Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 86 (2006) 3890–3903
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

A low complexity multicarrier PAR reduction approach based on subgradient optimization ☆

Alper Tunga Erdogan*

Electrical-Electronics Engineering, Koc University, Rumelifeneri Yolu, Sariyer,34450 Istanbul, Turkey

Received 1 April 2005; received in revised form 8 February 2006; accepted 29 March 2006
Available online 2 June 2006

Abstract

We introduce a novel subgradient optimization-based framework for iterative peak-to-average power ratio (PAR) reduction for multicarrier systems, such as wireless orthogonal frequency division multiplexing (OFDM) and wireline discrete multitone (DMT) very high-speed digital subscriber line (DMT-VDSL) systems. The proposed approach uses reserved or unused tones to minimize the peak magnitude of the DMT symbol vector where these tone values are iteratively updated through a subgradient search. The algorithms obtained through this framework have very simple update rules, and therefore, low computational complexities in general. Since the approach is based on the direct update of some frequency domain parameters, the power spectral density (PSD) level constraints that exist in the communications standards can easily be incorporated into the algorithms. This feature also enables simple compensation for the effects of transmit filter on PAR. Furthermore, we can locate the Active Set PAR reduction method for real baseband signals as a special case of the subgradient approach and provide its natural extension to handle complex baseband DMT signals. In addition to the peak level cost function, we also introduce the K-peak energy cost function which is also used to develop effective subgradient algorithms as illustrated by the simulation examples.

Keywords: Active-set; ADSL; Multicarrier; Discrete multitone (DMT); Orthogonal frequency division multiplexing (OFDM); Peak-to-average power ratio (PAR); PAPR; PAR reduction; PMEPR; Subgradient; Subdifferential; Tone reservation; VDSL

1. Introduction

Multicarrier scheme offers various advantages especially in terms of providing an easy means of counteracting the frequency selective effects of the broadband channels. For this reason, it has been the choice and the candidate for several wire-line (e.g., ITU ADSL, very high-speed digital subscriber line (VDSL) Standards) and wireless (e.g., IEEE 802.11a) standards.

One major drawback of the multicarrier scheme is the high effective dynamic range of the modulated signal. Considering the limitations of analog front ends of transmitters in terms of their linear operation ranges, high peak-to-average power ratio (PAR) nature of multicarrier signals causes severe challenges for implementation. A sizeable amount of research has been done to address this problem. (see for example [1–5] and the references therein). The major goal has been to produce low complexity algorithms and schemes to reduce the high dynamic

☆This work was supported in part by TUBITAK Career Award, Contract No:104E073.
*Tel.: +90 212 338 1490; fax: +90 212 338 1548.
E-mail address: alper@stanfordalumni.org.

0165-1684/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2006.03.030

range to a reasonable level with no or the minimum amount of bandwidth loss. Among the existing methods, the tone reservation method proposed by Tellado [1] and Gatherer and Polley [6] has been an effective approach for reducing the PAR level of the multicarrier signal. The method is based on the use of the unused carriers (due to low signal to noise ratio in those carriers) or the carriers reserved on purpose for the minimization of the peak level of each symbol. This scheme has the additional advantage that it is transparent to the receiver, i.e., no specific PAR reduction information needs to be communicated to the receiver.

One of the earliest algorithms was provided in [6], where the reduction in peak level is achieved by iterative projections on two convex sets: the convex set of signals with peak level less than a pre-specified level and the convex set of signals with some pre-specified values at the information tones. These iterations would converge to a point in the intersection of these convex sets if it is non-empty. We can also iterate the pre-specified peak level to minimize the peak level by repeating POCS for different target levels. The algorithm doesn't minimize the peak level, instead, it tries to push the peak below a selected level. Due to the simple structure of the projection rules to the sets involved, the algorithm has simple updates. However, the number of updates required for a satisfactory peak level tends to be high.

In [1], one of the proposed approaches was the minimization of the signal-to-clipping error energy ratio cost function again for a pre-determined clipping level where Tellado proposed the use of gradient-based updates on the reserved tone levels for the minimization of this cost function. This approach may also require a large number of updates, in addition, as in the approach of [6], this approach also require a wise and possibly iterative choice of the clipping level.

Another approach proposed in [1] is the direct minimization of the peak level (without any reference to a pre-determined clipping level). If we use the notation ρ for the vector of reserved tones for PAR reduction, this problem can be formulated as

$$\text{minimize } J(\rho) = \|\mathbf{x}\|_\infty, \tag{1}$$

i.e., the minimization of the infinity norm of the time domain discrete multitone (DMT) symbol vector $\mathbf{x}$. The non-smooth nature of this cost function probably makes it not so popular and this feature may be considered as an obstacle for the low complexity algorithm development. For the real baseband DMT case, Tellado proposed [1] the use of well-known linear programming (LP)-based relaxation for the minimization of the infinity norm of an affine form, which has the same optimal point. Similarly, for the complex baseband case, the minimization of the cost function in (1) can be casted as quadratically constrained quadratic program (QCQP)by introduction of some slack variables [1].

The conventional methods for the solution of the LP and QCQP equivalents of the minimization of (1) have polynomial complexity in general which makes them less desirable for the implementation, as this optimization task is to be repeated for every DMT symbol during the nominal operation. For a lower complexity solution of the LP problem corresponding to the real baseband case, Krongold [7] made use of the active set approach [8]. For the purpose of using the LP-based active set approach for the complex baseband case, Krongold used the polygon approximations of circles in complex plane such that the approximated version of the problem in (1) can still be casted as still an LP problem [7].

In real-life applications one cannot choose arbitrary values for the reserved tones, as there are power spectral density (PSD) constraints for different applications imposed by the standards (see for example the ITU standard document for ADSL [9]) for the spectral compatibility reasons. Therefore, the average power levels of the reserved tones are constrained by the PSD mask levels. These average power constraint levels can be converted to peak magnitude level constraints because of algorithmic convenience, in which case we can write the peak minimization problem as the constrained optimization problem

$$\begin{aligned} &\text{minimize} \quad J(\rho), \\ &\text{subject to} \quad \rho \leqslant \beta, \end{aligned} \tag{2}$$

where β is the positive vector of peak magnitude constraints. The problem in (2) can no longer be casted as an LP equivalent problem even in the real baseband case (whereas LP-based approximations are still possible), and it can be casted as QCQP for both real and complex baseband cases [10]. In [10], an LP-based approximation for (2) where the constraints on reserved tones are imposed on the classical active set algorithm in a way that can lead to a suboptimal solution. However, the examples in

[10] show that this approach can produce reasonable results that are somewhat close to the performance levels that can be obtained by the true QCQP solutions.

In this article we present a subgradient-based framework to be used for algorithm development for the direct minimization of the cost function in (1) and also for similar non-smooth cost functions that can be used for PAR reduction. The proposed framework comes up with a family of algorithms, depending on the choices of step sizes, subgradient vector combinations and even the cost functions. Basic contributions and benefits of this approach can be listed as follows:

- The subgradient approach allows the direct treatment of the original problem in (1) instead of dealing with equivalent problems.
- This framework enables easy extensions such as the inclusion of constraints. For example, it is easy to come up with a low complexity and optimal solution for the PSD constrained peak minimization problem (2).
- It brings a new theoretical perspective for the tone reservation approach with a rich structure where high performance algorithms such as active set approach [7] can be categorized as a special case.
- The extension of the active set approach [7] for complex constellations can be implemented with low complexity and without any loss of performance due to approximations.
- Through a wise selection of cost functions (for example the cost functions introduced in Section 3.2), step sizes and subgradient combinations, we can produce iterative algorithms with lower complexity than the existing methods that have desirable performances for the most general complex baseband and PSD constrained cases.

