DOI: 10.22032/dbt.39969
DOI (proceedings): 10.22032/dbt.39936

Verband Deutscher Tonmeister e.V.

Sweet area size for the envelopment of a recursive and a non-recursive diffuseness rendering approach

M. Blochberger[1], F. Zotter[2], M. Frank[2]

[1] *Graz University of Technology, Austria, Email: matthias.blochberger@student.tugraz.at*

[2] *Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, Austria, Email: { zotter, frank }@iem.at*

Abstract

We compare the extent of the usable audience area of two algorithms that produce diffusely enveloping, multi-channel surround playback from a single-channel input. The FIR approach designs a set of random group-delay allpass filters to generate a set of minimally correlated playback signals. Canfield-Dafilou presented a frequency-dependent maximum group delay value as a constraint to keep audible artifacts small, in studio environments. To enlarge the audience area in which an enveloping and diffuse listening experience is achieved, we relax this constraint while having to accept an unavoidable impression of spaciousness and reverberation. Consequently, the FIR approach naturally competes with IIR feedback-delay network as alternative approach. We conduct listening experiments to reveal quality and effectiveness of both methods, in particular regarding sweet area size and sound quality.

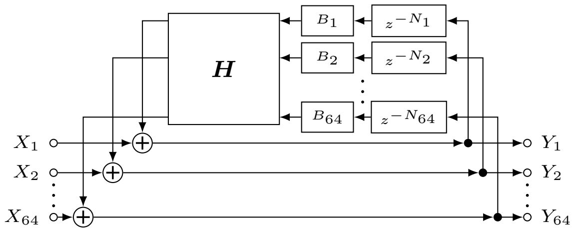
1. Introduction

Literature typically defines (e.g. Rumsey [1]) listener envelopment (LEV) as the auditory perception that spacious sound appears to arriving from all the surrounding directions in the space. Often envelopment is defined in contrast to apparent source width (ASW), which refers to the impression of a localized sound, however appearing to be wide. Others refer to measures related to late reverberation to calculate a value for the envelopment as seen in [2, 3].

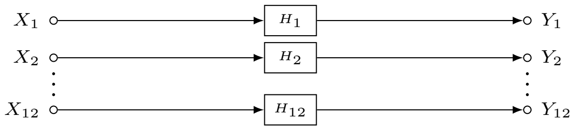
Multichannel audio playback could use room models, virtual microphones or measurements [4] and auralize them in order to produce the sensation of envelopment. By contrast, this paper regards decorrelation algorithms as less physical concept rendering diffuse sound fields on loudspeakers, such as those presented in [5–10]. Typically, such algorithms produce a set of signals by filtering a single-channel input, so that the set of signals is audio-technically considered to be uncorrelated. The expectation is that feeding those signals to surrounding loudspeakers produces envelopment.

Nevertheless, such algorithms may only partly decorrelate the signals, as audio quality requires to maintain as much of the temporal structure of the original signal as possible. And yet, spacious impressions are often generated as a noticeable side effect. Two candidate algorithms compared here: (i) Canfield-Dafilou recently proposed a promising block-filter implementation designed by random group-delay all-pass filters with frequency-dependent limits [9]. On the other hand, (ii) the feedback-delay network [11,12], is a multi-channel IIR structure offering high onset fidelity and efficiency.

In this paper we are going to illuminate how well the exemplary algorithms perform in producing envelopment over an extended listening area. Zotter/Frank [13] and Frank [14, 15] contained ideas to investigate the sweet-area size of decorrelated/diffuse and enveloping sounds, which we are going to extend, here.

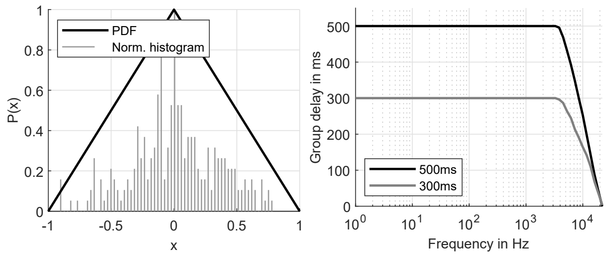


(a) Block diagram of the FDN used for the experiments.

