PAPR Reduction for OFDM Signals: A Monte Carlo Tree Search Method

Shuai Ye, Tianyu Wang, and Shaowei Wang
School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China
E-mail: dz20230027@smail.nju.edu.cn, {tianyu.alex.wang, wangsw}@nju.edu.cn

2021 IEEE/CIC International Conference on Communications in China (ICCC) | 978-1-6654-4385-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCC52777.2021.9580407

Abstract—**In this paper, we propose a low complexity partial transmit sequence (PTS) method to reduce the peak-to-average power ratio (PAPR) of OFDM signals, which is based on Monte Carlo tree search. Compared to the computation prohibitive exhaustive search, our proposed method can achieve a better tradeoff between PAPR reduction and computing load by cutting down the number of candidate phase vectors to the square of subblocks. Numerical results show that our proposed method reduces the computational complexity of the classic reduced complexity PTS method by 35 times at the cost of 1.16 dB PAPR performance degradation when the subblock number is 8, or achieves a 0.2 dB PAPR performance improvement at half of the computation cost of the classic iterative flipping PTS method.**

Index Terms—**Monte Carlo tree search, OFDM, partial transmit sequence, peak-to-average power ratio.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) is widely used as a high data rate transmission scheme in 4G wireless communication systems [1–4]. However, a major disadvantage of the OFDM system is that the OFDM symbols generally possess high peak-to-average power ratio (PAPR), which can drive the OFDM symbols into the nonlinear region of the power amplifier at the transmitter and lead to harmful in-band distortion and out-of-band radiation. Filtered-OFDM (F-OFDM) is considered as a promising candidate waveform to meet the requirement of the fifth generation mobile communication technology, which can suppress out-of-band emission and support asynchronous transmission. However, F-OFDM still faces the high PAPR problem due to the orthogonal transmission [5].

Existing PAPR reduction techniques include clipping, tone injection, coding, selective mapping and partial transmit sequence (PTS) [6–8], among which the PTS is a distortionless scheme that maintains the original data transmission rate at the cost of only a small amount of side information. Specifically, the PTS scheme divides the original symbol into orthogonal subblocks and weights each subblock with different phase factors. The objective is to find the optimal phase vector that minimizes the PAPR of the combined symbols. The classic PTS (C-PTS) method employs an exhaustive search to find the optimal phase vector and calculates the PAPR of each candidate phase vector with all samples of the original OFDM signal [9]. When the number of subblocks and the length of OFDM signals are large, the C-PTS method suffers from high computational complexity inevitably.

The phase vector needs to be calculated for each coming OFDM symbol, i.e., around every 70 ms in LTE systems, to maintain a low PAPR of the transmit signal. To alleviate system computation burden, various modified-PTS algorithms have been proposed [10]. Reducing the number of candidate phase vectors is a feasible way to meet this end [11, 12]. The iterative flipping PTS (I-PTS) method uses a greedy algorithm to sequentially determine phase factors of each subblock, such that the number of candidate phase vectors is reduced to be proportional to the subblock number [11]. Simultaneous flips of multiple phase factors can improve the PAPR performance at the cost of additional computation load [12]. These methods greatly reduce the complexity level but with limited PAPR performance. Another way is to decrease the complexity of PAPR calculation [13–15]. The reduced complexity PTS (RC-PTS) method uses dominant time domain samples to estimate the PAPR value of each candidate phase vector, which reduces the computation load remarkably with slight PAPR performance degradation [13]. Sample selection metrics have been proposed to further improve the accuracy of PAPR estimation [14, 15]. The performance of these methods is close to that of the C-PTS but their computational complexity also grows exponentially with the subblock number.

In this paper, we introduce a Monte Carlo tree search (MCTS) based PTS method, referred to as M-PTS, to reduce the PAPR of OFDM signals. Specifically, we adopt the RC-PTS method to estimate the PAPR of each candidate phase vector, and then utilize the MCTS method to search for the optimal phase vector, in which the number of candidate phase vectors is proportional to the square of the subblock number. Actually, our proposed MCTS based phase vector search method can be combined with any dominant time domain samples PTS method to further reduce its complexity. Simulation results show that our proposed M-PTS method can achieve a promising tradeoff between PAPR performance and computational complexity when the subblock number is large.

