Adaptive Automatic Compensation of Transducer Nonlinearities Using Extremum-Seeking Control

Andreas Männchen[1], Georg Fischer[1], Tobias Fritsch[1], Matthias Fiedler[1], Daniel Beer[1]
[1] *Fraunhofer Institute for Digital Media Technology IDMT, 98693 Ilmenau, Germany,*
email: andreas.maennchen@idmt.fraunhofer.de

Introduction

Thanks to linear control theory, the dynamic behavior of linear time-invariant (LTI) systems can be controlled in well-established boundaries. Unfortunately, most real systems, such as transducers, are at least somewhat time-variant and nonlinear. Engineers often use LTI approximations of such systems in order to deal with them efficiently. In many applications, however—including high-quality audio reproduction—signal distortion resulting from time-variant nonlinear input-output transducer characteristics poses a significant problem and cannot be ignored. For this reason, this article deals with the adaptive compensation of transducer nonlinearities.

The paper is structured as follows: After some fundamentals of nonlinear systems, electrostatic loudspeakers (ESLs), and extremum-seeking control (ESC) in the next section, the subsequent parts first introduce a method for the adaptive compensation of transducer nonlinearities using ESC and then evaluate it using measurement and simulation data. The article ends with some concluding thoughts and a short summary.

Fundamentals

Before going into detail regarding the adaptive compensation of transducer nonlinearities, this section lays out some necessary fundamentals.

Static Nonlinearities

With respect to nonlinear input-output characteristics, the focus of this article lies on static nonlinearities, which are memoryless and, therefore, do not exhibit dynamic behavior. The output $f(x)$ of a static nonlinear system depends only on the current input x. One of many possible descriptions of static nonlinear systems can be given with a polynomial of degree $N > 1$ and coefficients k_i, as displayed in eq. (1):

$$f(x) = \sum_{i=0}^{N} k_i x^i \tag{1}$$

The nonlinear distortion effected by such nonlinear characteristics manifests as amplitude distortion in the time domain and as additional frequency components in the frequency domain (fig. 1). When using a signal that is sparse in the frequency domain the additional frequency components can easily be identified and quantified.

Electrostatic Loudspeakers

As stated earlier, nonlinear distortion is often undesired, especially in high-quality audio applications. There, transducer nonlinearities can impair audio reproduction to an unacceptable degree, for example, when loudspeakers are driven in the large signal domain [1]. Due to their ubiquity and complex nonlinear behavior much research has gone into the transducer nonlinearities of electrodynamic loudspeakers [2, 3]. This article, however, takes a look at ESLs, since they are well-suited to demonstrating the compensation of static transducer nonlinearities.

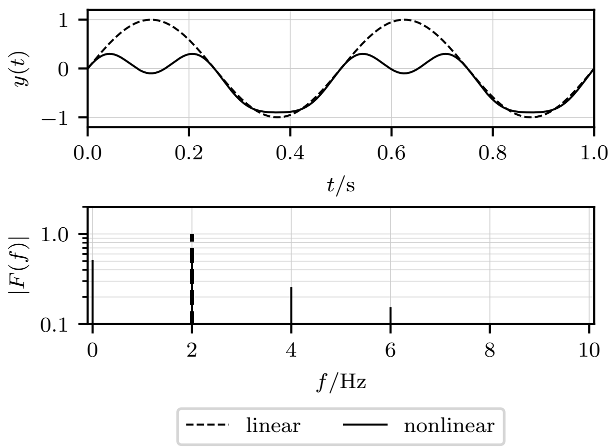


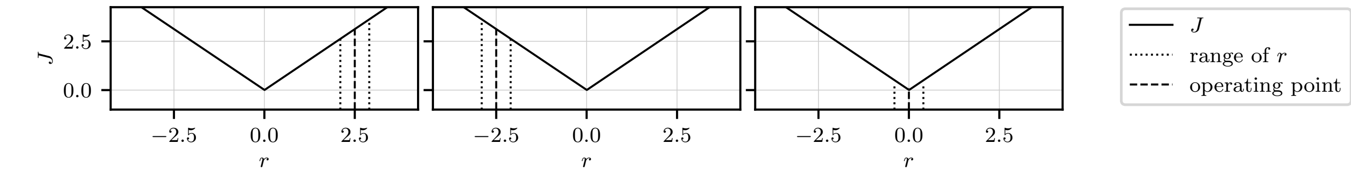
Figure 1: Effects of nonlinear distortion: nonlinear amplitude distortion in the time domain (top) and additional frequency components in the frequency domain (bottom).

