PERCEPTION-BASED NONLINEAR LOUDSPEAKER COMPENSATION THROUGH EMBEDDED CONVEX OPTIMIZATION

Bruno Defraene, Toon van Waterschoot, Moritz Diehl and Marc Moonen

Dept. E.E./ESAT, SCD-SISTA, KU Leuven
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

ABSTRACT

In this paper a novel nonlinear loudspeaker compensation technique is presented which is based on embedded convex optimization. The aim is to compensate for the linear as well as for the nonlinear perceptible distortions incurred in the loudspeaker. To this end, a psychoacoustic model is adopted and a convex optimization based problem formulation is set up. In order to solve the resulting convex optimization problems in a fast and reliable way, a projected gradient optimization method is proposed. From comparative objective evaluation experiments, it is concluded that the proposed nonlinear loudspeaker compensation technique indeed improves the average audio quality scores.

***Index Terms*—** Sound perception, loudspeaker compensation, nonlinear model, Hammerstein model, convex optimization.

.

1. INTRODUCTION

Achieving a high perceived audio quality is undoubtedly a main concern in the development of any audio reproduction system. In general, the loudspeakers in such a system have a non-ideal response introducing both linear and nonlinear distortions in the reproduced audio signal. This may result in a significant degradation of the perceived audio quality [1].

Loudspeaker compensation techniques aim at reducing the effects caused by the non-ideal loudspeaker characteristics. The idea is to apply a digital compensation operation in cascade with the audio reproduction channel to counteract the response errors and nonlinearities introduced by the loudspeakers. Traditionally, loudspeakers have been modeled by *linear* systems such as FIR filters, IIR filters, warped filters or Kautz filters. The aim of linear loudspeaker compensation (also known as equalization) techniques is then to identify/approximate and apply the inverse digital filter to the audio signal prior to playback in order to reduce the linear distortions [2].

However, the small and low-cost loudspeakers that are ubiquitous in mobile devices are also causing a high level of *nonlinear* distortion, especially at high playback levels. This nonlinear behaviour can be taken into account by using nonlinear loudspeaker models such as Hammerstein models, Wiener-Hammerstein models and Volterra models. The aim of nonlinear loudspeaker compensation techniques is then to invert the nonlinear system under consideration [3]. This is in general a computationally expensive approach.

In this paper, a novel nonlinear loudspeaker compensation technique is presented, which aims to compensate for the linear as well as for the nonlinear distortions incurred in the loudspeaker. The novelty compared to existing loudspeaker compensation techniques is twofold. Firstly, a convex optimization procedure is embedded into the algorithm to carry out the nonlinear loudspeaker compensation. Approaches based on *embedded convex optimization* have been successful in related signal processing applications, e.g. [4] [5]. Secondly, the proposed compensation technique is *perception-based*: a psychoacoustic model which captures knowledge about the human perception of sounds is employed, and allows to minimize the resulting perceptible distortion.

This paper is organized as follows. In Section 2, the nonlinear loudspeaker compensation technique is introduced in a framework of embedded convex optimization and the inclusion of a psychoacoustic model is discussed. In Section 3, a projected gradient optimization method is presented for solving the convex optimization problems at hand in a fast and reliable way. In Section 4, simulation results are given for a comparative audio evaluation experiment. Finally, in Section 5, some concluding remarks are presented.

2. NONLINEAR LOUDSPEAKER COMPENSATION

The aim of nonlinear loudspeaker compensation is to precompensate for the linear as well as for the nonlinear distortions incurred in the loudspeaker. Figure 1 shows the operation of

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) and PFV/10/002 (OPTEC), Concerted Research Action GOA-MaNet, the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 'Dynamical systems, control and optimization' (DYSCO) 2007-2011, Research Project IBBT, the EU via FP7-EMBOCON (ICT-248940), FP7-HD-MPC (INFSO-ICT-223854), and was supported by a Postdoctoral Fellowship of the Research Foundation Flanders (FWO-Vlaanderen, T. van Waterschoot). The scientific responsibility is assumed by its authors.

the proposed compensation technique. Frame-by-frame processing of the digital input audio signal $x[n]$ is applied, using non-overlapping input frames $x_m \in \mathbb{R}^N, m = 0, 1...M$. The loudspeaker is modeled by a Hammerstein model, i.e. a memoryless nonlinearity with a linear region $[-U, U]$, followed by a linear finite impulse response (FIR) filter with impulse response $h[n], n = 0...L-1$. Before it is fed into the loudspeaker, the input frame x_m passes through the nonlinear loudspeaker compensation block. For a given input frame x_m, the nonlinear loudspeaker compensation consists of the following steps:

1. Calculate the instantaneous global masking threshold $t_m \in \mathbb{R}^{\frac{N}{2}+1}$ of the input frame x_m using a psychoacoustic model (see Subsection 2.2).