The rest of this paper is organized as follows. In Section II, we review the PAPR in OFDM systems and give a brief introduction to the C-PTS, I-PTS and RC-PTS methods. The proposed MCTS based method is explained in detail in Section III. Numerical results and discussions are provided in Section IV. In Section V, we conclude the paper.

This work was partially supported by the National Natural Science Foundation of China (61801208, 61931023, U1936202).

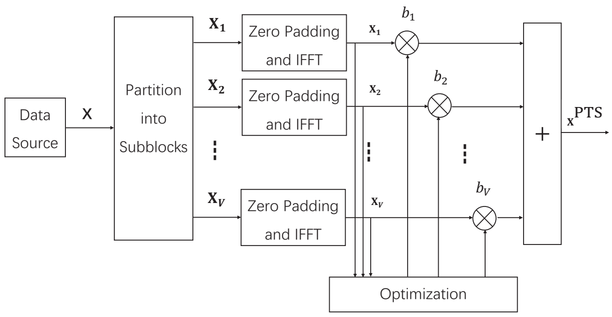

Fig. 1. Block diagram of the PTS scheme.

II. Problem Formulation

A. PAPR Definition

Let $\mathbf{X} = [X_0, X_1, \ldots, X_{N-1}]^T$ denote an OFDM symbol in the frequency domain, where N is the number of subcarriers and $X_i, 0 \leq i \leq N-1$ is the modulated signal of the i-th subcarrier. In time domain, the OFDM symbol can be seen as the sum of N narrow band signals corresponding to N subcarriers, which is given by

$$x(t) = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_i e^{j2\pi it/T}, \quad 0 \leq t \leq T, \tag{1}$$

where T is the length of the OFDM symbol in time domain.

The discrete time symbol $\mathbf{x} = [x_0, x_1, \ldots, x_{NL-1}]^T$ with sampling rate LN/T is then given by

$$x_m = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_i e^{j2\pi im/NL}, \quad 0 \leq m \leq NL-1, \tag{2}$$

where L is the oversampling ratio.

The PAPR of an OFDM signal is then defined as the ratio of the maximum instantaneous power to the average power over the entire OFDM symbol. Its value is usually approximated with the oversampled OFDM symbol, given by

$$\text{PAPR}(\mathbf{x}) = \frac{\max\limits_{0\leq m\leq NL-1} \left\{|x_m|^2\right\}}{E\left\{|x_m|^2\right\}}. \tag{3}$$

Such an approximation is accurate enough when the oversampling rate $L \geq 4$ as discussed in [16].

B. PTS Scheme

A general block diagram of the PTS scheme is shown in Fig. 1, where the OFDM symbol $\mathbf{X}$ is divided into V subblocks and each of them is denoted by $\mathbf{X}_v = [X_{v,0}, X_{v,1}, \ldots, X_{v,N-1}]^T, 1 \leq v \leq V$. Without loss of generality, we assume the number of subcarriers N is divisible by V, and set $M = N/V$ as the number of effective subcarriers of each subblock. For any subblock $\mathbf{X}_v$, its first $(v-1)M$ elements and last $(V-v)M$ elements are 0, i.e., $X_{v,i} = 0, i \in [0, (v-1)M-1] \cup [vM, N-1]$, and the middle elements are set to be equal to the elements of the original symbol in the corresponding subcarriers, i.e., $X_{v,i} = X_i, i \in [(v-1)M, vM-1]$. Thus, the original OFDM symbol can be represented by the sum of V subblocks,

$$\mathbf{X} = \sum_{v=1}^{V} \mathbf{X}_v. \tag{4}$$

Each subblock is individually oversampled with ratio L and transformed into time domain via an IFFT with length NL. The discrete time subblock v is $\mathbf{x}_v = [x_{v,0}, x_{v,1}, \ldots, x_{v,NL-1}]^T$, where each element $x_{v,m}$ is given by

$$x_{v,m} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_{v,i} e^{j2\pi im/NL}, \quad 0 \leq m \leq NL-1. \tag{5}$$

