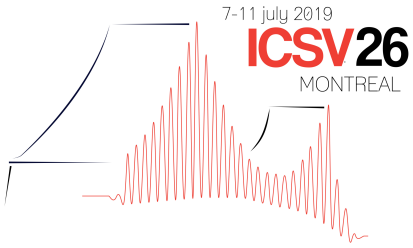
26th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
7-11 July 2019, Montreal



PSYCHOACOUSTIC OPTIMIZATION OF A FEEDBACK CONTROLLER FOR ACTIVE NOISE CANCELLING HEADPHONES

Piero Rivera Benois and Udo Zölzer
Helmut Schmidt University, Department of Signal Processing and Communications, Hamburg, Germany
email: piero.rivera.benois@hsu-hh.de

In this paper different cost functions are studied for the optimization of FIR filters as minimum variance controllers in active noise cancelling headphones. The resulting controllers are implemented in a headphones prototype and their attenuation performances are measured using a dummy-head. The measurements are then compared and evaluated in terms of psychoacoustic objective metrics, in order to establish which one of the cost functions relates more closely to the human perception of noise.

Keywords: ANC, headphones, feedback control, psychoacoustics.

1. Introduction

Active Noise Cancelling (ANC), also known as Active Noise Control, aims to reduce the sound pollution present in the environment by actively generating sound pressure waves that overlap destructively on a sweet-spot with the ones of the noise sources. Its origins can be dated back to Lueg's patent [1] on the control of acoustical tonal noise inside of a duct, while its particular application to headphones can be dated to the later work of Simshauser *et al.* in [2]. In order to achieve a perfect attenuation, the phase and the magnitude of the noise around the sweet-spot have to be matched. Different approaches can be used to achieve this, depending on the nature of the noise and the solution's context.

In the case of ANC headphones, the elements presented in Fig. 1a can be found. An external microphone measures a time-advanced reference of the disturbance $x(n)$. This signal is measured again as $d(n)$, after it has entered the ear-cup and reached the so-called error microphone. Thus, $P(z)$, known as the primary path, is defined as the changes in magnitude and phase that the disturbance suffers by means of the headphone's construction materials and its angle of incidence. The control signal $y'(n)$ generated by the ANC system is phase-inverted, fed to the speaker and measured again by the error microphone as $y(n)$. Hence, the loudspeaker and the error microphone define the so-called secondary path $S(z)$, which considers the characteristics of both elements plus the acoustic path between them. At the end, the control signal $y(n)$ and the disturbance $d(n)$ overlap destructively at the error microphone's position, and the residual error $e(n)$ is generated. ANC approaches that use the residual error $e(n)$ for calculating $y'(n)$ are called feedback control schemes. On the other hand, the ones that use the time-advanced reference signal $x(n)$ for calculating $y'(n)$ are called feedforward or forward control schemes. Adaptive implementations of feedforward controllers also utilize $e(n)$ to solve the underlying optimization problem.

Feedback systems are preferred over the feedforward ones in contexts where there is no economical way to provide a time-advanced reference to all noise sources, or for systems which have to perform upon

fast moving sources. Nowadays, the design and optimization of such systems are based on the mixed-sensitivity H_∞ method proposed by Zames in [3]. Successful applications of this method to the specific case of ANC headphones can be found in the literature, e. g. for analog [4] and digital [5] controller's circuitry. Here an emphasis on robust stability and performance is made, for which an uncertainty model of the secondary path is built from a set of measurements. This model is used to establish the variability over frequency that the secondary path has on different headphones' wearing situations. This variability is then used for building worst-case mathematical constraints for the controller's optimization based on an H_∞ cost function. One drawback of this approach is that a pole-zero model of each secondary path has to be matched to the measurements first, instead of using the measured impulse responses directly. Another drawback is that the optimized controller has to have a state-space formulation.

