Evaluation of a Decoupled Feedforward-Feedback Hybrid Structure for Active Noise Control in a Multi-Source Environment

Piero Rivera Benois*, Patrick Nowak*† and Udo Zölzer*
*Department of Signal Processing and Communications, Helmut Schmidt University, Germany
†Acoustics and Protection of Soldier Group, French-German Research Institute of Saint-Louis, France
Email: piero.rivera.benois@hsu-hh.de

***Abstract*—Feedforward control structures for active noise control headphones suffer from challenging processing delay constraints, due to the small distances between transducers and their varying relative orientation to the noise sources. On top of that, the context of its usage comprehends a multi-source environment, which prevents to provide the system with a strongly correlated and time-advanced reference for all incoming signals. Combining a feedforward structure with a feedback structure gives the opportunity to complement the attenuation capabilities of the system with a control strategy that is not based on a time-advanced reference, but on the measured residual error of the noise cancellation. If both structures are connected in a certain way, the controllers can be designed or adapted individually, resulting in an independent contribution to the overall attenuation. In the literature such a decoupled feedforward-feedback hybrid structure has been suggested and simulated to evaluate its capabilities to attenuate signals that are completely uncorrelated with the reference. In this work, a real-time FPGA implementation of it is presented and evaluated in a reverberant multi-source environment using a real prototype.**

***Index Terms*—ANC headset, ANC headphones, decoupled design, hybrid control, multiple noise sources.**

I. Introduction

Active Noise Control (ANC) headphones offer their users the possibility to attenuate the environmental noise that surrounds them, while they listen to music. The attenuation they produce is a combined effect of the passive attenuation characteristics of its construction materials and the active attenuation generated to attenuate the remaining noise that effectively penetrates into the earcup. The generation of the control signals may be based on the analog or digital filtering of the sound measured by means of microphones outside and/or inside of the headphone. Control schemes that filter the sound measured outside of the earcup are denominated as feedforward approaches. Those approaches which instead use the sound inside of the earcup for the filtering are denominated as feedback approaches.

The combination of feedforward and feedback approaches into a hybrid structure provides the possibility to combine the performance of the single control structures into one and compensate for their individual limitations. Particularly attractive for the feedback system is the Internal Model Controller (IMC), because the controller can be derived using Wiener filter theory [1] or on-line, using adaptive filters based on FxLMS [2]. If connected as proposed in [3] and [4], both systems combine completely decoupled from each others. This means that the optimum solution of one subsystem can be searched without influencing the optimum of the other, and that the overall transfer function will consist of the multiplication of the ones of the subsystems.

In this work, the system proposed in [3] is implemented on an FPGA platform following the filtering strategy suggested in [5], and studied with the use of an ANC headphone prototype under multi-source circumstances inside of a reverberant room. Special attention is paid to understand how optimum solutions found under the excitation of single sources can be used to anticipate the performance under multi-source scenarios. Between the aspects that are taken into account are complete and partial overlapping frequency bands and the different angle of incidence of the excitation signals.

In the following section, the hybrid structure is described and the decoupling effect together with its advantages is explained in detail. Afterwards, the setup and prototype used for the evaluation are described. The results are then presented and evaluated. At the end, conclusions are drawn from the results obtained.

II. Hybrid Decoupled Structure

The hybrid decoupled control structure presented in a simplified diagram in Fig. 1, aims to combine the attenuation capabilities of a feedforward FxLMS control structure (See Fig. 2a) and an adaptive IMC structure based on an FxLMS adaption algorithm (See Fig. 2b). In the diagram, $x(n)$ represents the incoming noise signal measured by the reference microphone outside the earcup. This signal is filtered by the construction materials of the earcup, represented by $P(z)$, to form $d(n)$. Simultaneously, $x(n)$ is filtered by the controller $W_f(z)$, and added with the control signal coming from the IMC, $\hat{y}_b(n)$. The resulting control signal, $\hat{y}(n)$, is then played through the secondary path, $S(z)$, which represents the digital-to-analog converter, power amplifier, headphones drive, and acoustic path until the control signal reaches the sweetspot as $y(n)$. The sweetspot of the cancellation is defined by the position of the error microphone. In this place, $d(n)$ and $y(n)$ overlap destructively and leave a residual cancellation error $e(n)$. This

