Real-Time Perception-Based Clipping of Audio Signals Using Convex Optimization

Bruno Defraene, *Student Member, IEEE*, Toon van Waterschoot, *Member, IEEE*, Hans Joachim Ferreau, Moritz Diehl, *Member, IEEE*, and Marc Moonen, *Fellow, IEEE*

***Abstract*—Clipping is an essential signal processing operation in many real-time audio applications, yet the use of existing clipping techniques generally has a detrimental effect on the perceived audio signal quality. In this paper, we present a novel multidisciplinary approach to clipping which aims to explicitly minimize the perceptible clipping-induced distortion by embedding a convex optimization criterion and a psychoacoustic model into a frame-based algorithm. The core of this *perception-based clipping* algorithm consists in solving a convex optimization problem for each time frame in a fast and reliable way. To this end, three different structure-exploiting optimization methods are derived in the common mathematical framework of convex optimization, and corresponding theoretical complexity bounds are provided. From comparative audio quality evaluation experiments, it is concluded that the perception-based clipping algorithm results in significantly higher objective audio quality scores than existing clipping techniques. Moreover, the algorithm is shown to be capable to adhere to real-time deadlines without making a sacrifice in terms of audio quality.**

***Index Terms*—Audio signal processing, clipping, convex optimization, psychoacoustics, real-time.**

I. Introduction

In many real-time audio applications, the amplitude of a digital audio signal is not allowed to exceed a certain maximum level. This amplitude level restriction can be imposed for different generic or application-specific reasons. First, it can relate to an inherent limitation of the adopted digital representation of the signal. In this case, audio signal samples exceeding the allowable maximum amplitude level will either *wrap-around* or *saturate*, depending on the digital signal processing (DSP) system architecture [1]. In both modes, the result will be a significant degradation of the audio signal's sound quality. Secondly, the maximum amplitude level can be imposed in order to prevent the audio signal from exceeding the reproduction capabilities of the subsequent power amplifier and/or electroacoustic transducer stages. In fact, an audio signal exceeding this maximum amplitude level will not only result in a degradation of the sound quality of the reproduced audio signal (e.g. due to amplifier overdrive and loudspeaker saturation), but could possibly also damage the audio equipment. Thirdly, in music production applications, the amplitude level restriction is often set deliberately as part of a mastering/mixing process. Lastly, in hearing aid applications, the maximum amplitude level restriction is necessary to preserve a high listening comfort, as impulsive noises in the vicinity of the hearing aid user will sound uncomfortably loud if the audio signal amplitude is not properly limited.

In order to preserve a high sound quality of the reproduced audio signal and a high user listening comfort in the above mentioned applications, it is of paramount importance to instantaneously limit the digital audio signal with respect to the allowable maximum amplitude level. *Clippers* (or infinite limiters) are especially suited for this purpose: these alter incoming signal sample amplitudes such that no sample amplitude exceeds the maximum amplitude level (referred to as *clipping level* from here on) [2, Sec. 5.2]. Most existing clipping[1] techniques are governed by a static input-output characteristic, acting onto the input audio signal on a sample by sample basis by mapping a range of input amplitudes to a reduced range of output amplitudes. Depending on the sharpness of this input-output characteristic, one can distinguish between two types of clipping techniques: *hard clipping* and *soft clipping* [3], where the input-output characteristic exhibits an abrupt ("hard") or gradual ("soft") transition from the linear zone to the nonlinear zone respectively.

However, such a clipping operation itself introduces different kinds of unwanted distortion into the audio signal: odd harmonic distortion components, intermodulation distortion components and aliasing distortion components [4]. In a series of listening experiments performed on normal hearing subjects [5] and hearing-impaired subjects [6], it is concluded that the application of hard clipping and soft clipping to audio signals has a significant negative effect on perceptual sound quality scores, irrespective of the subject's hearing acuity. To our best knowledge, there have been no previous research efforts on improving the perceptual sound quality of existing clipping techniques. It

Manuscript received July 19, 2011; revised February 11, 2012 and May 25, 2012; accepted July 03, 2012. Date of publication July 31, 2012; date of current version October 01, 2012. This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC), the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization," 2007–2011), and Concerted Research Action GOA-MaNet. The scientific responsibility is assumed by its authors. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bryan Pardo.

The authors are with the Department of Electrical Engineering, ESAT-SCD (SISTA), KU Leuven, B-3001 Leuven, Belgium (e-mail: bruno.defraene@esat.kuleuven.be; toon.vanwaterschoot@esat.kuleuven.be; joachim.ferreau@esat.kuleuven.be; moritz.diehl@esat.kuleuven.be; marc.moonen@esat.kuleuven.be).

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

Digital Object Identifier 10.1109/TASL.2012.2210875

[1]In this work, we use the word "clipping" to denote the deliberate operation of bounding the samples of a digital audio signal to a predefined maximum amplitude level. This should not be confused with the undesired "analog clipping phenomenon" as it can subsequently occur in various analog audio devices.

1558-7916/$31.00 © 2012 IEEE

is worthwhile to point out, however, recent research on the related problems of *audio declipping* and *audio imputation*, where the aim is to restore the missing values in clipped audio signals [7], [8].

In this paper, we propose a novel, multidisciplinary approach to clipping, aimed at minimizing the perceptible clipping-induced distortion. The proposed perception-based clipping algorithm combines aspects of digital signal processing, optimization theory and psychoacoustics. Hence, two algorithmic ingredients are novel compared to existing approaches:

- *Psychoacoustics*: incorporating knowledge about the human perception of sounds is indispensable for achieving minimally perceptible clipping-induced distortion. In other audio processing applications, the application of psychoacoustic principles and models has proven to be successful, e.g. in perceptual audio coding [9] and audio signal requantization [10].
- *Embedded convex optimization*: in an increasing number of signal processing applications, convex optimization is embedded directly into a signal processing algorithm in order to carry out nonlinear processing on the signal itself (as opposed to its more conventional use for e.g. linear filter design) [11]. In this framework, clipping of an audio signal will be formulated as a sequence of constrained convex optimization problems regularly spaced in time, aimed at minimizing perceptible clipping-induced distortion. Real-time operation of such a scheme obviously calls for application-tailored optimization methods able to solve instances of the optimization problem at hand in a fast and reliable way. Therefore, we will spend extensive attention to three different structure-exploiting optimization methods and their comparative performance.

In previous work, a perception-based approach to clipping has been presented and was seen to significantly outperform existing clipping techniques in terms of objective sound quality scores [12]. This approach has been refined by incorporating a projected gradient optimization method for solving the constrained optimization problems under consideration [13]. In this paper, the main ideas presented in [12], [13] will be reviewed and expanded, thereby introducing the following novel contributions:

- A new and significantly faster projected gradient optimization method will be proposed for solving the constrained optimization problems at hand, achieving an optimal linear convergence rate. By using this method, the perception-based clipping algorithm can effectively be applied in real time for very high solution accuracies.
- The different optimization methods will be rigorously described in the common mathematical framework of convex optimization. Advantages and disadvantages of the different optimization methods will be discussed, and theoretical complexity bounds will be derived in order to objectively compare their performance.
- The psychoacoustic principles and psychoacoustic model underpinning the perception-based clipping approach will be elaborated in detail.
- A thorough comparative objective perceptual evaluation of the proposed perception-based algorithm and existing clipping algorithms will be performed using two different objective measures of audio quality.

The paper is organized as follows. In Section II, clipping is formulated as a sequence of constrained convex optimization problems, and the inclusion of a psychoacoustic model is discussed in detail. In Section III, three different application-tailored convex optimization methods are proposed for solving the optimization problems at hand in a fast and reliable way, and corresponding theoretical complexity bounds are given. In Section IV, results are presented from a comparative audio quality evaluation of different clipping techniques, and an algorithmic complexity assessment of different optimization methods is performed. Finally, in Section V, some concluding remarks are presented.

II. Perception-Based clipping

A. General Description of the Algorithm

The goal of a clipping algorithm is to restrict the amplitude of a digital audio signal $x[n]$ to a given amplitude range $[L, U]$ (where $L \leq 0$ and $U \geq 0$), while keeping the clipped output signal $y[n]$ perceptually as close as possible to the input signal $x[n]$. In a perception-based clipping algorithm [12], the aim for *maximal perceptual similarity* is explicitly fulfilled, by

- incorporating into the algorithm knowledge about the human perception of sounds through the use of a psychoacoustic model.
- embedding into the algorithm the solution of an optimization problem, aimed at minimizing clipping-induced distortion.

Fig. 1 schematically depicts the actual operation of the perception-based clipping algorithm presented in [12]. The digital input audio signal $x[n]$ is segmented into frames of N samples,[2] with an overlap length of P samples between successive frames. The processing of one frame x_m consists of the following steps:

1) Calculate the instantaneous global masking threshold $t_m \in \mathbb{R}^{(N/2)+1}$ of the input frame $x_m \in \mathbb{R}^N$, using part of the ISO/IEC 11172-3 MPEG-1 Layer 1 psychoacoustic model 1 [14]. The instantaneous global masking threshold of a signal gives the amount of distortion energy (dB) in each frequency bin that can be masked by the signal.
2) Calculate the optimal output frame $y_m^* \in \mathbb{R}^N$ as the solution of a constrained optimization problem to be defined in Section II-B.
3) Apply a trapezoidal window to the optimal output frame y_m^* and sum the optimal output frames to form a continuous output audio signal $y^*[n]$.

B. Optimization Problem Formulation

The core of the perception-based clipping technique consists in calculating the solution of a constrained optimization problem for each frame. From the knowledge of the input frame x_m and its instantaneous properties, the optimal output frame y_m^* is calculated. Let us define the optimization variable of the problem as y_m, i.e. the output frame. A necessary constraint on the output frame y_m is that the amplitude of the output samples cannot exceed the upper and lower clipping levels U

[2]Note that N is assumed an even number from here on.

