



Liu, X., Zhang, X., Zhang, L., Xiao, P., Wei, J., Zhang, H. and Leung, V. C.M. (2020) PAPR reduction using iterative clipping/filtering and ADMM approaches for OFDM-based mixed-numerology systems. *IEEE Transactions on Wireless Communications*, (doi:10.1109/TWC.2020.2966600).

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/207659/>

Deposited on: 13 January 2020

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

PAPR Reduction Using Iterative Clipping/Filtering and ADMM Approaches for OFDM-Based Mixed-Numerology Systems

Xiaoran Liu, Xiaoying Zhang, Lei Zhang, *Senior Member, IEEE*, Pei Xiao, *Senior Member, IEEE*, Jibo Wei, *Member, IEEE*, Haijun Zhang, *Senior Member, IEEE*, and Victor C. M. Leung, *Fellow, IEEE*

Abstract—Mixed-numerology transmission is proposed to support a variety of communication scenarios with diverse requirements. However, as the orthogonal frequency division multiplexing (OFDM) remains as the basic waveform, the peak-to-average power ratio (PAPR) problem is still cumbersome. In this paper, based on the iterative clipping and filtering (ICF) and optimization methods, we investigate the PAPR reduction in the mixed-numerology systems. We first illustrate that the direct extension of classical ICF brings about the accumulation of inter-numerology interference (INI) due to the repeated execution. By exploiting the clipping noise rather than the clipped signal, the noise-shaped ICF (NS-ICF) method is then proposed without increasing the INI. Next, we address the in-band distortion minimization problem subject to the PAPR constraint. By reformulation, the resulting model is separable in both the objective function and the constraints, and well suited for the alternating direction method of multipliers (ADMM) approach. The ADMM-based algorithms are then developed to split the original problem into several subproblems which can be easily solved with closed-form solutions. Furthermore, the applications of the proposed PAPR reduction methods combined with filtering and windowing techniques are also shown to be effective.

Index Terms—OFDM, mixed-numerology, PAPR reduction, clipping and filtering, optimization, ADMM.

I. INTRODUCTION

Future wireless communication systems will address unprecedented challenges to cope with a high degree of heterogeneity in terms of services, applications, use cases, deployment scenarios and mobility levels [1], [2]. In order to

This work was supported in part by the National Natural Science Foundation of China (61931020, 61822104, 61771044), the U.K. Engineering and Physical Sciences Research Council (EP/S02476X/1, EP/P03456X/1) and the Fundamental Research Funds for the Central Universities (RC1631, FRF-TP-19-002C1).

X. Liu, X. Zhang and J. Wei are with the Department of Electronic Science, National University of Defense Technology, Changsha 410073, China (e-mail: {liuxiaoran10, zhangxiaoying, wjbhw}@nudt.edu.cn).

L. Zhang is with the School of Engineering, University of Glasgow, Glasgow, G12 8QQ, U.K. (e-mail: lei.zhang@glasgow.ac.uk).

X. Pei is with the 5G Innovation Centre, and the Institute for Communication Systems, University of Surrey, Guildford GU2 7XH, U.K. (e-mail: p.xiao@surrey.ac.uk).

H. Zhang is with the Beijing Advanced Innovation Center for Materials Genome Engineering, Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing 100083, China (e-mail: haijunzhang@ieee.org).

V. C. M. Leung is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: vleung@ieee.org).

support multiform families of communication scenarios and applications, enhanced mobile broadband (eMBB), massive machine type communications (mMTC) and ultra reliable and low latency communications (URLLC) have been categorized as three major usage scenarios for the 5G cellular wireless communication. However, these scenarios with diverse technical requirements generally lead to different physical layer designs. For example, the subcarrier spacing in mMTC scenario is expected to be small in order to support a massive number of devices with limited spectrum resource [3]. URLLC, on the other hand, designed for communication between devices/machines with high reliability and low end-to-end delay, can find its applications in vehicular communication, factory automation, remote surgery, etc [4], and large subcarrier spacing (and accordingly, smaller symbol duration) is thus preferred for the stringent latency. Therefore, the traditional “one-fit-all” numerology design cannot simultaneously satisfy all these requirements [5].

The Third Generation Partnership Project (3GPP) has discussed the 5G new radio (NR) access technology that provides guidelines for the design of waveform and numerology [6]. Mixed-numerology is one of the important features proposed to efficiently support multiple services on the same carrier but with different subcarrier spacing of the orthogonal frequency division multiplexing (OFDM). Compared with multiplexing in the time domain [7], aligning different numerologies in frequency domain has better forward compatibility and inclusive support for different latency services [8]. Therefore, the frequency domain multiplexing that divides the system bandwidth into several subbands and assigns different numerologies to each subband has received widespread support [9].

However, multiplexing different OFDM numerologies in frequency domain gives rise to several new challenges in system design, such as inter-numerology interference [10] and numerology scheduling [11]. Additionally, high peak-to-average-power ratio (PAPR) inherited from multicarrier modulation is still a major drawback in OFDM-based mixed-numerology systems. In order to avoid the nonlinear distortion caused by imperfect power amplifier (PA) [12], the PAPR problem should be carefully addressed for mixed-numerology systems [13].

A. Related Works

For the conventional OFDM system, a variety of PAPR reduction techniques have been proposed over the past decades.

Those techniques can be roughly classified into three categories, i.e., multiple signaling probabilistic, coding techniques, and signal distortion [14]. The first category is able to generate several candidate permutations of the OFDM signal and choose the one with the lowest PAPR, whose representative techniques include selective mapping (SLM) [15], partial transmit sequence (PTS) [16], and tone reservation (TR) [17]. The coding techniques modifies its inherence to provide both the capability of error detection/correction and PAPR reduction [18].

Compared with other PAPR reduction techniques, the distortion based techniques can be implemented transparently to the existing standards, which is more attractive in current communication systems [19]. The simplest signal distortion technique may be the iterative clipping and filtering (ICF) method that reduces PAPR in an intuitive manner. The clipping procedure directly confines the amplitude of OFDM signal to a preset threshold and then the filtering procedure suppresses the consequent out-of-band emission (OOBE). Although the classical ICF method has relatively low computational complexity which mainly costs the fast Fourier transform (FFT) and inverse FFT (IFFT) operations, multiple repetitions are generally required to achieve the desired PAPR reduction [20]. Recently, with the exploitation of optimization method, the distortion based technique has been studied to pursue the optimal system performance in terms of PAPR and error vector magnitude (EVM). In [21], the PAPR minimization problem with the EVM constraint is formulated as a second order conic programming (SOCP) for the first time. Subsequently, the SOCP is extended to solve an EVM optimization task subject to a deterministic PAPR constraint and a spectral mask constraint [22], which is solved by using the interior-point method. In addition, the filtering procedure of ICF is formulated as an optimization problem to find the optimal filter coefficients in [23], which is called optimized ICF (OICF) method. This method is further considered for simplification in [24]. Recently, the alternating direction method of multipliers (ADMM), as a powerful first-order optimization technique, is widely used in distributed optimization and statistical learning [25]. As the ADMM is well suited to large-scale convex optimization, the authors in [26] firstly exploit the ADMM to efficiently solve the PAPR minimization in large-scale multiple-input multiple-output (MIMO) systems. In [27], the ADMM is introduced to the PAPR optimization problem with theoretical convergence result.

Considering the mixed-numerology transmissions, the existing PAPR reduction techniques cannot be directly applied due to the following reasons. First, since PAPR is measured for the composite signal before PA, representing the sum of all subbands signals, the traditional PAPR reduction technique performed on the separate subbands possibly will not level down the PAPR of the composite signal. Furthermore, unlike other multi-bands OFDM systems, such as noncontiguous-OFDM (NC-OFDM) [28] and carrier aggregation [29], in which the subcarriers are still orthogonal to each other and the signal can be modulated by one IFFT/FFT module similar to the conventional OFDM, the mixed-numerology systems are equipped with multiple scalable IFFT/FFT modules where the

subband signals are generated individually. How to operate on the separate subbands to achieve the overall PAPR reduction becomes an open question. To the best of the authors knowledge, the mixed-numerology systems are different and lead to the research and development of PAPR reduction techniques that are not currently available in the literature.

B. Contributions

In this paper, we consider the signal distortion techniques for PAPR reduction in the mixed-numerology systems. It will be shown that the direct extension of the classical ICF method to the mixed-numerology system results in INI accumulation. A new noise-shaped ICF (NS-ICF) method is designed to elaborately avoid the increase of INI. Meanwhile, by formulating the PAPR reduction of composite signals as a multi-block convex program, a simple reformulation of the constrained convex problem falls in the applicable scope of the ADMM. As emphasized in [25], the philosophy of ADMM is a “decomposition-coordination procedure”, in which the solutions to small local subproblems are coordinated to find a solution to the large global problem. Interestingly, the generating process of the composite signal is implicitly similar to the procedure of ADMM. The PAPR optimization performed on each subband will eventually achieve the global optimized signal. We show that the ADMM approach is efficient for the PAPR reduction of composite signal and leads to tractable and scalable algorithms for the cases of any number of numerologies.

The contributions of this paper are summarized as follows:

- We first build a system model and analyze the PAPR problem for the OFDM-based mixed-numerology system. As the composite signal is composed of multiple individual OFDM signals, we illustrate that the high PAPR problem is still troublesome. It is also shown that independently applying PAPR reduction technique to separate subbands is incapable of reducing the PAPR of composite signal.
- In order to avoid the INI accumulation caused by repeated execution, we propose a new clipping-based method, named NS-ICF, as the benchmark for PAPR reduction, which exploits the clipping noise instead of the clipped signal. The analytical expression is provided to prove that the NS-ICF method does not introduce extra INI.
- We then pursue the optimal performance with respect to the EVM and PAPR. A distortion minimization problem with a constraint of PAPR is then formulated as a convex program. Reformulated as a favorable separable structure emerging in both the objective function and the constraints, the original problem can be efficiently solved through ADMM approach. The proposed ADMM-based algorithms split the original problem into several subproblems whose optimal solution can be determined analytically. Moreover, the computational complexity in each iteration is comparable to that of the NS-ICF method.
- The proposed PAPR reduction methods are also considered in combination with the filtering and windowing

techniques. We show that the idea of NS-ICF can be easily used in filtered-OFDM (F-OFDM) in which the clipping noise is processed by the identical subband filter in time domain, and the windowing operation is included into the optimization formulation that can be solved by the proposed ADMM-based algorithms.