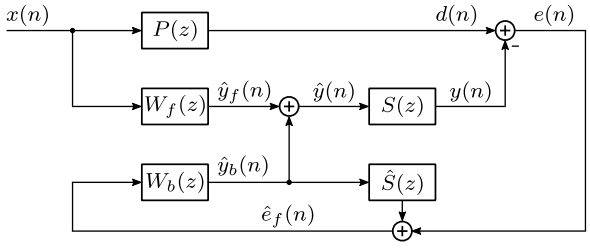
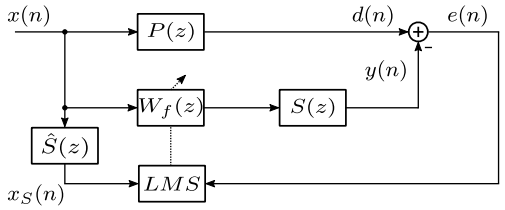
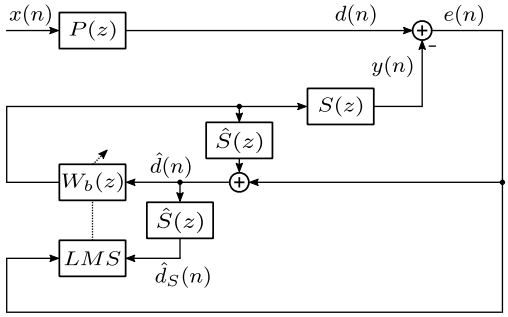


Fig. 1. Block diagram of the hybrid structure with the feedforward controller $W_f(z)$ and IMC controller $W_b(z)$.

signal is used together with the IMC control signal $\hat{y}_b(n)$ and an estimation of the secondary path $\hat{S}(z)$ to estimate the residual error left only by the feedforward controller, $e_f(n)$. Subsequently, $e_f(n)$ is filtered by the IMC controller, $W_b(z)$, to generate the complementary control signal $\hat{y}_b(n)$. Finally, the feedforward control signal $\hat{y}_f(n)$ and $\hat{y}_b(n)$ are added together again to form the next sample of the control signal.



(a) Feedforward



(b) IMC

Fig. 2. Block diagrams of the adaptive control structures: (a) feedforward control and (b) IMC control.

If the Z-transformed signals are taken and the main equations defined by the structure

$$D(z) = P(z) \cdot X(z), \tag{1}$$

$$\hat{Y}_f(z) = W_f(z) \cdot X(z), \tag{2}$$

$$\hat{Y}_b(z) = W_b(z) \cdot (E(z) + \hat{S}(z) \cdot \hat{Y}_b(z)), \tag{3}$$

and

$$E(z) = -S(z) \cdot (\hat{Y}_f(z) + \hat{Y}_b(z)) + D(z) \tag{4}$$

are used, then its transfer function

$$\frac{E(z)}{X(z)} = (P(z) - S(z) \cdot W_f(z)) \frac{1 - \hat{S}(z) \cdot W_b(z)}{1 - W_b(z)(S(z) - \hat{S}(z))} \tag{5}$$

can be derived. If the condition $\hat{S}(z) = S(z)$ is met, then the whole simplifies to

$$\frac{E(z)}{X(z)} = (P(z) - S(z) \cdot W_f(z)) \cdot (1 - \hat{S}(z) \cdot W_b(z)), \tag{6}$$

from which the left term is the feedforward transfer function and the right term the one of the IMC.

To evaluate the effect of the IMC controller into the effective secondary path of the feedforward structure (3) is used in (4) and $D(z)$ is set to zero

$$E(z) = -S(z) \cdot \left(Y_f(z) + \frac{W_b(z) \cdot E(z)}{1 - \hat{S}(z) \cdot W_b(z)} \right). \tag{7}$$

Then, the secondary path seen from the feedforward structure

$$\frac{E(z)}{Y_f(z)} = \frac{-S(z)}{1 + W_b(z)(S(z) - \hat{S}(z))} \tag{8}$$

can be simplified through the condition $\hat{S}(z) = S(z)$ to

$$\frac{E(z)}{Y_f(z)} = -S(z). \tag{9}$$

This means that after any change in $W_b(z)$, there is no further need for updating $\hat{S}(z)$ in the feedforward FxLMS (See Fig. 2a). So the feedforward and the IMC control structures can be adapted in any way, without breaking the condition $\hat{S}(z) = S(z)$.