2. Calculate a compensated input frame $v_m^* \in \mathbb{R}^N$ as the solution of a convex optimization problem, such that the resulting output frame y_m^* is perceptually as close as possible to x_m (see Subsection 2.1).

2.1. Optimization problem formulation

The core of the proposed nonlinear loudspeaker compensation technique consists in calculating the solution of a convex optimization problem for each input frame. The optimal compensated input frame v_m^* is calculated from the knowledge of the input frame x_m and its instantaneous masking threshold t_m. A necessary constraint on v_m^* is that the amplitude of the samples be contained within the linear region $[-U, U]$ of the memoryless nonlinearity. The objective function reflects the amount of perceptible distortion added between y_m and x_m. The optimization problem is then formulated as an inequality constrained frequency domain weighted L2-distance minimization, i.e.[1]

$$\begin{aligned} v_m^* &= \underset{v_m \in \mathbb{R}^N}{\arg\min}\ f(v_m) \quad \text{s.t.} \quad -u \leq v_m \leq u \\ &= \underset{v_m \in \mathbb{R}^N}{\arg\min}\ \frac{1}{2N} \sum_{i=0}^{N-1} w_m(i) |Y_m(e^{j\omega_i}) - X_m(e^{j\omega_i})|^2 \\ &\qquad\qquad \text{s.t.} \quad -u \leq v_m \leq u \end{aligned} \tag{1}$$

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 vector $u = U.1_N$ contains the upper amplitude level of the linear region (with $1_N \in \mathbb{R}^N$ a vector of ones), and $w_m(i)$ are the weights of a perceptual weighting function to be defined in subsection 2.2.

[1] Superscripts T and H denote the transpose and the Hermitian transpose, respectively.

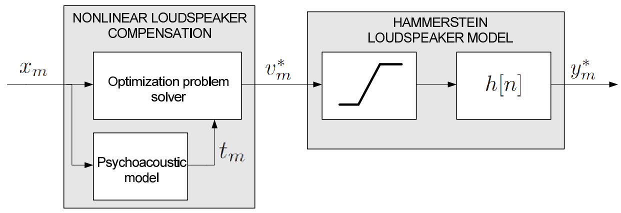

Fig. 1. Perception-based nonlinear loudspeaker compensation: schematic overview

The optimization problem (1) can be rewritten as follows,

$$\begin{aligned} v_m^* &= \underset{v_m \in \mathbb{R}^N}{\arg\min}\ \frac{1}{2} (y_m - x_m)^H \underbrace{D^H W_m D}_{\triangleq Q_m} (y_m - x_m) \\ &\qquad \text{s.t.} \quad -u \leq v_m \leq u \\ &= \underset{v_m \in \mathbb{R}^N}{\arg\min}\ \frac{1}{2} (Hv_m + \tilde{H}v_{m-1}^* - x_m)^T Q_m (Hv_m + \tilde{H}v_{m-1}^* - x_m) \\ &\qquad \text{s.t.} \quad -u \leq v_m \leq u \\ &= \underset{v_m \in \mathbb{R}^N}{\arg\min}\ \frac{1}{2} v_m^T \underbrace{H^T Q_m H}_{\text{Hessian } A_m} v_m + \underbrace{(H^T Q_m^T (\tilde{H}v_{m-1}^* - x_m))^T}_{\text{Gradient } b_m} v_m \\ &\qquad \text{s.t.} \quad -u \leq v_m \leq u \end{aligned} \tag{2}$$

where $D \in \mathbb{C}^{N \times N}$ is the unitary Discrete Fourier Transform (DFT) matrix, $W_m \in \mathbb{R}^{N \times N}$ is a diagonal weighting matrix with positive diagonal elements $w_m(i)$, obeing the symmetry property $w_m(i) = w_m(N-i)$ for $i = 1, 2, ..., \frac{N}{2} - 1$, and the matrices $H \in \mathbb{R}^{N \times N}$ and $\tilde{H} \in \mathbb{R}^{N \times N}$ are defined as

$$H = \begin{bmatrix} h[0] & 0 & \dots & \dots & \dots & \dots & 0 \\ h[1] & h[0] & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ h[L-1] & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & h[L-1] & \dots & h[1] & h[0] \end{bmatrix} \tag{3}$$

$$\tilde{H} = \begin{bmatrix} 0 & \dots & 0 & h[L-1] & \dots & h[2] & h[1] \\ 0 & \dots & 0 & 0 & h[L-1] & \dots & h[2] \\ \vdots & & & & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & h[L-1] \\ \vdots & & & & & & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix} \tag{4}$$