C. Organization and Notations

The rest of this paper is organized as follows. Section II establishes the system model for mixed numerologies transmission and briefly reviews the PAPR problem. In Section III, we redesign the classical ICF method and apply it to the mixed-numerology system. Then, the optimization model of PAPR reduction is formulated and two algorithms are proposed in Section IV. The applications of proposed PAPR methods with the filtering and windowing techniques are discussed in Section V. Section VI presents the simulation results and discussion. The conclusions are drawn in Section VII.

Notations: In this paper, vectors and matrices are denoted by lowercase and uppercase bold letters, and $(\cdot)^H$ symbolize the Hermitian conjugate operation. We use $\|\cdot\|_\infty$ and $\|\cdot\|_2$ to denote the ∞ -norm and 2-norm of a vector, respectively. \mathbf{I}_N and $\mathbf{O}_{m \times n}$ refer to $N \times N$ identity matrix and $m \times n$ zero matrix, respectively. $\text{blkdiag}(\mathbf{A}, N)$ denotes a diagonal matrix generated by N repetitions of \mathbf{A} . ∇ stands for the gradient operator.

II. PRELIMINARIES

We start by introducing the system model and then summarize the fundamental PAPR issue arising in the OFDM-based mixed-numerology systems.

A. System Model of Mixed-Numerology Transmission

We consider a mixed-numerology OFDM system where the system bandwidth is divided into several subbands with different subcarrier spacings. For a subband with K_i continuous subcarriers using numerology i , the continuous-time OFDM symbol can be obtained by

$$s_i^u(t) = \frac{1}{\sqrt{K_i}} \sum_{k=0}^{K_i-1} x_i^u(k) e^{j2\pi(kf_i + \Delta F_i)(t - T_{CP,i} - (u-1)T_i)}, \quad uT_i \leq t < (u+1)T_i, \quad (1)$$

where $x_i^u(k)$ is the independent identically distributed complex modulation symbol with zero mean and unit power on the k -th subcarrier of the u -th OFDM symbol. f_i denotes the subcarrier spacing of numerology i and ΔF_i is frequency offset of the i -th subband. The signal duration is $T_i = T_{CP,i} + T_{sy,s,i}$, where $T_{CP,i}$ denotes the duration of cyclic prefix (CP) and $T_{sy,s,i} = 1/f_i$.

There is a commonly accepted family of numerologies that defines the subcarrier spacings as follows [2], [9]

$$f_i = 2^{v_i} f_1, \quad T_{CP,i} = T_{CP,1}/2^{v_i}, \quad i = 1, 2, \dots, M, \quad (2)$$

where v_i is an integer and M is the number of OFDM numerologies. Equation (2) implies that any subcarrier spacing

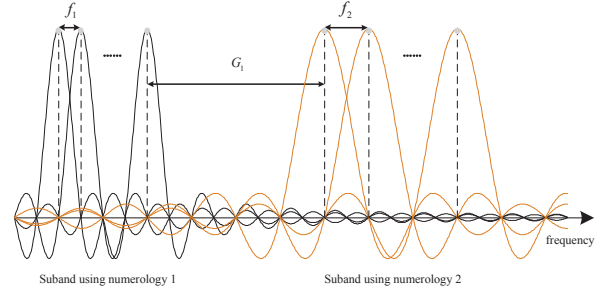


Fig. 1. An example of mixed numerologies transmission with two subbands.

is an integer divisible by all smaller subcarrier spacing. By this mean, the waveform is scalable in the sense that the subcarrier spacing of OFDM can be chosen according to (2). For example, in the current version of 5G NR [30], the subcarrier spacing of OFDM can be chosen according to 15×2^{v_i} kHz where 15 kHz is the subcarrier spacing used in Long Term Evolution (LTE). With this scaling approach, different OFDM numerologies can be implemented easily by using different-scaled FFT under the same sampling clock rate, which is also illustrated as single rate system [1].

Without loss generality, we assume that $\Delta F_1 = 0$ and the guard band between i -th and $(i+1)$ -th subbands is G_i . Then, the system bandwidth can be roughly expressed as

$$B = \sum_{i=1}^{M-1} (K_i f_i + G_i) + K_M f_M. \quad (3)$$

We also use the concept of generalized synchronized system whose symbol period is equivalent to the least common multiple (LCM) of duration of all subbands [1]. It benefits simplified system design since only limited symbols need to be considered in a processing window, and every LCM symbol has the same overall performance [1]. Specifically, the signal duration has the relationship of $T_1 = 2^{v_2} T_2 = \dots = 2^{v_M} T_M$. For example, if $M = 2$, one LCM symbol is compounded of one OFDM symbol of numerology 1 and two OFDM symbols of numerology 2, i.e.,

$$z^u(t) = \begin{cases} \eta_1 s_1^u(t) + \eta_2 s_2^{2u-1}(t), & uT_1 \leq t < uT_1 + T_2 \\ \eta_1 s_1^u(t) + \eta_2 s_2^{2u}(t - T_1/2), & uT_1 + T_2 \leq t < (u+1)T_1 \end{cases} \quad (4)$$

where $z^u(t)$ is the u -th LCM symbol and η_i denotes the power adjusting factor for the i -th subband. Note that if the total transmit power is normalized to M and evenly distributed on each subband, then we have $\eta_i = 1$. Without loss of generality, we focus on the time interval $[0, T_1)$ and omit the superscript of $z(t)$.

B. PAPR Problem

According to the definition of PAPR which is the ratio of the peak power of the signal to its average power, we can express the PAPR of the composite signal $z(t)$ as

$$\text{PAPR} = \frac{\max_{t \in [0, T_1)} |z(t)|^2}{P_{av}}, \quad (5)$$

where $P_{av} = E\{|z(t)|^2\}$ is the average power of $z(t)$. Nevertheless, analog signals are not amiable for calculation and analysis. To address this issue, the discrete samples of $z(t)$ are usually used to compute the PAPR. To approximate the PAPR of the continuous-time signal accurately, J -time oversampling is usually considered [31]. In addition, for the convenience of FFT implementation and the uniform of sampling rate, the JN_i -point IFFT are commonly used to modulate the oversampled OFDM signal, where $N_i = 2^{\lceil \log_2(B/f_i) \rceil}$. Then, the oversampled signal $s_i^{u_i}(n)$ can be efficiently computed by

$$s_i^{u_i}(n) = \frac{1}{\sqrt{JN_i}} \sum_{k=0}^{K_i-1} x_i^{u_i}(k) e^{j \frac{2\pi(n-L_{CP,i})(k+\Delta k_i)}{JN_i}}, \quad (6)$$

$$0 \leq n \leq L_{sys} - 1,$$

where $\Delta k_i = \Delta F_i/f_i$, $L_{sys} = JN_i + L_{CP,i}$ and $u_i = 1, 2, \dots, 2^{v_i}$. In order to facilitate the algebraic operation, we use the matrix representations instead. The oversampled signal (6) can also be expressed as

$$\mathbf{s}_i^{u_i} = \mathbf{P}_i \mathbf{D}_i \mathbf{x}_i^{u_i}, \quad (7)$$

where $\mathbf{x}_i^{u_i} = [x_i^{u_i}(0), x_i^{u_i}(1), \dots, x_i^{u_i}(K_i - 1)]^T$. $\mathbf{D}_i \in \mathbb{C}^{JN_i \times K_i}$ is the first K_i columns of the JN_i -points normalized and frequency-shifted IFFT matrix which is shifted by Δk_i , and $\mathbf{P}_i = [\mathbf{0}_{L_{CP,i} \times (JN_i - L_{CP,i})}, \mathbf{I}_{L_{CP,i}}; \mathbf{I}_{JN_i}]$ denotes the matrix of CP insertion.

Then, for an LCM symbol compounded of 2^{v_i} symbols of numerology i (for $i = 1, 2, \dots, M$), we denote the composite symbol as

$$\mathbf{z} = \mathbf{F}_1 \mathbf{x}_1 + \mathbf{F}_2 \mathbf{x}_2 + \dots + \mathbf{F}_M \mathbf{x}_M, \quad (8)$$

where $\mathbf{F}_i = \text{blkdiag}(\eta_i \mathbf{P}_i \mathbf{D}_i, 2^{v_i})$ and $\mathbf{x}_i = [\mathbf{x}_i^1; \mathbf{x}_i^2; \dots; \mathbf{x}_i^{2^{v_i}}]$ for $i = 1, 2, \dots, M$.

Thus, the PAPR of composite signal $z(n)$ is given by

$$\text{PAPR} = \frac{\max_{n=0,1,\dots,L_{sys}} |z(n)|^2}{\frac{1}{L_{sys}} \sum_{n=0}^{L_{sys}-1} |z(n)|^2} = \frac{\|\mathbf{z}\|_\infty^2}{\frac{1}{L_{sys}} \|\mathbf{z}\|_2^2}. \quad (9)$$

Note that the envelope of the composite signal \mathbf{z} is generated from the superposition of all the signals from each subband. We can infer that the composite signal sampled at the Nyquist rate still converges to a complex Gaussian random process for large N_i . It can be explained by the fact that the sum of Gaussian distributed random variables obeys Gaussian distribution. Therefore, the envelope of composite signal becomes a Rayleigh random variable [31], which implies that the high PAPR problem still exists in the mixed-numerology systems and the PAPR reduction techniques should be considered for the composite signal rather than the individual subband signal.

In the following sections, to clearly illustrate our ideas and without loss generality, we only consider the mixed-numerology transmission with $M = 2$ numerologies as shown in Fig. 1. However, the proposed PAPR reduction methods can be straightforwardly extended to the $M > 2$ cases.

III. ITERATIVE CLIPPING AND FILTERING METHOD IN MIXED-NUMEROLOGY SYSTEM

The main idea of ICF is to hard limit the amplitude of signals beyond the clipping threshold and subsequently eliminate the clipping noise outside the band. To characterize the fundamental properties, we start from reviewing the classical ICF method in conventional OFDM system. Then we present the different characteristics of ICF in the mixed-numerology system in full detail and propose a new clipping-based method as a benchmark for PAPR reduction techniques.