III. Measurement Setup

The overview of the measurement setup is presented in Fig. 3. A Neumann KU100 dummy-head is placed inside of a room designed for audio-listening with dimensions 4.80 x 4.20 x 2.05 m. Surrounding it, two Genelec 8030B speakers are placed 1 m away from it. One faces the dummy-head directly from the front and the other one from the right side. Both speakers are connected to an RME Fireface UCX audio interface (bottom front), which is used to generate the excitation signals for the measurements. An ANC headphones prototype based on a Beyerdynamic DT 770 PRO headset customized with inner and outer electret microphones on its right earcup is utilized as system under test. Preamplifiers with a gain of 20 dB are used to better match the expected dynamic range of the microphones signal with the one of the dSpace MicroLabBox ($\pm$10 V). Additional to that, a Behringer Powerplay Pro-8 headphones power amplifier is used.

A 2048 samples long impulse response of the secondary path, $S(z)$, is measured beforehand using the method proposed in [6]. A fixed-point VHDL implementation of the hybrid decoupled structure running at a sampling frequency of 48 kHz is programmed following the filtering strategy suggested in [5]. The length of the adaptive filters is chosen to be equal to 2048 samples with a word-length of 64 bits, from which 48 bit are dedicated for the decimal part. The step-sizes $\mu_f = 4 \cdot 10^{-12}$ for the feedforward FxLMS and $\mu_b = 3 \cdot 10^{-12}$ for the adaptive IMC substructures are chosen after a manual iteration on the trade-off between adaption speed and stability. Please note that the data coming from the ADC units is a 16 bit

Fig. 3. Measurement setup overview. Dummy head wearing the ANC headphone prototype (center), connected to the dSpace platform (in white on the bottom) and headphone amplifier (in black on the bottom), and the external loudspeakers controlled by an RME Fireface UCX (bottom front) audio interface to simulate noise sources.

(signed) fixed-point integer number, which is the main reason for such small μ values.

Around the hybrid structure filtering process, control logic is programmed to switch on and off the adaption of each one of the substructures and choose between feedforward control, IMC control, and hybrid control. The system is additionally programmed to deliver 48 times per second a 4096 taps long Hann windowed short-time FFT of the measured residual error signal $e(n)$ (see Fig. 1). The transformed values are then time-averaged

$$E_{avg}(m,\omega) = \alpha \cdot E(m,\omega) + (1-\alpha) \cdot E_{avg}(m-1,\omega), \quad (10)$$

with $\alpha = 0.001$, for evaluation and visualization purposes.

IV. Multi-Source Evaluation and Results

The measurement setup described in the previous section is used to study the performance of the hybrid decoupled structure under multi-source scenarios. To better understand the results obtained in such a complex context, the results obtained under multi-source excitation are compared with the ones of their respective single-source sub-cases. Special emphasis is made on understanding how the performance of the sub-structures degrade when the sources completely or partially overlap in frequency bands, to then explain the performance reached by the hybrid control structure.

For evaluation purposes, uniformly distributed white noise is filtered with three different filters, whose magnitude responses are presented in Fig. 4. The 0-1000 Hz band (dashed in black) is generated by a Butterworth lowpass filter with $f_{pass} = 1000$ Hz, $f_{stop} = 1200$ Hz, and 80 dB attenuation at the stop-band. The 0-500 Hz and 500-1000 Hz power-complementary sub-bands (solid red and solid yellow, respectively) are generated by further filtering the lowpass-filtered signal with a filter bank. The filter bank is designed following the procedure described in [7], based on a 7^{th}-order Butterworth IIR filter with $f_c = 500$ Hz.

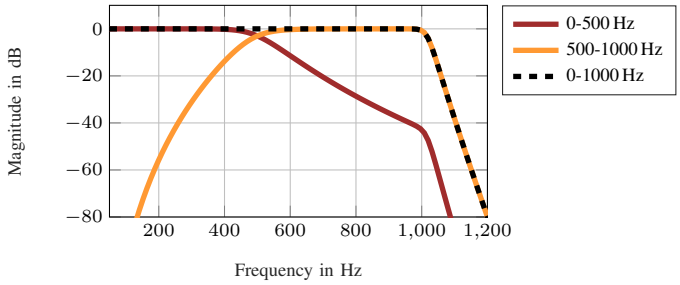


Fig. 4. Magnitude responses of the filters used for generating the evaluation signals: In dashed black, the magnitude response of first lowpass filter; In solid red and solid yellow, the magnitude response of the first and second sub-band of the filter bank.