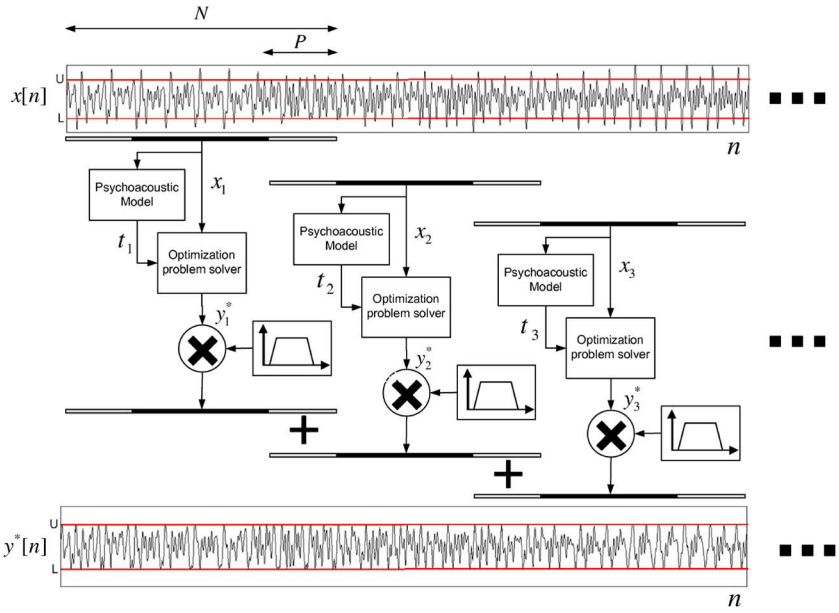


Fig. 1. Schematic overview of the presented perception-based clipping technique.

and L. The objective function f we want to minimize must reflect the amount of perceptible distortion added between y_m and x_m. We can thus formulate the optimization problem as an inequality constrained frequency domain weighted L2-distance minimization, i.e.

$$\begin{aligned} y_m^* &= \arg\min_{y_m \in \mathbb{R}^N} f(y_m) \\ &\text{s.t.} \quad \begin{cases} h_i(y_m) = U - y_m(i) \geq 0 \\ h_{N+i}(y_m) = y_m(i) - L \geq 0, \end{cases} \quad i = 0, \ldots, N-1 \\ &= \arg\min_{y_m \in \mathbb{R}^N} \frac{1}{2N} \sum_{i=0}^{N-1} w_m(i) \left| Y_m(e^{j\omega_i}) - X_m(e^{j\omega_i}) \right|^2 \\ &\quad \text{s.t.} \quad l \leq y_m \leq u \end{aligned} \tag{1}$$

where $\omega_i = (2\pi i)/N$ represents the discrete frequency variable, $X_m(e^{j\omega_i})$ and $Y_m(e^{j\omega_i})$ are the discrete frequency components of x_m and y_m respectively, the vectors $u = U\mathbb{1}_N$ and $l = L\mathbb{1}_N$ contain the upper and lower clipping levels respectively (with $\mathbb{1}_N \in \mathbb{R}^N$ an all ones vector), and $w_m(i)$ are the weights of a perceptual weighting function to be defined in Section II-C. Notice that in case the input frame x_m does not violate the inequality constraints, the optimization problem (1) trivially has the solution $y_m^* = x_m$ and the input frame is transmitted unaltered by the clipping algorithm.

Formulation (1) of the optimization problem can be rewritten as follows[3]

$$\begin{aligned} y_m^* &= \arg\min_{y_m \in \mathbb{R}^N} \frac{1}{2}(y_m - x_m)^H D^H W_m D (y_m - x_m) \\ &\quad \text{s.t.} \quad l \leq y_m \leq u \\ &= \arg\min_{y_m \in \mathbb{R}^N} \frac{1}{2} y_m^H \underbrace{D^H W_m D}_{\triangleq H_m} y_m + \underbrace{(-D^H W_m D x_m)}_{\triangleq g_m = -H_m x_m}^H y_m \\ &\quad \text{s.t.} \quad l \leq y_m \leq u \end{aligned} \tag{2}$$

[3]In this text, the superscripts T and H denote the transpose and the Hermitian transpose, respectively.

where $D \in \mathbb{C}^{N \times N}$ is the unitary Discrete Fourier Transform (DFT) matrix defined as

$$D \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & e^{-j\omega_1} & e^{-j\omega_2} & \ldots & e^{-j\omega_{N-1}} \\ 1 & e^{-j\omega_2} & e^{-j\omega_4} & \ldots & e^{-j\omega_{2(N-1)}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\omega_{N-1}} & e^{-j\omega_{2(N-1)}} & \ldots & e^{-j\omega_{(N-1)(N-1)}} \end{bmatrix} \tag{3}$$

and $W_m \in \mathbb{R}^{N \times N}$ is a diagonal weighting matrix with positive weights $w_m(i)$, being the symmetry property $w_m(i) = w_m(N-i)$ for $i = 1, 2, \ldots, (N/2) - 1$,

$$W_m = \begin{bmatrix} w_m(0) & 0 & 0 & \ldots & 0 \\ 0 & w_m(1) & 0 & \ldots & 0 \\ 0 & 0 & w_m(2) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & w_m(N-1) \end{bmatrix} \tag{4}$$

We remark that the objective function in (2) is a quadratic function and that the constraint functions are affine, hence optimization problem (2) constitutes a *quadratic program* (QP). Note also that the choice for a quadratic error in the objective function was made in order to strike a balance between perceptual relevance on one hand, and mathematical elegance and suitability of the objective function from an optimization point of view on the other hand. With this trade-off in mind, the use of a quadratic error criterion was preferred over other considered alternatives.

C. Perceptual Weighting Function

In order for the objective function in optimization problem (1) to reflect the amount of perceptible distortion added between input frame x_m and output frame y_m, the perceptual weighting function w_m must be constructed judiciously. The rationale behind applying signal-dependent weights in the summation (1) is the psychoacoustic fact that distortion at certain frequencies is

more perceptible than distortion at other frequencies, and that the relative perceptibility is mostly signal-dependent. Two phenomena of human auditory perception are responsible for this,

- The *absolute threshold of hearing* is defined as the required intensity (dB) of a pure tone such that an average listener will just hear the tone in a noiseless environment. The absolute threshold of hearing is a function of the tone frequency and has been measured experimentally [15].
- *Simultaneous masking* is a phenomenon where the presence of certain spectral energy (the masker) masks the simultaneous presence of weaker spectral energy (the maskee), or in other words, renders it imperceptible.

Combining both phenomena, the instantaneous global masking threshold of a signal gives the amount of distortion energy (dB) at each frequency bin that can be masked by the signal. In this framework, we consider the input frame x_m to act as the masker, and $y_m - x_m$ as the maskee. In other words, we make the assumption that when the ear is presented with the output frame y_m, it is in fact presented with the input frame x_m and the distortion $y_m - x_m$ simultaneously, and that the simultaneous masker-maskee relationship between both signals can be exploited. By selecting the weights $w_m(i)$ in (1) to be exponentially decreasing with the value of the global masking threshold of x_m at frequency bin i, the objective function effectively reflects the amount of perceptible distortion introduced.[4] This can be specified as

$$w_m(i) = \begin{cases} 10^{-\alpha t_m(i)} & \text{if } 0 \leq i \leq \frac{N}{2} \\ 10^{-\alpha t_m(N-i)} & \text{if } \frac{N}{2} < i \leq N-1 \end{cases} \quad (5)$$

where t_m is the global masking threshold of x_m (in dB). Appropriate values of the compression parameter α have been determined to lie in the range 0.04–0.06.

The instantaneous global masking threshold t_m of an input frame x_m is calculated using part of the ISO/IEC 11172-3 MPEG-1 Layer 1 psychoacoustic model 1. A complete description of the operation of this psychoacoustic model is out of the scope of this paper (we refer the reader to [9] and [14]). We will outline the relevant steps in the computation of the instantaneous global masking threshold and illustrate the result of each step on an example audio frame (see Fig. 2):

1) *Spectral analysis and SPL normalization*: In this step a high-resolution spectral estimate of the input frame is calculated, with spectral components expressed in terms of sound pressure level (SPL). After a normalization operation and a Hann window operation on the input signal frame, the PSD estimate is obtained through a 512-point Fast Fourier Transform (FFT). Fig. 2(a) shows the time-domain input signal, Fig. 2(b) shows the resulting spectral estimate.
2) *Identification of tonal and non-tonal maskers*: It is known from psychoacoustic research that the tonality of a masking component has an influence on its masking properties [16]. For this reason it is important to discriminate between tonal maskers (defined as local maxima of the signal spectrum) and non-tonal maskers. The output of the FFT is used to determine the relevant tonal and non-tonal maskers in the spectrum of the audio signal. In a first phase, tonal maskers are identified at local maxima of the PSD: energy from three adjacent spectral components centered at the local maximum is combined to form a single tonal masker. In a second phase, a single non-tonal masker per critical band is formed by addition of all the energy from the spectral components within the critical band that have not contributed to a tonal masker.
3) *Decimation of maskers*: In this step, the number of maskers is reduced using two criteria. First, any tonal or non-tonal masker below the absolute threshold of hearing is discarded. Next, any pair of maskers occurring within a distance of 0.5 Bark is replaced by the stronger of the two. Figs. 2(c) and 2(d) respectively depict the identified tonal and non-tonal maskers, after decimation.
4) *Calculation of individual masking thresholds*: an individual masking threshold is calculated for each masker in the decimated set of tonal and non-tonal maskers, using fixed psychoacoustic rules. Essentially, the individual masking threshold depends on the frequency, loudness level and tonality of the masker. Fig. 2(e) and 2(f) show the individual masking thresholds associated with tonal and non-tonal maskers, respectively.
5) *Calculation of global masking threshold*: Finally, the global masking threshold is calculated by a power-additive combination of the tonal and non-tonal individual masking thresholds, and the absolute threshold of hearing. This is illustrated in Fig. 2(g).

[4]Indeed, the terms in the summation of objective function (1) can then be seen to resemble to distortion-to-masker power ratios.

III. Optimization Methods

The core of the perception-based clipping algorithm described in Section II is formed by the solution of an instance of optimization problem (2) for every frame x_m. Looking at the relatively high sampling rates (e.g. 44.1 kHz for CD-quality audio) and associated frame rates under consideration, it is clear that real-time operation of the algorithm calls for application-tailored optimization methods to solve the optimization problems in a fast and reliable way. In this section, we will discuss three different structure-exploiting optimization methods, whose common ground is the notion of *convex optimization*. Before doing so, we will first look at the properties of the optimization problem at hand in this framework of convex optimization.

A. Convex Optimization Framework

Convex optimization is a subfield of mathematical optimization that studies a special class of mathematical optimization problems, namely convex optimization problems. This class can be formally defined as follows.

