0.04\pi$  when  $A = 9$  dB. A rigid justification of small  $\gamma_i \Delta t_i$  requires the joint cumulative distribution function (cdf) of  $\gamma_i \tau_i$ , which, unfortunately, is difficult to derive. However, an upper bound for the phase change of  $f_i(t)$  can be found.

Since  $|f_i(t)|$  is close to zero when  $|\Delta t_i|$  is close to  $\tau_i/2$ , we may look at the phase change within the 6-dB width of  $f_i(t)$ , which is defined as the time duration that  $|f_i(t)|$  is no less than half of its maximum magnitude, and is equal to  $\tau_i/\sqrt{2}$ . By using the Chebyshev inequality, we have

$$\Pr\left[|\gamma_i \tau_i/\sqrt{2}| \geq \delta\right] \leq \frac{\sigma_0^2}{\delta^2}$$

where  $\sigma_0^2$  is the variance of  $\gamma_i \tau_i / \sqrt{2}$ , and  $\delta > 0$ . However, by using the Cauchy–Schwarz inequality

$$\sigma_0^2 = E \left\{ \frac{1}{2} \gamma_i^2 \tau_i^2 \right\} \leq \frac{1}{2} \sqrt{E \{ \gamma_i^4 \} E \{ \tau_i^4 \}}.$$

Denote the right-hand side of this inequality as  $\sigma_1^2$ . It is calculated by using the results of Appendix B as

$$\sigma_1^2 = \frac{2\sqrt{6}\sigma^3}{A^3} \sqrt{2 - E_1 \left( \frac{A^2}{2\sigma^2} \right) \frac{A^2}{\sigma^2} e^{A^2/2\sigma^2}}$$

where

$$E_1(x) = \int_1^\infty \frac{e^{-tx}}{t} dt.$$

Thus, we have

$$\Pr \left[ |\gamma_i \tau_i / \sqrt{2}| \geq \delta \sigma_1 \right] \leq \frac{1}{\delta^2}.$$

For example, by letting  $\delta = 3$ , the probability that the phase change within the 6-dB width of  $f_i(t)$  is larger than  $0.12\pi$  for  $A/\sqrt{2}\sigma = 6$  dB, or  $0.03\pi$  for  $A/\sqrt{2}\sigma = 9$  dB, is less than or equal to  $1/9$ . Therefore, we omit the phase term  $\gamma_i \tau_i$  and approximate the clipping pulse as a constant phase parabolic function

$$\begin{aligned} f_i(t) &= f_i(t_i + \Delta t_i) \\ &\approx \left( -\frac{1}{2} b_i \Delta t_i^2 + \frac{1}{8} b_i \tau_i^2 \right) e^{j\theta_i}, \quad -\frac{\tau_i}{2} \leq \Delta t_i < \frac{\tau_i}{2}. \end{aligned} \quad (21)$$

*Remark 1:* In our approximation, we implicitly assume that  $f_i(t)$  has only one local maximum (at  $t = t_i$ ). In other words,  $\ddot{r}(t_i)$  is always negative. Appendix C shows that  $\Pr[\dot{r}(t_i) > 0 | \dot{r}(t_i) = 0, r(t_i) \geq A] \rightarrow 0$  when  $A \rightarrow \infty$ , and in practical OFDM systems,  $\Pr[\dot{r}(t_i) > 0 | \dot{r}(t_i) = 0, r(t_i) \geq A] \approx 0$  unless  $A$  is very small. On the other hand, some other papers (e.g., [24]) approximate  $f_i(t)$  by expanding  $r(t)$  at  $t = t_i - (\tau_i/2)$ , where  $r(t_i - (\tau_i/2)) = A$ , and  $\dot{r}(t_i - (\tau_i/2)) \geq 0$ . Then,  $\tau_i$  becomes

$$\tau_i \approx -\frac{2\dot{r}(t_i - \tau_i/2)}{\ddot{r}(t_i - \tau_i/2)}$$

with an assumption that  $\ddot{r}(t_i - (\tau_i/2)) < 0$ . Although such an assumption holds for  $A \rightarrow \infty$ , simulation results show that it is frequently violated, even for  $A = 6$  dB.

### B. Frequency Domain Analysis of the Clipping Noise

The frequency spectrum of the Nyquist-rate sampled discrete-time real clipping noise is given in [24]. For the continuous-time complex clipping noise, the frequency spectrum of  $f_i(t)$  is the Fourier transform of (21), i.e.,

$$F_i(\omega) = e^{j(\theta_i - \omega t_i)} \frac{b_i \tau_i}{\omega^2} \left( \text{sinc} \frac{\omega \tau_i}{2} - \cos \frac{\omega \tau_i}{2} \right) \quad (22)$$

where  $\text{sinc } x = \text{sinc } x/x$ .  $F_i(\omega)$  distributes over the whole frequency band from  $\omega = -\infty$  to  $\infty$ .

Considering multiple pulses, the PSD of the clipping noise is

$$\begin{aligned} S_f(\omega) &= \frac{1}{T} E \left\{ |F(\omega)|^2 \right\} \\ &= \frac{1}{T} E \left\{ \sum_{i=1}^{N_p} |F_i(\omega)|^2 \right\} + \frac{1}{T} E \left\{ \sum_{i=1}^{N_p} \sum_{\substack{k=1 \\ k \neq i}}^{N_p} F_i(\omega) F_k^*(\omega) \right\}. \end{aligned}$$

Since  $N_p$  is a random variable, we cannot directly exchange the order of summation and expectation. However, by definition, we have

$$\begin{aligned} \frac{1}{T} E \left\{ \sum_{i=1}^{N_p} |F_i(\omega)|^2 \right\} &= \frac{1}{T} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n \sum_{i=1}^{N_{p,l}} |F_{i,l}(\omega)|^2 \\ &= \frac{1}{T} \lim_{n \rightarrow \infty} \frac{\sum_{l=1}^n N_{p,l}}{n} \\ &\quad \times \frac{\sum_{l=1}^n \sum_{i=1}^{N_{p,l}} |F_{i,l}(\omega)|^2}{\sum_{l=1}^n N_{p,l}} \\ &= \frac{\bar{N}_p}{T} E \left\{ |F_i(\omega)|^2 \right\} = \lambda_A E \left\{ |F_i(\omega)|^2 \right\} \end{aligned}$$

where the subscript  $l$  represents the  $l$ th trial. Therefore

$$S_f(\omega) = \lambda_A E \left\{ |F_i(\omega)|^2 \right\} + \frac{1}{T} E \left\{ \binom{N_p}{2} \right\} E \left\{ F_i(\omega) F_k^*(\omega) \right\}$$

where  $i \neq k$ . Note that

$$F_i(\omega) F_k^*(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_i(\hat{t}) f_k^*(\hat{t}) e^{-j\omega(\hat{t}-\tilde{t})} d\hat{t} d\tilde{t}.$$

$E \{ F_i(\omega) F_k^*(\omega) \}$  is determined by

$$\begin{aligned} E \{ f_i(\hat{t}) f_k^*(\tilde{t}) \} &= E \left\{ (r(t_i + \Delta t_i) - A) e^{j\theta(t_i + \Delta t_i)} \right. \\ &\quad \left. \times (r(t_k + \Delta t_k) - A) e^{-j\theta(t_k + \Delta t_k)} \right\}. \end{aligned}$$

However,  $r(t)$  and  $\theta(t)$  are independent. Thus, we have

$$\begin{aligned} E \{ f_i(\hat{t}) f_k^*(\tilde{t}) \} &= E \left\{ (r(t_i + \Delta t_i) - A) (r(t_k + \Delta t_k) - A) \right\} \\ &\quad \times E \left\{ e^{j\theta(t_i + \Delta t_i)} e^{j\theta(t_k + \Delta t_k)} \right\}. \end{aligned}$$

We show in Appendix A that, when  $(t_i + \Delta t_i)$  and  $(t_k + \Delta t_k)$  belong to different clipping pulses (which is true in our case),  $x(t_i + \Delta t_i)$  and  $x(t_k + \Delta t_k)$  are approximately independent, and, thus,  $\theta(t_i + \Delta t_i)$  and  $\theta(t_k + \Delta t_k)$  are uncorrelated. Then,  $E \{ F_i(\omega) F_k^*(\omega) \} = 0$ , and

$$S_f(\omega) = \lambda_A E \left\{ |F_i(\omega)|^2 \right\}. \quad (23)$$

The out-of-band radiation will be eliminated by filtering. Therefore, we are interested in the in-band clipping noise.

When  $A$  is large, generally,  $\omega\tau_i/2$  is small for  $|\omega| \leq \pi N/T$ . Thus, we may approximate  $F_i(\omega)$  as

$$F_i(\omega) \approx e^{j(\theta_i - \omega t_i)} \frac{b_i \tau_i^3}{12} \quad (24)$$

by using  $\text{sinc } x - \cos x \approx x^2/3$  [24]. Since  $F_i(\omega)$  does not depend on the frequency  $\omega$ ,  $S_f(\omega)$  is (approximately) constant over the OFDM band.

We may write  $b_i$  and  $\tau_i$  as a function of  $x_R(t_i)$ ,  $x_I(t_i)$ ,  $\dot{x}_I(t_i)$ ,  $\ddot{x}_R(t_i)$ , and  $\ddot{x}_I(t_i)$ . It is easy to find the joint pdf of these random variables. However, a closed-form expression of  $E\{b_i^2 \tau_i^6\}$  cannot be obtained.

### C. Clipping Noise PSD

In this subsection, we calculate the in-band clipping noise PSD by exploiting a result in [30]. Define  $y(t) = r^2(t)/\sigma^2$ ,  $\lambda = \dot{\sigma}^2/\sigma^2$ , and  $u = A^2/\sigma^2$ . If  $y(t)$  up-crosses the level  $u$  at  $t = 0$ , it has been shown that, with probability 1,  $y(t)$  around  $t = 0$  can be written as [30, Th. 2.2]

$$y(t) = -\lambda ut^2 + 2z\sqrt{\lambda ut} + u \quad (25)$$

when  $u \rightarrow \infty$ , where  $z$  is a Rayleigh random variable with a pdf

$$p(z) = ze^{-z^2/2}, \quad z > 0.$$

Also, the time duration  $\tau$  between this up-crossing and the successive down-crossing is [30, Th. 3.1]  $\tau = 2z/\sqrt{\lambda u}$  with probability 1 when  $u \rightarrow \infty$ . Since, most probably,  $\tau$  is very small for large  $A/\sigma$  in practical OFDM systems, we may use (25) to approximate the whole clipping pulse. Expanding  $r(t) = \sigma\sqrt{y(t)}$  using its Taylor series at  $t = 0$ , we have

$$\begin{aligned} r(t) &\approx \sigma\sqrt{u} + \sqrt{\lambda}\sigma z t - \frac{1}{2}\lambda\sigma \left(\sqrt{u} + \frac{z^2}{\sqrt{u}}\right) t^2 \\ &\approx \sigma\sqrt{u} + \sqrt{\lambda}\sigma z t - \frac{1}{2}\lambda\sigma\sqrt{u} t^2 \\ &= -\frac{\dot{\sigma}^2 A}{2\sigma^2} \left(t - \frac{\tau}{2}\right)^2 + \frac{A\dot{\sigma}^2 \tau^2}{8\sigma^2} + A, \quad 0 \leq t \leq \tau. \end{aligned} \quad (26)$$

The second step is obtained since  $z^2/\sqrt{u} \ll \sqrt{u}$  when  $u \rightarrow \infty$ .

*Remark 2:* In the proof of [30, Th. 2.2],  $y(t/\sqrt{u})$  is first approximated by approximating  $R(t)$  and its first derivative  $\dot{R}(t)$ , where  $R(t)$  is the correlation function of  $x_R(t)$  or  $x_I(t)$ , as polynomials of  $t$  with orders no larger than 2. Then, (25) is obtained for  $u \rightarrow \infty$  by letting all terms that contain  $u^{-v}$ , where  $v > 0$ , be zero. In this paper, it is easy to check that the same approximation of  $r(t)$  as in (26) is also obtained if  $y(t)$  is approximated by approximating  $R(t)$  and  $\dot{R}(t)$  with polynomials of orders higher than 2.

Now, we can approximate a clipping pulse  $f_k(t)$  that occurs in  $t_k \leq t \leq t_k + \tau_k$  as

$$\begin{aligned} f_k(t) &= |r(t_k + \Delta t_k) - A| e^{j(\theta(t_k + \Delta t_k))} \\ &\approx \left(-\frac{\dot{\sigma}^2 A}{2\sigma^2} \left(\Delta t_k - \frac{\tau_k}{2}\right)^2 + \frac{A\dot{\sigma}^2 \tau_k^2}{8\sigma^2}\right) e^{j(\theta_k + \eta_k \Delta t_k)} \end{aligned}$$

where  $0 \leq \Delta t_k \leq \tau$ ,  $\theta_k = \theta(t_k)$ , and

$$\eta_k = \frac{\dot{x}_I(t_k)A - x_I(t_k)\dot{r}(t_k)}{A|x_R(t_k)|}.$$

Appendix E shows that when  $A$  is large,  $\eta_k$  has the same distribution as the  $\gamma_i$  used in the previous subsection. Thus,  $\eta_k \Delta t_k$  is, most probably, small and can be ignored. Following the same procedure of the previous subsection, we have

$$F_k(\omega) \approx \frac{A\dot{\sigma}^2 \tau_k^3}{12\sigma^2} e^{j(\theta_k - \omega(t_k + \tau_k/2))}.$$

Then, the in-band clipping noise PSD is given by

$$S_f(\omega) = \lambda_A E \left\{ |F_k(\omega)|^2 \right\} = \frac{16\sqrt{2}e^{-A^2/2\sigma^2}}{\pi\sqrt{3\pi}(A/\sigma)^3} S_x \quad (27)$$

where  $S_x = 2\sigma^2/W$  is the PSD of the OFDM signal  $x(t)$ . For example, when  $A/\sqrt{2}\sigma = 6$  dB, the PSD of the in-band clipping noise is  $-27$  dB lower than that of the input OFDM signal.

### D. Filtered Clipping Noise

The in-band clipping noise falls on reserved and data tones. Whereas the clipping noise on reserved tones must be kept for a PAR reduction, the clipping noise on data tones and out-of-band radiation must be filtered such that the clipping noise would not interfere with the data symbols as well as communications on neighboring frequency bands. In this subsection, we first consider reserving  $N_r$  consecutive tones around the center frequency, i.e.,  $\mathcal{R} = \{-(N_r/2), -(N_r/2) + 1, \dots, (N_r/2) - 1\}$ . Other distributions of reserved tones will be discussed later. The result of the case that the clipping noise contains only one dominant pulse is similar to that in [5].