A. Evaluation of Complete Overlapping Sources

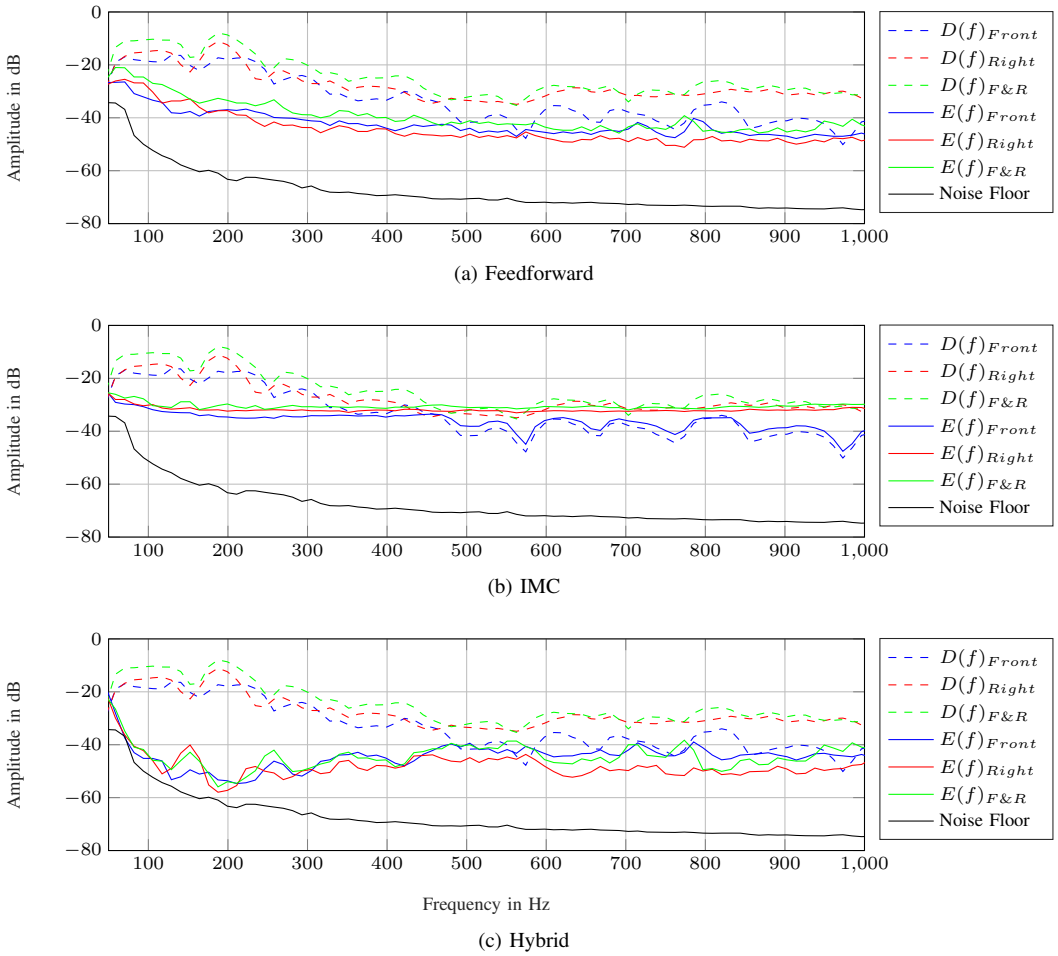
To evaluate the scenario of complete overlapping sources, the lowpass-filtered signal (dashed back curve in Fig. 4) is used as shown in Fig. 5. First, the signal is played simultaneously through both speakers. The sub-structures sequentially adapt for 5 min each, under the excitation of the overlapping signals. Their individual residual error signals are measured. The complete hybrid structure is then switched on and the final residual error is measured. The same procedure is repeated playing the excitation signal from the front speaker and then from the right speaker.

(a) Front and right simultaneously (b) Front: 0-1000 Hz (c) Right: 0-1000 Hz

Fig. 5. Procedure utilized to study complete overlapping of two noise sources

The results obtained with complete overlapping noise sources are presented in Fig. 6. The plots are 4096 taps long double-sided magnitude FFT spectra of the error signal represented in dB. The individual values are scaled up by a factor 32 for convenience, and the frequency range of interest (50-1000 Hz) is presented. Additionally to the residual error signals denoted with $E(f)$, a related $D(f)$ is plotted in the same color to show what is the noise level present inside of the earcup when there is no control signal. A measurement in complete silence is additionally provided, for having a reference of the noise floor of the channel.