Each time domain subblock $\mathbf{x}_v$ is then multiplied by a constant phase factor $b_v = e^{j\phi_v}, \phi_v \in [0, 2\pi)$, and the corresponding phase vector is denoted by $\mathbf{b} = (b_1, b_2, \ldots, b_V)$. The outcome of the time domain signal is then given by

$$\mathbf{x}^{\text{PTS}} = \sum_{v=1}^{V} b_v \mathbf{x}_v. \tag{6}$$

The phase factors are chosen from a predefined set Φ that divides 2π phase interval into W equal sections, which is formally given by

$$\Phi = \left\{e^{j2\pi w/W} | 0 \leq w \leq W-1\right\}. \tag{7}$$

Therefore, the optimal PTS phase vector problem is formulated as follows,

$$\min_{\mathbf{b}} \quad \text{PAPR}(\mathbf{x}^{\text{PTS}}) \tag{8}$$

$$s.t. \quad b_v \in \Phi, 1 \leq v \leq V \tag{9}$$

The C-PTS method uses brute search to find the optimal phase vector that minimizes the PAPR of the outcome signals. Note that we can always set the phase factor of the first subblock $b_1 = 1$ without loss of optimality. The number of phase vectors searched in C-PTS is then given by W^{V-1}. The I-PTS method starts by setting all the phase factors as 1, i.e., $b_v = 1, 1 \leq v \leq V$ [11]. Then, it sequentially determines the $V-1$ phase factors from $v = 2$ to $v = V$ in a greedy fashion. For any subblock v, given the current phase factors of all the other $V-1$ subblocks, the myopic optimal phase factor in Φ that minimizes the PAPR of the current outcome OFDM signal is selected. The number of phase vectors searched in the I-PTS method is then given by $W(V-1)$.

The RC-PTS method defines a metric $Y_m = \sum_{v=1}^{V} |x_{v,m}|^2$, for each time domain sample point $m \in \{0, 1, \ldots, NL-1\}$. The set of selected points is then given by $S_Y = \{m | Y_m \geq h(\gamma)\}$, where $h(\gamma)$ is the threshold function. We denote by U the number of selected points. Thus, the v-th subblock with dominant time domain samples is given by

$$\hat{\mathbf{x}}_v = \left[x_{v,q_1}, x_{v,q_2}, \ldots, x_{v,q_U}\right]^T, \tag{10}$$

where $q_i \in S_Y$ for all $i \in \{1, 2, \ldots, U\}$. The dominant time domain signal and the corresponding weighted signal are given by $\hat{\mathbf{x}} = \sum_{v=1}^{V} \hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}^{\text{PTS}} = \sum_{v=1}^{V} b_v \hat{\mathbf{x}}_v$, respectively. We denote $P_\gamma = \Pr\{Y_m \geq h(\gamma)\}$ as the probability that the metric Y_m is

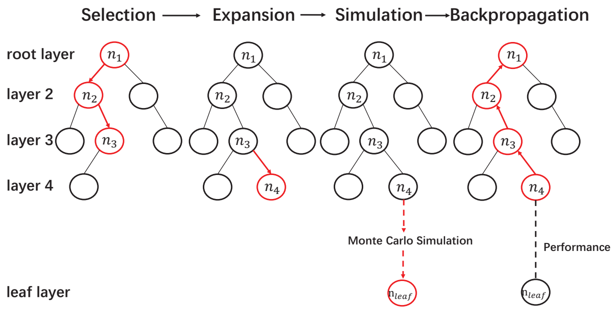

Fig. 2. A four-step loop in the MCTS method.

above the threshold $h(\gamma)$. The computational complexity of PAPR calculation is then reduced from LN to $P_\gamma LN$.

The transmission of side information is another research aspect of the PTS technique. For example, the number of phase vectors is W^{V-1} in the C-PTS method and this needs to transmit $\log_2(W^{V-1})$ bit side information per signal. In this paper, we assume the side information can be transmitted correctly for all PTS methods. Thus the OFDM signals are recovered without causing the bit error ratio degradation at the receiver side.

III. PAPR Reduction with M-PTS

A. MCTS Preliminaries

MCTS is a family of algorithms that have been successfully used in problems with massive searching spaces, e.g., the games of go and chess [17]. The basic idea of MCTS is to formulate a tree structure to represent the searching space, while in each exploration step along the tree structure, to use a large number of well-designed Monte Carlo simulations to estimate the unknown evaluation function of the child nodes. For a specific problem, the basic process of MCTS can be further specified and strengthened, e.g., AlphaGo [18] strengthens the MCTS using a deep neural network to deal with the massive searching space in go game.

