Peak-to-Average Power Ratio Reduction in OFDM Systems Using an Adaptive Differential Evolution Algorithm

Saurav Ghosh[1], Subhrajit Roy[1], Swagatam Das[1], Ajith Abraham[2, 3] and Sk. Minhazul Islam[1]

[1]Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata 700 032, India
[2]Faculty of Computer Science and Electrical Engineering, VSB – Technical University of Ostrava, Czech Republic
[3]Machine Intelligence Research Labs (MIR Labs), Seattle, WA, USA
ajith.abraham@ieee.org

Abstract—**Orthogonal Frequency Division Multiplexing (OFDM) has emerged as very popular wireless transmission technique in which digital data bits are transmitted at a high speed in a radio environment. But the high peak-to-average power ratio (PAPR) is the major setback for OFDM systems demanding expensive linear amplifiers with wide dynamic range. In this article, we introduce a low-complexity partial transmit sequence (PTS) technique for diminishing the PAPR of OFDM systems. The computational complexity of the exhaustive search technique for PTS increases exponentially with the number of sub-blocks present in an OFDM system. So we propose a modified Differential Evolution (DE) algorithm with novel mutation, crossover as well as parameter adaptation strategies (MDE_*p*BX) for a sub-optimal PTS for PAPR reduction of OFDM systems. MDE_*p*BX is utilized to search for the optimum phase weighting factors and extensive simulation studies have been conducted to show that MDE_*p*BX can achieve lower PAPR as compared to other significant DE and PSO variants like JADE, SaDE and CLPSO.**

Keywords-Peak-to-average power ratio; Partial transmit sequence; Differential Evolution; p-best crossover; Parameter adaptation

I. Introduction

Orthogonal frequency division multiplexing (OFDM) [1] is basically a Multi-Carrier Digital Modulation technique in which a transmitted signal (serial digital data stream) is split up into multiple parallel data streams, which are modulated onto multiple adjacent carriers (subcarriers) within the allotted bandwidth. This multiplexing method is renowned for giving rise to high-speed digital data transmission. But the major defect regarding OFDM systems is the high peak-to-average power ratio (PAPR). Many PAPR reduction approaches have already been proposed such as clipping [2] and peak windowing, block coding [3], scrambling [4], nonlinear commanding transform schemes [5,6]. The most effective and efficient PAPR reduction method is the partial transmit sequence (PTS) approach in which the signal sub-blocks obtained by splitting the input data block are multiplied by phase weighting factors and then summed up to produce an alternative transmit without any loss of information. The phase weighting factors should be selected in a way such that the ultimate PAPR is minimized. But the computational overhead associated with exhaustive search of the ordinary PTS technique increases exponentially with number of sub-blocks, so the exhaustive search method becomes rigorous, time-consuming and thus practically unrealizable for large number of sub-blocks.

With the arrival of evolutionary computation techniques researchers have opted for designing suitable evolutionary algorithms to replace the inefficient exhaustive search method. The Differential Evolution algorithm [16, 19] has emerged has a very competitive form of evolutionary computing more than a decade ago. The computational steps employed by a DE algorithm are similar in spirit to any standard Evolutionary Algorithm (EA). However, unlike the traditional EAs, DE-variants perturb the current-generation population members with the scaled differences of randomly selected and distinct population members. Therefore, no separate probability distribution has to be used for generating the offspring. The success of DE was demonstrated at the First International Contest on Evolutionary Optimization in May 1996, which was held in conjunction with the 1996 IEEE International Conference on Evolutionary Computation (CEC) [18]. DE finished third at the First International Contest on Evolutionary Optimization (1st ICEO), which was held in Nagoya, Japan. DE turned out to be the best evolutionary algorithm for solving the real-valued test function suite of the 1st ICEO (the first two places were given to non-evolutionary algorithms, which are not universally applicable but solved the test-problems faster than DE). Since the late 1990s, DE started to find several significant applications to the optimization problems arising from diverse domains of science and engineering.

In this article, we propose new mutation and crossover operators for DE as well as a simple but effective scheme of adapting two of its most important control parameters with an objective of achieving improved PAPR reduction performance in OFDM systems. We shall refer to this new adaptive DE algorithm as MDE_*p*BX (Modified DE with *p*-best Crossover). The MDE_*p*BX algorithm is applied to search for the optimal combination of phase weighting factors so as to reduce the PAPR of the OFDM signal. Numerous simulations have been done to show that MDE_*p*BX can bring about better PAPR reduction compared to different sub-optimum PTS techniques such as Iterative PTS technique (IPTS) and various DE and PSO variants, namely an adaptive DE with current-to-*p*best/1 mutation scheme and optional external archive (JADE) [7],

978-1-4244-7833-0/11/$26.00 ©2011 IEEE

Self-Adaptive DE (SaDE) [8] and Comprehensive Learning Particle Swarm Optimizer (CLPSO) [9].

The rest of the paper is organized as follows. Section II gives a description of PAPR in OFDM systems and PTS scheme. The basic differential evolution algorithm and the proposed MDE_*p*BX algorithm is discussed in sections III and IV in details. Sections V and VI describes the MDE_*p*BX-based PTS technique and the Iterative PTS (IPTS) technique. Section VII compares the computational complexity of different sub-optimum PTS techniques. Simulations and results are discussed in section VIII and section IX concludes the paper.