It can be seen in Fig. 6a through the comparison between $D(f)_{Front}$ and $D(f)_{Right}$, that the passive attenuation of the materials varies significantly with the incoming noise direction, producing specifically in the 500-1000 Hz frequency range a better passive attenuation over noise coming from the



(a) Feedforward

(b) IMC

(c) Hybrid

Fig. 6. Error signal measured spectra under (a) feedforward control, (b) IMC control, and (c) hybrid control. Disturbance signal is lowpass-filtered uniformly distributed pseudo random noise, which played at 1 m distance first from the front, then from the right side of the dummy-head perspective, and then simultaneously from both directions.

front. Nevertheless, the feedforward control structure compensates quite strongly for the lack of passive attenuation, producing around 15-25 dB attenuation levels, but vanishing in the lowest frequencies around 50 Hz. Interesting is to see that, although being $E(f)_{Right}$ constantly lower than $E(f)_{Front}$, the behavior inverts in the region below 200 Hz, probably because of the earcup vibration caused by the sound pressure wave coming from the right side.

The curve of the overlapping noise sources, $D(f)_{F\&R}$, shows a constructive superposition of the individual signals, $D(f)_{Front}$ and $D(f)_{Right}$, with no recognizable strong notches. Its corresponding residual noise, $E(f)_{F\&R}$, shows to be equal or higher as the single-source scenarios, produced because of the need to find a single set of FIR coefficients to attenuate two signals coming from different directions. Moreover, all three residual error signals show to reach a similar optimum, which does not related to the excitation signal frequency content, but instead to the noise floor. This is a sign of strong correlated noise floors between $x(n)$ and $e(n)$ (see Fig. 2a), which then manifest as multiplicative noise in the update equation of the coefficients in the adaptive filter

$$\underline{w}_f(n+1) = \underline{w}_f(n) + \mu_f \cdot \underline{x_S}(n-1) \cdot e(n). \qquad (11)$$

In Fig. 6b, the same behavior in the residual errors $E(f)_{Right}$ and $E(f)_{F\&R}$ is observed. In this case, the minimum reachable residual error is higher, because $\hat{d_S}(n)$ is derived from $e(n)$ (see Fig. 2b), and this increases the cross-correlation and therefore the multiplicative noise in the coefficients update equation

$$\underline{w}_b(n+1) = \underline{w}_b(n) + \mu_b \cdot \underline{d_S}(n-1) \cdot e(n). \qquad (12)$$

Interesting is the case of $E(f)_{Front}$, where its energy content beyond this minimum residual error is attenuated, and the rest falls inside of a non-care region, in some extent similar to the water-bed effect [2]. This may suggest that, the stronger passive attenuation can be seen as a disadvantage in such a constellation, forcing the excitation signals to be loud enough during the adaption, so some attenuation may still be possible afterwards.

The results presented in Fig. 6c show how the hybrid structure combines the attenuation capabilities of its individual

substructures. The residual errors $E(f)_{Front}$, $E(f)_{Right}$, and $E(f)_{F\&R}$ show now to have similar behaviors, because the residual error of the sub-structures showed a strong dependency on the noise floor, instead of a dependency on the noise levels that reached the headphone. However, if only the attenuation values are analyzed, a clear advantage can be seen in the situations where the noise that entered the cup is higher.

B. Evaluation of Partial Overlapping Sources

To evaluate the scenario of partial overlapping sources, the power-complementary sub-bands (solid red and solid yellow curves in Fig. 4) are used as shown in Fig. 7. For starting, the lower sub-band is played through the right speaker, while at the same time the higher sub-band is played through the front speaker. Analog to the previous evaluation, the sub-structures are sequentially adapted for 5 min each, under the excitation of both sources. Their individual residual error signals are then measured. The complete hybrid structure is then switched on and the resulting residual error is measured. The same procedure is repeated playing only one sub-band at a time from its respective speaker.

(a) Front and right simultaneously (b) Front: 500-1000 Hz (c) Right: 0-500 Hz

Fig. 7. Procedure utilized to study partial overlapping of two noise sources

In Fig. 8 the results obtained with partial overlapping sources are presented. In Fig. 8b can be clearly seen how the sub-bands share the same shape in their pass-band regions as their counterparts in Fig. 6, and outside these regions they decay with a slope similar to the one expected from Fig. 4. Moreover, $D(f)_{F\&R}$ matches perfectly the passband of the sub-bands, with a transition band between 420 and 700 Hz, in which mainly destructive superposition occurs.

