www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

PAN ARCHIVES OF ACOUSTICS
Vol. **38**, No. 1, pp. 27–37 (2013)

DOI: 10.2478/aoa-2013-0004

Efficient Phantom Source Widening

Franz ZOTTER, Matthias FRANK

Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz
Inffeldgasse 10/3, 8010 Graz, Austria; e-mail: {zotter, frank}@iem.at

(*received July 9, 2012; accepted December 6, 2012*)

We present a highly efficient filter structure to create power-complementary filter pairs for phantom source widening. It either introduces frequency-dependent phase or amplitude differences in a pair of loudspeaker signals. We evaluate how the perceptual effect is influenced by off-center listening positions in a standard $\pm 30^\circ$ loudspeaker setup. The evaluation of the phantom source widening effect is based on measurements of the inter-aural cross-correlation coefficient (IACC), which is justified by its pronounced correlation to the perceived phantom source width in prior listening test results.

Keywords: source widening, phantom source, decorrelation, IACC.

1. Introduction

A pair of loudspeakers symmetrically arranged with regard to the listener and driven by the same signal evokes a perceived auditory event in the middle between the loudspeakers. The location of this auditory event, the so-called phantom source (Wendt, 1963), is generally known to be influenced by time-delay and/or level differences of the loudspeaker signals (Wendt, 1963; Plewa, Kleczkowski, 2011). In general, it appears to have a width (Kin, Plaskota, 2011) that can be different to the width of a real sound source (Zotter *et al.*, 2011). The literature about (pseudo-)stereophony (Gerzon, 1993; Schröder, 1958; Orban, 1970b), stereo decorrelation (Kendall, 1995; Potard, Burnett, 2004; Bouéri, Kyriakakis, 2004), and parametric spatial audio (Laitinen *et al.*, 2012; Potard, 2006; Szczerba *et al.*, 2011) describes that nonuniform amplitude differences or phase differences over frequency are suitable for controlling the phantom source width. Phantom source widening provides a way of shaping the spatial salience of a sound signal with only little effect on its location, reverberation, or coloration.

A recently presented sinusoidal-phase all-pass method (Zotter *et al.*, 2011) could be tested in varying parameters. It seemed to be quite efficient and effective according to listening tests that were done at the central listening spot. Moreover, a pronounced negative correlation between perceived source width and the inter-aural coherence (measured by the inter-aural cross-correlation coefficient, IACC) could be experimentally proven. Nevertheless, the question remains how good the method works for outside the central listening spot and how much audible spectral corruption is produced.

Among the known methods, not all are phase-based. The known amplitude-based approaches are not similar to the phase-based method (Zotter *et al.*, 2011) in various ways (response structure, computational efficiency, power-complementarity). If a more similar amplitude-based phantom source widening algorithm was found, it would be useful to pose the question, whether amplitude-based or phase-based widening would be favorable.

This article therefore reviews suitable deterministic phantom source widening strategies in Sec. 2 with regard to their differences, considering computational effort, power-complementarity, and phasiness documented by Gerzon (1993). A further simplified version of the phase-based method from (Zotter *et al.*, 2011) is presented in Sec. 3, making its implementation easier and its highly efficient nature clearly obvious. As alternative, Sec. 4 presents a novel amplitude-based method that uses the same filter structure as the phase-based one and re-uses its coefficients. Section 5 is a review of the experimental results from (Zotter *et al.*, 2011) and the IACC as a predictor, and it discusses the dependency of the IACC on the inter-channel cross-correlation coefficient (ICCC). The prediction is applied to laterally shifted listening positions in Sec. 6 with both the phase-based and amplitude-

based method, using the IACC (ISO, 2009) measured with a dummy head. Hereby, the question can be dealt with, how good both phantom source wideners work in comparison to each other and outside the central listening position. Additionally, Sec. 6 observes in how far the algorithms introduce perceivable spectral corruption by measuring composite binaural third-octave level deviations. The deviations from the levels of the unaltered phantom source indicate possible coloration effects.

2. Review of phantom source widening filter pairs

Raised cosine filter responses are surprisingly simple to achieve. A straightforward implementation is done as a pair of discrete-time systems. In the z-transform domain, a delay by a positive time-shift $\delta(t-\mathrm{T})$ is written as $z^{-\mathrm{N}}$ (Oppenheim *et al.*, 1999), with $\mathrm{N} = \mathrm{T}\,f_s$, the adjustable time-delay T, and the sample rate f_s,

$$H_{1,2}(z) = \frac{1}{2} \pm \frac{\widehat{\phi}}{2}\left[z^{\mathrm{N}} + z^{-\mathrm{N}}\right] \tag{1}$$

and eventually yields the linear time-invariant (LTI) frequency responses of a raised cosine with a pair of signs

$$H_{1,2}(\omega) = \frac{1}{2} \pm \widehat{\phi}\,\cos(\omega \mathrm{T}). \tag{2}$$

To avoid out of phase signals, the parameter $\widehat{\phi}$ should be used within the interval $-1/2 \le \widehat{\phi} \le 1/2$. Although this filter pair would produce widened phantom sources, it is not surprising that it was not discussed a lot. Its result is perceived as severely corrupted and colored when $|\widehat{\phi}|$ is close to one half. This is largely because the frequency responses are only amplitude complementary, $H_1(\omega) + H_2(\omega) = 1$, but not power-complementary.

The so-called *Lauridsen network* (Schröder, 1958; Orban, 1970a,b) is another simple pair of comb filters aiming at creating a widened phantom source. Combining positive and negative time-shifts with different signs, the pair of Lauridsen filters

$$H_{1,2}(z) = \frac{1}{2}\left[z^{\mathrm{N}} \pm z^{-\mathrm{N}}\right] \tag{3}$$

has Fourier transforms that are obtained after inserting $z = e^{\mathrm{i}\omega/f_s}$, $\mathrm{N} = \mathrm{T}\,f_s$, and applying Euler's identities

$$H_1(\omega) = \cos(\omega \mathrm{T}), \qquad H_2(\omega) = \mathrm{i}\sin(\omega \mathrm{T}). \tag{4}$$

These filters are LTI and power-complementary,

$$|H_1(\omega)|^2 + |H_2(\omega)|^2 = 1, \tag{5}$$

due to $\cos^2 + \sin^2 = 1$. One could expect that their effect is less annoying. However, this is not the case: their effect was perceived as even more annoying. Gerzon (1992) reported it to be phasey and assumed that this is due to the phase shift of 90° between the filter responses.

Frequency-dependent amplitude variation that is time-invariant has been considered by Gerzon (1992, 1993) in order to reduce the effects of phasiness. Among various other phase-based, even time-variant methods he proposed, this is done by a frequency dependent and power-complementary amplitude variation between the channels using the parameter $-\pi/4 \le \widetilde{\varphi}(\omega) \le \pi/4$

$$\begin{aligned} H_1(\omega) &= \cos(\pi/4 + \widetilde{\varphi}(\omega)), \\ H_2(\omega) &= \sin(\pi/4 + \widetilde{\varphi}(\omega)). \end{aligned} \tag{6}$$

Gerzon noted that this part of his idea would be more complicated to implement than all-passes. The impulse responses need to be made causal and would tend to be long. In his opinion, the curve $\widetilde{\varphi}(\omega)$ should describe a to-and-fro contour on a Bark frequency scale.