II. DESCRIPTION OF PAPR IN OFDM SYSTEM AND PTS SCHEME

A. *Definition of Peak-to-average power ratio(PAPR)*

For OFDM system implementation, Inverse Fast Fourier Transform (IFFT) is usually being utilized to modulate multiple sub-band signals in an OFDM signal. The complex modulating (baseband) signal in an OFDM system consisting of N subcarriers can be expressed as

$$X(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_n \exp\left(j2\pi n \Delta f t\right), \quad 0 \le t \le NT \qquad (1)$$

where $X(t)$ is the transmitted signal, X_n is the data symbol at n[th] subcarrier, $j = \sqrt{-1}$ is the complex symbol and $\Delta f \left(= \frac{1}{T}\right)$ is the subcarrier frequency spacing. The PAPR of $X(t)$ is outlined as

$$PAPR = \frac{\max\limits_{0 \le t \le T} |X(t)|^2}{E\left[|X(t)|^2\right]} \qquad (2)$$

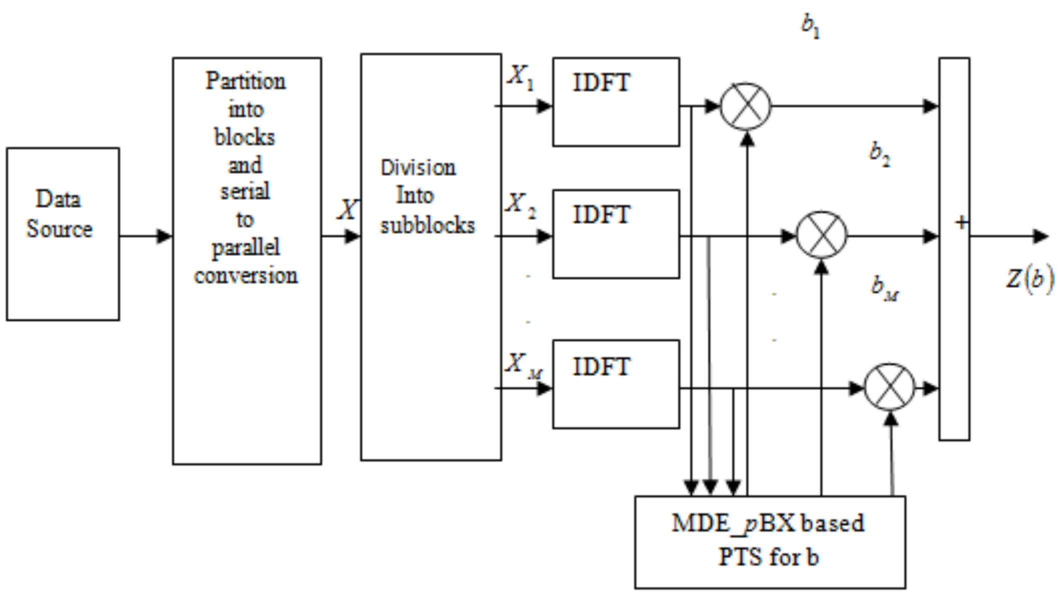

Figure 1: The structure of an OFDM transmitter with MDE_*p*BX-based PTS scheme (IDFT stands for Inverse Discrete Fourier Transform)

where $\max|X(t)|^2$ is the peak power of the OFDM signal and $E[\bullet]$ is the average power. Nowadays, most systems are dealing with discrete time signals and thus many PAPR reduction techniques are concerned with the amplitude of samples of $X(t)$.

B. *The Partial Transmit Sequence(PTS) scheme*

In the PTS scheme of PAPR reduction, a data block of N symbols fed as input is split into disjoint sub-blocks. The subcarriers in each sub-block are multiplied by a phase weighting factor of that sub-block. The principle structure of PTS scheme is shown in Figure 1 as that in [11]. The phase weighting factors are chosen in such a manner that the PAPR of the entire signal is minimized. So PAPR reduction is basically an optimization process. The input data block is represented as a vector $\vec{X} = [X_1, X_2, \ldots\ldots, X_N]^T$. The vector X is segmented into M separate sub-blocks shown as $\vec{X} = \sum_{i=1}^{M} \vec{X}_i$. There is a postulate that the sub-blocks are of equal size whose weighted sum combination can be represented as

$$Z(b) = \sum_{i=1}^{M} b_i \vec{X}_i \qquad (3)$$

Our goal is to search for the optimal phase weighting vector $\vec{b} = [b_1, b_2, \ldots., b_i]$ that will minimize the PAPR of $Z(b)$.

The optimal phase weighting factor b_i can be derived from a extensive matching of all possible

b^{M-1} sequences. But the exhaustive search will be extremely time-consuming and thus we will employ the optimization process on the cost function defined as

$$\hat{b} = \arg\min_{b} \left\{ \max \left| \sum_{i=1}^{M} b_i \vec{X}_i \right| \right\} \quad (4)$$

The amount of PAPR reduction is proportional to the number of phase weighting factor. If the number of phase weighting factor is large, the number of parallel addition processor and the number of phase weighting factor sequences are searched to find the optimum combination of phase weighting factors will be increased incorporating huge complexity in the system. Then we cannot assume that the candidate signals are independent in PTS. The correlation among candidate signals deteriorates the PAPR reduction performance in PTS. The correlation among candidate signals is governed by two factors-one is the sub-block partition style as described in [6] and the other is the value of phase weighting factor set. So it is possible to alter these two factors to produce candidate signals with diminished correlation, so as to prevent degradation of the PAPR reduction performance.

