



HAL
open science

Compression of Clipped OFDM IQ Samples for Cloud Radio Access Network

Aya Shehata, Philippe Mary, Matthieu Crussière

► **To cite this version:**

Aya Shehata, Philippe Mary, Matthieu Crussière. Compression of Clipped OFDM IQ Samples for Cloud Radio Access Network. 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Sep 2021, Helsinki, Finland. hal-03475354v2

HAL Id: hal-03475354

<https://hal.science/hal-03475354v2>

Submitted on 10 Dec 2021

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.

Compression of Clipped OFDM IQ Samples for Cloud Radio Access Network

Aya Shehata, Philippe Mary and Matthieu Crussière

Univ Rennes, INSA Rennes, CNRS, IETR-UMR 6164, 4 F-35000 Rennes, France

Email: aya.shehata@insa-rennes.fr

Abstract—In this paper, we analyze the impact of reducing the Peak to Average Power Ratio (PAPR) of a time domain Orthogonal Frequency Division Multiplexing (OFDM) signal before its compression, as required for the transmission of IQ samples in a Centralized Radio Access Network architecture. The analysis is performed considering a clipping function for PAPR reduction. First, the distribution of the clipped OFDM IQ samples is obtained which in turn allows to derive the threshold and codebook levels of a non-uniform quantizer optimized for clipped IQ samples and the number of representation bits for each quantized levels in the entropy coding stage. Then, we provide a closed-form expression for the Modulation Error Ratio of the compressed and clipped OFDM signal. The simulation results show that the Gaussian quantizer applied to a clipped signal is not quite robust compared to a quantizer that has been optimized w.r.t. the distribution of clipped IQ samples, especially in the low-rate quantization region.

Index Terms—C-RAN, Clipping, Quantization, Entropy coding, IQ Compression, MER

I. INTRODUCTION

Cloud Radio Access Network (C-RAN) is an attractive architecture for next generation wireless networks. The C-RAN architecture is deporting the baseband processing in a common server in the cloud, called base band unit (BBU), and transmits the generated digitized IQ samples to the remote radio unit, i.e. the transmitter, via the fronthaul link [1], [2]. This architecture implies that the physical layer including the baseband processing and network layer functionalities are moved to the BBU [3]. A fully-centralized solution offers high flexibility, low maintenance and operational costs. On the other hand, the fronthaul link should have a large bandwidth to transport the high resolution IQ samples generated by the BBU, and may become the bottleneck of such an architecture. Optical fiber offers a large bandwidth and can be a natural solution for IQ transportation, but with the drawback of very high cost though. Hence, IQ data compression before their transmission over the capacity-constrained fronthaul link is an interesting approach to reduce the data rate over the fronthaul link for the C-RAN architecture.

Several data rate compression techniques have recently been investigated, each of them presenting a trade-off between the achievable compression performance, design complexity, computational delay, and signal distortion [4]. Quantization and entropy coding are significant approaches in compressing the data rate over the fronthaul link. In [5], non-uniform quantization based on an iterative gradient algorithm was proposed.

Lloyd-based non-uniform quantizer and entropy coding were used in [6] to exploit the temporal characteristics of the signal. Vector quantization was studied in [7] to efficiently exploit the correlation between samples for an improved compression rate. However, this gain was at the cost of much higher computational complexity compared to scalar quantization.

On the other hand, the high Peak to Average Power Ratio (PAPR) of the Orthogonal Frequency Division Multiplexing (OFDM) signal remains a major drawback in the implementation of the OFDM system. Many dedicated algorithms were introduced in the literature to reduce the PAPR of the signal. In this paper, we are interested in clipping, the most straightforward PAPR reduction technique despite being a distortion approach itself [8].

At the output of the IFFT modulator, OFDM samples are generally assumed to be Gaussian distributed, thanks to the central limit theorem, and hence quantization and entropy coding techniques are optimized on this basis. However, implementing a PAPR reduction technique changes the IQ samples distribution. To the best of the authors knowledge, no studies on the robustness of the Gaussian-optimized quantizer have been performed considering any of the PAPR reduction procedures before compression or optimizing the compression techniques based on the distribution induced by the PAPR reduction.

Thus, our contribution in this paper can be summarized as follows. First, the distribution of the clipped IQ samples is derived, taking into account that the clipping is performed at the amplitude of the signal. The obtained IQ distribution is further exploited to optimize the compression blocks, i.e. quantizer and entropy coding, and to obtain a closed-form expression of the modulation error ratio (MER), hereby modeling the joint effect of the clipping and quantization operations. Then, we show through numerical simulations that the compression techniques optimized for Gaussian distribution are not optimal when applied to clipped samples and better performance on the MER can be obtained when using a quantizer optimized with the true distribution of clipped samples.