Frequency-varying Vector-Base Amplitude Panning (Pulkki, 1997; Laitinen *et al.*, 2012) on a pair of loudspeakers located at $\pm\alpha$ can be expressed by Eq. (6) using energy-normalized panning (*tangent law*, Pulkki (1997)). The angle $\widetilde{\varphi}(\omega)$ is dependent on the panning angle that lies between the loudspeakers $-\alpha \le \varphi(\omega) \le \alpha$

$$\widetilde{\varphi}(\omega) = \arctan\left\{\frac{\tan[\varphi(\omega)]}{\tan\alpha}\right\}. \tag{7}$$

Laitinen *et al.* (2012) proposed to employ a triangular panning curve over frequency, which lies between $-\widehat{\phi} \le \varphi(\omega) \le \widehat{\phi}$ and has a random starting parameter φ_0 for low frequencies,

$$\varphi(\omega) = \widehat{\phi}\left[1 - \frac{2}{\pi}\arccos[\cos(\omega \mathrm{T} + \varphi_0)]\right]. \tag{8}$$

Normally implementations of the resulting frequency responses are done by employing block convolution, which allows full freedom of adjustment, considering the width parameter $\widehat{\phi}$, the loudspeaker angles $\pm\alpha$, and the parameter for the zero frequency φ_0. If only two loudspeakers are used, the curve $\varphi(\omega)$ is limited to lie between $-\alpha \le \varphi(\omega) \le \alpha$. However, if more than one loudspeaker pair is available, the curve $\varphi(\omega)$ may run throughout a single loudspeaker pair. Nevertheless, the implementation of such elaborated real-valued, all-positive linear-phase filters requires a certain level of computational complexity.

All-pass approaches with frequency-dependent phase or time-delays were proposed by various authors (Orban, 1970a,b; Gerzon, 1992, 1993; Kendall, 1995; Potard, Burnett, 2004; Bouéri,

Kyriakakis, 2004; Potard, 2006) who applied IIR all-passes or random-phase Fourier-based FIR filters. A recent work (Zotter *et al.*, 2011) investigated a sinusoidal-phase all-pass with a sparse and efficient FIR implementation in the time-domain. Hereby, a static cosine contour of the inter-channel time-delay with regard to frequency was achieved, which reliably and efficiently produced widened phantom sources. Due to the convincing results of the listening test, this method is a reference here, and it will be simplified to a reasonable approximation in the following section.

3. Phase-based: efficient sinusoidal phase all-pass pair

The widening method in Zotter *et al.* (2011) inserts a cosine-modulated inter-channel time-delay with the periodic modulation interval of $\Delta f = 1/\mathrm{T}$ over frequency and a peak magnitude of $2\widehat{\tau}$ between the pair of channels. Despite the algorithm applied the time-delay modulation depth $\widehat{\tau}$ as a parameter, a closer inspection of the listening test results in Sec. 5 reveals that rather the phase modulation depth $\widehat{\phi} = \widehat{\tau}/\mathrm{T}$ is a relevant parameter. In particular, the width indicated by the listeners and the IACC was related to the product $\widehat{\tau}\,\Delta f = \widehat{\tau}/\mathrm{T} = \widehat{\phi}$, see also Table 1. Therefore $\widehat{\phi}$ is favored as parameter here, also because it appears to control the ICCC independently of the time-delay T.

Table 1. Parameters for both phase-based and amplitude-based widening, resulting identical ICCC and $\mathrm{IACC_{E3}}$ values for the central listening position.

$\widehat{\phi}$	$\Delta\phi_{\max} = 2\widehat{\phi}$	$\Delta \mathrm{L}_{\max}$	ICCC	P:$\mathrm{IACC_{E3}}$	A:$\mathrm{IACC_{E3}}$
0.00	0.0°	0.0 dB	1.0	0.88	0.88
0.31	35.6°	5.8 dB	0.9	0.78	0.78
0.45	52.0°	9.3 dB	0.8	0.68	0.67
0.57	66.3°	13.6 dB	0.7	0.58	0.57
0.66	77.2°	19.0 dB	0.6	0.50	0.48

The FIR responses from Zotter *et al.* (2011) use cylindrical Bessel functions and are re-written in the z-transform with $J_{-l}(\widehat{\phi}) = (-1)^l\, J_l(\widehat{\phi})$ and $\mathrm{N} = \mathrm{T}\, f_s$,

$$H_{1,2}(z) = \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} J_l(\widehat{\phi}) \Big[z^{\pm l\mathrm{N}} + (-1)^l z^{\mp l\mathrm{N}} (1 - \delta_{l,0}) \Big]. \tag{9}$$

The Kronecker delta $\delta_{l,0}$ was inserted to avoid a factor of two for $l = 0$; it is zero for $l \neq 0$ and 1 otherwise. The Fourier transform of the equation was demonstrated in Zotter *et al.* (2011) and simply is

$$H_{1,2}(\omega) = \frac{1}{\sqrt{2}} e^{\pm \mathrm{i}\widehat{\phi}\, \sin(\omega \mathrm{T})}. \tag{10}$$

The impulse responses are sparse and real-valued, and with small arguments $\widehat{\phi} \leq \pi/2$, the infinite sum can be easily truncated. According to the appended series expansion Eq. (23), the filter pair is simplified with reasonable accuracy by the equation

$$\sqrt{2}\, H_{1,2}(z) = g_0 \pm g_1 \left[z^{\mathrm{N}} - z^{-\mathrm{N}} \right] + g_2 \left[z^{2\mathrm{N}} + z^{-2\mathrm{N}} \right] \tag{11}$$

with the weights

$$g_0 = 1 - \frac{\widehat{\phi}^2}{4}, \qquad g_1 = \frac{\widehat{\phi}}{2} - \frac{\widehat{\phi}^3}{16}, \qquad g_2 = \frac{\widehat{\phi}^2}{8}. \tag{12}$$

Figure 1 shows the signal flow graph for causal phase-based phantom source widening applied on a mono signal, and its frequency responses are depicted in Fig. 2. The power $|H_1|^2 + |H_2|^2$ of the simplified filter pair deviates from unity by less than 0.1 dB, which

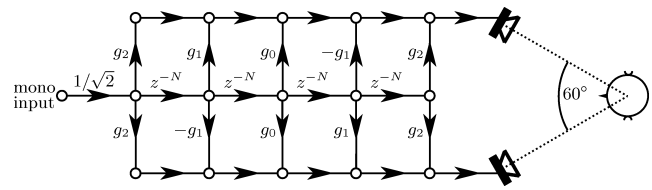


Fig. 1. Signal flow graph of a simplified and causal phase-based phantom source widener with extremely low computational demands. Suitable weights g_l are specified in Eq. (12) and $\mathrm{N} = \mathrm{T}\, f_s$.

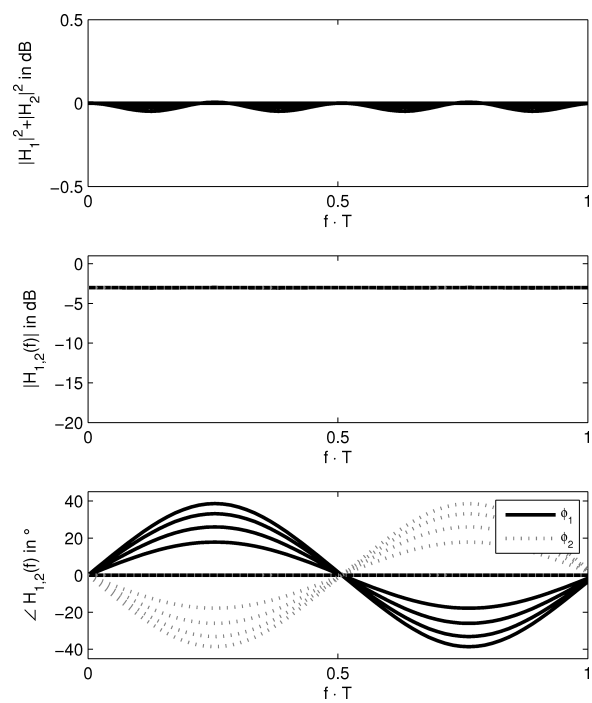


Fig. 2. The nearly constant power $|H_1|^2 + |H_2|^2$ and magnitudes $|H_{1,2}|$ in dB are shown, and the sinusoidally varying phase responses $\angle H_{1,2}$ of the parametric phase-based widening filters of Eqs. (11) and (12) for all $\widehat{\phi}$ settings in Table 1. The frequency period $1/\mathrm{T}$ is adjusted by $\mathrm{T} = \mathrm{N}/f_s$.

is considered to be sufficiently small (Karjalainen *et al.*, 1999). The frequency period of the sinusoidal phase is adjustable by the time-delay as 1/T.