An alternative to H_∞ was proposed by Rafaely in [6] on the optimization of an Internal Model Controller [7] by means of an H_2/H_∞ design method. The robust stability against the uncertainty in the secondary path is the same as in the H_∞ method, but a cost function based on the 2-norm is used instead. Moreover, the pole-zero system modeling is avoided by the use of measured impulse responses, and the state-space controller is substituted by a finite impulse response (FIR) filter. This formulation is adapted in this work, in order to optimize the coefficients of a controller in a Minimum Variance Control (MVC) scheme instead. As shown by Rafaely, if the problem can be formulated as a convex optimization task, then a solution based on sequential quadratic programming (SQP) yields a global optimum subject to the designed constraints. Based on this, alternative cost functions are evaluated here, which aim to incorporate different aspects of the human perception of sound.

In the following section, the general problem of feedback control in ANC headphones is described. Based on the observations made, constraints for robust stability and performance are derived. Afterwards, three cost functions are introduced, which are used for optimizing the coefficients of an MVC controller. Later on, the cost functions together with the stability and performance constraints are used to optimize three controllers. The performance of each one of them is evaluated based on psychoacoustic objective metrics derived from recordings made using a dummy-head wearing the ANC headphones prototypes. At the end, conclusions are drawn based on the measured performance.

2. The Feedback Control Problem

A conventional SISO feedback system, as the one presented in Fig. 1b, is designed to control the *plant* $S(z)$ through the use of a command signal $c(n)$, by means of a controller $W_{\mathrm{mvc}}(z)$. The output signal $y(n)$ of this series connection is subject to the external disturbance $d(n)$, which generates an effective system output $e(n)$. A negative feedback loop is used to calculate a correction factor by subtracting $e(n)$ from $c(n)$, which is fed to the controller $W_{\mathrm{mvc}}(z)$.

Under ideal circumstances, the effective system's output $e(n)$ follows closely the command signal $c(n)$, implying that the transfer function

$$T(z) = \frac{E(z)}{C(z)} = \frac{S(z) \cdot W_{mvc}(z)}{1 + S(z) \cdot W_{mvc}(z)}, \tag{1}$$

also known as complementary sensitivity function, is close to unity. Furthermore, the transfer function

$$H(z) = \frac{E(z)}{D(z)} = \frac{1}{1 + S(z) \cdot W_{mvc}(z)}, \tag{2}$$

the so-called sensitivity function, is close to zero, thus rejecting the disturbance $d(n)$ in the effective system output $e(n)$. For achieving this, a controller $W_{\mathrm{mvc}}(z)$ is designed to maximize the denominator of $H(z)$, respecting certain stability and performance constraints.

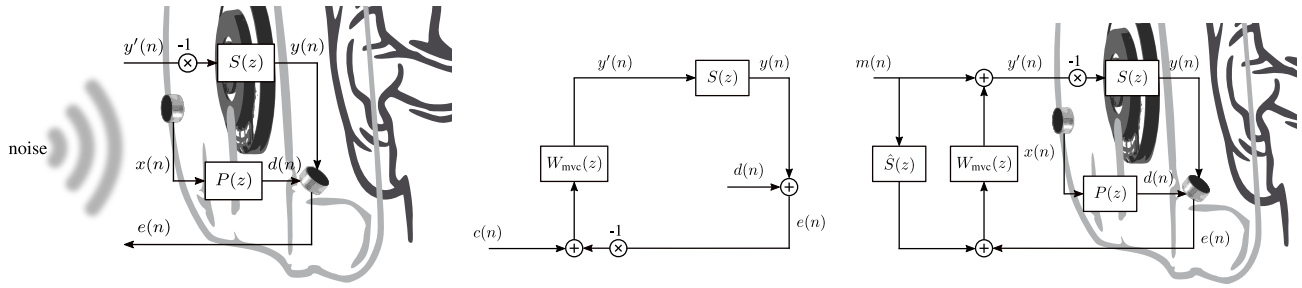


Figure 1: Feedback control systems: (a) The general ANC headphone without control algorithm. (b) The conventional feedback control system with controller $W_{\text{mvc}}(z)$, plant $S(z)$, command signal $c(n)$ and disturbance $d(n)$. (c) The feedback ANC headphones system with additional estimated plant $\hat{S}(z)$ and music signal $m(n)$, as in [5].

