



**HAL**  
open science

# PAPR Analysis as a Ratio of Two Random Variables: Application to Multicarrier Systems with Low Subcarriers Number

Vincent Savaux, Yves Louët

► **To cite this version:**

Vincent Savaux, Yves Louët. PAPR Analysis as a Ratio of Two Random Variables: Application to Multicarrier Systems with Low Subcarriers Number. *IEEE Transactions on Communications*, 2018, 66 (11), pp.5732-5739. 10.1109/TCOMM.2018.2857460 . hal-01847274

**HAL Id: hal-01847274**

**<https://hal.science/hal-01847274>**

Submitted on 7 Aug 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# PAPR Analysis as a Ratio of Two Random Variables: Application to Multicarrier Systems with Low Subcarriers Number

Vincent Savaux and Yves Louët

**Abstract**—The distribution of the peak to average power ratio (PAPR) of multicarrier systems is studied in the context of signals featuring low subcarriers number. Considering the mean power of the signal as a random variable, a general cumulative distribution function of the PAPR is suggested. We show that the new formulation is valid not only for low subcarriers number, but also for larger subcarriers number. Thus, an asymptotic analysis using a large subcarriers number proves a perfect concordance between the suggested approach and the state-of-the-art derivations only valid for large subcarriers number. All the theoretical developments are supported through simulations, and we discuss the case of oversampled signals as well.

**Index Terms**—PAPR, CCDF, Multicarrier, OFDM.

## I. INTRODUCTION

Multicarrier signals are well known to be prone to high power fluctuations due to the inherent summation of independent information carried on different tones. The most common way to quantify these power fluctuations is to define a metric taking into account its maximum value relatively to its mean. This gave birth to the peak factor definition namely the peak to average power ratio (PAPR) parameter highly discussed for decades especially in the context of multicarrier systems, such as orthogonal frequency division multiplexing (OFDM) (see [1] and references therein). PAPR is defined as the ratio of the maximum power and the mean power of a signal. An accurate PAPR derivation of multicarrier signal may be difficult given the initial statistic of the signal to be considered. A straightforward way to circumvent this issue is to directly upper bound the PAPR value regardless of any statistical hypothesis. As a result, a PAPR upper bound is found to be equal to  $N \cdot f(M)$ ,  $N$  being the useful subcarriers number of the OFDM systems and  $f(M)$  a real value depending on the digital modulation constellation size  $M$  (it is easily shown that  $f(M)$  equals  $3 \frac{M-1}{M+1}$ ). However this upper bound appears to be very far from reality and almost never reached due to the random behavior of the signal. As a consequence, the only way to provide a thorough PAPR model is to derive its distribution function or similarly its complementary cumulative distribution function (CCDF) viewing PAPR as a random variable due to the random character of the signal itself.

Manuscript submitted in March 2018

Vincent Savaux is with IRT b<>com, Rennes, France (e-mails: [vincent.savaux@b-com.com](mailto:vincent.savaux@b-com.com), Phone: +33 256 35 82 16.)

Yves Louët is with CentraleSupélec and IETR, Rennes, FR (e-mail: [yves.louet@centralesupelec.fr](mailto:yves.louet@centralesupelec.fr))

Considering a large subcarriers number  $N$ , a conventional analysis assumes that each time sample of the OFDM signal follows a complex Gaussian distribution. The overall signal can then be modeled as a random vector of  $N$  independent Gaussian samples. This has been first considered in [2], [3] where the CCDF of the PAPR is given in a very simple form and remains one of the most popular derivation. But the latter derivation ((2) in the paper) remains only valid for an oversampling factor equal to 1, which makes the formula inaccurate for oversampled signals simulating continuous phenomena. The reason is that oversampling leads to correlated samples. Still in [2], [3], the authors extended their work to provide an approximation of the PAPR CCDF valid for large oversampling factors, but this approximation is empirical and yields some discrepancies with simulations for low subcarriers number ( $N < 128$ ).

To extend [2], authors suggested more thorough derivations valid for lower oversampling values (4 or 8) [4]–[7] and different subcarriers numbers. These works are based either on the level crossing rate analysis of the peak distribution (since the envelop of the OFDM signal can be approximated as a Gaussian process), or theoretical links between peaks of sampled and continuous signals, called extreme value theory. However, these suggested expressions of PAPR CCDF are close to the simulation results only for  $N \geq 64$ .

Nevertheless the recent interest in OFDM systems with low subcarriers number (e.g. narrowband internet of things (NB-IoT) signal with 12 subcarriers [8]) should push us to revisit the aforementioned PAPR derivations. The reason lies in the PAPR denominator expression: it could not be approximated as the signal power expectation as previously, since the ergodic condition is not valid anymore. Even though the central limit theorem remains valid (down to 10 carriers [9]), the denominator itself has to be reconsidered as a random variable and no more as a constant value (the mean power).

Thus, the main contributions of this paper are summarized as follows.

- 1) We derive a general PAPR cumulative function valid for large and low subcarriers numbers, down to 12 subcarriers. Our results are supported by original theoretical developments. The main point is that the denominator of PAPR CCDF is viewed as a random variable.
- 2) We show that an asymptotic analysis for large  $N$  leads to existing results [2], [3], which confirms the validity of our approach. Moreover, the first and second order moments of PAPR are derived from our suggested

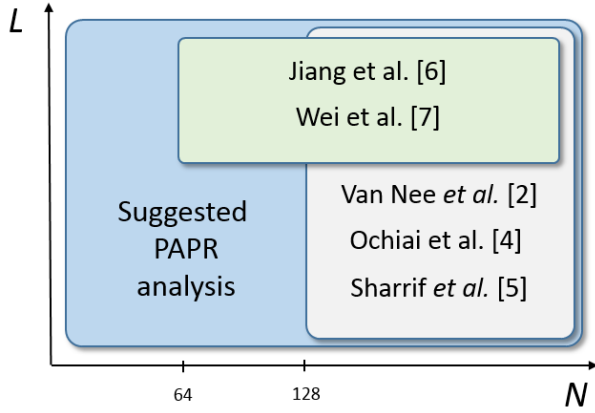


Fig. 1. Overview of the suggested PAPR analysis compared with state-of-the-art, according to subcarriers  $N$  and oversampling factor  $L$ .