To filter the clipping noise, we use an ideal low-pass filter with the passband  $[-\omega_c, \omega_c]$ , where

$$\omega_c = 2\pi f_c = 2\pi \frac{N_r}{2T} = \pi R W.$$

The filtered clipping noise is then given by

$$\hat{f}(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \sum_{i=1}^{N_p} F_i(\omega) e^{j\omega t} d\omega. \quad (28)$$

Substituting (24) into (28), the filtered clipping noise may be expressed as

$$\begin{aligned} \hat{f}(t) &= \hat{f}(t_i + \Delta t_i) = \sum_{i=1}^{N_p} \hat{f}_i(t) \\ &= \sum_{i=1}^{N_p} e^{j\theta_i} \frac{b_i \tau_i^3 f_c}{6} \text{sinc } 2\pi f_c \Delta t_i. \end{aligned} \quad (29)$$

Without loss of generality, we consider the clipping pulse  $f_i(t)$  and assume that it occurs at  $t_i = 0$  and has the phase  $\theta_i = 0$ . Its filtered version is given by

$$\hat{f}_i(t) = \frac{b_i \tau_i^3 f_c}{6} \text{sinc } 2\pi f_c t. \quad (30)$$

Several observations can be made by comparing  $f_i(t)$  with  $\hat{f}_i(t)$ .

- 1)  $f_i(t)$  and  $\hat{f}_i(t)$  reach their peaks at the same time instant  $t = 0$ .
- 2)  $f_i(t)$  and  $\hat{f}_i(t)$  have the same direction within the pulse duration of  $f_i(t)$ .
- 3) The mainlobe duration of  $\hat{f}_i(t)$  is much wider than that of  $f_i(t)$ . Note that the mainlobe duration of  $\hat{f}_i(t)$  can be calculated as  $\hat{\tau} = 2T/N_r$ . By using (16), the ratio of the average clipping pulse duration  $\bar{\tau}$  over  $\hat{\tau}$  is

$$\frac{\bar{\tau}}{\hat{\tau}} = \sqrt{\frac{3}{2\pi}} \frac{N_r}{N(\frac{A}{\sigma})} \ll 1$$

when  $A/\sigma$  is large.

- 4) The sidelobe peaks of  $|\hat{f}_i(t)|$  are given by

$$|\hat{f}_i(T_k)| \approx \frac{b_i \tau_i^3 f_c}{3(2k+1)\pi} = \frac{2}{(2k+1)\pi} |\hat{f}_i(t)|_{\max} \quad (31)$$

where  $k = 1, 2, 3, \dots$ , and  $T_k$  is the sidelobe peak occurrence time, i.e.,

$$T_k \approx \frac{(2k+1)T}{2N_r} = \frac{2k+1}{2RW}. \quad (32)$$

For example, the peak of the first sidelobe is only 21.22% as that of the mainlobe.

- 5) The maximum of  $\hat{f}_i(t)$  is much less than that of  $f_i(t)$ . In fact

$$|\hat{f}_i(t)|_{\max} = \alpha \tau_i |f_i(t)|_{\max}$$

where  $\alpha$  is defined as

$$\alpha = \frac{4}{3} f_c = \frac{2RW}{3} \quad (33)$$

and the expectation of  $\alpha \tau_i$  is

$$E\{\alpha \tau_i\} = \alpha \bar{\tau} = R \frac{2\sqrt{2}}{\sqrt{3\pi}} \frac{\sigma}{A} \ll 1$$

when  $A/\sigma$  is large. For example,  $\alpha \bar{\tau} \approx 1.63\%$  when  $A/\sqrt{2}\sigma = 6$  dB, and  $R = 5\%$ .

Point 5 explains why the peak regrows after filtering. Recall that the clipped signal is  $\tilde{x}(t) = x(t) - f(t)$ , and after filtering, it becomes  $\hat{x}(t) = x(t) - \hat{f}(t)$ . Whereas  $f(t)$  is chosen such that  $|\tilde{x}(t_i)| = A$  at the peaks of  $x(t)|_{t=t_i}$ , the clipped and filtered signal  $\hat{x}(t_i) > A$  since  $|\hat{f}_i(t)|_{\max} < |f_i(t)|_{\max}$ .

### E. Iterative Clipping and Filtering: Single Clipping Pulse

To suppress the peak regrowth, clipping and filtering may be repeated until a suitable criterion is met. We first temporarily assume that the clipping noise at the first iteration  $f^{(1)}(t)$  consists of only one dominant clipping pulse  $f_i^{(1)}(t)$  (with pulse duration  $\tau_i^{(1)}$ ) that is much larger than other clipping pulses. In this case, the clipped and filtered OFDM signal after the first iteration  $\hat{x}^{(1)}(t) \approx x(t) - \hat{f}_i^{(1)}(t)$ , where  $\hat{f}_i^{(1)}(t)$  is the filtered version of  $f_i^{(1)}(t)$ , can be partitioned into three parts, as follows.

- 1)  $|t| \leq \tau_i^{(1)}/2$ : Within this range,  $\hat{f}_i^{(1)}(t)$ ,  $f_i^{(1)}(t)$ , and  $x(t)$  have the same phase, and  $|\hat{f}_i^{(1)}(t)| < |f_i^{(1)}(t)|$ . Therefore

$$|\hat{x}^{(1)}(t)| \approx |x(t) - \hat{f}_i^{(1)}(t)| > |x(t) - f_i^{(1)}(t)| = A.$$

In other words, after passing  $\hat{x}^{(1)}(t)$  through the SL, a clipping pulse, which is denoted as  $f_i^{(2)}(t)$ , occurs in the second clipping iteration at the same position as  $f_i^{(1)}(t)$ . By applying Taylor series expansion to (30) and since  $(\omega_c \tau_i^{(1)}/2) \ll 1$ , we can approximate the filtered clipping pulse  $\hat{f}_i^{(1)}(t)$  as a constant, i.e.,

$$\hat{f}_i^{(1)}(t) \approx \frac{b_i (\tau_i^{(1)})^3 f_c}{6} = |f_i^{(1)}(t)|_{\max}, \quad |t| \leq \frac{\tau_i^{(1)}}{2}.$$

Then, the clipping pulse at the second iteration  $f_i^{(2)}(t)$  can be written as

$$\begin{aligned} f_i^{(2)}(t) &= f_i^{(1)}(t) - \hat{f}_i^{(1)}(t) \\ &\approx -\frac{1}{2} b_i t^2 + \frac{1}{8} b_i (\tau_i^{(1)})^2 - \frac{b_i (\tau_i^{(1)})^3 f_c}{6} \end{aligned} \quad (34)$$

which is also a parabolic arc with reduced magnitude. By solving  $f_i^{(2)}(t) = 0$ , the time duration of  $f_i^{(2)}(t)$  can be found as

$$\tau_i^{(2)} = \tau_i^{(1)} \sqrt{1 - \frac{4}{3} \tau_i^{(1)} f_c} = \tau_i^{(1)} \sqrt{1 - \alpha^{(1)} \tau_i^{(1)}}$$

where  $\alpha^{(1)}$  is the  $\alpha$  defined in (33).

- 2)  $(\tau_i^{(1)}/2) < |t| < T_2$ , where  $T_2$  is given in (32): In this range,  $|x(t)| < A$  since only one clipping pulse exists. However, depending on the phase of  $\hat{f}_i^{(1)}(t)$ ,  $|\hat{x}^{(1)}(t)|$  may be greater than  $A$ . In other words, new clipping pulses may be generated in the second clipping iteration. However, since  $|\hat{f}_i^{(1)}(t)|_{\max} \ll |f_i^{(1)}(t)|_{\max}$ , these new clipping pulses are very small compared to the clipping pulse  $f_i^{(2)}(t)$ , and their effects can be ignored.

- 3)  $|t| > T_2$ : Since the peaks of  $\hat{f}_i^{(1)}(t)$  decay with the rate of  $1/t$ , we can see that, in this range,  $\hat{x}^{(1)}(t) \approx x(t)$ . Therefore, no clipping pulses exist at  $|t| > T_2$  in the second clipping iteration.

The successive clipping and filtering iterations repeat this procedure. Therefore, we conclude that for the case of only one dominant clipping pulse, in the  $l$ th ( $l = 2, 3, \dots$ ) clipping

and filtering iteration,  $f_i^{(l-1)}(t)$  shrinks to  $f_i^{(l)}(t)$ , and some new pulses possibly appear. Here,  $f_i^{(l-1)}(t)$  and  $f_i^{(l)}(t)$  are the dominant clipping pulses at the  $(l-1)$ th and  $l$ th iterations, respectively. Until  $f_i^{(l)}(t)$  is comparable to the new pulses, the latter can be omitted, and the former can be written as

$$f_i^{(l)}(t) = -\frac{1}{2}b_i t^2 + \frac{1}{8}b_i \left(\tau_i^{(l)}\right)^2, \quad i = 1, 2, 3, \dots \quad (35)$$

where

$$\tau_i^{(l)} = \tau_i^{(l-1)} \sqrt{1 - \alpha \tau_i^{(l-1)}}, \quad i = 2, 3, 4, \dots \quad (36)$$

and  $\alpha$  is defined in (33). Moreover, the filtered clipping pulse in the  $l$ th iteration is given by

$$\hat{f}_i^{(l)}(t) = \frac{b_i \left(\tau_i^{(l)}\right)^3 f_c}{6} \text{sinc } 2\pi f_c t. \quad (37)$$

Thus, the filtered clipping noise that is generated in the  $l$ th iteration is proportional to that generated in the first iteration. Therefore, we have the following.

In iterative clipping and filtering, the total filtered clipping noise is proportional to that generated in the first iteration. Define  $\beta$  as follows:

$$\beta \triangleq \frac{\text{total filtered clipping noise after } K \text{ iterations}}{\text{filtered clipping noise generated in the first iteration}}.$$

If only one dominant clipping pulse exists, then

$$\beta = \frac{\sum_{l=1}^K \hat{f}_i^{(l)}(t)}{\hat{f}_i^{(1)}(t)} = \frac{\sum_{l=1}^{K-1} \left(\tau_i^{(l)}\right)^3}{\left(\tau_i^{(1)}\right)^3}. \quad (38)$$

It is difficult to find  $\bar{\beta}$ , the mean of  $\beta$ . However, an estimation of  $\bar{\beta}$  can be obtained when  $K$  is not large. When  $A$  is large,  $\alpha \tau_i^{(l)} \ll 1$ . Then,  $\sqrt{1 - \alpha \tau_i^{(l)}}$  can be treated as a constant.  $\bar{\beta}$  can then be estimated by replacing  $\tau_i^{(1)}$  with its mean  $\bar{\tau}$ . Thus

$$\bar{\beta} \approx \frac{1 - (1 - \alpha \bar{\tau})^{3K/2}}{1 - (1 - \alpha \bar{\tau})^{3/2}}. \quad (39)$$

We will use  $\bar{\beta}$  in the constant-scaling algorithm proposed in the next section.

*Remark 3:* By noting that  $\pi \tau_i W$  is very small with high probability when  $A$  is large, it is easy to show that (38) and (39) also apply to other reserved tone sets.

#### F. Iterative Clipping and Filtering: Multiple Clipping Pulses

In OFDM systems, the (unfiltered) clipping noise is usually a series of parabolic pulses. In each clipping and filtering iteration, the filtered clipping noise at  $t_i - \tau_i^{(l)}/2 \leq t \leq t + \tau_i^{(l)}/2$  is the mainlobe of  $\hat{f}_i^{(l)}(t)$  plus the mainlobes (if they are close to  $t = t_i$ ) or sidelobes (if they are far from  $t = t_i$ ) of all other filtered pulses at  $t_i - \tau_i^{(l)}/2 \leq t \leq t + \tau_i^{(l)}/2$ .

If a clipping pulse contributes an impact that is stronger than its  $k$ th sidelobe, where  $k$  is properly chosen such that the  $k$ th sidelobe is relatively large and cannot be omitted, it must occur within the time interval  $t_i - T_k \leq t \leq t_i + T_k$ . Equation (64) in Appendix A gives the probability that two or more clipping pulses occur within a time interval. Substituting (32) into (64), we can see that such a probability is independent to  $N$ . In other words, with a fixed probability, the number of clipping pulses that occur within the time interval  $t_i - T_k \leq t \leq t_i + T_k$  is independent to  $N$ . However, while  $N \rightarrow \infty$ , the average number of clipping pulses also goes to  $\infty$  [cf. (14)]. Therefore, the effect of most other filtered pulses at  $t_i - \tau_i^{(l)}/2 \leq t \leq t + \tau_i^{(l)}/2$  can be omitted. We only need to consider the pulses that are close to the mainlobe of  $\hat{f}_i^{(l)}(t)$ .

After the  $l$ th iteration, the peak-reduced OFDM signal at  $t = t_i$  becomes

$$\hat{x}^{(l+1)}(t_i) = \left( A + \left| f_i^{(l)}(t_i) \right| - \left| \hat{f}_i^{(l)}(t_i) \right| \right) e^{j\theta^{(l)}(t_i)} - \sum_{m \neq i} \hat{f}_m^{(l)}(t_i)$$

where  $\theta^{(l)}(t_i)$  is the phase of  $\hat{x}^{(l)}(t_i)$ . When  $A$  is very large, for example,  $A \geq 9$  dB, all  $f_m^{(l)}(t)$  are far apart from  $f_i^{(l)}(t)$ , and  $|\hat{f}_m^{(l)}(t_i)| \approx 0$  for all  $m \neq i$ . In this case, our conclusions in the previous section hold. It is easy to show that, with  $l \rightarrow \infty$ , the clipping noise at the  $l$ th iteration  $f^{(l)}(t) \rightarrow 0$ , and the peak of  $x(t)$  at  $t = t_i$  is reduced to  $A$ .