The MCTS searches for the optimal solution by sequentially building a path from the root node to a leaf node. Specifically, it decides the best child node in the next layer at each time by building a search tree that is rooted at the current node. The search tree is built iteratively by performing a four-step loop for a predefined amount of times: *selection*, *expansion*, *simulation* and *backpropagation*, which are described as follows.

- *Selection*: Select the most urgent expandable node in the search tree by using a node selection policy.
- *Expansion*: Expand the selected node by adding a randomly selected child node that is not in the search tree.
- *Simulation*: Run one Monte Carlo simulation from the child node to a leaf node and calculate the performance of the corresponding solution.
- *Backpropagation*: Update the values of each node in the search tree that belongs to the simulated path according to its performance.

The four-step loop is illustrated in Fig. 2. In the selection step, n_3 is selected as the most urgent expandable node. Then, in the expansion step, we expand n_3 by adding the child node n_4. In the simulation step, a randomly simulated path from n_4 to a leaf node n_{leaf} is built and the corresponding performance is calculated. In the backpropagation step, the values of node n_1, n_2, n_3 , n_4 are updated according to the performance of the simulated path. The MCTS algorithm is summarized in Algorithm 1, where r is the current root node, n_s is the selected node for expansion, n_c is the added child node, and Δ represents the performance of the simulated path.

Algorithm 1 MCTS Algorithm

Input: Root node r_0
Output: Path p

initialize $p = (r_0)$, $r = r_0$
while r is not a leaf **do**
  while within computational budget **do**
    $n_s \leftarrow$ selection(r)
    $n_c \leftarrow$ expansion(n_s)
    $\Delta \leftarrow$ simulation(n_c)
    backpropagation(Δ)
  end while
  $p \leftarrow$(p,bestchild(r))
  r = bestchild(r)
end while
return p

B. M-PTS Algorithm

We first use the RC-PTS method to generate the subblocks $\hat{\mathbf{x}}_v$ with dominant time domain samples. Then we use MCTS to search for the optimal phase vector instead of using an exhaustive search. Each candidate phase vector can be represented by a path from the root to a leaf in a W-ary tree Tr^{PTS} with depth V, where each node in the path represents the phase factor of the corresponding subblock.

In the selection step, we adopt upper confidence bound (UCB) policy [19, 20]. For any search tree $\text{Tr} \subseteq \text{Tr}^{\text{PTS}}$ rooted at node r, the policy starts from the root node and selects the child node in the next layer with the largest UCB value, until it reaches a leaf of Tr or a node that has at least one unexpanded child node in Tr^{PTS}. For any node $n \in \text{Tr}$, the UCB value is calculated based on its reward value $Q(n)$ and expansion value $E(n)$, as well as the expansion value $E(n')$ of its father node n'. $Q(n)$ and $E(n)$ are initialized by 0 when node n is first added in an expansion step and updated iteratively at each following backpropagation step. The UCB value of node n is formally given by

$$\text{UCB}(n) = \frac{Q(n)}{E(n)} + c\sqrt{\frac{2lnE(n')}{E(n)}}, \tag{11}$$

where c is a tunable weight factor that balances the exploitation part $Q(n)/E(n)$ and the exploration part $\sqrt{2lnE(n')/E(n)}$. We denote the selected node by n_s.

In the expansion step, a randomly chosen child node n_c of the selected node n_s is added to the current search tree Tr,

Algorithm 2 M-PTS Method for PAPR Reduction

Input: Subblocks $\hat{\mathbf{x}}_v$ with dominant time domain samples
Output: Phase vector $\mathbf{b}$

```
build a W-ary tree Tr^PTS with depth V rooted at node r_0
initialize b_1 = 1, r = r_0
build a search tree Tr ⊆ Tr^PTS with root node r
initialize Q(r) = 0, E(r) = 0
for v = 2 : V do
    for i = 1 : g(v) do
        initialize n_s = r
        while all children of n_s in Tr^PTS are expanded do
            n_s ← the child of n_s with the largest UCB
              value as given in (11)
        end while
        n_c ← randomly select an unexpanded child node
          of n_s in Tr^PTS
        add n_c to the current search tree Tr and initialize
          Q(n_c) = 0, E(n_c) = 0
        simulate a path p from n_c to a leaf in Tr^PTS and
          calculate reward Δ using (12)
        update Q(n) and E(n) using (13) and (14) for all
          nodes that belong to both Tr and p
    end for
    r ← the child node of r with largest Q(n)/E(n)
    b_v ← the phase factor represented by node r
end for
return b=(b_1, b_2, ..., b_V)
```

which implies that a randomly chosen phase factor from Φ is considered for the corresponding subblock.