The application of this principle to ANC headphones involves either setting the command signal $c(n)$ to zero, if only noise attenuation is desired, or integrating the music signal $m(n)$ in the signal flow, as shown in Fig. 1c. Here, an estimated secondary path $\hat{S}(z)$ helps generating the transfer function

$$\frac{E(z)}{M(z)} = -S(z)\frac{1+\hat{S}(z)\cdot W_{mvc}(z)}{1+S(z)\cdot W_{mvc}(z)}. \tag{3}$$

If $\hat{S}(z) = S(z)$ is chosen, then the transfer function simplifies to $\frac{E(z)}{M(z)} = -S(z)$, and $m(n)$ reaches the error microphone uncolored by the controller $W_{\text{mvc}}(z)$. If the headphones are well designed, the effect of the secondary path $S(z)$ does not need further treatment. In frequency regions where $\hat{S}(z) \neq S(z)$, the transfer function depends on the gain of the controller $W_{mvc}(z)$. More specifically, in frequency regions where $W_{mvc}(z) \gg 1$, the transfer function takes the value $-\hat{S}(z)$, and in the opposite case, where $W_{mvc}(z) \ll 1$, the transfer function approximates to $-S(z)$. Thus, the transfer function moves between the asymptotes $\hat{S}(z)$ and $S(z)$, avoiding an extreme coloration.

2.1 Secondary path variability and multiplicative uncertainty model

In the context of ANC headphones, the secondary path $S(z)$ varies from person to person, because of anatomical reasons, and may change drastically upon time, because of varying wearing situations. In order to incorporate this variability in the controller design method, a so-called multiplicative uncertainty model [10] is built. A multiplicative uncertainty model is based on a set of I secondary path finite impulse responses, measured under different wearing situations and using different human subjects. The finite impulse responses $\mathbf{s}_i(n)$ of length L_s and $i \in \{0, \cdots, I-1\}$ are Z-transformed $S_{\text{i}}(z) = \sum_{n=0}^{L_s-1} \mathbf{s}_{\text{i}}(n) z^{-n}$ and evaluated in the *positive* frequencies of interest by following the well-known substitution rule $z = e^{j\Omega_k}$, with $\Omega_k = \frac{2\pi}{N}k$ and $k \in \{0, 1, \ldots, N/2-1\}$. The result is used to calculate the secondary path's maximum multiplicative magnitude deviation over frequency

$$W_2(\Omega_k) = \max_i \Big| \frac{S_i(e^{j\Omega_k}) - S_0(e^{j\Omega_k})}{S_0(e^{j\Omega_k})} \Big|, \tag{4}$$

being $S_0(e^{j\Omega_k})$ the nominal secondary path impulse response, which can be chosen as the impulse response $\mathbf{s}_i(n)$ that produces the $W_2(\Omega_k)$ with the smallest maximal magnitude over frequency. $W_2(\Omega_k)$ is used then to model the secondary path as

$$S(e^{j\Omega_k}) = S_0(e^{j\Omega_k})\Big(1 + W_2(\Omega_k)\cdot\Delta(e^{j\Omega_k})\Big), \tag{5}$$

being Δ the subset of complex numbers that fulfills the condition $\left|\Delta(e^{j\Omega_k})\right| \leq 1, \forall\Omega_k$. Thus, leaving the phase deviation unbounded.