On the other hand, for moderate  $A$ , some of  $\hat{f}_m^{(l)}(t_i)$  may be relatively large and cannot be omitted. In this case, we decompose  $\hat{f}_m^{(l)}(t_i)$  as

$$\hat{f}_m^{(l)}(t_i) = \left( \hat{f}_{m,I}^{(l)}(t_i) + j \hat{f}_{m,Q}^{(l)}(t_i) \right) e^{j\theta^{(l)}(t_i)}$$

where  $\hat{f}_{m,I}^{(l)}(t_i)$  and  $\hat{f}_{m,Q}^{(l)}(t_i)$  are the inphase and quadrature components along the direction of  $\hat{x}^{(l)}(t_i)$ , respectively. By noting that with high probability,  $A$  is much larger than any clipping pulse,  $\hat{f}_{m,Q}^{(l)}(t_i)$  can be omitted when calculating  $\hat{x}^{(l+1)}(t_i)$ . Then, in the  $(l+1)$ th iteration, the clipping pulse at  $t = t_i$  is

$$f_i^{(l+1)}(t_i) \approx \left( \left| f_i^{(l)}(t_i) \right| - \left| \hat{f}_i^{(l)}(t_i) \right| - \sum_{m \neq i} \hat{f}_{m,I}^{(l)}(t_i) \right) e^{j\theta^{(l)}(t_i)}.$$

Depending on the sign of  $\sum_{m \neq i} \hat{f}_{m,I}^{(l)}(t_i)$ , the peak reduction may be strengthened or weakened. Moreover,  $|f_i^{(l+1)}(t_i)| > |f_i^{(l)}(t_i)|$ , i.e., the peak is *increased*, when  $\sum_{m \neq i} \hat{f}_{m,I}^{(l)}(t_i) < -|\hat{f}_i^{(l)}(t_i)|$ .

*Remark 4:* In the multiple clipping pulse case, the validity of (38) and (39) depends on  $A$ . That is, (38) and (39) are valid when  $A$  is large. Otherwise, the estimation error of these two equations is relatively large. However, the above analysis is still valid, and the total filtered clipping noise is still proportional to that generated in the first iteration until  $A$  is small such that the width of the mainlobe of  $F_i(\omega)$  is comparable to or smaller than  $W$ .

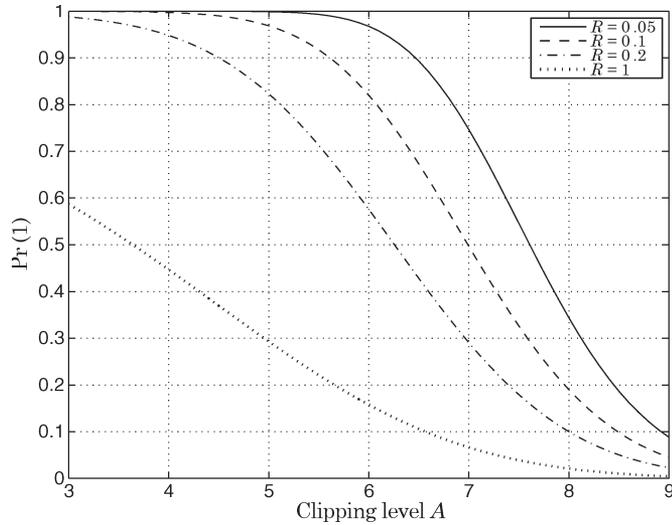


Fig. 1. Probability that more than one clipping pulse occur with a time duration of  $T_4$ , given that a clipping pulse has already occurred in this time interval, where  $\sigma = 1/\sqrt{2}$ .

*Remark 5:* A wider mainlobe of a filtered clipping pulse implies that its magnitude would be interfered by more neighboring clipping pulses, which, in turn, implies worse PAR reduction performance. Therefore, it is desired to make the filtered clipping pulse close to an impulse function, i.e., be of smallest width of mainlobe and lowest sidelobes.

It may be worthwhile to check at what level of  $A$  (38) and (39) are valid. Such a level of  $A$ , denoted as  $A_{\text{thres}}$ , depends on the choice of reserved-tone set. Reference [7] proves that a close-to-optimum reserved-tone set can be found from a small number of randomly selected reserved-tone sets, which implies that consecutive reserved tones usually lead to nonoptimal solutions. Therefore, we may use the consecutive reserved tones to find  $A_{\text{thres}}$ . In other words, if (38) and (39) are valid for consecutive reserved tones when  $A \geq A_{\text{thres}}$ , they are also valid for most other reserved-tone sets when  $A \geq A_{\text{thres}}$ .

Equation (32) indicates that the width of the mainlobe and sidelobes of the filtered clipping pulse  $\hat{f}_i(t)$  is only determined by the tone-reservation ratio  $R$  and the OFDM bandwidth  $W$ . Therefore, we may use  $T_k$  as a reference such that the tail of  $\hat{f}_i(t)$  beyond  $T_k$  is small and can be ignored. From (31), the peak of the fourth sidelobe is only 7% of that of the mainlobe. Therefore, if any two pulses are apart by at least  $T_4$  seconds, (38) and (39) are valid. Substituting  $T_4$  into (66) in Appendix A, the probability that more than one clipping pulse occur with a time duration of  $T_4$ , given that a clipping pulse has already occurred in this time interval, is

$$\Pr(1) = 1 - e^{-T_4 \lambda_A}.$$

Fig. 1 illustrates the relationship of  $\Pr(1)$  and  $A$  for  $R = 0.05, 0.1, \text{ and } 0.2$ , where  $\sigma = 1/\sqrt{2}$  and  $T_4$  are used. We also include  $R = 1$  as a reference, which we refer to as the iterative clipping and filtering technique [3], where no tone is reserved, and the clipping noise is distributed over the whole OFDM band.

For  $R = 0.05, 0.1, \text{ and } 0.2$ , (38) and (39) can be used with small approximation error when  $A \geq 9, 8.5, \text{ and } 8$  dB, respectively. However, if no tone is reserved, and the clipping noise is distributed over the whole OFDM band ( $R = 1$ ), the two equations are valid when  $A \geq 6.5$  dB.

*G. Effect of the Reserved-Tone Position on PAR Reduction*

Let us consider using an ideal bandpass filter with passbands on reserved tones  $\mathcal{R}$  only<sup>4</sup> to filter the  $i$ th clipping pulse  $f_i(t)$ . We assume that  $f_i(t)$  occurs at  $t_i = 0$  and has a phase  $\theta_i = 0$ . The filtered clipping pulse is given by

$$\hat{f}_i(t) = \frac{1}{2\pi} \int_{\mathcal{R}} \frac{b_i(\tau_i)^3}{12} e^{j\omega t} d\omega = \frac{b_i(\tau_i)^3}{12T} h(t) \quad (40)$$

where

$$h(t) = \text{sinc} \frac{\pi t}{T} \sum_{k \in \mathcal{R}} e^{j2\pi k \frac{t}{T}} \quad (41)$$

is the impulse response of the bandpass filter, and  $-T \leq t \leq T$  since the time duration of an OFDM symbol is  $T$ . When  $\mathcal{R} = \{-N_r/2, -(N_r/2) + 1, \dots, (N_r/2) - 1\}$ , (40) reduces to (30).

To make  $\hat{f}_i(t)$  close to an impulse function, we only need to consider the magnitude response  $|h(t)|$ . Some conclusions can be obtained from (41).

- 1) The mainlobe of  $h(t)$  is of maximum width when  $N_r = 1$ , implying the worst PAR reduction capability.
- 2) For consecutive reserved tones, the mainlobe of  $h(t)$  is of minimum width when  $N_r = N$ , implying the best PAR reduction capability.

If we choose another set of reserved tones  $\mathcal{R}'$  that is a shift of  $\mathcal{R}$ , i.e.,  $\mathcal{R}' = \mathcal{R} + n_s$ , where  $\mathcal{R}$  is of any kind, and  $n_s$  is an integer, then the impulse response of  $\mathcal{R}'$  is

$$\begin{aligned} |h'(t)| &= \left| \text{sinc} \frac{\pi t}{T} \sum_{k \in \mathcal{R}'} e^{j2\pi k \frac{t}{T}} \right| \\ &= \left| e^{j2\pi n_s \frac{t}{T}} \text{sinc} \frac{\pi t}{T} \sum_{k \in \mathcal{R}} e^{j2\pi k \frac{t}{T}} \right| = |h(t)|. \end{aligned} \quad (42)$$

That is,  $h'(t)$  has the same magnitude response as  $h(t)$ .

- 3) From point 2, shifting  $\mathcal{R}$  could not change the PAR reduction capability.

In (41), the term  $\text{sinc}(\pi t/T)$  renders the envelope of  $h(t)$ , and the width of the mainlobe of  $h(t)$  is determined by

$$g(t) = \sum_{k \in \mathcal{R}} e^{j2\pi k \frac{t}{T}}.$$

<sup>4</sup>Although  $\mathcal{R}$  is a discrete set, here, we slightly misuse  $\mathcal{R}$  to represent the reserved tones in the continuous frequency domain for simplicity. In this case, each item  $i \in \mathcal{R}$  represents a frequency band with width  $1/T$  and central frequency  $i/T$ .

<sup>5</sup>The actual range of  $t$  is  $-t_k \leq t \leq T - t_k$ . Considering  $0 \leq t_k \leq T$ , we have  $-T \leq t \leq T$ .

Let  $\mathbf{b} = [b_{-N/2}, b_{-(N/2)+1}, \dots, b_{(N/2)-1}]$  be the indicator of reserved tones, i.e.,  $b_k = 1$  when  $k \in \mathcal{R}$  and is 0 otherwise. Then

$$g(t) = \sum_{k=-N/2}^{N/2-1} b_k e^{j2\pi k \frac{t}{T}}.$$

Since the phase of  $h(t)$  is irrelevant to our consideration, we can focus on  $|g(t)|^2$ . However

$$|g(t)|^2 = \sum_{n=-N+1}^{N-1} \Psi(n) e^{j2\pi nt/T} \quad (43)$$

where

$$\Psi(n) = \begin{cases} \sum_{k=-\frac{N}{2}-n}^{\frac{N}{2}-1-n} b_k b_{k+n}, & n \geq 0 \\ \sum_{k=-\frac{N}{2}-n}^{\frac{N}{2}-1} b_k b_{k+n}, & n < 0 \end{cases} \quad (44)$$

is the aperiodic autocorrelation function of  $\mathbf{b}$ . Note that the width of  $\Psi(n)$  is inversely proportional to that of  $|g(t)|^2$ .

- 4) For fixed  $N_r$ , a contiguous set of reserved tones leads to a small PAR reduction because that set has the narrowest width of  $\Psi(n)$  compared to other sets. A randomly chosen reserved-tone set usually leads to better PAR reductions than a consecutive reserved-tone set [7].

To find the optimal positioning scheme, we may optimize  $\mathcal{R}$  such that the maximum peak in  $|t| > \bar{\tau}/2$  is minimized, i.e.,

$$\min_{\mathcal{R}} \max_{T \geq |t| > \frac{\bar{\tau}}{2}} |h(t)| \quad (45)$$

which is equivalent to

$$\begin{aligned} & \min_{\mathbf{b}} \max_{T \geq |t| > \frac{\bar{\tau}}{2}} \left| g(t) \operatorname{sinc} \frac{\pi t}{T} \right| \\ & \text{subject to: } \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} b_k = N_r. \end{aligned} \quad (46)$$

$\operatorname{sinc}(\pi t/T)$  is only a weighting factor, and  $g(t)$  is periodic with a period of  $T$ . Moreover,  $|g(t)| = |g(-t)|$  and  $|\operatorname{sinc}(\pi t/T)| = |\operatorname{sinc}(-(\pi t/T))|$  since  $b_i$  are integers. Then, (46) is equivalent to

$$\begin{aligned} & \min_{\mathbf{b}} \max_{T/2 \geq |t| > \bar{\tau}/2} \left| g(t) \operatorname{sinc} \frac{\pi t}{T} \right| \\ & \text{subject to: } \sum_{k=-N/2}^{N/2-1} b_k = N_r. \end{aligned} \quad (47)$$

Sampling the objective function with the sampling frequency  $JN/T$ , where  $J$  is the oversampling factor, (47) becomes

$$\begin{aligned} & \min_{\mathbf{b}} \max_{n_r \leq n \leq JN/2-1} \left| g(n) \operatorname{sinc} \frac{\pi n}{JN} \right| \\ & \text{subject to: } \sum_{k=-N/2}^{N/2-1} b_k = N_r \end{aligned} \quad (48)$$

where

$$g(n) = \sum_{k=-N/2}^{N/2-1} b_k e^{j2\pi \frac{kn}{JN}}, \quad n = 0, 1, \dots, \frac{JN}{2} - 1$$

and  $n_\tau = \lceil J\sqrt{(6/\pi)}(\sigma/A) \rceil$ , with  $\lceil x \rceil$  representing the minimum integer that is greater than  $x$ . If the factor  $\operatorname{sinc}(\pi n/JN)$  is omitted, and let  $n_\tau = J$ , (48) reduces to the optimization of  $\mathcal{R}$  proposed in [7].

The search space of (48) is  $\binom{N}{N_r}$ , which is prohibitively large when  $N$  and  $N_r$  are large. Alternatively, we can use the random set optimization [7] to find a suboptimal solution  $\mathcal{R}^*$ , i.e., selecting the best from  $M$  randomly generated reserved-tone sets  $\mathcal{R}_1, \dots, \mathcal{R}_M$ .

## V. NEW TONE-RESERVATION ALGORITHMS

Using this analysis, we now propose two new tone-reservation algorithms for the PAR reduction. The key idea is that since the clipping noise in all the iterations are similar, the total filtered clipping noise can be approximated by scaling the filtered clipping noise that is generated in the first iteration.

### A. Constant-Scaling Tone-Reservation Algorithm

If a relatively high PAR is tolerable, but low computational complexity is a must, we propose a constant-scaling tone-reservation algorithm. This algorithm scales the filtered clipping noise by a constant factor  $\bar{\beta}$  and subtracts the scaled clipping noise from the original OFDM symbol. This algorithm can be stated as follows:

#### Algorithm 1 (constant scaling)

##### Initialization:

Note that this stage only needs to run once.

- 1) Choose a relatively high clipping threshold  $A$  and randomly choose the reserved-tone set  $\mathcal{R}$  or set it up using random set optimization.
- 2) Choose a  $K$  and calculate  $\bar{\beta}$  using (39).