Definition 1 (Convex Optimization Problem): A convex optimization problem is one of the form

$$\min_{y \in \mathbb{R}^N} f(y) \quad \text{s.t.} \quad h_i(y) \leq 0, \quad i = 0, \ldots, K-1$$
$$c_j^T y = d_j, \quad j = 0, \ldots, M-1 \quad (6)$$

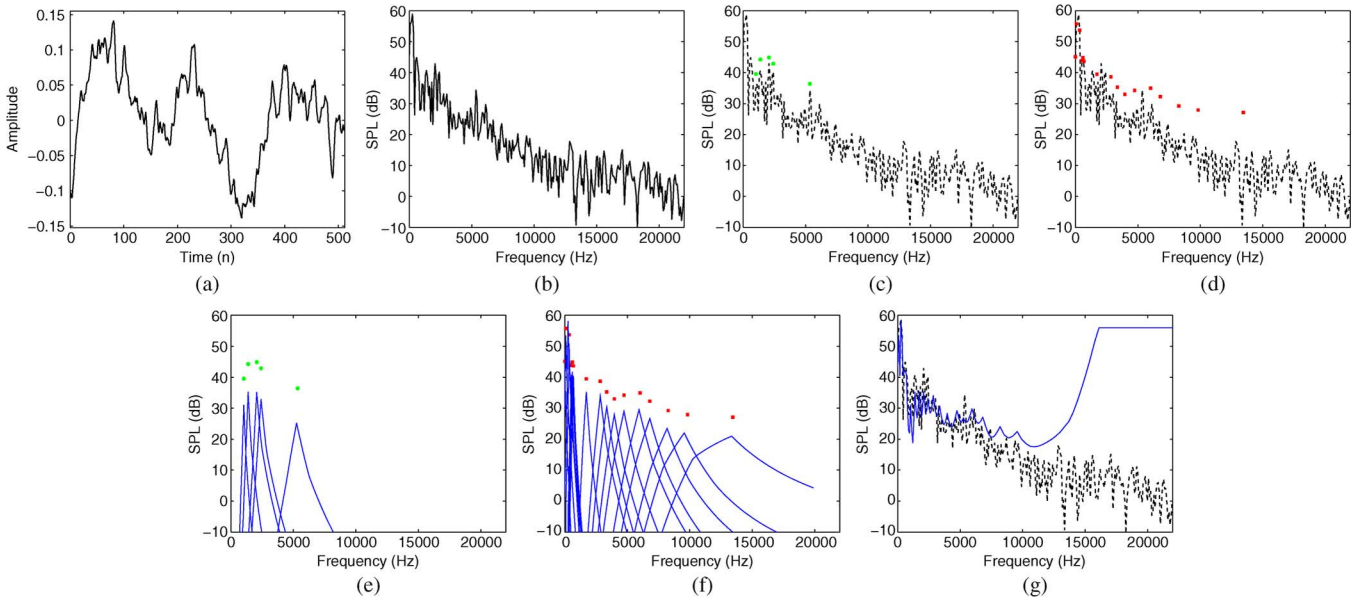


Fig. 2. Different steps in the computation of the global masking threshold using the ISO/IEC 11172-3 MPEG-1 Layer 1 psychoacoustic model 1: (a)-(b) Time domain and normalized frequency domain representations of the input audio signal (c)-(d) Tonal maskers (circles), non-tonal maskers (squares) and input frequency spectrum (dotted line) (e)-(f) Individual masking thresholds related to tonal and non-tonal maskers respectively (g) Global masking threshold (solid line) and input frequency spectrum (dotted line).

in which the objective function f and the constraint functions $h_0, \ldots, h_{K-1}$ are convex, which means they satisfy

$$f(ax+by) \leq af(x)+bf(y) \tag{7}$$
$$h_i(ax+by) \leq ah_i(x)+bh_i(y) \tag{8}$$

for all $x, y \in \mathbb{R}^N$ and all $a, b \in \mathbb{R}$ with $a+b=1, a \geq 0, b \geq 0$.

In particular, for quadratic programs, the next definition holds.

Definition 2 (Convex QP): A quadratic program (QP) is *convex* if and only if the Hessian matrix $\nabla^2 f(y)$ is positive semi-definite. It is *strictly convex* if and only if $\nabla^2 f(y)$ is positive definite.

A fundamental property of convex optimization problems is that any local minimizer is a global minimizer. For strictly convex problems, it will also be the unique minimizer.

We will now show that optimization problem (2) is a convex quadratic program, thereby looking a bit deeper into the structure of the Hessian matrix H_m.

Definition 3 (Circulant Matrix): A circulant matrix C is a square matrix having the form

$$C = \begin{bmatrix} c_0 & c_{N-1} & & \cdots & & c_1 \\ c_1 & c_0 & c_{N-1} & & & \vdots \\ c_2 & c_1 & c_0 & \ddots & & \\ & c_2 & c_1 & \ddots & & c_{N-1} \\ \vdots & & & \ddots & & \\ c_{N-1} & \cdots & & c_2 & c_1 & c_0 \end{bmatrix} \tag{9}$$

where each row is a cyclic shift of the row above it.

Theorem 1 (Diagonalization of Circulant Matrices [17], [18]): A circulant matrix $C \in \mathbb{C}^{N \times N}$ is diagonalized by the unitary DFT matrix D defined in (3), i.e.

$$C = D^H \Psi D \tag{10}$$

where $\Psi = \mathrm{diag}(\psi)$ contains the eigenvalues of C, which are obtained as the DFT of the first column of C,

$$\psi(l) = \sum_{k=0}^{N-1} c_k e^{-j\omega_{lk}}, \quad l = 0, \ldots, N-1 \tag{11}$$

We can now state the following important properties of the Hessian matrix $H_m = D^H W_m D$ in (2):

Theorem 2: The Hessian matrix $H_m = D^H W_m D \in \mathbb{R}^{N \times N}$ in optimization problem (2) is real, symmetric, positive definite and circulant.

Proof: From (10) in Theorem 1, we readily see that H_m is *circulant* and W_m contains its eigenvalues. By definitions (4), (5), the elements of W_m are real and have even symmetry, so by (11) the first column of H_m will also be real and have even symmetry. From this we can see that H_m is *real* and *symmetric*. As a symmetric matrix is positive definite if and only if all of its eigenvalues are positive, it remains to remark that the elements of W_m are positive by construction (5) to conclude that H_m is *positive definite*. ■

Corollary 1: Optimization problem (2) is a strictly convex quadratic program.

Convex optimization problems can be solved reliably and efficiently by using special methods for convex optimization. In particular, different iterative optimization methods for solving (strictly) convex QPs have been presented. Essentially, three classes of methods can be distinguished. A first class are the *projected gradient methods*, where only first-order information is used for solving the optimization problem. These methods are

conceptually simple and computationally cheap, but typically suffer from slow convergence, sometimes preventing their use in real-time applications. However, it is generally possible to perform a convergence analysis and to establish useful polynomial computational complexity bounds for these algorithms [19]. A second class are *interior-point methods*: these rely on heavier computational tasks, but have a better convergence rate. Some interior-point methods are polynomial time, but the complexity bounds are generally far off from practically observed ones [20]. A third class are *active set methods*: these have a good performance in practice, but suffer from the drawback that in general no polynomial complexity bounds can be given [21].

In the remainder of this section, we will propose three different optimization methods tailored to QP (2), wherein the structure of the optimization problem will be exploited.

- In Section III-B, an active set type of method will be proposed which exploits the fact that only a small subset of the constraints will influence the final solution.
- In Sections III-C and III-D, two projected gradient methods will be proposed which exploit the circulant structure of the Hessian matrix and the geometry of the convex feasible set.

B. Optimization Method 1: Dual Active Set Strategy

In [12], it was experimentally shown that general-purpose QP solvers are largely inadequate to solve instances of QP (2) in real time. Therefore, an active set optimization method was proposed that efficiently solves the *dual optimization problem* of (2).

Definition 4 (Dual Optimization Problem): For any primal optimization problem of the form (6), the dual optimization problem is defined as the convex maximization problem

$$\max_{\lambda\in\mathbb{R}^K,\nu\in\mathbb{R}^M} \underbrace{\inf_{y\in\mathbb{R}^N} \underbrace{\left[\underbrace{f(y)-\sum_{i=0}^{K-1}\lambda(i)h_i(y)-\sum_{j=0}^{M-1}\nu(j)\left(c_j^T y-d_j\right)}_{\triangleq\mathcal{L}(y,\lambda,\nu)} \right]}_{\triangleq q(\lambda,\nu)}}_{\text{s.t.}\quad \lambda\geq 0} \quad (12)$$

where λ and ν are the vectors of *Lagrange multipliers* associated to the inequality constraints and equality constraints respectively, $\mathcal{L}(y,\lambda,\nu)$ is the *Lagrangian function* and $q(\lambda,\nu)$ is the *Lagrange dual function*.

A primal optimization problem and its associated dual optimization problem are related in interesting ways. In general, the dual optimization problem can be used to obtain a lower bound on the optimal value of the objective for the primal problem. This primal-dual relationship is known as *weak duality* [22].

Theorem 3 (Weak Duality): The optimal objective value p^* of any primal optimization problem (6) and the optimal objective value d^* of the associated dual optimization problem (12) are related as follows,

$$d^* \leq p^* \quad (13)$$

One of the main advantages of convex optimization problems is the fact that the dual optimization problem can be used directly for solving the primal optimization problem. This primal-dual relationship is known as *strong duality* [22].

Theorem 4 (Strong Duality): If the primal optimization problem (6) is convex and it has a strictly feasible point, then primal optimization problem (6) and dual optimization problem (12) have the same optimal objective value,

$$d^* = p^* \quad (14)$$

Since in Section III-A optimization problem (2) was shown to be a strictly convex quadratic program, and zero is a strictly feasible point, it is clear from Theorem 4 that it has a strong duality relationship with its dual counterpart. We formulate this dual optimization problem as follows. First, the Lagrangian function $\mathcal{L}(y_m,\lambda_{m,u},\lambda_{m,l})$ is given by

$$\begin{aligned}\mathcal{L}(y_m,\lambda_{m,u},\lambda_{m,l}) &= f(y_m)-\lambda_{m,u}^T h_u(y_m)-\lambda_{m,l}^T h_l(y_m)\\ &= \frac{1}{2}(y_m-x_m)^T H_m(y_m-x_m)\\ &\quad +\lambda_{m,u}^T(y_m-u)+\lambda_{m,l}^T(l-y_m)\end{aligned} \quad (15)$$

where $\lambda_{m,u},\lambda_{m,l}\in\mathbb{R}^N$ denote the vectors of Lagrange multipliers associated to the upper clipping level constraints $h_u(y_m)=[h_0(y_m)\ldots h_{N-1}(y_m)]^T$ and the lower clipping level constraints $h_l(y_m)=[h_N(y_m)\ldots h_{2N-1}(y_m)]^T$, respectively. Then, the Lagrange dual function equals

$$\begin{aligned}q(\lambda_{m,u},\lambda_{m,l}) &= \inf_{y_m\in\mathbb{R}^N}\mathcal{L}(y_m,\lambda_{m,u},\lambda_{m,l})\\ &= -\frac{1}{2}(\lambda_{m,u}-\lambda_{m,l}-H_m x_m)^T H_m^{-1}\\ &\quad\times(\lambda_{m,u}-\lambda_{m,l}-H_m x_m)\\ &\quad -\lambda_{m,u}^T u+\lambda_{m,l}^T l+\frac{1}{2}x_m^T H_m x_m\end{aligned} \quad (16)$$

Finally, the dual optimization problem can be formulated as

$$\begin{aligned}\lambda_m^* &= \arg\max_{\lambda_m\in\mathbb{R}^{2N}} q(\lambda_m)\quad \text{s.t.}\lambda_m\geq 0\\ &= \arg\min_{\lambda_m\in\mathbb{R}^{2N}} \frac{1}{2}\lambda_m^T\underbrace{A^T H_m^{-1}A}_{\triangleq\tilde{H}_m}\lambda_m+\left(\underbrace{B^T\mathbb{1}_N-A^T x_m}_{\triangleq\tilde{g}_m}\right)^T\lambda_m\\ &\quad \text{s.t.}\quad \lambda_m\geq 0\end{aligned} \quad (17)$$

where $\lambda_m\in\mathbb{R}^{2N}$, $A\in\mathbb{R}^{N\times 2N}$ and $B\in\mathbb{R}^{N\times 2N}$ are defined as $\lambda_m=[\lambda_{m,u}^T\ \lambda_{m,l}^T]^T$, $A=[I_N|-I_N]$, and $B=[UI_N|-LI_N]$, with $I_N\in\mathbb{R}^{N\times N}$ the identity matrix.