ESLs are constructed similarly to capacitors. In their most basic form of assembly, called single-sided, they consist of two electrodes: an immobile, often perforated stator electrode and a moving, non-perforated membrane electrode. With some simplifying assumptions, an electrical input voltage u between the electrodes can be said to elicit a mechanical force F according to eq. (2):

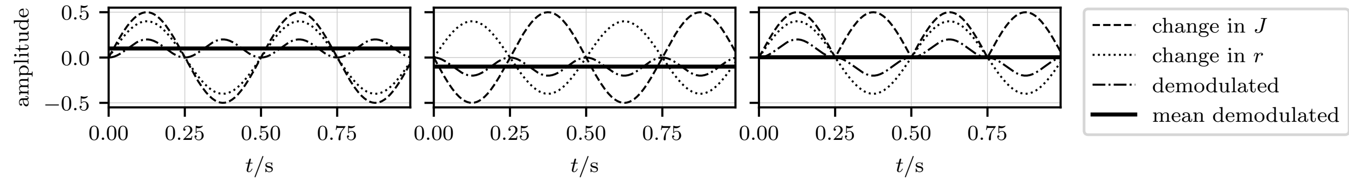
$$F = \frac{Qu}{2d} = \frac{Cu^2}{2d}, \qquad [\mathrm{N}] \tag{2}$$

where Q denotes the amount of electric charge in the electrodes, d their distance, and C is the capacitance of the system [4, p. 109]. This force results in a membrane displacement and, ultimately, a sound pressure p in the loudspeaker scenario, where $u = u_{\mathrm{signal}} + u_{\mathrm{bias}}$ consists of an AC part u_{signal} and some DC bias voltage u_{bias}. The force-voltage characteristic described by eq. (2) represents a static quadratic nonlinearity that may change over time if any of its parameters vary, e.g., due to environmental influences.

With ESLs, there are two main approaches to dealing with nonlinear distortion. The first one is to apply a large



(a) Range of the sinusoidal modulaton of parameter r around operating points lying above (left), below (center), or at the optimum (right) of the objective function J.



(b) Relevant ESC signals over time, with the mean demodulated signal approximating the local gradient of objective function J in operating points lying above (left), below (center), or at the optimum (right).

Figure 2: Illustration of the working principle of ESC.

u_{bias} to improve the operating point on the force-voltage characteristic [5, pp. 176–187]. There, small changes in u_{signal} lead to relatively large changes in F, allowing the system to be driven in the small signal domain, minimizing nonlinear distortion. This approach carries some disadvantages, e.g., the costly electrical components needed to provide a large u_{bias}. The second, more widely used approach, is the so-called push-pull configuration that practically eliminates nonlinear distortion [4, pp. 110–112]. However, it necessitates the assembly of a second stator electrode.

Extremum-Seeking Control

The approach for an automatic compensation of transducer nonlinearities proposed in this article uses ESC to adapt to varying operating points. Historically, ESC has most often been used in industrial applications [6], but occasionally, it has found its use in acoustics as well [7].

ESC is a form of online optimization that can find an extremum of an objective function by estimating and following its local gradient [8]. Classic ESC approaches find the optimum by estimating the gradient of the objective function J through parameter modulation, called the perturbation method [6]. The principle is illustrated in fig. 2, where the local gradient of the objective function J is estimated by sinusoidally modulating its parameter r with a certain frequency f_{mod}. If J is chosen correctly, at least some change in J directly results from the variation of r. If those changes are in phase to one another, as shown in the bottom subfigures, the gradient can be approximated by the mean of the demodulated signal, which is the product of the change in J and the change in r. Following this gradient, ESC converges at the optimum, where the gradient becomes zero. The convergence behavior of ESC has been examined for various implementations and plant dynamics [9, 10].