##### Runtime:

- 1) Distribute  $(N - N_r)$  input symbols to data tones  $\mathcal{R}^c$  and calculate the corresponding time-domain signal  $x_n$  using (2). Note that oversampling may be required.
- 2) If  $\text{PAR} > A$ , go to step 3; otherwise, transmit  $x_n$  and terminate.
- 3) Clip  $x_n$  to the threshold  $A$  to find the clipping noise  $f_n$  using the discrete-time version of (9).
- 4) Filter  $f_n$  subject to tone-reservation constraints and obtain the peak-canceling signal  $c_n$ :
  - a) Convert  $f_n$  to the frequency domain to obtain  $F_k = \text{DFT}\{f\}$ , where  $\mathbf{f} = [f_0, \dots, f_{JN-1}]$ .
  - b) Obtain the filtered clipping noise  $\hat{\mathbf{F}}$  by keeping the items of  $F_k$  for  $k \in \mathcal{R}$ , and set other  $F_k$  to zero.
  - c) The peak reduction signal  $C_k$  is then  $C_k = -\bar{\beta}\hat{F}_k$ . Note that scaling the filtered noise in the frequency

domain has lower complexity than scaling it in the time domain.

- d) Convert  $C_k$  to time domain to obtain  $c_n$  using an IDFT operation.
- 5) Calculate the PAR-reduced OFDM signal as  $\hat{x}_n = x_n + c_n$ , and transmit it.

*Remark 6:* Calculating the PAR and finding  $f_n$  require the calculation of  $|x_n|$ , which is costly if all  $|x_n|$  are calculated. In this paper, we propose a method to reduce such cost. Note that we only need to calculate  $|x_n|$  for those  $|x_n| \geq A$ . A necessary condition of  $|x_n| \geq A$  is as follows:

$$\left( |x_{n,R}| \geq \frac{A}{\sqrt{2}} \text{ or } |x_{n,I}| \geq \frac{A}{\sqrt{2}} \right) \text{ and } (|x_{n,R}| + |x_{n,I}| \geq A) \quad (49)$$

where  $x_{n,R}$  and  $x_{n,I}$  are the real and imaginary parts of  $x_n$ , respectively. Thus, we only need to calculate  $|x_n|$  for the samples satisfying (49). We will later show that the number of samples satisfying (49) is small. This method can also be used in other PAR reduction techniques, such as SLM and PTS.

### B. Adaptive-Scaling Tone-Reservation Algorithm

When a large PAR reduction is required, we propose an adaptive-scaling tone-reservation algorithm. Instead of using fixed  $\beta$ , this algorithm calculates  $\beta$  for each OFDM symbol.

A discrete-time-domain description of the algorithm is given here. Filtering the clipping noise  $f_n$  to  $\mathcal{R}$ , we obtain the filtered clipping noise  $\hat{f}_n = \text{IDFT}\{\hat{\mathbf{F}}\}$ . The PAR reduced signal  $\hat{x}_n$  can be written as follows:

$$\hat{x}_n = x_n - \beta \hat{f}_n = Ae^{j\theta_n} + f_n - \beta \hat{f}_n \quad (50)$$

where  $\theta_n$  is the phase of  $x_n$ . Our task is to minimize the out-of-range power  $P$ , i.e., the total power of those  $|\hat{x}_n| > A$ . The objective function is

$$\min_{\beta} P \quad (51)$$

where

$$P = \sum_{|\hat{x}_n| > A} (|\hat{x}_n| - A)^2. \quad (52)$$

Equation (52) can be rewritten as follows:

$$P = \sum_{n \in \mathcal{S}_1} (|\hat{x}_n| - A)^2 - \sum_{\substack{n \in \mathcal{S}_1 \\ |\hat{x}_n| \leq A}} (|\hat{x}_n| - A)^2 + \sum_{n \in \mathcal{S}_2} (|\hat{x}_n| - A)^2 \quad (53)$$

where  $\mathcal{S}_1 = \{n : |f_n| > 0\}$  is the index set of all clipping pulses, and  $\mathcal{S}_2 = \{n : |f_n| = 0 \text{ and } |\hat{x}_n| > A\}$ . Since clipping pulses are parabolic arcs, the power of any clipping pulse is a

monotonic function of its peak amplitude. Minimizing (53) is equivalent to minimizing

$$\hat{P} = \underbrace{\sum_{n \in \mathcal{S}_p} (|\hat{x}_n| - A)^2}_{P_1} - \underbrace{\sum_{\substack{n \in \mathcal{S}_p \\ |\hat{x}_n| \leq A}} (|\hat{x}_n| - A)^2}_{P_2} + \underbrace{\sum_{n \in \mathcal{S}_p^+} (|\hat{x}_n| - A)^2}_{P_3} \quad (54)$$

where  $\mathcal{S}_p = \{n : n \in \mathcal{S}_1, |x_n| > |x_{n-1}|, \text{ and } |x_n| \geq |x_{n+1}|\}$  is the index set of the peaks of  $f_n$ , and  $\mathcal{S}_p^+ = \{n : n \in \mathcal{S}_2, |\hat{x}_n| > |\hat{x}_{n-1}| \text{ and } |\hat{x}_n| \geq |\hat{x}_{n+1}|\}$  is the index set of the peaks of newly generated pulses whose amplitudes are larger than  $A$ . That is, if  $\hat{P}$  is minimized,  $P$  is also close-to-optimally minimized.

Equation (54) implies that the optimal  $\beta$ , denoted as  $\beta^{(\text{opt})}$ , must minimize the peaks of  $x_n$  and prevent any large newly generated pulses, which, in turn, implies that  $\beta$  should not be large. Consequently,  $P_2$  and  $P_3$  are small, and their difference can be omitted. Therefore, we have

$$\begin{aligned} \hat{P} &\approx P_1 = \sum_{n \in \mathcal{S}_p} \left| \hat{x}_n - Ae^{j\theta_n} \right|^2 \\ &= \sum_{n \in \mathcal{S}_p} \left| f_n - \beta \hat{f}_n + A(e^{j\theta_n} - e^{j\hat{\theta}_n}) \right|^2 \end{aligned} \quad (55)$$

where  $\hat{\theta}_n$  is the phase of  $\hat{x}_n$ .

Since  $\beta^{(\text{opt})}$  is not large, we can see that  $|x_n| = |Ae^{j\theta_n} + f_n| \gg |\beta \hat{f}_n|$ , i.e.,  $\beta \hat{f}_n$  could not significantly change the phase of  $x_n$ . Therefore,  $\hat{\theta}_n \approx \theta_n$ , and

$$\hat{P} \approx \sum_{n \in \mathcal{S}_p} |f_n - \beta \hat{f}_n|^2. \quad (56)$$

The optimal solution is

$$\beta^{(\text{opt})} = \frac{\Re \left[ \sum_{n \in \mathcal{S}_p} f_n \hat{f}_n^* \right]}{\sum_{n \in \mathcal{S}_p} |\hat{f}_n|^2} \quad (57)$$

where  $\Re[x]$  represents the real part of  $x$ , and  $(\cdot)^*$  represents the complex conjugate.

If  $\beta$  can be a complex number, the optimal solution is

$$\beta_c^{(\text{opt})} = \frac{\sum_{n \in \mathcal{S}_p} f_n \hat{f}_n^*}{\sum_{n \in \mathcal{S}_p} |\hat{f}_n|^2}.$$

However, note that  $\beta^{(\text{opt})}$  is mainly determined by some dominant peaks. Most likely, if  $f_n$  is a dominant peak,  $\hat{f}_n$  is also large, and the phase of  $\hat{f}_n$  is close to that of  $f_n$ . Therefore, the imaginary part of  $\beta_c^{(\text{opt})}$  is small and can be omitted.

To further reduce the complexity, some small samples of  $f_n$  can be excluded from (57), which, however, may degrade the PAR reduction performance.

Now, the adaptive-scaling tone-reservation algorithm can be summarized as follows, where the omitted parts are the same as those in Algorithm 1.

*Algorithm 2 (adaptive scaling)*

*Initialization:*

1) Set up  $A$ ,  $\mathcal{R}$ , and the maximum number of iterations  $L$ .

*Runtime:*

•••

- 4) Construct  $\mathcal{S}_p$  (the index set of the peaks of  $f_n$ ) and find  $c_n$  as follows.
  - a) Find the filtered clipping noise (in the frequency domain)  $\hat{\mathbf{F}}$  as in Algorithm 1.
  - b) Convert  $\hat{\mathbf{F}}$  to the time domain to obtain  $\hat{f}_n = \text{IDFT}(\hat{\mathbf{F}})$ .
  - c) Find  $\beta^{(\text{opt})}$  using (57).
  - d) The peak reduction signal is  $c_n = -\beta^{(\text{opt})}\hat{f}_n$ .
- 5) Calculate the PAR reduced OFDM signal as  $\hat{x}_n = x_n + c_n$ . If  $\text{PAR} > A$ , and the iteration number is less than  $L$ , go to step 3. Otherwise, transmit  $\hat{x}_n$  and terminate the algorithm.

**C. Complexity Comparison**

1) *Complexity Analysis of the Proposed Algorithms:* We now measure the complexity of our algorithms using the number of real multiplications. A complex multiplication is counted as three real multiplications [31]. Only the runtime complexity is considered, and the complexity of the initialization stage can be omitted since it occurs only once. In the runtime stage of both algorithms, steps 1 and 2 are not counted either because all OFDM systems must execute step 1, and all PAR reduction techniques require at least one iteration of these two steps.

In step 3,  $f_n$  can be calculated as  $f_n = x_n(1 - (A/|x_n|))$ , where  $n \in \mathcal{F} = \{n : |x_n| > A\}$ , and  $N_f$  is the size of  $\mathcal{F}$ . The complexity of this step is determined by the cost of calculating  $|x_n|$  and  $f_n$ .

By applying condition (49) to exclude small samples, the complexity of calculating  $|x_n|$  is small. The number of samples that satisfy (49) is

$$\bar{N}_c = JN \left( 1 - (\text{erf}(A/2\sigma))^2 - P1 \right)$$

where

$$P1 = \int_{A/\sqrt{2}}^A \frac{2\sqrt{2}}{\sigma\sqrt{\pi}} \text{erf}\left(\frac{A-x}{\sigma\sqrt{2}}\right) e^{-x^2/2\sigma^2} dx.$$

For example,  $\bar{N}_c \approx 0.0038JN$ ,  $0.057JN$ , and  $0.10JN$  when  $A/\sqrt{2}\sigma = 9, 6$ , and  $5$  dB, respectively.

The complexity of calculating  $f_n$  for  $n \in \mathcal{F}$  is  $2N_f$  real multiplications and  $N_f$  real divisions. However,  $N_f$  is a function of  $N$ . To see this, we calculate the mean of  $N_f$  as  $\bar{N}_f = \bar{N}_p \bar{r} f_s$ , where  $f_s = JN/T$  is the sampling frequency, and  $\bar{N}_p$  is the average number of pulses in an OFDM signal duration calculated in (14). Then, we have

$$\bar{N}_f = JN e^{-A^2/2\sigma^2}. \tag{58}$$

$N_f$  may change after the first iteration. However, note that the OFDM signal after the first iteration is  $\hat{x}_n = x_n + c_n$ . The  $N_f$  for  $\hat{x}_n$ , denoted as  $\hat{N}_f$ , is  $\hat{N}_f = N_f - N_1 + N_2$ , where  $N_1$  is the number of samples that are higher than  $A$  in the first iteration but are lower than  $A$  after the first iteration, and  $N_2$  is the number of samples of new peaks (i.e., samples that are lower than  $A$  in the first iteration but are higher than  $A$  after the first iteration).

Because  $c_n$  is very small (its average power is usually no larger than a fraction of a decibel), only those peaks of  $x_n$  that are slightly smaller or higher than  $A$  will contribute to  $N_1$  or  $N_2$ . Using (58), it is easy to check that  $N_1$  and  $N_2$  are small numbers, and their difference can be omitted. Thus,  $N_f$  is roughly constant in all iterations. We estimate the complexity of calculating  $f_n$  as  $2\bar{N}_f$  real multiplications and  $\bar{N}_f$  real divisions, i.e.,  $\mathcal{O}(N)$  with a small constant of proportionality. For example,  $\bar{N}_f = 3.5 \times 10^{-4}JN$ ,  $0.019JN$ , and  $0.042JN$  for  $A/\sqrt{2}\sigma = 9, 6$ , and  $5$  dB, respectively.

The complexity of our algorithms is mainly determined by the  $JN$ -point DFT/IDFT pair and weighting the clipping noise in step 4. The latter requires  $2JN$  real multiplications.

Using the decimation-in-time split-radix FFT algorithm [32], the complexity of a  $JN$ -point DFT with  $N_f$  nonzero inputs and  $N$  in-band outputs (other outputs are not needed) can be calculated as follows [33]:

$$\begin{aligned} \mathcal{M}_{JN} &= \mathcal{M}_{JN/2} + 2\mathcal{M}_{JN/4} \\ &\quad + \max(0, \min(6N_f, 3JN/2 - 8)) \end{aligned} \tag{59}$$

$$\mathcal{M}_J = 0 \tag{60}$$

$$\mathcal{M}_{2J} = \max(0, \min(3N_f, 3J/2 - 4)) \tag{61}$$

where  $\mathcal{M}_k$  denotes the number of real multiplications for calculating a  $k$ -point DFT. Thus, the average complexity  $\mathcal{M}_{\text{DFT}}$  of the DFT in step 4 can be calculated by using (59)–(61) replacing  $N_f$  by  $\bar{N}_f$ . Since the  $JN$ -point IDFT in step 4 has  $N_r$  nonzero inputs and  $JN$  outputs, its complexity  $\mathcal{M}_{\text{IDFT}}$  can be calculated using (59)–(61) replacing  $N_f$ ,  $N$ , and  $J$  by  $N_r$ ,  $JN$ , and  $1$ .

With the above discussion, the complexity of our algorithms for  $L$  iterations ( $L = 1$  for constant scaling) is

$$\mathcal{M} = L(2\bar{N}_f + 2\bar{N}_c + 4\bar{N}_p + \mathcal{M}_{\text{DFT}} + \mathcal{M}_{\text{IDFT}} + 2JN)$$

real multiplications and  $L(\bar{N}_f + 1)$  real divisions.

The number of iterations  $L$  that adaptive scaling requires to reach a fixed PAR (i.e., independent to  $N$ ) does not depend on  $J$  and  $N$ . Based on our analysis, the strength of clipping pulses and the distances between clipping pulses are independent of  $J$  and  $N$ . Also, note that the adaptive-scaling algorithm compensates *all* large peaks (higher than  $A$ ) by using the scaled clipping noise in each iteration. Thus, once other parameters (e.g.,  $A$ ,  $R$ , and the reserved tone set) are fixed the PAR reduction obtained in each iteration is also fixed, irrespective of  $J$  and  $N$ . Therefore, the complexity of adaptive scaling for achieving fixed PAR is  $\mathcal{O}(N \log_2 N)$ . Moreover, since the inputs/outputs of the DFT/IDFT used in our algorithms are

sparse, the complexity of the DFT/IDFT may be further reduced to  $\mathcal{O}(N)$  by using a proper wavelet transform [34].

