Constant Modulus Algorithm for Peak-to-Average Power Ratio (PAPR) Reduction in MIMO OFDM/A

Seyran Khademi, *Student Member, IEEE*, and Alle-Jan van der Veen, *Fellow, IEEE*

***Abstract*—A new peak-to-average power ratio (PAPR) reduction approach for MIMO-OFDM/A is developed based on the well-known constant modulus algorithm (CMA). This combines two ideas: 1) time domain signals from "resource blocks" (consisting of several subcarriers) may be linearly combined using precoding weights, transparent to the receiver; 2) the precoding weights can be designed to minimize the modulus variations of the resulting signal, leading generally to a reduction in PAPR. This technique is compatible with various beamforming modes in single antenna and MIMO systems. Simulation results show a noticeable improvement relative to the Partial Transmit Sequences (PTS) technique with significantly less complexity.**

***Index Terms*—Beamforming, convex optimization, multiple input multiple output (MIMO), orthogonal frequency division multiplexing (OFDM), partial transmit sequence (PTS).**

I. Introduction

OFDM is known as one of the most favorable modulation techniques for communication over frequency selective wireless channels, and is widely used in telecommunication standards. A well-known drawback of OFDM is that the amplitude of the time domain signal varies strongly with the transmitted symbols modulated on the subcarriers in the frequency domain, resulting in a 'peaky' signal. If the maximum amplitude of the time domain signal is too large, it pushes the transmit amplifier into a non-linear region which distorts the signal resulting in a substantial increase in the error rate at the receiver. Over the past decade, an extensive amount of literature has been dedicated to Peak to Average Power Ratio (PAPR) reduction techniques. These techniques are associated with costs in terms of bandwidth or/and transmit power. Also, most of them require modifications to both the transmitter and the receiver which makes them non-compliant to existing standards. Multiple signal representation methods, such as PTS and selected mapping (SLM) are among the most cited techniques [1], [2]. Extension of these algorithms to multiple antenna (MIMO) systems is not straightforward. Another combined precoding and PAPR reduction technique has been proposed for multiuser MIMO systems with sorted Tomlinson-Harashima precoding (sTHP). For more details and further developed techniques on MIMO-OFDM peak reduction see [3] and references therein.

A new technique called CP-PTS is proposed in [4] which is adaptable for different beamforming schemes in standard point to point or multiuser MIMO systems. In this technique, the OFDM subcarriers are grouped into blocks and the phase of each block is changed in a manner similar to the PTS method but without the drawback of sending explicit side information. As long as each block is multiplied with only one phase coefficient, the receiver will perceive this as a channel effect and will compensate for it during the channel equalization process [5]. An extension of CP-PTS to MIMO-OFDM systems is introduced in[6]. In both cases, a sequential quadratic programming (SQP) algorithm is used to solve the phase optimization problem. The computational complexity of this algorithm can be prohibitive for high data rate and/or low latency communication links. The PAPR weights need to be determined again for every OFDM data block, hence the underlying algorithm should be sufficiently efficient to enable a real-time processing.

In this letter, the same configuration as CP-PTS is used but instead of solving a non-convex optimization problem, an alternative problem formulation is proposed based on a cost function used in constant modulus algorithms (CMAs). Accordingly, the block-iterative SDCMA algorithm [7] is used to find the precoding PAPR weights. The resulting computational complexity is linear in the number of subcarriers. Furthermore, to make sure that the BER performance of the system is not affected by the PAPR precoding an additional constraint is appended to the CMA objective function which requires the weights to be on the unit circle. Like CP-PTS, the proposed technique is transparent to the receiver; this means that it only affects the base station (BS) and it does not require any signal processing in the mobile station (MS).

The proposed method does not function if the channel estimation exploits the smooth changes of the channel coefficients over the complete OFDM block. However, this assumptions is not valid in the modern multiuser systems based on RB assignment [5], [8].