In the simulation step, a path from the root to a leaf in Tr^{PTS} is built by Monte Carlo simulation, which implies that a candidate phase vector is considered for all subblocks, and the performance Δ is given by the difference of PAPR between the original symbol $\hat{\mathbf{x}}$ and the outcome symbol $\hat{\mathbf{x}}^{\text{PTS}}$, i.e.,

$$\Delta = \text{PAPR}(\hat{\mathbf{x}}) - \text{PAPR}(\hat{\mathbf{x}}^{\text{PTS}}). \quad (12)$$

In the backpropagation step, for any node n in Tr that belongs to the simulated path, the reward and expansion values are updated as follows:

$$Q(n) = Q(n) + \Delta, \quad (13)$$
$$E(n) = E(n) + 1. \quad (14)$$

After a predefined number of iterations, the child node with the largest average reward $Q(n)/E(n)$ of the current root r is returned as the best child.

For any search tree that returns the phase factor of subblock v, the computational budget of the corresponding exploration step is designed as follows:

$$g(v) = k(V - v) + W, \quad (15)$$

where $k \geq 1$ is the complexity parameter. The budget design ensures that the last subblock V can explore W phase factors, which is sufficient to find the optimal solution, while at the same time, maintains a low computational complexity. The number of phase vectors searched in the proposed M-PTS method is given by $\sum_{v=2}^{V} g(v) = (k/2)V^2 + (W - 3k/2)V - (W - k)$. The process of finding the optimal phase vector by the M-PTS method is summarized in Algorithm 2.

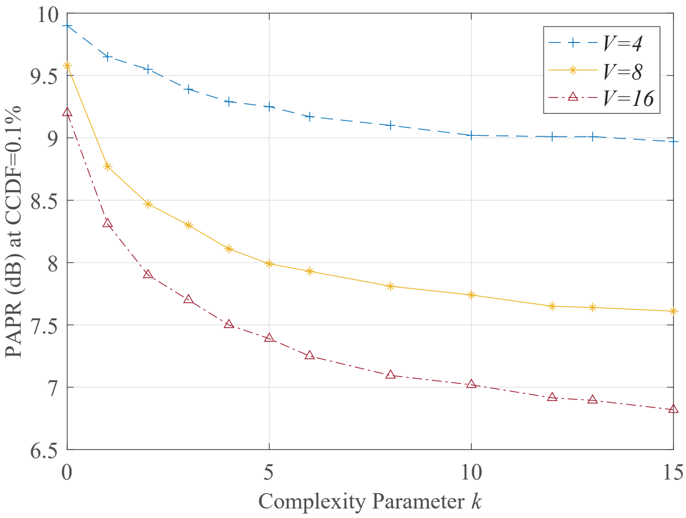

Fig. 3. PAPR of the M-PTS method at CCDF= 0.1% with QPSK modulation and different complexity parameter k.

C. Computational Complexity

Here, we use the number of complex multiplications to represent the computational complexity of different PTS methods. The C-PTS and I-PTS methods use all samples of the original OFDM signal to calculate the PAPR values. Thus, the number of complex multiplication required by each candidate phase vector is given by LN. Note that the number of candidate phase vectors of C-PTS and I-PTS are given by W^{V-1} and $W(V-1)$, respectively. We have the total number of complex multiplications are LNW^{V-1} and $LNW(V-1)$ for C-PTS and I-PTS, respectively. The RC-PTS method performs an exhaustive search on the dominant time domain samples to find the optimal phase vector, where the number of candidate phase vectors is W^{V-1} and each candidate phase vector requires $P_\gamma LN$ complex multiplications for PAPR calculation. Note that there are additional VLN complex multiplications for the calculation of metric Y_m. We have the overall number of complex multiplications for the RC-PTS method is given by $VLN + P_\gamma LNW^{V-1}$.