2) *Complexity Analysis of the Active-Set Algorithm:* If the  $(2G)$ -agon approximation ( $G = 4, 8, \dots$ ) is used, the complexity of the active-set algorithm with  $L$  iterations is [33]

$$\mathcal{M}_{\text{Act.}} \approx \frac{1}{2}(L^2 + 5L - 2)GJN$$

real multiplications and  $2LG\bar{N}_f$  real divisions.<sup>6</sup>

Note that the active-set algorithm proceeds only one peak (outside the active set) in each iteration. To reach the fixed PAR  $A$ , all the peaks above  $A$  must be compensated by the peak-canceling signal. Since the number of such peaks is proportional to  $N$ , the required number of iterations is also proportional to  $N$ . That is, the complexity of the active-set algorithm for reaching a fixed PAR is  $\mathcal{O}(N^2)$ .<sup>7</sup>

### VI. SIMULATION RESULTS

In this section, the simulation results verify the estimation of  $\bar{\beta}$  that is used in our constant-scaling algorithm. We also compare our proposed algorithms with the active-set algorithm [1] by simulation. We use  $N = 512$ ,  $J = 4$ , and  $10^6$  uniformly distributed 64-quadratic-amplitude-modulation symbols as input to the OFDM system. The clipping threshold  $A$  is measured in decibels with respect to the average signal power before the PAR reduction.

#### A. Theoretical and Actual Values of $\bar{\beta}$

First, we compare the theoretical value of  $\bar{\beta}$  calculated using (39), which is denoted  $\bar{\beta}_{\text{Theo}}$ , and the actual value obtained by simulation, which is denoted  $\bar{\beta}_{\text{Simu}}$ . Here,  $\bar{\beta}_{\text{Simu}}$  is calculated as

$$\bar{\beta}_{\text{Simu}} = E \left\{ \arg \min_{\beta} \sum_{k \in \mathcal{R}} \left| \hat{F}_k^{(L)} - \beta \hat{F}_k \right|^2 \right\}$$

and  $\hat{F}_k$  and  $\hat{F}_k^{(L)}$  are the in-band clipping noise terms at the first and  $L$ th iterations, respectively.

Fig. 2 compares  $\bar{\beta}_{\text{Theo}}$  and  $\bar{\beta}_{\text{Simu}}$ , where the tone-reservation ratio  $R = 4.88\%$ ,  $9.96\%$ , and  $19.92\%$ , respectively, the number of clipping and filtering iterations  $L = 20$ , and the set of reserved tones  $\mathcal{R}$  is randomly selected at the initialization stage.  $\bar{\beta}_{\text{Theo}}$  matches  $\bar{\beta}_{\text{Simu}}$  when  $A \geq 9$  dB. The difference becomes larger when  $A < 9$  dB, but may still be acceptable for  $A \geq 6$  dB [the relative difference defined as  $(\bar{\beta}_{\text{Theo}} - \bar{\beta}_{\text{Simu}})/\bar{\beta}_{\text{Simu}}$  is less than 10%]. Therefore, we can use  $\bar{\beta}_{\text{Theo}}$  to evaluate the PAR reduction performance and use it in our constant-scaling algorithm.

The approximation error in  $\bar{\beta}_{\text{Theo}}$  mainly depends on the strength of the mainlobe (outside the pulse duration) and sidelobes of filtered pulses, which is determined by  $R$  and the

<sup>6</sup>We omit the cost of finding a suitable optimization direction, which requires to solve a set of  $l$  linear equations with  $l$  variable, where  $l$  is the iteration number.

<sup>7</sup>Similarly, the complexity of the controlled clipper algorithm [7] is also  $\mathcal{O}(N^2)$ .

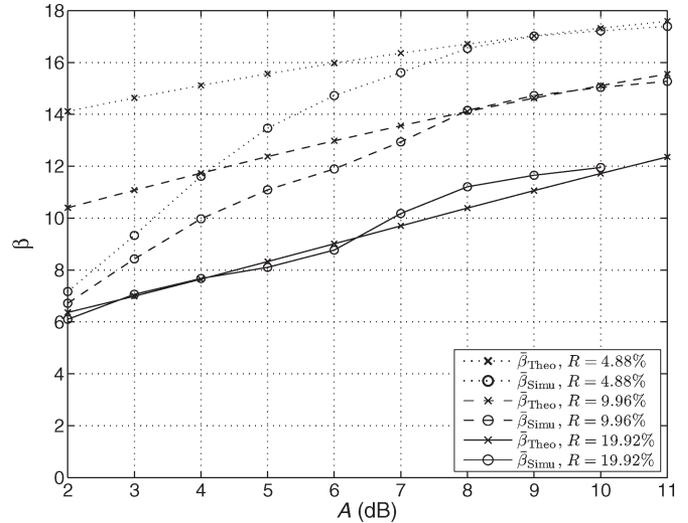


Fig. 2. Comparison of  $\bar{\beta}_{\text{Theo}}$  and  $\bar{\beta}_{\text{Simu}}$ .

TABLE I  
COMPLEXITY IN THE NUMBER OF REAL MULTIPLICATIONS/DIVISIONS OF CONSTANT-SCALING (CS), ADAPTIVE-SCALING (AS), AND ACTIVE-SET (ACT. SET) ALGORITHMS, WHERE  $N = 512$  AND  $J = 4$

	$R = 4.88\%$ , $A = 6.22$ dB	$R = 19.92\%$ , $A = 4.96$ dB
CS/AS, $L = 1$	22350.68 / 32.16	29273.32 / 89.93
AS, $L = 3$	67052.05 / 96.48	87819.96 / 269.78
AS, $L = 16$	357610.95 / 514.58	468373.12 / 1438.84
Act. Set, $L = 2$	49152.00 / 498.58	49152.00 / 1422.84
Act. Set, $L = 9$	507904.00 / 2243.62	507904.00 / 6402.80
2048-point FFT	16388 / 0	16388 / 0

selection of reserved tones, and the distances between clipping pulses, which are determined by  $A$ . When the tone-reservation ratio is large, e.g.,  $R = 19.92\%$  in Fig. 2, the mainlobe outside the pulse duration and the sidelobes of filtered pulses are small, and the approximation error is small. On the other hand, for small  $R$  (e.g.,  $R = 4.88\%$  or  $R = 9.96\%$  in Fig. 2), the distances between clipping pulses are large when  $A$  is large (e.g.,  $A \geq 6$  dB in Fig. 2), and the approximation error is small.

#### B. Performance of the Proposed Algorithms

We now compare the constant-scaling and adaptive-scaling algorithms with the active-set algorithm. Here, we use the modified PAR definition (6). The PAR complementary cdf (CCDF), defined as  $F(\xi_0) = \Pr[\xi > \xi_0]$ , is used to indicate the clip probability. Constant scaling, where the parameter  $K$  is used to calculate  $\bar{\beta}$ , requires only *one* iteration. The complexity of the algorithms we simulated is listed in Table I, where  $L$  is the number of iterations that are required for the adaptive-scaling and active-set algorithms. As a reference, the complexity of a  $JN$ -point FFT is also listed in this table.

Fig. 3 compares these algorithms for clipping threshold  $A = 6.22$  dB and a tone-reservation ratio of  $R = 4.88\%$ . The set of reserved tones  $\mathcal{R}$  is randomly selected. The PAR of original OFDM (no reserved tones) and that of OFDM with null reserved tones (reserved tones are set to 0) are also plotted.

Setting the reserved tones to 0 clearly does not reduce the PAR. For a  $10^{-4}$  clip probability, the constant-scaling algorithm

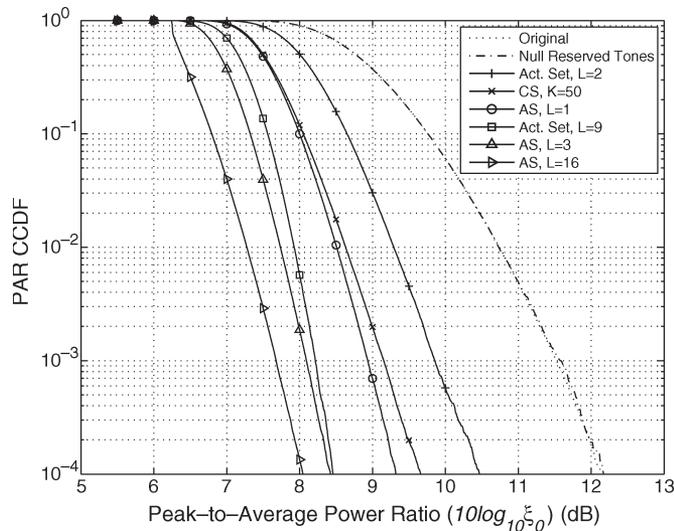


Fig. 3. PAR reduction comparison of constant-scaling (CS), adaptive-scaling (AS), and active-set (Act. Set) algorithms, where  $R = 4.88\%$ ,  $A = 6.22$  dB, and  $\mathcal{R}$  is randomly selected.

that approximates 50 clipping and filtering iterations obtains a 2.5-dB PAR reduction, which is about 0.8 dB larger than that of the active-set algorithm with two iterations but is 0.3 dB smaller than that of the adaptive-scaling algorithm with one iteration. Adaptive-scaling algorithm with three iterations obtains the same PAR reduction as an active-set algorithm with nine iterations for a  $10^{-4}$  clip probability but with only 13% complexity of the latter. With 16 iterations, the adaptive-scaling algorithm yields a 4.1-dB PAR reduction (0.4 dB larger than an active-set algorithm with nine iterations) for a  $10^{-4}$  clip probability. Note that its complexity is only 70% of that of the active-set algorithm with nine iterations.

Fig. 4 compares these algorithms for 4.96-dB clipping, 19.92% tone reservation, and a randomly selected set of reserved tones. Again, setting the reserved tones to 0 does not reduce the PAR. An adaptive-scaling algorithm with one iteration obtains a 4-dB PAR reduction, which is about the same as a constant-scaling algorithm that approximates 20 clipping and filtering iterations, 2.4 dB larger than an active-set algorithm with two iterations, and 0.6 dB larger than an active-set algorithm with nine iterations. With three and 16 iterations, adaptive-scaling algorithm obtains 5.3- and 6.2-dB reductions (1.8 and 2.7 dB larger than an active-set algorithm with nine iterations) with 17% and 90% complexities of an active-set algorithm with nine iterations, respectively.

Table II lists the average power increase. The larger the PAR reduction is, the larger the average power increases. However, the largest average power increase is only 0.44 dB. Therefore, the power increase would not significantly increase the BER. Note that the active-set algorithm has a negligible average power increase. However, since its PAR reduction is much smaller than that of the adaptive-scaling algorithm, its BER performance is worse than the latter.

Taking into account the PAR reduction, complexity, and power increase, constant scaling may be a good choice if low complexity is desired. If a large PAR reduction is desired,

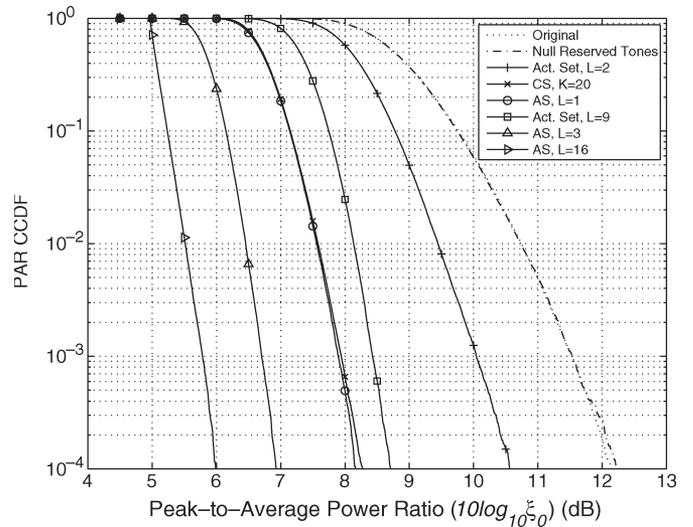


Fig. 4. PAR reduction comparison of constant-scaling (CS), adaptive-scaling (AS), and active-set (Act. Set) algorithms, where  $R = 19.92\%$ ,  $A = 4.96$  dB, and  $\mathcal{R}$  is randomly selected.

TABLE II  
AVERAGE POWER INCREASE (IN DECIBELS) OF CONSTANT-SCALING (CS), ADAPTIVE-SCALING (AS), AND ACTIVE-SET (ACT. SET) ALGORITHMS, WHERE THE RESERVED TONE SET IS RANDOMLY SELECTED

	$R = 4.88\%$ , $A = 6$ dB	$R = 19.92\%$ , $A = 4$ dB
CS, $K = 50$ for $R = 4.88\%$ , $K = 20$ for $R = 19.92\%$	0.14	0.13
AS, $L = 1$	0.11	0.13
AS, $L = 3$	0.23	0.23
AS, $L = 16$	0.44	0.39
Act. Set, $L = 2$	0.02	0.005
Act. Set, $L = 9$	0.11	0.03

adaptive-scaling algorithms with three and 16 iterations may be good choices for 4.88% and 19.92% reserved-tone cases.

Fig. 5 compares the BER performance of the adaptive-scaling and active-set algorithms, where 19.92% randomly selected tones are reserved for the PAR reduction, and the clipping threshold is  $A = 4.96$  dB. The OFDM signal is first processed using the adaptive-scaling algorithm with 16 iterations or the active-set algorithm with nine iterations, respectively. The peak reduced signal is passed through a solid-state power amplifier (SSPA) with a limited linear range and an additive white Gaussian noise channel. The input/output relationship of the SSPA can be written as [2]

$$y(t) = \frac{|x(t)|}{\left(1 + \left(\frac{|x(t)|}{C}\right)^{2p}\right)^{\frac{1}{2p}}} e^{j\phi(t)}$$

where  $x(t) = |x(t)|e^{j\phi(t)}$  is the input, and  $y(t)$  is the output of the SSPA. The SSPA approaches the conventional SL as  $p \rightarrow \infty$ . For the simulations, the SSPA parameters are  $p = 3$  and saturation point  $C = 5.46$  dB. In Fig. 5, the BERs of the OFDM signal without a PAR reduction over the same SSPA and over an ideal SSPA with an infinite linear range are also included for reference.