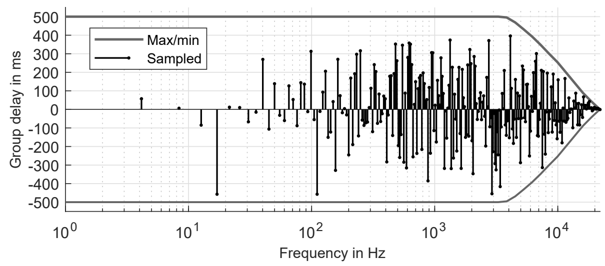


(b) Block diagram of GDL FIR filter set.

Fig. 1: The FDN used in the experiment with the mixing matrix $\boldsymbol{H}$ in the feedback loop and the GDL FIR approach has 12 filter implemented as FFT.



(a) Probability density function (b) Maximum group delay curves



(c) Example of the maximum group delay with the sampled values at ERB-spaced frequencies ($M = 512$, $N = 512$), in ms.

Fig. 2: The triangular probability density function ensures more values close to zero. The maximum group delay curves have a roll off above 4kHz as it provided better sounding impulse responses. The sampled values on the interval $[-1\,1]$ are scaled to the min/max interval of group delay values.

2. FDN (IIR)

Feedback-delay networks are multi-channel recursions, e.g. 64 channels, denoted in the z-domain as

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_{64} \end{bmatrix} = \boldsymbol{H}\,\mathrm{diag}\left\{\begin{matrix} B_1 \\ \vdots \\ B_{64} \end{matrix}\right\}\mathrm{diag}\left\{\begin{matrix} z^{-N_1} \\ \vdots \\ z^{-N_{64}} \end{matrix}\right\}\begin{bmatrix} Y_1 \\ \vdots \\ Y_{64} \end{bmatrix} + \begin{bmatrix} X_1 \\ \vdots \\ X_{64} \end{bmatrix}$$
$$\text{w. } B_i = g_{\mathrm{lo}}^{N_i} B_{\mathrm{lo}} + g_{\mathrm{mid}}^{N_i} B_{\mathrm{mid}} + g_{\mathrm{hi}}^{N_i} B_{\mathrm{hi}}, \tag{1}$$

their feedback matrix $\boldsymbol{H}$ is unitary, the channel delays N_i are individual, and the channel gains in 3 bands g_{lo}, g_{mid}, g_{hi} correspond to the desired attenuation per sample (Fig. 1(a)) The network employed in the study is the IEM FdnReverb [16]. It uses a selection of increasing prime numbers to specify the sample delays N_i. For efficiency and optimal mixing, the unitary matrix $\boldsymbol{H}$ is implemented as Fast Walsh-Hadamard transform with 64 multiplications (normalization) and $8 \cdot 64$ sums/differences, see Rochesso [12], which leaves the 3-band IIR filters in the 64 bands as the only costly operation.

As is, the FDN would start with a strong attack after which the signal decays in a $10^{-3t/T_{60}}$ shape. In the implementation here, the single-channel input is fed to the inputs $X_1, \ldots, X_{12}$ to get a delayed feed-forward signal to every loudspeaker from the outputs $Y_1, \ldots, Y_{12}$.

The IEM FdnReverb plugin moreover allows to set a slow onset, which is accomplished by subtracting two 64-channel FDNs. From the main FDN with the desired decay time, another one is subtracted that exhibits a fast decay, which becomes a $(1 - 10^{-3t/T_{\mathrm{onset}}})$-shaped onset. In the slow-onset implementation here, the single-channel input is fed to the input X_1 to get a delayed feed-forward signal to every loudspeaker from the outputs $Y_1, \ldots, Y_{12}$.

These FDN input setups were used in all experiments except number 1 (see Sec. 4).

3. Random GDL (FIR)

The FIR approach uses impulse responses generated from random group delay curves in the same manner as Canfield-Dafilou and Abel proposed in [9].

For each filter, M random values are drawn from a symmetric triangular probability density function (Fig. 2(a)) provided by MATLABs *makedist* method. Having the maximum probability at the zero value ensures a higher amount of zero/small values. These values, spaced on a Moore-Glasberg ERB [17] warped frequency scale, yield the group delay curve. The generated curve is scaled by maximum and minimum values predefined for each frequency (Fig. 2(c)).

To avoid instability of the listening impression stemming from certain phase differences, the group delay curves are modified to limit the differences at low frequencies. One of the generated curves is taken as common ground while the others are scaled to fall into the interval $\left[-(4f)^{-1} \quad (4f)^{-1}\right]$ in relation to that common curve. A cross fade between the modified curves and originally generated curves happens from 600 Hz to 1400 Hz.