In this paper, an innovative sub-blocks partition scheme is proposed and its performance is assayed. The scheme is somewhat a concatenation of pseudo-random and interleaved partition schemes. In the proposed technique, sub-bands in an OFDM signal are partitioned into multiple separate sub-blocks and signals are assigned randomly in a partial sub-band of each sub-block. These partial sub-bands signals are replicated and designated to the remaining sub-bands repetitively to yield a complete sub-block.

III. CLASSICAL DIFFERENTIAL EVOLUTION

DE is a simple real-coded evolutionary algorithm. It works through a simple cycle of stages, which are detailed below.

A. *Initialization of the Parameter Vectors*

DE [16, 19] searches for a global optimum point in a D-dimensional continuous hyperspace. It begins with a randomly initiated population of NP D dimensional real-valued parameter vectors. Each vector, also known as *genome/chromosome*, forms a candidate solution to the multi-dimensional optimization problem. We shall denote subsequent generations in DE by $G = 0,1...,G_{\max}$. Since the parameter vectors are likely to be changed over different generations, we may adopt the following notation for representing the i-th vector of the population at the current generation:

$$\vec{X}_{i,G} = [x_{1,i,G}, x_{2,i,G}, x_{3,i,G}, \ldots\ldots, x_{D,i,G}]. \quad (5)$$

The initial population (at $G = 0$) should cover the entire search space as much as possible by uniformly randomizing individuals within the search space constrained by the prescribed minimum and maximum bounds: $\vec{X}_{\min} = \{x_{1,\min}, x_{2,\min}, \ldots, x_{D,\min}\}$ and $\vec{X}_{\max} = \{x_{1,\max}, x_{2,\max}, \ldots, x_{D,\max}\}$.

Hence we may initialize the j-th component of the i-th vector as:

$$x_{j,i,0} = x_{j,\min} + rand_{i,j}[0,1] \cdot (x_{j,\max} - x_{j,\min}), \quad (6)$$

where rand is a uniformly distributed number lying within the range $[0,1]$ and is instantiated independently for each component of the i-th vector.

B. *Mutation with Difference Vectors*

After initialization, DE creates a *donor* vector $\vec{V}_{i,n}$ corresponding to each population member or *target* vector $\vec{X}_{i,G}$ in the current generation through mutation. Five most frequently referred mutation strategies implemented in the public-domain DE codes available online *at http://www.icsi.berkeley.edu/~storn/code.html* are listed below:

"DE/rand/1": $\vec{V}_{i,G} = \vec{X}_{r_1^i,G} + F \cdot (\vec{X}_{r_2^i,n} - \vec{X}_{r_3^i,G}).$ (7)

"DE/best/1": $\vec{V}_{i,G} = \vec{X}_{best,G} + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}).$ (8)

"DE/target-to-best/1":

$$\vec{V}_{i,G} = \vec{X}_{i,G} + F \cdot (\vec{X}_{best,G} - \vec{X}_{i,G}) + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}). \quad (9)$$

"DE/best/2":

$$\vec{V}_{i,G} = \vec{X}_{best,G} + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}) + F \cdot (\vec{X}_{r_3^i,G} - \vec{X}_{r_4^i,G}). \quad (10)$$

"DE/rand/2":

$$\vec{V}_{i,G} = \vec{X}_{r_1^i,G} + F \cdot (\vec{X}_{r_2^i,G} - \vec{X}_{r_3^i,G}) + F \cdot (\vec{X}_{r_4^i,G} - \vec{X}_{r_5^i,G}). \quad (11)$$

The indices $r_1^i, r_2^i, r_3^i, r_4^i$, and r_5^i are mutually exclusive integers randomly chosen from the range [1, NP], and all are different from the index i. These indices are randomly generated once for each donor vector. The scaling factor F is a positive control parameter for scaling the difference vectors. $\vec{X}_{best,G}$ is the best individual vector with the best fitness (i.e. lowest objective function value for minimization problem) in the population at generation G.

C. *Crossover*

To enhance the potential diversity of the population, a crossover operation comes into play after generating the donor vector through mutation. The donor vector exchanges its components with the target vector $\vec{X}_{i,G}$ under this operation to form the *trial* vector $\vec{U}_{i,G} = [u_{1,i,G}, u_{2,i,G}, u_{3,i,G}, \ldots, u_{D,i,G}]$. In this article we focus on the widely used binomial crossover that is performed on each of the D variables whenever a randomly generated number between 0 and 1 is less than or equal to the Cr value. In this case, the number of parameters inherited from the donor has a (nearly) binomial distribution. The scheme may be outlined as:

$$u_{j,i,G} = \begin{cases} v_{j,i,G}, & \text{if } (rand_{i,j}[0,1) \le Cr \text{ or } j = j_{rand} \\ x_{j,i,G}, & \text{otherwise,} \end{cases} \quad (12)$$

where, as before, $rand_{i,j}[0,1)$ is a uniformly distributed random number, which is called anew for each j-th component of the i-th parameter vector. $j_{rand} \in [1,2,....,D]$ is a randomly chosen index, which ensures that $\vec{U}_{i,G}$ gets at least one component from $\vec{V}_{i,G}$.