Computation of y_m^* from λ_m^* is then straightforward,

$$\begin{aligned}y_m^* &= -H_m^{-1}(A\lambda_m^*-H_m x_m)\\ &= x_m-H_m^{-1}A\lambda_m^*\end{aligned} \quad (18)$$

The dual optimization problem formulated in (17), (18) can be solved efficiently by exploiting the fact that only a small subset of the large number $(2N)$ of inequality constraints are expected to influence the solution. Under the assumption of a moderate

clipping factor, indeed only a small number of samples in the input frame x_m are expected to exceed the clipping levels. An iterative external active set strategy is adopted, where the following steps are executed in each iteration (see Algorithm 1):[5]

1) Check which inequality constraints are violated in the previous solution iterate. In case no inequality constraints are violated, the algorithm terminates.
2) Add the violated constraints to an active set $\mathbb{S}$ of constraints to be monitored.
3) Solve a small-scale QP corresponding to (17) with those Lagrange multipliers corresponding to constraints not in the active set $\mathbb{S}$ set to zero.
4) Compute the new solution iterate by evaluating (18).

Algorithm 1 Dual active set strategy

Input $x_m \in \mathbb{R}^N, L, U, W$

Output $y_m^* \in \mathbb{R}^N$

$y_m^0 = x_m$

$\mathbb{V}_m^0 = \{i | h_i(y_m^0) < 0\}$

$\mathbb{S}_m^0 = \mathbb{V}_m^0$

$k = 0$

while $\mathbb{V}_m^k \neq \emptyset$ **do**

  $\lambda_m^{k+1}(i) = 0, i \notin \mathbb{S}_m^k$

  Calculate $\lambda_m^{k+1}(i), i \in \mathbb{S}_m^k$ as solution of small-scale QP defined in (17)

  $y_m^{k+1} = x_m - H_m^{-1} A \lambda_m^{k+1}$ [using (18)]

  Collect index set of violated inequality constraints $\mathbb{V}_m^{k+1} = \{i | h_i(y_m^{k+1}) < 0\}$

  Extend active set $\mathbb{S}_m^{k+1} = \mathbb{S}_m^k \cup \mathbb{V}_m^{k+1}$

  $k = k + 1$

end while

$y_m^* = y_m^k$

Using this strategy of dualizing and iteratively adapting an appropriate subset of inequality constraints, a QP dimensionality reduction is achieved which brings along a significant computational complexity reduction. In effect, the solution of QP (2) is found by solving a small number of small scale QPs (17) instead of by solving the full scale QP at once. From simulations, it is concluded that 4 external iterations generally suffice for solving an instance of optimization problem (2). In comparison to general-purpose dense QP solvers, the method achieves a reduction of computation time with a factor ranging from 10 up to 200. Moreover, for clipping factors[6] higher than 0.95, the method could potentially be used in a real-time clipping context [12].

Although computation times are reduced considerably, it has to be remarked that this optimization method still has a few shortcomings, possibly preventing it to be used reliably in real-time audio applications:

- First of all, the method's computational complexity is seen to increase with increasing number of violated constraints in the input frame x_m. That is, the computational complexity increases with decreasing clipping factors, making it impossible to run the method in real time for low clipping factors.
- Secondly, it is practically impossible to derive certifying polynomial complexity bounds for the optimization method.
- Lastly, the iterative optimization method cannot be stopped early (i.e. before convergence to the exact solution) to provide an approximate solution of the optimization problem.

C. Optimization Method 2: Projected Gradient Descent

In this subsection, we present a projected gradient optimization method that deals with the different issues raised in Section III-B concerning applicability in real time. First, Section III-C-1 gives a general description of the method. Then, in Section III-C-2, the selection of an appropriate stepsize is discussed. Finally, in Section III-C-3, the computation of approximate solutions is discussed and theoretical algorithmic complexity bounds are derived.

1) Description of the Method: Projected gradient methods are a class of iterative methods for solving optimization problems over convex sets. In each iteration, first a step is taken along the negative gradient direction of the objective function, after which the result is orthogonally projected onto the convex feasible set, thereby maintaining feasibility of the iterates [23]. A low computational complexity per iteration is the main asset of projected gradient methods, provided that the orthogonal projection onto the convex feasible set and the gradient of the objective function can easily be computed.

For optimization problem (2), both these elements can indeed be computed at an extremely low computational complexity, by exploiting the structure of the Hessian matrix and the convex feasible set. The main steps to be performed in the $(k+1)$th iteration of the proposed projected gradient method are as follows:

- Take a step with stepsize s_m^k along the negative gradient direction:

$$\tilde{y}_m^{k+1} = y_m^k - s_m^k \nabla f\left(y_m^k\right) \tag{19}$$

  where using (2),

$$\begin{aligned} \nabla f\left(y_m^k\right) &= H_m\left(y_m^k - x_m\right) \\ &= D^H W_m D\left(y_m^k - x_m\right) \end{aligned} \tag{20}$$

  and where the stepsize s_m^k will be defined in Section III-C-2. It is clear from (20) that the gradient computation can be performed at a very low computational complexity, by sequentially applying a DFT (multiplication by D), an element-wise weighting (multiplication by W_m), and an IDFT (multiplication by D^H) to the vector $(y_m^k - x_m)$. An alternative interpretation is that we perform a matrix-vector multiplication of the circulant matrix H_m with the vector $(y_m^k - x_m)$. By exploiting the computational efficiency of the FFT

[5]We introduce in this algorithm the notation $[\cdot]_m^k$ for the kth iterate of the mth frame:

[6]Clipping factor CF is defined as 1-(fraction of signal samples exceeding the upper or lower clipping level).

algorithm, the gradient computation thus has a complexity of $O(N \log_2 N)$.

- Project $\tilde{y}_m^{k+1}$ orthogonally onto the convex feasible set Q of (2), which is defined as

$$Q = \{y_m \in \mathbb{R}^N | l \leq y_m \leq u\} \tag{21}$$

The feasible set can be thought of as an N-dimensional box. An orthogonal projection $\Pi_Q(\tilde{y}_m^{k+1})$ onto this N-dimensional box boils down to performing a simple componentwise hard clipping operation (with lower bound L and upper bound U), i.e.

$$y_m^{k+1} = \Pi_Q\left(\tilde{y}_m^{k+1}\right) = \arg\min_{z \in Q} \frac{1}{2} \left\| z - \tilde{y}_m^{k+1} \right\|_2^2 \tag{22}$$

where

$$y_m^{k+1}(i) = \begin{cases} L & \text{if} \quad \tilde{y}_m^{k+1}(i) < L \\ \tilde{y}_m^{k+1}(i) & \text{if} \quad L \leq \tilde{y}_m^{k+1}(i) \leq U \\ U & \text{if} \quad \tilde{y}_m^{k+1}(i) > U, \end{cases} \quad i = 0 \ldots N-1 \tag{23}$$

2) Stepsize Selection: Several rules for selecting stepsizes s_m^k in projected gradient methods have been proposed in literature, e.g. fixed stepsizes, diminishing stepsizes, or line search rules [23]. Here a fixed stepsize is used, thereby avoiding the additional computational complexity incurred by line searches. In [19], it is shown that by choosing a fixed stepsize

$$s_m^k = \frac{1}{C_m}, \quad \forall k \geq 0 \tag{24}$$

with C_m the Lipschitz constant of the gradient ∇f of (1) on the set Q (for frame m), a limit point of the sequence $\{y_m^k\}$ obtained by iteratively applying (19) and (23) is a stationary point. Because of the convexity of f, it is a local minimum and hence a global minimum.

In order to establish the Lipschitz constant C_m of our problem, we introduce the next lemma.

Lemma 1: Let function f be twice continuously differentiable on set Q. The gradient ∇f is Lipschitz continuous on set Q with Lipschitz constant C if and only if

$$\left\| \nabla^2 f(z) \right\| \leq C, \quad \forall z \in Q \tag{25}$$

In other words, the Lipschitz constant of the gradient ∇f can be seen as an upper bound to the curvature of the objective function f. Using this lemma, we can easily show that the Lipschitz constant C_m is computed as

$$\begin{aligned} C_m &= \|H_m\| \\ &= \max_{1 \leq i \leq N} \beta_i(H_m) \\ &= \max_{1 \leq i \leq N} \beta_i(D^H W_m D) \\ &= \max_{0 \leq i \leq N-1} w_m(i) \end{aligned} \tag{26}$$

where $\beta_i(H_m)$, $i = 1 \ldots N$, denote the eigenvalues of the Hessian matrix H_m.

Algorithm 2 Projected gradient method

Input $x_m \in \mathbb{R}^N$, $y_m^0 \in Q, L, U, W_m$
Output $y_m^* \in \mathbb{R}^N$
1: $k = 0$
2: Calculate Lipschitz constant C_m [using (26)]
3: **while** convergence is not reached **do**
4: $\quad \tilde{y}_m^{k+1} = y_m^k - (1/C_m)\nabla f(y_m^k)$ [using (20)]
5: $\quad y_m^{k+1} = \Pi_Q(\tilde{y}_m^{k+1})$ [using (23)]
6: $\quad k = k + 1$
7: **end while**
8: $y_m^* = y_m^k$

3) Algorithmic Complexity and Approximate Solutions: The proposed projected gradient optimization method is summarized in Algorithm 2. Clearly, the computational complexity of one iteration is seen to be extremely low. Moreover, the shortcomings of the optimization method 1 presented in Section III-B are dealt with:

- It is possible to solve the optimization problem inexactly by stopping the iterative optimization method before convergence to the exact solution y_m^* is reached. The iterates y_m^k of the proposed projected gradient method are feasible by construction. Moreover, the sequence $\{f(y_m^k)\}$ can be proved to be monotonically decreasing. Hence, stopping the method after any number of iterations η will result in a feasible point y_m^η for which $f(y_m^\eta) \leq f(y_m^0)$. We can then define the obtained solution accuracy as $\epsilon = f(y_m^\eta) - f(y_m^*)$.
- It is possible to derive polynomial upper and lower bounds on the *algorithmic complexity*, i.e. the number of necessary iterations of the optimization method as a function of the obtained solution accuracy ϵ, as we will show next.