The organization of the article is as follows: Section 2 outlines the data model corresponding to the tone reservation approach for the PAR reduction problem. The subgradient-based PAR reduction approach is introduced in Section 3. In this section, the issues of step-size selection, PSD level constraints and the active-set approach connection are covered. Furthermore, we introduce new non-smooth cost functions as alternatives to the cost function in (1) and the corresponding subgradient algorithms. The discussion of the complexities of the algorithms are provided also in Section 3. In Section 4, the examples illustrating the proposed algorithms' performance are provided. Finally, Section 5 is the conclusion. At the end of the article, we include an Appendix where we provide a brief summary of subgradient optimization algorithms.

2. Multicarrier data model and PAR minimization by tone reservation

The baseband samples for a multicarrier (DMT or orthogonal frequency division multiplexing (OFDM)) symbol is given by

$$x_n = \sum_{k=-(N/2)+1}^{N/2} \frac{X_k e^{j2\pi kn/NL}}{\sqrt{N \cdot L}} \quad n = -P \cdot L, \ldots, N \cdot L - 1, \tag{3}$$

where N is the FFT size (without oversampling), P is the number of prefix samples, X_k is the information signal at the kth carrier and L is the oversampling factor. Let

$$\mathbf{x} = [\, x_0 \quad x_1 \quad \ldots \quad x_{N \cdot L-1} \,]^{\mathrm{T}} \tag{4}$$

be the vector formed by the DMT symbol samples (excluding prefix), we can write

$$\mathbf{x} = \mathbf{F}\mathbf{X}, \tag{5}$$

where $\mathbf{F} = [F_{lm}]_{N \cdot L \times N}$ is the IFFT matrix with

$$F_{lm} = \tfrac{1}{\sqrt{N \cdot L}} e^{j2\pi(l-1)(m-(N/2))/N \cdot L}, \quad \begin{matrix} l = 1, \ldots, N \cdot L \\ m = 1, \ldots, N \end{matrix} \tag{6}$$

and

$$\mathbf{X} = [\, X_{-(N/2)+1} \quad \ldots \quad X_0 \quad X_1 \quad \ldots \quad X_{N/2} \,]^{\mathrm{T}} \tag{7}$$

is the vector of transmission tones. The PAR of the corresponding DMT symbol is defined as

$$PAR = \frac{\|\mathbf{x}\|_\infty^2}{E(x_n^2)}. \tag{8}$$

The tone reservation method makes use of some reserved or unused tones to reduce the peak level of the symbol $\|\mathbf{x}\|_\infty$ [1,6]. This would be equivalent to the minimization of PAR when we take the denominator of (8) to be equal to the average energy contribution from only the data carrying tones. If we assume that Q tones are used for the PAR reduction purposes with the indexes $\{l_1, l_2, \ldots, l_Q\}$, we can decompose the expression

in (5) as

$$\mathbf{x} = \mathbf{\Gamma}\rho + \underbrace{\mathbf{U}\varphi}_{\gamma} \tag{9}$$

where $\rho = [\,X_{l_1} \quad \cdots \quad X_{l_Q}\,]^{\mathrm{T}}$ is the vector containing tones to be used for the PAR reduction and $\mathbf{\Gamma}$ is the matrix containing the corresponding columns of $\mathbf{F}$, and φ is the vector containing the information carrying tones and $\mathbf{U}$ is the corresponding partial FFT matrix.

As a result we can write the peak minimization problem in (1) more explicitly as the convex optimization problem [1]

$$\underset{\rho}{\text{minimize}}\ \|\mathbf{x}\|_\infty = \|\mathbf{\Gamma}\rho + \gamma\|_\infty, \tag{10}$$

which is the minimization of the infinity norm of a vector that is an affine function of the reserved tone values. As we noted in the introduction, (10) can be casted as LP in the real baseband case and QCQP in the complex baseband case. The PSD constrained version of (10) can be casted as QCQP for both real and complex baseband cases.

In this article, we provide a new framework for the iterative solution of the convex optimization problem in (10) and its PSD constrained version, for both real and complex baseband cases, based on the subgradient optimization methods. This framework enables the development of low complexity algorithms for the direct iterative manipulation of frequency domain parameters. In the Appendix, we provide a short review of subgradient methods, which forms a background for our treatment of subgradient-based PAR reduction framework in the next section.

3. Subgradient-based PAR reduction

The subgradient approach introduced in the Appendix can be used to build a framework for the development of low complexity algorithms. We'll first show the application of the subgradient optimization to the conventional peak minimization problem, i.e., the infinity norm minimization in (1). It will be shown that this approach will yield a family of algorithms, where different algorithms can be developed based on the selection and combination methodology of the subgradients and the step size choices. The active-set approach [7], can be shown to be a special case of subgradient-based algorithms as discussed in Section 3.1.1. This observation easily leads to the simple extension of the active-set approach for complex constellations without any approximations and unnecessary complexity increase. Another convenience of the subgradient approach in terms of simple handling of the PSD constraints in (2) is discussed in Section 3.1.2. Later in Section 3.2, we'll introduce an alternative cost function, K-peak energy, which leads to the development of subgradient-based algorithms which have better performance and lower complexity than the existing approaches as later illustrated by the examples in Section 4. Finally, in Section 3.3, we discuss the relative complexities of the subgradient-based algorithms.

3.1. Subgradient-based peak minimization

In order to obtain a low complexity solution to the PAR minimization problem in (1), we first start by finding the subdifferential set corresponding the convex but non-smooth cost function

$$f(\rho) = \|\mathbf{\Gamma}\rho + \gamma\|_\infty, \tag{11}$$

and for that we refer to the following two properties of subgradients [11]:

- Given a set of convex functions $g_i(\rho)$, $i = 1 \ldots m$, with the corresponding subdifferential sets $\partial g_i(\rho)$, if
$$g(\rho) = \max_{i=1\ldots m} g_i(\rho), \tag{12}$$
then
$$\partial g(\rho) = \mathbf{Co}\left\{\bigcup\{\partial g_i(\rho) | g_i(\rho) = g(\rho)\}\right\}, \tag{13}$$
where **Co** represents the convex hull operation.
- Given $h(\rho) = g(A\rho + b)$, where g is convex, then $\partial h(\rho) = \{A^H\mathbf{y} | \mathbf{y} \in \partial g(\rho)\}$.

Using these properties one can show that the subdifferential set of function f in (11) is given by

$$\partial f(\rho) = \mathbf{Co}\left\{\left\{\frac{x_k}{|x_k|}\mathbf{\Gamma}^H_{k,:} | |x_k| = f(\rho)\right\}\right\}, \tag{14}$$

where $\mathbf{\Gamma}_{k,:}$ is the kth row of $\mathbf{\Gamma}$.