An advantage of ESC lies in it being model-free, enabling it to work without prior knowledge of the plant dynamics or the characteristics of the objective function [11, p. 6]. Simple variants of ESC can, however, only find local extrema, not necessarily global ones [8], which needs to be considered when designing an objective function.

Proposed Method

Based on the fundamentals presented in the previous section, this article proposes a digital signal processing (DSP) method for the adaptive compensation of ESL nonlinearities. Its goal is to enable lower DC bias voltages in conjunction with large signal domain driving of the transducer at low levels of nonlinear distortion. It is based on a parametric predistortion function that is adapted to varying operating points by real-time optimization of its parameters through ESC.

Compensation of Static Nonlinearities

DSP techniques can be used to compensate for transducer nonlinearities. While the major goal of linearizing the input-output characteristic is the same as with the hardware approaches mentioned above, the DSP method calls for some further conditions that make sense when dealing with single-sided ESLs: The operating point of the system set by u_{bias} and the maximum operating range of the input signal set by u_{signal} should stay the same when applying the compensation. At the same time, the AC part of the mechanical output force F or the output sound pressure p should stay as large as possible.

The predistortion function described by eqs. (3) and (4) fulfills these conditions, when it is combined with a DC offset removal and a scaling function.

$$\tilde{x} = \operatorname{sgn}(r)\sqrt{|r|} + \frac{x}{\sqrt{|r|}} \tag{3}$$

$$\overline{\overline{x}} = \operatorname{sgn}(\tilde{x})\,\sqrt[n]{|\tilde{x}|} \tag{4}$$

It applies an approximately inverse characteristic, parameterized by $r \neq 0$, on an input signal $-1 \leq x \leq 1$, resulting in intermediate signals $\tilde{x}$ and $\overline{\overline{x}}$. This predistortion is based on the assumption that the static nonlinearity can be described by a simple polynomial of degree

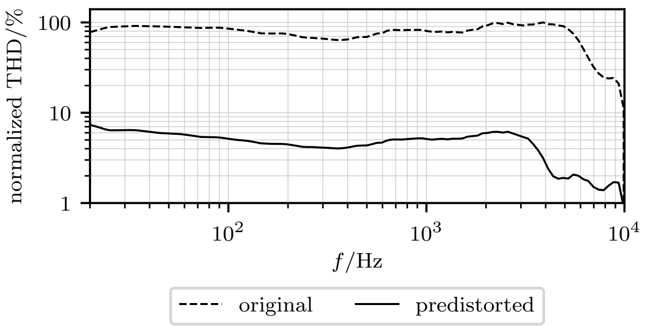


Figure 3: Measurement results for the predistortion applied to a MEMS ESL: normalized THD without (original) and with the predistortion (predistorted) over frequency.

$n > 1$, as formulated in eq. (5):

$$y = (a + bx)^n, \tag{5}$$

with output y, input x, as well as parameters a and b that determine the operating point and range on the nonlinear characteristic. The more closely the nonlinear characteristic fits this description, the better the predistortion given in eqs. (3) and (4) can linearize it with an optimally chosen value of r.

Adaptivity via Extremum Seeking

In order to apply the predistortion without first manually finding an optimal value for r, it needs to work adaptively. This also enables the compensation of time-variant static nonlinearities. The ESC paradigm lends itself to this introduction of adaptability into parameterized control systems [12].

In the application considered here, the objective function J must be a measure of nonlinear distortion to be minimized with respect to r. This nonlinearity measure ideally works for a wide range of audio signals, from single sinusoids to broadband material like music. Therefore, the total noncoherent distortion (TNCD), as popularized in [13], is chosen as the objective function. It is calculated based on a known input stimulus and the nonlinear system response.

Furthermore, the parameter modulation frequency f_{mod} must be chosen wisely. Ideally, f_{mod} does not approach the frequencies of any natural changes in J or variance of J due to any external disturbance—f_{mod} is crucial to the robustness of the ESC. In addition, it affects the convergence speed of the extremum seeking. Generally, a high f_{mod} is desirable [14, p. 6]. Since this is an audio application and the parameter modulation period must be considerably longer than the TNCD window length of a few milliseconds, a value of $f_{\mathrm{mod}} = 2\,\mathrm{Hz}$ is chosen.