4. Amplitude-based: efficient cosine of raised cosine pair

A simple way to obtain a power-complementary zero-phase filter for widening would be to apply the square root to the raised cosine filter Eq. (1), see appended Eq. (24). A similar power-complementary response pair that is easier to approximate is obtained by insertion of $\widetilde{\varphi} = \widehat{\phi}\,\cos(\omega \mathrm{T})$ into Eq. (6), which yields

$$H_{1,2}(\omega) = \cos\Big[\pi/4 \pm \widehat{\phi}\,\cos(\omega \mathrm{T})\Big]\,. \tag{13}$$

This response pair is related to the phase-based one of Eq. (10) as formulated in the appended Eqs. (25) and (26), and it is therefore surprisingly simple. As Eq. (9), it also involves the Bessel functions as coefficients

$$\sqrt{2}\,H_{1,2}(z) \;=\; \sum_{l=0}^{\infty} J_l(\widehat{\phi})\,\sqrt{2}\cos\Big(\frac{\pi}{4} \pm \frac{\pi}{2}\,l\Big) \cdot \big[z^{l\mathrm{N}} + z^{-l\mathrm{N}}(1-\delta_{l,0})\big]\,, \tag{14}$$

but now two sign alteration patterns

$$\langle +,-,-,+,+,-,-,\ldots\rangle \quad \text{and} \quad \langle +,+,-,-,+,+,\ldots\rangle$$

expressed by $\sqrt{2}\cos\left(\frac{\pi}{4} \pm \frac{\pi}{2}\,l\right)$ equally determine the signs of the symmetric delays $z^{\pm l\,\mathrm{N}}$. Because this is the only difference, the infinite sum is truncated as in the section above, and the coefficients are approximated with Eq. (12),

$$\sqrt{2}\;H_{1,2}(z) \;=\; g_0 \mp g_1\,\big[z^{\mathrm{N}} + z^{-\mathrm{N}}\big] - g_2\,\big[z^{2\mathrm{N}} + z^{-2\mathrm{N}}\big]\,. \tag{15}$$

Note that also the filter structure is exactly the same as for the phase-based method, and it even creates the same decorrelation, see appended Eq. (28).

Figure 3 shows a causal signal flow graph for the new amplitude-based method of phantom source widening applied on a mono signal. The frequency responses are depicted in Fig. 4 and exhibit the deviation from a power-complementary response by less than 0.1 dB, as before. There is a curve for each $\widehat{\phi}$ value of Table 1.

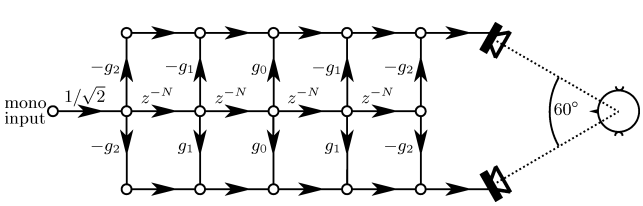


Fig. 3. Signal flow graph of an efficient and causal amplitude-based phantom source widener. Suitable weights g_l are specified in Eq. (12) and $\mathrm{N} = \mathrm{T}\,f_s$.

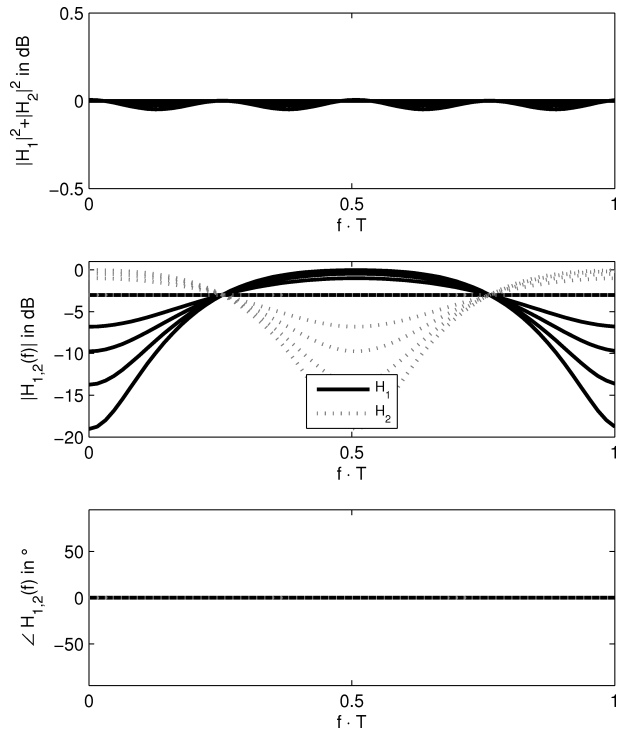


Fig. 4. The nearly constant power $|H_1|^2 + |H_2|^2$, the cosine of raised cosine magnitude responses $|H_{1,2}|$ in dB, and the constant phase responses $\angle H_{1,2}$ of the parametric amplitude-based widening filters of Fig. 3 for all parameters $\widehat{\phi}$ in Table 1. The frequency period 1/T is adjusted by $\mathrm{T} = \mathrm{N}/f_s$.

5. Evaluation at the central listening position

This section discusses the effect of the presented efficient widening algorithms in terms of known experimental data. These data are well predicted by the IACC, which is affected by the ICCC of the two playback channels. Both technical measures are introduced and applied to compare both widening effects.

5.1. Experimental results from a previous listening test

The phase-based approach was tested with a listening experiment in a preceding article (Zotter *et al.*, 2011). The analysis presented here uses experimental data thereof that was obtained for pink noise processed with the phase-based widening structure using the parameters[1] given in Table 2.

[1] In the print version of Zotter *et al.* (2011) the parameters $\widehat{\phi}$ and T were expressed equivalently as $\widehat{\tau} = \widehat{\phi}\,\mathrm{T}$ and $\Delta f = 1/\mathrm{T}$. However, we listed erroneous values of $\widehat{\tau}$ in contrast to those actually employed in the experiment. Retrieval of T and $\widehat{\phi}$ from the original stimuli pairs $X_{1,2}$, i.e. the phase of $\dfrac{X_1(\omega)\,X_2^*(\omega)}{\|X_1(\omega)\|\,\|X_2(\omega)\|}$, clearly revealed that exactly half the stated value was employed. The erratum was forwarded to the DAFx-11 editors and the online publication includes a dated erratum remark.

Table 2. Listening test conditions of the phase-based approach that was evaluated in a former article (Zotter *et al.*, 2011). For the present article only the conditions C2, C4, C6, C7, and C8 are interesting. The $(\widehat{\tau}, \Delta f)$ parameters from the previous nomenclature were re-expressed using $\widehat{\phi} = \widehat{\tau}/\Delta f$ (mind errata in (Zotter *et al.*, 2011)) and $\mathrm{T} = 1/\Delta f$.

	$\widehat{\phi}$	T	ICCC($\widehat{\phi}$)	$\mathrm{IACC_{E3}}$
C2	0.0	–	1.00	0.824
C4	0.3	5.0 ms	0.91	0.690
C6	0.6	2.5 ms	0.67	0.554
C7	0.6	5.0 ms	0.67	0.521
C8	0.9	1.7 ms	0.34	0.468

Experimental setup. The experiment was set up in a 11 m×11 m×5 m room with a mean reverberation time of $\mathrm{RT_{60}} = 470$ ms, using 2 Genelec 8020 loudspeakers at ±30°. The listening distance was 2 m, which is within the effective critical distance of the room. Originally, 8 conditions were compared in a full pairwise comparison that was performed twice within each stimulus set: one generated from 5 s pink noise, the other from 22 s male speech. Participants were familiarized with the noise stimuli that were leveled to 65 dB(A), took about 15 min to finish their 56 comparisons tasks (different random order for each participant) with noise, took an intermission, and then the entire procedure was repeated from the start but with speech stimuli. The subjects responded by either selecting the sound in each pair that was perceived as wider or stating that both were equal, using three marked keys of a keyboard on their lap. Seamless switching within the comparison pair while listening was possible at any time. All 11 participants were members of a trained listening panel and had participated in an experiment about source width before (Frank *et al.*, 2011). Below, only results for the noise stimuli using the phase-based FIR algorithm with settings of $\mathrm{T} = \{1.7,\ 2.5,\ 5\}$ ms were used, as those for speech are similar anyway. The condition C5 with $\mathrm{T} = 0.8$ ms was excluded as it caused other effects as well, C1 (mono) and C3 (IIR widening) are excluded as they do not match the topic of the present article.