Based on this subdifferential set, the PAR minimization algorithm steps can be outlined as

- Determine the time samples for which the peak magnitude level is achieved.
- The complex conjugate of the corresponding rows of $\mathbf{\Gamma}$ scaled by the magnitude normalized time samples are the subgradients.
- Possible iteration directions are the convex combinations of the negative of these subgradients.

Therefore, possible search directions are the opposite of the scaled versions of the vectors which define the mapping between the frequency parameters to be adjusted and the peak values. If J is the set of time instants for which maximum magnitude is achieved, i.e., $J=\{k||x_k|=\|\mathbf{x}\|_\infty\}$, then a possible search direction for the subgradient projection algorithm is

$$\mathbf{d}=-\sum_{k\in J}\xi_k\underbrace{\frac{x_k}{|x_k|}}_{s_k}\mathbf{\Gamma}^H_{k,:}, \tag{15}$$

where $\sum_{k\in J}\xi_k=1$ and $\xi_k\geqslant 0$.

As a result, we can write the update for the subgradient-based PAR minimization algorithm as follows

$$\rho^{(i+1)}=\rho^{(i)}-\mu^{(i)}\sum_{k\in J^{(i)}}\xi_k^{(i)}s_k^{(i)}\mathbf{\Gamma}^H_{k,:} \tag{16}$$

where

- $J^{(i)}$ is the set of indexes for which maximum magnitude output is achieved at the ith iteration,
- $k\in\{0,\ldots,\Omega-1\}$ is an index from the index set $J^{(i)}$,
- $\mu^{(i)}$ is the step size at the ith iteration.

Note that (16) is a flexible update rule, where different variations are possible depending on the selection of the step-size rule and also on the selection of $\xi^{(i)}$s which determines the inclusion and the weight of different subgradient vectors corresponding to different search directions. Changing convex combination in (16) with a linear combination is also another possible approach which brings extra flexibility in terms of obtaining alternative search directions.

As discussed in the Appendix, there are various possible choices for $\mu^{(i)}$ such as

- Constant step size: $\mu^{(i)}=\mu$.
- Normalized constant step size: $\mu^{(i)}=\mu/\|\mathbf{d}^{(i)}\|_2$.
- A zero-limit-divergent-sum (ZLDS) step size satisfying Polyak's conditions in (46) and (47): $\mu^{(i)}=\mu_0/i$.
- A normalized ZLDS: $\mu^{(i)}=\mu_0/\|\mathbf{d}^{(i)}\|_2^2 i$. Note that since $\|\mathbf{d}^{(i)}\|_2^2$ is bounded from above and below, the ZLDS rule is satisfied.
- The relaxation rule:

$$\mu^{(i)}=\alpha^{(i)}\frac{|x^{(i)}_{l^{(i)}}|-\hat{f}^{*(i)}}{\|\mathbf{d}^{(i)}\|_2^2} \tag{17}$$

as in the relaxation rule of (48), where $\alpha^{(i)}\in[0,2)$. Here a reasonable choice for $\hat{f}^{*(i)}$ is given by

$$\hat{f}^{*(i)}=|x^{(0)}_{l^{(0)}}|10^{-\Psi/20}, \tag{18}$$

where Ψ is the target PAR reduction level (in dB). Alternatively, one could use the adaptive target level methods (e.g., [12–14]) to determine $\hat{f}^{*(i)}$.

For convenience, one may choose $\xi_l=1$ for some $l\in J$ and $\xi_k=0$ for $k\neq l$ in which case the search direction simplifies to

$$\mathbf{d}=-s_l\mathbf{\Gamma}^H_{l,:}, \tag{19}$$

where we can rewrite the PAR minimization algorithm as

$$\rho^{(i+1)}=\rho^{(i)}-\mu^{(i)}s^{(i)}_{l^{(i)}}\mathbf{\Gamma}^H_{l^{(i)},:}, \tag{20}$$

where $l^{(i)}$ is a peak location at the ith index.

Note that for the case when we would like to take transmit filter's frequency selective effect into account, we can modify the algorithm following the same methodology as in [7,15]. In that case the only required change in the algorithm would be to modify the $\mathbf{\Gamma}$ matrix in a way to reflect transmit filter's effect.

3.1.1. Active set approach as a special case

In the literature, probably the most efficient existing algorithm for the tone reservation-based peak minimization is the active set method of [7]. In this section, we show that the active set approach for real baseband signals can be classified as a special case of the subgradient approach for the problem in (1). This observation is useful in terms of obtaining the direct extension of the active set approach for the complex baseband case, instead of using polygon approximation-based approach in [7] that leads to unnecessary complexity increase and performance loss.

In the active set approach of [7], a new time domain signal is formed at each iteration, which is a linear combination of the projections of the delta functions at the peak locations, where the projection is to the set of time signals spanned by PAR reduction tones, i.e.,

$$\mathbf{p}^{(i)}=\sum_{k\in J^{(i)}}\alpha_k\mathbf{p}_k, \tag{21}$$

where

- $J^{(i)}$ is the set of peak locations at iteration i (Active Set),
- $\mathbf{p}_k$ is the projection of $\delta(n-k)$ to the set of time-domain signals spanned by PAR reduction tones,
- α_k is the linear weighting corresponding to $\mathbf{p}_k$.

Here, α_ks are selected such that $\mathbf{p}^{(i)}$ has the same magnitude at all peak locations and the signs of $\mathbf{p}^{(i)}$ at the peak locations are the same as the signs of the corresponding peaks. The active set approach updates the time domain signal $\mathbf{x}$ using $p^{(i)}$:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \mu^{(i)}\mathbf{p}^{(i)}. \tag{22}$$

The step size $\mu^{(i)}$ is selected such that the current peak levels (active set) are reduced to a level where it is balanced by a new peak, and this new peak will be the new member of the active set. Therefore, $\mu^{(i)}$ can be determined by checking the magnitude variations due to $\mu^{(i)}\mathbf{p}^{(i)}$ at all time points.

We note that the set of peak locations $J^{(i)}$ is the index set of active constraints of the LP equivalent problem for the real baseband case. This set is also used to form the subdifferential set in (14), where the subgradients are obtained from the normals of the "active" hyperplanes of the piecewise linear cost function $f(\rho)$. Furthermore, the $\mathbf{p}_k$ functions are nothing but the inverse fourier transform of the zero extended rows of $\mathbf{\Gamma}$ matrix. Therefore, the time domain update of the active set approach is equivalent to the frequency domain update of the subgradient approach. This shows that the active set approach derived for the LP equivalent problem for the real baseband case can be classified as a special case of the subgradient approach by replacing the convex combination requirement in (15) with the linear combination, i.e., we remove the constraints $\sum_{k\in J^{(i)}}\xi_k^{(i)} = 1$ and $\xi_k^{(i)} \geqslant 0$. Therefore, the frequency domain version of the active set approach can be obtained through the use of the subgradient framework:

The active set approach constrains the update level at the peak locations to be equal in magnitude and in order to reflect this requirement, we choose $\xi_k^{(i)}$s in (16) such that

$$\mathbf{\Gamma}_{J^{(i)},:}\sum_{k\in J^{(i)}} s_k^{(i)}\xi_k^{(i)}\mathbf{\Gamma}_{k,:}^H = \underbrace{\begin{bmatrix} s_{l_1^{(i)}} \\ s_{l_2^{(i)}} \\ \vdots \\ s_{l_m^{(i)}} \end{bmatrix}}_{\mathbf{s}^{(i)}}, \tag{23}$$

$$\Rightarrow \mathbf{\Gamma}_{J^{(i)},:}\mathbf{\Gamma}_{J^{(i)},:}^H \underbrace{\begin{bmatrix} s_{l_1}^{(i)} & 0 & \dots & 0 \\ 0 & s_{l_2}^{(i)} & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & s_{l_{m^{(i)}}^{(i)}}^{(i)} \end{bmatrix}}_{\mathbf{S}^{(i)}} \xi^{(i)} = \mathbf{s}^{(i)}, \tag{24}$$

from which we can write

$$\mathbf{S}^{(i)^H}\mathbf{\Gamma}_{J^{(i)},:}\mathbf{\Gamma}_{J^{(i)},:}^H\mathbf{S}^{(i)}\xi^{(i)} = \mathbf{1}_{m^{(i)}}, \tag{25}$$

where $m^{(i)}$ is the number of active set elements and $l_1^{(i)}, l_2^{(i)}, \dots, l_{m^{(i)}}^{(i)}$ are the indexes of the active set elements at the i^{th} iteration. Note that $\xi^{(i)}$ obtained as the solution of (25) may contain negative elements and that is the reason for the relaxation of the convex combination condition in (16).

Once ξ is found the remaining step is the determination of the step size $\mu^{(i)}$. The step size is selected as the minimum value at which the magnitude of the active set elements becomes equal to the magnitude of a time point which is not a member of active set. We can write this equivalence relation as

$$(f(\rho^{(i)}) - \mu^{(i)})^2 = |x_m^{(i)} - \mu^{(i)}\underbrace{\mathbf{\Gamma}_{m,:}\mathbf{d}^{(i)}}_{q_m}|^2, \quad m \notin J^{(i)}, \tag{26}$$

$$\Rightarrow (1-|q_m|^2)\mu^{(i)^2} - 2(f(\rho^{(i)}) - \Re e\{x_m^{(i)}q_m^{(i)*}\})\mu^{(i)} + (f(\rho^{(i)})^2 - |x_m^{(i)}|^2) = 0. \tag{27}$$

The step values for which the magnitude of the time sample located at m becomes equal to the magnitude of active set members are given by the roots of (27). The quadratic equation in (27) has two positive roots when $|q_m^{(i)}| < 1$. The smallest of these two roots is given by

$$\mu_m^{(i)} = \frac{f(\rho^{(i)}) - \Re e\{x_m^{(i)}q_m^{(i)*}\} - \sqrt{(f(\rho^{(i)}) - \Re e\{x_m^{(i)}q_m^{(i)*}\})^2 - (1-|q_m|^2)(f(\rho^{(i)})^2 - |x_m^{(i)}|^2)}}{1-|q_m^{(i)}|^2}. \tag{28}$$

When $|q_m^{(i)}|>1$ the quadratic equation will have one positive and one negative root. The positive root in this case will again be given by (28). We should note that in the case of real baseband DMT signals (28) reduces to

$$\mu_m^{(i)} = \frac{f(\rho^{(i)}) - x_m^{(i)} q_m^{(i)} - |q_m^{(i)} f(\rho^{(i)}) - x_m^{(i)}|}{1 - |q_m^{(i)}|^2}. \tag{29}$$

In order to reduce the step size complexity one can follow the assumption in [7] and simplify the step size computation further.

As a result, we can write the desired step size as

$$\mu^{(i)} = \min_{m \notin J^{(i)}} \mu_m^{(i)}. \tag{30}$$

Note that once the step-size is determined, the update for the time values can be obtained by

$$x_m^{(i+1)} = x_m^{(i)} - \mu^{(i)} q_m, \tag{31}$$

where q_m values were pre-computed in the step-size determination phase.

We should also note that the frequency domain version of the active set approach is not based on the LP equivalent or the approximation of the problem in (1), it is directly based on the minimization of this cost function using linear combinations of its subgradients. As a result, it provides the natural extension of the real baseband active set approach for more general complex constellations. The additional computational burden and the potential performance loss due to casting an LP problem to complex case via use of polygon approximation of circular boundaries in complex plane [7] is avoided.

3.1.2. Incorporation of PSD level constraints

As we previously noted in the introduction, the PAR minimization problem that we deal with in practice is actually a constrained problem, where the PAR reduction tone values cannot be selected arbitrarily and they should obey the PSD mask constraint set by the standards (see for example [9]). Therefore, the problem that we need to deal with is (2) rather than (1). As we noted before, we cannot write an LP equivalent for (2), however, we can obtain a QCQP equivalent problem and the conventional methods to solve QCQP have high complexities that make them unsuitable for real-time implementation.

The effects of the PSD constraints for the tone-reservation methods was recently addressed in [16] where a bound (the so called $A_{\max}$-bound) for the PAR reduction performance under PSD constraints was proposed. In [10], a method to incorporate PSD constraints into the active set approach is proposed. In this reference, within the LP modeling (real baseband case) and approximation (complex baseband case) frameworks, the active set approach is modified in a manner such that the step size level is corrected in case of spectral constraint violation of any tone after a classical active set update. Furthermore, during the iterations some tone values may get frozen at the constraint level (till the end of iterations) and the algorithm continues with the active set process applied on the remaining tones. Satisfying PSD constraints in frequency domain and peak balance requirement in time domain can become two contradictory requirements, and freezing some tones and continuing active set with the remaining tones provides a solution to get around this problem. However, this approach results in a suboptimal solution for (2) since the frozen tones do not necessarily have the PSD constraint level at the optimal point. In addition, this treatment of constraints may slow down the convergence, as the step that is used for updating all tones is reduced in case one or more tones exceed the constraint level. Therefore, exceeding the constraint levels at some tones causes slower update in the other tones.

Through the use of subgradient framework with PSD constraints, which is outlined below, a better performance can be obtained with less complexity as illustrated in Section 4. Within the subgradient framework, the PSD constraints can be easily imposed without sacrificing from optimality, therefore, the optimal solution can be achieved in the limit.

To consider the modification to the subgradient-based updates, we first note that the PSD mask constraints impose an average power constraint for each tone. These average power constraints can be converted to instantaneous magnitude constraints such that the set

$$S = \{\rho \,||\rho_i| \leqslant \beta_i i = 1, \ldots, Q\} \tag{32}$$

would define the feasible set of values that ρ can take, where β_i is the magnitude constraint for the ith component of ρ . Since S is a convex set, we can incorporate the PSD level requirements into our algorithm easily by including the projection to the constraint set as suggested by (45). As a result, after the update given by (16), we apply the following

projection rule

$$\mathbf{P}_S\{\rho_k\} = \begin{cases} \rho_k & |\rho_k| \leqslant \beta_k, \\ \frac{\beta_k}{|\rho_k|}\rho_k & |\rho_k| > \beta_k, \end{cases} \tag{33}$$

to each component of ρ. We should note that, as outlined in the Appendix, subgradient iterations with the projection to the convex constraint set outlined above will converge to the optimal solution for the PSD-constrained peak minimization problem (2), i.e., there will be no loss of optimality.