The rest of the paper is organized as follows. Section II describes the system model. Section III details the analytical study of the clipped signal IQ distribution and the derivation of the MER. In Section IV, we present numerical simulations that evaluate the accuracy of our theoretical findings and emphasize the gain in system MER performance achieved by

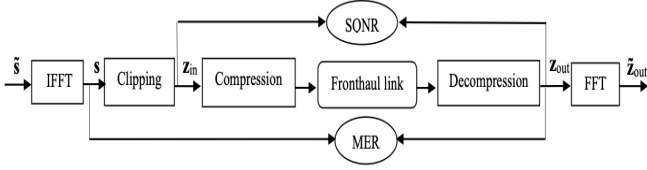


Figure 1. Downlink system model of a C-RAN system.

clipped optimized compression techniques. Finally, Section V concludes the paper.

II. SYSTEM MODEL

The downlink OFDM transmission chain is considered as depicted in Fig. 1, with N_{fft} subcarriers and M -QAM modulation. IQ samples are clipped in order to reduce PAPR and they are compressed before being transmitted through the fronthaul link. After the IFFT block, $\mathbf{s} \in \mathbb{C}^{N_{\text{fft}} \times 1}$ represents the complex OFDM signal expressed as

$$s(k) = \frac{1}{\sqrt{N_{\text{fft}}}} \sum_{n=0}^{N_{\text{fft}}-1} \tilde{s}(n) e^{j2\pi k \frac{n}{N_{\text{fft}}}}, \quad \forall k = [1, \dots, N_{\text{fft}}] \quad (1)$$

where, $\tilde{\mathbf{s}} = [\tilde{s}(0), \dots, \tilde{s}(N_{\text{fft}} - 1)] \in \mathbb{C}^{N_{\text{fft}} \times 1}$ is a vector of QAM symbols modulating the subcarriers. The vector \mathbf{z}_{in} contains the clipped IQ samples before the compression and \mathbf{z}_{out} after the decompression. For N_{fft} sufficiently large, the real and imaginary parts of a complex sample $s(k)$, $\forall k \in \{1, \dots, N_{\text{fft}}\}$, denoted by $x_s(k) = \text{Re}(s(k))$ and $y_s(k) = \text{Im}(s(k))$, respectively, are zero-mean Gaussian distributed. Hence, the magnitude $r_s(k) = |s(k)|$ is Rayleigh distributed and the phase $\theta_s(k) = \angle s(k)$ is uniformly distributed in $[-\pi, \pi]$.

A. Clipping PAPR Reduction Technique

Clipping consists in limiting the high amplitude peaks of the signal to a certain threshold V_{max} without affecting the phase of the signal. It is a widely used PAPR reduction technique due to its high simplicity and reduction gain. The clipped time domain signal $z_{\text{in}}(k)$ is

$$z_{\text{in}}(k) = \begin{cases} s(k), & \text{if } |s(k)| < V_{\text{max}} \\ V_{\text{max}} e^{j\theta_s}, & \text{if } |s(k)| \geq V_{\text{max}}. \end{cases} \quad (2)$$

The amplitude of the signal at the input of the clipper is distributed as

$$f_{R_s}(r) = \frac{r}{\sigma_s^2} e^{-\frac{r^2}{2\sigma_s^2}}, \quad r \in [0, +\infty[\quad (3)$$

where R_s and σ_s are the random variable (RV) representing the amplitude of the input signal and its standard deviation, respectively. Clipping is applied on the signal amplitude as in (2) and it modifies the amplitude distribution such as [9]

$$f_{R_{\text{in}}}(r) = \begin{cases} f_{R_s}(r), & \text{if } r < V_{\text{max}} \\ e^{-\Lambda} \delta(r - V_{\text{max}}), & \text{if } r = V_{\text{max}} \end{cases} \quad (4)$$

where R_{in} is the RV representing the amplitude of the signal after clipping. $\delta(r)$ is the Dirac distribution, $e^{-\Lambda}$ is the probability that the amplitude of \mathbf{s} exceeds V_{max} , i.e. $\int_{V_{\text{max}}}^{\infty} f_{R_s}(r) dr = e^{-\Lambda}$, and $\Lambda = \frac{V_{\text{max}}^2}{2\sigma_s^2}$ is the clipping ratio. Despite the simplicity of the clipping technique, it causes potentially high signal distortions.