CCDF, and perfectly match those found in the state of the art [4].

In addition, simulations results validate the theoretical developments, which shows the interest to reconsider PAPR expression for low subcarriers number. Furthermore, we discuss the extension of PAPR CCDF for low  $N$  value and large oversampling rate. Fig.1 positions the suggested work according to the number of subcarriers  $N$  and the oversampling factor  $L$ . Such as highlighted in Fig. 1, the suggested CCDF is relevant for low  $N$  value, as it was a pending issue, but it remains valid for larger  $N$  value as well.

The remainder of the paper is organized as follows. Section II presents the problem statement. Section III provides the new PAPR CCDF, first validated in Section IV to match the state-of-the-art results in an asymptotic analysis, and second in Section V through simulation results, which cover both signals at Nyquist rate and oversampled signals. Several concluding remarks are made in Section VI. Main proofs are given in Appendix to maintain the flow of the paper.

## II. PROBLEM STATEMENT

Let us consider a signal  $x(t)$  sampled at Nyquist rate  $1/t_s$  over an observation window of duration  $Nt_s$ , with  $t_s$  the sampling time. It is assumed that the samples  $x_n, n = 0, \dots, N-1$  are independent and identically distributed (iid) and obey a complex white Gaussian distribution  $\mathcal{CN}(0, \sigma^2)$ , where  $\sigma^2$  is the variance of  $x_n$ , i.e.  $\sigma^2 = \mathbb{E}\{|x_n|^2\}$ . This hypothesis fits the multicarrier signal such as OFDM or filter bank multicarrier (FBMC) [10], according to the central limit theorem for  $N > 10$  [9]. Note that in that case,  $N$  is usually equal to the number of samples per symbol, which corresponds to the number of subcarriers of the signal at Nyquist rate. According to these assumptions, the PAPR of the signal can be expressed as

$$\lambda = \frac{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}{\frac{1}{N} \sum_{n=0}^{N-1} |x_n|^2}. \quad (1)$$

It is worth mentioning that, in the literature, the mean square in the denominator expression is almost systematically replaced

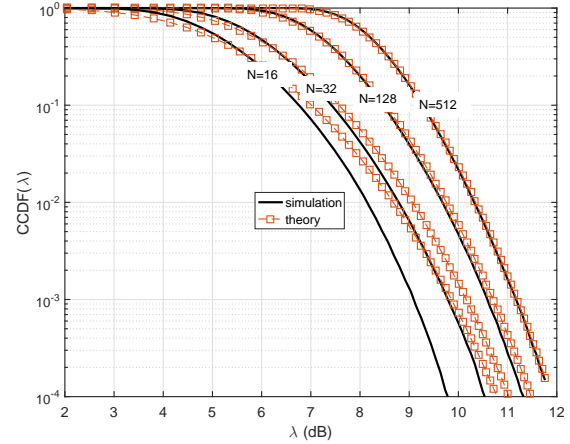


Fig. 2. CCDF of PAPR versus  $\lambda$  for  $N \in \{16, 32, 128, 512\}$ , comparison between theory in (2) and simulations.

by the expectation as  $\frac{1}{N} \sum_{n=0}^{N-1} |x_n|^2 \approx \mathbb{E}\{|x_n|^2\} = \sigma^2$ , assuming large  $N$ . This approximation allows for a simplification of the analytical developments, as the CCDF of the PAPR is derived from the numerator of (1). Since the elements  $|x_n|^2$  are iid and obey a Chi-squared ( $\chi^2$ ) distribution with two degrees of freedom, the CCDF is written as:

$$\begin{aligned} CCDF(\lambda) &= 1 - \prod_{n=0}^{N-1} \mathbb{P}\left(\frac{1}{\sigma^2} |x_n|^2 \leq \lambda\right) \\ &= 1 - \left(1 - e^{-\lambda}\right)^N, \end{aligned} \quad (2)$$

where  $\mathbb{P}(\cdot)$  means "probability of an event". However (2) holds when  $N$  is large enough. In fact, as shown in Fig. 2, the CCDF in (2) holds for  $N \geq 128$ . Therefore, a new expression of the CCDF should be derived for lower  $N$  values, i.e. when the mean  $\frac{1}{N} \sum_{n=0}^{N-1} |x_n|^2$  cannot be approximated by the expectation. We hereby propose a more general expression of the PAPR fitting low  $N$  values. Then, we present an asymptotic analysis of the PAPR when  $N$  tends to infinity, which shows that the known results of the literature can be obtained from the suggested general expression.

## III. SUGGESTED CCDF

For clarity purpose, we denote by  $n^*$  the index corresponding to the maximum value of  $|x_n|^2$ , i.e.

$$n^* = \arg \max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2. \quad (3)$$

Accordingly, we define the set of indexes  $\Xi = \{0, 1, \dots, N-1\}$ , and the subset  $\Xi^*$  which does not contain  $n^*$ , i.e.  $\Xi^* \cup \{n^*\} = \Xi$  and  $\Xi^* \cap \{n^*\} = \emptyset$ .

When  $N$  is low, the numerator and denominator of the PAPR in (1) cannot be considered as independent variables, since the maximum is included in the sum. Therefore we should rewrite the denominator of PAPR by using  $\Xi^*$  subset as

$$\frac{1}{N} \sum_{n=0}^{N-1} |x_n|^2 = \frac{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}{N} + \frac{1}{N} \sum_{n \in \Xi^*} |x_n|^2, \quad (4)$$

where the sum in the right side of (4) does not contain the maximum value. Thus, the general expression of the CCDF is obtained by substituting (4) into (1):

$$\begin{aligned} CCDF(\lambda) &= \mathbb{P} \left( \frac{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}{\frac{1}{N} \sum_{n=0}^{N-1} |x_n|^2} \geq \lambda \right) \\ &= \mathbb{P} \left( \frac{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}{\frac{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}{N} + \frac{1}{N} \sum_{n \in \Xi^*} |x_n|^2} \geq \lambda \right) \\ &= \mathbb{P} \left( \underbrace{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}_X \geq \underbrace{\frac{\lambda}{1 - \frac{\lambda}{N}}}_{\tilde{\lambda}} \underbrace{\frac{1}{N} \sum_{n \in \Xi^*} |x_n|^2}_Y \right), \end{aligned} \quad (5)$$

where  $\tilde{\lambda}$ ,  $X$ , and  $Y$  have been defined for clarity purpose. Note that (5) holds only if  $\lambda < N$ , which is a reasonable condition since we assume that  $N$ , although low-valued, is large enough to make the central limit theorem valid, in case of multicarrier signal. In order to definitely address this issue in the rest of the developments, we arbitrarily assume that  $N \geq 12$ , which is consistent with the considered CCDF range (see Fig. 2), as well as with the condition  $\lambda < N$  imposed by  $\tilde{\lambda}$  in (5).