Evaluation

The method for an adaptive compensation of static transducer nonlinearities proposed above is evaluated with respect to its performance regarding the minimization of nonlinear distortion, its robustness for different operating points and input signals, as well as its stability.

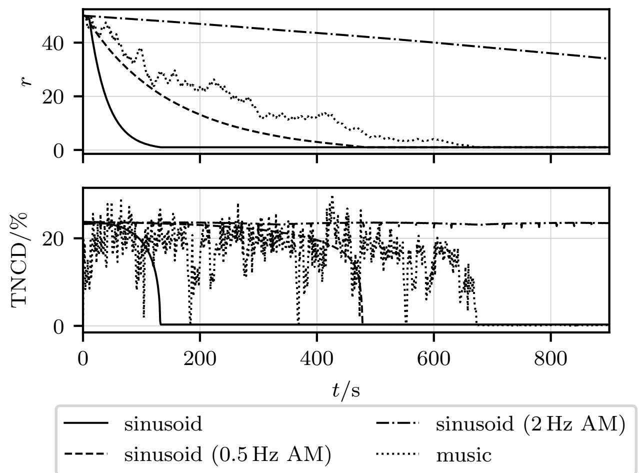


Figure 4: Simulation results for the ESC-based method, using $a = 100$ and $b = 100$ with four different stimuli: predistortion parameter r (top) and TNCD (bottom) over time.

Methods

First, the predistortion algorithm itself is evaluated based on measurement data. The measurements were performed on a single-sided MEMS ESL, where a PiezoDrive PX200 amplifier provided the voltage $u = u_{\mathrm{bias}} + u_{\mathrm{signal}}$ and the output sound pressure p was measured using a GRAS RA0401 ear simulator. Nonlinear distortion is evaluated as total harmonic distortion (THD) in p according to IEEE 519 (only up to the eleventh harmonic or 20 kHz, whichever is lower). The THD values are all normalized so that the highest value overall becomes 100 %.

The second part of the evaluation focuses on the robustness and stability of the adaptive behavior and is based on simulation data. The nonlinear system model is described by eq. (5) with $n = 2$. Its behavior was simulated with varying values of the parameters a and b, resulting in different operating points and ranges. Robustness of the ESC convergence was tested with four different stimuli, comprising a single sinusoid at 1 kHz, two amplitude-modulated (AM) sinusoids at 1 kHz with modulation frequencies 0.5 Hz and 2 Hz, respectively, and a musical piece. Stability analysis was performed by using three different control gains K for the ESC. The initial value of r was always 50, corresponding to about 0.25 % THD when applying the predistortion to a memoryless LTI system.

Results

Fig. 3 shows that the predistortion works well on the MEMS ESL. THD is reduced by more than 90 % compared to the orginal value at all frequencies, using a value of r found via a grid search based on prior measurements.

Continuing with the ESC-based adaptive method, the simulation results in fig. 4 show that the approach can robustly handle various kinds of stimuli at a quite strongly nonlinear operating point ($a = 100$, $b = 100$). Convergence is achieved within minutes for the pure sinusoid, the 0.5 Hz AM sinusoid, and the music signal, while the 2 Hz AM sinusoid leads to a much longer convergence

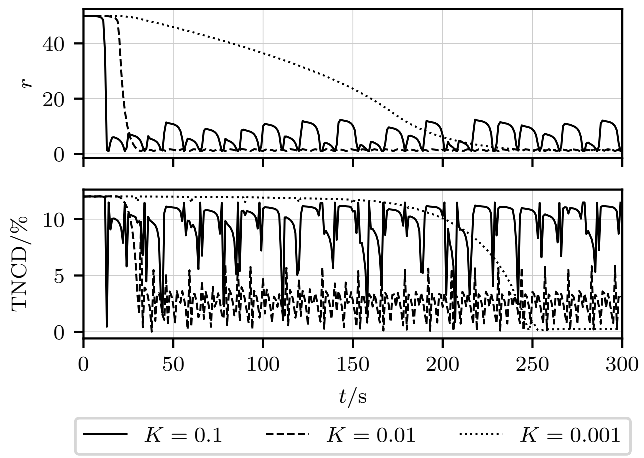


Figure 5: Simulation results for the ESC-based method, using $a = 10$ and $b = 5$ with a sinusoidal stimulus and three different values of the control gain K: predistortion parameter r (top) and TNCD (bottom) over time.

time. This is explained by the coincidence of the AM frequency with f_{mod}, making the 2 Hz AM signal critical. While all the sinusoid stimuli lead to strictly monotonic convergence, the changing signal statistics of the musical piece result in a trajectory that sometimes changes direction. Nevertheless, the ESC finds the optimum value of r and TNCD is strongly reduced.