B. Compression Techniques

a) *Scalar Quantization*: Each IQ sample is quantized to one of the N quantization levels, each represented over R bits. Scalar N -level non-uniform quantizer (NUQ) can be optimally designed for a given distribution of samples in the Minimum Mean Square Error (MMSE) sense based on the well known Lloyd algorithm (LA) [10]. According to LA, decision thresholds t_i , $\forall i = [1, \dots, N - 1]$ are

$$t_i = \frac{q_i + q_{i+1}}{2}, \quad (5)$$

where, t_0 and t_N are set to the minimum and maximum possible values of the signal, respectively. Quantization levels q_i , $\forall i = [1, \dots, N]$, are the centroid of each decision region, e.g. for the in-phase component

$$q_i = \frac{\int_{t_{i-1}}^{t_i} x f_X(x) dx}{\int_{t_{i-1}}^{t_i} f_X(x) dx}, \quad (6)$$

where $f_X(x)$ is the Probability Density Function (PDF) associated to the in-phase component $x(k) = \text{Re}(z_{\text{in}}(k))$, $\forall k$, of the clipped signal.

b) *Entropy Coding*: Entropy Coding (EC) is a lossless compression technique, where additional compression is gained by eliminating the redundancy from the distribution of the quantization levels. A large number of bits is used to represent levels with a low probability of occurrence while a lower number of bits for the levels of higher probability. Huffman coding [11] is the most practical EC technique approaching the Shannon's lossless source coding theorem. Thus, the average codeword length assigned to the i^{th} quantization level is

$$L_{q_i} = -\log_2 \int_{t_{i-1}}^{t_i} f_X(x) dx. \quad (7)$$

C. Performance Metrics

Rate R and MER are the metrics used to evaluate the performance of the studied model. R is defined as the average number of bits required to represent a single IQ sample. MER indicates the distortion experienced by the signal and gives a figure of merit for system performance by comparing the actual location of a received sample to its ideal location. It is defined as the ratio between the power of the original signal over the distortion introduced to the original signal. Therefore, the MER of the decompressed signal \mathbf{z}_{out} is expressed as follows:

$$\text{MER} = \frac{\mathbb{E}[|S|^2]}{\mathbb{E}[|S - Z_{\text{out}}|^2]}, \quad (8)$$

with $\mathbb{E}(\cdot)$ is the expectation operator, S and Z_{out} are the RV representing the IQ samples before the compression and after

decompression, respectively. Through our analysis, signal to quantization noise ratio (SQNR) will be the metric used to evaluate the performance of the compression techniques, i.e. of the quantizer, since EC is distortion-less. In that case, \mathbf{z}_{in} will be considered as the reference signal and SQNR is defined as

$$\text{SQNR} = \frac{\mathbb{E}[|Z_{\text{in}}|^2]}{\mathbb{E}[|Z_{\text{in}} - Z_{\text{out}}|^2]}, \quad (9)$$

where Z_{in} is the RV representing the clipped samples.

III. ANALYTICAL STUDY OF THE OPTIMIZED COMPRESSION TECHNIQUES FOR A CLIPPED SIGNAL

In this section, we first derive the distribution of the clipped IQ samples, i.e. $f_X(x)$ and $f_Y(y)$ for the In-phase and Quadrature-phase parts, respectively. Second, we derive the parameters of the NUQ and EC blocks for the new distribution of the signal with reduced PAPR. Finally, based on the derived IQ distribution, we provide both MER and SQNR closed-form expressions for system evaluation considering clipping and quantization distortions.

A. Distribution of the Clipped IQ samples

As reminded in II-A, clipping modifies the signal amplitude without affecting the phase. Thus, $\forall k \in \{1, \dots, N_{\text{fit}}\}$, $r_{\text{in}}(k) = |z_{\text{in}}(k)|$ and $\theta_{\text{in}}(k) = \angle z_{\text{in}}(k)$ are samples drawn from two independent random variables, i.e. R_{in} and Θ_{in} , respectively. Hence, the joint PDF of $(R_{\text{in}}, \Theta_{\text{in}})$ is the product of the marginal PDFs, i.e. $f_{R_{\text{in}}, \Theta_{\text{in}}}(r, \theta) = f_{R_{\text{in}}}(r)f_{\Theta_{\text{in}}}(\theta)$, which yields:

$$f_{R_{\text{in}}, \Theta_{\text{in}}}(r, \theta) = \begin{cases} \frac{r}{2\pi\sigma_s^2} e^{-\frac{r^2}{2\sigma_s^2}}, & \text{if } r < V_{\text{max}} \\ \frac{e^{-\Lambda}}{2\pi} \delta(r - V_{\text{max}}), & \text{if } r = V_{\text{max}} \end{cases} \quad (10)$$

The conversion from polar to Cartesian coordinates of the complex samples is such that $(x, y) = \varphi(r, \theta)$ with:

$$\begin{aligned} x &= r \cos \theta & r &= \sqrt{x^2 + y^2} \\ y &= r \sin \theta & \theta &= \varphi_2^{-1}(x, y), \end{aligned} \quad (11)$$

where, φ is the one-to-one mapping defined from the set $D = [0, V_{\text{max}}] \times [-\pi, \pi]$ to $\Delta =]-V_{\text{max}}, V_{\text{max}}[$, and $\varphi_2^{-1}:]-V_{\text{max}}, V_{\text{max}}[\mapsto [-\pi, \pi]$ with:

$$\varphi_2^{-1}(x, y) = \begin{cases} \tan^{-1}\left(\frac{y}{x}\right), & \text{if } x > 0 \\ \tan^{-1}\left(\frac{y}{x}\right) - \pi, & \text{if } x < 0, y < 0 \\ \pi - \tan^{-1}\left(\frac{y}{x}\right), & \text{if } x < 0, y \geq 0 \\ \pi/2 & \text{if } x = 0, y > 0. \\ -\pi/2 & \text{if } x = 0, y < 0. \end{cases} \quad (12)$$

