PAPR Reduction in Wireless OFDM Systems

Arun Kumar[1] , Dr. Sandip Vijay[2] , Ashish Kr. Gupta[3] , and Dr. R.K.Singh[4]

[1]Doon College of Engineering & Technology, Dehradun, Uttarakhand, India, arun_aec7271@yahoo.co.in
[2,3] Dehradun Institute of Technology, Dehradun, Uttarakhand, India, vijaysandip@gmail.com
[4] Uttrakhand Technical University, Dehradun, Uttarakhand, India, rksingh123@yahoo.com

ABSTRACT- Non-linear programming has been extensively used in wireless telecommunication systems design. PAPR reduction is implemented in the handheld devices and low complexity is a major objective. High PAPR is one of the major disadvantages of OFDM system which is resulted from large envelope fluctuation of the signal. Our proposed technique to reduce the PAPR is based on constellation shaping that starts with a larger constellation of points, and then the points with higher energy are removed. The constellation shaping algorithm is combined with peak reduction, with extra abilities defined to reduce the signal peak. This method, called MMSE-Threshold, has a significant improvement in PAPR reduction with low computational complexity. The peak reduction formulated into a quadratic minimization problem is subsequently optimized by the semi-definite programming algorithm, and the simulation results show that the PAPR of semi-definite programming algorithm (SDPA) has noticeable improvement over MMSE-Threshold while SDPA has higher complexity. Results are also presented for the PAPR minimization by applying optimization techniques such as hill climbing and simulated annealing. The simulation results indicate that for a small number of sub-carriers, both hill climbing and simulated annealing result in a significant improvement in PAPR reduction, while their degree of complexity can be very large.

Keywords – PAPR, OFDM, CDMA, AWGN, SDP

I. INTRODUCTION

Non-linear and quadratic optimization techniques have always been important research problems in wireless communications. One criterion in optimization of wireless communication systems is minimum mean square error (MMSE). This paper deals with non-linear optimization techniques with minimum mean square error criterion. We examine peak to average power ratio (PAPR) reduction in orthogonal frequency division multiplexing (OFDM) systems. PAPR reduction is implemented in the handheld devices and low complexity is a major objective. We look at peak to average power ratio (PAPR) reduction of an orthogonal frequency division multiplexing (OFDM) system. OFDM used in fourth-generation wireless technology is a multi-carrier multiplexing technique. OFDM technology allows many users to transmit in an allocated band by sub-dividing the available bandwidth into many narrow bandwidth carriers. The narrow bandwidth carriers result in the signal having a high tolerance to multi-path delay spread, because the delay spread must be very long to cause significant inter-symbol interference. The major disadvantage of OFDM technology is when all the signal peaks happen at the same time which results in a signal with a large peak. This problem is called peak to average power ratio (PAPR). A lot of research is devoted to PAPR reduction techniques categorized as, scrambling, coding, unused spectrum cancellation, nonobjective constellation, tone reservation and tone injection. An overview of most important PAPR reduction techniques is given in [1].

Since both OFDM and CDMA deal with high PAPR, we compare OFDM technology with the CDMA technology, to emphasize the advantages of OFDM over CDMA and why OFDM technology is suggested for fourth-generation systems. That is why in this paper we have just focused on PAPR reduction in OFDM systems.

II. PAPR DEFINITION

The peak of a signal x(t) is given by the maximum of its envelope $|x(t)|$. However, for a continuous random process, max $|x(t)|$ can reach infinity provided that the observation interval is long enough. Even in a discrete random process where max $|x(t)|$ is bounded, the maximum may occur at a very low probability. Therefore, a more useful definition of peak is in probability terms given by definition. A signal x(t) is said to have a peak x_p at cut-off probability Pc if

$$\Pr\left[|x(t)| < x_p\right] = P_c \quad (1)$$

Therefore, the peak to average ratio (PAR) of a random process x(t) can be specified by its histogram. The PAR definition in (2) refers to the probability density function (PDF) generated from its time samples (i.e. collect R samples of x(t), plot its histogram, and the PDF is obtained as R tends to infinity). PAR usually refers to a discrete-time measurement using $\{x_g\}$ which is equal to $x_{g\,(total)}$. Given an N-dimensional sample signal x of an OFDM system, where N is the number of sub-carriers, the PAR of OFDM is defined as,

$$PAR\big(x(t)\big) = \frac{max\left|x_g(t)\right|^2}{E_x\left[\frac{1}{N}|x(t)|^2\right]} \quad (2)$$