The choice of the instantaneous magnitude β_k strictly at the mask level may be unnecessarily conservative and lower the achievable PAR reduction level. Therefore, we may consider to set β_k levels to be above the PSD level as suggested in [10].

We should note that we cannot apply the projection rule in (33) to the active set method, as the individual manipulation of the PAR reduction tones with these projections would cause the violation of the peak balance rule of the active-set approach.

3.2. The largest K-peak energy approach

Another potential benefit of the subgradient optimization framework is that it can be used for handling alternative non-smooth cost functions that can be proposed for PAR reduction. For example, we can introduce the following cost function

$$J_K(\rho) = \sum_{l=1}^{K} |x_{i_l}|^2, \tag{34}$$

where $|x_{i_1}| \geqslant |x_{i_2}| \geqslant \cdots \geqslant |x_{i_{NL}}|$ and the ordered tuple $(i_{l_1}, i_{l_2}, \ldots, i_{l_{NL}})$ is a permutation of $(0, 1, \ldots, NL-1)$. Note that this cost function corresponds to the total energy of the K largest magnitude components of the vector $\mathbf{x}$ (which we will refer as K-peaks). We can show that the cost function J_K is a convex function of ρ: if we consider the cost function evaluated at $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in C^{NL}$ and $\lambda \in [0, 1]$:

$$J_K(\mathbf{x}) = \sum_{l=1}^{K} |\lambda x_{i_l} + (1-\lambda) y_{i_l}|^2 \tag{35}$$

$$\leqslant \lambda \sum_{l=1}^{K} |x_{i_l}|^2 + (1-\lambda) \sum_{l=1}^{K} |y_{i_l}|^2 \tag{36}$$

$$\leqslant \lambda J_K(\mathbf{x}) + (1-\lambda) J_K(\mathbf{y}). \tag{37}$$

Therefore, J_K is a convex function of $\mathbf{x}$. Since $\mathbf{x}$ is an affine function of ρ, J_K is also a convex function of ρ. Note that the inequality (36) is due to the convexity of the absolute-squared function, and the inequality in (37) directly follows from the definition of J_K.

We should note that the idea of taking the K largest magnitude values into consideration also appeared in [17], where group of peaks of the uncompensated DMT symbol with a magnitude above a certain prespecified clipping level are reduced through a non-iterative peak cancellation algorithm. Despite this similarity in terms of taking group of largest peaks into account, our consideration is based on the minimization of cost function of the largest K-peaks of the compensated DMT symbol.

By incorporating the largest K magnitudes to the cost function, we are able to take the effect of potential peaks that can appear during the iterative search into account. This choice results in the use of multiple rows of Γ matrix in each iteration of the search, without any need for the equivalence of the corresponding output magnitudes. This property further results in faster convergence as illustrated in Section 4 with the use of a simple step size rule. Note that the effect of potential peaks is taken into account during the step size calculation step of the active set method whereas this consideration is pushed to the cost function with the introduced approach. Furthermore, the projection rule in (33) can be applied to impose PSD constraints, as there is no constraint about maintaining equivalence between some time domain peaks.

In order to obtain a subgradient-based algorithm to minimize J_K, we first write down the subdifferential set corresponding to J_K which is given by

$$\partial J_K(\rho) = \mathbf{Co}\left\{ \sum_{l=1}^{K} \mathbf{x}_{i_l} \mathbf{\Gamma}^H_{i_l,:} \;\middle|\; \forall (i_1, i_2, \ldots, i_{NL}) \text{ such that } |x_{i_1}| \geqslant |x_{i_2}| \geqslant \cdots \geqslant |x_{i_K}| \right.$$
$$\left. \text{where } (i_{l_1}, i_{l_2}, \ldots, i_{l_{NL}}) \text{ is a permutation of } (0, 1, \ldots, NL-1) \right\}.$$

One important property of the vectors in the above convex hull expression is that they are weighted sums of (the hermitian transpose of) the rows of $\mathbf{\Gamma}$ where the weights are the corresponding peak levels. Therefore, each vector is the combination of multiple search directions corresponding to different peak levels where higher weights are given to the directions corresponding to the larger peaks. This property turns out to be useful in terms of obtaining effective search directions and enables use of very simple step size selection rules such as ZLDS as illustrated by the examples in Section 4.

The idea of K-peak energy cost function can be generalized to a larger set of cost functions such as

$$J_K(\rho) = \sum_{l=1}^{K} w_l |x_{i_l}|^p, \tag{38}$$

where i_ls are the ordering indexes, w_ls are the non-negative weights and p is the integer exponent. For example, for $p=1$ and $w_l = 1$ this would correspond to K-peak total magnitude function.

3.3. Complexity Analysis

If we consider the basic subgradient update algorithm given by (16), this equation would require $(Q+1)M$ complex multiplies and QM complex additions where M is the number of non-zero ξs, or the number of active set components being used. The complexity is dominated not by this expression, but by the update of the time domain vector. However, we should note that we don't need to take the IFFT of ρ to compute the time domain contributions. Instead, since we are using the rows of $\mathbf{\Gamma}$ for the update, IFFT of the zero extended version of the rows are nothing but the cyclic shifted versions of the same kernel $\mathbf{p}$ as used in the active-set approach. Therefore, these time shifted kernels can be combined using the same weights, which would lead to computational complexity of NLM additions and multiplications. In addition, the determination of the peak point would require NL comparisons which can be performed using NL additions. (In the complex case additional NL multiplications are required for the computation of magnitudes.)

If we consider the active set approach for the real baseband case, according to [7], the first iteration requires NL multiplications and additions, the second iteration requires NL multiplications and $2NL$ additions, and the ith iteration (for $i>2$) requires $(i+1)NL$ multiplications and iN additions. In addition, the step size computation with full search over time samples requires additional $2NL$ adds and NL multiplications (when we apply simplifications in computation with sign balance assumption and use of table look-up as proposed in [7]). If only significant magnitude time samples are used for the step size computation, to reduce the complexity even further as suggested in [7], then only NL comparisons are needed for the determination of the significant samples, and therefore the cost for the step size computation is NL additions. To summarize the total cost of N_I active-set iterations is

$$3NL + NL\sum_{i=3}^{N_I} i = \left(\frac{N_I^2}{2} + \frac{N_I}{2}\right)NL \quad \text{additions}, \tag{39}$$

$$2NL + NL\sum_{i=3}^{N_I}(i+1) = \left(\frac{N_I^2}{2} + \frac{3N_I}{2} - 3\right)NL \quad \text{multiplications}, \tag{40}$$

ignoring the cost of step size computation, which would cause additional $2N_INL$ additions and N_INL multiplications in the full search case (with sign balance condition and table look-up based simplifications) and N_INL additions in the reduced search case (which includes all simplifications proposed in [7]).

For the complex baseband case, if we use the subgradient-based extension of the active set approach, the complexity will be the same as above where real additions and multiplications needs to be replaced by complex counterparts (which can be roughly approximated as a factor of 4 overall complexity increase compared to the real case). If the polygon approximation-based active set method of [7] is used, then we need a factor of 8 complexity increase in the $8-ogonal$ approximation case and a factor of 16 complexity increase in the $16-ogonal$ approximation case. Therefore, for the complex baseband peak minimization, the polygonal approximation-based active set method has a factor of 2 or 4 more computational requirement than the subgradient extension of the active set approach.