According to the change of variables Theorem, the joint PDF of the couple (X, Y) is

$$f_{X, Y}(x, y) = |J(x, y)| f_{R_{\text{in}}, \Theta_{\text{in}}}(r(x, y), \theta(x, y)), \quad (13)$$

where $|J(x, y)|$ is the module of the Jacobian of the mapping φ computed as

$$|J(x, y)| = \left| \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} \right| = \frac{1}{\sqrt{x^2 + y^2}}, \quad \forall (x, y) \neq (0, 0) \quad (14)$$

Thus, substituting (10) and (14) into (13), the marginal PDF of the Cartesian coordinate RV is obtained as

$$f_X(x) = \int_{-V_{\text{max}}}^{V_{\text{max}}} \left(\frac{1}{2\pi\sigma_s^2} e^{-\frac{x^2+y^2}{2\sigma_s^2}} + \frac{e^{-\Lambda}}{2\pi\sqrt{x^2+y^2}} \delta(\sqrt{x^2+y^2} - V_{\text{max}}) \right) dy. \quad (15)$$

Finally, using the integral identity in [12, eq. (1), Sec. 3.321, p. 336] and adequately rearranging the terms leads to the following distribution of the In-phase component $f_X(x)$ ¹, i.e.

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_s^2}} e^{-\frac{x^2}{2\sigma_s^2}} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2\sigma_s^2}}\right) + \frac{e^{-\Lambda}}{2\pi} \frac{1}{\sqrt{V_{\text{max}}^2 - x^2}}, \quad x \in]-V_{\text{max}}, V_{\text{max}}[\quad (16)$$

where $\text{erf}(u) = (2/\sqrt{\pi}) \int_0^u \exp(-t^2) dt$ is the error function. From this distribution, it comes that without clipping, i.e. when $V_{\text{max}} \rightarrow \infty$ and thus $\Lambda \rightarrow \infty$, the second term vanishes and the error function tends to one. This ends with the zero-mean Gaussian distribution of the OFDM signal, which is expected from the elimination of the clipping distortion effect.

B. Derivation of Optimized Quantizer and EC Parameters

Based on the PDF of the clipped IQ samples given in (16), the codebook quantization levels for the real and imaginary parts can be derived by injecting (16) into (6) using adequate variable changes along with the integral identity in [12, eq. (1), Sec. 3.321, p. 336]. This yields,

$$q_i = \frac{\frac{\sigma_s}{\sqrt{2\pi}} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2\sigma_s}}\right) \left[e^{-\frac{t_{i-1}^2}{2\sigma_s^2}} - e^{-\frac{t_i^2}{2\sigma_s^2}} \right] + \frac{e^{-\Lambda}}{2\pi} \left[\sqrt{V_{\text{max}}^2 - t_{i-1}^2} - \sqrt{V_{\text{max}}^2 - t_i^2} \right]}{\text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2\sigma_s}}\right) \left[Q\left(\frac{t_{i-1}}{\sigma_s}\right) - Q\left(\frac{t_i}{\sigma_s}\right) \right] + e^{-\frac{\Lambda}{2\sigma_s^2}} \left[\arcsin\left(\frac{t_{i-1}}{V_{\text{max}}}\right) - \arcsin\left(\frac{t_i}{V_{\text{max}}}\right) \right]}, \quad (17)$$

where $Q(u) = (1/\sqrt{2\pi}) \int_u^\infty \exp(-v^2/2) dv$ is the Q-function. Decision thresholds are the mid-points between quantization levels according to (5). The quantizer is followed by the entropy coder where the average codeword length assigned to a quantization level is derived for the clipped IQ PDF based on (7) as:

$$L_{q_i} = -\log_2 \left(\text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2\sigma_s}}\right) \left[Q\left(\frac{t_{i-1}}{\sigma_s}\right) - Q\left(\frac{t_i}{\sigma_s}\right) \right] + e^{-\frac{\Lambda}{2\sigma_s^2}} \left[\arcsin\left(\frac{t_{i-1}}{V_{\text{max}}}\right) - \arcsin\left(\frac{t_i}{V_{\text{max}}}\right) \right] \right). \quad (18)$$

Without clipping, i.e. when $V_{\text{max}} \rightarrow \infty$ and thus $\Lambda \rightarrow \infty$, the error function tends to one and the exponential term tends to zero in (17) and (18) and we end up with the parameters of a quantizer and an entropy coder optimized for a zero-mean Gaussian distribution analyzed in [13].