2.2 Robust stability and nominal performance

Given the redefinition of the secondary path in Eq. (5), a controller can be designed to be robust stable upon changes in the secondary path. This can be determined by evaluating the controller's Z-transformed impulse response $\mathbf{w}_{\text{mvc}}(n)$ of length L_w in the frequencies of interest Ω_k, and calculating the open-loop transfer function

$$O(e^{j\Omega_k}) = W_{\text{mvc}}(e^{j\Omega_k}) \cdot S_0(e^{j\Omega_k}) + W_2(\Omega_k) \cdot \Delta(e^{j\Omega_k}) \cdot W_{\text{mvc}}(e^{j\Omega_k}) \cdot S_0(e^{j\Omega_k}). \tag{6}$$

By using the Nyquist stability criterion for stable plants and discrete-time controllers (which is the case with these secondary paths and controllers based on FIR filters [11]), a Nyquist plot like the one in Fig. 2 of the *nominal* open-loop transfer function $O_0(e^{j\Omega_k}) = W_{\text{mvc}}(e^{j\Omega_k}) \cdot S_0(e^{j\Omega_k})$ is drawn. By following the curve from low to high frequencies, the number of unstable poles in Eq. (2) can be determined by counting the number of clockwise encirclements that the curve does around the Nyquist point $(-1, 0)$. If no clockwise encirclement occur, then the system is *nominal* stable and the analysis can be further carried out. In a second step the discs $O_r(e^{j\Omega_k}) = W_2(\Omega_k) \cdot \Delta(e^{j\Omega_k}) \cdot O_0(e^{j\Omega_k})$ with frequency-dependent radii $|W_2(\Omega_k) W_{\text{mvc}}(e^{j\Omega_k}) S_0(e^{j\Omega_k})|$ are added on top of the nominal open-loop curve. Thus, $O(e^{j\Omega_k}) = O_0(e^{j\Omega_k}) + O_r(e^{j\Omega_k})$. Finally, if no disc touches the Nyquist point $(-1, 0)$, then the system is *robust* stable [6]. This graphical analysis can be expressed mathematically as the inequality

$$|W_2(\Omega_k) O_0(e^{j\Omega_k})| - |1 + O_0(e^{j\Omega_k})| < 0. \tag{7}$$

Another design parameter is the maximum disturbance amplification that the system is allowed to pro-

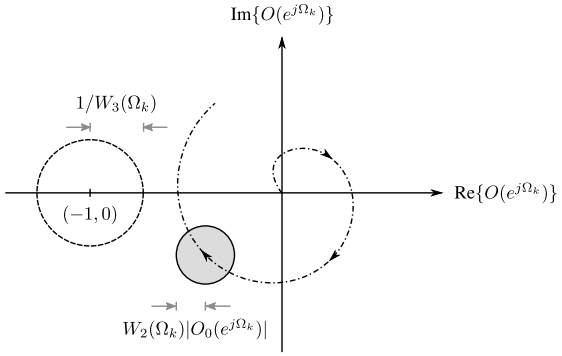


Figure 2: Nyquist plot of the open-loop system $O(e^{j\Omega_k})$ evaluated in the *positive* frequencies $0 \leq \Omega_k < \pi$, with uncertainty $W_2(\Omega_k)$ and specified performance $W_3(\Omega_k)$.

duce. In other words, what is the maximum value that the sensitivity function $H(z)$ can take. For this purpose, a maximum amplification over frequency $1/W_3(\Omega_k)$ is defined. This is equivalent to draw a circle centered in the Nyquist point with a frequency-dependent radius $W_3(\Omega_k)$, as shown in Fig. 2. The system can be designed to have a *nominal* performance, under which no point of the nominal open-loop transfer function $O_0(e^{j\Omega_k})$ is allowed to enter the circle $W_3(\Omega_k)$ drawn for its respective frequency. Alternatively, the system can be designed to have a *robust* performance. This would mean that no element in the discs $O_r(e^{j\Omega_k})$ centered in their respective $O_0(e^{j\Omega_k})$ curve points are allowed to enter the circle drawn by $W_3(\Omega_k)$. This graphical principles can be described mathematically for the robust performance, with the inequality

$$|W_3(\Omega_k)| - |1 + O_0(e^{j\Omega_k})| + |W_2(\Omega_k) O_0(e^{j\Omega_k})| < 0. \tag{8}$$

One consideration for the digital implementation is to limit the gain of the controller. This is done not only for security purposes, but also to avoid the amplification of the ADC's noise floor to a level that can be heard. This can be done by comparing the magnitude response of the controller to the one of a maximum gain curve as in the inequality

$$|W_{\text{mvc}}(\Omega_k)| - |W_4(\Omega_k)| < 0, \tag{9}$$

where $W_4(\Omega_k)$ can be chosen as an inverted ITU-R 468 noise weighting curve. Please note that if we had chosen to use the Discrete-Time Fourier Transform (DTFT), the inequalities in Eqs. (7) and (8) would have been expressed in terms of ∞-norms and they would have taken a different form than the inequalities used here, which are based on evaluated single frequencies.