It is remarked 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). As it can be shown that the Hessian matrix $A_m = H^T Q_m H$ in (2) is guaranteed to be real and positive definite, the optimization problem is a strictly convex quadratic program.

2.2. 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.

A first phenomenon is the *absolute threshold of hearing*, which 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. A second phenomenon is *simultaneous masking*, 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, consider the input frame x_m to act as the masker, and $y_m - x_m$ as the maskee. 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. 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 by using part of the ISO/IEC 11172-3 MPEG-1 Layer 1 psychoacoustic model 1 [6]. The major steps in its computation are the following:

1. *Spectral analysis and SPL normalization*: A high-resolution spectral estimate of the input frame is calculated. After a normalization operation and a Hann windowing operation on the input signal frame, the PSD estimate is obtained through a 512-point Fast Fourier Transform (FFT).

2. *Identification of tonal and non-tonal maskers*: The output of the FFT is used to determine relevant tonal and non-tonal maskers in the spectrum of the audio signal.

3. *Calculation of individual masking thresholds*: an individual masking threshold is calculated for each masker in a 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.

4. *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.

3. FAST PROJECTED GRADIENT OPTIMIZATION

The core of the nonlinear loudspeaker compensation technique described in Section 2 consists in the solution of successive instances of the convex quadratic optimization problem (2). The focus of this section is on the class of *projected gradient methods* for solving the convex QPs. In Subsection 3.1, the standard projected gradient method is introduced. In Subsection 3.2, an improved projected gradient method is presented, which reaches an optimal convergence.

3.1. Standard projected gradient method

In each iteration of the standard projected gradient method, 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. 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.

Introducing the notation v_m^k for the kth iterate of the mth frame, the main steps in the $(k+1)$th iteration of the projected gradient method can be written as follows:

- Take a step of stepsize s_m^k along the negative gradient direction :
$$\tilde{v}_m^{k+1} = v_m^k - s_m^k \nabla f(v_m^k) \quad (6)$$
where the gradient is computed as
$$\nabla f(v_m^k) = H^T Q_m^T (H v_m + \tilde{H} v_{m-1}^* - x_m) \quad (7)$$

- Project $\tilde{v}_m^{k+1}$ orthogonally onto the convex feasible set Ω of (2), which is defined as
$$\Omega = \{v_m \in \mathbb{R}^N | -u \leq v_m \leq u\} \quad (8)$$

An orthogonal projection $\Pi_\Omega(\tilde{v}_m^{k+1})$ onto Ω can be shown to come down to performing a simple componentwise hard clipping operation (with lower bound $-U$ and upper bound U), i.e.

$$v_m^{k+1} = \Pi_\Omega(\tilde{v}_m^{k+1}) = \underset{v_p \in \Omega}{\arg\min} \frac{1}{2} \|v_p - \tilde{v}_m^{k+1}\|_2^2 \quad (9)$$

Algorithm 1 Optimal projected gradient method

Input $x_m \in \mathbb{R}^N$, $v^*_{m-1} \in \mathbb{R}^N$, $v^0_m = c^0_m \in \Omega$, $\gamma^0_m \in (0, 1)$, U, h, W_m
Output $v^*_m \in \mathbb{R}^N$