Using any group delay function $\tau(f)$ over frequency f the phase is calculated by

$$\phi(f) = -2\pi \int_0^f \tau(f)\mathrm{d}f. \tag{2}$$

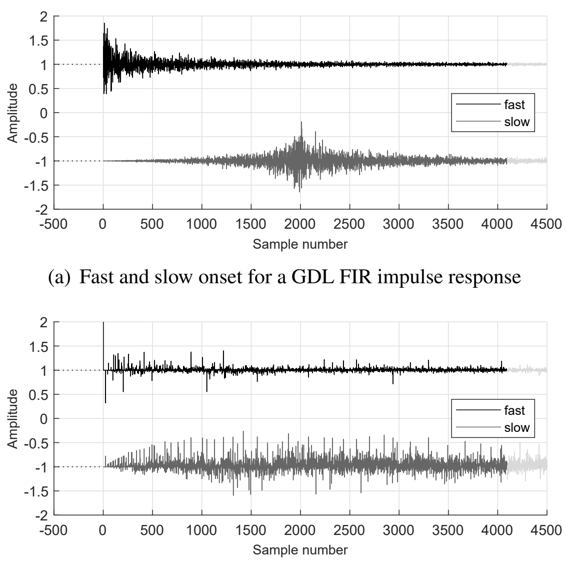
Evaluating the integral for frequencies on the interval $\left[0 \quad \frac{f_s}{2}\right]$ where f_s is the sampling frequency. Since the positions of the sample values on the frequency scale are not equidistant, a resampling to the equidistant bins $k = 0, \ldots, N/2$ at $f = \frac{k\, fs}{2N}$ has to be done first, in this case linear interpolation provides sufficient precision. Integration is numerically solved by $\phi[k] = -\sum_{k'=0}^{k} \tau[k'] \frac{2\pi\, f_s}{N}$, and mirroring about the origin yields the phase of the desired skew-symmetric spectrum. Lastly the the inverse fast Fourier transformation yields a real valued impulse response

$$h[n] = \mathcal{IFFT}\left\{e^{j\phi[k]}\right\}. \tag{3}$$

The resulting all-pass impulse responses are symmetric in time which means the slow onsets can be disturbing. An option to mitigate this problem is to truncate the impulse responses at chosen points. The best option for the preservation of transients is to truncate at the center symmetry of the impulse response. For slower onsets it can be truncated at earlier points. In general this leads to better onsets but compromises the all-pass frequency response. The randomized nature of the impulse responses leads to random deviations in the frequency responses, which are cancelling out when using a sufficient amount of impulse responses.

The impulse responses generated by this method and used in this study are of two lengths and onset types. A maximum group delay value curve of 300 ms as well as 500 ms both with a decrease in high frequencies above 4 kHz (Fig. 2(b)) with a fast onset and an onset of 2000 samples (Fig. 3(a)). The slow onset is formed by truncating the symmetric impulse response 2000 samples (45 ms) earlier and multiplying a fade in function of the form $f(n) = 2\frac{n}{L} - \left(\frac{n}{L}\right)^2$ with the first 2000 samples where $n = 0 \ldots L$ is the sample number and $L = 2000$ the length of the fade in.

All GDL impulse responses used for the following experiments were generated using the parameters $M = 131072$, $N = 131072$, $fs = 44100Hz$.



(a) Fast and slow onset for a GDL FIR impulse response

(b) Fast and slow onset for a FDN IIR impulse response

Fig. 3: Fast and slow onsets for truncated GDL impulse responses and a FDN. The slow onsets for the FDN approach are fit to the peak location of GDL at 45 ms.

4. Off-center envelopment (Exp. 1)

Beranek [3] defines envelopment as perceived presence of all sound arrival directions in a room. Choisel/Wickelmaier [18] define it as the perception of a sound that *wraps around you* giving you the impression of *being immersed in it* in contrast to *being outside of it*. Most literature defines it in a similar manner, although a common accepted definition is non-existent.

In our first experiment, listeners were asked which loudspeaker directions did not appear to contribute to the surround sound playback, at two off-center listening positions.

4.1. Method

Both FDN and GDL algorithms were used to produce 12 independent impulse responses to feed the 12 horizontal loudspeakers of the IEM CUBE (10×11 m, 500 ms reverberation time; Fig. 4). The parameters of the GDL approach are as described in Sec. 3, the parameters of the FDN are listed in Tab. 1 which were chosen to fit the corresponding GDL conditions by ear. The onset was varied between $T_{\text{onset}} = \{0,\, 45\}$ ms, yielding 8 conditions in total.