For the proposed M-PTS method, the number of candidate phase vectors is reduced to $\sum_{v=2}^{V} g(v)$. Note that the M-PTS method also uses dominant time domain samples for PAPR estimation. The number of complex multiplications required by each candidate phase vector is given by $P_\gamma LN$ and the additional multiplications for the calculation of metric Y_m is given by VLN. Therefore, the overall number of complex multiplications for the proposed M-PTS method is given by $VLN + P_\gamma LN \sum_{v=2}^{V} g(v)$.

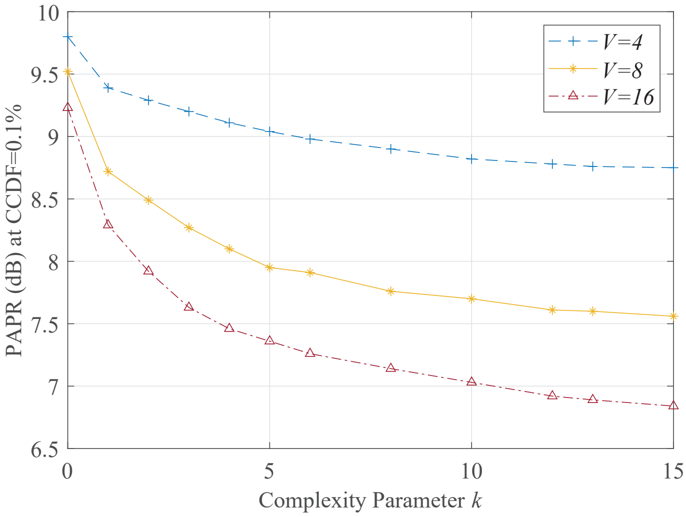

Fig. 4. PAPR of the M-PTS method at CCDF= 0.1% with 16-QAM modulation and different complexity parameter k.

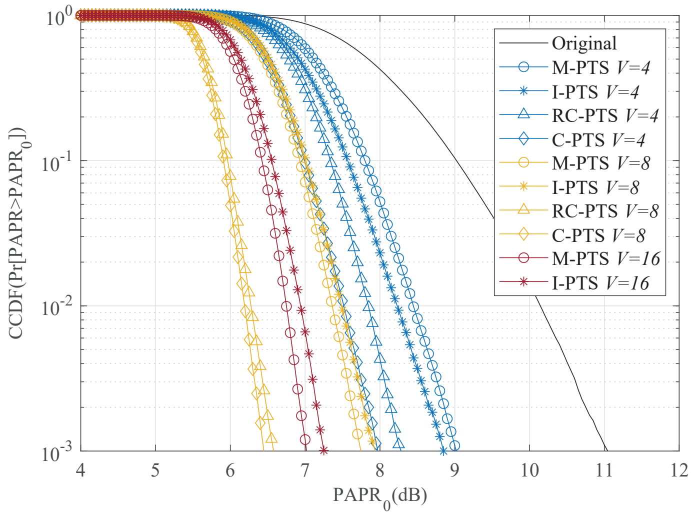

Fig. 5. CCDFs of different PTS methods with QPSK modulation.

TABLE I. PAPR at CCDF=0.1% and the corresponding computational complexity of different PTS methods with QPSK modulation.

Methods	$V = 4$		$V = 8$		$V = 16$	
C-PTS	7.99 dB	1	6.44 dB	2.6×10^{2}	-	1.7×10^{7}
RC-PTS	8.28 dB	9.8×10^{-2}	6.58 dB	9.1	-	5.9×10^{5}
I-PTS	8.85 dB	1.9×10^{-1}	7.94 dB	4.4×10^{-1}	7.25 dB	9.4×10^{-1}
M-PTS	9.02 dB	8.5×10^{-2}	7.74 dB	2.6×10^{-1}	7.02 dB	8.5×10^{-1}

IV. Numerical results

We compare our proposed M-PTS method with the C-PTS, I-PTS and RC-PTS methods. Specifically, we consider an OFDM system consists $N = 128$ subcarriers. Both the QPSK modulation and the higher order 16-QAM modulation are used. Each OFDM symbol is oversampled with oversampling ratio $L = 4$. For each subblock, the phase factor is selected from a set of $W = 4$ different values. For the RC-PTS method, we set $\gamma = 0.5$ to achieve a tradeoff between PAPR performance and computational complexity [13], and we have $P_\gamma = 0.035$. For the proposed M-PTS method, we also set $\gamma = 0.5$ and the weight factor is set as $c = 1/\sqrt{2}$ [19].