Calculate Lipschitz constant C_m [using (13)]
Calculate convexity parameter μ_m [using (14)]
$\kappa_m = \frac{C_m}{\mu_m}$
$k = 0$
while convergence is not reached **do**
  $\tilde{v}^{k+1}_m = c^k_m - \frac{1}{C_m}\nabla f(c^k_m)$ [using (7)]
  $v^{k+1}_m = \Pi_\Omega(\tilde{v}^{k+1}_m)$ [using (10)]
  Calculate γ^{k+1}_m from $(\gamma^{k+1}_m)^2 = (1 - \gamma^{k+1}_m)(\gamma^k_m)^2 + \kappa_m \gamma^{k+1}_m$
  $\delta^k_m = \frac{\gamma^k_m(1-\gamma^k_m)}{(\gamma^k_m)^2+\gamma^{k+1}_m}$
  $c^{k+1}_m = v^{k+1}_m + \delta^k_m(v^{k+1}_m - v^k_m)$
  $k = k + 1$
end while
$v^*_m = v^k_m$

where

$$v^{k+1}_m(i) = \begin{cases} -U & \text{if} \quad \tilde{v}^{k+1}_m(i) < -U \\ \tilde{v}^{k+1}_m(i) & \text{if} \quad -U \leq \tilde{v}^{k+1}_m(i) \leq U \\ U & \text{if} \quad \tilde{v}^{k+1}_m(i) > U. \end{cases} \quad , i = 0...N-1 \tag{10}$$

3.2. Optimal projected gradient method

In this subsection, a projected gradient optimization method is presented that reaches an optimal convergence for the class of convex optimization problems with *strongly convex* objective functions. This method was first proposed in [7] and variants of the method have been applied in diverse applications, e.g. real-time clipping of audio signals [8].

Algorithm 1 summarizes the optimal projected gradient optimization method. In each iteration, a standard projected gradient step is performed on a potentially infeasible weighted sum of two previous feasible iterates. Knowledge of the Lipschitz constant C_m of the gradient ∇f and the convexity parameter μ_m of f on the set Ω is assumed. In order to establish C_m and μ_m for optimization problem (2), the next two lemmas are proposed.

Lemma 3.1. *(cfr. [7]) Let function f be twice continuously differentiable on set Ω. The gradient ∇f is Lipschitz continuous on set Ω with Lipschitz constant C if and only if*

$$||\nabla^2 f(z)|| \leq C \ , \ \forall z \in \Omega \tag{11}$$

Lemma 3.2. *(cfr. [7]) Let function f be twice continuously differentiable on set Ω. The function f is strongly convex on set Ω with convexity parameter μ if and only if there exists $\mu > 0$ such that*

$$\nabla^2 f(z) \geq \mu I \ , \ \forall z \in \Omega \tag{12}$$

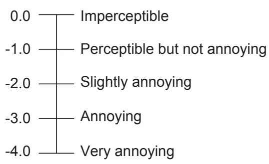

Fig. 2. The ITU-R five-grade impairment scale

Using Lemma 3.1, it is proved that the Lipschitz constant C_m can be computed as

$$\begin{aligned} C_m = ||A_m|| &= \max_{1\leq i\leq N} \lambda_i(A_m) \\ &= \max_{1\leq i\leq N} \lambda_i(H^H D^H W_m DH) \end{aligned} \tag{13}$$

where $\lambda_i(A_m),\ i = 1...N$, denote the eigenvalues of A_m.

Using Lemma 3.2, it is proved that the convexity parameter μ_m can be computed as

$$\begin{aligned} \mu_m &= \min_{1\leq i\leq N} \lambda_i(A_m) \\ &= \min_{1\leq i\leq N} \lambda_i(H^H D^H W_m DH). \end{aligned} \tag{14}$$

4. SIMULATION RESULTS

For audio quality evaluation purposes, a test database consisting of 8 audio excerpts was compiled (see Table 1 for details). A first set of excerpts (numbers 1-5) was extracted from different commercial audio CDs. A second set of excerpts (numbers 6-8) was extracted from an ITU CD-ROM associated to Recommendation BS.1387-1 [14].

The loudspeaker was modeled with a Hammerstein model, using a hard clipping nonlinearity with linear region $[-U, U]$ and a 5th-order lowpass filter with impulse response $h = \begin{bmatrix}1 & 0.7 & 0.5 & 0.3 & 0.1\end{bmatrix}^T$. Each audio signal in the test database was passed through the Hammerstein loudspeaker model, with and without performing a preceding perception-based nonlinear loudspeaker compensation step. The following parameters were used for nonlinear loudspeaker compensation: $N = 512$, $\alpha = 0.04$, and using κ_m iterations of Algorithm 1 for all instances of optimization problem (2), such that the solution accuracy $\epsilon = f(v^{\kappa_m}_m) - f(v^*_m) = 10^{-12}$. Simulations were performed for five different degrees of nonlinear loudspeaker distortion P_{LD}. The parameter U was set such that the amplitude of the input audio signal $x[n]$ exceeded the linear region $[-U, U]$ for P_{LD}={90, 95, 97, 98, 99}% of the samples.