Let us now investigate the distributions of  $X$  and  $Y$ , denoted by  $f_X$  and  $f_Y$ , respectively. From (2), it can be deduced that the distribution of  $X$  is the derivative of the CDF

$$\prod_{n=0}^{N-1} \mathbb{P}(|X|^2 \leq x) = \left(1 - e^{-\frac{x}{\sigma^2}}\right)^N, \quad (6)$$

where it can be noted that, unlike (2),  $|X|^2$  is not normalized by  $\sigma^2$ , since the denominator is not considered as constant anymore. Thus, we obtain the corresponding distribution:

$$f_X(x) = \frac{N}{\sigma^2} (1 - e^{-\frac{x}{\sigma^2}})^{N-1} e^{-\frac{x}{\sigma^2}}. \quad (7)$$

The derivation of the distribution of  $Y$  requires more developments. On one hand, it can be noted that  $Y$  is the sum of  $N - 1$  independent elements  $|x_n|^2$ , which are not normalized by  $\sigma^2$ . On the other hand, each element in the sum can be rewritten as  $|x_n|^2 = |Re(x_n)|^2 + |Im(x_n)|^2$ , where  $Re$  and  $Im$  stand for the real and the imaginary parts, respectively. Since  $Re(x_n)$  (resp.  $Im(x_n)$ ) is a zero-mean real Gaussian variable, then we can deduce that  $f_Y$  is similar to a  $\chi^2$  distribution with  $2N - 2$  degrees of freedom. In Appendix A, it is proved that the distribution of  $Y$  is expressed as

$$f_Y(y) = \frac{(N-1)^{N-1}}{\sigma_{y,N}^{2(N-1)} \Gamma(N-1)} y^{N-2} e^{-\frac{(N-1)y}{\sigma_{y,N}^2}}, \quad (8)$$

where  $\Gamma(\cdot)$  is the Gamma function [11], and the parameter  $\sigma_{y,N}^2$  is equal to

$$\sigma_{y,N}^2 = \sigma^2 \underbrace{\left(1 - \sum_{k=0}^{N-1} \frac{(-1)^k \binom{N-1}{k}}{(k+1)^2}\right)}_{\Theta_N}. \quad (9)$$

$\Theta_N$  in (9) has been defined for clarity purpose.

It is worth mentioning that  $X$  and  $Y$  in (5) are uncorrelated, but not independent, since we know that  $X \geq \frac{N}{N-1} Y$ . Therefore, the joint distribution  $f_{X,Y}$  may be hard to derive, then we make the following approximation:

$$\begin{aligned} CCDF(\lambda) &= \mathbb{P}(X \geq \tilde{\lambda} Y, \tilde{\lambda} Y \geq 0) \\ &\approx \int_0^{+\infty} f_Y(y) \int_{\tilde{\lambda} y}^{+\infty} f_X(x) dx dy \\ &= 1 - \int_0^{+\infty} (1 - e^{-\frac{\tilde{\lambda} y}{\sigma^2}})^N f_Y(y) dy. \end{aligned} \quad (10)$$

Given the binomial theorem  $(1 - e^{-\frac{\tilde{\lambda} y}{\sigma^2}})^N = \sum_{k=0}^N \binom{N}{k} (-1)^k e^{-\frac{k\tilde{\lambda} y}{\sigma^2}}$  to develop (10) with (8), we obtain:

$$\begin{aligned} CCDF(\lambda) &\approx 1 - \sum_{k=0}^N \frac{(N-1)^{N-1} \binom{N}{k} (-1)^k}{(\Theta_N \sigma^2)^{(N-1)} \Gamma(N-1)} \\ &\quad \times \int_0^{+\infty} y^{N-2} e^{-\frac{(N-1)y}{\Theta_N \sigma^2} - \frac{k\tilde{\lambda} y}{\sigma^2}} dy. \end{aligned} \quad (11)$$

By substituting  $z = y(\frac{k\tilde{\lambda}}{\sigma^2} + \frac{N-1}{\Theta_N \sigma^2})$  in the exponential, we recognize the Gamma function [12], then (11) yields

$$\begin{aligned} CCDF(\lambda) &\approx 1 - \sum_{k=0}^N \frac{\binom{N}{k} (-1)^k (N-1)^{N-1}}{\Theta_N^{N-1} (\frac{N-1}{\Theta_N} + k\tilde{\lambda})^{N-1}} \\ &= - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k}{(1 + \frac{k\Theta_N \tilde{\lambda}}{N-1})^{N-1}}. \end{aligned} \quad (12)$$

It must be emphasized that (12) can be considered as a general CCDF expression of PAPR, which holds for low subcarriers number (or equivalently, for low samples number). In order to show that this expression still holds for large  $N$  value as well, we derive in Section IV an asymptotic analysis of PAPR when  $N$  tends to infinity.

#### IV. ASYMPTOTIC PAPR ANALYSIS

For clarity purpose, we remove the notation  $\lim_{N \rightarrow +\infty}$  throughout this section, as we assume that  $N$  is large in all developments. Also, we note  $\sim$  instead of  $\underset{N \rightarrow +\infty}{\sim}$ . First, we show that the CCDF (12) asymptotically tends to the usual expression (2), and second, we derive the mean and variance of PAPR.