Results. The individual repetitions were averaged, and the Thurstone scales (Thurstone, 1994) were recalculated based on the full pairwise comparison response matrices for the selected conditions of Table 2. The scales were calculated for each participant's individual responses to estimate inter-subjective confidence intervals, and one overall scale was built upon a response matrix pooled from all participants as a consistency check. Figure 5 shows the significant effect of $\widehat{\phi}$ on the perceived source width scale. Conditions C6 and C7 only differ in the amount of delay T and yield significantly equal results.

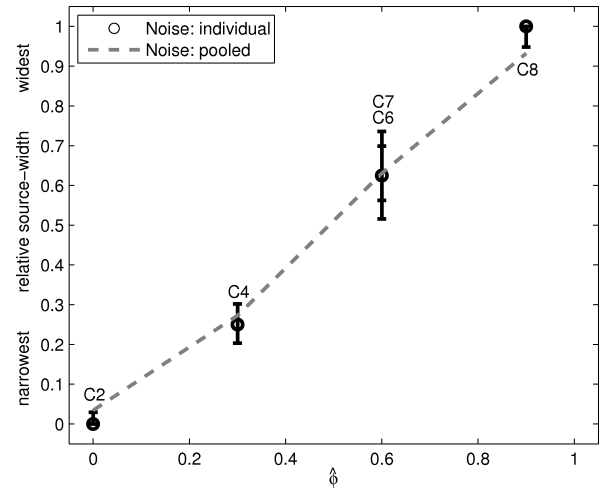


Fig. 5. Thurstone scales for noise conditions C2, C4, C6, C7, and C8 (Zotter *et al.*, 2011): median and inter-subjective 95% confidence interval using participants' individual scales, compared to the over-all scale using answers pooled from all participants. The scales are plotted as a function of $\widehat{\phi}$.

5.2. Technical prediction

Inter-Aural Cross-Correlation Coefficient (IACC). Originally conceived to describe the apparent source width in room acoustics, the IACC also suitably describes the width change of the phantom source after the application of widening. It is defined as the maximum of the inter-aural cross-correlation function (IACF), cf. ISO (2009),

$$\mathrm{IACF}(\tau) = \frac{\int_{t_1}^{t_2} x_{\mathrm{left}}(t)\, x_{\mathrm{right}}(t+\tau)\,\mathrm{d}t}{\sqrt{\left[\int_{t_1}^{t_2} x_{\mathrm{left}}^2(t)\,\mathrm{d}t\right]\left[\int_{t_1}^{t_2} x_{\mathrm{right}}^2(t)\,\mathrm{d}t\right]}}, \tag{16}$$

$$\mathrm{IACC} = \max_{|\tau|\le 1\,\mathrm{ms}} |\mathrm{IACF}(\tau)|, \tag{17}$$

which is calculated from the binaural signal pair $x_{\mathrm{left}}(t)$, $x_{\mathrm{right}}(t)$. In our evaluation, the signal pair was replaced by the binaural impulse response pair due to an impulse at the system input, and the early part ($t_1 = 0$ ms and $t_2 = 80$ ms) is of particular interest, averaged over the octave bands 500 Hz, 1 kHz, and 2 kHz as defined in Hidaka *et al.* (1995), denoted as $\mathrm{IACC_{E3}}$. The $\mathrm{IACC_{E3}}$ exhibits a 97 . . . 98% statistical correlation to the perceived source width across all conditions of the experiment (speech, noise, other filtering structure). Therefore, we assume that the phantom source width can be estimated by measuring the $\mathrm{IACC_{E3}}$.

In order to document the ear signals of the experiment, they were recorded using a B&K 4128C Head-and-Torso simulator at the listening position, and binaural loudspeaker impulse responses were measured.

These, $h_{1,\mathrm{left}}(t)$, $h_{2,\mathrm{left}}(t)$, $h_{1,\mathrm{right}}(t)$, $h_{2,\mathrm{right}}(t)$, can be used to calculate ear signals due to any pair of loudspeaker signals $s_{1,2}(t)$ by convolution

$$x_{\mathrm{left}}(t) = h_{1,\mathrm{left}}(t) \star s_1(t) + h_{2,\mathrm{left}}(t) \star s_2(t),$$

and

$$x_{\mathrm{right}}(t) = h_{1,\mathrm{right}}(t) \star s_1(t) + h_{2,\mathrm{right}}(t) \star s_2(t).$$

Inter-Channel Cross-Correlation Coefficient (ICCC). The phantom source widening algorithms above directly control the correlation between the loudspeaker signals $s_{1,2}(t)$. The ICCC is the maximum of the inter-channel cross-correlation function (ICCF)

$$\mathrm{ICCF}(\tau) = \frac{\int_{-\infty}^{\infty} s_1(t) s_2(t+\tau)\,\mathrm{d}t}{\sqrt{\left[\int_{-\infty}^{\infty} s_1^2(t)\,\mathrm{d}t\right]\left[\int_{-\infty}^{\infty} s_2^2(t)\,\mathrm{d}t\right]}}, \tag{18}$$

$$\mathrm{ICCC} = \max_{|\tau|\le \mathrm{P}} |\mathrm{ICCF}(\tau)|. \tag{19}$$

To obtain a generalized estimation, the signal pair $s_{1,2}(t)$ is replaced by the impulse response pair $h_{1,2}(t)$ of phantom source widening. In the tested cases $0 \le \widehat{\phi} < \pi/2$, this yields for the period $\mathrm{P} > 0$ according to the appended Eq. (28)

$$\mathrm{ICCC} = J_0(2\widehat{\phi}), \tag{20}$$

which approximately equals $\cos(\widehat{\phi}\sqrt{2})$. Table 2 shows the ICCC of Eq. (20) for the test conditions: The bigger the modulation depth parameter $\widehat{\phi}$ of the phantom source widener the smaller the ICCC.

Relation between IACC and ICCC at the central listening position. Judging from the values in Table 2, not only the IACC but also the ICCC seems to be an excellent predictor for the perceived width. Nevertheless, the ICCC only depends on the algorithmic parameter $\widehat{\phi}$. Therefore, it is unable to estimate changes under more general circumstances, under which room and setup will have a natural impact on the apparent source width. However, in the acoustic situation of the experiment, theI CCC indirectly maps to the $\mathrm{IACC_{E3}}$ in a monotonic curve, see solid curve in Fig. 6. Linear regression (dashed curve) yields

$$\mathrm{IACC_{E3}} \approx 0.77 \cdot \mathrm{ICCC} + 0.03. \tag{21}$$

The relationship given in Fig. 6 is still valid if, instead of the phantom-source widening algorithms, variably correlated white noise signals are convolved with the binaural loudspeaker responses (dash-dot curve). This was also verified for another room (Sec. 6, Fig. 7a) with dominant direct sound (dotted curve). Noise signals of

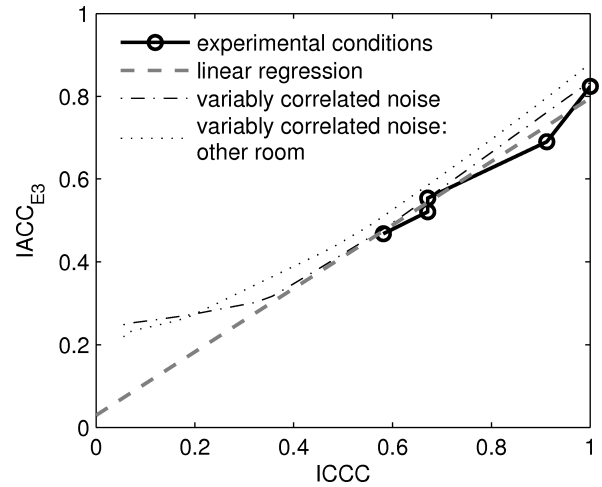


Fig. 6. The inter-aural cross-correlation coefficient (IACC) monotonically depends on the inter-channel cross-correlation coefficient (ICCC) in the experiment (Zotter *et al.*, 2011). The black curve is the relation for phase-based phantom source widening, the dash-dot curve for a variably correlated, white stereo noise. The dashed line is a linear regression, and the dotted line expresses the same tendency for stereo noise in a different room.