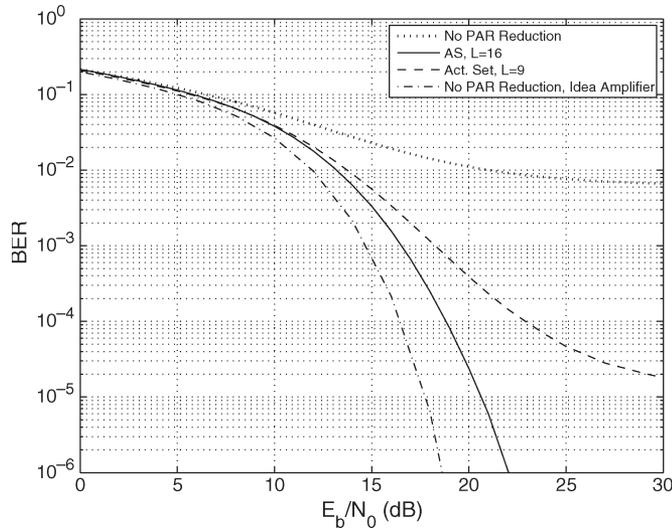


Fig. 5. BER comparison of adaptive-scaling (AS) and active-set (Act. Set) algorithms, where  $R = 19.92\%$ ,  $A = 4.96$  dB,  $\mathcal{R}$  is randomly selected, and  $C = 5.46$  dB.

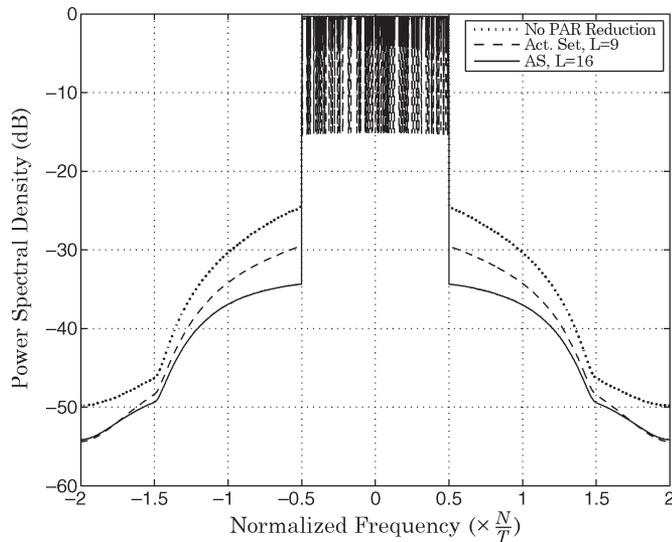


Fig. 6. Out-of-band radiation comparison of adaptive-scaling (AS) and active-set (Act. Set) algorithms, where  $R = 19.92\%$ ,  $A = 4.96$  dB,  $\mathcal{R}$  is randomly selected, and  $C = 5.46$  dB.

With the ideal SSPA, the OFDM system has a BER of  $10^{-6}$  when  $E_b/N_0 = 18.6$  dB. However, if the SSPA with  $C = 5.46$  dB is used, the BER plateaus at  $6 \times 10^{-3}$ . With the adaptive-scaling algorithm and 16 iterations, the BER is  $10^{-6}$  at  $E_b/N_0 = 22.0$  dB. On the other hand, the active-set algorithm with nine iterations results in a BER floor of  $2 \times 10^{-5}$ .

Fig. 6 compares the radiation out of the OFDM frequency band  $[-N/2T, N/2T]$  for the adaptive-scaling and active-set algorithms. The simulation parameters are the same as above. If no PAR reduction is used, the out-of-band radiation is  $-24.5$  dB. With the active-set algorithm (nine iterations), the out-of-band radiation is reduced to  $-29.5$  dB. However, by using the adaptive-scaling algorithm (16 iterations), the out-of-band radiation is further reduced to  $-34.5$  dB.

### VII. CONCLUSION

In this paper, using a parabolic approximation of clipping pulses, we analyzed the peak regrowth and the flat power spectrum of the in-band clipping noise of the OFDM signal subject to tone reservation with iterative clipping and filtering. We showed that the clipping noise obtained after several clipping and filtering iterations is approximately proportional to that generated in the first iteration; we also derived the constant of proportionality via the level-crossing theory. We have also proposed a constant-scaling algorithm and an adaptive-scaling algorithm for tone reservation. These algorithms scale the filtered clipping noise by a constant or an adaptively calculated factor to generate a peak-canceling signal. The simulation results show that the PAR and the complexity of the proposed algorithms are lower than those of the active-set algorithm.

### APPENDIX A

#### TIME INDEPENDENCE OF DIFFERENT CLIPPING PULSES

It can be shown that  $x(t)$  is a cyclostationary process. Then, the correlation of  $x(t)$  of time  $t_i$  and  $t_k$  only depends on the time difference  $\Delta t = t_i - t_k$ . The correlation coefficient of  $x(t)$  and  $x(t + \Delta t)$  is given by

$$\begin{aligned} \rho_x(\Delta t) &= \frac{\frac{1}{2}E\{x(t)x^*(t + \Delta t)\}}{\frac{1}{2}E\{x(t)x(t)^*\}} \\ &= \frac{\sin(\pi N\Delta t/T)}{N \sin(\pi\Delta t/T)} e^{-j\pi\Delta t/T}. \end{aligned} \quad (62)$$

Strictly speaking,  $\rho_x(n)$  is a  $\delta$  function, and the samples of  $x(n)$  are independent only at the Nyquist sampling rate  $T/N$ . However, we will show that, in the continuous-time domain, the possibility that two or more clipping pulses fall within a small time interval and have large correlation is small and can be omitted. Thus, clipping pulses that occurred at different times can be effectively treated as independent.

Without loss of generality, let  $\Delta t = nT/N$ , where  $n$  is a real number, and  $0 \leq n < N$ . When  $n$  is small<sup>8</sup> compared to  $N$ , then

$$|\rho_x(nT/N)| = \left| \frac{\sin(\pi n)}{N \sin(\pi n/N)} \right| = \left| \frac{\sin(\pi n)}{\pi n} \right| \quad (63)$$

when  $N \rightarrow \infty$ . For example,  $|\rho_x(nT/N)| \approx 0.071$  when  $n = 4.5$ .

The probability that two or more clipping pulses occur within a time interval of  $4.5T/N$  is small. It has been shown in [29] that the up-crossing time of  $x(t)$  is Poisson distributed when  $N \rightarrow \infty$ , i.e.,

$$\lim_{A \rightarrow \infty} \Pr[U_{n,A}(0, \Delta t) = k] = \frac{(\Delta t \lambda_A)^k e^{-\Delta t \lambda_A}}{k!} \quad (64)$$

<sup>8</sup>We do not need to consider large  $n$  since at large  $n$ ,  $\rho_x(n)$  is virtually 0.

TABLE III  
GIVEN THAT A CLIPPING PULSE OCCURS IN  $(0, 4.5T/N)$ , THE PROBABILITY THAT MORE THAN  $k$  ( $k \geq 1$ ) CLIPPING PULSES OCCUR IN THE SAME TIME INTERVAL  $(0, 4.5T/N)$

A (dB)	3	6	9
Pr(1)	0.59	0.16	$4.6 \times 10^{-3}$
Pr(2)	0.22	$1.3 \times 10^{-2}$	$1.1 \times 10^{-5}$
Pr(3)	$6.0 \times 10^{-2}$	$7.4 \times 10^{-4}$	$1.6 \times 10^{-8}$
Pr(4)	$1.3 \times 10^{-2}$	$3.1 \times 10^{-5}$	$1.9 \times 10^{-11}$
Pr(5)	$2.2 \times 10^{-3}$	$1.1 \times 10^{-6}$	$1.7 \times 10^{-14}$

where  $U_{n,A}(0, \Delta t)$  is the number of up-crossings that a  $\chi^2$  random process

$$Y(t) = X_1^2(t) + X_2^2(t) + \dots + X_n^2(t) \quad (65)$$

up-crosses a fixed level  $A/\sigma > 0$ , with  $\sigma = E\{X_k^2(t)\}$  for all  $k$ , during the time interval  $(0, \Delta t)$ , and  $\lambda_A$  is the up-crossing rate. Based on our assumption of large  $A$ , each up-crossing corresponds to a parabolic clipping pulse. Then, given the condition that a clipping pulse occurs in  $(0, \Delta t)$ , the probability that more than  $m$  ( $m \geq 1$ ) clipping pulses occur in the same time interval  $(0, \Delta t)$ , which is denoted as  $\text{Pr}(m)$ , is

$$\begin{aligned} \text{Pr}(m) &= \text{Pr}[U_{n,A}(0, \Delta t) > m | \text{A clipping pulse occurs}] \\ &= \text{Pr}[U_{n,A}(0, \Delta t) > m - 1] \\ &= 1 - \sum_{l=0}^{m-1} \frac{(\Delta t \lambda_A)^l e^{-\Delta t \lambda_A}}{l!}. \end{aligned} \quad (66)$$

$\text{Pr}(m)$  is independent to  $N$  when  $\Delta t = nT/N$ . Table III lists  $\text{Pr}(m)$  for a different clipping threshold  $A$ , where  $n = 4.5T/N$ . We observe that when  $A$  is large, the chance that two or more clipping pulses fall within the same time interval  $(0, 4.5T/N)$  is small. In other words, most clipping pulses are apart from each other with a large ‘‘distance,’’ and then,  $|\rho_x(\Delta t)|$  between these pulses are small and can be approximated as 0. Since  $x(t)$  is Gaussian,  $x(t)$  and  $x(t + \Delta t)$  are also independent with respect to  $\Delta t$ .

Table III also includes a low clipping threshold case, where  $A = 3$  dB. In this case,  $\text{Pr}(1)$  and  $\text{Pr}(2)$  are relatively large at  $n = 4.5$ . However, taking into account that  $|\rho_x(\Delta t)|$  is only 0.071 at  $n = 4.5$ , we may still treat the clipping pulses in this case as uncorrelated to simplify the power spectrum estimation.

APPENDIX B

CONDITIONAL PDF AND MOMENTS OF  $\gamma_i$

In this section, we find the conditional pdf and moments of  $\gamma_i = \dot{x}_I(t_i)/|x_R(t_i)|$ , given  $\dot{r}(t_i) = 0$  and  $r(t_i) \geq A$ . For ease of notation, we drop off the subscript  $i$  and the time index  $t_i$  in the following analysis. Strictly speaking, we also need a condition of  $\ddot{r} \leq 0$ . However, Appendix C shows that  $\text{Pr}[\ddot{r} | \dot{r} = 0, r \leq A] \approx 0$  unless  $A$  is small.

First, we show that  $x_R, x_I, \dot{x}_R$ , and  $\dot{x}_I$  are independent when  $N$  is large. Note that  $\dot{x}_R = (1/2)(\dot{x} + \dot{x}^*)$  and  $\dot{x}_I =$

$(1/2j)(\dot{x} - \dot{x}^*)$ , where

$$\dot{x}(t) = \frac{dx(t)}{dt} = \frac{1}{\sqrt{N}} \sum_{k=-N/2}^{N/2-1} \frac{j2\pi k}{T} X_k e^{j2\pi kt/T}. \quad (67)$$

Also, note that  $E\{X_k X_l\} = 0$  for any  $k$  and  $l$ . Then,  $\dot{x}_R$  and  $\dot{x}_I$  are i.i.d. Gaussian processes with zero mean and variance given by

$$\dot{\sigma}^2 = \frac{4\pi^2 \sigma^2}{NT^2} \sum_{k=-N/2}^{n/2-1} k^2 = \frac{(N^2 + 2)\pi^2 \sigma^2}{3T^2}.$$

When  $N$  is large,  $\dot{\sigma}^2 \approx (\pi^2 N^2 \sigma^2 / 3T^2) = (\pi^2 / 3) W^2 \sigma^2$ , which agrees with (12).

Using (1) and (67), we have

$$E\{x_R \dot{x}_R\} = E\{x_I \dot{x}_I\} = 0.$$

On the other hand

$$E\{x_R \dot{x}_I\} = -E\{x_I \dot{x}_R\} = \frac{2\pi \sigma^2}{NT} \sum_{k=-N/2}^{N/2-1} k = -\frac{\pi \sigma^2}{T}$$

and their correlation coefficients are

$$\rho_{x_R \dot{x}_I} = -\rho_{x_I \dot{x}_R} = -\frac{3}{\sqrt{N^2 + 2}}$$

which are zero when  $N \rightarrow \infty$  (less than 0.014 when  $N \geq 128$ ). Therefore,  $x_R, x_I, \dot{x}_R$ , and  $\dot{x}_I$  are independent when  $N$  is large.