3. Feedback Controller Optimization

In this section, the design of the controller is formulated as a convex optimization problem by following the strategy suggested in [6]. This guarantees that any local minimum found with sequential quadratic programming yields also the global one. The convex cost functions presented here are based on the maximization of the weighted denominator of Eq. (2)

$$C(z) = W_1(z) \cdot \Big(1 + W_{\text{mvc}}(z) \cdot S_0(z)\Big), \tag{10}$$

where $S_0(z)$ is the nominal secondary path chosen for the multiplicative uncertainty model of $S(z)$, and $W_1(z)$ as the estimated disturbance signal's energy content weighted with the ITU-R 468 noise weighting curve. The first cost function variant is defined as the sum of squared values of $C(z)$, when it is uniformly evaluated along the first half of the unit-circle as

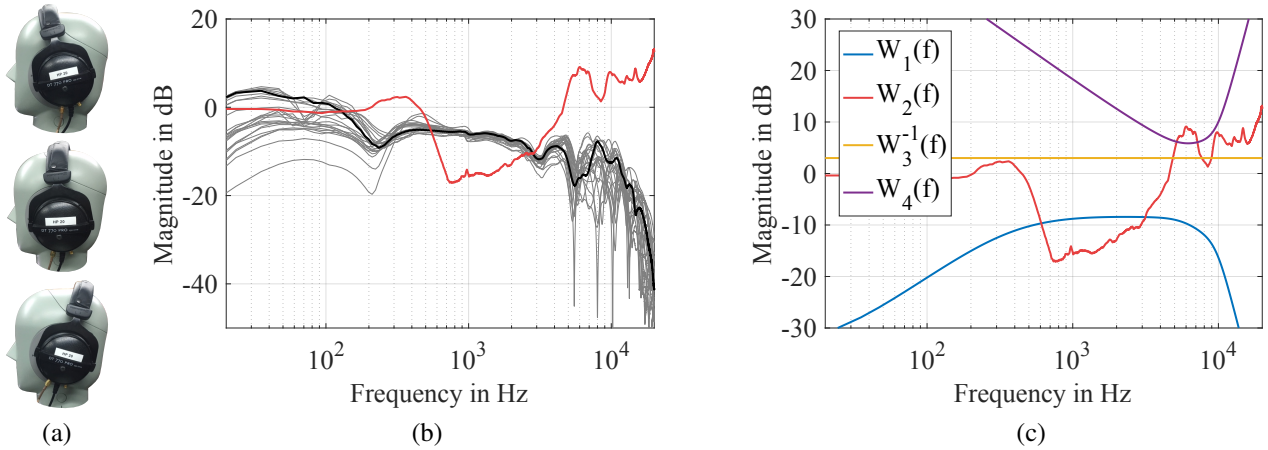
$$C_1(\mathbf{w}_{\text{mvc}}) = \min_{\mathbf{w}_{\text{mvc}}} \left(-\frac{2}{N} \sum_{k=0}^{N/2-1} \left| C(e^{j\Omega_k}) \right|^2 \right). \tag{11}$$