¹The distribution of the quadrature-phase component, $f_Y(y)$, derives from the same steps.

C. Derivation of Asymptotic MER and SQNR

MER in (8) can be shown to be equal to

$$\text{MER} = \frac{\mathbb{E}[|S|^2]}{\mathbb{E}[|S - Z_{\text{out}}|^2]} = \frac{P_s}{D_C + D_Q}, \quad (19)$$

where $P_s = 2\sigma_s^2$ is the power of a complex Gaussian distributed signal with standard deviation σ_s , D_C and D_Q are the power of clipping and quantization distortions, respectively, which are considered as statistically independent [14].

Let us first derive the clipping distortion D_C which can be calculated as follows:

$$D_C = \mathbb{E}[|S - Z_{\text{in}}|^2] = \mathbb{E}[(R_s - R_{\text{in}})^2], \quad (20)$$

Expanding the expectation operator and taking the clipping definition (2) into consideration leads to

$$D_C = \mathbb{E}[(R_s - R_{\text{in}})^2] = \int_{V_{\text{max}}}^{\infty} (r_s - V_{\text{max}})^2 f_{R_s}(r) dr. \quad (21)$$

After expanding the squared term and using the integral identity in [12, eq. (9), Sec. 3.381, p. 346], D_C can be expressed as:

$$D_C = 2\sigma_s^2 \Gamma\left(2, \frac{V_{\text{max}}^2}{2\sigma_s^2}\right) - 2^{\frac{3}{2}} \sigma_s V_{\text{max}} \Gamma\left(\frac{3}{2}, \frac{V_{\text{max}}^2}{2\sigma_s^2}\right) + V_{\text{max}}^2 e^{-\frac{V_{\text{max}}^2}{2\sigma_s^2}}, \quad (22)$$

where, $\Gamma(u, v) = \int_v^{\infty} e^{-t} t^{u-1} dt$ is the upper incomplete Gamma function. On the other hand, the asymptotically rate-distortion function for a non-uniform scalar quantizer is derived in terms of the quantizer input signal PDF and number of quantization levels N in [15] and known as "Panter and Dite formula":

$$D(R) \sim \frac{2^{-2R}}{12} \left(\int_{-\infty}^{\infty} \sqrt[3]{f_X(x)} dx \right)^3. \quad (23)$$

As shown previously in III-A, the in-phase and the quadrature-phase of a clipped complex signal at the input of the quantizer are identical. Thus, the quantization distortion is the same on both components, and for complex samples can be derived as

$$D_Q = \frac{2^{-2R}}{6} \underbrace{\left(\int_{-\infty}^{\infty} \sqrt[3]{f_X(x)} dx \right)^3}_I. \quad (24)$$

The evaluation of (24) using the PDF of the clipped IQ samples given in (16) leads to the following integral derivation,

$$I = \int_{-V_{\text{max}}}^{V_{\text{max}}} \left(\frac{1}{\sqrt{2\pi}\sigma_s} e^{-\frac{x^2}{2\sigma_s^2}} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2}\sigma_s}\right) + \frac{e^{-\Lambda}}{2\pi} \frac{1}{\sqrt{V_{\text{max}}^2 - x^2}} \right)^{\frac{1}{3}} dx \quad (25)$$

Let us assume the following notations

$$a = \frac{1}{\sqrt{2\pi}\sigma_s} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2}\sigma_s}\right), \quad b = \frac{1}{2\sigma_s^2}, \quad \text{and} \quad c = \frac{e^{-\Lambda}}{2\pi} \quad (26)$$

Thus, using the generalized Newton binomial theorem [16], I can be developed as

$$\begin{aligned} I &= \int_{-V_{\text{max}}}^{V_{\text{max}}} \sum_{k=0}^{\infty} \binom{\frac{1}{3}}{k} \left(a e^{-bx^2} \right)^{\frac{1}{3}-k} \left(\frac{c}{\sqrt{V_{\text{max}}^2 - x^2}} \right)^k dx \\ &= \sum_{k=0}^{\infty} \binom{\frac{1}{3}}{k} \int_{-V_{\text{max}}}^{V_{\text{max}}} \frac{\left(a e^{-bx^2} \right)^{\frac{1}{3}-k} c^k}{\left(\sqrt{V_{\text{max}}^2 - x^2} \right)^k} dx, \end{aligned} \quad (27)$$

where $\binom{r}{k}$, is the generalized binomial coefficient defined as $\binom{r}{k} = \frac{(r)_k}{k!}$, and $(r)_k = \prod_{n=0}^{k-1} (r-n)$, is the falling factorial. Rearranging the terms in I leads to

$$I = \sum_{k=0}^{\infty} \binom{\frac{1}{3}}{k} c^k a^{\frac{1-3k}{3}} \underbrace{\int_{-V_{\text{max}}}^{V_{\text{max}}} \frac{e^{-\frac{b(1-3k)}{3}x^2}}{\left(\sqrt{V_{\text{max}}^2 - x^2} \right)^k} dx}_{I_1}. \quad (28)$$