If we now look at the complexity of the subgradient algorithms for K-peak energy function: If we assume that the search vector has the form

$$\mathbf{d} = \sum_{k=1}^{K} x_{i_k} \mathbf{\Gamma}^H_{i_k,:}, \tag{41}$$

then the update of the time signal would require NLK additions and multiplications at each iteration. If we use a ZLDS step-size rule, which turns out to have sufficient performance as illustrated in examples of Section 4, then we can ignore the cost of step size computation as it will require about one multiplication per iteration. As a result the total complexity of time sample computation for N_I iterations would be

$$N_I KNL \quad \text{additions}, \tag{42}$$

$$N_I KNL \quad \text{multiplications}. \tag{43}$$

In addition to this we need $N_I KNL$ additions for determining K largest peak locations (the worst case cost).

The complexity expressions for the basic subgradient approach and K-peak approach are linear on the number of iterations, whereas the active set approach is quadratic on the number of iterations. This is due to the fact that the number of correction vectors per iteration grows with the number of iterations for the active set approach whereas it remains constant for the basic subgradient-based approach and K-peak approach. However, we cannot directly conclude that the proposed algorithms have lower complexity unless we compare their complexities in connection with PAR reduction performances as a function of iterations, which is done in the next section.

4. Examples

In the first example, we illustrate the benefit of the subgradient framework for the PSD constrained PAR minimization problem by considering the setting provided in [10]. In this setting, a (frequency multiplexed) downlink ADSL scenario is considered where the transmitter under consideration is a real baseband multicarrier system with $N = 512$ and tones 33–255 are used for data transmission. Data tones are assumed to be loaded with 10 bits (i.e., 1024 QAM constellation). We also assume that an oversampling factor of $L = 4$ is used. We consider the use of 24 random tones in 33–255 range where instantaneous reserved tone magnitudes are constrained by the PSD constraint level (as in Section 4.2.3 of [10]). We used the subgradient search algorithm for the basic subgradient update rule give in (20) followed by the constraint set projection step in (33). For the step size selection rule, a simple ZLDS scheme with $\mu^{(i)} = 0.3/(i\|\mathbf{d}^{(i)}\|_2^2)$ where $\mathbf{d}^{(i)}$ is the $Q \times 1$ search vector at the ith step. The computational load of calculating $\|\mathbf{d}^{(i)}\|_2^2$ is negligible in general and for this specific example $\|\mathbf{d}^{(i)}\|_2^2$ is even equal to a constant which is Q. We made this choice as the computational complexity of the ZLDS rule is very low and we can obtain a satisfactory performance level.

In Fig. 1, PAR CCDFs corresponding to the different iterations of the proposed algorithm and the Amax bound are shown. (PAR is defined as the ratio of the peak energy to the average energy of the data carrying tones.) In the same figure we show the PAR CCDF corresponding to four iterations of the PSD constrained active set approach of [10] together with minimum PAR bound (QCQP solution) curve obtained from Fig. 9 of [10] are shown. It is clear from this figure that the six iterations of the proposed algorithm has better performance than the four iterations of the PSD-constrained active set approach of [10]. If we compare the complexities: the four iterations of the active set approach requires $14NL$ additions and $11NL$ multiplications in the simplified step size search case, $18NL$ additions and $15NL$ multiplications in the full step size search case. The six iterations of the subgradient approach requires $12NL$ additions (including peak search) and $6NL$ multiplications. Therefore, a better performance level is achieved with less complexity. Note that the presented complexity numbers refer to the unconstrained case for both algorithms and ignores the additional load due to imposition of PSD constraints. For this example, it turns out that the subgradient algorithm requires 2.42 total tone corrections for six iterations on average due to the imposition of the PSD constraints which means an additional load of $2.42NL$ additions and multiplications. Even when we assume that there is no complexity increase in the active set approach due to the PSD constraints , the complexity of the subgradient approach is still less than the active set approach. In general the PSD constrained active set approach of [10] has higher complexity than the original active set approach: for each iteration where a tone is frozen at the PSD level, time domain kernel needs to be recalculated for the remaining free tones [10] (as it is not practical to store all possible time kernels referring to different reserved tone combinations) which corresponds to additional NL additions and multiplications for that iteration. However, for this example, the active set process was terminated once a tone's value exceeded the constraint level [10].

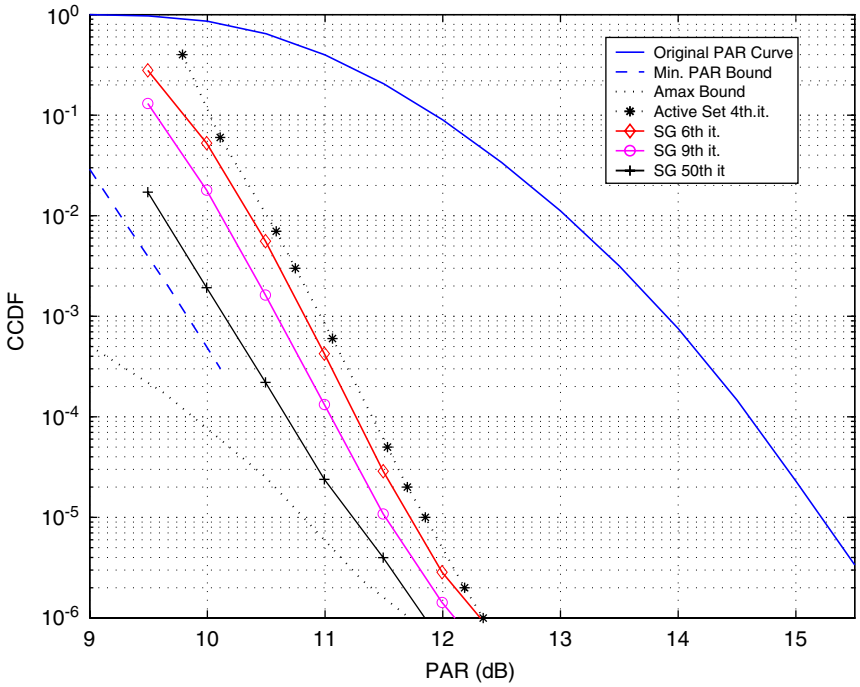

Fig. 1. PAR CCDFs for downlink ADSL with 24 random PAR reduction tones. The reserved tone magnitudes are at the PSD constraint level.

Due to the preserved optimality property of the proposed approach, the PAR CCDF curve gets closer to the bound curves as we increase the number of iterations. The PAR CCDF curves for 9 and 50 iterations of the algorithm are shown in Fig. 1 to illustrate this behavior.

As the next example, we illustrate the use of subgradient approach for the complex baseband signals, where we consider an OFDM signal with 256 carriers. As in the setting corresponding to Fig. 13 of [7], we assume that the tones $\{5, 25, 54, 102, 125, 131, 147, 200, 204, 209, 247\}$ are the reserved tones for peak reduction and the data carriers are loaded with QPSK signals. We used $L = 4$ for the algorithm and $L = 8$ for the measurement of peak (as in Fig. 13 of [7]).