For this experiment only, the inputs of the FDN were fed by a mono source encoded in 5th-order ambisonics in order to feed signal into multiple inputs simultaneously (Azimuth: 24°, Elevation 38°). As a loop sound, *Joyride* from IEM OpenData Archive [19] was used.

8 participants aged from 23 to 39 years took part in the experiment. Every participant did the task alone and had a remote control to switch between the 8 conditions. For each condition, the participant could look into any direction and was asked to write on a piece of paper which of the 12 loudspeaker directions did not appear to contribute at the laterally off-center listening positions −3 m, −1 m, 1 m, and 3 m. 4 listeners were able to finish the task at both, the left and right off-center listening positions within 30 min, and 4 did so for only the left off-center listening position (average time for either left or right positions 22 min).

Listeners reported difficulty in estimating the presence of a loudspeaker direction whenever the perceived sound appeared

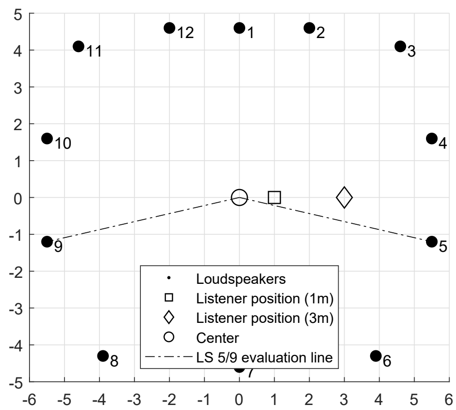
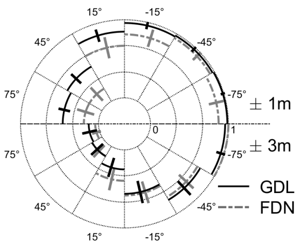


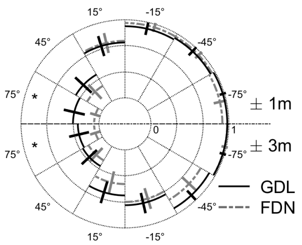
Fig. 4: Layout of the IEM CUBE with loudspeaker positions, listening positions for experiment 1 and evaluation lines for experiment 2.

Stim.	Parameter			
	Room size	Rev. Time [s]	Fade-In [s]	Gain [dB]
300ms *so.*	4	0.7	0.11	0.0
300ms *fo.*	4	0.7	0	0.0
500ms *so.*	4	1.0	0.11	0.0
500ms *fo.*	4	1.0	0	0.0

Tab. 1: Settings for the IEM FdnReverb for experiment Nr. 1. The Filter Gain parameter is applied to both bands for a flat filter response. (*so.* = slow onset; *fo.* = fast onset)



(a) 300 ms conditions



(b) 500 ms conditions

Fig. 5: The average perceived loudspeaker activity from binary responses of 12 directions are shown in terms of the means (circular segments) and 95% confidence intervals (radial segments). Sectors with significant differences are marked ($*$). Listeners were laterally off-center by either 1 m or 3 m, left and right offsets were mirrored and averaged, and front-back responses were assumed symmetrically pooled (upper half plane: graph for 1 m , lower half plane: graph for 3 m), with short (a) and long (b) decorrelation networks, FDN and GDL.

to be closer than the loudspeaker in distance.

4.2. Results

The evaluation was done based on the small set of responses, which was not extended because of the difficulty and duration of the task. As listeners freely changed their look direction, responses for back loudspeakers were mapped to the front, and responses for left-off-center listening positions were mapped to such for right-off-center positions. In this way, 8 responses acquired 12 directions for left offsets and 4 for right offsets, give us $2 \times (8 + 4) = 24$ responses for 6 frontal directions, per condition. Pair-wise tests on the pooled data indicated hard/soft onset not to be significant as factor, so responses for the hard/soft onset conditions were pooled, yielding 48 data points per direction. By contrast FDN/GDL and short/long response lengths were found to improve the ratings.