Fig. 3 shows the PAPR of the proposed M-PTS method using QPSK modulation at CCDF= 0.1% as a function of the complexity parameter k for subblock number $V = 4, 8, 16$. It shows that the 0.1% PAPR of the M-PTS method decreases with k and it flattens out when $k > 10$. The same trend can be seen in Fig. 4, where the OFDM symbols are 16-QAM modulated. This is because as the PAPR value decreases, it is getting more difficult to find a better phase vector that can achieve a lower PAPR. Thus, the marginal contribution of k decreases and the curves flatten out when k is large. In the following simulations, we set $k = 10$ so as to achieve a tradeoff between PAPR reduction performance and computational complexity.

The 0.1% PAPR of different PTS methods with QPSK modulation and their corresponding computational complexities are listed in Table I. Fig. 5 shows the complete CCDF curves of these PTS methods. The values of computational complexities are normalized by defining C-PTS with $V = 4$ as 1. The proposed M-PTS method always achieves the lowest computational complexity for any subblock number V. Note that due to the high computational complexity, it is impossible to run the C-PTS and RC-PTS methods with $V = 16$.

Compared with the I-PTS method, the proposed M-PTS method achieves 55%, 41% and 10% computational complexity reduction for $V = 4, 8$ and 16, respectively. When the subblock number is $V = 4$, the proposed M-PTS method suffers a 0.17 dB PAPR performance degradation. When $V = 8$ and $V = 16$, the M-PTS method achieves 0.2 dB and 0.23 dB PAPR performance improvement, respectively.

Compared with the RC-PTS method, the proposed M-PTS method can reduce the computational complexity by 1.2, 35 and 6.9×10^{5} times for $V = 4, 8$ and 16, respectively. While the corresponding cost is the degradation of PAPR performance, which are given by 0.74 dB and 1.16 dB for the subblock number $V = 4$ and $V = 8$, respectively.

Based on numerical results, we can conclude conservatively that our proposed M-PTS method can achieve an efficient tradeoff between PAPR performance and computational complexity. Specifically, if the computational complexity is limited below 1 (i.e., lower than the C-PTS method with $V = 4$), the M-PTS method achieves the lowest 0.1% PAPR 7.02 dB with $V = 16$, while the I-PTS, RC-PTS and C-PTS methods can only achieve 7.25 dB, 8.28 dB and 7.99 dB 0.1% PAPR by setting $V = 16$, 4 and 4, respectively. Other settings that can achieve lower PAPRs than 7.02 dB (i.e., RC-PTS and C-PTS with $V = 8$) will introduce a high computational complexity cost (i.e., 11 and 305 times increase) with only marginal PAPR reduction (i.e., 0.44 dB and 0.58 dB).

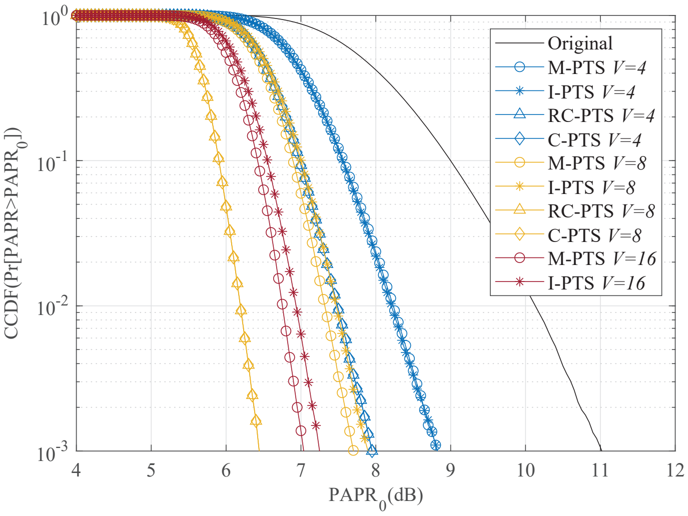

Fig. 6. CCDFs of different PTS methods with 16-QAM modulation.