II. Transmit Signal Model

Similar to [6] we consider a generic MIMO-OFDM/A downlink scenario with one base station (BS) employing M_t antennas. An OFDM block with N subcarriers is transmitted from each antenna. The N subcarriers include N_u useful subcarriers surrounded by two guard bands with zero energy. The useful subcarriers are further grouped into M resource blocks (RBs) each consisting of $N_b = N_u/M$ subcarriers. Data of one or more users is placed in these RBs and mapped into the space-time domain using an inverse discrete Fourier transform (IDFT) and space-time block coding (STBC). To allow channel estimation at the receivers (mobile stations), each RB also contains several pilot subcarriers that act as training symbols. The

Manuscript received January 09, 2013; revised March 04, 2013; accepted March 08, 2013. Date of current version April 04, 2013. This work was supported in part by the STW under Contract 10551 (FASTCOM). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Richard K. Martin.

The authors are with the Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, Delft, The Netherlands (e-mail: s.khademi@tudelft.nl, a.j.vanderveen@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2013.2254114

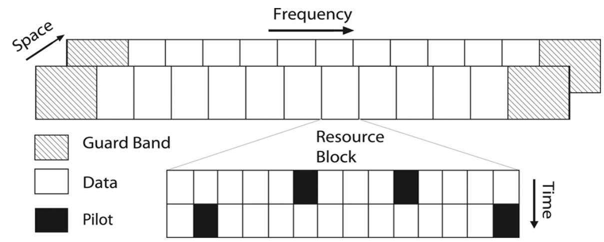

Fig. 1. Data structure of an OFDM block for a MIMO-OFDM/A downlink.

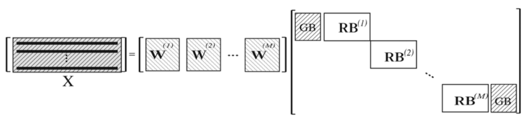

Fig. 2. Beamformed MIMO transmit data in frequency domain.

transmit signal model is illustrated in Fig. 1. It is compatible with the WiMAX standard [8].

Let us first describe the MIMO transmit data model in the frequency domain; for simplicity we consider only a single time block from now on. The data in the q-th RB is a matrix $\mathbf{D}^{(q)} \in \mathbb{C}^{M_t \times N_b}$, it is premultiplied with a corresponding beamforming matrix $\mathbf{W}^{(q)} \in \mathbb{C}^{M_t \times M_t}$, $q = 1, \cdots, M$, resulting in transmit sequences $\mathbf{X}^{(q)} = \mathbf{W}^{(q)H}\mathbf{D}^{(q)}$. Together with guard intervals, they are collected in a matrix $\mathbf{X} \in \mathbb{C}^{M_t \times N}$, where the M_t rows of this matrix represent the N symbols to be transmitted from the M_t antennas. The data model is

$$\mathbf{X} = \mathbf{W}^H\mathbf{D}, \tag{1}$$

where $\mathbf{W} = [\mathbf{W}^{(1)H}, \cdots, \mathbf{W}^{(M)H}]^H$, and $\mathbf{D} \in \mathbb{C}^{MM_t \times N}$ is a block-diagonal matrix with structure as in Fig. 2, which includes guard intervals as well. Matrix $\mathbf{X}$ represents the spatial data in the frequency domain.

The time-domain MIMO-OFDM transmit data model is obtained by taking the IDFT of the beamformed data matrix $\mathbf{X}$, resulting in

$$\mathbf{Y} = \mathbf{X}\mathbf{F}^H = \mathbf{W}^H\mathbf{D}\mathbf{F}^H, \tag{2}$$

where $\mathbf{F}^H \in \mathbb{C}^{N \times N}$ denotes the IDFT matrix, and $\mathbf{Y} \in \mathbb{C}^{M_t \times N}$ contains the resulting transmit OFDM sequences for each of the M_t antennas. Let us further denote the time-domain data matrix $\mathbf{B} = \mathbf{D}\mathbf{F}^H$; this is a full matrix. Accordingly, the beamformed OFDM block can be expressed as