A. ICF method

The clipping and filtering operation are iteratively performed as described in [20]. To be specific, the clipping procedure is performed by

$$\bar{s}(n) = \begin{cases} Ae^{j\theta(n)}, & |s(n)| > A \\ s(n), & |s(n)| \leq A \end{cases} \quad (10)$$

where $\bar{s}(n)$ is the clipped signal and $\theta(n)$ denotes the phase of $s(n)$. A is the clipping level calculated in each iteration according to the predefined clipping ratio (CR), which is defined as [14]

$$\text{CR} = 20 \log_{10} \left(\frac{A}{\frac{1}{\sqrt{JN}} \|\mathbf{s}\|_2} \right) \text{ (dB)}. \quad (11)$$

We can see that CR is calculated by the preset PAPR threshold. Note that clipping is a nonlinear process that leads to both in-band distortion and OOB. In order to eliminate the OOB, the filtering approach is applied to the clipped signal in frequency domain. The filter design is based on a rectangular window of which the frequency response is given by [20]

$$H(k) = \begin{cases} 1, & 1 \leq k \leq N \\ 0, & N + 1 \leq k \leq JN \end{cases}. \quad (12)$$

What should be further illustrated is that the clipping noise is defined as the difference between the clipped and original signals, which is expressed as [32]

$$d(n) = \bar{s}(n) - s(n) = \begin{cases} (A - |s(n)|) e^{j\phi(n)}, & |s(n)| > A \\ 0, & |s(n)| \leq A \end{cases}. \quad (13)$$

Additionally, the power spectrum density (PSD) of the clipping noise is evenly distributed over the system band [33] where the in-band power distorts the transmitted symbol and the out-of-band power leads to low spectrum efficiency. Although the filtering approach can eliminate the OOB, the peak regrowth would occur as a side effect. Therefore, repeated clipping and filtering are always required to achieve the desired PAPR.

B. ICF method for Mixed-Numerology Transmitted Signal

As shown in Fig. 2(a), we can readily extend the ICF method to the mixed-numerology system by clipping the composite signal \mathbf{z} and filtering the clipped version in individual subbands, where $\bar{\mathbf{s}}_i$ is obtained by removing the beginning $L_{CP,i}$ samples of $\bar{\mathbf{z}}$ with the matrix $\mathbf{C}_i = [\mathbf{0}_{JN_i \times L_{CP,i}}, \mathbf{I}_{JN_i}]$.

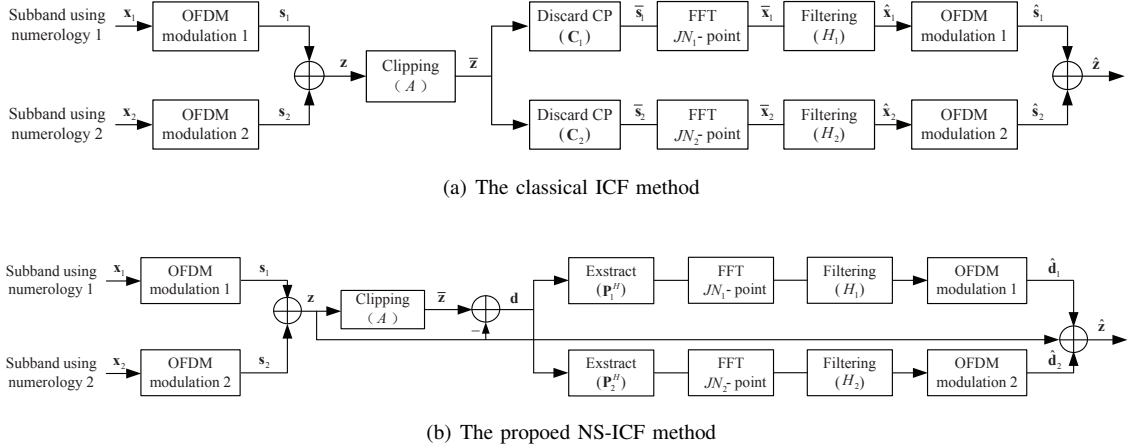


Fig. 2. The diagrams of the ICF and NS-ICF methods.

Then, by using JN_i -point FFT to transform \bar{s}_i to the frequency domain, we have the clipped symbol \bar{x}_i . The following filtering $\mathbf{H}_i = \text{diag}(H_i(1), \dots, H_i(JN_i))$ is applied to \bar{x}_i , from which the components located in each subband are extracted and the out-of-band components are set to zero. After frequency filtering, \hat{x}_1 and \hat{x}_2 are modulated to OFDM symbols that are then added together to be transmitted or put into the next execution.

However, the output symbol of the filter contains not only the distorted original symbol but the interference from the other subband. For example, \hat{x}_1 can be expressed as

$$\begin{aligned} \hat{x}_1 &= \mathbf{D}_1^H \bar{s}_1 = \mathbf{D}_1^H \mathbf{C}_1 (\mathbf{z} + \mathbf{d}) \\ &= \mathbf{x}_1 + \underbrace{\mathbf{D}_1^H \mathbf{C}_1 \mathbf{F}_2 \mathbf{x}_2}_{\text{INI}} + \underbrace{\mathbf{D}_1^H \mathbf{C}_1 \mathbf{d}}_{\text{in-band distortion}}, \end{aligned} \quad (14)$$

where the FFT operation is represented by \mathbf{D}_1^H . As \mathbf{D}_1 only contains K_1 columns of normalized IFFT matrix to modulate the used subcarriers, filtering operation \mathbf{H}_1 can be omitted in (14). We can observe that the original symbol is distorted by the in-band clipping noise which is denoted as the third term in the left of the equation from (14). The second term is the INI introduced by the subband using numerology 2. Since the orthogonality is no longer held for the waveforms using different numerologies, the leakage energy will be caught by the other subband. The output composite signal is then given by

$$\begin{aligned} \hat{\mathbf{z}} &= \mathbf{F}_1 \hat{x}_1 + \mathbf{F}_2 \hat{x}_2 \\ &= \mathbf{z} + \underbrace{\mathbf{F}_1 \mathbf{D}_1^H \mathbf{C}_1 \mathbf{F}_2 \mathbf{x}_2 + \mathbf{F}_2 \mathbf{D}_2^H \mathbf{C}_2 \mathbf{F}_1 \mathbf{x}_1}_{\text{INI}} \\ &\quad + \underbrace{(\mathbf{F}_1 \mathbf{D}_1^H \mathbf{C}_1 + \mathbf{F}_2 \mathbf{D}_2^H \mathbf{C}_2) \mathbf{d}}_{\text{in-band distortion}}. \end{aligned} \quad (15)$$

Although the INI can be reduced by broadening the guard band, the repetition progress will inevitably accumulate the INI that remains in the output signal at each execution. The interference level is increasingly reinforced as the ICF method is repeatedly executed. In addition, the high amplitudes at the beginning $L_{CP,i}$ samples of the signal period is not considered

after the CP removal, so that high peaks are probably still in existence. Therefore, the straightforward application of the classical ICF method to the mixed-numerology systems seems infeasible.

In order to avoid INI and inherit the idea of the ICF method, we then focus on the clipping noise. As shown in Fig. 2(b), the clipping noise rather than the clipped signal are used for frequency-domain filtering in the individual subbands. Moreover, the beginning $L_{CP,i}$ samples of the clipping noise are added to the end of the signal using matrix \mathbf{P}_i^H .

The filtered clipping noise is then written as

$$\hat{\mathbf{d}}_i = \mathbf{F}_i \mathbf{D}_i^H \mathbf{P}_i^H \mathbf{d} = \mathbf{F}_i \mathbf{F}_i^H \mathbf{d}, \quad i = 1, 2. \quad (16)$$

By this mean, the output composite signal can be obtained by

$$\hat{\mathbf{z}} = \mathbf{z} + \hat{\mathbf{d}}_1 + \hat{\mathbf{d}}_2 = \mathbf{z} + (\mathbf{F}_1 \mathbf{F}_1^H + \mathbf{F}_2 \mathbf{F}_2^H) \mathbf{d}. \quad (17)$$

Obviously, the terms of INI no longer appear in (17) compared with (15). On account that the filtering is applied to the clipping noise, we call it the noise-shaped ICF method. Interestingly, while the NS-ICF method is equivalent to the classical ICF method in the conventional single numerology cases, the classical ICF method has to be redesigned as the NS-ICF method for the mixed-numerology systems.

Note that the NS-ICF method is very simple and straightforward and can serve as a benchmark for the distortion based PAPR reduction techniques. Due to the peak regrowth, it generally requires many times of execution to achieve the desired PAPR. In addition, the distortion caused by amplitude clipping is not considered. In the next section, the optimization method is proposed to achieve both the PAPR reduction and distortion control.

IV. PAPR REDUCTION USING OPTIMIZATION METHOD

In this section, we first look back the OICF method and then establish the optimization problem for the PAPR reduction of composite signal. As a common approach to solve optimization problems in the form of SOCP, interior-point method has been applied for the PAPR reduction problem [22], [23].

However, prohibitively high computation complexity is often required for the system with large number of subcarrier. Thus, algorithm efficiency is of extreme importance for practical application. By reformulating the optimization problem as a linearly constrained separable convex programming, whose objective function is separable into multiple individual convex functions without coupled variables, we show that the resulting model falls into the applicable scope of the well-known ADMM approach and can be efficiently solved by splitting it into some more tractable subproblems.

A. OICF method

The OICF method aims to minimize the in-band distortion under the PAPR constraint. The symbol-wise distortion can be quantified by the EVM, which is defined as the square root of the ratio of the mean error vector power to the mean reference power, i.e.,

$$\text{EVM}_{\mathbf{x}} = \sqrt{\frac{\frac{1}{N} \sum_{k=0}^{N-1} |x(k) - \hat{x}(k)|^2}{\frac{1}{N} \sum_{k=0}^{N-1} |x(k)|^2}} = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}, \quad (18)$$

where \mathbf{x} is the original OFDM symbol and $\hat{\mathbf{x}}$ denotes the modified symbol generated by PAPR reduction technique. Then the OICF method can be formulated as [23]

$$\min_{\hat{\mathbf{x}} \in \mathbb{C}^N} \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}, \quad (19a)$$

$$s.t. \quad \hat{\mathbf{s}} = \text{IFFT}(\hat{\mathbf{x}}), \quad (19b)$$

$$\frac{\|\hat{\mathbf{s}}\|_\infty}{\|\hat{\mathbf{s}}\|_2 / \sqrt{L_{\text{sys}}}} \leq \sqrt{\text{CR}} = \gamma. \quad (19c)$$

Note that the problem (19) considers the optimized OFDM symbol instead of filter coefficients as the optimization variable, which is an equivalent form to determine the optimal filter design [24]. To relax the non-convex constraint (19c) to its convex counterpart that makes the problem solvable, $\|\hat{\mathbf{s}}\|$ is replaced with the original version $\|\mathbf{s}\|$. Then, the constraint (19c) is rewritten as

$$\|\hat{\mathbf{s}}\|_\infty \leq \gamma \|\mathbf{s}\|_2 / \sqrt{L_{\text{sys}}}. \quad (20)$$

With this modification, the problem (19) is formulated as an SOCP, to which many well-known optimization algorithms such as interior-point method can be applied.