adjustable ICCC and equal variance are created from uncorrelated noise signals $a(t)$, $b(t)$ of equal variance, see appendix, cf. Blauert, Lindemann (1986), by

$$\begin{aligned} s_1(t) &= a(t), \\ s_2(t) &= \mathrm{ICCC}\, a(t) + \sqrt{1-\mathrm{ICCC}^2}\, b(t). \end{aligned} \tag{22}$$

Phantom source width controllability. At the central listening position, the $\mathrm{IACC_{E3}}$ directly increases with the ICCC. Their linear relationship has been shown before. According to measurements done with the setup described in the next section (Fig. 7a), this is true for both amplitude- and phase-based widening. The corresponding $\mathrm{IACC_{E3}}$ values were measured to differ by 0.02 at most (rightmost columns in Table 1). Therefore amplitude and phase modification is expected to achieve similar widening in listening experiments; an expectation that goes along with our listening experience. For controlling the phantom source width, we ideally wish to achieve at least a just noticeable difference (JND) in the $\mathrm{IACC_{E3}}$ when increasing the ICCC by 0.1. As the JND for the $\mathrm{IACC_{E3}}$ is roughly 0.05 (Blau, 2002), this means that we desire a slope of at least $\dfrac{\Delta\mathrm{IACC_{E3}}}{\Delta\mathrm{ICCC}} \ge 0.5$.

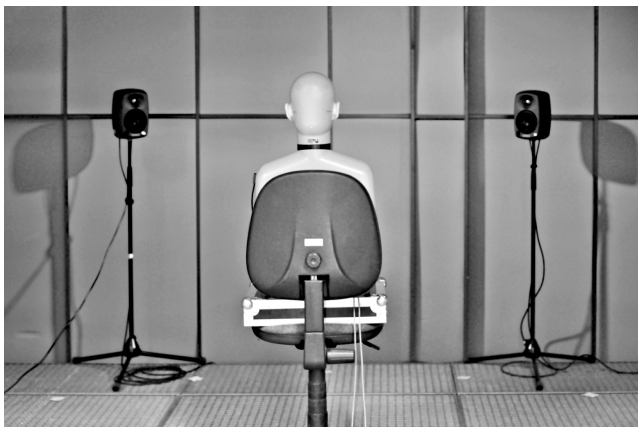
6. Evaluation at off-center listening positions

For the purpose of further evaluation, a series of binaural measurements have been taken in a different room and at various off-center positions. In particular, the measurement data shall clarify whether the

amplitude-based widening method is more robust to lateral shifts.

Measurement setup. A pair of Genelec 8020 loudspeakers has been set up in a standard ±30° stereo arrangement with 1.8 m distance to the central listening position in a 3.7 m×3.7 m×3.2 m room with a reverberation time of RT_{30} = 50 ms. The binaural impulse responses of the loudspeakers have been taken with the B&K dummy head. The responses of the filters in Figs. 1 and 3 convolved with binaural responses allow to predict the binaural signals for several settings of the algorithms. The series of binaural impulse responses included shifts of the dummy head to lateral positions by a shift d ranging from −90 cm to 90 cm in 10 cm steps, e.g., for $d = 0$ cm in Fig. 7a and $d = 90$ cm in Fig. 7b.

a)



b)

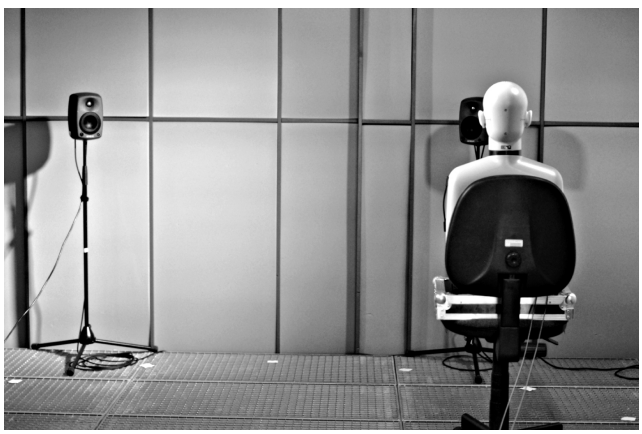


Fig. 7. The experimental setup for evaluating the effect of phantom source widening on IACC employs a dummy head in an acoustically damped room. The head is also employed at side-wise shifted locations $d = \{-90, -80, \ldots, +80, +90\}$ cm to gather information about off-center listening locations: a) central position, b) 90 cm shifted position.

Results for IACC at different shifts and ICCCs. Figure 8 shows which $IACC_{E3}$ values are accessible at different shifts, playing back correlated and correlated noise, i.e. ICCC = $\{0, 1\}$. This implies that any attempt to widen the phantom source could be ineffective at large shifts $d \geq 40$ cm.

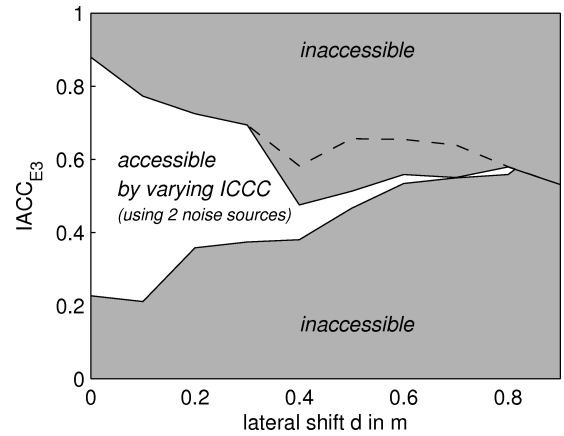
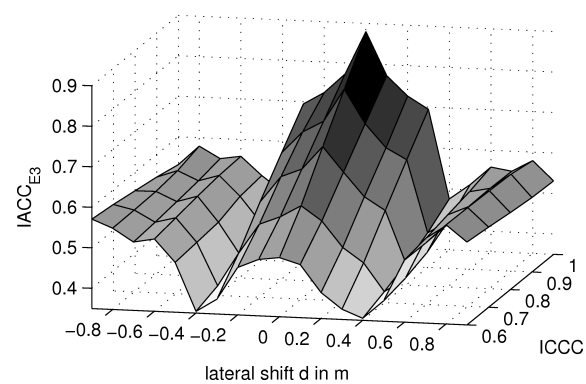


Fig. 8. $IACC_{E3}$ values that can be created at different lateral shifts 0 cm $\leq d \leq$ 90 cm by controlling the ICCC between 0 and 1. (The dashed curve is for a lag within $|\tau| \leq 2$ ms or more in Eq. (17).)