In Fig. 8a, the residual error signal $E(f)_{F\&R}$ of the feed-forward sub-structure shows a behavior that can be divided into three frequency regions: In the low frequency region between 50 and 420 Hz, $E(f)_{F\&R}$ follows the behavior of $E(f)_{Right}$ as their signals were louder than the minimum residual error, whereas the $D(f)_{Front}$ suffers important amplification, because of being in the non-care region below the minimum residual error; In the middle frequency region between 420 and 700 Hz, $E(f)_{F\&R}$ gets leveraged to the minimum residual error, mainly through attenuation, but also with some amplification around the 550 Hz; In the high frequency region between 700 and 1000 Hz, $E(f)_{F\&R}$ follows now the behavior of $E(f)_{Front}$, whereas $D(f)_{Right}$ drops below into the non-care region, suffering similar amplification values as $D(f)_{Front}$ in the lower frequency region.

In Fig. 8b, it can be seen that the level of $D(f)_{Front}$ felt completely under the minimum residual error, and this prevented the adaption algorithm to work correctly. Therefore, $D(f)_{Front}$ is neither significantly attenuated nor amplified. One detail worth to mention is that the level reached by $E(f)_{Right}$ in its sub-band is slightly lower than its counterpart in Fig. 6b. Although the difference is small (around 1-3 dB), it could show an advantage of the adaption algorithm in attenuating smaller sub-bands.

In Fig. 8c, the results measured with the hybrid structure are presented. Similarly to the results seen in the last subsection, the best results are concentrated in the low frequency region, where the noise levels inside the earcup are the highest. Although attenuation is still measured in the middle and high frequency regions, the behavior is not constant and strongly varies together with the noise level present coming from the front.