B. Optimization Formulation

We now consider the distortion in an LCM symbol. Referring to the conception of EVM (18), we define the EVM of a composite signal with M different numerologies as the square root of total normalized distortion power on every sub-symbol, which is

$$\text{EVM}_{\mathbf{z}} = \sqrt{\sum_{i=1}^M \frac{1}{2^{v_i}} \left(\sum_{v=1}^{2^{v_i}} \text{EVM}_{\mathbf{x}_i^v}^2 \right)}. \quad (21)$$

In the case that only the first and second numerologies are used, the EVM of the composite signal \mathbf{z} can be written as

$$\text{EVM}_{\mathbf{z}} = \sqrt{\frac{\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2}{\|\mathbf{x}_1\|_2^2} + \frac{\|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2}{\|\mathbf{x}_2\|_2^2}}. \quad (22)$$

To minimize the distortion of an LCM symbol and constrain the PAPR, we formulate the following optimization problem

$$\min_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2} \sqrt{\frac{\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2}{\|\mathbf{x}_1\|_2^2} + \frac{\|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2}{\|\mathbf{x}_2\|_2^2}}, \quad (23a)$$

$$s.t. \quad \frac{\|\mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2\|_\infty}{\frac{1}{\sqrt{L_{\text{sys}}}} \|\mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2\|_2} \leq \gamma, \quad (23b)$$

where $\hat{\mathbf{x}}_1 \in \mathbb{C}^{K_1}$, $\hat{\mathbf{x}}_2 \in \mathbb{C}^{K_1}$. Similar to (19), the constraint (23b) is also a non-convex inequity that makes numerical solution of (23) difficult. Following the same procedure as mentioned in (20), $\frac{1}{\sqrt{JN_1}} \|\mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2\|_2$ can be approximated by $\frac{1}{\sqrt{JN_1}} \|\mathbf{z}\|_2$. Then the problem (23) is rewritten in the following form

$$\min_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2} \sqrt{\frac{\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2}{\|\mathbf{x}_1\|_2^2} + \frac{\|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2}{\|\mathbf{x}_2\|_2^2}}, \quad (24a)$$

$$s.t. \quad \|\mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2\|_\infty \leq \gamma \frac{1}{\sqrt{L_{\text{sys}}}} \|\mathbf{z}\|_2, \quad (24b)$$

which is also an SOCP that can be solved by resorting to some open source software like CVX [34].

In order to find more efficient algorithm, we continue with this optimization problem (24). Since the problem $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$ has the same optimal solution as $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ [35], we can square the objective function and then obtain a sum of two separable convex functions. Moreover, an auxiliary variable $\hat{\mathbf{z}} \in \mathbb{C}^{L_{\text{sys}}}$ can be introduced to the optimization problem, which has the linear relationship with $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$, i.e., $\hat{\mathbf{z}} = \mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2$. Taking these aspects into consideration, the original problem can be then reformulated as

$$\min_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{z}}} \frac{1}{2\sigma_1^2} \|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2 + \frac{1}{2\sigma_2^2} \|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2 \quad (25a)$$

$$s.t. \quad \hat{\mathbf{z}} = \mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2, \quad (25b)$$

$$\|\hat{\mathbf{z}}\|_\infty \leq \gamma \frac{1}{\sqrt{L_{\text{sys}}}} \|\mathbf{z}\|_2, \quad (25c)$$

where $\sigma_i^2 = \|\mathbf{x}_i\|_2^2$ for $i = 1, 2$. Note that the optimization objective (25a) is the sum of two separate convex functions and the variables obey a linear constraint (25b). This separable structure in both the objective function and constraints in (25) enables us to partition the original problem in the light of the ADMM. By this mean, the minimization problem of (25a) can be decomposed into three smaller ones which solve the variables $\hat{\mathbf{x}}_1$, $\hat{\mathbf{x}}_2$ and $\hat{\mathbf{z}}$ separately in the consecutive order.

C. O-ADMM Algorithm Framework

The basic idea of the ADMM is to search for a saddle point of the augmented Lagrangian function rather than directly solve the original constrained optimization problem. We first give the augmented Lagrangian function of the problem (25)

$$L_\rho(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{z}}, \mathbf{y}) = \frac{1}{2\sigma_1^2} \|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2 + \frac{1}{2\sigma_2^2} \|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2 \\ + \langle \mathbf{y}, \mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2 - \hat{\mathbf{z}} \rangle + \frac{\rho}{2} \|\mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2 - \hat{\mathbf{z}}\|_2^2, \quad (26)$$

where $\rho > 0$ is the penalty parameter, $\mathbf{y} \in \mathbb{C}^{L_{\text{sys}}}$ is the Lagrangian multiplier, and the inner product is defined as $\langle \mathbf{a}, \mathbf{b} \rangle \doteq \text{Re} \{ \mathbf{a}^H \mathbf{b} \}$. The proposed O-ADMM algorithm carries out the following iterative steps,

$$\hat{\mathbf{x}}_1^{l+1} = \arg \min_{\mathbf{x}_1} L_\rho(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2^l, \hat{\mathbf{z}}^l, \mathbf{y}^l), \quad (27a)$$

$$\hat{\mathbf{x}}_2^{l+1} = \arg \min_{\hat{\mathbf{x}}_2} L_\rho(\hat{\mathbf{x}}_1^{l+1}, \hat{\mathbf{x}}_2, \hat{\mathbf{z}}^l, \mathbf{y}^l), \quad (27b)$$

$$\hat{\mathbf{z}}^{l+1} = \arg \min_{\hat{\mathbf{z}}} L_\rho(\hat{\mathbf{x}}_1^{l+1}, \hat{\mathbf{x}}_2^{l+1}, \hat{\mathbf{z}}, \mathbf{y}^l), \quad (27c)$$

$$\mathbf{y}^{l+1} = \mathbf{y}^l + \rho (\mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} + \mathbf{F}_2 \hat{\mathbf{x}}_2^{l+1} - \hat{\mathbf{z}}^{l+1}), \quad (27d)$$

where l is the iteration number. In each iteration, the O-ADMM algorithm consists of $\hat{\mathbf{x}}_1$ -minimization step, $\hat{\mathbf{x}}_2$ -minimization step, $\hat{\mathbf{z}}$ -minimization step, and dual variable update. Although each step of (27a), (27b) and (27c) needs to solve an optimization problem, we will show that each has a simple closed-form solution and can be solved efficiently.

Based on the augmented Lagrangian function $L_\rho(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{z}}, \mathbf{y})$, the problem (27a) becomes equivalent to

$$\min_{\hat{\mathbf{x}}_1} \frac{1}{2\sigma_1^2} \|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2 + \frac{\rho}{2} \left\| \mathbf{F}_1 \hat{\mathbf{x}}_1 + \mathbf{F}_2 \hat{\mathbf{x}}_2^l - \hat{\mathbf{z}}^l + \frac{\mathbf{y}^l}{\rho} \right\|_2^2. \quad (28)$$

By setting the derivative of the augmented Lagrangian $L_\rho(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2^l, \hat{\mathbf{z}}^l, \mathbf{y}^l)$ to zero, we can obtain the closed-form solution of $\hat{\mathbf{x}}_1$ -minimization as

$$\hat{\mathbf{x}}_1^{l+1} = \left(\frac{1}{\sigma_1^2} \mathbf{I}_{K_1} + \eta_1^2 \rho \mathbf{F}_1^H \mathbf{F}_1 \right)^{-1} \left(\frac{1}{\sigma_1^2} \mathbf{x}_1 - \mathbf{v}_1^l \right), \quad (29)$$

where

$$\mathbf{v}_1^l = \rho \mathbf{F}_1^H \left(\mathbf{F}_2 \hat{\mathbf{x}}_2^l - \hat{\mathbf{z}}^l + \frac{\mathbf{y}^l}{\rho} \right). \quad (30)$$

Similarly, the $\hat{\mathbf{x}}_2$ -minimization step can be solved by

$$\hat{\mathbf{x}}_2^{l+1} = \left(\frac{1}{\sigma_2^2} \mathbf{I}_{K_1} + \eta_2^2 \rho \mathbf{F}_2^H \mathbf{F}_2 \right)^{-1} \left(\frac{1}{\sigma_2^2} \mathbf{x}_2 - \mathbf{v}_2^l \right), \quad (31)$$

where

$$\mathbf{v}_2^l = \rho \mathbf{F}_2^H \left(\mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} - \hat{\mathbf{z}}^l + \frac{\mathbf{y}^l}{\rho} \right). \quad (32)$$

For the $\hat{\mathbf{z}}$ -minimization step, we first denote

$$\mathbf{u}^l = \mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} + \mathbf{F}_2 \hat{\mathbf{x}}_2^{l+1} + \frac{1}{\rho} \mathbf{y}^l. \quad (33)$$

Then, the subproblem (27d) is equivalent to

$$\min_{\hat{\mathbf{z}} \in \mathbb{C}^{JN_1}} \|\mathbf{u}^l - \hat{\mathbf{z}}\|_2^2, \quad (34a)$$

$$\text{s.t.} \quad \|\hat{\mathbf{z}}\|_\infty \leq \gamma \frac{1}{\sqrt{L_{\text{sys}}}} \|\mathbf{z}\|_2. \quad (34b)$$

Clearly, the problem (34) searches for a vector $\hat{\mathbf{z}}$ that is nearest to the vector \mathbf{u} and lies within the PAPR threshold, which can be split into L_{sys} one-dimension subproblems as

$$\min_{\hat{z}^{l+1}(n) \in \mathbb{C}} |u^l(n) - \hat{z}(n)|^2, \quad (35a)$$

$$\text{s.t.} \quad \hat{z}(n) \leq \gamma \frac{1}{\sqrt{L_{\text{sys}}}} |z(n)|_2. \quad (35b)$$

Algorithm 1 O-ADMM algorithm

Initialization:

Set $k = 0$ and initialize $(\hat{\mathbf{x}}_1^0, \hat{\mathbf{x}}_2^0, \hat{\mathbf{z}}^0, \hat{\mathbf{y}}^0)$.

Select the parameters (ρ, γ) .

Calculate $A = \gamma \|\mathbf{z}\|_2 / \sqrt{L_{\text{sys}}}$.

Runtime:

1: **Repeat** l

2: Process the $\hat{\mathbf{x}}_1$ -minimization step.

(1) Calculate $\mathbf{v}_1^l = \rho \mathbf{F}_1^H (\mathbf{F}_2 \hat{\mathbf{x}}_2^l - \hat{\mathbf{z}}^l + \frac{\mathbf{y}^l}{\rho})$.

(2) Calculate

$$\hat{\mathbf{x}}_1^{l+1} = \left(\frac{1}{\sigma_1^2} \mathbf{I}_{K_1} + \eta_1^2 \rho \mathbf{F}_1^H \mathbf{F}_1 \right)^{-1} \left(\frac{1}{\sigma_1^2} \mathbf{x}_1 - \mathbf{v}_1^l \right).$$

3: Process the $\hat{\mathbf{x}}_2$ -minimization step.