D. Selection

The next step of the algorithm calls for *selection* to determine whether the target or the trial vector survives to the next generation i.e. at $G = G + 1$. The selection operation is described as:

$$\begin{aligned} \vec{X}_{i,G+1} &= \vec{U}_{i,G}, \quad \text{if } f(\vec{U}_{i,G}) \le f(\vec{X}_{i,G}) \\ &= \vec{X}_{i,G}, \quad \text{if } f(\vec{U}_{i,G}) > f(\vec{X}_{i,G}), \end{aligned} \quad (13)$$

where $f(\vec{X})$ is the objective function to be minimized. Note that throughout the article, we shall use the terms *objective function value* and *fitness* interchangeably. But, always for minimization problems, a lower objective function value will correspond to higher fitness.

IV. PROPOSED MDE_*p*BX ALGORITHM

In this section, we describe MDE_*p*BX and discuss the various features of the algorithm such as the mutation scheme called DE/current-to-gr_best/1, a *p*-best crossover scheme and rules for adapting the control parameters F and Cr in each iteration.

A. DE/current-to-gr_best/1

DE/current-to-best/1 is one of the widely used mutation schemes in DE as it incorporates the useful information of the best solution (with highest objective function value for maximization problems) resulting in fast convergence by guiding the evolutionary search towards a specific point in the search space. Due to such exploitative behavior the algorithm may lose its global exploration capabilities and converge to a locally optimal point in the search space. To avoid such difficulties in this article we propose a less greedy and more explorative variant of the DE/current-to-best/1 mutation strategy termed as DE/current-to-gr_best/1 which utilizes the best vector of a dynamic group of q% of the randomly selected population members for each target vector. Now the population does not get attracted towards a specific point in the search space, rather it moves towards different specific points and explores the landscape much better. The new scheme may be formulated as

$$\vec{V}_{i,G} = \vec{X}_{i,G} + F \cdot (\vec{X}_{gr\_best,G} - \vec{X}_{i,G} + \vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}), \quad (14)$$

where $\vec{X}_{gr\_best,G}$ is the best solution of q % members randomly selected from the present population whereas $\vec{X}_{r_1^i,G}$ and $\vec{X}_{r_2^i,G}$ are two distinct vectors picked up randomly from the current population. Under this scheme, the target solutions are not always attracted towards the same best position found so far by the entire population and this feature is helpful in avoiding premature convergence at local optima.

B. The p-best Crossover

The crossover operation in MDE_*p*BX is named *p*-best crossover where for each donor vector, a vector is randomly chosen from the p top-ranking individuals (in accordance with their objective function values) in the current population and then normal binomial crossover is carried out as per equation (12) between the donor vector and the randomly selected *p*-best vector to produce the trial vector of same index. By means of this innovative crossover scheme the information contained in the top ranking individuals of the population is incorporated into the trial vector resulting in fast convergence. The parameter p is reduced in a linear fashion with iterations in the following manner:

$$p = ceil\left[\frac{Np}{2} \cdot \left(1 - \frac{G-1}{G_{\max}}\right)\right], \quad (15)$$

where Np is the population size, G is the current generation number, $G_{\max}$ is the maximum number of generations, $G = [1,2,...,G_{max}]$, and $ceil(y)$ is the 'ceiling' function returning the lowest integer greater than its argument y. The reduction routine of p favors exploration at the beginning of the search and exploitation during the later stages by gradually reducing the elitist portion of the population, with a randomly selected member from where the component mixing of the donor vector is allowed for generation of the trial vector.

C. Parameter Adaptation

The parameter adaptation schemes in MDE_*p*BX are guided by the knowledge of the successful values of F and Cr that were able to generate better offspring (trial vectors) in the last generation.

Scale Factor adaptation: At every generation, the scale factor F_i *of* each individual target vector is independently generated as:

$$F_i = Cauchy(F_m, 0.1), \quad (16)$$

where $Cauchy(F_m, 0.1)$ is a random number sampled from a Cauchy distribution with location parameter F_m and scale parameter 0.1. The value of F_i is regenerated if $F_i \le 0$ or

$F_i > 1$. Denote $F_{success}$ as the set of the successful scale factors, so far, of the current generation generating better trial vectors that are likely to advance to the next generation. Also let $mean_A(F_{G-1})$ is the simple arithmetic mean of all scale factors associated with population members in generation $G-1$. Location parameter F_m of the Cauchy distribution is initialized to be 0.5 and then updated at the end of each generation in the following manner:

$$F_m = w_F \cdot F_m + (1 - w_F) \cdot mean_{Pow}(F_{success}) \tag{17a}$$

The weight factor w_F is let vary randomly between 0.8 and 1 in the following way:

$$w_F = 0.8 + 0.2 * rand(0,1), \tag{17b}$$

where $rand(0,1)$ stands for a uniformly distributed random number in (0, 1) and $mean_{Pow}$ stands for power mean [10] given by:

$$mean_{Pow}(F_{success}) = \sum_{x \in F_{Success}} \left(x^n / |F_{success}| \right)^{1/n}, \tag{18}$$

with $|F_{Success}|$ denoting the cardinality of the set $F_{Success}$. We took n = 1.5 as it gives best results on a wide variety of tested problems. Small random perturbations to the weight terms of F_m and $mean_{pow}$ puts slightly varying emphasis on the two terms each time an F is generated, and improves the performance of MDE_*p*BX as revealed through our parameter tuning experiments.