Similarly, the peak to average power ratio (PAPR) is usually defined on the continuous time signal x(t) as,

$$PAPR\big(x(t)\big) = \frac{max\,|x(t)|^2}{E_x\left[\frac{1}{NT_{total}}\int_0^{NT_{total}}|x(t)|^2dt\right]} \quad (3)$$

In general, $PAPR\big(x(t)\big) \geq PAR\big(x(t)\big)$ and therefore evaluating performance in the discrete-time domain may lead to optimistic values. Peak power limitations are usually placed at the power amplifier, which limits the continuous time signals; therefore, PAPR is generally the more relevant metric in practice.

III. Proposed PAPR Reduction Method

In conventional shaping, one tries to minimize the average energy of the constellation for a given number of points from a given packing. The price to be paid for shaping involves: (i) an increase in the constellation-expansion ratio (CER), (ii) an increase in the peak to average power ratio (PAPR), and iii) an increase in the addressing complexity, which is the assignment of the data bits to the constellation points. In shaping, one starts with a number of points greater than what is required for a specific bit rate, with the objective of providing some degree of flexibility in the selection of the final constellation. In traditional shaping, this flexibility is used to select the points of the least average energy, but it could also be used to select the points with a low average energy and at the same time result in a small value for the peak power along the time dimensions. This result can be achieved with a signal constellation that has several choices of points (in a multi-dimensional space) available for a given input bit label. In this case, the transmitter side will select the constellation point that results in a small average energy and at the same time has a small peak power (among the possible choices corresponding to the given binary input) as shown in Fig.1. We mix the objectives of reducing the peak and the average energy in the selection of the constellation. This requires finding a proper cost function to incorporate the combined effects of the peak and the average energy, while allowing for an efficient search procedure.

IV. Definition OF The System for the Average Energy Reduction

Shaping is a method for reducing the average energy required to transmit data relative to the average energy required for an unshaped (cube) constellation while maintaining the minimum distance between constellation points.

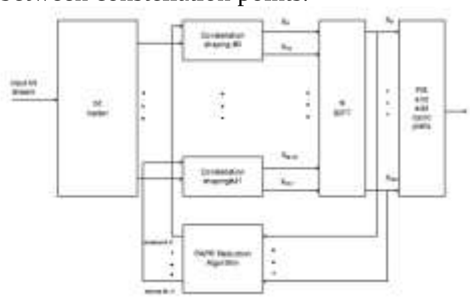

Figure 1: OFDM transmitter structure with constellation shaping and PAPR reduction algorithm

This reduction in energy is measured by shaping gain, which is achieved with a larger constellation size compared to an unshaped constellation where the increase in constellation size is given by the constellation expansion ratio (CER) as shown in Fig.2 CER is defined as the ratio of the number of points per 2-D of a shaped constellation to the minimum required number of points per 2-D to achieve the same overall rate in an unshaped constellation.

This is then mapped by the quadrature amplitude modulation to N complex frequency points. Each of these points corresponds to a sub-carrier. With redundancy in the addressing scheme, the constellations of the modulated points can be shaped to reduce the overall signal energy. The modulated points are passed through the OFDM IFFT modulator, whose outputs, after parallel-to-serial conversion, represent N Nyquist rate time domain complex samples of the baseband OFDM waveform.