(1) Calculate $\mathbf{v}_2^l = \rho \mathbf{F}_2^H (\mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} - \hat{\mathbf{z}}^l + \frac{\mathbf{y}^l}{\rho})$.

(2) Calculate

$$\hat{\mathbf{x}}_2^{l+1} = \left(\frac{1}{\sigma_2^2} \mathbf{I}_{K_1} + \eta_2^2 \rho \mathbf{F}_2^H \mathbf{F}_2 \right)^{-1} \left(\frac{1}{\sigma_2^2} \mathbf{x}_2 - \mathbf{v}_2^l \right).$$

4: Process the $\hat{\mathbf{z}}$ -minimization step.

(1) Calculate $\mathbf{u}^l = \mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} + \mathbf{F}_2 \hat{\mathbf{x}}_2^{l+1} + \frac{1}{\rho} \mathbf{y}^l$.

(2) Calculate

$$\hat{z}^{l+1}(n) = \begin{cases} Ae^{j\angle u^l(n)}, & |u^l(n)| > A \\ u^l(n), & |u^l(n)| \leq A \end{cases}.$$

5: Update the Lagrangian dual variable $\hat{\mathbf{y}}$.

Calculate $\mathbf{y}^{l+1} = \mathbf{y}^l + \rho (\mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} + \mathbf{F}_2 \hat{\mathbf{x}}_2^{l+1} - \hat{\mathbf{z}}^{l+1})$.

6: **until** reach the stop-criterion. Then output $\hat{\mathbf{z}}^l$.

We can find that the optimal solution of (35) is a clipping procedure of vector \mathbf{u}^l . Hence the $\hat{\mathbf{z}}$ -minimization step can be easily obtained by

$$\hat{z}^{l+1}(n) = \begin{cases} Ae^{j\angle u^l(n)}, & |u^l(n)| > A \\ u^l(n), & |u^l(n)| \leq A \end{cases}, \quad (36)$$

where $A = \gamma \|\mathbf{z}\|_2 / \sqrt{L_{\text{sys}}}$. The proposed ADMM-based algorithm is detailed in Algorithm 1 and referred to as the origin-ADMM algorithm.

D. CU-ADMM Algorithm Framework

Due to the convex approximation of the constraint (23b), the O-ADMM algorithm is probably unable to reach the PAPR threshold after executing once. Several repetitions are generally needed to make the constraint (24b) approximate to (23b). In this subsection, inspired by the projection operation in the $\hat{\mathbf{z}}$ -minimization step, we propose an improved algorithm in which the constraint of (34b) is updated in each iteration and the rest of steps in O-ADMM are reused.

Instead of using the original composite signal \mathbf{z} to calculate the clipping level A in (36), we exploit the $\hat{\mathbf{z}}^l$ to update A^l before the $\hat{\mathbf{z}}$ -minimization step, which is written by

$$A^{l+1} = \gamma \frac{1}{\sqrt{L_{\text{sys}}}} \|\hat{\mathbf{z}}^l\|_2. \quad (37)$$

With this modification, the constraint (25c) is set according to the instantaneous composite signal $\hat{\mathbf{z}}^l$, which implies that the original non-convex constraint (23b) is approximated by a

Algorithm 2 CU-ADMM algorithm

Initialization:

- Set $k = 0$ and initialize $(\hat{\mathbf{x}}_1^0, \hat{\mathbf{x}}_2^0, \hat{\mathbf{z}}^0, \hat{\mathbf{y}}^0)$.
 Select the parameters (ρ, γ) .

Runtime:

- 1: **Repeat** l
 - 2: Update the constraint.
 Calculate $A^{l+1} = \gamma \|\hat{\mathbf{z}}^l\|_2 / \sqrt{L_{\text{sys}}}$.
 - 3: Process the $\hat{\mathbf{x}}_1$ -minimization step like O-ADMM algorithm.
 - 4: Process the $\hat{\mathbf{x}}_2$ -minimization step like O-ADMM algorithm.
 - 5: Process the $\hat{\mathbf{z}}$ -minimization step.

$$\hat{\mathbf{z}}^{l+1}(n) = \begin{cases} A^{l+1} e^{j\angle u^l(n)}, & |u^l(n)| > A^{l+1} \\ u^l(n), & |u^l(n)| \leq A^{l+1} \end{cases}$$
 - 6: Update the Lagrangian dual variable $\hat{\mathbf{y}}$ like O-ADMM algorithm.
 - 7: **until** reach the stop-criterion. Then output $\hat{\mathbf{z}}^l$.
-

series of iteratively updated convex constraints. Therefore, the $\hat{\mathbf{z}}$ -minimization step can be rewritten as

$$\hat{\mathbf{z}}^{l+1}(n) = \begin{cases} A^{l+1} e^{j\angle u^l(n)}, & |u^l(n)| > A^{l+1} \\ u^l(n), & |u^l(n)| \leq A^{l+1} \end{cases}, \quad (38)$$

where A^{l+1} is calculated as (37). We can see that the feasible region (25c) will gradually shrink as the iteration proceeds and the final optimized output $\hat{\mathbf{z}}$ will closely satisfy the PAPR constraint when the algorithm reaches the stop criterion. The proposed algorithm is referred as the constraint update-ADMM and summarized in Algorithm 2.

E. Computational Complexity

For the proposed O-ADMM algorithm, the computation complexity of calculating A in the initialization stage is $\mathcal{O}(L_{\text{sys}})$. In the $\hat{\mathbf{x}}_1$ -minimization step, the main complexity depends on the calculation of \mathbf{v}_1^l , in which the matrix multiplication is equivalent to a JN_1 -FFT operation. The computation complexity of JN_1 -point FFT is $\mathcal{O}(JN_1 \log_2(JN_1))$. The following calculation of $\hat{\mathbf{x}}_1^l$ contains an inverse matrix, i.e., $(1/\sigma_1^2 \mathbf{I}_{K_1} + \eta_1^2 \rho \mathbf{F}_1^H \mathbf{F}_1)^{-1}$. However, we can see that it can be calculated offline in advance. Similarly, the calculation of \mathbf{v}_2^l in $\hat{\mathbf{x}}_2$ -minimization step roughly requires twice JN_2 -point FFT operation, which leads to $\mathcal{O}(2JN_2 \log_2(JN_2))$ computation complexity. Thus, the computational cost to obtain $\hat{\mathbf{x}}_1^{l+1}$ and $\hat{\mathbf{x}}_2^{l+1}$ is dominant by JN_1 and JN_2 -point FFT operations, i.e. $\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$. In the $\hat{\mathbf{z}}$ -minimization step, the calculation of \mathbf{u}^l equally requires once JN_1 -IFFT and twice JN_2 -IFFT operations. The incurred complexities are $\mathcal{O}(JN_1 \log_2(JN_1))$ and $\mathcal{O}(2JN_2 \log_2(JN_2))$, respectively. From (36), we can also see that every entry \hat{z}_i^{l+1} in $\hat{\mathbf{z}}^{l+1}$ can be easily computed through once multiplication. In addition, since $\mathbf{F}_1 \hat{\mathbf{x}}_1^{l+1} + \mathbf{F}_2 \hat{\mathbf{x}}_2^{l+1}$ has been calculated in the $\hat{\mathbf{z}}$ -minimization step, the dual variable $\hat{\mathbf{y}}^{l+1}$ can be obtained merely through L_{sys} complex multiplications. According to the analysis above, we can conclude that the total computational cost in each iteration of O-ADMM is about

TABLE I
 COMPLEXITY COMPARISON OF THE INTERIOR-POINT METHOD,
 ADMM-BASED ALGORITHMS AND NS-ICF METHOD

Algorithm	Computational Complexity
interior-point method	$\mathcal{O}((N_1 + 2N_2)^3)$
O-ADMM/CU-ADMM	$\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$
NS-ICF	$\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$

$\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$. On the other hand, if the SOCP model in (24) is solved by standard interior-point methods, the computation complexity of solving the optimization problem is $\mathcal{O}((N_1 + 2N_2)^3)$ [23]. Because of $N \gg J$ in most cases, the proposed O-ADMM algorithm is highly efficient to solve the original optimization problem (24) compared with traditional interior-point methods.

As for the CU-ADMM algorithm, the only difference from O-ADMM is that the constraint in the $\hat{\mathbf{z}}$ -minimization step is updated in each iteration. This extra computational cost is $\mathcal{O}(L_{\text{sys}})$, which is far less than the FFT implementation. Hence, the computation complexity of CU-ADMM is also $\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$.

For comparison, we also consider the complexity of the NS-ICF method. The calculation of the clipping level incurs a complexity of $\mathcal{O}(L_{\text{sys}})$. Besides, The clipping procedure also leads to $\mathcal{O}(L_{\text{sys}})$. One can also find that the filtering procedure results in the primary computation complexity, which contains once JN_1 -point and twice JN_2 -point FFT/IFFT pair operations. Consequently, the computation complexity of NS-ICF is roughly $\mathcal{O}(JN_1 \log_2(JN_1) + 2JN_2 \log_2(JN_2))$. The complexity comparison of the interior-point method, the proposed ADMM-based algorithms and the NS-ICF method is summarized in Table I. Based on the analysis above, we can see that using interior-point method to solve SOCP has the highest computational cost. On the other hand, the complexities of the O-ADMM and CU-ADMM algorithms are much lower than interior-point method, in which each iteration has comparable computation complexity to the NS-ICF method.

F. Convergence Issue and Algorithm Summary

The ADMM approach has been widely used to solve structured convex optimization problems. However, most of the applications of ADMM is limited to the case of two separable variables, such as [26], [27], in which the algorithm has the theoretical linear convergence rate. The convergence of the ADMM with more than two variables remains an open question for a long time. In our cases, the mixed-numerology system always deal with $M \geq 2$ subbands, which leads to more than two variables in the optimization models (25). Fortunately, the recent work [36] shows that under three easily verifiable conditions the convergence of the ADMM with more than two variables still holds. As we can see from the (25), the objective function involves M 2-norm functions that can be easily proven to be strongly convex and Lipschitz continuous. Meanwhile, the linear constraint (25b) shows that the coefficient matrices of $\hat{\mathbf{x}}_1$, $\hat{\mathbf{x}}_2$ and $\hat{\mathbf{z}}$ have full column rank.