$$\mathbf{Y} = \mathbf{W}^H\mathbf{B}. \tag{3}$$

Denote the total power (or energy) in the data matrix $\mathbf{D}$ by $P_d := \|\mathbf{D}\|_F^2 = \|\mathrm{vec}(\mathbf{D})\|^2 =: \alpha N_t$, where $N_t = NM_t$.Function $\mathrm{vec}(\mathbf{D})$ creates a column vector whose elements are the columns of the matrix $\mathbf{D}$. N_t is the total number of subcarriers or samples to be sent from all M_t antennas, and α is defined as the average transmit power per sample (including the zero power guard bands). If we assume that the beamforming matrix $\mathbf{W}$ consists of *orthonormal* matrices $\mathbf{W}^{(q)}$, then applying beamforming and the IDFTdoes not change the total transmit power.

III. Proposed Precoding Scheme

The IDFT operation in (2) leads to a large dynamic range of the resulting time-domain OFDM signal. PAPR is a common metric to measure the distortion caused by probable high peak of the OFDM signal and for a MIMO-OFDM block $\mathbf{Y}$ we define

$$\mathrm{PAPR}(\mathbf{Y}) = \frac{\alpha N_t \|\mathrm{vec}(\mathbf{Y})\|_\infty^2}{\|\mathrm{vec}(\mathbf{Y})\|_2^2}. \tag{4}$$

Clearly, the lowest PAPR is achieved for a constant modulus (CM) signal, for which the infinity norm is equal to the average power of the sequence.

The main idea in [4], [6] is to design a precoding matrix to transform the OFDM symbols in $\mathbf{Y}$ to a favorable signal $\mathbf{S}$ with lower PAPR (ideally a CM signal). This precoding matrix $\mathbf{\Omega}$ needs to fulfill the following requirements:

1) Reduce the dynamic range of the OFDM block,
2) Preserve the beamforming property,
3) Be transparent to the receiver,
4) Not impact the bit error rate (BER).

To satisfy the second and third constraint, we are allowed to premultiply each RB, $\mathbf{D}^{(q)}$, with a diagonal scaling matrix $\mathbf{\Omega}^{(q)}$. To the receiver, this will appear as a fading channel effect. To not affect the BER, the scaling should be unimodular (phase only). Equivalently, a diagonal (unimodular) precoding matrix $\mathbf{\Omega} \in \mathbb{C}^{MM_t \times MM_t}$ is applied to $\mathbf{D}$. The resulting MIMO-OFDM transmit matrix (replacing $\mathbf{Y}$) is

$$\mathbf{S} = \mathbf{W}^H\mathbf{\Omega}\mathbf{D}\mathbf{F}^H. \tag{5}$$

If we define $\boldsymbol{\omega} = \mathrm{vecdiag}(\mathbf{\Omega})$, then the PAPR reduction problem is to design $\boldsymbol{\omega}$ as

$$\min_{\boldsymbol{\omega}} \|\mathrm{vec}(\mathbf{S})\|_\infty^2 \quad \text{s.t.} \quad \|\mathrm{vec}(\mathbf{S})\|_2^2 = P \tag{6}$$

where $P = \alpha N_t$ is a fixed total transmit power. This problem is not convex because nonlinear equality constraints can rarely be expressed in a convex form. The approach in [4], [6] was to solve a series of quadratic convex subproblems iteratively. Although this does not solve the original problem in (6) exactly, the results were excellent compared to other techniques, and attractive as the method is transparent to the receiver and does not distort the transmit signals. Unfortunately, this approach is yet too complex for real time applications.