Crossover probability adaptation: At every generation the crossover probability Cr_i of each individual vector is independently generated as:

$$Cr_i = Gaussian(Cr_m, 0.1), \tag{19}$$

where $Gaussian(Cr_m, 0.1)$ is a random number sampled from a Gaussian distribution according with mean Cr_m and standard deviation 0.1. Cr_i is truncated if it falls outside the interval [0, 1]. Denote $Cr_{success}$ as the set of all successful crossover probabilities Cr_i 's at the current generation. The mean of the normal distribution Cr_m is initialized to be 0.6 and then updated at the end of each generation as:

$$Cr_m = w_{Cr} \cdot Cr_m + (1 - w_{Cr}) \cdot mean_{Pow}(Cr_{success}), \tag{20a}$$

with the weight being uniformly randomly fluctuating between 0.9 and 1:

$$w_{Cr} = 0.9 + 0.1 * rand(0,1). \tag{20b}$$

The power mean is calculated as:

$$mean_{Pow}(Cr_{success}) = \sum_{x \in Cr_{Success}} \left(x^n / |Cr_{success}| \right)^{1/n}, \tag{21}$$

where $|Cr_{Success}|$ denotes the cardinality of the set $Cr_{Success}$. We took n = 1.5 here also.

D. *Explanation of Parameter Adaptation*

Earlier theoretical studies on DE [14, 15] have indicated that the scale factor F has a big role in controlling the population diversity and the explorative power of DE. During the adaptation of F_m the usage of Power mean leads to higher value of F_m that accounts for larger perturbation to the target vectors, thus avoiding premature convergence at local optima. The essence of $F_{success}$ is that, it memorizes the successful scale factors in the current generation, thereby glorifying the chance of creating better donor vectors as more and more target vectors are processed. F_m is used as a location parameter of Cauchy distribution, which diversifies the values of F more as compared to the traditional normal distribution. The fact that the Cauchy distribution has a far wider tail than traditional Gaussian distribution is beneficial when the global optima is far away from the current search point as the values of F taken from the tail region give sufficient perturbation so that premature convergence can be avoided. The adaptation of Cr_m is also based on the record of recent successful crossover probabilities and use them to guide the generation of new Cr_i 's. So, for the adaptation of Cr_m the usage of $Cr_{success}$ again records the successful Cr values, thus generates better individuals as offspring, which are more likely to survive. A normal distribution with mean Cr_m and standard deviation of 0.1 is used to generate the Cr values. The usage of power mean instead of arithmetic mean in adaptation of Cr_m leads to higher values of Cr, which eliminates the implicit bias of Cr towards small values during its self-adaptation. Here, the Cauchy distribution is avoided and Gaussian distribution is selected because the long tail property of the former is not needed in case of the crossover probability adaptation. If the Cauchy distribution were used, the long tail property of the Cauchy distribution may lead to excessive higher values of Cr, which would have to be truncated to unity. Consequently, values of Cr would become independent of the Cauchy distribution. But the usage of the Gaussian distribution provides the opportunity to generate most of the Cr values within unity because of its short tail property.

V. MDE_*p*BX-BASED SUB-OPTIMAL PTS TECHNIQUE

In PTS, the signal sub-blocks are phase-shifted by distinct phase weight factors to produce multiple candidate signals and then combined so as to choose the optimal PAPR signal. In this section, the searching sequence of PTS is worked out as a $b \times M$ dimensions combinatorial optimization problem. The total search space covering the rotational phase weight factors or angles is divided into identically spaced angles with a particular phase increment. These discrete angles are termed as trial angles. The total discrete angle space $[0 \;\; 2\pi]$ is discredited with a 10° increment. Since the MDE_*p*BX algorithm operates on a real-valued fitness space, i.e. the particle positions representing the phase weight factors or angles are real numbers. So there should be a conversion method to transform the real valued positions to the positive integer valued angles. We propose a simple conversion scheme in which the modulus of the real valued position is taken and truncating the real valued position to the nearest integer value. For the contestant algorithms (JADE, SaDE and CLPSO) the same conversion scheme is being applied so that any difference in their performance may be attributed to their internal search operators only.

VI. ITERATIVE BASED PTS (IPTS) TECHNIQUE

An iterative PTS (IPTS) technique is another popular sub-optimum PTS technique defined in [17]. Here the phase weighting factors are considered as binary quantities. In the first step, b_i is assumed to have a value 1 for all i and calculate the PAPR of the combined OFDM signal as shown in (4). In the next step we will invert the first phase weighting factor b_1 and recalculate the PAPR value with $b_1 = 0$. If the new PAPR value is less than the previous one, the current value of b_1 is retained as a component of the final sequence, otherwise the previous value is retained. The search complexity in IPTS reduces significantly since only binary phase weighting factors are considered but at the cost of average PAPR performance. The computational complexity of IPTS can be shown to be $O(M)$, where M is the number of sub-blocks.