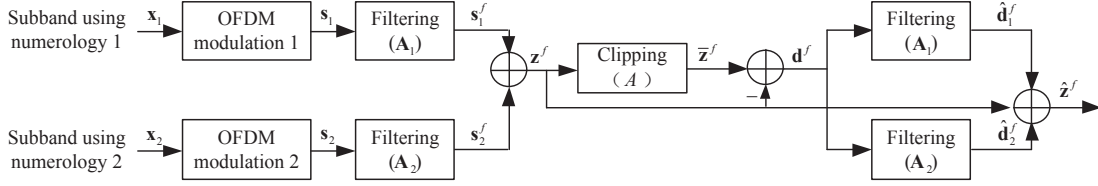


Fig. 3. Illustration of the NS-ICF method with filtering.

The conditions in [36] are then satisfied and our proposed O-ADMM is hence convergent as well. Interested readers are referred to [36] for the detailed discussion. As for the CU-ADMM, the step of the constraint update is heuristic and the theoretical proof for this modification is not straightforward. However, the simulation results show that the convergence can be always ensured.

According to the iterative steps of the proposed algorithms, we can see that the subproblems of \hat{x}_1 and \hat{x}_2 are performed on the subbands signals and the subproblem of \hat{z} produces the optimized composite signal. From this perspective, the ADMM framework can be seen as the generating procedure of PAPR reduced composite signal. Therefore, the proposed ADMM-based algorithms can be extended to the cases that have any number of numerologies.

V. APPLICATIONS WITH FILTERING AND WINDOWING TECHNIQUES

In order to improve the spectrum utilization efficiency as well as suppress the OOB, filtering and windowing techniques are proposed as two main waveform enhancements in 5G NR [13]. Filtering operation is performed on each subband with predesigned frequency response, which is also known as F-OFDM [19]. Windowed-OFDM (W-OFDM) introduces a well-chosen window function to be multiplied with the boundaries of each OFDM symbol [37]. F-OFDM and W-OFDM inherit the benefits of traditional OFDM while achieving the spectrum confinement, both of which are simple and can be transparent to the receiver. However, F-OFDM and W-OFDM maintain the high PAPR problem and thus suffer from PA nonlinearity. Nonlinear distortion also results in spectral regrowth that degrades the expected spectrum confinement performance [13].

In this section, we will show that the basic idea behind the proposed PAPR reduction methods can be easily extended to the variants of OFDM. In combination with filtering and windowing techniques, the proposed PAPR reduction methods are able to preserve the suppression of OOB while reducing the PAPR.

A. NS-ICF with Filtering

We assume that the i -th subband-wise filter impulse response is $\mathbf{a}_i = [a_i(0), a_i(1), \dots, a_i(L_f - 1)]$, where L_f is the filter length. In addition, the prototype filter design used in each subband is assumed to be same. We let the energy of

\mathbf{a}_i be normalized to unity, i.e., $\sum_{l=0}^{L_f-1} |a_i(l)|^2 = 1$. In order for matrix representation, we use an $(L_{sys} + L_f - 1) \times L_{sys}$ dimension Toeplitz matrix \mathbf{A}_i consisting of the filter impulse response to realize linear convolution. Then, the filtered signal in the i -th subband can be expressed as

$$\mathbf{s}_i^f = \mathbf{A}_i \mathbf{F}_i \mathbf{x}_i. \quad (39)$$

Similarly, for $M = 2$, the composite signal \mathbf{z}_i^f composed of filtered subbands signals can be written as

$$\mathbf{z}^f = \mathbf{A}_1 \mathbf{F}_1 \mathbf{x}_1 + \mathbf{A}_2 \mathbf{F}_2 \mathbf{x}_2. \quad (40)$$

The idea behind the NS-ICF method is to use the clipping noise instead of clipped signal to the filter, which avoids INI accumulation during the repeated execution. Combined with F-OFDM, the filtering procedure in NS-ICF can be realized through subband-wise filter. Fig. 3 illustrates the processing structure of NS-ICF for F-OFDM. It can be seen that after the clipping procedure, clipping noise is applied to individual subbands filters. The filtered clipping noise is obtained as $\hat{\mathbf{d}}_i^f = \mathbf{A}_i^f \mathbf{d}^f$ for $i = 1, 2$. Thus, the output composite signal is

$$\hat{\mathbf{z}}^f = \mathbf{z}^f + (\mathbf{F}_1 \mathbf{A}_1 + \mathbf{F}_2 \mathbf{A}_2) \mathbf{d}^f. \quad (41)$$

Similar to (17), we observe that the INI accumulation is also avoided in (41).

Compared with the original NS-ICF method, the filtering procedure employs the time-domain filter of F-OFDM. Thus, the computational cost mainly depends on the time domain convolution that leads to $\mathcal{O}(L_f L_{sys})$ complexity. For F-OFDM, since the filter length can be set as half of the signal duration to achieve fast frequency roll-off, the computation complexity of F-OFDM is significantly larger than original OFDM. Hence, we find that although the NS-ICF method combined with filtering technique provides better spectrum confinement, the computational cost is greatly increased.

B. Optimization Model with Windowing

The transmitter windowing $w(n)$ is commonly generated by the raised cosine shape function in [37]. The windowing function has unit response in the middle and smoothly converges to zero on the two sides. The roll-off factor is $\beta \in [0, 1)$ that controls the length of roll-off portion of the window, i.e., $L_{roff} = \beta L_{sys}$. L_{roff} samples at the beginning of the CP are weighted by the increasing window while an extra cyclic postfix with L_{roff} samples added after the OFDM symbol is multiplied by the decreasing window.

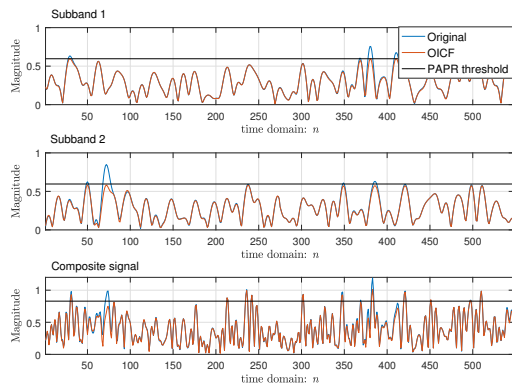


Fig. 4. Signal magnitude when using OICF method on individual subbands.

When transmitter windowing is applied, subband signal can be expressed as

$$\mathbf{s}_i^w = \mathbf{W}_i \mathbf{P}_i^w \mathbf{D}_i \mathbf{x}_i, \quad (42)$$

where the diagonal matrix $\mathbf{W}_i = \text{diag}(w_i(0), \dots, w_i(L_{\text{sys}} + L_{\text{roff}} - 1))$ denotes the windowing operation and $\mathbf{P}_i^w = [\mathbf{P}_i; \mathbf{I}_{L_{\text{roff}}}, \mathbf{0}_{L_{\text{roff}} \times (L_{\text{sys}} - L_{\text{roff}})}]$ performs the cyclic extension at both the prefix and postfix. Then, by denoting $\mathbf{F}_i^w = \text{blkdiag}(\eta_i \mathbf{W}_i \mathbf{P}_i^w \mathbf{D}_i, 2^{v_i})$, the composite signal for $M = 2$ can be written as

$$\mathbf{z}^w = \mathbf{F}_1^w \mathbf{x}_1 + \mathbf{F}_2^w \mathbf{x}_2. \quad (43)$$

We can find that the only difference between (43) and (8) is the coefficient matrix \mathbf{F}_i^w . Thus, this expression can be easily used in the optimization model (25) to take the place of the constraint (25b), which can be formulated as

$$\min_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{z}}^w} \frac{1}{2\sigma_1^2} \|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_2^2 + \frac{1}{2\sigma_2^2} \|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|_2^2 \quad (44a)$$

$$s.t. \quad \hat{\mathbf{z}}^w = \mathbf{F}_1^w \hat{\mathbf{x}}_1 + \mathbf{F}_2^w \hat{\mathbf{x}}_2, \quad (44b)$$

$$\|\hat{\mathbf{z}}^w\|_\infty \leq \gamma \frac{1}{\sqrt{L_{\text{sys}}}} \|\mathbf{z}^w\|_2. \quad (44c)$$

Note that the constraint (44b) still satisfies a linear relationship among $\hat{\mathbf{x}}_1$, $\hat{\mathbf{x}}_2$ and $\hat{\mathbf{z}}^w$. Therefore, the proposed O-ADMM and CU-ADMM algorithms can also be applied to solve this optimization problem (44). In addition, the windowing operation only needs to multiply the samples residing on the edges of signal by the increasing and decreasing window function, so that the windowing technique generally has negligible complexity overhead compared with the filtering technique. Hence, the computation complexities of ADMM-based algorithms combined with windowing technique are similar to those without windowing.

VI. SIMULATION RESULTS

In this section, we present simulation results to evaluate the performance of the proposed NS-ICF method and optimization method (the O-ADMM and CU-ADMM algorithms). In the simulations, we consider two adjacent subbands with different selection of numerology as illustrated in Fig. 2, i.e., $f_1 = 1/2f_2$. The CP length takes 7% of the symbol period. The

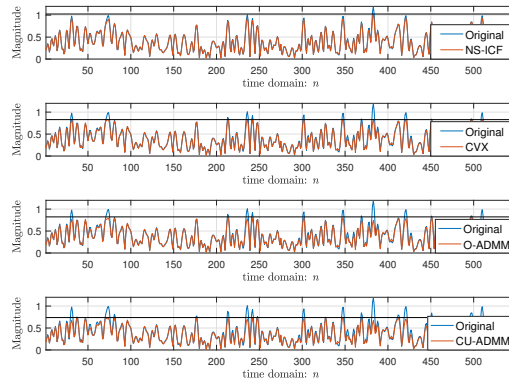


Fig. 5. Magnitude of original and PAPR-reduced composite signals using the proposed NS-ICF method, CVX, O-ADMM and CU-ADMM algorithm.

first subband modulates the data symbols on 56 subcarriers, while the second subband has 28 data subcarriers. Meanwhile, the reserved guard band is $G_1 = 8f_1$ between the subband 1 and subband 2. The oversampling rate is set as $J = 4$ and the QPSK modulation is adopted by both of subbands. The penalty parameter for the proposed ADMM-based algorithms is $\rho = 0.25$. The CR is set to 5dB. The transmission power on every subband is equal to each other and normalized to 1, which makes the total power evenly allocated on the available bandwidth. In our simulation, we randomly generate 5000 LCM symbols.