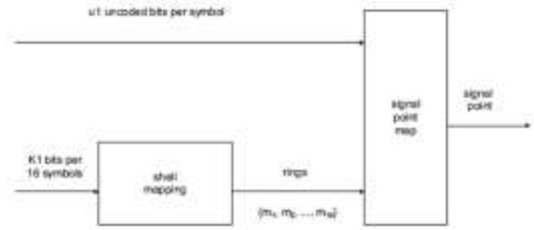

Figure 2: Constellation Shaping

Using the combination of addressing and QAM modulation allows us to reduce the energy of the signal. Performing the constellation shaping for all N sub-carriers simultaneously is difficult, and therefore we divide the N sub-carriers into m sub-spaces each with k sub-carriers (N = mk), and the constellation shaping is performed individually for each sub-space. The block diagram of the system is presented in Fig. 1.

V. Peak Reduction Algorithm With Low Complexity

One of the problems of the OFDM system is the disproportionate peaks compared to the signal average. We propose a technique that works in conjunction with constellation shaping to reduce the amplitude of these peaks. The N transmitted sub-carriers are composed of m = N=k sub-spaces. In each sub-space, we designate one bit as a dummy bit that does not carry any data. This gives us flexibility to select values of the dummy bits to reduce the signal peak in the time domain.

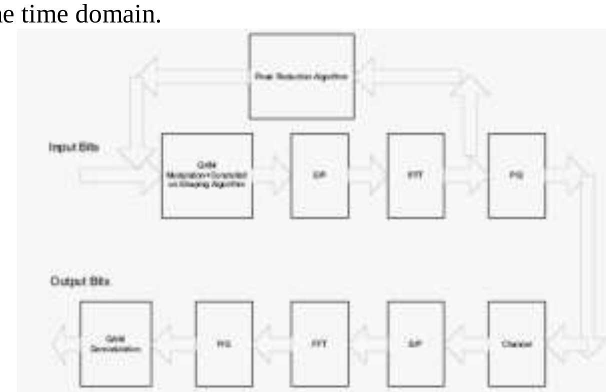

Figure 3: System Block Diagram

Since these dummy bits are used in the shaping algorithm, they can have a significant effect on peak reduction. Some preliminary results are given in .To generate the time signal, we first set all the dummy bits in all the sub-spaces to 0 and perform the addressing, modulation, and the IFFT for all N points. The result is a time domain base band vector of N complex points: $z = z_1, \ldots, z_N$. We define a complex clipping function as

$y = \text{clip}(z, T)$, such that a point z_i in vector z will be mapped to point y_i in vector y according to

$$y_i = \begin{cases} T e^{j arg(Z_i)} & |Z_i| \geq T \\ Z_i & otherwise \end{cases} \quad (4)$$

We define the error vector e for a given threshold T as

$$e = z - clip(z,T) \quad (5)$$

where z is the time domain signal vector for all the dummy bits set to 0. The addressing operates on the individual sub-spaces, and therefore changing a dummy bit in one sub-space changes the frequency domain points in that sub-space with no effect on the points in the other sub-spaces. We can therefore construct a set of orthogonal vectors, with each vector corresponding to one sub-space. Since the IFFT is an orthogonal transform the orthogonal vectors in the frequency domain result in new orthogonal vectors in the time domain. We can construct a vector €i for each sub-space i. To create this vector, we set the dummy bit to 0 in the sub-space i and perform the addressing and QAM modulation for that sub-space. Then, we set the points in other sub-spaces to zero, resulting in a frequency domain vector s_i. Afterward, we set the dummy bit in sub-space i to 1 and zero in other sub-spaces and perform the same operation to obtain the frequency domain vector si. The corresponding time domain vector €i is defined as

$$\epsilon_i = IFFT(\overline{S_i} - S_i) \quad (6)$$

Each vector ϵ_i corresponds to the change of the time domain vector by changing the dummy bit in sub-space i. Using the vectors €i, we can define
the transmitted time domain signal w for given dummy bit values $\alpha_1, \ldots, \alpha_m$ as

$$w = z - \sum_{i=1}^{m} \alpha_i \epsilon_i \quad (7)$$

and based on definition of z, we have

$$Z = IFFT(\overline{S_i}) \quad (8)$$

From (6) and (8), we can rewrite (7) in the following

$$w = IFFT(\overline{S_i}) - \sum_{i=1}^{m} \alpha_i IFFT(\overline{S_i} - S_i) \quad (9)$$