VII. COMPUTATIONAL COMPLEXITY

While calculating the computational complexity of the proposed MDE_*p*BX technique, the multiplications and the generation of random number should be considered as the required computations. The computational complexity of the optimal PTS technique is exponential with the number of sub-blocks of the order $O(b^{M-1})$. With the population size P_p and the allotted maximum number of generations G_n the computational complexity of the MDE_*p*BX technique is $O(P_p G_n)$. Now both $O(P_p)$ and $O(G_n)$ can be assumed to possess a linear relationship with the dimensionality of the problem associated with M number of sub-blocks. As a result, on the whole, the computational complexity of the MDE_*p*BX-based PTS technique is $O(M^2)$. So a large value of M results in a more complicated search landscape and demands higher values of P_p and G_n. Table 1 shows a comparison of the computational complexity between OPTS and sub-OPTS (IPTS, CLPSO-PTS, SaDE-PTS, JADE-PTS and MDE_*p*BX-PTS) along with the PAPR values calculated and averaged over 25 independent runs when $CCDF = 10^{-4}(dB)$ (CCDF is defined in section VIII) under the conditions as stated in Figure 4.

Table 1: Comparison of computational complexity and PAPR of OPTS and other sub-OPTS techniques

Method	Total computational complexity	$PAPR\left(CCDF = 10^{-4} dB\right)$
OPTS	$O(b^{M-1})$	6.25
IPTS	$O(M)$	9
CLPSO-PTS	$O(M^2)$	8.8
SaDE-PTS	$O(M^2)$	7
JADE-PTS	$O(M^2)$	6.8
MDE_*p*BX-PTS	$O(M^2)$	6.5

A close scrutiny of Table 1 reveals that among the sub-OPTS techniques, the computational cost associated with IPTS is the least but its PAPR performance is poor as compared to the other sub-OPTS techniques. The best PAPR performance is achieved by MDE_*p*BX but not at the cost of incurring huge computational overhead as compared to IPTS. So MDE_*p*BX-PTS is able to achieve a proper trade-off between PAPR reduction and computational complexity associated with OFDM systems to some extent.

VIII. EXPERIMENTAL RESULTS

Since PAPR is a random variable, we have to calculate its statistical properties by means of a complementary cumulative distribution function (CCDF). CCDF [12,13] of the PAPR denotes the probability that the PAPR of a data block exceeds a given threshold. CCDF is the most frequently used performance metric for PAPR reduction methods. The modulation type used in this experiment is Quadrature Phase Shift Keying (QPSK) with $N = 128$ subcarriers. The phase weighting factors $b = [0\ 2\pi]$ have been used. 10000 random OFDM frames have been generated in order to generate the CCDF of the PAPR. The sampling rates for an accurate PAPR need to be increased by four times. The CDF of the amplitude of a signal sample is given by $CDF = 1 - \exp(PAPR_0)$. As a performance metric, the parameter of CCDF is defined as:

$$CCDF = P_r\left(PAPR > PAPR_0\right) \quad (22)$$

$$P_r(PAPR > PAPR_0) = 1 - P_r\left(PAPR \le PAPR_0\right) \quad (23)$$

$$\text{So, } CCDF = 1 - (1 - \exp(-PAPR_0))^N \quad (24)$$

There is an assumption in the above equation that the N time domain signal samples are mutually independent and uncorrelated. But this assumption is violated when oversampling is applied. Also, this equation does not hold for a small number of subcarriers due to violation of Gaussian assumption. The population size is kept as the same value

($P_p = 20$) for all the algorithms for a relative fair comparison. The parameter q in the mutation scheme DE/target-to-gr_best/1 of MDE_*p*BX is kept as 1/4th of the population size. The reason for setting such a value for the group size *q* is that if *q* is on par with population size, the probability that the best of randomly chosen *q*% vectors is similar to the globally best vector of the entire population will be high and the proposed mutation scheme DE/current-to-gr_best/1 basically becomes identical to the DE/current-to-best/1 scheme. This drives most of the vectors towards a specific point in the search space resulting in premature convergence. The parameter *p* in *p*-best crossover is linearly decreased with generations as shown in equation (15). For the contestant algorithms, we follow the parameter settings in the original paper of SaDE and CLPSO. For JADE, we set the parameters to be fixed: *p*=0.05 and *c*=0.1. For all the competitor algorithms, the control parameters are kept at their optimal values so that a fair comparison is made. All the performances are calculated and averaged over 25 independent runs. In Figure 2, results of the CCDF of the PAPR for MDE_*p*BX are simulated for the OFDM system with $N = 256$ subcarriers, in which M = 16 sub-block employing random partition and the phase weight factor b uniformly distributed random variable are used for PTS. As we can see that the CCDF of the PAPR is gradually promoted upon increasing the numbers of generations due to the limited phase weighting factor. Upon increasing the number of generations, CCDF of the PAPR is also improved. As evident from figure 2, MDE_*p*BX is able to attain OPTS technique performance under relatively small number of generations ($G_n = 30$).