A. PAPR Reduction Verification

Firstly, we examine the time-domain composite signal when using the OICF method on separate subbands. The original signal is also depicted for comparison. As shown in Fig. 4, the first and second plots are the signal amplitudes in subband 1 and 2, respectively. It can be observed that the peaks can be reduced compared with the original signal in each subband. Nonetheless, when combining two subbands' signals, several peaks still rise high in the third sub-figure of Fig. 4. This demonstrates that the PAPR reduction techniques individually applied to each subband do not work in the mixed-numerology systems. In Fig. 5, we show the signal amplitudes generated by the NS-ICF method and optimization method. The optimization problem for the SOCP in (24) is solved by the proposed O-ADMM algorithm and CVX individually. We can see that all methods are able to reduce the PAPR of composite signal. Specifically, the horizontal lines indicate that the NS-ICF method achieves PAPR of 7dB, while the optimization methods using CVX and O-ADMM attain exactly the same PAPR of 5.5dB. Furthermore, it is of interest that the proposed CU-ADMM algorithm achieves the best PAPR performance, which reaches to the desired PAPR threshold 5dB within one execution.

B. Algorithm Efficiency of O-ADMM and CU-ADMM

To evaluate the convergence of the proposed O-ADMM and CU-ADMM algorithms, we plot the EVM of composite

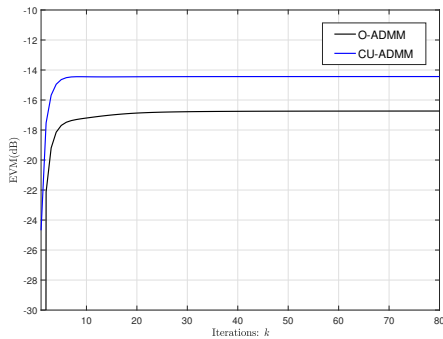


Fig. 6. Objective function with respect to the number of iterations.

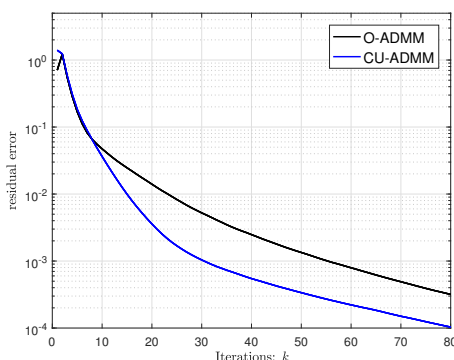


Fig. 7. Primal residual with respect to the number of iterations.

signal with respect to the number of iterations. In Fig. 6, we can see that both O-ADMM and CU-ADMM algorithms can converge with no more than 10 iterations. We can also find that the resulting EVM of CU-ADMM is slightly worse than O-ADMM due to the iterative update of the constraint in the $\hat{\mathbf{z}}$ -minimization step. In Fig. 7, the primal residual error is defined as $\|\mathbf{F}_1 \hat{\mathbf{x}}_1^l + \mathbf{F}_2 \hat{\mathbf{x}}_2^l - \hat{\mathbf{z}}^l\|_2^2$. From the convergence curves, one can observe that the residual error decreases rapidly in the first few times and this bend becomes gradually flat after 10 iterations. Meanwhile, the CU-ADMM algorithm shows better convergence rate than O-ADMM. According to the two figures, the stopping criterion can be determined carefully.

Apart from the convergence, the low computation complexity of the proposed ADMM-based algorithms is also one major reason that leads to high efficiency. Table II compares the running time of the ADMM-based algorithm with 10 iterations to that of using the CVX toolbox. The simulation environment is based on MATLAB. According to the execution times shown in Table II, the proposed ADMM-based algorithms are nearly 1000 times time-saving compared with the SOCP model using CVX, which is of significant benefits for practical implementation. In addition, the elapsed time of CU-ADMM is slightly higher than that of O-ADMM, which is probably due to the constraint update step.

TABLE II
EXECUTION TIME COMPARISON OF O-ADMM, CU-ADMM, AND CVX METHODS

Algorithm	O-ADMM	CU-ADMM	CVX
Time elapse (ms)	6.1	6.2	6249

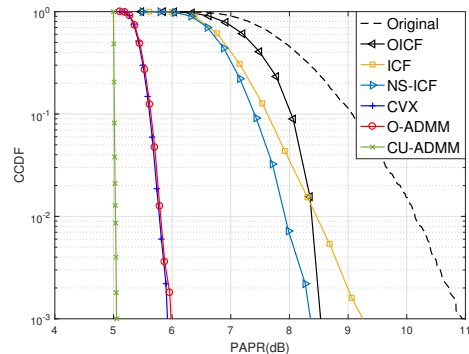


Fig. 8. PAPR performance of the ICF, NS-ICF methods, O-ADMM and CU-ADMM algorithms, CVX, and OICF method when the CR is set to 5dB.

C. PAPR Comparison

In this subsection, we use the complementary cumulative distribution function (CCDF) to evaluate the PAPR reduction performance. It is defined as the probability that the PAPR of signal exceeds a given PAPR value, which is expressed as

$$\text{CCDF} = \Pr(\text{PAPR} > \text{PAPR}_0), \quad (45)$$

where $\Pr(\cdot)$ denotes the probability function.

Fig. 8 shows the CCDF curves of different PAPR reduction methods with one execution. The CCDF of the original composite signal is also considered as a reference. As expected, the OICF method working on the separate subbands has the worst PAPR performance, which only has 2dB PAPR reduction compared to the original signal. Besides, the direct application of the ICF method leads to a slow fall of CCDF curve. Meanwhile, it can be seen that the proposed NS-ICF method leads to limited improvement of PAPR performance. Due to the peak regrowth, the NS-ICF method, like the classical ICF method, usually has to be executed many times to gradually approach the desired PAPR. The optimization method, however, achieves significant performance of PAPR reduction and shows steep curves of CCDF. To be more specific, when solving the optimization problem (23), the proposed O-ADMM algorithm and the CVX toolbox have visually overlapped CCDF curves, which are very close to the preset PAPR threshold. However, due to the high efficiency of the ADMM approach, the O-ADMM algorithm is far more computationally efficient. Furthermore, we can observe that the proposed CU-ADMM algorithm shows a cut-off CCDF curve right on the desired PAPR. It shows that updating the constraint of $\hat{\mathbf{z}}$ -minimization step at each iteration leads to the further improvement on PAPR performance.

Fig. 9 plots the CCDFs of the PAPR using NS-ICF method (executing 1, 6, and 12 times), and the proposed O-ADMM algorithm (executing 1 and 2 times). It can be seen that both the NS-ICF method and the O-ADMM algorithm can

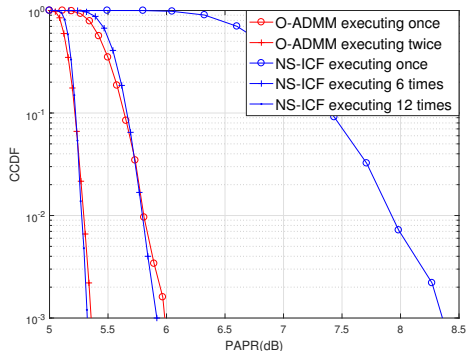


Fig. 9. PAPR performance of the NS-ICF and optimization methods with different times of execution.

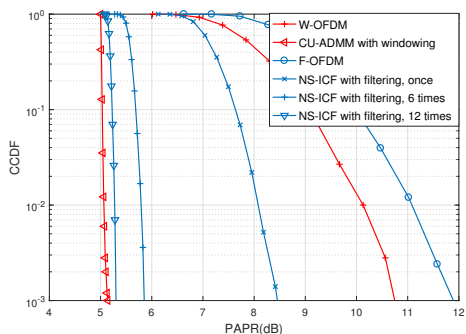


Fig. 10. PAPR performance of the NS-ICF method with filtering and CU-ADMM algorithm with windowing.

gradually reduce the PAPR to the desired level with being executed repeatedly. However, for the O-ADMM algorithm, the PAPR reduction is about 5.9dB at a probability of 10^{-3} after executing once and about 5.3dB after executing twice, while the NS-ICF method requires 6 and 12 times of execution to achieve a comparable PAPR reduction performance. Thus we can reach a similar conclusion as in [23] that the optimization method requires fewer executions to approach the PAPR threshold.

To illustrate the applications of proposed PAPR reduction methods with spectrum confinement techniques, F-OFDM and W-OFDM are considered in our simulations. Subband-wise filters are designed based on the root raised cosine windowed Sinc function with 128 length for F-OFDM. The raised cosine window with $\beta = 0.04$ is applied for W-OFDM. Meanwhile, we use the NS-ICF method for F-OFDM and the CU-ADMM algorithm for W-OFDM. From Fig. 10, it can be observed that the proposed methods are still effective to reduce the PAPR of composite signal. The NS-ICF method results in more than 6dB PAPR reduction after 12 times executions. Besides, as expected, the CU-ADMM algorithm achieves a steep drop at the PAPR of 5dB.

In order to show the influence of PAPR reduction methods on the OOB performance of F-OFDM and W-OFDM, we further consider passing the composite signal through the solid state power amplifier (SSPA) [14]. The smoothing factor of SSPA model is set to 3 and the input back-off (IBO) is 5dB. The power

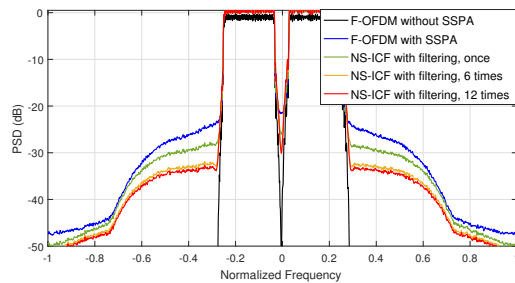


Fig. 11. OOB performance of the NS-ICF method with filtering.

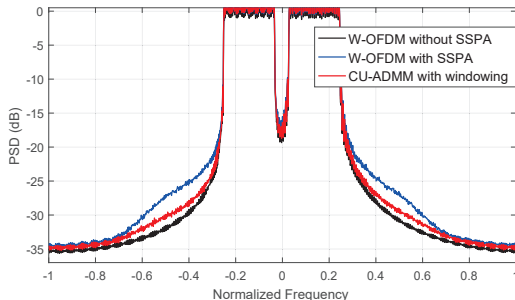


Fig. 12. OOB performance of the CU-ADMM algorithm with windowing.

spectrum densities (PSD) are depicted in Fig. 11 and Fig. 12, which are computed by means of periodogram. From Fig. 11, we can see that the OOB performance of F-OFDM is obviously degraded at the output of SSPA. When using the proposed NS-ICF method, it can achieve near 10dB lower OOB than that without PAPR reduction after 12 times of execution. As for W-OFDM, the proposed CU-ADMM algorithm leads to a similar OOB performance compared to the W-OFDM without SSPA in Fig. 12. Note that spectrum confinement is of great importance in 5G communications. Thus, it is beneficial to exploit the PAPR reduction techniques to avoid the nonlinearity of PA, which definitely saves the property of low OOB brought by spectrum confinement techniques.