Fig. 5 shows the averaged binary directional response (loudspeaker direction perceived to contribute or not) of the 6 frontal response directions on a linear radial scale from 0 to 1. Upper half planes show perceived directional activity for 1 m off-center positions and lower half planes for 3 m. Whereas more reverberant responses of the decorrelator networks generally produce a slight increase in directional activity, we can

Stim.	Parameter			
	Room size	Rev. Time [s]	Fade-In [s]	Gain [dB]
300ms *so.*	8	0.6	0.10	0.0dB
300ms *fo.*	8	0.6	0	0.0dB
500ms *so.*	8	1.0	0.10	-2.6dB
500ms *fo.*	8	1.0	0	-2.6dB

Tab. 2: Settings for the IEM FdnReverb for experiment 2 and 3. The Filter Gain parameter is applied to both bands for a flat filter response. (*so.* = slow onset; *fo.* = fast onset)

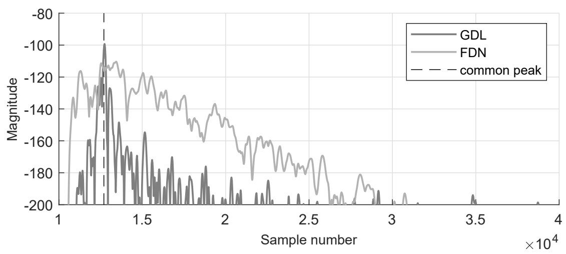


Fig. 6: The squared impulse responses are summed up to fit the onset peak of FDN to the onset peak of GDL.

also observe in Fig. 5 that distant loudspeakers may only produce a perceived activity up to $\frac{1}{2}$ according to our results.

For a directional activity rated with $\geq \frac{1}{2}$, we can find a coverage of $\pm 120°$ measured from the closest loudspeaker at 1 m off-center, or $\pm 90°$ at 3 m. Apparently, already slight off-center positions can exclude distant loudspeakers from a clear audibility in an all-enveloping playback setting.

In greater detail, the GDL algorithm appears produce significantly better results in some cases (Student's T-test). In particular, GDL produces significantly more loudspeaker activity with long group delay time (500 ms) at the 75° direction at both the 1 m ($p = 0.0001$) and the 3 m ($p = 0.001$) off-center position and weakly outperforms at 1 m/45° ($p = 0.0882$), 3 m/15° ($p = 0.0512$) and 3 m/$-45°$ ($p = 0.0324$). The short group delay condition (300 ms) weakly outperforms the FDN counterpart at 1 m/$-75°$ ($p = 0.0579$), 1 m/$-15°$ ($p = 0.0569$). On the other hand, the FDN algorithm has a weak advantage at 3 m/15° ($p = 0.0702$). All other condition-direction combinations show no significant differences.

From these results, we might conclude that envelopment in the standard definition (presence of all directions) is infeasible for an extended audience area in surround playback using individual loudspeakers. While this might appear counter-intuitive, e.g. when regarding the apparent success in [15], plausibility of an enveloping auditory scene might not depend on a detailed presence of all directions. The presence of dry and direct as well as lateral reverberant sound might already be satisfactory.

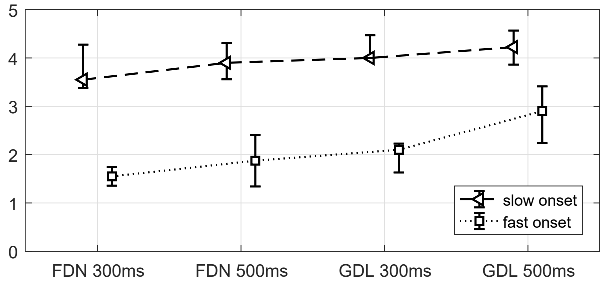


Fig. 7: Rating of the lateral limit of the plausible-reproduction sweet area in meters, with regard to direct frontal and reverberant enveloping/lateral sound, in medians and 95% confidence intervals.

5. Sweet-area size (Exp. 2)

As envelopment might be too strict a criterion to evaluate the sweet-area size, our second experiment considers direct sound from the frontal loudspeaker together with the diffuse sound from all the 12 horizontal loudspeakers.

5.1. Method

The 8 conditions were the same GDL impulse responses as in experiment 1 as described in Sec. 3 and the FDN parameters as listed in Tab. 2. The direct sound from the frontal loudspeaker was set at a level of -6 dB compared to the diffuse sound field at the center position.