In Fig. 2, the PAR CCDFs for octagonal approximation-based active set approach (taken from Fig. 13 of [7]), the proposed complex extension of the active set approach and the K-peak energy approach with $K = 2$ are shown. In the K-peak energy approach we again used ZLDS rule with $\mu^{(i)} = 5/(i\|\mathbf{d}^{(i)}\|_2^2)$. It is clear from this figure that both the complex extension of the active set approach and the K-peak energy approach have better performances than the octagonal approximation-based active set approach of [7]. If we compare the complexities: five iterations of the octagonal approximation-based active set approach requires $160NL$ real additions and $136NL$ real multiplications assuming simplified search case, the proposed complex extension of the active set approach has about half of the complexity of the octagonal approximation, which is $80NL$ real additions and $68NL$ real multiplications (for the simplified search case). Five iterations of K-peak energy approach requires about $55NL$ real additions and $50NL$ real multiplications, and six iterations of K-peak energy approach requires $66NL$ real additions and $60NL$ real multiplications. (For the two-peak energy algorithm, at each iteration, peak detection requires $2NL$ real multiplications and $3NL$ real additions and the computation of time update requires $8NL$ real multiplications and $8NL$ real additions which adds up to $10NL$ multiplications and $11NL$ real additions per iteration.). Therefore, the K-peak energy approach has the least complexity and the complexity of the proposed complex extension of the active set approach is less than the polygonal approximation-based active set approach of [7].

For the final example, we consider the constrained version of the previous example to illustrate

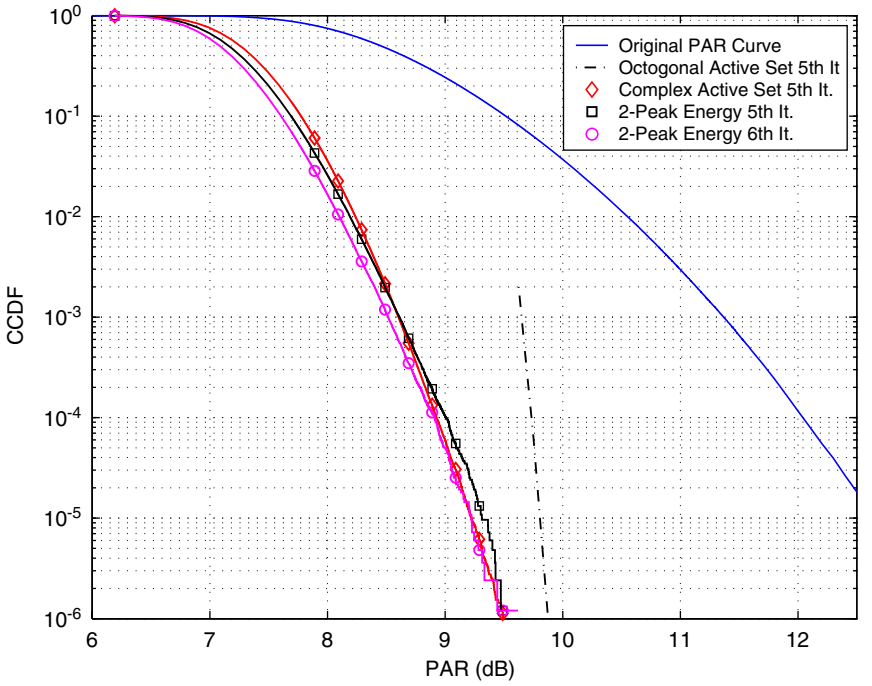

Fig. 2. PAR CCDF for 256-Carrier OFDM with 11 reserved tones. $L = 4$ is used for the algorithm and $L = 8$ is used for the performance evaluation.

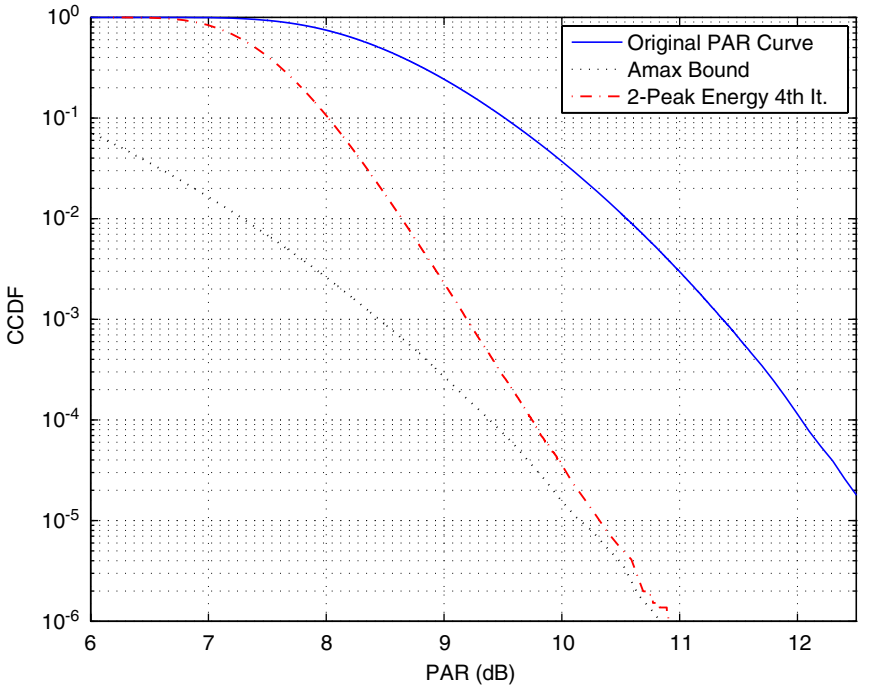

Fig. 3. PAR CCDF for 256-Carrier OFDM with 11 reserved tones, where reserved tones are constrained by a level 50% above the PSD level. $L = 4$ is used for the algorithm and $L = 8$ is used for the performance evaluation.

the most general complex baseband case with PSD constraints, where we assume that the reserve tones are assumed to be magnitude constrained by a level which is 50% above the PSD level. We again used two-peak energy algorithm with the same ZLDS rule as the previous example. The corresponding CCDF curve along with the Amax bound are shown in Fig. 3. According to this figure, the CCDF curve corresponding to the four iterations of the algorithm is very close to the bound curve for clipping probabilities less than 10^{-5} which is the region of interest.

5. Conclusion

We presented the subgradient optimization-based framework to be used for low complexity algorithm development for multicarrier PAR reduction. The introduced framework provides a flexible approach in terms of the selection of search directions, step sizes and possibly non-smooth cost functions for the construction of efficient algorithms. Furthermore, the PSD constraints and complex baseband signals can be conveniently handled without any loss of performance. The existing active-set approach for real baseband signals can be located as a special case of the proposed subgradient framework and can be easily extended to the complex baseband systems which allows the replacement of polygon approximation-based active set approach with it less complex version. We illustrated that, within this framework we can develop algorithms which can achieve better performance-complexity tradeoff than the existing approaches.