Thus, with proper variable substitution along with the integral property in [12, eq. (1), Sec. 3.383, p. 347], I_1 is solved as

$$I_1 = \frac{1}{b^{\frac{1-k}{2}}} \left[B\left(\frac{2-k}{2}, \frac{1}{2}\right) (dV_{\text{max}}^2)^{\frac{1-k}{2}} \times {}_1F_1\left(\frac{1}{2}; \frac{3-k}{2}; -dV_{\text{max}}^2\right) \right] \quad \text{with, } k \leq 1, \quad (29)$$

where $d = \frac{b(1-3k)}{6\sigma_s^2}$, $B(u, v) = 2 \int_0^1 t^{u-1} (1-t^2)^{v-1} dt$ is the Beta function, and ${}_1F_1$ is the confluent hyper geometric function. Finally, substituting I_1 into I , filling I into (24), and with further simplification of the expression, D_Q can finally be expressed as follows:

$$D_Q = \frac{2^{-2R}}{6} \left[\frac{2\sqrt[3]{a}V_{\text{max}}}{\sqrt{6b}\sigma_s} {}_1F_1\left(\frac{1}{2}; \frac{3}{2}; -\frac{V_{\text{max}}^2}{6\sigma_s^2}\right) + \frac{-26cV_{\text{max}}}{25\sqrt{3}\sqrt[3]{a^2}\sigma_s} {}_1F_1\left(\frac{1}{2}; 1; \frac{V_{\text{max}}^2}{3\sigma_s^2}\right) \right]^3. \quad (30)$$

Finally, it comes that the MER depends only on the input signal standard deviation σ_s , clipping amplitude threshold V_{max} , and the number of bits per sample R .

SQNR allows to assess the distortion caused by the quantizer only and it expresses as

$$\text{SQNR} = \frac{\mathbb{E}[|Z_{\text{in}}|^2]}{\mathbb{E}[|Z_{\text{in}} - Z_{\text{out}}|^2]} = \frac{P_{z_{\text{in}}}}{D_Q}, \quad (31)$$

with D_Q given by (30). $P_{z_{\text{in}}}$ is the power of a complex clipped IQ sample, computed with the distribution in (16), i.e.

$$P_{z_{\text{in}}} = 2 \int_{-V_{\text{max}}}^{V_{\text{max}}} x^2 f_X(x) dx. \quad (32)$$

Using the integral identity in [12, eq. (5), Sec. 3.321, p. 336] for the exponential function, $P_{z_{\text{in}}}$ is expressed as

$$P_{z_{\text{in}}} = \frac{2\sigma_s^2}{\sqrt{\pi}} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2}\sigma_s}\right) \left[\sqrt{\frac{\pi}{2}} \text{erf}\left(\frac{V_{\text{max}}}{\sqrt{2}\sigma_s}\right) - \frac{V_{\text{max}}}{\sqrt{2}\sigma_s} e^{-\frac{V_{\text{max}}^2}{2\sigma_s^2}} \right] + \frac{V_{\text{max}}^2}{4} e^{-\frac{V_{\text{max}}^2}{2\sigma_s^2}}. \quad (33)$$

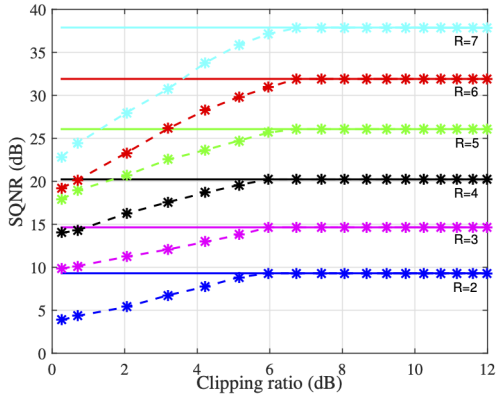


Figure 2. SQNR vs clipping ratio of a Gaussian quantizer applied on a clipped signal (dashed lines) w.r.t. the number of bits.

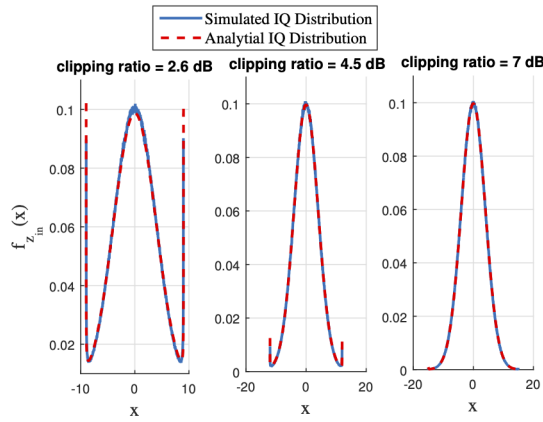


Figure 3. Simulated and analytical PDF of clipped PAPR-reduced IQ samples for different clipping ratios.