For the class of convex optimization problems with smooth objective functions, a general lower bound on the algorithmic complexity was derived that holds for all iterative methods that use only first-order (i.e. gradient) information [19, Ch. 2]:

Theorem 5: For any starting point $y_m^0 \in \mathbb{R}^N$, for any first-order projected gradient method, and for any closed convex set Q,there exists a convex, continuously differentiable function f with Lipschitz continuous gradient with constant C_m, such that we have

$$f\left(y_m^k\right) - f\left(y_m^*\right) \geq \frac{3 C_m \left\| y_m^0 - y_m^* \right\|^2}{32(k+1)^2}, \quad \forall k \leq \frac{1}{2}(N-1) \tag{27}$$

Also, for the same problem class, an upper bound on the algorithmic complexity of the projected gradient method used in Algorithm 2 was derived [19, Ch. 2]:

Theorem 6: Let f be a convex, continuously differentiable function with Lipschitz continuous gradient with Lipschitz constant C_m, and let Q be a convex feasible set. Then the projected

gradient method described in Algorithm 2 with fixed stepsize $1/C_m$ generates a sequence $\{y_m^k\}$ which converges as follows:

$$f\left(y_m^k\right) - f\left(y_m^*\right) \leq \frac{2C_m \left\|y_m^0 - y_m^*\right\|^2}{k+4}, \quad \forall k \qquad (28)$$

This method is said to have a *sublinear rate of convergence*. As the general algorithmic complexity lower bound (27) for first-order optimization methods, and the specific algorithmic complexity upper bound (28) for the projected gradient optimization method described in Algorithm 2 differ by an order of magnitude, it can be concluded that the proposed optimization method is not optimal in terms of convergence.

Algorithm 3 Optimal projected gradient method

Input $x_m \in \mathbb{R}^N$, $y_m^0 = c_m^0 \in Q$, $\gamma_m^0 \in (0,1)$, L, U, W_m
Output $y_m^* \in \mathbb{R}^N$
1: Calculate Lipschitz constant C_m [using (26)]
2: Calculate convexity parameter μ_m [using (29)]
3: $\kappa_m = C_m/\mu_m$
4: $k = 0$
5: **while** convergence is not reached **do**
6: $\tilde{y}_m^{k+1} = c_m^k - (1/C_m)\nabla f(c_m^k)$
7: $y_m^{k+1} = \Pi_Q(\tilde{y}_m^{k+1})$ [using (23)]
8: Calculate γ_m^{k+1} from $\left(\gamma_m^{k+1}\right)^2 = (1-\gamma_m^{k+1})(\gamma_m^k)^2 + \kappa_m\gamma_m^{k+1}$
9: $\delta_m^k = \gamma_m^k(1-\gamma_m^k)/((\gamma_m^k)^2 + \gamma_m^{k+1})$
10: $c_m^{k+1} = y_m^{k+1} + \delta_m^k(y_m^{k+1} - y_m^k)$
11: $k = k+1$
12: **end while**
13: $y_m^* = y_m^k$

D. Optimization Method 3: Optimal Projected Gradient Descent

From the complexity bounds given in Section III-C it might be expected that it is theoretically possible for a first-order optimization method to have a better convergence rate than optimization method 2. Indeed, if there exists a first-order method for which the complexity upper bound is proportional to the complexity lower bound for a given problem class, this method could be called *optimal* for that problem class [19]. In this subsection, we present a projected gradient optimization method that reaches an optimal convergence for the class of convex optimization problems with *strongly convex* objective functions. This method was first proposed in [19] and variants of the method have been applied in diverse applications, e.g. for real-time model predictive control [20].

In Section III-C, only Lipschitz continuity of the gradient of the convex objective function was assumed, where the Lipschitz constant indicates an upper bound on the curvature of the objective function. Strong convexity assumes that in addition also a lower bound on the curvature of the objective function can be found, determined by the *convexity parameter* μ_m.

Lemma 2: Let function f be twice continuously differentiable on set Q. The function f is strongly convex on set Q with convexity parameter μ if and only if there exists $\mu > 0$ such that

$$\nabla^2 f(z) \geq \mu I, \quad \forall z \in Q$$

Using the former lemma, we can prove that for our objective function the convexity parameter μ_m can be computed as

$$\begin{aligned}\mu_m &= \min_{1\leq i\leq N} \beta_i(H_m) \\ &= \min_{1\leq i\leq N} \beta_i(D^H W_m D) \\ &= \min_{0\leq i\leq N-1} w_m(i)\end{aligned} \qquad (29)$$

The ratio $\kappa_m = C_m/\mu_m$ is called the *condition number*.

Algorithm 3 summarizes the optimal projected gradient optimization method. We note the following differences compared to Algorithm 2:

- Knowledge of the convexity parameter μ_m is incorporated.
- In each iteration, a standard projected gradient step is performed on a potentially infeasible weighted sum of two previous feasible iterates.

It is again possible to derive polynomial upper and lower bounds on the *algorithmic complexity*, i.e. the number of necessary iterations of this optimization method as a function of the solution accuracy ϵ. For the class of convex optimization problems with strongly convex objective functions, a general lower bound on the algorithmic complexity was derived that holds for all iterative first-order methods [19, Ch. 2]:

Theorem 7: For any starting point $y_m^0 \in \mathbb{R}^N$, for any first-order projected gradient method, and for any closed convex set Q, there exists a strongly convex, continuously differentiable function f with Lipschitz constant C_m and convexity parameter $\mu_m = C_m/\kappa_m$ (where $\kappa_m > 1$), such that we have

$$f\left(y_m^k\right) - f\left(y_m^*\right) \geq \frac{\mu_m}{2}\left(\frac{\sqrt{\kappa_m}-1}{\sqrt{\kappa_m}+1}\right)^{2k}\left\|y_m^0 - y_m^*\right\|^2, \quad \forall k \qquad (30)$$

Also, for the same problem class, an upper bound on the algorithmic complexity of the projected gradient method used in Algorithm 3 was derived [19, Ch. 2]:

Theorem 8: Let f be a strongly convex, continuously differentiable function with Lipschitz constant C_m and convexity parameter $\mu_m = C_m/\kappa_m$ (where $\kappa_m > 1$), and let Q be a convex feasible set. Then the projected gradient method described in Algorithm 3 generates a sequence $\{y_m^k\}$ which converges as follows:

$$\begin{aligned}&f\left(y_m^k\right) - f\left(y_m^*\right) \\ &\leq C_m \min\left\{\left(1-\sqrt{\frac{\mu_m}{C_m}}\right)^k, \frac{4}{(k+2)^2}\right\}\left\|y_m^0 - y_m^*\right\|^2\end{aligned} \qquad (31)$$

This method is said to have a *linear rate of convergence*.

From (30), the minimum number of necessary iterations to find y_m^k satisfying $f(y_m^k) - f(y_m^*) \leq \epsilon$ is

$$k_m^{min} = \frac{\sqrt{\kappa_m}-1}{4}\left(\log\frac{1}{\epsilon} + \log\frac{\mu_m}{2} + 2\log\left\|y_m^0 - y_m^*\right\|\right) \qquad (32)$$

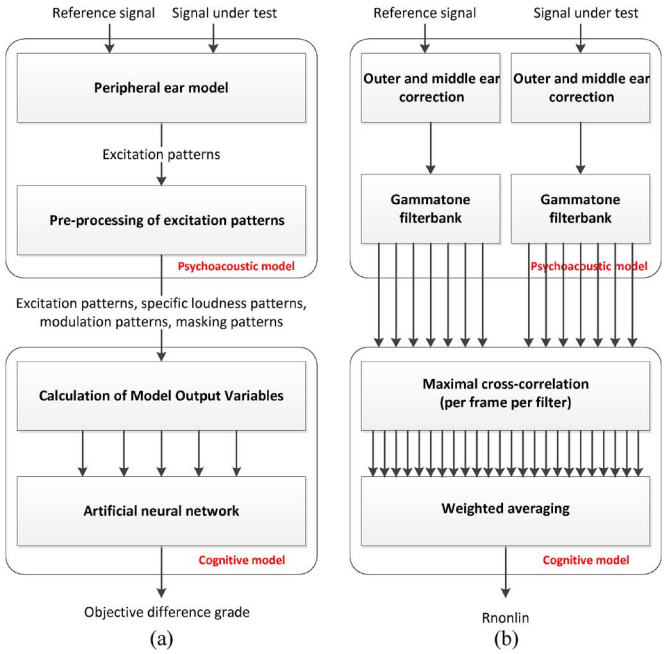


Fig. 3. Block diagrams of the used objective measures of perceived audio quality: (a) PEAQ basic version (adapted from [25]) (b) Rnonlin.

TABLE I
ALGORITHMIC AND ARITHMETIC COMPLEXITY OF OPTIMIZATION METHODS 2 AND 3

Algorithm	Algorithmic complexity	Arithmetic complexity per iteration
Method 2	$\frac{8C_m NU^2}{\epsilon} - 4$	$\frac{51}{9} N \log_2 N - \frac{61}{18} N$
Method 3	$\sqrt{\kappa_m} \log \frac{4C_m NU^2}{\epsilon}$	$\frac{51}{9} N \log_2 N - \frac{7}{18} N$

From (31), the maximum number of necessary iterations to find y_m^k satisfying $f(y_m^k) - f(y_m^*) \leq \epsilon$ is

$$k_m^{max} = \sqrt{\kappa_m} \left(\log \frac{1}{\epsilon} + \log C_m + 2 \log \left\| y_m^0 - y_m^* \right\| \right) \quad (33)$$

Hence, the main term in the upper bound estimate (33), $\sqrt{\kappa_m} \log(1/\epsilon)$, is proportional to the lower bound (32), proving this optimization method to be an optimal first-order method for the class of strongly convex optimization problems.

Table I summarizes the computational complexity results of optimization methods 2 and 3. For both methods, the algorithmic complexity as well as the arithmetic complexity per iteration (in terms of number of real additions and multiplications) is given. The algorithmic complexity results are straightforwardly found from upper bounds (28) and (31) by incorporating $\|y_m^0 - y_m^*\| = 2\sqrt{N}U$ as the (worst-case) maximum distance between two points in the feasible set Q defined in (21),[7] where it is assumed that $U = -L$. In the arithmetic complexity computations, estimates were used for the arithmetic complexity of a FFT with power-of-two length N, as derived in [24]: namely, $(17/9)N \log_2 N - (89/27)N + O(\log_2 N)$ for a real-data FFT, and $(34/9)N \log_2 N - (124/27)N + O(\log_2 N)$ for a complex-data FFT. In conclusion, optimization method 3 has a significantly better $O(\log(1/\epsilon))$ algorithmic complexity compared to the $O(1/\epsilon)$ algorithmic complexity of optimization method 2, and this for a negligibly higher arithmetic complexity per iteration.