This cost function builds a quadratic dependency upon the controller coefficients $\mathbf{w}_{\text{mvc}}$ and is tightly related with the 2-norm of its DTFT counterpart. One problem found with this cost function is that it does not include the human perception of frequency distance. So for example, an attenuation bandwidth between 10 kHz and 11 kHz would have the same impact in the cost function as another one between 20 Hz and 1020 Hz, although the perceived bandwidth of attenuation is way higher in the second case. The cost function $C(z)$ is alternatively evaluated in exponentially spaced frequencies within the lower and upper band limits of the one-third octave bands, following the substitution $\Omega_k^{\text{exp}} = 2\pi \frac{1000}{fs} 2^{\frac{-18.5K+k}{3K}}$, with $k \in \{0, 1, \ldots, 32K\}$. For this purpose K can be chosen ≥ 1, for generating $32K$ frequencies within the one-third octave bands range. This re-spacing of the evaluation points produces that each one-third octave band has the same amount of frequency taps, and therefore the same impact in the sum of the resulting cost function

$$C_2(\mathbf{w}_{\text{mvc}}) = \min_{\mathbf{w}_{\text{mvc}}} \left(-\frac{1}{32K+1} \sum_{k=0}^{32K} \left| C(e^{j\Omega_k^{\text{exp}}}) \right|^2 \right). \tag{12}$$

In order to match the perceived attenuation over frequency that the controller under optimization produces, it is proposed in [12] to sum over dB values, as in the proposed third cost function

$$C_3(\mathbf{w}_{\text{mvc}}) = \min_{\mathbf{w}_{\text{mvc}}} \left(-\frac{1}{32K+1} \sum_{k=0}^{32K} 10 \cdot \log_{10} \left| C(e^{j\Omega_k^{\text{exp}}}) \right|^2 \right). \tag{13}$$



(a) (b) (c)

Figure 3: (a) The three wearing situations used during the secondary path measurements. (b) 24 secondary paths measured with eight subjects using the three wearing situations (in gray), the nominal secondary path (in black), and the resulting W_2 over frequency (in red). (c) The weighting functions utilized for the controller optimization.

However, because of the calculation of the logarithm over the squared values, the quadratic dependency with respect to the controller coefficients $\mathbf{w}_{\text{mvc}}$ is broken. This means that a local minima found with this cost function is not automatically the global one.

Together with the cost functions, the constraints are also re-defined to yield a convex programming problem. Because of this, the inequality in Eq. (7) for robust stability is re-written as the constraint $|W_2(\Omega_k)O_0(e^{j\Omega_k})|^2 - |1 + O_0(e^{j\Omega_k})|^2 < 0$, the inequality in Eq. (8) for robust performance yields the constraint $|W_3(\Omega_k)|^2 - |1 + O_0(e^{j\Omega_k})|^2 + |W_2(\Omega_k)O_0(e^{j\Omega_k})|^2 < 0$, and the Eq. (9) for limiting the controller's gain are re-written in its respective squared terms as $|W_{\text{mvc}}(\Omega_k)|^2 - |W_4(\Omega_k)|^2 < 0$. These changes do not alter the constraints, but introduces a quadratic dependency with respect to $\mathbf{w}_{\text{mvc}}$ [6].

4. Results and Evaluation

In order to evaluate the proposed cost functions, a multiplicative uncertainty model is built from the impulse response measurement of the secondary path following the procedure described in [13], under the three different wearing positions presented in Fig. 3a, using two dummy-heads, and six human subjects. In Fig. 3b, the frequency responses of the resulting 24 secondary path measurements are presented. The W_2 with the smallest maximum magnitude is presented in red, while its respective S_0 is presented in black. In Fig. 3c the weighting curves used during the optimization are presented. W_1 is chosen as the PSD of a low-passed filtered white noise, weighted with the ITU-R 468 curve. The maximum sensitivity function's disturbance amplification W_3 is chosen as 0 dB for $f = 0$ Hz and 3 dB otherwise. The maximum controller gain over frequency W_4 is chosen as an inverted ITU-R 468 curve, which relative level is repeatedly changed until the noise floor of the optimal controllers is not heard when using the prototype. The length of the controllers' impulse response $\mathbf{w}_{\text{mvc}}(n)$ is chosen to be 512 coefficients. The Z-transforms of the impulse responses are all evaluated in 2048 frequencies, following the described distributions. The optimizations are performed using the SQP algorithm provided by MATLAB for function minimization under constraints. The initial point of the optimization is chosen as a unit impulse with amplitude 0.05. The optimizations are carried out during 20 iterations, after which the optimization is stopped and the controller coefficients are used for the measurements.