IV. SIMULATION RESULTS AND ANALYSIS

Through our simulations an 8-MHz DVB-T2 frame structure with 32K FFT mode is considered. We use 10^3 randomly generated OFDM symbols with 32768 subcarriers conveying 64-QAM modulation symbols.

A. Gaussian Quantizer Performance Evaluation

Our goal is to assess the robustness of Gaussian-optimized compression techniques when applied to a clipped signal. The results are given thanks to the SQNR in order to emphasize the effect of the quantization block.

Fig. 2 shows the simulated SQNR as a function of the clipping ratio Λ for different resolutions. The solid lines represent the optimal performance of the Gaussian quantizer applied on a Gaussian signal. The dashed lines represent the SQNR obtained when the Gaussian quantizer is applied on the clipped IQ samples. We first conclude that using a Gaussian quantizer on a clipped signal causes a severe degradation on SQNR. Furthermore, the performance loss increases as the resolution of the quantizer increases, i.e. as R increases. Finally and as expected from (16), the higher the clipping

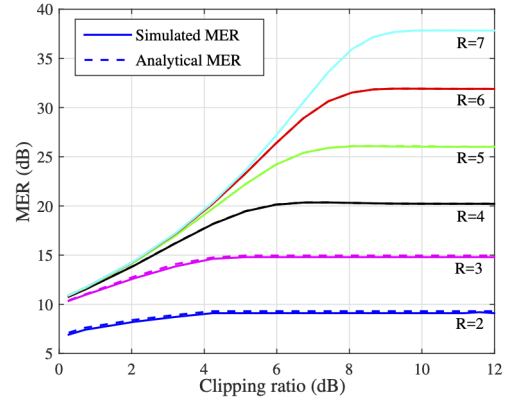


Figure 4. Simulated and analytical MER vs clipping ratio of optimized compression techniques for different resolutions.

threshold Λ , the smaller the performance loss. This is due to the convergence of the distribution of clipped IQ samples to a zero-mean Gaussian distribution when the clipping threshold is increased, as shown in Fig. 3 from left to right. Thus, the SQNR curves reach a constant ceiling at high clipping thresholds.

B. Validity of the MER Expression

In the first place, we validate the obtained clipped IQ samples distribution in Fig. 3 for different clipping ratios of 2.6 dB, 4.5 dB, and 7 dB respectively. The analytical curves perfectly match the simulated ones which proves the consistency and accuracy of the obtained IQ PDF.

Fig. 4 shows the analytical and simulated MER in function of the clipping ratio Λ for different resolutions. The analytical expression of MER perfectly matches the simulation results at high resolutions, i.e. $R \geq 4$, due to the accuracy of the asymptotic quantization distortion formula in (23) [17]. Moreover, the difference between the theoretical and simulated MER does not exceed 0.25 dB at low resolution, which validates the proposed MER expression.

C. Comparison of Gaussian and Optimized Quantizers

Fig. 5 shows MER results for both Gaussian and clipped optimized compression techniques, i.e. quantizer and EC, obtained from the proposed analytical study as a function of the clipping ratio Λ for different resolutions. First of all, MER increases when increasing the clipping ratio and at high clipping ratio only the quantization distortion affects the system performance. Second, the higher the clipping ratio, the closer the signal approaches the zero-mean Gaussian distribution. Thus, the two types of compression techniques merge to reach a ceiling, which corresponds to the MER obtained when only the quantization distortion of a Gaussian-distributed signal at certain rate is considered. This explains the 6-dB gap between different rates at high clipping ratios and verifies the famous 6-dB quantization rule [17].

Finally, we observe that using optimized compression techniques on a clipped PAPR-reduced signal produces an im-

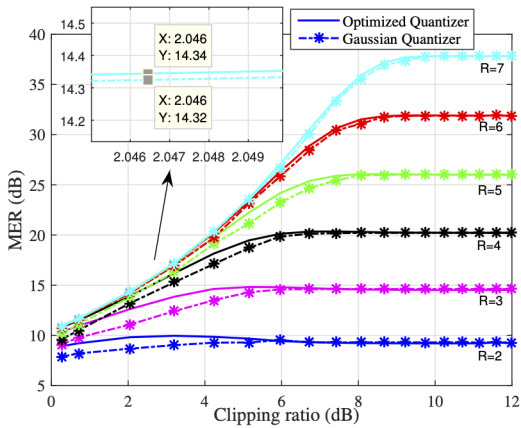


Figure 5. MER of Gaussian and optimized compression techniques vs clipping ratio for different resolutions.