Finally, fig. 5 reveals the results of a rudimentary stability analysis at a moderately nonlinear operating point ($a = 10$, $b = 5$). As is expected, higher values of K result in a fast approach to the optimum within seconds, but also lead to self-oscillations of the ESC that never settles at the optimum. At $K = 0.001$ convergence takes about four minutes, but the optimum value of r is found and held in a perfectly stable manner. This trade-off between convergence speed and stability must be considered when setting the ESC parameters.

Conclusion

Based on the results, ESC appears to be a quick and easy-to-implement method to introduce adaptability into a control system. It can be applied to the automatic compensation of static transducer nonlinearities, even for broadband stimuli such as music, with an adequate objective function J. However, its parameters, such as f_{mod} and K, have to be chosen carefully to guarantee robust and stable behavior. It will be interesting to extend the proposed approach for an adaptive compensation of nonlinearities to other types of transducers, e.g., dielectric polymer actuators, and application domains.

Summary

This article proposed a signal processing method for the adaptive automatic compensation of transducer nonlinearities that enables large signal domain driving of ESLs at low levels of nonlinear distortion. A predistortion of the input signal by parameterized inversions of polynomial characteristics was made adaptive by real-time optimization of its parameter through ESC with a cost function that estimates the amount of nonlinear distortion. This approach was evaluated with respect to its performance, robustness, and stability for different operating points and stimuli, showing promising results.

References

[1] C.-T. Tan, B. C. J. Moore, and N. Zacharov, "The Effect of Nonlinear Distortion on the Perceived Quality of Music and Speech Signals," *Journal of the Audio Engineering Society*, vol. 51, no. 11, pp. 1012–1031, 2003.

[2] A. J. M. Kaizer, "Modeling of the Nonlinear Response of an Electrodynamic Loudspeaker by a Volterra Series Expansion," *Journal of the Audio Engineering Society*, vol. 35, no. 6, pp. 421–433, 1987.

[3] W. Klippel, "Loudspeaker Nonlinearities – Causes, Parameters, Symptoms," in *Audio Engineering Society Convention 119*, 2005.

[4] J. Borwick, ed., *Loudspeaker and Headphone Handbook.* Focal Press, 3rd ed., 2012.

[5] F. Hunt, *ELECTROACOUSTICS: The Analysis of Transduction, and Its Historical Background.* Acoustical Society of America, 2. print ed., 1982.

[6] J. Sternby, *A Review of Extremum Control.* Technical Reports TFRT-7161, Department of Automatic Control, Lund Institute of Technology (LTH), 1979.

[7] R. Maury, M. Kœnig, L. Cattafesta, P. Jordan, and J. Delville, "Extremum-Seeking Control of Jet Noise," *International Journal of Aeroacoustics*, vol. 11, no. 3-4, pp. 459–473, 2012.

[8] D. Nešić, T. Nguyen, Y. Tan, and C. Manzie, "A non-gradient approach to global extremum seeking: An adaptation of the Shubert algorithm," *Automatica*, vol. 49, no. 3, pp. 809–815, 2013.

[9] M. Krstić and H.-H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems," *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.

[10] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.

[11] C. Zhang and R. Ordóñez, *Extremum-Seeking Control and Applications: A Numerical Optimization-Based Approach.* Springer-Verlag London Ltd, 2012.

[12] N. Killingsworth and M. Krstić, "Auto-tuning of PID controllers via extremum seeking," in *Proceedings of the 2005 American Control Conference*, vol. 4, pp. 2251–2256, 2005.

[13] S. Temme and P. Brunet, "A New Method for Measuring Distortion Using a Multitone Stimulus and Noncoherence," *Journal of the Audio Engineering Society*, vol. 56, no. 3, pp. 176–188, 2008.

[14] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum-Seeking Control.* Wiley Interscience, 2003.