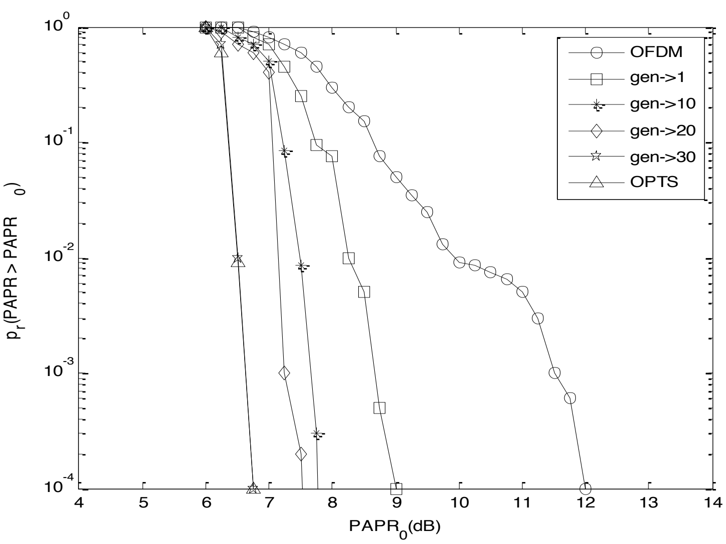

Figure 2: CCDF of MDE_*p*BX-based PTS technique for different G_n, when $N = 256, M = 16$ and $b = 4$.

Figure 3 compares the performance of the MDE_*p*BX-based PTS technique for different values of M. The value of M takes four values 2, 4, 8 and 16.

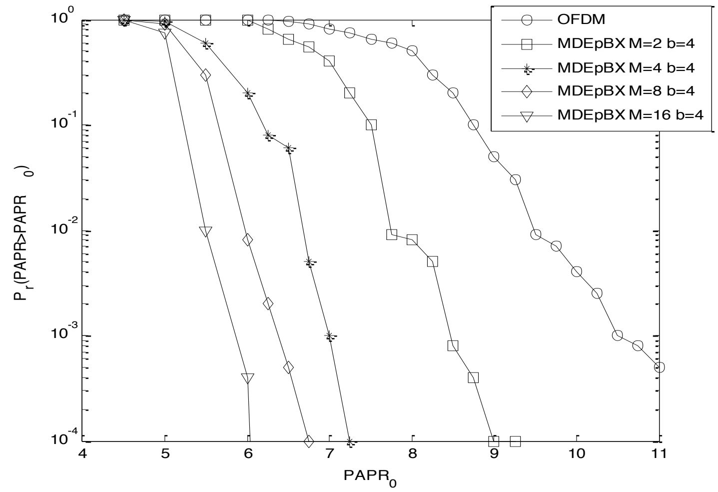

Figure 3: CCDF of the PAPR with the PTS technique searched by MDE_*p*BX technique when $N = 128$, $M = 2$, 4, 8 and 16.

It is evident from Figure 2 that the performance of MDE_*p*BX-based PTS technique is better for larger M since larger numbers of particles are searched for larger M in every update of the phase weighting factors. As the number of sub-blocks and the set of phase weighting factor are increased, the performance of the PAPR reduction becomes better. Figure 4 shows the CCDFs of the PAPR of QPSK-modulated OFDM signals in OPTS and sub-OPTS (i.e. CLPSO-PTS, SaDE-PTS, JADE-PTS and MDE_*p*BX-PTS), respectively, for M = 4 and b = 2 (M is the number of sub-blocks). Clearly, the PAPR reduction performance of MDE_*p*BX-PTS is better than the rest of the algorithms and the performance for PAPR reduction of the proposed MDE_*p*BX-PTS scheme is almost same as that of the near optimal PTS.

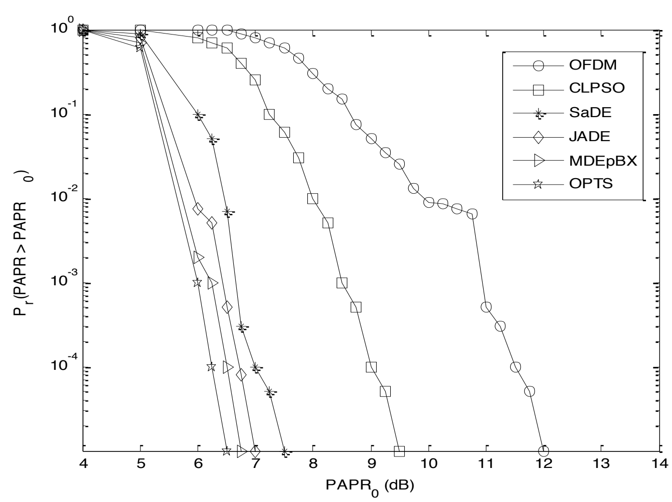

Figure 4: CCDFs of OPTS, sub-OPTS with 128 subcarriers, QPSK modulation and over sampling factor $L = 4$.

In other words, we can say that the MDE_*p*BX-based sub-optimal PTS technique is able to outperform the other sub-optimal PTS techniques based on JADE, SaDE and CLPSO and capable of attaining almost optimal PTS technique. For instance, in Figure 4 given that $CCDF = 10^{-4}$, the PAPR of the normal OFDM is near about 11 dB, and those of CLPSO-PTS, SaDE-PTS, JADE-PTS, MDE_*p*BX-PTS and OPTS are near about 8.8, 7, 6.8, 6.5, and 6.25 dB for M = 4, respectively. The superior performance of MDE_*p*BX-PTS can be attributed to the modifications in different algorithmic components of MDE_*p*BX, namely the less greedy and explorative mutation scheme DE/current-to-gr_best/1 to avoid premature convergence, innovative parameter adaptation schemes for *F* and *Cr* guided by the knowledge of the successful values of *F* and *Cr* in the last generation to increase the robustness of the proposed algorithm and the novel exploitative *p*-best crossover scheme to improve the convergence speed by incorporating the useful genetic information contained in the top-ranking individuals into the trial vector. It is to be noted that performance of MDE_*p*BX is not highly sensitive to the parameters *p* and *q*.