D. EVM Analysis

We now show the comparison of distortion for different PAPR reduction methods. The root mean-square (RMS) EVM is used here to evaluate the distortion statistically, which is defined as

$$\text{RMSEVM} = \sqrt{E[\text{EVM}^2]}, \quad (46)$$

where the expectation is calculated by averaging 5000 LCM symbols. The target PAPR threshold is 5dB for the CU-ADMM algorithm and it is carefully chosen for ICF, NS-ICF and O-ADMM so that the same PAPR can be achieved with one execution.

From Table III, we can see that the optimization method achieves better EVM performance than the ICF and NS-ICF methods. The O-ADMM algorithm achieves the EVM of -14.24dB, while the CU-ADMM only results in negligible increase of EVM, which is -14.03dB. The NS-ICF method

TABLE III
PERFORMANCE COMPARISON OF THE ICF, NS-ICF METHODS AND
OPTIMIZATION METHOD (USING O-ADMM, CU-ADMM ALGORITHMS).

Algorithm	EVM(dB) in Subband 1	EVM(dB) in Subband 2	EVM(dB) of LCM symbol
ICF	-13.75	-13.74	-10.73
NS-ICF	-15.50	-15.51	-12.50
O-ADMM	-17.25	-17.25	-14.24
CU-ADMM	-17.04	-17.04	-14.03

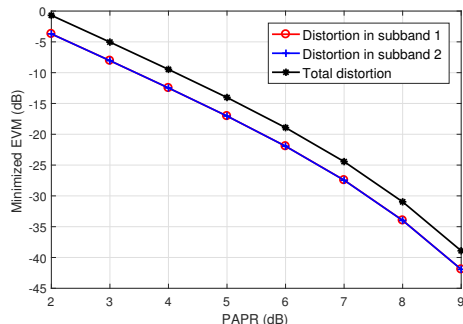


Fig. 13. Relation between the target PAPR and EVM performance using the CU-ADMM algorithm.

generally causes EVM deterioration, which is 1.74dB worse than the O-ADMM algorithm. Moreover, the EVM performance of the ICF method is further degraded due to the INI accumulation. Additionally, we can see that the distortion is fairly uniform on both subbands, each of which has similar EVM performance to the other.

As the CU-ADMM algorithm can be roughly regarded as an optimizer that achieves the minimal EVM with only one execution, we can explore the relationship between EVM and PAPR. From Fig. 13, it can be observed that the EVM drops rapidly as the target PAPR increases. When we set a higher PAPR, the EVM deterioration is alleviated since fewer amplitudes of the signal need to be changed. In addition, this method leads to a fair EVM performance for each subband. With the help of CU-ADMM, we can wisely choose system parameters satisfying the PAPR and EVM requirements.

VII. CONCLUSIONS

In this paper, the PAPR reduction techniques are investigated for the mixed-numerology transmissions. We first consider the classical clipping based approach and show that the direct application of the ICF method to composite signal will result in the INI accumulation. By exploiting the clipping noise rather than the clipped signal, the NS-ICF method is able to reduce the high peaks without increasing the INI. Next, we take the EVM distortion into consideration and formulate the task into an EVM minimization problem subject to the PAPR constraint. Then, taking advantage of the separable structure of the convex programming, two low complexity algorithms, named as O-ADMM and CU-ADMM, are developed to efficiently solve this problem. As its decomposing and coordinating procedure can be associated with the structure of mixed numerology, the ADMM framework provides a tractable and scalable solution for the PAPR reduction problem

with any number of numerologies. Furthermore, the proposed PAPR reduction methods can be easily extended to the F-OFDM and W-OFDM. In combination with the filtering and windowing techniques, the proposed PAPR reduction methods achieve the suppression of OOB as well as alleviate the spectrum regrowth caused by PA nonlinearity.

Our future work will focus on the joint optimization of the OOB and PAPR using the ADMM framework. For example, the combination of the optimization method and windowing technique discussed in Section V does not consider the windowing function as the optimization objective. Hence, we can further take into account of other possible optimized designs.

REFERENCES

- [1] L. Zhang, A. Ijaz, P. Xiao, and R. Tafazolli, "Multi-service system: An enabler of flexible 5G air interface," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 152–159, Oct. 2017.
- [2] A. A. Zaidi, R. Baldemair, V. Moles-Cases, N. He, K. Werner, and A. Cedergren, "OFDM numerology design for 5G new radio to support IoT, eMBB, and MBSFN," *IEEE Communications Standards Magazine*, vol. 2, no. 2, pp. 78–83, Jun. 2018.
- [3] A. Ijaz, L. Zhang, M. Grau, A. Mohamed, S. Vural, A. U. Qaddus, M. A. Imran, C. H. Foh, and R. Tafazolli, "Enabling massive IoT in 5G and beyond systems: PHY radio frame design considerations," *IEEE Access*, vol. 4, pp. 3322–3339, 2016.
- [4] 3rd Generation Partnership Project (3GPP), "Study on scenarios and requirements for next generation access technologies," 3GPP Tech. Rep., TR 38.913, Oct. 2016.
- [5] C. I. S. Han, Z. Xu, S. Wang, Q. Sun, and Y. Chen, "New paradigm of 5G wireless internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 474–482, Mar. 2016.
- [6] 3rd Generation Partnership Project (3GPP), "Technical Specification Group Radio Access Network; NR; Physical Layer; General Description (Release 15)," 3GPP Tech. Spec., TS 38.201, Dec. 2017.
- [7] A. Sahin and H. Arslan, "Multi-user aware frame structure for OFDMA based system," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Quebec, Canada, Sep. 2012, pp. 1–5.
- [8] L. Zhang, A. Ijaz, P. Xiao, A. Qaddus, and R. Tafazolli, "Subband filtered multi-carrier systems for multi-service wireless communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1893–1907, Mar. 2017.
- [9] P. Guan, D. Wu, T. Tian, J. Zhou, X. Zhang, L. Gu, A. Benjebbour, M. Iwabuchi, and Y. Kishiyama, "5G field trials: OFDM-based waveforms and mixed numerologies," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1234–1243, Jun. 2017.
- [10] A. Yazar and H. Arslan, "Flexible multi-numerology systems for 5G new radio," *Journal of Mobile Multimedia*, vol. 14, no. 2, pp. 367–394, Oct. 2018.
- [11] —, "Reliability enhancement in multi-numerology-based 5G new radio using INI-aware scheduling," *EURASIP Journal on Wireless Communications and Networking*, vol. 110, no. 1, pp. 1–14, May 2019.
- [12] D. Lim, S. Heo, and J. No, "An overview of peak-to-average power ratio reduction schemes for OFDM signals," *J. Commun. Networks*, vol. 11, no. 3, pp. 229–239, Jun. 2009.
- [13] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.
- [14] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, 4th Quart. 2013.
- [15] R. W. Bauml, R. F. H. Fischer, and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electron. Lett.*, vol. 32, no. 22, pp. 2056–2057, Oct. 1996.
- [16] S. H. Muller and J. B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electron. Lett.*, vol. 33, no. 5, pp. 368–369, Feb. 1997.
- [17] J. Tellado and J. M. Cioffi, "Peak power reduction for multicarrier transmission," in *Proc. IEEE GLOBECOM*, vol. 99, Sydney, Australia, Nov. 1998, pp. 5–9.

- [18] B. S. M. Sabbaghian, Y. Kwak and V. Tarokh, "Near Shannon limit and low peak to average power ratio turbo block coded OFDM," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2042–2045, Aug. 2011.
- [19] X. Zhang, L. Chen, J. Qiu, and J. Abdoli, "On the waveform for 5G," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 74–80, Nov. 2016.
- [20] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *Electron. Lett.*, vol. 38, no. 5, pp. 246–247, Feb. 2002.
- [21] A. Aggarwal and T. H. Meng, "Minimizing the peak-to-average power ratio of OFDM signals using convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3099–3110, Aug. 2006.
- [22] Q. Liu, R. J. Baxley, X. Ma, and G. T. Zhou, "Error vector magnitude optimization for OFDM systems with a deterministic peak-to-average power ratio constraint," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 418–429, Jun. 2009.
- [23] Y. Wang and Z. Luo, "Optimized iterative clipping and filtering for PAPR reduction of OFDM signals," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 33–37, Jan. 2011.
- [24] X. Zhu, W. Pan, H. Li, and Y. Tang, "Simplified approach to optimized iterative clipping and filtering for PAPR reduction of OFDM signals," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1891–1901, May 2013.
- [25] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [26] H. Bao, J. Fang, Q. Wan, Z. Chen, and T. Jiang, "An ADMM approach for PAPR reduction for large-scale MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7407–7418, Aug. 2018.
- [27] Y. Wang, Y. Wang, and Q. Shi, "Optimized signal distortion for PAPR reduction of OFDM signals with IFFT/FFT complexity via ADMM approaches," *IEEE Trans. Signal Process.*, vol. 67, no. 2, pp. 399–414, Jan. 2019.
- [28] S. Pagadarai, A. Kliks, H. Bogucka, and A. M. Wyglinski, "Non-contiguous multicarrier waveforms in practical opportunistic wireless systems," *IET Radar, Sonar Navigation*, vol. 5, no. 6, pp. 674–680, Jul. 2011.
- [29] Z. Khan, H. Ahmadi, E. Hossain, M. Coupechoux, L. A. Dasilva, and J. J. Lehtomki, "Carrier aggregation/channel bonding in next generation cellular networks: methods and challenges," *IEEE Network*, vol. 28, no. 6, pp. 34–40, Nov. 2014.
- [30] 3rd Generation Partnership Project (3GPP), "New radio (NR); base station transmission and reception (Release 15)," 3GPP Tech. Spec., TS 38.104, Dec. 2017.
- [31] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 282–289, Feb. 2001.
- [32] —, "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.
- [33] L. Wang and C. Tellambura, "Analysis of clipping noise and tone-reservation algorithms for peak reduction in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1675–1694, May 2008.
- [34] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming (web page and software)*, Oct. 2008. [Online]. Available: <http://stanford.edu/~boyd/cvx>
- [35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [36] T. Lin, S. Ma, and S. Zhang, "On the global linear convergence of the ADMM with multiblock variables," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1478–1497, Jul. 2015.
- [37] A. A. Zaidi, R. Baldemair, H. Tullberg, H. BJORKEGREN, L. Sundstrom, J. Medbo, C. Kilinc, and I. D. Silva, "Waveform and numerology to support 5G services and requirements," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 90–98, Nov. 2016.