[7]The length of an N-agonal in an N-dimensional hypercube with side length $2U$ equals $2\sqrt{N}U$.

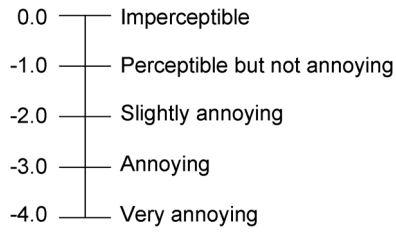


Fig. 4. The ITU-R five-grade impairment scale.

IV. Simulation Results

A. Comparative Evaluation of Perceived Audio Quality

For audio quality evaluation purposes, a test database consisting of 24 audio excerpts was compiled (16 bit mono, 44.1 kHz). The excerpts were selected so as to cover different music styles, melodic and rythmic textures, instrumentations and dynamics, as well as different speech types (see Table II for details). A first set of excerpts (numbers 1–11) was extracted from different commercial audio CDs. A second set of excerpts (numbers 12–16) was extracted from an ITU CD-ROM associated to Recommendation BS.1387-1, which contains a database (DB3) used for validating the conformance of implementations to this recommendation [25]. A third set of excerpts (numbers 17–20) was extracted from the HINT speech database [26]. A fourth set of excerpts (numbers 21–24) was extracted from the VoxForge speech corpus [27].

Each audio signal in the test database was processed by three different clipping algorithms:

- *Hard symmetrical clipping*, where the input-output characteristic is defined as

$$y[n] = \begin{cases} x[n], & |x[n]| \leq U \\ \operatorname{sgn}(x[n])\, U, & |x[n]| > U \end{cases} \quad (34)$$

- *Soft symmetrical clipping* [3], where the input-output characteristic is defined as a linearized hyperbolic tangent function which is linear for inputs below a parametric amplitude level r,

$$y[n] = \begin{cases} x[n], & |x[n]| \leq r \\ \operatorname{sgn}(x[n]) \left[(U-r) \tanh\left(\frac{|x[n]|-r}{U-r} \right) + r \right], & |x[n]| > r \end{cases} \quad (35)$$

  here used with parameter setting $r = 0.8U$.
- *Perception-based clipping* as described in this paper, with parameter values $N = 512$, $P = 128$, $\alpha = 0.04$, and application of optimization method 3 with a solution accuracy of $\epsilon = 10^{-12}$ for all instances of (2).

This processing was performed for eight clipping factors {0.70, 0.80, 0.85, 0.90, 0.95, 0.97, 0.98, 0.99}.[8] For each of a resulting

[8]Note that the clipping factor above which a normal-hearing listener does not perceive hard clipping distortion has been subjectively evaluated to be higher than 0.99 for speech [39], and 0.997 for music [40].

TABLE II
AUDIO EXCERPTS DATABASE USED FOR COMPARATIVE AUDIO QUALITY EVALUATION OF CLIPPING ALGORITHMS

Nr.	Name	Texture	Composition	Style	Duration [s]	$\text{Sample}_{\text{start}}$	$\text{Sample}_{\text{end}}$	Origin
1	poulenc.wav	polyphonic	instrumental	classical	17.8	400000	1183000	[28]
2	rhcp.wav	polyphonic	instrumental	rock	9.8	468996	900000	[29]
3	pierle.wav	polyphonic	instrumental+vocal	pop	11.7	2234000	2750000	[30]
4	mascagni.wav	polyphonic	instrumental+vocal	classical	16.1	1	711000	[31]
5	chopin.wav	monophonic	instrumental	classical	17.8	50000	836200	[32]
6	kraftwerk.wav	polyphonic	instrumental	electronic	17.2	7480000	8240000	[33]
7	baloji.wav	polyphonic	instrumental+vocal	hip hop	14.4	8084000	8719000	[34]
8	arcade.wav	polyphonic	instrumental+vocal	pop	13.6	1	600000	[35]
9	strokes.wav	polyphonic	instrumental	rock	13.6	1	600000	[36]
10	beethoven.wav	monophonic	instrumental	classical	15.4	1580000	2260000	[37]
11	albeniz.wav	monophonic	instrumental	classical	11.8	7800000	8320000	[38]
12	breftri.wav	monophonic	instrumental	classical	19.7	1	869675	[25]
13	crefsax.wav	monophonic	instrumental	classical	10.9	1	479026	[25]
14	freftr1.wav	monophonic	instrumental	classical	13.1	1	577887	[25]
15	grefcla.wav	monophonic	instrumental	classical	6.9	1	302534	[25]
16	mrefcla.wav	monophonic	instrumental	classical	7.4	1	326500	[25]
17-20	hint{01-04}.wav	monophonic	speech	male	4×17.0	1	749700	[26]
21-24	silvia{01-04}.wav	monophonic	speech	female	4×8.0	1	352800	[27]

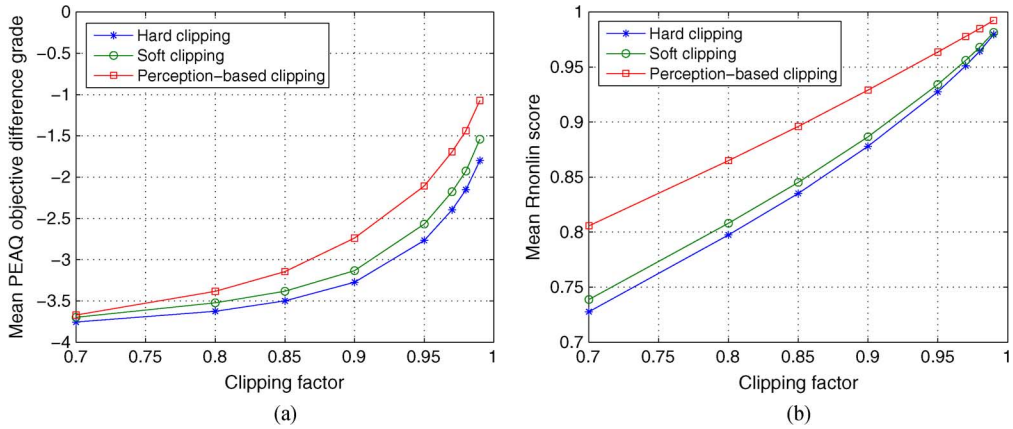


Fig. 5. Comparative evaluation of different clipping techniques in terms of objective perceived audio quality: (a) mean PEAQ ODG and (b) mean Rnonlin scores for signals processed by hard clipping, soft clipping and perception-based clipping as a function of the clipping factor. (a) PEAQ basic version; (b) Rnonlin.

total of $24 \times 8 \times 3 = 576$ processed audio signals, two *objective measures of perceived audio quality* were calculated, which aim to predict the subjective audio quality score that would be attributed to the processed audio signal by an average human listener. Taking a reference signal (i.e. the clean signal) and a signal under test (i.e. the processed signal) as an input, such an objective measure of perceived audio quality is calculated through sequential application of a psychoacoustic model and a cognitive model, and resultingly attributes a perceived audio quality score to the signal under test with respect to the reference signal (see Fig. 3).

A first objective measure of perceived audio quality was calculated using the Basic Version of the PEAQ (*Perceptual Evaluation of Audio Quality*) recommendation [25]. A block diagram representation of this method is shown in Fig. 3(a). The resulting Objective Difference Grade (ODG) predicts the basic audio quality of the signal under test with respect to the reference signal, and has a range between 0 and -4, corresponding to the ITU-R five grade impairment scale depicted in Fig. 4.

A second objective measure used here was specifically designed to predict the perceived audio quality of nonlinearly distorted signals and is described in [41]. A block diagram representation of this method is shown in Fig. 3(b). The resulting *Rnonlin* score is a perceptually relevant measure of distortion. Rnonlin values have a range between 0 and 1 and are seen to decrease for increasing perceptible distortion (i.e. with decreasing audio quality).

The results of these simulations are shown in Fig. 5. In Fig. 5(a), the average PEAQ ODG score over all 24 audio signals is plotted as a function of the clipping factor, and this for the three different clipping techniques. Analogously, Fig. 5(b) shows the results for the Rnonlin measure. The obtained results for both audio quality measures are seen to be in accordance with each other. Logically, we observe a monotonically increasing average audio quality score for increasing clipping factors. Soft clipping is seen to result in slightly higher average objective audio quality scores than hard clipping. Clearly, the perception-based clipping technique is seen to result in significantly higher average objective audio quality scores than the other clipping techniques, and this for all considered clipping factors. In Table III, the full simulation results per individual audio excerpt are provided for two selected clipping factors (0.99 and 0.90). For each audio excerpt, the highest score for each objective measure is highlighted.

In order to infer statistically significant conclusions on the comparative audio quality performance of the three clipping algorithms under study, a statistical analysis was performed on the obtained set of PEAQ ODG and Rnonlin scores. Let us represent

TABLE III
OBJECTIVE AUDIO QUALITY SCORES FOR HARD CLIPPING (H), SOFT CLIPPING (S) AND PERCEPTION-BASED CLIPPING (P-B) FOR TWO SELECTED CLIPPING FACTORS {0.99, 0.90}: PEAQ ODG AND RNONLIN. HIGHEST SCORES PER AUDIO EXCERPT ARE HIGHLIGHTED