Figure 9 shows the particular dependency of the $IACC_{E3}$ on the ICCC for all measured positions and both algorithms using a time-delay of T = 5 ms. Table 1 summarizes the settings of both phantom source wideners to produce an ICCC of values between 0.6

a)



b)

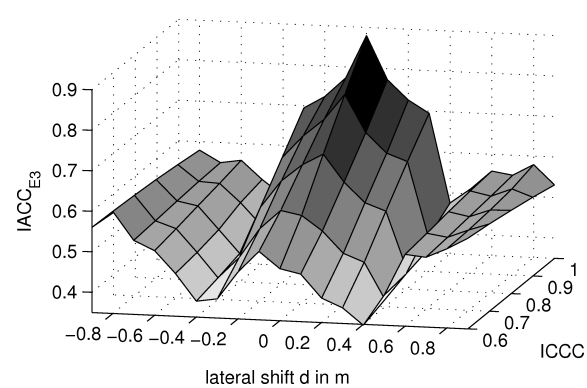
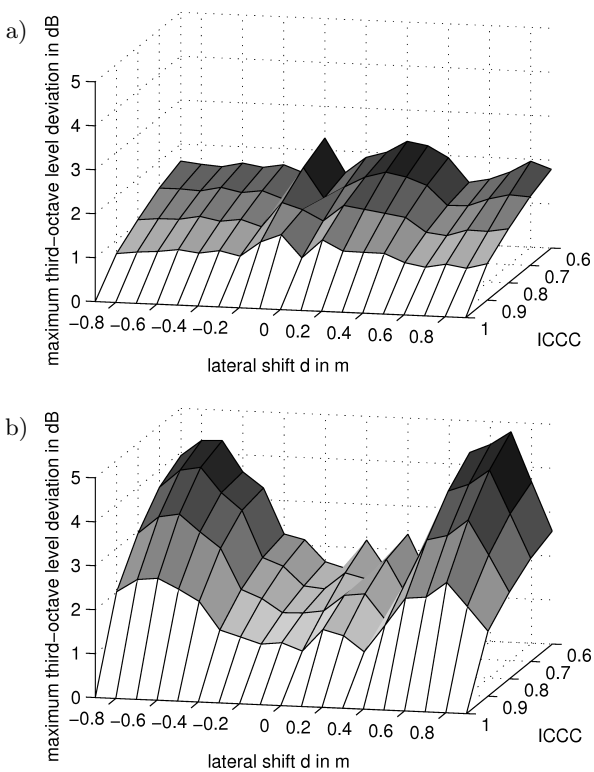


Fig. 9. $IACC_{E3}$ measured for phantom source widening using lateral listening position shifts $d = \{-90, -80, \ldots, +80, +90\}$ cm. The ICCC is variable from 0.6 to 1 in 0.1 steps and should control the $IACC_{E3}$: a) phase-based, T = 5 ms; b) amplitude-based, T = 5 ms.

and 1, in steps of 0.1. The ICCC value is independent of the particular time-delay value T. For the central listening position $d = 0$ cm, the diagram shows the linear controllability already observed above. The desired slope of $\frac{\Delta \mathrm{IACC_{E3}}}{\Delta \mathrm{ICCC}} \geq 0.5$ for controllability is achieved for all $|d| < 40$ cm. Thus the phantom source widening is expected to work within this range of listening positions. This is true for all investigated settings and also various time-delays T = {1.6, 2.5, 5} ms of both filter structures. However, for $|d| \geq 40$ cm the $\mathrm{IACC_{E3}}$ is obviously not controlled by the ICCC; for both algorithms. As the $\mathrm{IACC_{E3}}$ is related to the perceived source width, the perceived width at the farthest lateral listening positions seems to be indifferent to ICCC, as in Fig. 8. Nevertheless, different ICCCs seem to produce other perceivable differences according to informal listening experience.

Coloration: Maximum third-octave level difference. In order to observe whether the algorithms introduce perceivable spectral coloration when decreasing the ICCC, i.e. decorrelation, binaural third-octave levels are also measured. Levels are determined after summation of the third-octave powers across the both ear signals (Ono *et al.*, 2001), which was done for each band from 200 Hz to 12.5 kHz, here. The differences of the resulting third-octave levels to those of the unaltered phantom source should stay small. According to Karjalainen *et al.* (1999), a third-octave level deviation of 1 dB is audible in loudspeaker equalization.

The maximum of the third-octave level deviations is plotted in Figs. 10a to 10f for varying ICCC



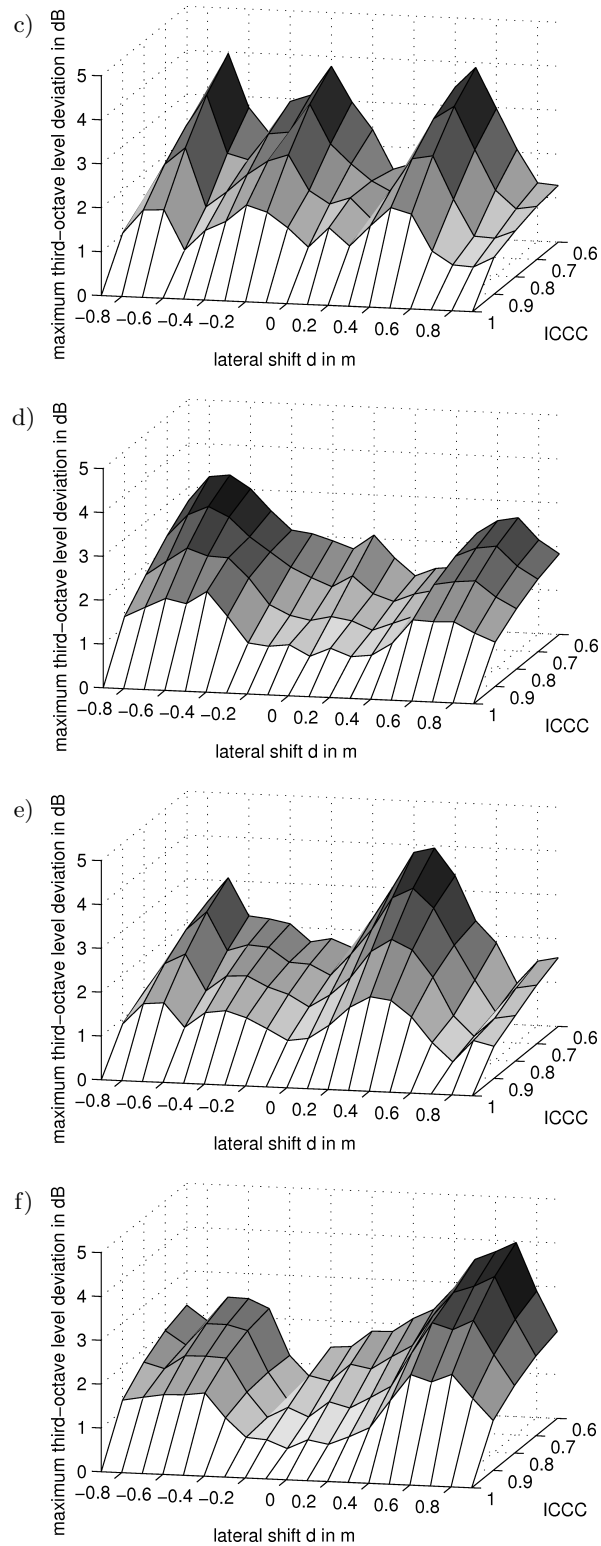


Fig. 10. Maximum third-octave level deviation with regard to normal phantom source (ICCC=1) for (a), (c), (e) phase-based widening and (b), (d), (f) amplitude-based widening using T = {1.6, 2.5, 5} ms. Listening positions vary from −90 cm to 90 cm in 10 cm steps and the adjusted ICCC from 0.6 to 1 in 0.1 steps: a) phase-based, T = 1.6 ms, b) amplitude-based, T = 1.6 ms, c) phase-based, T = 2.5 ms, d) amplitude-based, T = 2.5 ms, e) phase-based, T = 5 ms, f) amplitude-based, T = 5 ms.

and lateral shifts d, for both algorithms and T = {1.6, 2.5, 5} ms. Clearly, the maximum deviation as a measure of sound coloration increases with the amount of decorrelation produced between the channels. Coloration also depends on the listening position. For small time-delays, the coloration of the phase-based approach is quite low, but there might be some phasiness perceived. For larger time-delay settings, both approaches become similar.

7. Conclusion

We presented a highly efficient filter structure for phantom source widening based on a frequency-dependent modification of either phase or amplitude. Both ways of controlling the structure are power-complementary at a high accuracy, yield the same inter-channel cross-correlation coefficient (ICCC), and use the same number of operations (4 delays of equal size, 8 additions, 10 multiplications depending on 3 variable weights). The algorithmic parameter $\widehat{\phi}$ allows to adjust the correlation of a mono signal used in stereo to a desired degree, with the relation ICCC $\approx \cos(\widehat{\phi}\sqrt{2})$. Thinkable target applications are, e.g., stereo effect processors in audio mixing, the research of spatial hearing, and auditory interfaces in which the salience of sounds shall be controlled by spatial sharpness or width.