### A. Asymptotic CCDF

As preliminary results, it is straightforward to show that  $\tilde{\lambda} \sim \lambda$ , and  $\Theta_N \sim 1$ . This equivalence can be proved by reminding that  $\sigma_{y,N}^2 = \sigma^2 \Theta_N$  on one hand, and from (27),  $\sigma_{y,N}^2 = \sigma^2 - \mathbb{E}\left\{\frac{|X_{i,m}|^2}{N}\right\} \sim \sigma^2$  on the other hand. Then, for any  $x \in \mathbb{R}$  we have  $(1 + \frac{x}{N})^N \sim e^x$ . The substitution of the previous equivalences into (12), in addition to the binomial theorem, lead to

$$\begin{aligned} CCDF(\lambda) &\sim - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k}{e^{k\lambda}} \\ &= 1 - \sum_{k=0}^N \frac{\binom{N}{k} (-1)^k}{e^{k\lambda}} \\ &= 1 - \left(1 - e^{-\lambda}\right)^N. \end{aligned} \quad (13)$$

We recognize (2), which shows the relevance of the general expression of the CCDF in (12), which holds for large  $N$  values.

### B. Asymptotic Mean and Variance of PAPR

1) *Mean of PAPR*: It must be noticed that the mean and variance of PAPR cannot be directly derived from (12) since it involves integrals with respect to  $\lambda$  from 0 to infinity. However, it has been stated that (5) holds only if  $\lambda < N$ . Therefore, we make the following approximation: we suppose that  $N$  is large enough so that the numerator and the denominator of the PAPR are uncorrelated, but the denominator is still considered as a random variable. Then, we can rewrite (5) and redefine the variable  $Y$  as

$$CCDF(\lambda) = \mathbb{P}\left(\underbrace{\max_{n \in \llbracket 0, N-1 \rrbracket} |x_n|^2}_X \geq \lambda \frac{1}{N} \underbrace{\sum_{n=0}^{N-1} |x_n|^2}_Y\right) \quad (14)$$

where  $X$  has the same distribution as previously (see (7)), and  $Y$  obeys a  $\chi^2$  distribution with  $2N$  degrees of freedom and parameter  $\sigma^2$ . By following the same reasoning as in (10)-(12), the CCDF from (14) can be derived as follows:

$$CCDF(\lambda) = - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k}{\left(1 + \frac{k\lambda}{N}\right)^N}. \quad (15)$$

We then deduce the distribution of PAPR, denoted  $f_\lambda$  as

$$\begin{aligned} f_\lambda(\lambda) &= \frac{\partial}{\partial \lambda} (1 - CCDF(\lambda)) \\ &= - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k k}{\left(1 + \frac{k\lambda}{N}\right)^{N+1}}. \end{aligned} \quad (16)$$

The mean, denoted by  $\mu_{PAPR}$ , whose derivation is proved in Appendix B, is expressed as

$$\begin{aligned} \mu_{PAPR} &= \int_0^{+\infty} \lambda f_\lambda(\lambda) d\lambda \\ &= - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k N^2}{k(N-1)(N-2)} \\ &\sim \gamma + \ln(N), \end{aligned} \quad (17)$$

where  $\gamma$  is the Euler-Mascheroni constant. It must be emphasized that (17) matches the results presented in [13], where the usual PAPR CCDF (2) is used.

2) *Variance of PAPR*: The variance of the PAPR, denoted by  $\nu_{PAPR}$ , is obtained from (15)-(17) as

$$\nu_{PAPR} = \underbrace{\int_0^{+\infty} \lambda^2 f_\lambda(\lambda) d\lambda}_{\tilde{\nu}_{PAPR}} - \mu_{PAPR}^2. \quad (18)$$

In the following, we focus on the development of  $\tilde{\nu}_{PAPR}$ . Similarly to (17), the asymptotic  $\tilde{\nu}_{PAPR}$  value is expressed as

$$\begin{aligned} \tilde{\nu}_{PAPR} &= - \sum_{k=1}^N \frac{\binom{N}{k} (-1)^k 2N^3}{k^2(N-1)(N-2)(N-3)} \\ &\sim \frac{\pi^2}{6} + (\gamma + \ln(N))^2, \end{aligned} \quad (19)$$

from which we finally deduce:

$$\nu_{PAPR} \sim \frac{\pi^2}{6}. \quad (20)$$

We prove (19) in Appendix C. Once again, this result matches those in [13], which validates the suggested PAPR form.

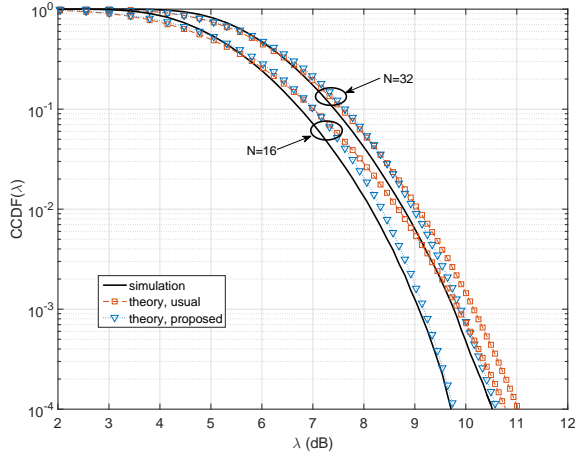
## V. SIMULATIONS AND DISCUSSION

### A. Simulations Results

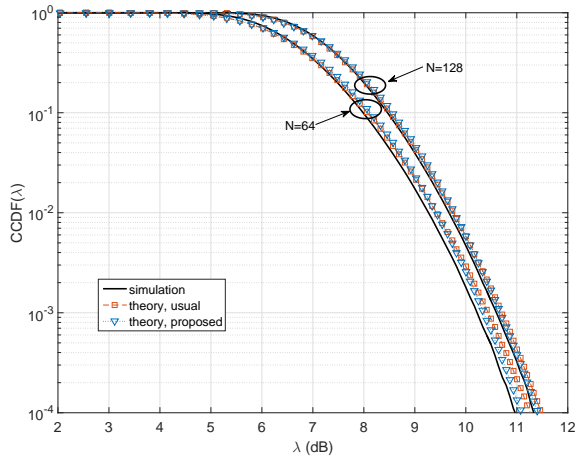
1) *Nyquist Rate Signal*: Simulations have been carried out to show the relevance of the proposed theoretical developments regarding PAPR of multicarrier signal with low subcarriers number ( $N \in \{16, 32, 64, 128\}$ ). In all simulations, a 16-quadratic amplitude modulation (QAM) constellation has been used. Note that higher order constellations leads to very similar results. The simulation results have been averaged over 10000 independent runs.