IV. Proposed CMA Approach

A. Formulation as a Constant Modulus Problem

Using properties of Kronecker products, we can rewrite $\mathbf{S}$ in (5) as

$$\mathbf{s} = \mathrm{vec}(\mathbf{S}) = (\bar{\mathbf{B}} \circ \mathbf{W})^H \mathrm{vecdiag}(\mathbf{\Omega}) =: \mathbf{A}\boldsymbol{\omega}, \tag{7}$$

where $\mathbf{A} \in \mathbb{C}^{N_t \times MM_t}$, $\mathbf{D}\mathbf{F}^H = \mathbf{B} \in \mathbb{C}^{MM_t \times N}$, $\bar{\mathbf{B}}$ denotes the complex conjugate of $\mathbf{B}$, and $\circ$ denotes the Khatri-Rao product (column-wise Kronecker product). The $\mathrm{vecdiag}(\mathbf{D})$ creates a column vector whose elements are the main diagonal of the matrix $\mathbf{D}$. The optimization problem (6) becomes

$$\min_{\boldsymbol{\omega}} \|\mathbf{A}\boldsymbol{\omega}\|_\infty^2 \quad \text{s.t.} \quad \|\mathbf{A}\boldsymbol{\omega}\|_2^2 = \alpha N_t \tag{8}$$

We now propose an alternative formulation of this problem, by replacing the infinity norm by the average deviation of the OFDM block from a constant modulus signal. Ideally, the resulting $\mathbf{S}$ will be close to a CM signal, and hence have close-to-optimal PAPR. The corresponding cost function is

$$J(\boldsymbol{\omega}) = \left\| \mathbf{A}\boldsymbol{\omega} \odot \overline{(\mathbf{A}\boldsymbol{\omega})} - \alpha \mathbf{1}_{N_t} \right\|_2^2 = \sum_{n=1}^{N_t} \left(\boldsymbol{\omega}^H \mathbf{a}_n \mathbf{a}_n^H \boldsymbol{\omega} - \alpha \right)^2.$$

Here, the vector $\mathbf{a}_n^H$, $n = 1, \cdots, N_t$ represents the n-throw of matrix $\mathbf{A}$, the column vector $\mathbf{1}_{N_t}$ is a vector with all entries equal to 1 and dimension N_t, and $\odot$ denotes the Schur-Hadamard product (pointwise multiplication).

This formulation is similar to the well-known "CMA(2,2)" cost function for adaptive blind equalization or blind beamforming, and can be solved efficiently using available iterative algorithms. The matrix $\mathbf{A}$ plays the role of the data matrix in the usual CMA context, whereas $\boldsymbol{\omega}$ plays the role of the beamforming vector. The original CMA cost function is expressed in terms of an expectation operator; the present "deterministic" formulation is similar to the Steepest Descent CMA (SDCMA) in [7].

B. Steepest-Descent CMA (SDCMA)

The SDCMA is a block-iterative algorithm in which we act on the full data matrix $\mathbf{A}$ and update $\boldsymbol{\omega}$ until it converges. The derivation of the block SDCMA is straightforward when the statistical expectation in original formula in [7] is replaced by an average over a block. For the i-th iteration, we start from the current estimate $\boldsymbol{\omega}^i$ and compute:

$$\hat{\mathbf{s}}^i = \mathbf{A}\boldsymbol{\omega}^i \tag{9}$$

$$\mathbf{e}^i = (\hat{\mathbf{s}}^i \odot \bar{\hat{\mathbf{s}}}^i) - \alpha \mathbf{1}_{N_t} \tag{10}$$

$$\hat{\mathbf{s}}_e = \hat{\mathbf{s}}^i \odot \mathbf{e}^i \tag{11}$$

$$\boldsymbol{\omega}^{i+1} = \boldsymbol{\omega}^i - \mu \nabla J(\boldsymbol{\omega}^i) = \boldsymbol{\omega}^i - \mu \mathbf{A}^T \hat{\mathbf{s}}_e. \tag{12}$$

Here, μ is a suitable step size, and $\hat{\mathbf{s}}_e$ is the update error. The maximal step size μ could be defined as a scale independent parameter in relation to the signal power in $\mathbf{A}$. To keep the solution unchanged as $\mathbf{A}$ scales, μ needs to be divided by factor α^2, $\mu = \mu'/\alpha^2$. For convergence, the algorithm is initialized with $\boldsymbol{\omega}^0 = \mathbf{1}$ (although other choices are possible). The algorithm should be run until the cost function $J(\boldsymbol{\omega})$converges; in practice convergence is fast and the algorithm is run for a fixed small number of iterations.