We took dummy head measurement at the central listening position to show the linear proportionality of the inter-channel coherence to the inter-aural coherence measured by the IACC, whose relation to the perceived phantom source width is known from former experimentation. Moreover, we could show that the controllability of the phantom source width has to do with the listening position. It strongly influences the slope in the dependency of the inter-aural on the inter-channel coherence.

By measurements at laterally shifted positions, we showed the ability of both amplitude and phase decorrelation to control the phantom source width for displaced listening positions within $|d| \le 30$ cm. This is largely independent of the size of the algorithmic time-delay within the interval of 1.6 . . . 5 ms. Neither of the algorithms controls the phantom source width outside the range of lateral shifts, according to the inter-aural coherence.

Both amplitude and phase decorrelation seem to cause sound coloration that increases with the parameter $\widehat{\phi}$. The corresponding third-octave level differences to an unaltered phantom source are maximally 5 dB in all bands between 200 Hz and 12.5 kHz. This dependency is observed for all measured listening positions albeit the coloration of both algorithms behaves differently in its detail. For small time-delays the phase-based method seems to be better.

Appendix

Sinusoidal-phase all-pass. The series expansion of the cylindrical Bessel function inserted into Eq. (9) yields

$$\begin{aligned} H_1(z) &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} J_l(\widehat{\phi}) \left[z^{l\mathrm{N}} + (-1)^l z^{-l\mathrm{N}} (1-\delta_{l,0}) \right] \\ &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} \sum_{n=|l|}^{\infty} \widehat{\phi}^n \frac{(-1)^{\frac{n}{2}-\frac{l}{2}}\, \delta_{n \bmod 2, l \bmod 2}}{\left(\frac{n}{2}-\frac{l}{2}\right)! \left(\frac{n}{2}+\frac{l}{2}\right)!\, 2^n} \\ &\quad \cdot \left[z^{l\mathrm{N}} + (-1)^l z^{-l\mathrm{N}} (1-\delta_{l,0}) \right] \\ &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} g_l \left[z^{l\mathrm{N}} + (-1)^l z^{-l\mathrm{N}} (1-\delta_{l,0}) \right]. \end{aligned} \tag{23}$$

In the range $0 \le \widehat{\phi} \le \pi/4$, a reasonable approximation with $-0.1 \le |H_{1,2}(\omega)| \le 0.02$ dB is obtained for $|l| \le 2$ and $n \le 3$ with the weights in Eq. (12).

Root raised cosine. In order to achieve the square root response of the raised cosine filter pair in Eq. (2), the binomial formula

$$(x+y)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^{n-k} y^k$$

is used to represent

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \frac{(-1)^n (2n-1)!!}{(1-2n) n! 2^n} x^n.$$

Consequently, the application of the square-root on the z-domain expression in Eq. (1) can be re-expressed after applying the binomial formula and re-arranging the sums

$$\begin{aligned} H_1(z) &= \sqrt{1+\widehat{\phi}\,(z^{\mathrm{N}}+z^{-\mathrm{N}})}/\sqrt{2} \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \frac{(-1)^n (2n-1)!!}{(1-2n)\, n!\, 2^n} \left[\widehat{\phi}\,(z^{\mathrm{N}}+z^{-\mathrm{N}}) \right]^n \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \widehat{\phi}^n \frac{(-1)^n (2n-1)!!}{(1-2n)\, n!\, 2^n} \\ &\quad \cdot \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, z^{(2k-n)\mathrm{N}} \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \sum_{l=-n}^{n} \widehat{\phi}^n \frac{(-1)^n (2n-1)!!}{(1-2n)\, 2^n \left(\frac{n}{2}+\frac{l}{2}\right)! \left(\frac{n}{2}-\frac{l}{2}\right)!} \\ &\quad \cdot z^{l\mathrm{N}}\, \delta_{n \bmod 2, l \bmod 2} \\ &= \frac{1}{\sqrt{2}} \sum_{l=-\infty}^{\infty} \left[\sum_{n=|l|}^{\infty} \widehat{\phi}^n \right. \\ &\quad \cdot \frac{(-1)^n (2n-1)!!}{(1-2n)\, 2^n \left(\frac{n}{2}+\frac{l}{2}\right)! \left(\frac{n}{2}-\frac{l}{2}\right)!} \\ &\quad \left. \cdot\, \delta_{n \bmod 2, l \bmod 2} \right] z^{l\mathrm{N}}. \end{aligned} \tag{24}$$

The Kronecker delta $\delta_{n \bmod 2, l \bmod 2}$ was introduced in line three to ensure that the factorials stay integer as in the lines before. This is only possible if n and l are both exclusively either even or odd. A simple approximation is obtained by truncating the sum over l to $|l| \le 2$ and the sum over n to $n \le 2$ and the weights $\left[1 - \frac{\widehat{\phi}^2}{4}, \frac{\widehat{\phi}}{2} - \frac{3\,\widehat{\phi}^3}{16}, -\frac{\widehat{\phi}^2}{8}\right]$.

Cosine of raised cosine pair. The response

$$\cos\left[\pi/4 \pm \widehat{\phi}\,\cos(\omega \mathrm{T})\right]$$

of Eq. (13) can be re-written as

$$H_{1,2}(\omega) = \Big[e^{\mathrm{i}\pi/4}\, e^{\mathrm{i}\sin(\omega\mathrm{T}\pm\pi/2)} + e^{-\mathrm{i}\pi/4}\, e^{-\mathrm{i}\sin(\omega\mathrm{T}\mp\pi/2)}\Big]\Big/2, \tag{25}$$

and thus it is related to Eq. (10) by

$$H_{1,2}(\omega) = \Big\{e^{\mathrm{i}\pi/4}\,\widetilde{H}_1[\omega \pm \pi/(2\mathrm{T})] + e^{-\mathrm{i}\pi/4}\,\widetilde{H}_2[\omega \mp \pi/(2\mathrm{T})]\Big\}/\sqrt{2}.$$

Applying the corresponding frequency shift of $\pm\pi/(2\mathrm{N})$, the $\pi/4$ phase shift, and Euler's formula $\cos\alpha = (e^{\mathrm{i}\alpha} + e^{-\mathrm{i}\alpha})/2$ on $H(z) = \sum_{l=-\infty}^{\infty} J_l(\widehat{\phi})\, z^{-l\mathrm{N}}$, cf. Eq. (9), yields

$$\begin{aligned} H_{1,2}(z) &= \frac{1}{2}\sum_{l=-\infty}^{\infty} J_l(\widehat{\phi})\, z^{-l\mathrm{N}}\left[e^{\mathrm{i}\pi/4}(\pm\mathrm{i})^l + e^{-\mathrm{i}\pi/4}(\mp\mathrm{i})^l\right] \\ &= \sum_{l=-\infty}^{\infty} J_l(\widehat{\phi}) z^{-l\mathrm{N}} \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2}\,l\right). \end{aligned} \tag{26}$$

Using $J_{-l}(\widehat{\phi}) = (-1)^l J_l(\widehat{\phi})$ and $(-1)^l \cos\left(\frac{\pi}{4} \mp \frac{\pi}{2}\,l\right) = \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2}\,l\right)$, we find that the involved cosine expresses a factor $1/\sqrt{2}$ and the alternating sign sequence $\langle 1, \mp 1, -1, \pm 1\rangle$ for $\pm l = 0, 1, 2, 3$. The response pair is therefore symmetrical with regard to time, which is consistent with Eq. (13) being zero-phase. Simplified by multiplication with $\sqrt{2}$, we get

$$\begin{aligned} \sqrt{2}\,H_{1,2}(z) \;=\; & J_0(\widehat{\phi}) \mp J_1(\widehat{\phi})\left[z^{-\mathrm{N}} + z^{\mathrm{N}}\right] \\ & - J_2(\widehat{\phi})\left[z^{-2\mathrm{N}} + z^{2\mathrm{N}}\right] \\ & \pm J_3(\widehat{\phi})\left[z^{-3\mathrm{N}} + z^{3\mathrm{N}}\right] \\ & + J_4(\widehat{\phi})\left[z^{-4\mathrm{N}} + z^{4\mathrm{N}}\right] \mp \ldots \end{aligned} \tag{27}$$

Obviously, the same approximated Bessel functions can be used as weights as suggested two paragraphs above, but now with different signs. Due to the same mathematical origin, the hereby obtained filter pair yields the same accuracy of power-complementarity as suggested two paragraphs above.