Fig. 3 depicts the CCDFs of PAPR versus  $\lambda$  (in dB), for (a)  $N=16, 32$  subcarriers, and (b)  $N=64, 128$  subcarriers. For each subcarriers number configuration, three curves are compared: one obtained through simulations, and two obtained from theory, referred as "usual" in (2), and "proposed" in (12). It can be generally observed that the suggested theoretical PAPR CCDF is closer to simulations than the usual one, for any  $N \in \{16, 32, 64, 128\}$ . In particular, Fig. 3 shows that at  $CCDF(\lambda) = 10^{-4}$ , the difference between proposed CCDF and simulation is less than 0.1 dB, whereas compared to (2), it is equal to 1, 0.5, 0.3, and 0.2 for  $N = 16, 32, 64$ , and 128, respectively. This shows the relevance of the proposed PAPR analysis for  $N \leq 128$  compared to the usual one. Furthermore, it shows that the larger  $N$ , the closer to simulation the CCDFs.





(a) CCDF of PAPR versus  $\lambda$  (dB),  $N=16, 32$  subcarriers.



(b) CCDF of PAPR versus  $\lambda$  (dB),  $N=64, 128$  subcarriers.

Fig. 3. CCDF of PAPR versus  $\lambda$  (dB), comparison of simulated (exact) CCDF, and usual theoretical CCDF (2) and proposed theoretical CCDF (12), using  $N=16, 32, 64, 128$  subcarriers.

However, it can be noted that the proposed CCDF does not perfectly match the simulations, in particular in small  $\lambda$  ranges, where it is even slightly worse than CCDF in (2). This is mainly due to the approximation made in (10).

Fig. 4 shows the mean and variance of PAPR versus  $N$  from 16 to 2048. Theoretical results in (17) and (20) are compared with those obtained through simulations. It can be clearly observed that  $\mu_{PAPR}$  tends to  $\gamma + \ln(N)$ , and  $\nu_{PAPR}$  tends to  $\frac{\pi^2}{6}$  as  $N$  increases. This shows the relevance of the asymptotic mean and variance analysis of PAPR, and then confirms the results presented in [13].

2) *Oversampled Signal*: It must be emphasized that previous developments and results have been presented for multicarrier signals sampled at Nyquist rate. However, for more practical scenarios, the PAPR of oversampled signal should be considered. In order to fit oversampled signals with low subcarriers number, we propose two approximations of  $CCDF(\lambda)$ :

- Van Nee based approximation [2], *i.e.*  $N$  is substituted by

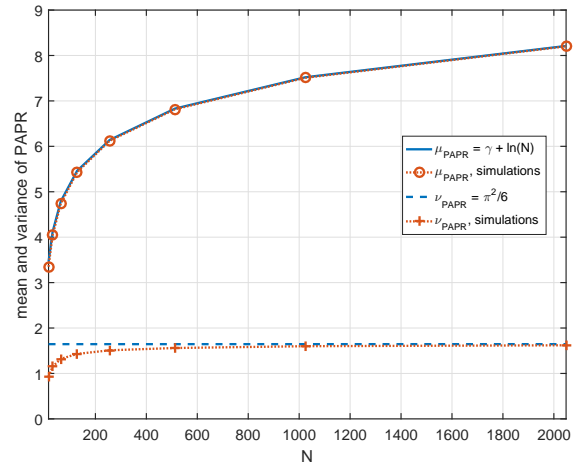


Fig. 4. Mean and variance of PAPR versus  $N$  for  $N = 16, 32, 64, 128, 256, 512, 1024, 2048$ , comparison between (17), (20), and simulations.

$[\alpha N]$  in (12), where  $[\cdot]$  is the nearest integer function. In that case, the coefficient  $\alpha$  is empirically set in order to fit the curves obtained by simulation. In the following, this approximation is referred as "adapted", since it is adapted from Van Nee approach in [2].

- Proposed approximation, where  $\tilde{\lambda}$  is substituted by  $\check{\lambda} = \frac{\beta \lambda}{1 - \frac{\beta \lambda}{N}}$  in (12), where  $\beta$  is a parameter which is empirically set. In the following, this approximation is referred as "proposed".

Furthermore, we compared the suggested approximation with CCDF expressions derived in the literature [4], [6], [7]. In [4], the authors expressed the CCDF of PAPR as:

$$CCDF(\lambda) = 1 - \exp\left(-N e^{-\lambda} \sqrt{\frac{\pi \lambda}{3}}\right). \quad (21)$$

In [6], [7], the authors based their developments on extreme value theory. The CCDF in [7] is expressed as

$$CCDF(\lambda) = 1 - \exp\left(-N e^{-\lambda} \sqrt{\frac{\pi \log(N)}{3}}\right), \quad (22)$$

and this result has been extended in [6] to more a realistic model where subcarriers may not be active, and with different power distribution. In that more general case, the CCDF of PAPR can be written as

$$CCDF(\lambda) = 1 - \exp\left(-2K e^{-\lambda} \sqrt{\frac{2\rho\pi\lambda K}{3N}}\right), \quad (23)$$

where  $K$  is half the number of activated subcarriers, and  $\rho$  is related to the power allocation.