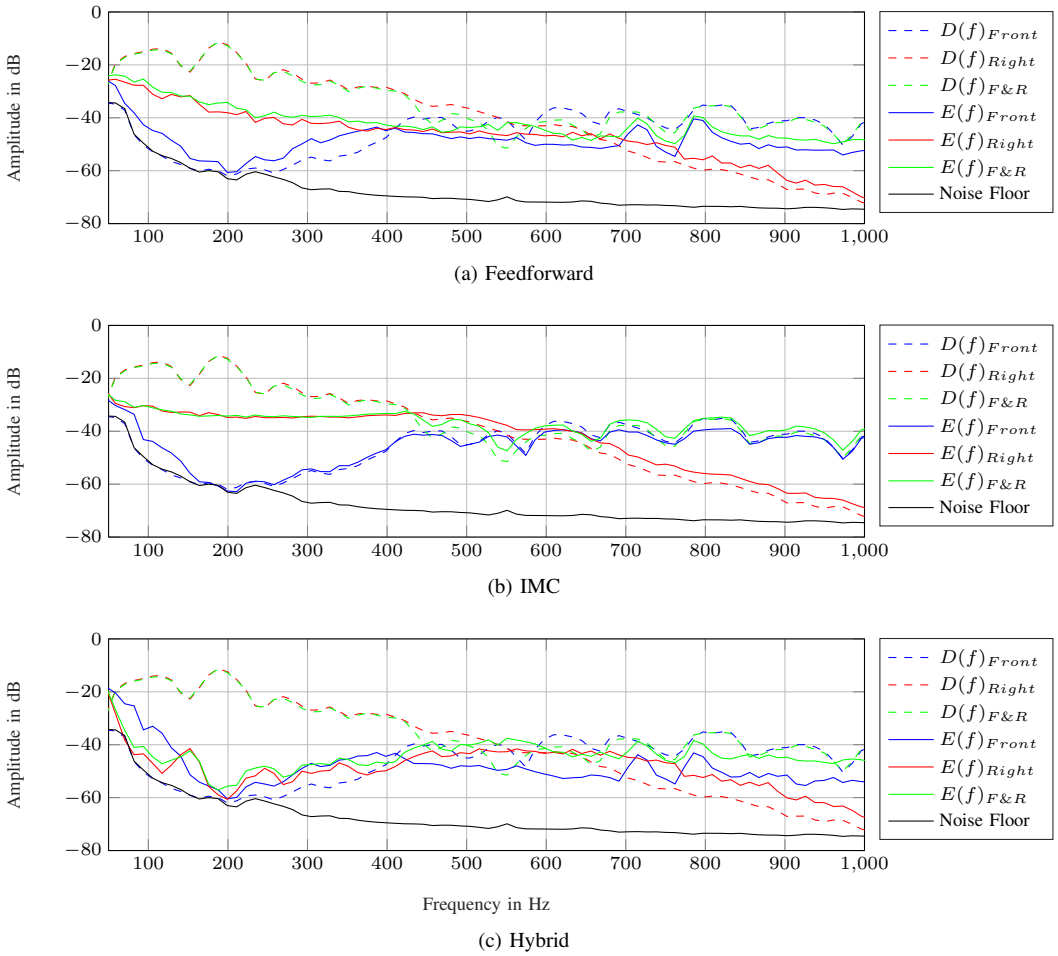
V. Conclusions

An ANC headphones prototype equipped with an FPGA implementation of a hybrid decoupled structure algorithm has been evaluated under a multi-source context inside of a reverberant room. The analysis of the results obtained with the hybrid structure were enriched by providing information about the contribution of the sub-structures to the whole, and comparing the multi-source results with the ones of their respective single-source sub-cases.

The analysis of complete overlapping sources provided a clue about how the cross-correlation between the noise floors of signals involved in the adaption algorithm limit the minimum reachable residual error. In this aspect, the IMC sub-structure showed to be more sensible to such limitations, because both inputs to the adaption algorithm come from the same physical channel.

The analysis of the partial overlapping regions showed that, if the signal level partially falls below the reachable minimum residual error, the adaption algorithm produces some amplification in that frequency range. If it is the case that the signal level completely falls below the reachable minimum, the adaption algorithm produces neither an important amplification nor attenuation. Additionally to this, a small attenuation improvement has been measured within a sub-band, when the system is excited during the adaption specifically in that narrow frequency range and not in a broader one.

All in all, by looking at the adaption results under partial and complete overlapping sources together, it can be concluded that the loudness of the excitation signals during the adaption process should be at least over the minimum reachable residual error for the adaption to work properly. If louder excitation signals for the adaption process are not possible, decreasing the noise floor in the analog channels and their cross-correlation can also provide an improvement.



(a) Feedforward

(b) IMC

(c) Hybrid

Fig. 8. Error signal measured spectra under (a) feedforward control, (b) IMC control, and (c) hybrid control. Disturbance signal is lowpass-filtered uniformly distributed pseudo random noise, divided into two subbands: the first subband 500-1000 Hz played from the front of the dummy-head, then the second subband 50-500 Hz played from the right side of the dummy head's perspective, and then both simultaneously played from their respective positions.

REFERENCES

[1] W.-K. Tseng, B. Rafaely, and S. J. Elliott, "Combined feedback-feedforward active control of sound in a room," *The Journal of the Acoustical Society of America*, vol. 104, no. 6, pp. 3417–3425, 1998. [Online]. Available: http://dx.doi.org/10.1121/1.423925

[2] S. J. Elliot, *Signal Processing for Active Control*, ser. Signal Processing and its Applications. London: Academic Press, 2001.

[3] L. Wu, X. Qiu, I. S. Burnett, and Y. Guo, "Decoupling feedforward and feedback structures in hybrid active noise control systems for uncorrelated narrowband disturbances," *Journal of Sound and Vibration*, vol. 350, pp. 1 – 10, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022460X15003508

[4] H. Foudhaili, "Kombinierte feedback- und adaptive Feedforward-Regelung für aktive Lärmreduktion in einem Kommunikations-Headset," Ph.D. dissertation, Aachen, 2008.

[5] P. Rivera Benois, P. Nowak, and U. Zölzer, "Fully Digital Implementation of a Hybrid Feedback Structure for Broadband Active Noise Control in Headphones," in *2017 Proceedings of the 24th International Congress on Sound and Vibration*, July 2017.

[6] P. Rivera Benois, P. Bhattacharya, and U. Zölzer, "Derivation Technique for Headphone Transfer Functions Based on Sine Sweeps and Least Squares Minimization," *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 253, no. 4, 2016.

[7] S. Kraft and U. Zölzer, "Time-Domain Implementation of a Stereo to Surround Sound Upmix Algorithm," in *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, September 2017.