To show that w is the transmitted time domain signal for selected value of α_i, we transfer w to frequency domain called w_f as

$$w = \overline{S_i} - \sum_{i=1}^{m} \alpha_i (\overline{S_i} - S_i) \quad (10)$$

Where

$$\begin{cases} if\ \alpha_i = 0, & w_f(i) = \overline{s_i}(i) \\ if\ \alpha_i = 1 & w_f(i) = s_i(i) \end{cases} \quad (11)$$

From equation (5) and (7), we can define the actual clipped error e_{act} as

$$e_{act} = z - \sum_{i=1}^{m} \alpha_i \epsilon_i - clip(z - \sum_{i=1}^{m} \alpha_i \epsilon_i, T) \quad (12)$$

Minimizing the actual error defined in (12) is a complicated problem because of non-linear clip function; therefore we define a new minimization problem in (13). The performance of this minimization problem in terms of PAPR reduction and symbol error rate are evaluated by simulation and the results are presented in section simulation results.

$$Minimise \|e - \textstyle\sum_{i=1}^{m} \alpha_i \epsilon_i\|^2 \quad (13)$$

Consider a system with only one sub-space i and the corresponding dummy bit value α_i in (13). The first case has $\alpha_i = 0$, where the objective function in (13) is $\|e\|^2$. In the second case, the dummy bit is set to one, i.e. $\alpha_i = 1$, the objective function in (13) becomes

$$\|e\|^2 + \|\epsilon_i\|^2 - 2Real(\epsilon_i^H . e)$$

We can therefore define a decision function γ(i) for each dummy bit α_i as:

$$\gamma_i = \|\epsilon_i\|^2 - 2Real(\epsilon_i^H . e) \quad (14)$$

where $€_i^H$ is the hermitian of €i. We select $\alpha_i = 0$ for $\gamma(i) \geq 0$ and $\alpha_i = 1$, otherwise. We iterate over all the dummy bits, and use (14) to select the value of each dummy bit. Then, new error vector is calculated from (7) and (5), and new values of α_i are obtained. This is done iteratively until no further reduction is achieved in (13). This technique is called MMSE Threshold and the peak reduction algorithm is summarized in Fig. 2. The complexity of MMSE-Threshold algorithm is O(N) scalar multiplications from (14) and the total complexity is the complexity of MMSE-Threshold algorithm plus the complexity of IFFT. Therefore the total complexity of MMSE-Threshold is O(N logN) .To evaluate the performance of MMSE-threshold which is based on (13), this equation is also minimized with one of the classical methods of quadratic programming called semi definite programming . Then to evaluate how good (13) estimates the PAPR reduction, the peak to average energy of w given in (7) is minimized in terms of α_i by two heuristic algorithms called hill climbing and simulated annealing . The minimization of the quadratic problem given in (13) is known as an unconstrained quadratic program (UQP) for binary variables. Many combinatorial optimization problems pertaining to graphs such as determining maximum cliques, maximum cuts, maximum vertex packing, minimum coverings, maximum independent sets, and maximum independent weighted sets are also formulated as the UQP problem. There are many non-linear optimization techniques that can be utilized to solve the quadratic problem defined in (13). From the classical methods, we select semidefinite programming for comparison with MMSE- Threshold to minimize. To describe semidefinite programming, first we define a standard quadratic problem as [2]

$$max\ \ x^T Qx;\ x \in \{-1, +1\}^{n_1} \quad (15)$$

where Q can be any symmetric matrix. Since $x^T Q_x$ = Trace $(x^T Q_x)$, (15) is equivalent to the following problem,

$$\max\ \ Trace(XQ)$$

$$s.t. X = xx^T, x \in R^{n_1};\ X_{i_1 i_1} = 1, i_1 = 1, \ldots\ldots, n_i \quad (16)$$