For this experiment, the parameters of the FDN implementation were fit to the GDL approach by the measure of the center time

$$T_s = \frac{\int_0^\infty t \cdot h^2(t)\mathrm{d}t}{\int_0^\infty h^2(t)\mathrm{d}t} \quad (4)$$

for the fast-onset stimuli and additionally fitting the onset peak of the FDN to the onset peak of GDL at 45 ms (Fig. 6). Again, the onset was varied between $T_{\mathrm{onset}} = \{0, 45\}$ ms, yielding 8 conditions in total and as a loop sound, *Joyride* from the IEM OpenData Archive [19] was used.

Listeners were individually asked to switch through 8 conditions with a remote control. For each condition, their task was to write on paper the lateral limits of the sweet area when walking the line from the central listening spot towards either loudspeaker 5 (at $110°$ right) or 9 (at $110°$ left) as depicted in Fig. 4. We gave them a definition of the sweet area as being laterally limited where either (i) the frontal sound would move outside the $\pm 30°$ range of the frontal loudspeakers 12-1-2, or (ii) the lateral reverberation begins to dominate and disturb the spatial impression.

8 persons between 24 and 55 years of age participated in the experiment with an average duration of 12 minutes. The lower time it took participants to complete this task, suggests that this task was much simpler than the previous one.

5.2. Results

Fig. 7 shows that overall that diffuse impulse responses with slow onset are most effective when desiring an extended sweet area for direct sound plus reverberation. A Wilcoxon signed rank test with Bonferroni-Holm correction confirms the significantly larger sweet area for each condition ($p < 0.009$). Only in case of the GDL with fast onset, long reverberation contributes to a enlarged sweet area ($p = 0.025$). In detail, the short GDL with slow onset weakly outperforms its FND counterpart ($p = 0.051$). The same holds for the long GDL with fast onset ($p = 0.087$).

6. Envelopment/transients (Exp. 3)

While the experiments above did not consider audio quality aspects and would not give much insight yet into which of the algorithms is more effective, our third experiment uses a multi-stimulus test setup for the central listening position (Fig. 4) to acquire a closer differentiation.

6.1. Method

The conditions were chosen as in the previous experiment investigating sweet area size, however without the frontal direct sound. The 12 surrounding loudspeakers were fed by the 8 different filter sets to switch between and the listeners were asked to comparatively rate in multi-stimulus trials: (i) the preservation of transients ($\min \ldots \max$) as an audio quality aspect, and (ii) the diffuse envelopment ($\min \ldots \max$) they perceived as a criterion of effectiveness. Both multi-stimulus tasks were rated twice per listener, each time with randomly arranged assignment of the stimuli to the 8 sliders. Like previously, *Joyride* from IEM OpenData Archive [19] was used.

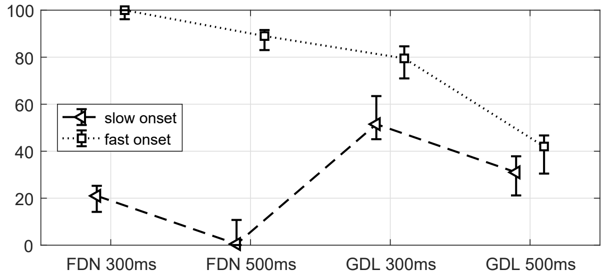
As this experiment happened in conjunction with the previous one, the same 8 persons aged 24 to 55 years participated. The average completion time was 18 minutes.

The participants reported difficulty in the diffuse envelopment task, especially whenever conditions appeared to be different in timbre (high frequencies) or onset, which occasionally made their decision difficult.

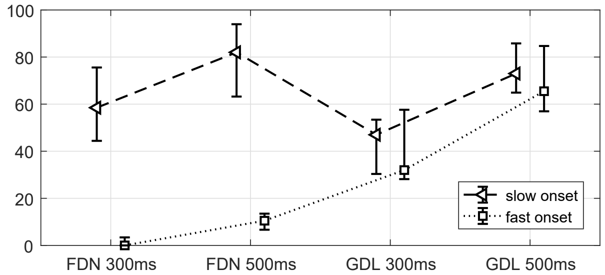
6.2. Results

Fig. 8(a) shows that the slow onset deteriorates the perceived preservation of transients for both FDN and GDL ($p < 0.01$), except for the long GDL ($p = 0.125$). This negative effect is worse in the FDN conditions compared to such with GDL. In general, quality decreases for longer decay times ($p < 0.011$). The fast onset conditions (dotted) are generally of higher quality, and there, the FDN conditions outperform the transient preservation of GDL ($p < 0.011$), especially for the long decay time. This might be due to the linear decay profile of the GDL responses which on the other hand appears to be beneficial under slow onset conditions, as GDL largely outperforms FDN in terms of diffuseness.