For each of a resulting total of 8×5×2=80 processed audio signals, the PEAQ (*Perceptual Evaluation of Audio Quality*) [14] measure was calculated. The PEAQ ODG (Objective

Table 1. Audio excerpts database used for comparative audio quality evaluation (16 bit mono at 44.1 kHz)

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	[9]
2	rhcp.wav	polyphonic	instrumental	rock	9.8	468996	900000	[10]
3	pierle.wav	polyphonic	instrumental+vocal	pop	11.7	2234000	2750000	[11]
4	chopin.wav	monophonic	instrumental	classical	17.8	50000	836200	[12]
5	kraftwerk.wav	polyphonic	instrumental	electronic	17.2	7480000	8240000	[13]
6	breftri.wav	monophonic	instrumental	classical	19.7	1	869675	[14]
7	crefsax.wav	monophonic	instrumental	classical	10.9	1	479026	[14]
8	grefcla.wav	monophonic	instrumental	classical	6.9	1	302534	[14]

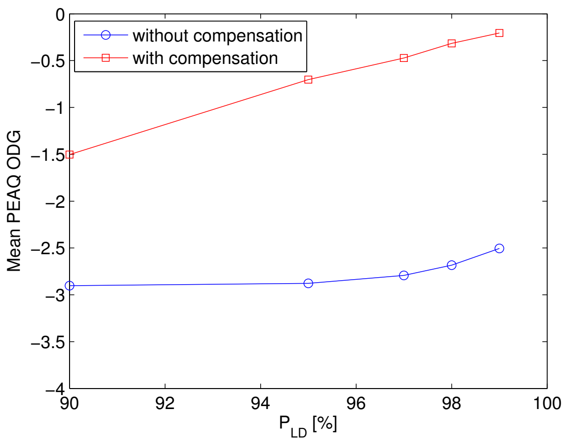

Fig. 3. Mean PEAQ ODG for audio signals processed with and without perception-based nonlinear loudspeaker compensation, as a function of the degree of nonlinear loudspeaker distortion P_{LD}.

Difference Grade) has a range between 0 and - 4, corresponding to the ITU-R impairment scale depicted in Figure 2.

In Figure 3, the average PEAQ ODG score over all 8 audio signals is plotted as a function of the degree of nonlinear loudspeaker distortion P_{LD}. A monotonically increasing average audio quality score is observed for increasing P_{LD}. Clearly, including the perception-based nonlinear loudspeaker compensation is seen to improve the average objective audio quality score significantly, and this for all considered degrees of nonlinear loudspeaker distortion[2].

5. CONCLUSION

In this paper a novel nonlinear loudspeaker compensation technique was presented. By including a psychoacoustic model and embedding convex optimization into the algorithm, it was possible to minimize the perceptible distortion. A fast projected gradient optimization method was proposed for solving the resulting convex optimization problems. Comparative objective evaluation experiments have shown that the proposed nonlinear loudspeaker compensation improves the average objective audio quality score, and this for different degrees of nonlinear loudspeaker distortion.

[2]Processed audio signals from this simulation are available online on http://homes.esat.kuleuven.be/~bdefraen/.

6. REFERENCES

[1] K. Lashkari, "A modified Volterra-Wiener-Hammerstein model for loudspeaker precompensation," in *Proc. 39th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 2005, pp. 344–348.

[2] M. Karjalainen and T. Paatero, "Equalization of loudspeaker and room responses using Kautz filters: Direct least squares design," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–13, 2007.

[3] M. Viberg K. Shi, G. T. Zhou, "Compensation for nonlinearity in a Hammerstein system using the coherence function with application to nonlinear acoustic echo cancellation," *IEEE Transactions on Signal processing*, vol. 55, no. 12, pp. 5853–5858, 2007.

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

[5] B. Defraene, T. van Waterschoot, H. J. Ferreau, M. Diehl, and M. Moonen, "Real-time perception-based clipping of audio signals using convex optimization," *submitted.*

[6] 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.

[7] Y. Nesterov, *Introductory lectures on convex optimization*, Springer, 2004.

[8] 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 2011*, Prague, Czech Republic, May 2011, pp. 333–336.

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

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

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

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

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

[14] International Telecommunications Union Recommendation BS.1387, "Method for objective measurements of perceived audio quality," 1998.