Fig. 6 shows the CCDFs of different PTS methods with 16-QAM modulation. The performance of these methods is similar to that of QPSK modulation with only two differences. One is that the RC-PTS method achieves the same performance as the C-PTS method. However, it is still hard to employ the RC-PTS method since its computational complexity grows exponentially with the subblock number. The other is that the proposed M-PTS method has the same performance as the I-PTS method for $V = 4$. The results in Fig. 6 show that the M-PTS method still achieves an efficient tradeoff between PAPR performance and computational complexity when higher order modulation is used.

V. Conclusion

In this paper, we proposed an M-PTS method for OFDM PAPR reduction, which utilizes the MCTS process to search for the optimal phase vector. The proposed M-PTS method can achieve an efficient tradeoff between PAPR performance and computational complexity. Numerical results show that, compared with the RC-PTS method, the proposed M-PTS method significantly reduces the computational complexity at the cost of a small PAPR performance degradation. Also, it outperforms the I-PTS method in both PAPR performance and computational complexity when the subblock number is large.

References

[1] 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.

[2] S. Wang, W. Shi, and C. Wang, "Energy-efficient resource management in OFDM-based cognitive radio networks under channel uncertainty," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3092–3102, Sep. 2015.

[3] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for ofdm-based cognitive radio networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3181–3191, Aug. 2013.

[4] X. Sun and S. Wang, "Resource allocation scheme for energy saving in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4407–4416, Aug. 2015.

[5] J. Yli-Kaakinen, T. Levanen, A. Palin, M. Renfors, and M. Valkama, "Generalized fast-convolution-based filtered-OFDM: Techniques and application to 5G new radio," *IEEE Trans. Signal Process.*, vol. 68, pp. 1213–1228, Feb. 2020.

[6] J. Hou, X. Zhao, F. Gong, F. Hui, and J. Ge, "PAPR and PICR reduction of OFDM signals with clipping noise-based tone injection scheme," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 222–232, Jan. 2017.

[7] S. B. Slimane, "Reducing the peak-to-average power ratio of OFDM signals through precoding," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 686–695, Mar. 2007.

[8] M. S. Ahmed, S. Boussakta, A. Al-Dweik, B. Sharif, and C. C. Tsimenidis, "Efficient design of selective mapping and partial transmit sequence using T-OFDM," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2636–2648, Mar. 2020.

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

[10] H. Chen and K. Chung, "A low complexity PTS technique using minimal trellis in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 817–821, Jan. 2018.

[11] L. J. Cimini and N. R. Sollenberger, "Peak-to-average power ratio reduction of an OFDM signal using partial transmit sequences," *IEEE Commun. Lett.*, vol. 4, no. 3, pp. 86–88, Mar. 2000.

[12] S. H. Han and J. H. Lee, "PAPR reduction of OFDM signals using a reduced complexity PTS technique," *IEEE Signal Process. Lett.*, vol. 11, no. 11, pp. 887–890, Nov. 2004.

[13] S. Ku, C. Wang, and C. Chen, "A reduced-complexity PTS-based PAPR reduction scheme for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2455–2460, Aug. 2010.

[14] Y. Cho, K. Kim, J. Woo, K. Lee, J. No, and D. Shin, "Low-complexity PTS schemes using dominant time-domain samples in OFDM systems," *IEEE Trans. Broadcast.*, vol. 63, no. 2, pp. 440–445, Jun. 2017.

[15] K. Lee, H. Kang, and J. No, "New PTS schemes with adaptive selection methods of dominant time-domain samples in OFDM systems," *IEEE Trans. Broadcast.*, vol. 64, no. 3, pp. 747–761, Sep. 2018.

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

[17] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of Monte Carlo tree search methods," *IEEE Trans. Comput. Intell. AI Games*, vol. 4, no. 1, pp. 1–43, Mar. 2012.

[18] D. Silver, A. Huang *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.

[19] L. Kocsis, C. Szepesvári, and J. Willemson, "Improved monte-carlo search," *Univ. Tartu, Estonia, Tech. Rep*, vol. 1, 2006.

[20] Y. Lin, T. Wang, and S. Wang, "UAV-assisted emergency communications: An extended multi-armed bandit perspective," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 938–941, May 2019.