In order to validate the suggested approximations, Figs. 5 and 6 show the CCDFs of PAPR versus  $\lambda$  for adapted and proposed approximations, respectively. The oversampling rate is equal to 20, then the PAPR of the signal is very close to that of the continuous signal. In Fig. 5, Van Nee based adaptation has been carried out with  $\alpha = 1.5$ . Furthermore,

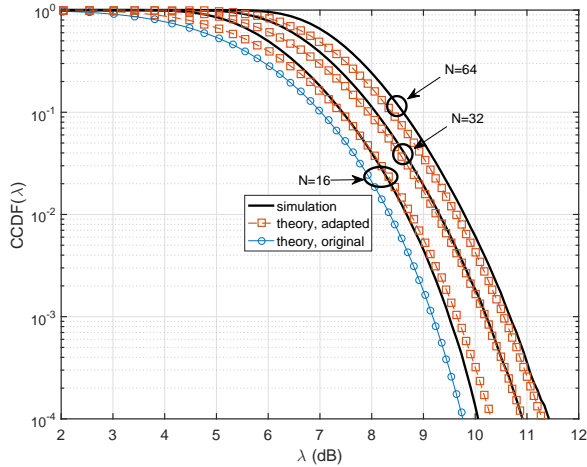


Fig. 5. CCDF of PAPR versus  $\lambda$  (dB) of oversampled OFDM signal using Van Nee based approximation. The "original" corresponds to Nyquist rate model in (12). The results have been obtained with an oversampling rate of 20,  $\alpha = 1.5$ .

the curve referred as "theory, original" corresponds to the multicarrier signal with  $N = 16$ , and sampled at Nyquist rate such as proposed in (12). It can be seen that the CCDF in (12) underestimates the actual CCDF of oversampled signal of about 0.4 dB. In addition, the Van Nee approximations do not perfectly match the curves obtained through simulations. Thus, a specific  $\alpha$  value should be set for every  $N$  value.

In Fig. 6, the proposed CCDF approximation has been carried out with  $\beta = 0.92$ . In that case, it can be observed that the proposed CCDFs match well the curves obtained through simulations for  $CCDF(\lambda) \leq 0.1$ . However, it must be noted that further simulations (not shown here) revealed that this approximation does not hold anymore for  $N > 128$  with  $\beta = 0.92$  ( $\beta$  should then be adapted), whereas Van Nee's does with  $\alpha = 2.8$  such as suggested in [2]. Thus, we can conclude that Van Nee's approximation can still be used for oversampled multicarrier signals with large  $N$  number ( $N \geq 128$ ), whereas the proposed one fits signals with low  $N$  number ( $N < 128$ ), even though it can also be used for larger  $N$  by adapting  $\beta$  value.

Fig. 7 compares the CCDF of PAPR of the proposed approximation with those of the aforementioned state-of-the-art [4], [6], [7], using a 16-QAM and the same oversampling rate as previously. The CCDF obtained through simulations has been plotted as a reference. Two subcarriers numbers are considered:  $N=16$ , in Fig. 7-(a), and  $N = 64$  in Fig. 7-(b). The curve of CCDF corresponding to (23) has been obtained by considering that all subcarriers are active, *i.e.*  $K = \frac{N}{2}$ . Note that in the case where  $\rho = 1$ , then (23) reduces to (21). Therefore, we arbitrarily set  $\rho = 0.6$  to distinguish (23) and (21). It can be observed in Fig. 7-(a) that the proposed approximation almost match the simulations, whereas the three other approximations deviate from simulation for  $\lambda \geq 7$  dB. However, in Fig. 7-(b), all the approximations are closer to the simulation. In particular, it is highlighted that, at  $CCDF(\lambda) = 0.0001$ , the proposed approximation, (22), and

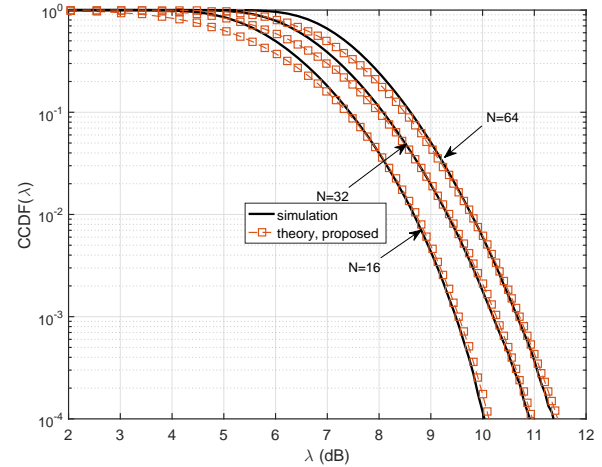


Fig. 6. CCDF of PAPR versus  $\lambda$  (dB) of oversampled OFDM signal using proposed approximation. The results have been obtained with an oversampling rate of 20,  $\beta = 0.92$ .

(23) are less than 0.2 dB far from the simulation. This shows the relevance of the suggested analysis and the general CCDF of PAPR expression (12) for low subcarriers number  $N < 64$ .

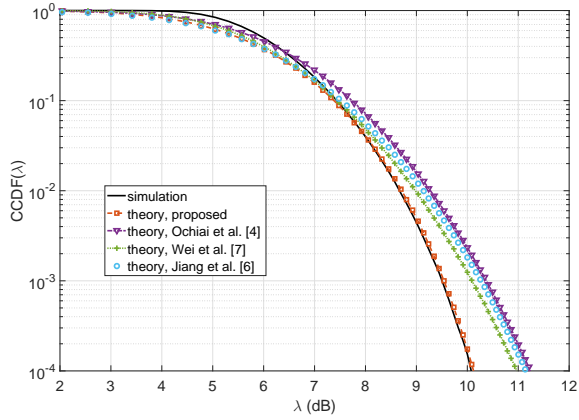
### B. Discussion

The various simulations results show the relevance of the suggested general expression of CCDF of PAPR, in particular for low subcarriers numbers  $12 \leq N \leq 64$ . However, it can be noted that (12) involves the derivation of two binomial coefficients per element of the sum. Therefore, the calculus of (12) may be not tractable in practice. In fact, if  $N$  is higher than 128, the derivation of the binomial coefficients exceeds the computing capacity of most of the present computers. However, it can be noted that  $\Theta_N$  in (9) quickly converges towards  $\frac{1}{N}(\gamma + \ln(N))$  (this is not proved in this paper), which reduces the computation cost of the CCDF for  $N > 32$ . Otherwise, computation tricks should be used to obtain the binomial coefficient in (12), *e.g.* by considering logarithmic additions instead of multiplications.