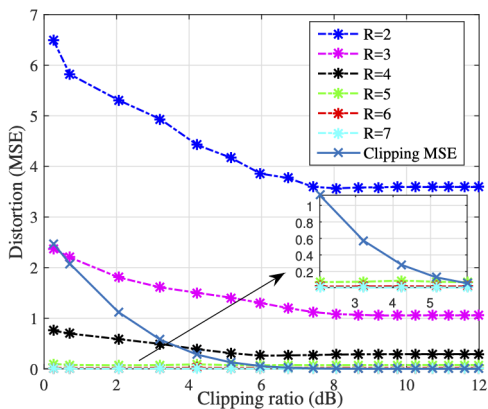


Figure 6. Clipping PAPR-reduction and Gaussian quantizer MSEs vs clipping ratio for different resolutions.

provement in MER performance at low resolutions and low clipping ratios, i.e. for a highly clipped signal. However, the improvement compared to purely Gaussian compression techniques decreases when resolution increases, reaching approximately about 0.02 dB when using on average 7 bits per IQ sample. This could be explained by exploring Fig. 6, which shows the clipping and Gaussian quantizer distortions separately for different resolutions. It can be observed that although the quantizer is not optimized for the clipped input signal, the quantization distortion is negligible compared to the clipping distortion at high rates. This is because large number of quantization levels in a finite quantization region effectively reduces the quantization distortion to almost 0.006 even if it is not an optimized quantizer. Thus, in both cases, clipping distortion becomes the only dominant fixed distortion compared to quantization distortion.

V. CONCLUSION

In this paper, we have studied the effect of reducing the PAPR of an OFDM signal using clipping before the compression in a C-RAN architecture. First, the distribution of the clipped IQ samples has been derived. This allowed to

optimize the codebook levels of a non-uniform quantizer and the average number of bits used in a Huffman encoder. Moreover, a closed-form expression for MER has been provided taking into account the joint effect of clipping and quantization operations. Simulation results emphasized the validity of our findings and showed that the Gaussian-optimized quantizer is not quite robust when applied to a clipped signal, especially in the low-rate quantization region. However, this gain has been shown to significantly decrease in the high-rate region.

ACKNOWLEDGEMENT

This work was financially supported by BPIfrance for the FUI25 CLOUDCAST project.

REFERENCES

- [1] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Communications surveys & tutorials*, vol. 17, no. 1, pp. 405–426, 2014.
- [2] A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things you should know about fronthaul," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1077–1083, 2015.
- [3] I. Chih-Lin, J. Huang, R. Duan, C. Cui, J. Jiang, and L. Li, "Recent progress on C-RAN centralization and cloudification," *IEEE Access*, vol. 2, pp. 1030–1039, 2014.
- [4] J. Lee, E. Hyun, and J.-Y. Jung, "A simple and efficient IQ data compression method based on latency, EVM, and compression ratio analysis," *IEEE Access*, vol. 7, pp. 117 436–117 447, 2019.
- [5] A. Vosoughi, M. Wu, and J. R. Cavallaro, "Baseband signal compression in wireless base stations," in *2012 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2012, pp. 4505–4511.
- [6] B. Guo, W. Cao, A. Tao, and D. Samarzija, "LTE/LTE-A signal compression on the CPRI interface," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 117–133, 2013.
- [7] H. Si, B. L. Ng, M. S. Rahman, and J. Zhang, "A novel and efficient vector quantization based CPRI compression algorithm," *IEEE Transactions on vehicular technology*, vol. 66, no. 8, pp. 7061–7071, 2017.
- [8] Y.-C. Wang and Z.-Q. Luo, "Optimized iterative clipping and filtering for PAPR reduction of OFDM signals," *IEEE Transactions on communications*, vol. 59, no. 1, pp. 33–37, 2010.
- [9] A. Cheaito, M. Crussière, J.-F. Hélar, and Y. Louët, "Quantifying the memory effects of power amplifiers: EVM closed-form derivations of multicarrier signals," *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 34–37, 2016.
- [10] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.
- [11] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.
- [12] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier, 2007.
- [13] A. Shehata, M. Crussière, and P. Mary, "Analysis of baseband IQ data compression methods for centralized RAN," in *2020 28th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1762–1766.
- [14] M. Bernhard, D. Rörich, T. Handte, and J. Speidel, "Analytical and numerical studies of quantization effects in coherent optical OFDM transmission with 100 gbit/s and beyond," *ITG-Fachtagung Photonische Netze*, pp. 34–40, 2012.
- [15] P. Panter and W. Dite, "Quantization distortion in pulse-count modulation with nonuniform spacing of levels," *Proceedings of the IRE*, vol. 39, no. 1, pp. 44–48, 1951.
- [16] C.-s. Liu, "The essence of the generalized newton binomial theorem," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 10, pp. 2766–2768, 2010.
- [17] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Springer Science & Business Media, 2012, vol. 159.