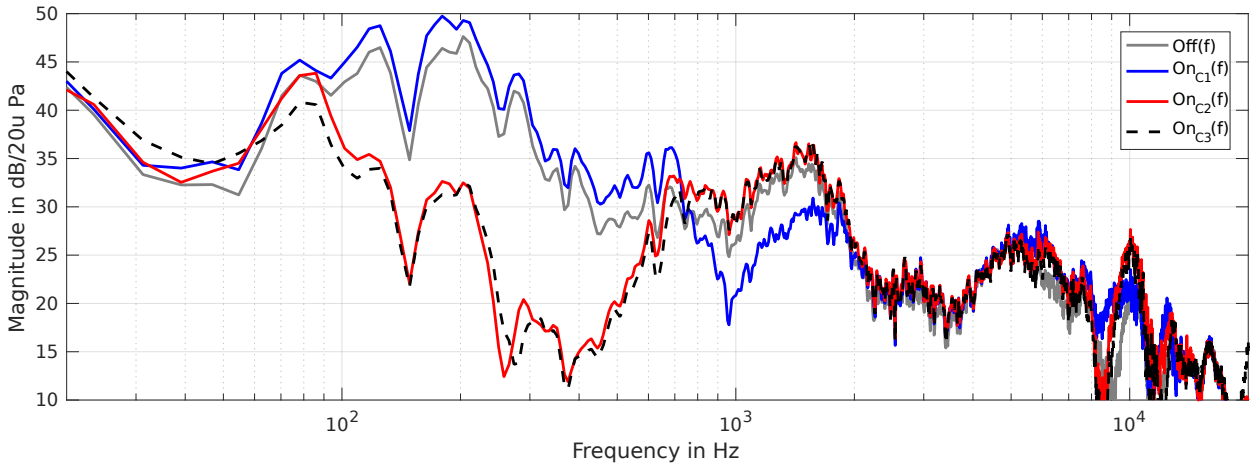


Figure 4: Pressure in dB/20u Pa measured at the dummy-head's left ear.

For the evaluation the setup described in [14] is used. The excitation signal is a Gaussian white noise amplitude-modulated with a modulation frequency of 70 Hz and a modulation degree of 30%. The pressure over frequency and objective metrics are based on the dummy-head recordings and calculated using the BK Connect. In Fig. 4, the pressure level over frequency on the dummy-head's left ear is presented. As a reference, the disturbance when the system is turned off is shown in gray. When the controller optimized with C_1 is turned on, a frequency range between approx. 800 Hz and 2 kHz shows an attenuation of around 5 dB, while the rest of frequencies show a constant amplification of 3 dB, as designed, with exception of the frequencies around 10 kHz, where the amplification is higher. This probably because the measurements are done at the dummy-head's mic and not at the error mic. On the other hand, when the controller derived with C_2 is used, higher attenuation levels of up to 20 dB in the frequency range between 90 and 650 Hz are measured. Similarly to the case of C_1, outside of the attenuation bandwidth an amplification of 3 dB is produced and an amplification around 10 kHz is measured. When the controller optimized with $C3$ is used, the attenuation bandwidth achieved by $C2$ increases towards the low frequencies down to 65 Hz.

In Fig. 5, the psychoacoustic objective metrics derived from the recordings are presented. It can be seen in Fig. 5a that the controller derived with C_1 produces the highest loudness, because it amplifies by 3 dB the most prominent frequencies around the 200 Hz. On the other hand, the controllers optimized with $C2$ and $C3$ produce a high attenuation in that region and therefore, a loudness improvement is measured. In Fig. 5b, the higher amplification measured around the 10 kHz shows an effect in the increment of the perceived sharpness, which is more prominent in the case of $C2$ and $C3$, because these controllers also attenuate the lower frequencies. In Fig. 5c it can be seen that the perceived roughness shows only small variations around the 4 asper.