We now find the joint pdf  $p(\dot{x}_I, x_R, x_I | \dot{r} = 0)$ . Since  $x_R, x_I, \dot{x}_R$ , and  $\dot{x}_I$  are independent, fixing  $x_R$  and  $x_I$  does not change the distribution of  $\dot{x}_R$  and  $\dot{x}_I$ . Note that  $\dot{r} = (1/r)(x_R \dot{x}_R + x_I \dot{x}_I)$ . Then, given  $x_R$  and  $x_I$ ,  $\dot{r}$  is also a Gaussian process with zero mean and variance  $\dot{\sigma}_r^2 = E\{\dot{r}^2\} = \dot{\sigma}^2$ . Since  $\dot{\sigma}_r^2$  is independent of  $x_R$  and  $x_I$ ,  $\dot{r}$  is independent of  $x_R, x_I$ , and  $r$ , and  $p(\dot{r}) = (1/(\sqrt{2\pi}\dot{\sigma}))e^{-\dot{r}^2/2\dot{\sigma}^2}$ . Given  $x_R$  and  $x_I$ , the correlation coefficient between  $\dot{x}_I$  and  $\dot{r}$  is  $\rho_{\dot{x}_I, \dot{r}} = x_I/r$ . Then

$$\begin{aligned} p(\dot{x}_I, \dot{r} | x_R, x_I) &= \frac{\sqrt{x_R^2 + x_I^2}}{2\pi \dot{\sigma}^2 |x_R|} \\ &\times \exp\left(-\frac{x_R^2 + x_I^2}{2\dot{\sigma}^2 x_R^2}\right) \\ &\times \left(\dot{x}_I^2 - 2\frac{x_I \dot{x}_I \dot{r}}{\sqrt{x_R^2 + x_I^2}} + \dot{r}^2\right) \end{aligned} \quad (68)$$

$$\begin{aligned} p(\dot{x}_I, x_R, x_I | \dot{r} = 0) &= \frac{p(\dot{x}_I, \dot{r} | x_R, x_I) p(x_R) p(x_I)}{p(\dot{r})} \Bigg|_{\dot{r}=0} \\ &= \frac{\sqrt{x_R^2 + x_I^2}}{(2\pi)^{3/2} \dot{\sigma} \sigma^2 |x_R|} \\ &\times \exp\left(-\frac{(x_R^2 + x_I^2) \dot{x}_I^2}{2\dot{\sigma}^2 x_R^2} - \frac{x_R^2 + x_I^2}{2\sigma^2}\right). \end{aligned} \quad (69)$$

We use the following transforms to obtain  $p(\gamma, r, \theta | \dot{r} = 0)$ :

$$\dot{x}_I = r\gamma | \cos \theta | + \dot{r} \sin \theta \tag{70}$$

$$x_R = r \cos \theta \tag{71}$$

$$x_I = r \sin \theta. \tag{72}$$

The Jacobian of the transformation from  $\dot{x}_I, x_R,$  and  $x_I$  to  $\gamma, r,$  and  $\theta$  is

$$\mathfrak{J} = r^2 | \cos \theta |. \tag{73}$$

Then, we have

$$p(\gamma, r, \theta | \dot{r} = 0) = p(\dot{x}_I, x_R, x_I | \dot{r} = 0) | \mathfrak{J} | \\ = \frac{r^2}{(2\pi)^{3/2} \dot{\sigma} \sigma^2} \exp \left( -\frac{r^2 \gamma^2}{2\dot{\sigma}^2} - \frac{r^2}{2\sigma^2} \right). \tag{74}$$

The pdf of  $\gamma$  conditioned on  $\dot{r} = 0$  and  $r \geq A$  is

$$p(\gamma | \dot{r} = 0, r \geq A) = \frac{\int_A^\infty \int_0^{2\pi} p(\gamma, r, \theta | \dot{r} = 0) d\theta dr}{\int_A^\infty p(r) dr} \\ = \frac{A^3}{4\sqrt{2\pi} \dot{\sigma} \sigma^2 \psi^3} \left( 2\psi e^{-\psi^2} + \sqrt{\pi} \operatorname{erfc}(\psi) \right) \\ \times e^{A^2/2\sigma^2} \tag{75}$$

where

$$\psi = \frac{A\sqrt{\gamma^2 \sigma^2 + \dot{\sigma}^2}}{\sqrt{2}\dot{\sigma}\sigma}.$$

The conditional cdf of  $\gamma$  cannot be written in closed form. However, since  $p(\gamma | \dot{r} = 0, r \geq A) = p(-\gamma | \dot{r} = 0, r \geq A)$ , the conditional mean  $m_\gamma = E\{\gamma | \dot{r} = 0, r \geq A\} = 0$ . The conditional variance  $\sigma_\gamma^2 = E\{\gamma^2 | \dot{r} = 0, r \geq A\}$  can be found by using (69).

We first transform  $p(\dot{x}_I, x_R, x_I | \dot{r} = 0)$  to  $p(\dot{x}_I, r, \theta | \dot{r} = 0)$ . The Jacobian of the transformation is  $\mathfrak{J} = r$ . Then

$$\sigma_\gamma^2 = \frac{\int_{-\infty}^\infty \int_A^\infty \int_0^{2\pi} \frac{\dot{x}_I^2}{r^2 \cos^2 \theta} p(\dot{x}_I, r, \theta | \dot{r} = 0) d\theta dr d\dot{x}_I}{\int_{-\infty}^\infty \int_A^\infty \int_0^{2\pi} p(\dot{x}_I, r, \theta | \dot{r} = 0) d\theta dr d\dot{x}_I} \\ = \frac{\pi^2 W^2}{6} E_1 \left( \frac{A^2}{2\sigma^2} \right) e^{A^2/2\sigma^2} \tag{76}$$

$$\sigma_\gamma^4 = \frac{\pi^4 W^4}{12} \left( E_1 \left( \frac{A^2}{2\sigma^2} \right) e^{A^2/2\sigma^2} - \frac{2\sigma^2}{A^2} \right). \tag{77}$$

We may also find the central moments of  $|\gamma|$  by using (69). The conditional mean of  $|\gamma|$  is

$$m_{|\gamma|} = E\{|\gamma| | \dot{r} = 0, r \geq A\} \\ = \frac{\int_{-\infty}^\infty \int_A^\infty \int_0^{2\pi} \frac{|\dot{x}_I|}{r |\cos \theta|} p(\dot{x}_I, r, \theta | \dot{r} = 0) d\theta dr d\dot{x}_I}{\int_{-\infty}^\infty \int_A^\infty \int_0^{2\pi} p(\dot{x}_I, r, \theta | \dot{r} = 0) d\theta dr d\dot{x}_I} \\ = \frac{\pi W}{\sqrt{3}} \operatorname{erfc} \left( \frac{A}{\sqrt{2}\sigma} \right) e^{A^2/2\sigma^2}. \tag{78}$$

The conditional variance of  $|\gamma|$  is  $\sigma_{|\gamma|}^2 = \sigma_\gamma^2 - m_{|\gamma|}^2$ .

TABLE IV  
 $m_{|\gamma|}$  AND  $\sigma_{|\gamma|}$  FOR DIFFERENT  $A$ . SIMULATED RESULTS ARE OBTAINED WITH  $N = 512, J = 128,$  AND QPSK INPUT SYMBOLS

$A/\sqrt{2}\sigma$ (dB)	3	6	9
$m_{ \gamma }$ ( $\times \pi W$ ) (Theoretical)	0.19	0.15	0.11
$m_{ \gamma }$ ( $\times \pi W$ ) (Simulated)	0.16	0.13	0.11
$\sigma_{ \gamma }$ ( $\times \pi W$ ) (Theoretical)	0.15	0.11	0.083
$\sigma_{ \gamma }$ ( $\times \pi W$ ) (Simulated)	0.12	0.10	0.078

When  $A \rightarrow \infty, m_{|\gamma|}$  and  $\sigma_{|\gamma|}$  are virtually zero. For practical  $A$ 's,  $m_{|\gamma|}$  and  $\sigma_{|\gamma|}$  are also small compared to the OFDM half bandwidth (in radians per second)  $\pi W$ . Table IV lists the theoretical and simulated values of  $m_{|\gamma|}$  and  $\sigma_{|\gamma|}^2$  for different  $A$ . Note that  $\gamma_i$  introduces a frequency shift to  $F_i(\omega)$ , which is the frequency spectrum of the clipping pulse  $f_i(t)$ . Also, note that  $F_i(\omega)$  reaches its maximum magnitude at  $\omega = 0$  if  $\gamma_i = 0$ . Then, we can measure the frequency at which  $F_i(\omega)$  reaches its maximum magnitude to obtain the simulated statistics of  $\gamma_i$ . In our simulation, we use  $N = 512, J = 128,$  and QPSK input symbols.<sup>9</sup>

APPENDIX C

$\Pr[\dot{r}(t_i) > 0 | \dot{r}(t_i) = 0, r(t_i) \geq A] \rightarrow 0$  WHEN  $A \rightarrow \infty$

In this section, we prove that  $\Pr[\dot{r}(t_i) > 0 | \dot{r}(t_i) = 0, r(t_i) \geq A] \rightarrow 0$  when  $A \rightarrow \infty$ . In the following, we drop off the time index  $t_i$  for ease of notation.

Since  $\ddot{x}_R = (1/2)(\ddot{x} + \ddot{x}^*)$  and  $\ddot{x}_I = (1/2j)(\ddot{x} - \ddot{x}^*)$ , where

$$\ddot{x}(t) = \frac{d\dot{x}(t)}{dt} = -\frac{1}{\sqrt{N}} \sum_{k=-N/2}^{N/2-1} \frac{4\pi^2 k^2}{T^2} X_k e^{j2\pi kt/T} \tag{79}$$

$\ddot{x}_R$  and  $\ddot{x}_I$  are i.i.d. Gaussian processes with zero mean and variance given by<sup>10</sup>

$$\ddot{\sigma}^2 = \frac{16\pi^4 \sigma^2}{NT^4} \sum_{k=-N/2}^{N/2-1} k^4 \approx \frac{\pi^4}{5} W^4 \sigma^2.$$

By using (1) and (79), we have

$$E\{x_R \ddot{x}_I\} = E\{x_I \ddot{x}_R\} = 0$$

$$E\{x_R \ddot{x}_R\} = E\{x_I \ddot{x}_I\} = -E\{\dot{x}_I^2\} = -\frac{\pi^2}{3} W^2 \sigma^2$$

$$\rho_{x_R \ddot{x}_R} = \rho_{x_I \ddot{x}_I} = -\frac{\sqrt{5}}{3}.$$

Thus, it is given that  $x_R$  and  $x_I, \ddot{x}_R$  and  $\ddot{x}_I$  are independent Gaussian processes with mean  $-(\pi^2/3)W^2 x_R$  (for  $\ddot{x}_R$ ) and  $-(\pi^2/3)W^2 x_I$  (for  $\ddot{x}_I$ ) and variances of  $(4/9)\ddot{\sigma}^2$ .

<sup>9</sup>We choose a large oversampling factor  $J$  to avoid the case that the clipping pulse has only one nonzero sample.

<sup>10</sup>We may also calculate  $\ddot{\sigma}^2$  as  $\ddot{\sigma}^2 = \int \omega^4 S(\omega) d\omega = (\pi^4/5)W^4 \sigma^2$  [22].

Note that  $\dot{x}_I = \gamma|x_R|$ . Given  $\dot{r} = (1/r)(x_R\dot{x}_R + x_I\dot{x}_I) = 0$ , we have  $\dot{x}_R = -x_I\dot{x}_I/x_R$ , and

$$\ddot{r} = \frac{1}{r} (r^2\gamma^2 + x_R\ddot{x}_R + x_I\ddot{x}_I). \quad (80)$$

Then, given  $x_R$ ,  $x_I$ ,  $\gamma$ , and  $\dot{r} = 0$ ,  $\ddot{r}$  is a Gaussian process with mean

$$\begin{aligned} \hat{m}_{\ddot{r}} &= E\{\ddot{r}|x_R, x_I, \gamma, \dot{r} = 0\} \\ &= \frac{1}{r} \left( r^2\gamma^2 - \frac{\pi^2}{3}W^2 (x_R^2 + x_I^2) \right) \\ &= r \left( \gamma^2 - \frac{\pi^2W^2}{3} \right) \end{aligned}$$

and variance

$$\sigma_{\ddot{r}}^2 = E\{\ddot{r}^2|x_R, x_I, \gamma, \dot{r} = 0\} - \hat{m}_{\ddot{r}}^2 = \frac{4\pi^4}{45}W^4\sigma^2.$$

Since  $\sigma_{\ddot{r}}^2$  is independent of  $x_R$ ,  $x_I$ , and  $\gamma$ , it is also the variance of  $\ddot{r}$  given  $\dot{r} = 0$  and  $r \geq A$ . Normalizing  $\ddot{r}$  by dividing it by  $\sigma$ , the variance of  $\ddot{r}/\sigma$  given  $\dot{r} = 0$  and  $r \geq A$  is  $\sigma_{\ddot{r}}^2/\sigma^2 = 4\pi^4W^4/45$ , which is a constant that is only related to the OFDM bandwidth  $W$ .

Note that  $\hat{m}_{\ddot{r}}$  is also the mean of  $\ddot{r}$  given  $r$ ,  $\gamma$ , and  $\dot{r} = 0$  since it depends on  $r$  instead of individual  $x_R$  and  $x_I$ . Then, we have

$$\begin{aligned} p(\ddot{r}|\gamma, r, \dot{r} = 0) &= p(\ddot{r}|\gamma, x_R, x_I, \dot{r} = 0) \\ &= \frac{1}{\sqrt{2\pi\sigma_{\ddot{r}}^2}} \exp\left(-\frac{(\ddot{r} - \hat{m}_{\ddot{r}})^2}{2\sigma_{\ddot{r}}^2}\right). \quad (81) \end{aligned}$$

The conditional mean of  $\ddot{r}/\sigma$  given  $\dot{r} = 0$  and  $r \geq A$  can be calculated as follows:

$$\begin{aligned} m_{\ddot{r}/\sigma} &= E\left\{\frac{\ddot{r}}{\sigma}|\dot{r} = 0, r \geq A\right\} \\ &= \frac{\int_{-\infty}^{\infty} \int_A^{\infty} \int_{-\infty}^{\infty} \int_0^{2\pi} \ddot{r}p(\ddot{r}, \gamma, r, \theta|\dot{r} = 0)d\theta d\gamma dr d\ddot{r}}{\sigma \int_A^{\infty} p(r)dr}. \end{aligned}$$

However, note that

$$\begin{aligned} &\int_{-\infty}^{\infty} \ddot{r}p(\ddot{r}, \gamma, r, \theta|\dot{r} = 0)d\ddot{r} \\ &= \int_{-\infty}^{\infty} \ddot{r}p(\ddot{r}|\gamma, r, \theta, \dot{r} = 0)p(\gamma, r, \theta|\dot{r} = 0)d\ddot{r} \\ &= \hat{m}_{\ddot{r}}p(\gamma, r, \theta|\dot{r} = 0) \end{aligned}$$

where  $p(\gamma, r, \theta|\dot{r} = 0)$  is calculated in (74). Therefore

$$\begin{aligned} m_{\ddot{r}/\sigma} &= \frac{\int_A^{\infty} \int_{-\infty}^{\infty} \int_0^{2\pi} \hat{m}_{\ddot{r}}p(\gamma, r, \theta|\dot{r} = 0)d\theta d\gamma dr}{\sigma \int_A^{\infty} p(r)dr} \\ &= -\frac{\pi^2W^2A}{3\sigma}. \end{aligned}$$

When  $A \rightarrow \infty$ ,  $m_{\ddot{r}/\sigma} \rightarrow -\infty$ . Since  $\ddot{r}/\sigma$  has a constant (conditional) variance for any  $A$ , we have  $\Pr[\ddot{r} > 0|\dot{r} = 0, r \geq A] \rightarrow 0$  when  $A \rightarrow \infty$ . Moreover, note that the OFDM bandwidth  $W$  is usually large (several megahertz). Then, unless  $A$  is very small,  $m_{\ddot{r}/\sigma} \ll 0$ , implying  $\Pr[\ddot{r} > 0|\dot{r} = 0, r \geq A] \approx 0$ .