The ICCC of both structures is obtained from the normalized frequency domain cross-correlation. The inter-channel cross-correlation function (ICCF) in the time domain becomes at a variable lag $\tau = n\mathrm{T}$, $n \in \mathbb{Z}$,

$$\begin{aligned} \mathrm{ICCF}[n\mathrm{T}] &= \frac{\int_0^{2\pi} H_1(\omega)\, H_2^*(\omega)\, e^{\mathrm{i}\, n\omega\mathrm{T}}\,\mathrm{d}(\omega\mathrm{T})}{\sqrt{\left[\int_0^{2\pi}|H_1(\omega)|^2\,\mathrm{d}(\omega\mathrm{T})\right]\left[\int_0^{2\pi}|H_2(\omega)|^2\,\mathrm{d}(\omega\mathrm{T})\right]}} \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{\mathrm{i}\,[2\widehat{\phi}\cos(\omega\,\mathrm{T}) + n\omega\mathrm{T}]}\,\mathrm{d}(\omega\mathrm{T}) \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sin\left[n\omega\mathrm{T} + \frac{\pi}{2} + 2\widehat{\phi}\cos(\omega\,\mathrm{T})\right]\mathrm{d}(\omega\mathrm{T}) \\ &= J_n(2\widehat{\phi}). \end{aligned} \tag{28}$$

Thereof, the ICCC is the maximal value within some time period P, i.e. $\mathrm{ICCC} = \max_{|n\mathrm{T}|\le P} J_n(2\widehat{\phi})$, which yields just $J_0(2\widehat{\phi})$ if $2\widehat{\phi} < \pi/2$ or P is short.

Variably correlated noise pair. Assume two zero-mean random signals a and b, $E\{a\} = E\{b\} = 0$, that are uncorrelated $E\{a\,b\} = 0$ and whose variance equals $\sigma^2 = E\{a^2\} = E\{b^2\}$; $E\{\cdot\}$ denotes the expected value. The cross-correlation of $s_1 = a$ with $s_2 = \gamma\,a + \sqrt{1-\gamma^2}\,b$ from Eq. (22) is easily shown to equal γ

$$\begin{aligned} \mathrm{ICCC} &= \frac{E\{s_1\,s_2\}}{\sqrt{E\{s_1^2\}E\{s_2^2\}}} \\ &= \frac{\gamma\,\sigma^2 + \cancel{\sqrt{1-\gamma^2}\,E\{a\,b\}}}{\sqrt{\sigma^2\left[\cancel{\sigma^2\gamma^2} + \sigma^2(1-\cancel{\gamma^2}) + \cancel{2\gamma\sqrt{1-\gamma^2}\,E\{a\,b\}}\right]}} \\ &= \gamma. \end{aligned} \tag{29}$$

Acknowledgments

We gratefully thank Alois Sontacchi and Georgios Marentakis for their contributions in our former paper on which this article is based. Our research is partly supported by the project AAP, which is funded by Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and the departments 3 and 14 of the Styrian Government. The Austrian Research Promotion Agency (FFG) conducted the funding under the Competence Centers for Excellent Technologies (COMET, K-Project), a program of the above-mentioned institutions.

A part of this work was presented at the 59th Open Seminar on Acoustics (Poland, Boszkowo, September 10–14. 2012).

References

1. Blau M. (2002), *Difference limens for measures of apparent source width*, [in:] *Forum Acusticum*, Sevilla, Spain.
2. Blauert J., Lindemann W. (1986), *Spatial mapping of intracranial auditory events for various degrees of interaural coherence*, The Journal of the Acoustical Society of America, **79**, 806–813.
3. Bouéri M., Kyriakakis C. (2004), *Audio signal decorrelation based on a critical band approach*, [in:] *Preprint 6291, 117th Conv. Audio Eng. Soc.*, San Francisco.
4. Frank M., Marentakis G., and Sontacchi A. (2011), *A simple technical measure for the perceived source width*, [in:] *Fortschritte der Akustik, DAGA*, Düsseldorf.
5. Gerzon M. A. (1992), *Signal processing for simulating realistic stereo images*, [in:] *Preprint 3423, 93rd Conv. Audio Eng. Soc.*, San Francisco.
6. Gerzon M. A. (1993), *Stereophonic signal processor generating pseudo stereo signals*, Patent, WO 93/25055.
7. Hidaka T., Beranek L. L., Okano T. (1995), *Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls*, The Journal of the Acoustical Society of America, **98**, 2, 988–1007.
8. ISO (2009), ISO 3382-1:2009: *Acoustics – measurement of room acoustic parameters – part 1: Performance spaces.*
9. Karjalainen M., Piirilä E., Järvinen A., Huopaniemi J. (1999), *Comparison of loudspeaker equalization methods based on DSP techniques*, J. Audio Eng. Soc., **47**, 1/2, 14–31.
10. Kendall G. S. (1995), *The decorrelation of audio signals and its impact on spatial imagery*, Computer Music J., **19**, 4, 71–87.
11. Kin M. J., Plaskota P. (2011), *Comparison of sound attributes of multichannel and mixed-down stereo recordings*, Archives of Acoustics, **36**, 2, 333–345.
12. Laitinen M.-V., Philajamäki T., Erkut C., Pulkki V. (2012), *Parametric time-frequency representation of spatial sound in virtual worlds*, ACM Trans. Appl. Percept., **3**, 2, 8:1–20.
13. Ono K., Pulkki V., Karjalainen M. (2001), *Binaural modeling of multiple sound source perception: Methodology and coloration experiments*, [in:] *Preprint 5446, 111th Conv. Audio Eng. Soc.*, New York.
14. Oppenheim A. V., Schafer R. W., Buck J. R. (1999), *Discrete-Time Signal Processing*, Prentice Hall, New Jersey.
15. Orban R. (1970a), *A rational technique for synthesizing pseudo-stereo from monophonic sources*, J. Audio Eng. Soc., **18**, 2, 157–164.
16. Orban R. (1970b), *Letters to the editor: Further thoughts on "a rational technique for synthesizing pseudo-stereo from monophonic sources"*, J. Audio Eng. Soc., **18**, 4, 443–444.
17. Plewa M., Kleczkowski P. (2011), *Choosing and configuring a stereo microphone technique based on localisation curves*, Archives of Acoustics, **36**, 2, 347–363.
18. Potard G., Burnett I. (2004), *Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays*, [in:] Proc. DAFx-04, 280–284, Napoli.
19. Potard G. (2006), *3D-audio object oriented coding*, PhD thesis, University of Wollongong.
20. Pulkki V. (1997), *Virtual sound source positioning using vector base amplitude panning*, J. Audio Eng. Soc., **45**, 6, 456–466.
21. Schröder M. R.(1958), *An artificial stereophonic effect obtained from a single audio signal*, J. Audio Eng. Soc., **6**, 2, 74–79.
22. Szczerba M., Oomen W., Therssen D. (2011), *Parametric audio based decoder and music synthesizer for mobile applications*, Archives of Acoustics, **36**, 2, 461–478.
23. Thurstone L. L. (1994), *A law of comparative judgment*, Psychological Review, **101**, 2, 266–270.
24. Wendt K. (1963), *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie*, PhD thesis, RWTH Aachen.
25. Zotter F., Frank M., Marentakis G., Sontacchi A. (2011), *Phantom source widening with deterministic frequency dependent time-delays*, [in:] DAFx-11 Proceedings, 307–312, Paris.