5. Conclusions

In this work an adapted H_2/H_∞ optimization method for the derivation of a feedback controller has been described. With this method, the coefficients of an FIR filter can be found which yield a robust stable system with robust performance upon changes in the secondary path. Based on the measurement results, it could be shown that the cost function C_3 based on one-third octave bands and accumulation over dB values produces the best overall performance. In a future work a listening test should be conducted considering realistic acoustic scenarios, e.g. inside of an aircraft, a coffee shop, or the subway.

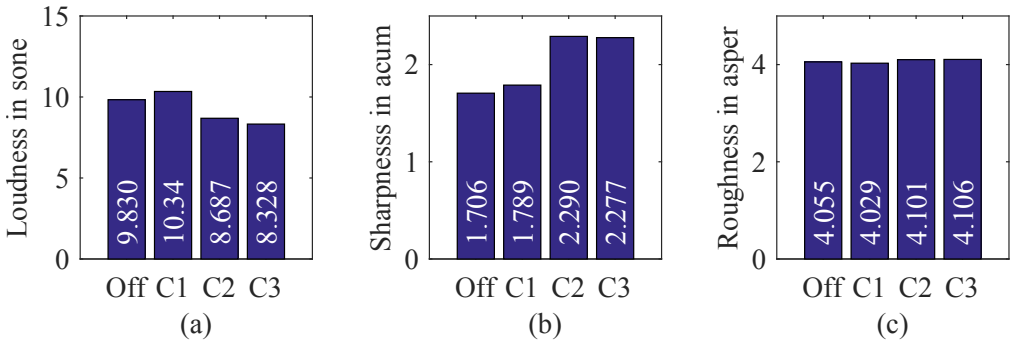


Figure 5: Psychoacoustic objective metrics calculated from the recordings done with the dummy-head using the controllers derived with the three cost functions (a) Loudness in sone, (b) Sharpness in acum, and (c) Roughness in asper.

REFERENCES

1. Lueg, P., (1936), *Process of silencing sound oscillations*. US Patent 2,043,416.
2. Simshauser, E. D. and Hawley, M. E. The noise cancelling headset—an active ear defender, *The Journal of the Acoustical Society of America*, **27** (1), 207–207, (1955).
3. Zames, G. Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses, *IEEE Transactions on Automatic Control*, **26** (2), 301–320, (1981).
4. Bai, M. and Lee, D. Implementation of an active headset by using the H_∞ robust control theory, *The Journal of the Acoustical Society of America*, **102** (4), 2184–2190, (1997).
5. Liebich, S., Anemüller, C., Vary, P., Jax, P., Rüschen, D. and Leonhardt, S. Active noise cancellation in headphones by digital robust feedback control, *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug, pp. 1843–1847, (2016).
6. Rafaely, B., *Feedback Control of Sound*, Ph.D. thesis, University of Southampton, Southampton, England, (1997).
7. Morari, M. and Zafiriou, E., *Robust process control*, Prentice-Hall, NJ (1989).
8. Gill, P., Murray, W. and Wright, M., *Practical optimization*, Academic Press (1981).
9. Fletcher, R., *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, USA, second edn. (1987).
10. Skogestad, S. and Postlethwaite, I., *Multivariable Feedback Control: Analysis and Design*, John Wiley & Sons, Inc., USA (2005).
11. Franklin, G. F., Workman, M. L. and Powell, D., *Digital Control of Dynamic Systems*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 3rd edn. (1997).
12. Pawełczyk, M. Analogue active noise control, *Applied Acoustics*, **63** (11), 1193–1213, (2002).
13. Rivera Benois, P., Bhattacharya, P. and Zölzer, U. Derivation Technique for Headphone Transfer Functions Based on Sine Sweeps and Least Squares Minimization, *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, **253** (4), (2016).
14. Rivera Benois, P., Veatriki Papantoni and Zölzer, U. Psychoacoustic Hybrid Active Noise Control Structure for Application in Headphones, *2017 Proceedings of the 25th International Congress on Sound and Vibration*, July, (2018).