#### APPENDIX D

GIVEN  $\dot{r} = 0$  AND  $r \geq A$ ,  $\gamma_i$  AND  $\tau_i$  ARE UNCORRELATED

We first find the conditional joint pdf  $p(\ddot{r}, \gamma, r|\dot{r} = 0)$ . By using (74) and (81), we have

$$\begin{aligned} p(\ddot{r}, \gamma, r|\dot{r} = 0) &= p(\ddot{r}|\gamma, r, \theta, \dot{r} = 0)p(\gamma, r, \theta|\dot{r} = 0)p(\theta) \\ &= \frac{r^2}{2\pi\sigma_{\ddot{r}}\sigma^2} \exp\left(-\frac{(\ddot{r} - m_{\ddot{r}})^2}{2\sigma_{\ddot{r}}^2} - \frac{r^2\gamma^2}{2\sigma^2} - \frac{r^2}{2\sigma^2}\right). \end{aligned}$$

Since  $\tau = \sqrt{-8(r-A)/\ddot{r}}$ , the conditional joint moments of  $\tau$  and  $\gamma$  can be found as

$$\begin{aligned} E\{\tau^m\gamma^n|\dot{r} = 0, r \geq A\} &= \frac{\int_A^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^0 \tau^m\gamma^n p(\ddot{r}, \gamma, r|\dot{r} = 0)d\ddot{r}d\gamma dr}{\int_A^{\infty} p(r)dr} \end{aligned}$$

where  $m$  and  $n$  are positive integers. However, note that  $p(\ddot{r}, \gamma, r|\dot{r} = 0)$  is symmetric to  $\gamma$ . Therefore

$$\int_{-\infty}^{\infty} \tau^m\gamma^n p(\ddot{r}, \gamma, r|\dot{r} = 0)d\gamma = 0$$

for any odd  $n$ , and, in turn,  $E\{\tau^m\gamma^n|\dot{r} = 0, r \geq A\} = 0$  for any odd  $n$ . Specifically,  $E\{\tau\gamma|\dot{r} = 0, r \geq A\} = 0$ . That is, given  $\dot{r} = 0$  and  $r \geq A$ ,  $\tau$  and  $\gamma$  are uncorrelated.

#### APPENDIX E

CONDITIONAL PDF AND MOMENTS OF  $\eta_k$

In this section, we find the conditional mean and variance of  $\eta_k = (\dot{x}_I(t_k)r(t_k) - x_I(t_k)\dot{r}(t_k))/(r(t_k)|x_R(t_k)|)$ , given  $\dot{r}(t_k) \geq 0$  and  $r(t_k) = A$ . In the following, we drop off the time index  $t_k$  for ease of notation.

Note that

$$p(x_R, x_I, \dot{x}_I, \dot{r}) = p(\dot{x}_I, \dot{r}|x_R, x_I)p(x_R)p(x_I)$$

where  $p(\dot{x}_I, \dot{r}|x_R, x_I)$  is defined in (68). We have

$$\begin{aligned} p(\dot{x}_I, \dot{r}|\theta) &= \frac{rp(x_R = r \cos \theta, x_I = r \cos \theta, \dot{x}_I, \dot{r})}{p(r)} \\ &= \frac{1}{4\pi^2\sigma^2|\cos \theta|} \exp\left(-\frac{\dot{x}_I^2 - 2\dot{x}_I\dot{r} \sin \theta + \dot{r}^2}{2\sigma^2 \cos^2 \theta}\right). \end{aligned}$$

Since  $\dot{x}_I = r\eta_k|\cos \theta + \dot{r} \sin \theta$ , the Jacobian of transforming  $p(\dot{x}_I, \dot{r}|\theta)$  to  $p(\eta_k, \dot{r}|\theta)$  is  $r|\cos \theta|$ . Then, the pdf of  $\eta_k$

conditioned on  $\dot{r} \geq 0$  and  $r = A$  is given by

$$\begin{aligned} p(\eta_k | \dot{r} \geq 0, r = A) &= \frac{\int_0^\infty \int_0^{2\pi} r |\cos \theta| p(\dot{x}_I = r\eta_k |\cos \theta| + \dot{r} \sin \theta, \dot{r}, \theta | r) d\theta d\dot{r}}{\int_0^\infty p(\dot{r}) d\dot{r}} \\ &= \frac{\sqrt{6}A}{2\pi\sqrt{\pi}W\sigma} \exp\left(-\frac{3A^2\eta_k^2}{2\pi^2W^2\sigma^2}\right). \end{aligned}$$

The conditional moments of  $\eta_k$  can be found, i.e.,

$$\begin{aligned} m_{\eta_k} &= E\{\eta_k | \dot{r} \geq 0, r = A\} = 0 \\ \sigma_{\eta_k}^2 &= E\{\eta_k^2 | \dot{r} \geq 0, r = A\} = \frac{\pi^2 W^2 \sigma^2}{3A^2} \\ \sigma_{\eta_k}^4 &= E\{\eta_k^4 | \dot{r} \geq 0, r = A\} = \frac{\pi^4 W^4 \sigma^4}{3A^4} \\ m_{|\eta_k|} &= E\{|\eta_k| | \dot{r} \geq 0, r = A\} = \sqrt{\frac{2\pi}{3}} \frac{W\sigma}{A} \end{aligned}$$

and the conditional variance of  $|\eta_k|$  is

$$\sigma_{|\eta_k|}^2 = \sigma_{\eta_k}^2 - m_{|\eta_k|}^2 = \left(\frac{\pi^2}{3} - \frac{2\pi}{3}\right) \frac{W^2 \sigma^2}{A^2}. \quad (82)$$

When  $A$  is large,  $\eta_k$  has approximately the same distribution as  $\gamma_i$  (calculated in Appendix B). In (75), we see that  $\psi \gg 1$  when  $A$  is large. Then

$$\sqrt{\pi} \operatorname{erfc} \psi \approx \frac{e^{-\psi^2}}{\psi} \ll 2\psi e^{-\psi^2}.$$

Thus, (75) can be approximated as

$$\begin{aligned} p(\gamma_i | \dot{r} = 0, r \geq A) &\approx \frac{A^3 \psi e^{-\psi^2} e^{A^2/2\sigma^2}}{2\sqrt{2\pi} \dot{\sigma} \sigma^2 \psi^3} \\ &\approx \frac{\sqrt{6}A}{2\pi\sqrt{\pi}W\sigma} \exp\left(-\frac{3A^2\gamma_i^2}{2\pi^2W^2\sigma^2}\right) \end{aligned}$$

which is the same as  $p(\eta_k | \dot{r} \geq 0, r = A)$ . This approximation is valid when  $3\gamma_i^2 \ll \pi^2 W^2$ . On the other hand, when  $3\gamma_i^2$  is comparable to or larger than  $\pi^2 W^2$ , both  $p(\gamma_i | \dot{r} = 0, r \geq A)$  and this approximation are virtually zero. Therefore, we may use this approximation for all  $\gamma_i$ .

*Remark 7:* By using the asymptotic expansion of  $\operatorname{erfc} x$  and  $E_i x$  [35], it is easy to verify that the moments of  $\eta_k$  are approximately the same as those of  $\gamma_i$  when  $A$  is large. Intuitively, we can explain this fact by observing that  $\eta_k$  is measured at the time that  $r(t)$  up-crosses level  $A$ , and  $\gamma_i$  is measured at the time that  $r(t)$  reaches its local peak after the up-crossing. Since these two time instances are very close when  $A$  is large, the statistics of  $\eta_k$  and  $\gamma_i$  are the same.

## REFERENCES

- [1] B. Krongold and D. Jones, "An active-set approach for OFDM PAR reduction via tone reservation," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 495–509, Feb. 2004.
- [2] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Boston, MA: Artech House, Mar. 2000.
- [3] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of OFDM," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 131–133, May 1998.
- [4] A. Gatherer and M. Polley, "Controlling clipping probability in DMT transmission," in *Proc. 31st Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 2–5, 1997, vol. 1, pp. 578–584.
- [5] L. Wang and C. Tellambura, "A simplified clipping and filtering technique for PAR reduction in OFDM systems," *IEEE Signal Process. Lett.*, vol. 12, no. 6, pp. 453–456, Jun. 2005.
- [6] S. Leung, S. Ju, and G. Bi, "Algorithm for repeated clipping and filtering in peak-to-average power reduction for OFDM," *Electron. Lett.*, vol. 38, no. 25, pp. 1726–1727, Dec. 2002.
- [7] J. Tellado, "Peak to average power reduction for multicarrier modulation," Ph.D. dissertation, Stanford Univ., Stanford, CA, Sep. 1999.
- [8] W. Henkel and V. Zmo, "PAR reduction revisited: An extension to Tellado's method," in *Proc. 6th Int. OFDM-Workshop*, Hamburg, Germany, Sep. 18–19, 2001, pp. 31–31-6.
- [9] S. H. Müller and J. B. Huber, "A novel peak power reduction scheme for OFDM," in *Proc. 8th IEEE Int. Symp. PIMRC*, Helsinki, Finland, Sep. 1–4, 1997, vol. 3, pp. 1090–1094.
- [10] C. Tellambura and A. Jayalath, "PAR reduction of an OFDM signal using partial transmit sequences," in *Proc. 54th IEEE VTC-Fall*, Atlantic City, NJ, 2001, vol. 1, pp. 465–469.
- [11] C. Tellambura, "Improved phase factor computation for the PAR reduction of an OFDM signal using PTS," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 135–137, Apr. 2001.
- [12] J. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed–Muller codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2397–2417, Nov. 1999.
- [13] K. Paterson, "Generalized Reed–Muller codes and power control in OFDM modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 104–120, Jan. 2000.
- [14] H. Ochiai and H. Imai, "Block coding scheme based on complementary sequences for multicarrier signals," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci. (Japan)*, vol. E80-A, no. 11, pp. 2136–2143, Nov. 1997.
- [15] H. Ochiai and H. Imai, "Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2270–2277, Nov. 2000.
- [16] H. Ochiai and H. Imai, "Performance analysis of deliberately clipped OFDM signals," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 89–101, Jan. 2002.
- [17] H. Ochiai, "Performance analysis of peak power and band-limited OFDM system with linear scaling," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1055–1065, Sep. 2003.
- [18] H. Saeedi, M. Sharif, and F. Marvasti, "Clipping noise cancellation in OFDM systems using oversampled signal reconstruction," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 73–75, Feb. 2002.
- [19] D. Kim and G. L. Stüber, "Clipping noise mitigation for OFDM by decision-aided reconstruction," *IEEE Commun. Lett.*, vol. 3, no. 1, pp. 4–6, Jan. 1999.
- [20] J. Tellado, L. M. C. Hoo, and J. M. Cioffi, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, Feb. 2003.
- [21] H. Chen and A. Haimovich, "Iterative estimation and cancellation of clipping noise for OFDM signals," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 305–307, Jul. 2003.
- [22] N. M. Blachman, "Gaussian noise—Part I: The shape of large excursions," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1396–1400, Nov. 1988.
- [23] M. Pätzold, *Mobile Fading Channels*. Hoboken, NJ: Wiley, 2002.
- [24] A. Bahai, M. Singh, A. Goldsmith, and B. Saltzberg, "A new approach for evaluating clipping distortion in multicarrier systems," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1037–1046, Jun. 2002.
- [25] C. Tellambura, "Use of m-sequence for OFDM peak-to-average power ratio reduction," *Electron. Lett.*, vol. 33, no. 15, pp. 1300–1301, Jul. 1997.
- [26] C. Tellambura, "Phase optimisation criterion for reducing peak-to-average power ratio in OFDM," *Electron. Lett.*, vol. 34, no. 2, pp. 169–170, Jan. 1998.
- [27] H. Rowe, "Memoryless nonlinearities with Gaussian inputs: Elementary results," *Bell Syst. Tech. J.*, vol. 61, no. 7, pp. 1519–1525, Sep. 1982.
- [28] W. Henkel, V. Azis, and S. Trautmann, "The analytic treatment of the error probability due to clipping—A solved problem?" in *Proc. ISITA*, Parma, Italy, Oct. 10–13, 2004.
- [29] K. Sharpe, "Some properties of the crossings process generated by a stationary  $\chi^2$  process," *Adv. Appl. Probab.*, vol. 10, no. 2, pp. 373–391, Jun. 1978.

- [30] M. Aronowich and R. J. Adler, "Extrema and level crossings of  $\chi^2$  processes," *Adv. Appl. Probab.*, vol. 18, no. 4, pp. 901–920, Dec. 1986.
- [31] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [32] H. Sorensen, M. Heideman, and C. Burrus, "On computing the split-radix FFT," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 1, pp. 152–156, Feb. 1986.
- [33] L. Wang, "Peak-to-average power ratio reduction in OFDM systems," Ph.D. dissertation, Univ. Alberta, Edmonton, AB, Canada, 2007.
- [34] H. Guo and C. S. Burrus, "Wavelet transform based fast approximate Fourier transform," in *Proc. IEEE ICASSP*, Munich, Germany, Apr. 21–24, 1997, vol. 3, pp. 1973–1976.
- [35] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1964.



**Luqing Wang** (S'03) received the B.Sc. degree in electronics and information systems from Jilin University, Changchun, China, in 1992 and the M.Sc. degree in electrical and computer engineering in 2002 from the University of Alberta, Edmonton, Canada, where he is currently working toward the Ph.D. degree.

From 1992 to 2000, he was an Electrical and Computer Engineer and a System Analyst.



**Chintha Tellambura** (M'97–SM'02) received the B.Sc. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.Sc. degree in electronics from the University of London, London, U.K., in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993.

From 1993 to 1994, he was with the University of Victoria and, from 1995 to 1996, with the University of Bradford, Bradford, U.K., as a Postdoctoral Research Fellow. From 1997 to 2002, he was with Monash University, Melbourne, Australia. He is currently with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada, as a Professor. His research interests include coding, communication theory, modulation, equalization, and wireless communications.

Dr. Tellambura is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the Area Editor of Wireless Communications Theory and Systems for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was the Chair of the 2005 Communication Theory Symposium in Global Telecommunications Conference, held in St. Louis, MO.