IX. CONCLUSIONS

Comparisons of the PAPR reduction performance for different sub-optimal PTS searching strategies based on evolutionary computation techniques have been considered in this study. In this paper, we propose a novel PTS based on MDE_*p*BX which is applied to search the optimal combination of phase weighting factors, which can achieve the OFDM system with low PAPR and does not incur huge computational overhead. The proposed three algorithmic components (DE/target-to-gr_best/1, *p*-best crossover and parameter adaptation) in MDE_*p*BX do not incorporate any additional computational cost because they are realized based on simple DE operators. The performance of the proposed scheme MDE_*p*BX-PTS is compared with IPTS and other sub-optimum PTS techniques based on JADE, SaDE and CLPSO. The complexity of the proposed technique is approximately $O\left(M^2\right)$, and it is evident from Section VIII that its performance is significantly better and more robust compared to the other compared evolutionary computation techniques. For many cases we have verified that the proposed MDE_*p*BX-based PTS technique has almost identical performance to that of the optimum PTS in the range of PAPR of most practical interest.

ACKNOWLEDGEMENTS

This work was supported by the Czech Science Foundation, under the grant no. GA102/09/1494.

REFERENCES

1. L.J. Cimini, Jr. "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing", *IEEE Trans Commun* 1985; COM-33:665–75.
2. X. Li, L.J. Cimini, Jr. "Effects of clipping and filtering of the performance of OFDM", *IEEE Commun Lett* 1998; 2:131–3.
3. A.E. Jones, T.A. Wilknson, S.K. Barton, "Block coding scheme for reduction of peak to mean envelope power ration of multicarrier transmission scheme", *Electron Lett* 1994; 30:2098–9.
4. K. Choi, K. Kang, S. Kim, "Peak power reduction scheme based on subcarrier scrambling for MC-CDMA system", *IEE Proc Commun* 2004; 151:39–43.
5. J. Tao, Z. Guangxi, "Nonlinear companding transform for reducing peak-to-average power ratio of OFDM signals", *IEEE Trans Broadcast* 2004; 50:342–6.
6. T. Jiang, W.B. Yao, P. Guo, Y.H. Song, D. Qu, "Two novel nonlinear companding schemes with interactive receiver to reduce PAPR in multi-carrier modulation systems", *IEEE Trans Broadcast* 2006; 52:268–73.
7. J. Zhang, and A. C. Sanderson, "JADE: Adaptive differential evolution with optional external archive", *IEEE Transactions on Evolutionary Computation*, Vol. 13, Issue 5, Page(s): 945-958, Oct. 2009.
8. A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization", *IEEE Transactions on Evolutionary Computation,* Vol. 13, Issue 2,, pp. 398-417, April, 2009.
9. J. J. Liang, A. K. Qin, Ponnuthurai Nagaratnam Suganthan, and S. Baskar, "Comprehensive Learning Particle Swarm Optimizer for Global Optimization of Multimodal Functions", *IEEE Trans. on Evolutionary Computation*
10. C. O. Imoru, "The power mean and the logarithmic mean", in *International Journal of Mathematics and Mathematical Sciences*, vol. 5, no. 2, pp. 337 – 343,1982.
11. Y-R. Tsai, S.J. Huang, "PTS with non-uniform phase factors for PAPR reduction in OFDM system", *IEEE Commun Lett* 2008; 12:20–2.
12. L. Yang, R.S. Chen, Y.M. Siu, K.K. Soo, "PAPR reduction of an OFDM signal by use of PTS with low computational complexity", *IEEE Trans Broadcast*, 2006; 52:83–6.
13. S.H. Han, J.H. Lee, "An overview of peak-to-average power ration reduction techniques for multicarrier transmission", *IEEE Wireless Commun*, 2005; 12:56–65.
14. D. Zaharie, "On the explorative power of differential evolution", 3rd *Int.Workshop on Symbolic and Numerical Algorithms on Scientific Computing*, Timişoara, Romania, Oct. 2 – 4, 2001.
15. S. Dasgupta, S. Das, A. Biswas, and A. Abraham, "On stability and convergence of the population-dynamics in differential evolution", *AI Communications,* Vol. 22, Issue 1, pp. 1-20, 2009.
16. R. Storn and K. Price, "Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces", *Journal of Global Optimization*, 11(4) 341–359, 1997.
17. Cimini Jr LJ, Sollenberger NR. "Peak-to-average power ratio reduction of an OFDM signal using partial transmits sequences", *IEEE Commun Lett 2000;4:86–8.*
18. R. Storn and K. V. Price, "Minimizing the real functions of the ICEC 1996 contest by differential evolution", *Proc. of the 1996 IEEE int. conf. on evolutionary computation*, Nagoya, Japan, pp. 842–844, 1996.
19. S. Das and P. N. Suganthan, "Differential evolution – a survey of the state-of-the-art", *IEEE Transactions on Evolutionary Computation*, Vol. 15, No. 1, pp. 4 – 31, 2011.