The constraint $X = xx^T$ implies that X is symmetric, positive semidefinite and rank-1. Due to the constraint $X = xx^T$, (16) is a non-convex optimization problem. If we remove the rank-1 constraint, we obtain the following relaxed problem known as semidefinite programming (SDP) problem

$$\max \; Trace(XQ); \; s.t. \; X \geq 0$$

$$X_{i_1 i_1} = 1, i_1 = 1, \dots\dots, n_i \tag{17}$$

where X ≥ 0 means that X is symmetric and positive semidefinite.

The quadratic problem defined in (13) is solved by semidefinite programming package given in [3]. First, the binary variables αis are changed to x(-1, 1) of semidefinite programming according to [4], then the solution of semidefinite programming is mapped to the solution of binary quadratic

problem by randomization technique defined in [2]. The complexity of semidefinite programming is estimated in polynomial time according to [2].

To evaluate how good (13) estimates the PAPR reduction ,the peak to average of w given in (7) is minimized in terms of by two α_i heuristic algorithms called hill climbing and simulated annealing. Hill climbing (discrete form of steepest descent algorithm) is a search technique which starts with a known solution and at each step examines all possible changes to the input parameters, dummy bits in our case, and selects the change that results in the best improvement. In this application, we set all the dummy bits to 0, perform the entire encoding process and calculate the PAPR. Then, we iterate over all the dummy bits and for each dummy bit we flip the value and perform the encoding and calculate the PAPR. For each flip, we must undo all previous flips, i.e. each new solution is different from the current one by one bit. We find the bit that results in the largest decrease, flip it and continue until no further improvement is observed. The complexity of hill climbing can be as large as that of the exhaustive search. The other optimization technique examined is the simulated annealing. Simulated annealing is a generalization of a Monte Carlo method for examining the equations of state and frozen states of a system. To apply the simulated annealing technique to our problem, we start with all dummy bits set to zero and the solution with temperature Temp. In this application, the PAPR is interpreted as the energy E of the system. Similar to the hill climbing method, all the possible bits are flipped, and for each possible dummy bit flip, ΔE is calculated as the difference in the PAPR in the two cases. If the change is negative, the new configuration is accepted and if it is positive, it is accepted with the probability of exp(-ΔE=(KBTemp), where KB is the Boltzman's constant. Over time, the temperature Temp will be slowly decreased until it reaches 0, where the simulated annealing becomes a hill climbing search. Once no further improvement is made, the algorithm terminates. The complexity of simulated annealing can be as large as that of the exhaustive search.

VI. Simulation Results

The simulation is done for QAM-256 in two different cases for MMSE-Threshold.

1- When the clipping threshold is set to zero, and starting point for the minimum PAPR search is set to all dummy bits equal to zero. In this case, the simulations are done for two different numbers of sub-carriers, N = 128 and N = 1024.

2- When the optimum value of threshold for minimum PAPR is determined by simulation and few random starting points are tried to select the minimum PAPR among them. In this case, the simulations are performed for two different numbers of sub-carriers, N = 256 and N = 1024.

In both of these cases, as the number of sub-carriers changes, PAPR changes considerably. In addition, simulation is done for different numbers of bits per dimension (QAM-4, QAM-16, QAM-64); however, there is a negligible change in PAPR. Therefore, the corresponding graphs are not presented. In MMSE-Threshold, we have considered one dummy bit to have the choice of selecting the set with lower PAPR. Simulation results show that increasing the number of dummy bits has a negligible effect on PAPR but a noticeable effect on the clipped energy above the threshold. Since MMSE-Threshold is based on a clipping algorithm and clipping is a non-linear operation, MMSE-Threshold does not guarantee that we can get the minimum PAPR, but the simulation results show a large improvement in PAPR for MMSE-Threshold compared to original OFDM. We minimize the error by (13) at an appropriate threshold level and 20 random starting points for case 2 of simulations for MMSE-Threshold. In Fig. 4 and Fig. 3, the time domain signal is obtained by assuming zero value for dummy bits in different sub-spaces and the threshold is set to zero. In these figures, cumulative distribution of PAPR in terms of PAPR values is depicted. From Fig. 4 and Fig. 3, it is clear that by the proposed MMSE-Threshold, we have gained an improvement of approximately 5dB and 4dB at 10^{-3} in PAPR reduction for N = 128 and N = 1024, respectively. Because the average energy in the proposed method has reduced 1dB, the PAPR graphs of the proposed method are shifted to the left with