To satisfy the power constraint in (6), we can simply scale the resulting $\boldsymbol{\omega}$ after convergence. If $\mathbf{s} = \mathbf{A}\boldsymbol{\omega}$ is indeed a constant modulus signal, then $\|\mathbf{s}\|_2^2 = \alpha N_t$, and the power constraint is inherently already satisfied. Thus, the scaling is expected to be close to 1 and could be omitted in practice (it has no effect on the cost function $\mathrm{PAPR}(\mathbf{S})$).

A difference with the standard CMA is that, here, a good solution does not necessarily exist. The usual application of CMA is for a linear combination of constant modulus sources for which, without noise, a perfect beamformer exists. The present situation could be said to correspond to a very noisy source separation situation. Note that, also for other methods, there are no existence results for PAPR reduction.

C. Unit-Circle CMA (UC-CMA)

In SDCMA, the computed $\boldsymbol{\omega}$ has no constraints and may have some small entries. These are equivalent to a (broad) null in the channel which will affect the BER performance. Ideally, we should restrict the entries of $\boldsymbol{\omega}$ to take only unimodular values: $\omega_m = e^{j\phi_m}$, $m = 1, \cdots, M$, and add this constraint to the optimization problem (8).

In order to restrict the solution to be on the unit circle, a normalization step is added to each iteration after(12):

$$\boldsymbol{\omega}^{i+1} = \boldsymbol{\omega}^{i+1} \oslash |\boldsymbol{\omega}^{i+1}| \tag{13}$$

where $\oslash$ denotes pointwise division, and $|\cdot|$ takes the absolute value of each entry of the vector argument. This alternative updating algorithm is called Unit Circle CMA (UC-CMA) since (13) projects the solution of CMA to a unit circle at each iteration.

D. Computational Complexity

The complexity of the SDCMA algorithm in (12) is dominated by the matrix products $\mathbf{A}^T\hat{\mathbf{s}}_e$ and $\mathbf{A}\boldsymbol{\omega}^i$. The resulting complexity is approximately $2NMM_t^2$ per iteration (linear in the number of subcarriers). UC-CMA has the same complexity.

In conventional PTS [2], each RB (sub-block in PTS context) is weighted with a phase shift in such a way that the summation of sub-blocks produce an OFDM sequence with a smaller PAPR [2]. The phase weights are selected by an exhaustive search among a discrete set of phases, and are sent as side information to the receiver. Accordingly, all combinations of the M available phase weights need to be calculated and then multiplied with an IDFT summation matrix, which has the same size as matrix B. Finally, one sequence with the least PAPR metric is chosen with the corresponding phase weights. The complexity of the exhaustive search is calculated for the simplest set of only two phases $\{\pm 1 = e^{\pm j\pi/2}\}$ and M RBs as $2^M NM$ multiplications and 2^M comparisons. For CP-PTS, the complexity is $O(M^3)$; the exact expression for complexity is derived in [4].

V. Simulation Results

In WiMAX, one RB spans $N_b = 14$ sub-carriers over two OFDM symbols in time, containing 4 pilots and 24 data symbols. For a 10 MHz system, there are a total of $M = 60$ RBs [8]. In agreement with this WiMAX setting, the proposed PAPR reduction technique is simulated for an OFDM block of size $N = 1024$ including $MN_b = 840$ data subcarriers with QPSK modulation and 92 guard subcarriers at each end of the band. The number of MIMO transmit antennas is either $M_t = 1$, 2 or 4, as will be indicated. The various techniques are evaluated using the complementary cumulative density function (CCDF), which denotes the probability that the PAPR of a data block exceeds the argument of the function. To avoid the PAPR underestimation, The algorithm is run with four times oversampling so the number of the samples processed in the simulations is $N' = 4N_t$.