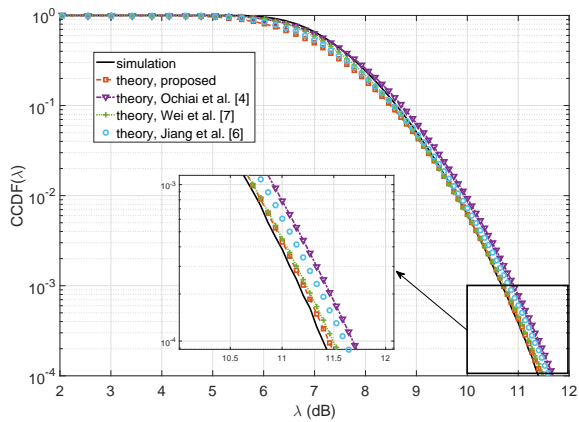
In practice, we conclude from the different simulations results that (12) should be used for very low subcarriers number, typically  $N < 128$  at Nyquist rate, and  $N < 64$  for oversampled signals. For higher  $N$  values, other approximations [2], [4], [6], [7] can be considered for simplicity matter.

## VI. CONCLUSION

In this paper, we presented a new and general expression of PAPR CCDF for multicarrier systems, which holds for large and low subcarriers number values, down to  $N \geq 12$ . This formula has been derived by considering the denominator of PAPR as a random variable, whereas it is usually assumed to be a constant. This led us to derive an original distribution function for this denominator. In addition, we showed that an asymptotic analysis of the PAPR leads to the existing results in the literature, in terms of CCDF, mean, and variance of PAPR. These developments, supported by various simulations results,



(a) CCDF of PAPR versus  $\lambda$  (dB),  $N=16$  subcarriers.



(b) CCDF of PAPR versus  $\lambda$  (dB),  $N=64$  subcarriers.

Fig. 7. CCDF of PAPR versus  $\lambda$  (dB), comparison of simulated (exact) CCDF, proposed approximation, and CCDF (21), (22), and (23) in [4], [7], and [6], using  $N=16, 64$  subcarriers.

validate the theoretical analysis. Furthermore, we discussed the issue of PAPR of oversampled signals with low subcarriers number, and we suggested an original approximation of PAPR CCDF for such signals. Future works will deal with more practical applications of the presented analysis, considering signals with low subcarriers number such as NB-IoT.

## APPENDIX

### A. Proof of (8)-(9)

Let  $\Omega_X = \{X_0, X_1, \dots, X_{N-1}\}$  be a set of  $N$  independent zero-mean complex white Gaussian variables with variance  $\sigma^2 = \mathbb{E}\{|X_n|^2\}$ . Then, the sum  $Z = \frac{1}{N} \sum_{n=0}^{N-1} |X_n|^2$  obeys a  $\chi^2$  distribution with  $2N$  degrees of freedom and parameter  $\sigma_Z^2$  (due to the fact that  $X_n$  are not normalized), *i.e.* the distribution of  $Z$  is written as

$$f_Z(z) = \frac{N^N}{\sigma_Z^{2N} \Gamma(N)} z^{N-1} e^{-\frac{Nz}{\sigma_Z^2}}, \quad (24)$$

where the parameter  $\sigma_Z^2$  can be straightforwardly deduced from (24). In fact, on one hand, we have

$$\mathbb{E}\{Z\} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{|X_n|^2\} = \sigma^2, \quad (25)$$

and, on the other hand, by property of random variables featuring continuous density functions, it can be shown that

$$\mathbb{E}\{Z\} = \int_0^{+\infty} z f_Z(z) dz = \sigma_Z^2, \quad (26)$$

which yields  $\sigma_Z^2 = \sigma^2$ . Note that the latter feature holds for any degree of  $\chi^2$  distribution. Thus, suppose that  $X_i, i \in [0, N-1]$ , is randomly chosen in  $\Omega_X$ , and let  $Z'$  be defined as  $Z' = Z - \frac{|X_i|^2}{N}$ . Then  $Z'$  obeys a  $\chi^2$  distribution with  $2N-2$  degrees of freedom with parameter  $\sigma_{Z'}^2 = \mathbb{E}\{Z - \frac{|X_i|^2}{N}\} = \frac{N-1}{N} \sigma^2$ .

Now, suppose that  $X_i$  is chosen such that  $|X_i|^2 = |X_{i,m}|^2$ , where  $|X_{i,m}|^2 = \max_{i \in [0, N-1]} |X_i|^2$  (the subscript  $m$  points out that  $|X_{i,m}|^2$  is the maximum value). Then, the variable  $Y = Z - \frac{|X_{i,m}|^2}{N}$  still obey a  $\chi^2$  with  $2N-2$  degrees of freedom with new parameter  $\sigma_Y^2$ , which is derived below. It must be noted that the distribution of  $X = |X_{i,m}|^2$  is given in (7), therefore the value of  $\sigma_Y^2$  is derived as follows:

$$\begin{aligned} \sigma_Y^2 &= \mathbb{E}\left\{Z - \frac{|X_{i,m}|^2}{N}\right\} \\ &= \sigma^2 - \frac{1}{N} \int_0^{+\infty} x f_X(x) dx. \end{aligned} \quad (27)$$

Then, the binomial theorem applied to  $f_X(x)$  leads to

$$\begin{aligned} f_X(x) &= \frac{N}{\sigma^2} (1 - e^{-\frac{x}{\sigma^2}})^{N-1} e^{-\frac{x}{\sigma^2}} \\ &= \sum_{k=0}^{N-1} \binom{N-1}{k} (-1)^k \frac{N}{\sigma^2} e^{-\frac{(k+1)x}{\sigma^2}}. \end{aligned} \quad (28)$$

Therefore, substituting (28) into (27) yields

$$\sigma_Y^2 = \sigma^2 - \frac{1}{N} \int_0^{+\infty} \sum_{k=0}^{N-1} \binom{N-1}{k} (-1)^k \frac{xN}{\sigma^2} e^{-\frac{(k+1)x}{\sigma^2}} dx. \quad (29)$$

Then we substitute  $w = \frac{(k+1)x}{\sigma^2}$  to recognize the Gamma function, which finally leads to (9), and then concludes the proof.