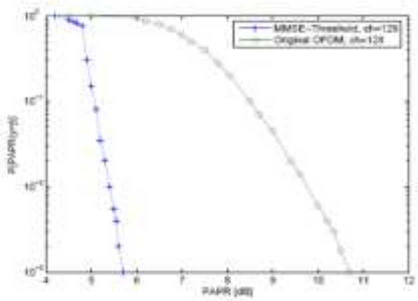

Figure 4: Probability that PAPR is grater than PAPR values on the x axis (N=128)

the amount of 1dB for a fair comparison. Fig.5 shows the energy clipped above threshold for different threshold levels. As it is clear from the graph, by increasing the number of dummy bits, total clipped energy decreases but it has a negligible effect on PAPR, which is shown in Fig 8 Therefore, we assume only one dummy bit in every sub-space. Fig. 7 shows the probability that the PAPR is larger than some defined value using a system with N = 256 for the following optimization techniques:

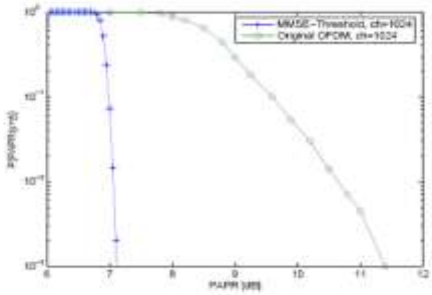

Figure 5: Probability that PAPR is grater than PAPR values on the x axis (N=1024)

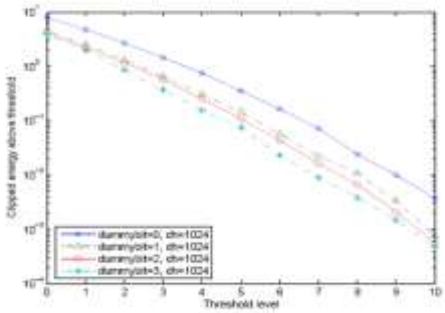

Figure 6: Clipped energy above threshold for different numbers of dummy bits (N=1024)

MMSE-Threshold, semidefinite programming, hill climbing, simulated annealing and exhaustive search. Fig. 7 indicates that MMSE-Threshold results in PAPR of 6.7 dB at the probability of 10^{-3}, which corresponds to an improvement of about 4dB over the original OFDM system, while SDPA has achieved noticeably better PAPR with higher complexity.

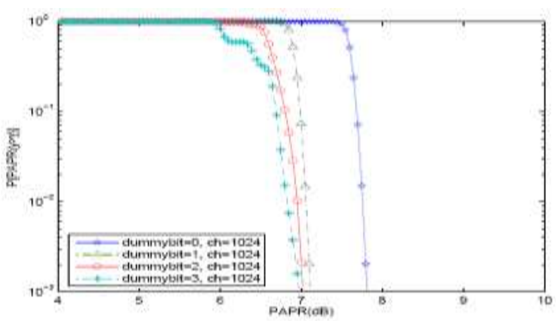

Figure 7: PAPR reduction for different numbers of dummy bits (N=1024)

The PAPR reduction for exhaustive search, hill climbing and simulated annealing are better than that of MMSE-Threshold technique because they have minimized the PAPR problem (not the quadratic problem defined in with much higher complexity compared to MMSE-Threshold. Fig.9 represents the PAPR results for a system with N = 1024 for the following optimization techniques: MMSE-Threshold, semi-definite programming, hill climbing and simulated annealing. The exhaustive search algorithm is not feasible due to the size of the problem in this case. MMSE- Threshold results in PAPR of 7:5 dB at the probability of 10^{-3}, which corresponds to an improvement of about 4 dB over the original OFDM system, while SDPA has noticeable improvement over that of MMSE-Threshold with higher complexity.