A total number of 10,000 OFDM blocks are randomly generated to produce the CCDF curves. For each block, a random complex fading channel is generated, and the beamforming matrices $\mathbf{W}$ are chosen as the right singular vectors of these channel matrices.

In Fig. 3, the CCDF performance is shown for SDCMA (various number of iterations), UC-CMA (50 iterations), and compared to CP-PTS [6] and the standard PTS [2]. The latter algorithm is simulated only for $M = 10$ RBs due to prohibitive computational complexity for larger M. In this simulation, $M_t = 1$ transmit antenna. The simulations show that the proposed techniques attain a PAPR reduction of up to 6 dB. Although 50 iterations are sufficient for good performance, another 0.5 dB is gained by increasing this to 500 iterations. UC-CMA (50 iterations) is worse by about 0.5 dB. The PAPR reduction for PTS is worse by 1 to 2 dB. The previously proposed CP–PTS outperforms PTS and SDCMA with 50 iterations, however a similar gain is reached by SDCMA with a larger number of iterations.

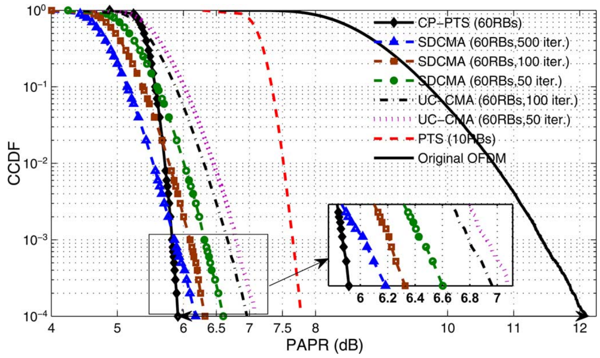

Fig. 3. Performance comparison for the proposed CMA PAPR reduction algorithm for various number of iterations and $\mu' = 0.05$, CP-PTS with 5 iterations, and conventional PTS with phase alphabet $\{\pm 1\}$ and $M = 10$.

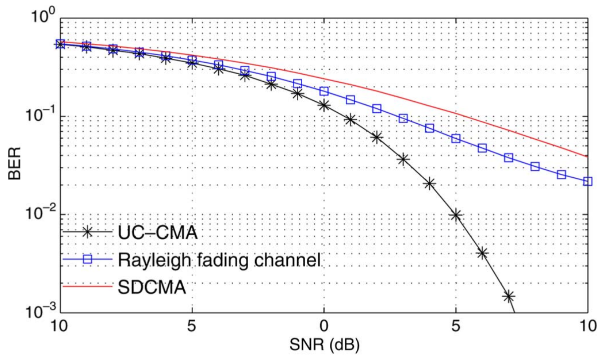

Fig. 4. BER performance of the proposed algorithms in comparison with AWGN and Raleigh fading channels for single antenna QPSK–OFDM system of size $N = 1024$ and $M = 60$.

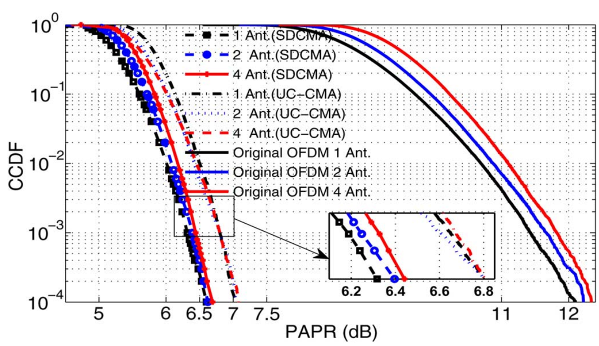

Fig. 5. PAPR reduction performance in MIMO-OFDM for both SDCMA and UC-CMA with 50 iterations and $\mu' = 0.05$.