Nr.	PEAQ						Rnonlin					
	CF=0.99			CF=0.90			CF=0.99			CF=0.90		
	H	S	P-B	H	S	P-B	H	S	P-B	H	S	P-B
1	-1.86	-1.79	-1.13	-3.85	-3.84	-3.58	99.14	99.15	99.57	87.72	87.87	93.22
2	-0.38	-0.31	-0.36	-0.93	-0.94	-0.84	99.29	99.38	99.56	92.68	93.54	95.99
3	0.13	0.11	-0.17	-1.26	-0.74	-1.34	99.93	99.87	99.97	95.35	95.88	97.61
4	-2.98	-2.10	-2.03	-3.86	-3.84	-3.77	98.62	98.77	99.41	88.60	89.68	93.37
5	-3.09	-2.21	-1.08	-3.89	-3.88	-3.71	96.97	97.01	98.94	79.89	81.52	89.46
6	-0.62	-0.52	-0.16	-2.55	-2.31	-1.52	99.12	99.22	99.71	91.29	91.76	95.94
7	-0.42	-0.39	-0.14	-1.89	-1.64	-1.07	98.53	98.63	99.55	90.64	91.18	94.73
8	-0.24	-0.21	-0.14	-1.77	-1.28	-2.00	99.41	99.46	99.70	93.36	94.14	95.45
9	-0.51	-0.44	-0.40	-3.75	-3.68	-3.51	98.93	99.03	99.52	88.36	88.99	93.46
10	-3.71	-1.65	-1.71	-3.82	-3.82	-3.41	98.32	98.99	99.51	90.02	91.58	94.79
11	-1.65	-0.67	-0.59	-3.70	-3.44	-1.82	97.63	98.22	99.26	85.43	87.10	92.34
12	-1.92	-1.93	-1.42	-3.80	-3.80	-3.69	98.67	98.66	99.50	90.68	90.68	94.07
13	-0.67	-0.46	-0.45	-3.21	-2.47	-2.25	97.86	98.31	99.44	85.07	86.08	91.13
14	-3.45	-3.46	-0.88	-3.85	-3.85	-3.60	96.35	96.25	99.54	87.84	87.75	94.37
15	-1.41	-1.37	-0.95	-3.56	-3.25	-3.17	97.41	97.40	99.02	87.35	88.21	93.16
16	-1.59	-1.50	-1.07	-3.64	-3.40	-3.35	97.46	97.46	99.03	87.62	88.32	93.07
17	-3.86	-3.78	-3.06	-3.91	-3.91	-3.83	95.47	95.93	98.14	82.61	83.99	88.90
18	-3.84	-3.73	-2.67	-3.91	-3.91	-3.85	96.54	96.93	98.49	82.49	83.84	88.94
19	-3.80	-3.51	-2.81	-3.90	-3.89	-3.78	96.12	96.73	98.42	82.42	84.24	89.14
20	-3.80	-3.61	-2.38	-3.90	-3.90	-3.81	96.61	97.08	98.55	84.42	85.78	90.25
21	-1.07	-1.10	-0.55	-3.57	-3.55	-2.16	98.31	98.37	99.24	88.26	88.77	92.37
22	-1.00	-0.98	-0.56	-3.42	-3.36	-1.99	98.06	98.22	99.12	88.88	89.43	92.64
23	-0.80	-0.76	-0.55	-3.40	-3.37	-2.02	97.91	98.07	99.19	88.38	88.99	92.85
24	-0.62	-0.60	-0.47	-3.18	-3.14	-1.63	98.25	98.34	99.27	87.86	88.52	92.44
Avg.	**-1.80**	**-1.54**	**-1.07**	**-3.27**	**-3.13**	**-2.74**	**97.95**	**98.14**	**99.24**	**87.80**	**88.66**	**92.90**

the audio quality scores resulting from hard clipping, soft-clipping and perception-based clipping for a given clipping factor by random variables S_H, S_S, and S_{PB}, respectively. Under the assumption that these random variables follow a normal probability distribution,[9] we tested the two following statistical hypotheses based on the sample data. The first null hypothesis H_0 and its alternative H_1 are formulated as follows,

$$H_0 : E[S_H] \geq E[S_{PB}] \tag{36}$$
$$H_1 : E[S_{PB}] > E[S_H] \tag{37}$$

The second null hypothesis H_0 and its alternative H_1 are formulated as follows,

$$H_0 : E[S_S] \geq E[S_{PB}] \tag{38}$$
$$H_1 : E[S_{PB}] > E[S_S] \tag{39}$$

These two statistical hypotheses were tested for all considered clipping factors, and for both audio quality measures. All statistical hypotheses were tested using one-tailed paired t-tests

[9]The validity of this assumption was verified for our sample data using the Jarque-Bera normality test [42] at significance level $\alpha_{JB} = 0.05$.

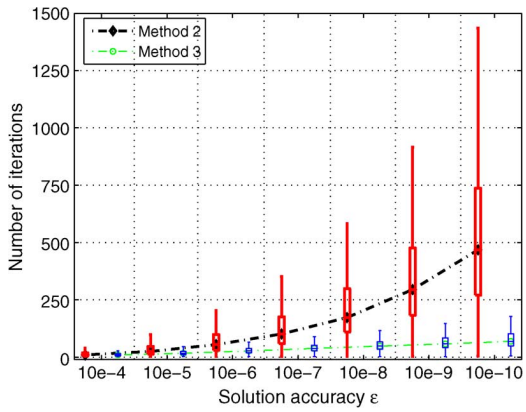


Fig. 6. Boxplot of number of iterations vs solution accuracy for optimization methods 2 and 3.

with significance level $\alpha_{TT} = 0.05$. The resulting one-sided P-values are synthesized in Table IV. For PEAQ scores, the first null hypothesis (36) can be rejected in favor of the alternative (37) for clipping factors of 0.80 and higher. The second null hypothesis (38) can be rejected in favor of the alternative (39) for clipping factors of 0.85 and higher. For Rnonlin scores, both

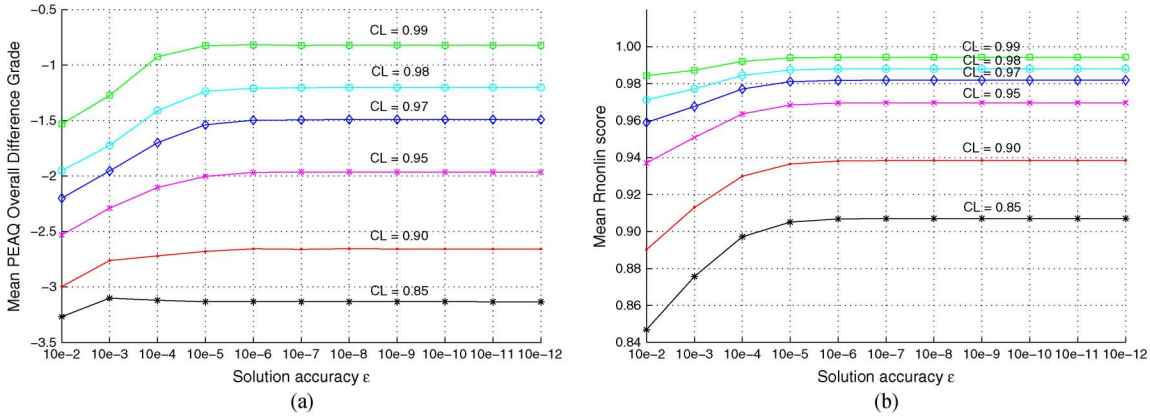


Fig. 7. Mean objective audio quality scores for different solution accuracies for perception-based clipping (a) PEAQ ODG (b) Rnonlin.

TABLE IV
P-VALUES FROM ONE-TAILED PAIRED T-TESTS ON AUDIO QUALITY SCORES. SIGNIFICANT P-VALUES WITH RESPECT TO $\alpha_{TT} = 0.05$ IN BOLD

H_0	$E[S_H] \geq E[S_{PB}]$		$E[S_S] \geq E[S_{PB}]$	
	PEAQ	Rnonlin	PEAQ	Rnonlin
CF=0.70	0.0766	**<0.0001**	0.3566	**<0.0001**
CF=0.80	**0.0042**	**<0.0001**	0.1106	**<0.0001**
CF=0.85	**0.0017**	**<0.0001**	**0.0398**	**<0.0001**
CF=0.90	**<0.0001**	**<0.0001**	**0.0026**	**<0.0001**
CF=0.95	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
CF=0.97	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
CF=0.98	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
CF=0.99	**<0.0001**	**<0.0001**	**0.0004**	**<0.0001**

null hypotheses can be rejected in favor of the alternative for all considered clipping factors. We can conclude that there is strong statistical evidence that the perception-based clipping technique will in general deliver signals with a higher perceptual audio quality compared to the other considered clipping techniques, and this for moderate to high clipping factors.

B. Experimental Evaluation of Algorithmic Complexity

In order to assess experimentally the validity of the theoretical algorithmic complexity bounds of the projected gradient optimization methods 2 and 3 described in Sections III-C and III-D, a simulation was conducted as follows. For optimization methods 2 and 3, the number of iterations needed to reach solution accuracies $\epsilon = \{10^{-4}, 10^{-5}, \ldots, 10^{-10}\}$ was determined for a subset of the instances of optimization problem (2) occurring in our test database of 24 audio signals. This was performed for six clipping factors {0.85, 0.90, 0.95, 0.97, 0.98, 0.99}.

In Fig. 6, the simulation results per optimization method are summarized in the form of a boxplot for each solution accuracy, depicting graphically the minimum, lower quartile, median, upper quartile and maximum values of the number of iterations. The dotted lines connect the median number of iterations of both optimization methods for different solution accuracies. We observe that the median (the same holds for the maximum) number of iterations follows a different curve depending on the optimization method: for optimization method 2, we observe an exponential growth for logarithmically decreasing solution accuracies, whereas for optimization method 3, the curve shows a linear growth. These simulation results provide evidence for the greatly improved algorithmic complexity of method 3 $(O(\log(1/\epsilon)))$ compared to method 2 $(O(1/\epsilon))$, as was derived theoretically in Sections III-C and III-D.

C. Applicability in Real-Time Context: Effect of Solution Accuracy ϵ on Perceived Audio Quality

In a real-time processing context, the number of clock cycles that can be spent on solving an instance of optimization problem (2) is strictly limited. In view of this, several computationally efficient convex optimization methods tailored to the optimization problem were presented in Section III. It was shown that the iterative projected gradient optimization methods 2 and 3 have the advantage that approximate solutions can be computed, which makes it possible to adhere to the imposed real-time deadlines. The question remains as to how approximately solving the optimization problems affects the perceived audio quality of the resulting output signal. In order to assess this effect, PEAQ ODG an Rnonlin scores were calculated for a subset of the signals in our test database, each of which was processed by perception-based clipping for optimization problem solution accuracies $\epsilon = \{10^{-2}, 10^{-3}, \ldots, 10^{-12}\}$. This was performed for six clipping factors {0.85, 0.90, 0.95, 0.97, 0.98, 0.99}.

In Figs. 7(a) and 7(b), the resulting mean PEAQ ODG and Rnonlin scores over all audio signals are plotted as a function of the solution accuracy, and this for all considered clipping factors. We observe that, according to both measures, the mean audio quality is affected negatively for low solution accuracies, increases nearly monotonically with increasing solution accuracies, and saturates at a solution accuracy that depends on the clipping factor. Irrespective of the clipping factor, we observe that no further improvement in mean audio quality scores is obtained for higher solution accuracies than 10^{-6}. Hence, $\epsilon = 10^{-6}$ can be put forward as an experimentally established sufficient solution accuracy for all considered clipping factors, such that no sacrifice in terms of audio quality is made.

For parameter values $N = 512$, $P = 128$ and sampling rate of 44.1 kHz, the real-time computation time limit for solving one instance of optimization problem (2) is equal to 8.7 ms. In our simulation setting,[10] this corresponds to an iteration limit of roughly 200 iterations for optimization methods 2 or 3 (neglecting here the small difference in arithmetic complexity per iteration between both methods as shown in Table I). Looking

[10]All simulations were performed on a GenuineIntel CPU @2826 Mhz

back at Fig. 6, we see that even for the worst-case instances in our test database, optimization method 3 meets the real-time iteration limit for solution accuracies up to 10^{-10}, which largely surpasses the required solution accuracy of 10^{-6}.