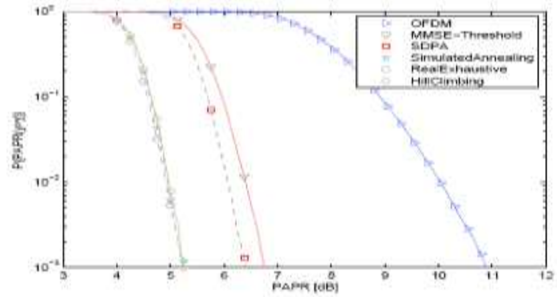

Figure 8: Probability that PAPR is grater than PAPR values on the x axis (N=256)

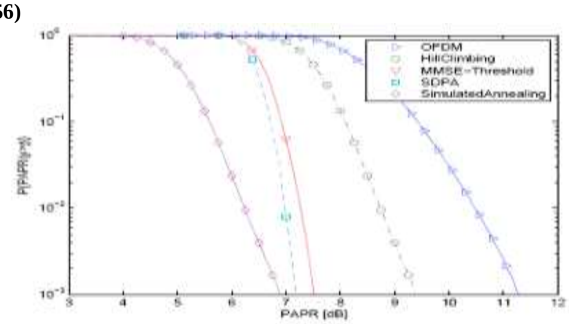

Figure 9: Probability that PAPR is grater than PAPR values on the x axis (N=1024)

The PAPR reduction for simulated annealing is better than that of MMSE-Threshold technique because it has minimized the PAPR problem (not the quadratic problem with much higher complexity compared to MMSE-Threshold. The hill climbing performance is significantly inferior compared to the simulated annealing technique because the number of sub-carriers is large. In terms of symbol error rate (SER), MMSE-Threshold has negligible degradation compared to the original OFDM signal in additive white Gaussian noise (AWGN) channel and the corresponding graph is not presented.

VIII. Conclusion

Simulation results show that the PAPR reduction of MMSE-Threshold is significant compared to the original OFDM. The PAPR reduction of SDPA is noticeably better than MMSE-Threshold, while the complexity of MMSE- Threshold is less than that of SDPA. MMSE-Threshold is compared with some state-of-the-art methods, and the simulation results show that our proposed technique has better or similar PAPR reduction compared to the state-of-the-art techniques and the proposed technique's complexity is low. In addition, the proposed method results in a 1dB reduction in average energy.

IX. References

[1] S. H. Han and J. H. Lee. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. IEEE wireless communications, 12(2):56{65, April 2005.

[2] W-K Ma, T. N. Davidson, K. M. Wong, Z-Q Luo, and P-C Ching. Quasi-maximum-likelihood multiuser detection using semi-de_nite relaxation with application to synchronous CDMA. IEEE transaction on signal processing, 50(4):912{922, April 2002.

[3] K. Fujisawa, M. Kojima, K. Nakata, and M. Yamashita. SDPA User'sManual. Department of Mathematical and Computing Sciences, Tokyo

Institute of Technology, Tokyo, Japan, 2004.

[4] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani. A randomization method for quasi maximum likelihood decoding. In Proceedings of the 9th Canadian workshop on information theory, pages 135{138, 2005.

[5] M. H. Magalhaes, R. Ballini, P. Molck, and F. Gomide. Combining forecasts for natural streamow prediction. In IEEE annual meeting of the fuzzy information, volume 1, pages 390{394, 2004.

[6] K. S. Man. Long memory time series and short term forecasts. International Journal of Forecasting, 19, issue 3:477-491, 2003.