Moreover, the CCDF curves in Fig. 3 show the superior performance of SDCMA in 90% and 99.9% of OFDM blocks in 50 and 500 iterations respectively, comparing to the CP-PTS.

The empirical CDF of $|\omega_q|$ values in SDCMA indicates the Rayleigh distribution of PAPR weights which affect the BER performance of the system. Fig. 4 shows the BER versus SNR curves for the QPSK-OFDM system without PAPR reduction in a randomly generated Rayleigh fading and AWGN channels compared to the scenarios that SDCMA and UC-CMA weights are applied at the transmitter. In SDCMA and UC-CMA the channel is assumed to be AWGN and the received vector is divided by $\boldsymbol{\omega}$ to equalize the PAPR weights. Where, in Rayleigh fading channel the received vector is divided by the frequency domain channel coefficients. in both cases, the perfect channel recovery is assumed. From Fig. 4, the effect of non-modified SDCMA is analogous to a Rayleigh fading channel in terms of BER performance so the same error correcting codes used for a fading channel can be applied here. As expected the UC-CMA does not influence the BER performance. This motivates the use of UC-CMA technique.

Fig. 5 shows the performance of SDCMA and UC-CMA for various number of transmit antennas, $M_t = 1, 2, 4$, and 50 iterations. It is seen that the performance is not a strong function of the number of antennas; small improvements are seen due to more available phase weights or degrees of freedom in the optimization problem.

To demonstrate computational complexity, Matlab runtimes on a standard 2011 laptop are shown in Fig. 6 as a function of M (number of RBs). In this simulation, $M_t = 1$. It is seen that the proposed CMA algorithms (using 50 iterations) are about a factor 50 faster than CP-PTS, whereas the complexity of PTS is growing exponentially with the number of RBs and is quickly not feasible anymore.

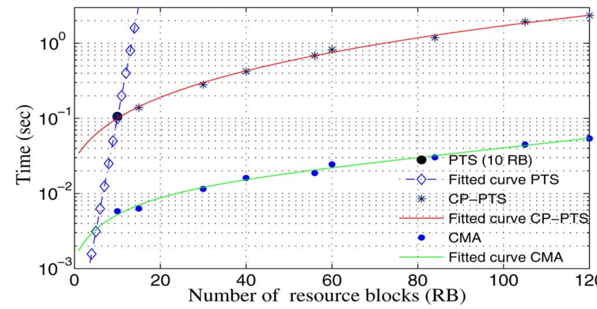

Fig. 6. Complexity comparison between CMA, CP-PTS and PTS using Matlab runtime evaluation.

References

[1] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun. Mag.*, vol. 12, pp. 56–65, Apr. 2005.

[2] S. Muller and J. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electron. Lett.*, vol. 33, pp. 368–369, Feb. 1997.

[3] C. Siegl and R. F. Fischer, "Selected basis for PAR reduction in multi-user downlink scenarios using lattice-reduction-aided precoding," *EURASIP J. Adv. Signal Process.*, Jul. 2011.

[4] S. Khademi, T. Svantesson, M. Viberg, and T. Eriksson, "Peak-to-average-power-ratio (PAPR) reduction in WiMAX and OFDM/A systems," *EURASIP J. Adv. Signal Process.*, Aug. 2011.

[5] Y. Shen and E. Martinez, "WiMAX channel estimation: Algorithms and implementations," *Application Note, Freescale*, 2007.

[6] S. Khademi, A.-J. van der Veen, and T. Svantesson, "Precoding technique for peak-to-average-power-ratio (PAPR) reduction in MIMO OFDM/A systems," in *, 2012 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2012, pp. 3005–3008.

[7] J. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, pp. 459–471, Apr. 1983.

[8] J. Wang, Z. Lan, R. Funada, and H. Harada, "On scheduling and power allocation over multiuser MIMO-OFDMA: Fundamental design and performance evaluation WiMAX systems," in *IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, 2009, pp. 2752–2756.