V. Conclusions

In this paper, we have presented a novel algorithm for real-time perception-based clipping of audio signals. By including a psychoacoustic model and embedding convex optimization into the algorithm, it is possible to explicitly minimize the perceptible distortion introduced by clipping. From comparative audio quality evaluation experiments, it has been concluded that the perception-based clipping algorithm results in significantly higher objective audio quality scores than standard clipping techniques, and this for moderate to high clipping factors. Furthermore, three optimization methods aimed at efficiently solving the convex optimization problems were derived. The reliable use of optimization method 3 in real-time applications was seen to be supported by theoretically derived complexity bounds as well as by simulation experiments.

In a broader view, the results presented in this paper suggest that embedded convex optimization is a very promising paradigm in real-time audio processing applications, with numerous potential applications, of which we point out e.g. speech enhancement and acoustic echo cancellation.

References

[1] G. A. Constantinides, P. Y. K. Cheung, and W. Luk, "Synthesis of saturation arithmetic architectures," *ACM Trans. Design Automation of Electron. Syst.*, vol. 8, no. 3, pp. 334–354, Jul. 2003.

[2] U. Zölzer *et al., DAFX: Digital Audio Effects*, U. Zölzer, Ed. New York: Wiley , May 2002.

[3] A. N. Birkett and R. A. Goubran, "Nonlinear loudspeaker compensation for hands free acoustic echo cancellation," *Electron. Lett.*, vol. 32, no. 12, pp. 1063–1064, Jun. 1996.

[4] F. Foti, "Aliasing distortion in digital dynamics processing, the cause, effect, and method for measuring it: The story of 'digital grunge!'," in *Preprints AES 106th Conv.*, Munich, Germany, May 999, Preprint no. 4971.

[5] C.-T. Tan, B. C. J. Moore, and N. Zacharov, "The effect of nonlinear distortion on the perceived quality of music and speech signals," *J. Audio Eng. Soc.*, vol. 51, no. 11, pp. 1012–1031, Nov. 2003.

[6] C.-T. Tan and B. C. J. Moore, "Perception of nonlinear distortion by hearing-impaired people," *Int. J. Audiol.*, vol. 47, pp. 246–256, May 2008.

[7] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. D. Plumbley, "A constrained matching pursuit approach to audio declipping," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Prague, Czech Republic, May 2011, pp. 329–332.

[8] J. Han, G. J. Mysore, and B. Pardo, "Audio imputation using the non-negative hidden Markov model," *Lecture Notes in Computer Science: Latent Variable Analysis and Signal Separation*, 2012.

[9] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–515, Apr. 2000.

[10] D. De Koning and W. Verhelst, "On psychoacoustic noise shaping for audio requantization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Hong Kong, Apr. 2003, pp. 413–416.

[11] J. Mattingley and S. Boyd, "Real-time convex optimization in signal processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 50–61, May 2010.

[12] B. Defraene, T. van Waterschoot, H. J. Ferreau, M. Diehl, and M. Moonen, "Perception-based clipping of audio signals," in *Proc. 2010 Eur. Signal Process. Conf. (EUSIPCO-2010)*, Aalborg, Denmark, Aug. 2010, pp. 517–521.

[13] B. Defraene, T. van Waterschoot, M. Diehl, and M. Moonen, "A fast projected gradient optimization method for real-time perception-based clipping of audio signals," in *Proc. ICASSP '11*, Prague, Czech Republic, May 2011, pp. 333–336.

[14] ISO/IEC, 11172-3 Information Technology—Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s—Part 3: Audio 1993.

[15] E. Terhardt, "Calculating virtual pitch," *Hearing Res.*, vol. 1, no. 2, pp. 155–182, 1979.

[16] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, 2nd ed. New York: Springer, 1999.

[17] P. Davis, *Circulant Matrices*. New York: Wiley, 1979.

[18] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Stanford, CA: Stanford Univ., 2001.

[19] Y. Nesterov, *Introductory Lectures on Convex Optimization*. New York: Springer, 2004.

[20] S. Richter, C. Jones, and M. Morari, "Real-time input-constrained MPC using fast gradient methods," in *Proc. Conf. Decision and Control (CDC)*, Shanghai, China, Dec. 2009, pp. 7387–7393.

[21] H. J. Ferreau, H. G. Bock, and M. Diehl, "An online active set strategy to overcome the limitations of explicit MPC," *Int. J. Robust Nonlinear Control*, vol. 18, pp. 816–830, Jul. 2008.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[23] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[24] S. G. Johnson and M. Frigo, "A modified split-radix FFT with fewer arithmetic operations," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 111–119, Jan. 2007.

[25] *Method for Objective Measurements of Perceived Audio Quality*, International Telecommunications Union Recommendation BS.1387, 1998.

[26] M. Nilsson, S. D. Soli, and J. A. Sullivan, "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Amer.*, vol. 95, no. 2, pp. 1085–1099, 1994.

[27] VoxForge, Free Speech Corpus, [Online]. Available: http://www.voxforge.org Jan. 2012

[28] F. Poulenc, "Sonata for flute and piano (Cantilena)," French Flute Music, Naxos 8.557328, 2005.

[29] Red Hot Chili Peppers, Californication, Californication, Warner Bros. Records 9362473862, 1999.

[30] An Pierlé & White Velvet, Mexico, An Pierlé & White Velvet, PIAS Recordings 941.0170.020, 2006.

[31] P. Mascagni, Cavalleria Rusticana, [Online]. Available: ftp://ftp.esat.kuleuven.be/pub/SISTA/bdefraen/reports/mascagni.wav

[32] F. Chopin, "Waltz op. 69 no. 2," Favourite Piano Works (Vladimir Ashkenazy), Decca 02894448302, 1995.

[33] Kraftwerk, Tour de France Etape 1, Minimum-Maximum, EMI Music 724356061620, 2005.

[34] Baloji, Entre les lignes, HUMO's Top 2007, EMI Music 5099950986125, 2007.

[35] A. Fire, Deep Blue, The Suburbs, Mercury 602527426297, 2010.

[36] T. Strokes, Last Nite, Is this it, RCA 07836804521, 2001.

[37] L. van Beethoven, Piano sonata op. 13, Beethoven Favourite Piano Sonatas, Deutsche Grammophon 028947797586, 2011.

[38] I. Albeniz, Iberia, book i, Lang Lang Live in Vienna, Sony Classical 886977190025, 2010.

[39] J. Kates and L. Kozma-Spytek, "Quality ratings for frequency-shqped peak-clipped speech," *J. Acoust. Soc. Amer.*, vol. 95, no. 6, pp. 3586–3594, 1994.

[40] J. Moir, "Just detectable distortion levels," *Wireless World*, vol. 87, no. 1541, 1981.

[41] C.-T. Tan, B. C. J. Moore, N. Zacharov, and V.-V. Mattila, "Predicting the perceived quality of nonlinearly distorted music and speech signals," *J. Audio Eng. Soc.*, vol. 52, no. 7–8, pp. 699–711, Jul. 2004.

[42] C. M. Jarque and A. K. Bera, "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Econon. Lett.*, vol. 6, no. 3, pp. 255–259, 1980.



Bruno Defraene (S'10) was born in Halle, Belgium, in 1986. He received the M.Sc. degree in electrical engineering from KU Leuven, Belgium, in 2009.

He is currently pursuing the Ph.D. degree at the Electrical Engineering Department of KU Leuven, and is a member of KU Leuven's Optimization in Engineering Center (OPTEC).

His research interests include audio signal processing, acoustical signal enhancement, audio quality assessment, and optimization for signal processing applications.



Toon van Waterschoot (S'04–M'12) was born in Lier, Belgium, on June 11, 1979. He received the Master's degree and the Ph.D. degree in Electrical Engineering, both from KU Leuven, Belgium, in 2001 and 2009, respectively.

Since 2011, he has been a Postdoctoral Research Fellow of the Research Foundation—Flanders (FWO) at KU Leuven, Belgium. In 2002, he spent a year as a Teaching Assistant with the Antwerp Maritime Academy, Belgium. From 2002 to 2003, and from 2008 to 2009, he was a Research Assistant with KU Leuven, Belgium, while from 2004 to 2007, he was a Research Assistant with the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT), Belgium. After his Ph.D. graduation, he was a Postdoctoral Research Fellow at KU Leuven, Belgium (2009–2010) and at Delft University of Technology, The Netherlands (2010–2011). Since 2005, he has been a Visiting Lecturer at the Advanced Learning and Research Institute of the University of Lugano, Switzerland, where he is teaching Digital Signal Processing.

He is currently serving as an Associate Editor for the EURASIP Journal on Audio, Music, and Speech Processing, and as a Nominated Officer for the European Association for Signal Processing (EURASIP). His research interests are in adaptive and distributed signal processing and parameter estimation, with application to acoustic signal enhancement, speech and audio processing, and wireless communications.



Hans Joachim Ferreau studied Mathematics and Computer Science at Heidelberg University, where he received a master degree in Mathematics in 2007. He recently received a Ph.D. degree in Engineering Science at the Electrical Engineering department of KU Leuven. His current research interests include the development of fast online QP solvers and efficient nonlinear model predictive control methods. He is also active in applying such methods to real-world control problems with practical relevance.



Moritz Diehl (M'11) studied Mathematics and Physics in Heidelberg and Cambridge University and did his Ph.D. in 2001 at Heidelberg's Interdisciplinary Center for Scientific Computing (IWR). Since 2006 he is professor at the Electrical Engineering Department of KU Leuven University and serves as the principal investigator of KU Leuven's Optimization in Engineering Center (OPTEC). His research focuses on optimization and control, spanning from numerical methods and algorithm development to applications. Main application areas are chemical and process control, mechatronics, and renewable energy systems. He currently supervises 8 Ph.D. students and 3 postdoctoral researchers and serves in the editorial boards of four international journals. Since 2011 he holds an ERC Starting Grant on the topic "Simulation, Optimization, and Control of High-Altitude Wind Power Generators."



Marc Moonen (M'94–SM'06–F'07) received the electrical engineering degree and the Ph.D. degree in applied sciences from KU Leuven, Belgium, in 1986 and 1990 respectively. Since 2004 he has been a Full Professor at the Electrical Engineering Department of KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing.

He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science. He received a journal best paper award from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with Geert Leus) and from Elsevier Signal Processing (with Simon Doclo).

He was chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), and a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and is currently President of EURASIP (European Association for Signal Processing).

He has served as Editor-in-Chief for the *EURASIP Journal on Applied Signal Processing* (2003–2005), and has been a member of the editorial board of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, *IEEE Signal Processing Magazine*, *Integration-the VLSI Journal*, *EURASIP Journal on Wireless Communications and Networking*, and *Signal Processing*. He is currently a member of the editorial board of *EURASIP Journal on Applied Signal Processing* and Area Editor for Feature Articles in *IEEE Signal Processing Magazine*.