### B. Proof of (17)

We use the following equivalence  $\binom{N}{k} \sim \frac{N^k}{k!}$  to simplify  $\mu_{PAPR}$  in (17):

$$\mu_{PAPR} \sim - \sum_{k=1}^N \frac{(-1)^k N^k}{kk!}. \quad (30)$$

Let  $E_1(z) = \int_z^{+\infty} \frac{e^{-t}}{t} dt, z \in \mathbb{C} \setminus \mathbb{R}^-$ , be the exponential integral function, defined in [12]. The series representation of  $E_1$  is



## REFERENCES

- [1] G. Wunder, R. F. H. Fischer, H. Boche, S. Litsyn, and J.-S. No, "The PAPR Problem in OFDM Transmission: New Directions for a Long-Lasting Problem," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 130 – 144, October 2013.
- [2] R. Van Nee and A. De Wild, "Reducing the peak-to-average power ratio of OFDM," in *proc. of VTC'98*, Ottawa, Ont., Canada, May 1998.
- [3] R. Van Nee and R. Prasad, *OFDM for wireless multimedia communications*. Artech House, January 2000.
- [4] H. Ochiai and H. Imai, "On the Distribution of the Peak to Average Power Ratio in OFDM Signals," *IEEE Transactions on Communications*, vol. 49, no. 2, pp. 282 – 289, February 2001.
- [5] M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "On the Peak to Average of OFDM Signals Based on Oversampling," *IEEE Transactions on Communications*, vol. 51, no. 1, pp. 72 – 78, January 2003.
- [6] T. Jiang, M. Guizani, H. Chen, W. Xiang, and Y. Wu, "Derivation of PAPR Distribution for OFDM Wireless Systems Based on Extreme Value Theory," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1298 – 1305, April 2008.
- [7] S. Wei, D. L. Goeckel, and P. E. Kelly, "A modern extreme value theory approach to calculating the distribution of the peak-to-average power ratio in OFDM systems," in *proc. of ICC'02*, New York, NY, April-May 2002, pp. 1686 – 1690.
- [8] 3GPP, "3GPP TS 36.211, Physical channels and modulation (Release 14)," 3GPP, Tech. Rep., March 2017.
- [9] A. Rudziński, "Normalized Gaussian Approach to Statistical Modeling of OFDM Signals," *Journal of Telecommunications and Information Technology*, vol. 1, no. 1, pp. 54 – 61, January 2014.
- [10] B. Farhang-Boroujeny, "OFDM Versus Filter Bank Multicarrier," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 92 – 112, May 2011.
- [11] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1970, ch. 6: Gamma Function and Related Functions, pp. 258 – 266.
- [12] —, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, 1970, ch. 5: Exponential Integral and Related Functions, pp. 227 – 237.
- [13] H. Ochiai and H. Imai, "Peak-Power Reduction Schemes in OFDM Systems : A Review," in *proc. of WPMC'98*, 1998, pp. 247 – 252.
- [14] M. S. Milgram, "The Generalized Integro-Exponential Function," *Mathematics of Computation*, vol. 44, no. 170, pp. 443 – 458, April 1985.
- [15] I. K. Abu-Shumays, "Transcendental Functions Generalizing the Exponential Integrals," Northwestern University, Evanston, Illinois, Tech. Rep., November 1973.

$$E_1(z) = -\gamma - \ln(z) - \sum_{k=1}^{+\infty} \frac{(-1)^k z^k}{k k!}. \quad (31)$$

Since  $\lim_{z \rightarrow +\infty} E_1(z) = 0$ , and reminding that  $N \rightarrow +\infty$ , then the substitution of  $z$  by  $N$  in (31) leads to  $\mu_{PAPR}$  in (17), which concludes the proof.

### C. Proof of (19)

Similarly to the reasoning in Appendix B, it should be first noticed that the first line in (19) becomes

$$\tilde{\nu}_{PAPR} \sim -2 \sum_{k=1}^N \frac{(-1)^k N^k}{k^2 k!}. \quad (32)$$

Then, we use the series expansion of the so-called generalized integro-exponential function (see (2.3) and (2.10) in [14]), denoted by  $E_s^j$ , such as hereby described:

$$\begin{aligned} E_s^j(z) &= \frac{1}{\Gamma(j+1)} \int_1^{+\infty} \ln(t)^j t^{-s} e^{-zt} dt \\ &= \sum_{\substack{k=0 \\ k \neq s-1}}^{+\infty} \frac{(-z)^k}{(s-1-k)^{j+1} k!} \\ &\quad + \frac{z^{s-1} (-1)^{j+s}}{(j+1)!} \sum_{k=0}^{j+1} \binom{j+1}{k} \ln(z)^{1+j-k} \Psi_{k,s}, \end{aligned} \quad (33)$$

where  $\Psi_{k,s}$  is the polygamma function [11], whose analytical values are provided in [14], Table 1, and in [15]. It can be noticed that, when  $s = j = 1$ :

$$E_1^1(z) = \sum_{k=1}^{+\infty} \frac{(-1)^k z^k}{k^2 k!} + \frac{1}{2} \sum_{k=0}^2 \binom{2}{k} \ln(z)^{2-k} \Psi_{k,1}, \quad (34)$$

where

$$\begin{aligned} \Psi_{0,1} &= 1 \\ \Psi_{1,1} &= \gamma \\ \Psi_{2,1} &= \gamma^2 + \frac{\pi^2}{6}. \end{aligned} \quad (35)$$

From the integral form of  $E_s^j(z)$  in (33) we have  $\lim_{z \rightarrow +\infty} E_1^1(z) = 0$ . Then considering  $z = N$  and substituting (32) into (34) finally leads to

$$\tilde{\nu}_{PAPR} \sim \frac{\pi^2}{6} + (\gamma + \ln(N))^2, \quad (36)$$

which concludes the proof.