Appendix

A review of subgradient methods

Let $f(\mathbf{w})$ be a convex and possibly non-differentiable function with domain S, where S is convex. The subdifferential of $f(\mathbf{w})$ at point $\mathbf{w}$ is defined as

$$\partial f(\mathbf{w}) = \{\mathbf{g} | f(\mathbf{y}) \geq f(\mathbf{w}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{w} \rangle \quad \forall \mathbf{y} \in S\}, \quad (44)$$

where $\langle \cdot, \cdot \rangle$ is the inner product. A vector $\mathbf{g}$ which is a member of $\partial f(\mathbf{w})$ is called a subgradient of $f(\mathbf{w})$ at $\mathbf{w}$.

The subgradients play an important role in obtaining the non-differentiable counterpart of the gradient descent algorithm for differentiable functions. The subgradient-based iterative algorithm is in the same form as the gradient descent algorithm except the gradient is replaced by a subgradient. We can write the update rule for the subgradient method as

$$\mathbf{w}^{(i+1)} = \mathbf{P}_S\{\mathbf{w}^{(i)} - \mu^{(i)}\mathbf{g}^{(i)}\}, \quad (45)$$

where $\mathbf{g}^{(i)}$ is a subgradient picked from the subdifferential set $\partial f(\mathbf{w}^{(i)})$, $\mu^{(i)}$ is the step size and $\mathbf{P}_S$ is the projection to the convex set S. Although the subgradient algorithm looks very much like the gradient descent algorithm, in the subgradient iteration it may happen that $f(\mathbf{w}^{(i+1)}) > f(\mathbf{w}^{(i)})$ for any $\mu^{(i)} > 0$ [18]. However, if the $\mu^{(i)}$ step size parameter sequence is properly chosen, $\mathbf{w}^{(i)}$ can be made to converge the optimal point $\mathbf{w}^*$.

One major result about the selection of the step size parameter $\mu^{(i)}$ is due to Polyak [19]: if

$$\lim_{i\to\infty} \frac{\mu^{(i)}}{\|\mathbf{g}^{(i)}\|} = 0, \quad (46)$$

$$\sum_{i=0}^{\infty} \frac{\mu^{(i)}}{\|\mathbf{g}^{(i)}\|} = \infty, \quad (47)$$

hold then $\lim_{i\to\infty} \mathbf{w}^{(i)} = \mathbf{w}^*$, which provides sufficient conditions for convergence.

Furthermore, if the step size satisfies

$$0 < \mu^{(i)} < 2\frac{(f(\mathbf{w}^{(i)}) - f^*)}{\|\mathbf{g}^{(i)}\|_2^2}, \quad (48)$$

where f^* is the minimum value of $f(\mathbf{w})$, then, due to the convex nature of $f(\mathbf{w})$, it is guaranteed that

$$\|\mathbf{w}^{(i+1)} - \mathbf{w}^*\|_2 < \|\mathbf{w}^{(i)} - \mathbf{w}^*\|_2 \quad \forall i, \quad (49)$$

i.e., the distance to the optimal vector decreases monotonically. The choice of μ as in (48) also provides a faster (geometric) rate of convergence [13]. However, it requires a priori knowledge of the optimal value of the function to be optimized which is not reasonable in most cases. The use of an estimate of f^*, instead of f^* have been investigated in several references (see for example [20–22]) such that the subgradient update rule takes the form

$$\mathbf{w}^{(i+1)} = \mathbf{P}_S\left\{\mathbf{w}^{(i)} - \alpha^{(i)}\frac{(f(\mathbf{w}^{(i)}) - \hat{f}^*)}{\|\mathbf{g}^{(i)}\|^2}\mathbf{g}^{(i)}\right\}. \quad (50)$$

References

[1] J. Tellado, Peak to average power reduction in multicarrier modulation, Ph.D. Thesis, Stanford University, 1999.

[2] M. Sharif, B. Hassibi, A deterministic algorithm that achieves the PMEPR of *clogn* for multicarrrier signals, Proc. ICASSP 2003 (2003) 540–543.

[3] V. Tarokh, H. Jafarkhani, On the computation and reduction of peak-to-average power ratio in multicarrier communications, IEEE Trans. Commun. 48 (2000) 37–44.
[4] S.H. Muller, J.B. Huber, OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences, Electron. Lett. 33 (1997) 368–369.
[5] T. Starr, J.M. Cioffi, P.J. Silverman, Understanding Digital Subscriber Line Technology, Prentice-Hall, Englewood Cliffs, NJ, 1999.
[6] A. Gatherer, M. Polley, Controlling clipping probability in dmt transmission, 1997 Conference Record of the 31st Asilomar Conference on Signals, Systems and Computers, vol. 1, 1997, pp. 578–584.
[7] B.S. Krongold, D.L. Jones, An active-set approach for OFDM PAR reduction via tone reservation, IEEE Trans. Signal Process. 52 (2004) 495–509.
[8] D. Luenberger, Linear and Nonlinear Programming, Springer, Berlin, 2003.
[9] ITU G.dmt (G.992.1): Asymmetric Digital Subscriber Line (ADSL) Transceivers, ITU Telecommunication Standardization Sector (ITU-T), 1999.
[10] N. Andgart, B. Krongold, P. Odling, A.A. Johansson, P. Borjesson, PSD-constrained PAR reduction for DMT/OFDM, EURASIP J. Appl. Signal Process. 10 (2004) 1498–1507.
[11] S. Boyd, L. Vandenberghe, Convex Optimization, Stanford University, Stanford, 2001.
[12] J. Goffin, K. Kiwiel, Convergence of a simple subgradient level method, Math. Programm. 85 (1999) 207–211.
[13] A. Nedic, D. Bertsekas, Convergence rate of incremental subgradient algorithms, in: S. Ursayev, P. M. Pardalos (Eds.), Stochastic Optimization: Algorithms and Applications.
[14] H. Sherali, G. Choi, C. Tuncbilek, A variable target method for nondifferentiable optimization, OR. Lett. 26 (2000) 1–8.
[15] A. Salvekar, C. Aldana, J. Tellado, J. Cioffi, Peak-to-average power ratio reduction for block transmission systems in the presence of transmit filtering, IEEE ICC 2001 (2001) 175–178.
[16] N. Petersson, A. Johansson, P. Odling, P. Borjesson, A performance bound on PSD-constrained PAR reduction, IEEE Int. Conf. Commun. 5 (2003) 3498–3502.
[17] P.O. Borjesson, H.G. Feichtinger, N. Grip, M. Isaksson, N. Kaiblinger, L. Persson, A low complexity par reduction method for dmt-vdsl, Proc. DSPCS 99 (1999) 164–169.
[18] D. P. Bertsekas, Nonlinear Programming, Athena Scientific, 1999.
[19] B.T. Polyak, A general method for solving extremal problems, Dokl. Akad. Nauk SSSR 174 (1967) 33–36.
[20] M. Held, P. Wolfe, H. Crowder, Validation of subgradient methods, Math. Program. 6 (1974) 62–88.
[21] B. Polyak, Minimization of unsmooth functionals, USSR Comput. Math. Math. Phys. 9 (1969) 509–521.
[22] E. Allen, R. Helgason, J. Kennington, B. Shetty, A generalization of Polyak's convergence result for subgradient optimization, Math. Programm. 37 (1987) 309–317.