The comparison of the diffuse envelopment rating of the sound fields in Fig. 8(b) shows an increase with the decay length ($p < 0.029$) and increase with onset time by tendency: In case of the FDN approach, the fast onset conditions perform significantly less effective ($p < 0.011$) than their slow counterparts. However, in case of the GDL approach, there



(a) Preservation of transients



(b) Diffuse envelopment

Fig. 8: Comparative rating of the diffusion networks in terms of quality (preservation of transients) and effectiveness (diffuse envelopment) at central listening position concerning medians and 95% confidence intervals.

are no significant differences for both the short and long decay time.

The GDL condition with 300 ms with slow onset appears to be a good compromise between preservation of transients (quality) and the perceived diffuse envelopment (effectiveness), and despite the larger efforts required, the GDL approach appears to be an interesting alternative.

7. Conclusion

In this contribution we compared two fundamentally different diffuseness and envelopment rendering approaches on a rather large loudspeaker setup, random-group-delay (GDL) allpasses implemented as FIR structures and feedback-delay networks (FDN) that are implemented IIR.

Independent of the algorithms and their settings, we found that it is challenging to produce diffuse envelopment across a large audience area in experiment 1 (Sec. 4), in particular in the strict understanding of sound being perceived to directionally arrive from everywhere. This is because the contribution of distant loudspeakers soon becomes imperceptible, even at relatively small distances to the central listening position. The standard definition of envelopment appears impractical in evaluation of sound fields for larger audiences. As a refinement of the method, future experiment might consider asking for the auditory distance of the enveloping sound into a given set of directions.

For diffuse rendering with direct sound, experiment 2 (Sec. 5) showed that independent of the approach, methods with slow onset have a crucial advantage over such with fast onset. Consistent rendering of envelopment is possible for a sweet area of about 4 m when using algorithms producing impulse responses with mentioned slow onsets, in contrast to such with fast onsets (2 m). Moreover, the random group-delay-based approach appears to produce a slightly larger sweet area.

For both algorithms, we showed in experiment 3 (Sec. 6) that they produce more envelopment when used with slow onset, in particular the FDN structure. Additionally, the proposed GDL structure with moderate time constants ($300ms$) can be recommended in terms of its pronounced effect, while its transient preservation is quite acceptable.

Acknowledgements

We thank the voluntary listeners for their participation in our experiments.

References

[1] F. Rumsey, “Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm,” *J. Audio Eng. Soc*, vol. 50, no. 9, pp. 651–666, 2002. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=11067

[2] G. A. Soulodre, M. C. Lavoie, and S. G. Norcross, “Investigation of listener envelopment in multichannel surround systems,” in *Audio Engineering Society Convention 113*, Oct 2002. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=11283

[3] L. L. Beranek, “Concert hall acoustics,” *J. Audio Eng. Soc*, vol. 56, no. 7/8, pp. 532–544, 2008. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14398

[4] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, “Spatial decomposition method for room impulse responses,” *J. Audio Eng. Soc*, vol. 61, no. 1/2, pp. 17–28, 2013. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=16664

[5] E. Weinstein, M. Feder, and A. V. Oppenheim, “Multichannel signal separation based on decorrelation,” Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, RLE Technical Report 573, 1992.

[6] G. S. Kendall, “The decorrelation of audio signals and its impact on spatial imagery,” *Computer Music Journal*, vol. 19, no. 24, 1995.

[7] G. Potard and I. Burnett, “Decorrelation techniques for the rendering of apparent sound source width in 3d audio displays,” in *Proc. of the 7th Int. Conf. on Digital Audio Effects*, 01 2004, pp. 280–208.

[8] M. Boueri and C. Kyriakakis, “Audio signal decorrelation based on a critical band approach,” in *Audio Engineering Society Convention 117*, 01 2004.

[9] E. K. Canfield-Dafilou and J. S. Abel, “A group delay-based method for signal decorrelation,” in

Audio Engineering Society Convention 144, May 2018. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=19508

[10] S. Schlecht, B. Alary, V. Välimäki, and E. Habets, "Optimized velvet-noise decorrelator," in *21st International Conference on Digital Audio Effects (DAFx-18)*, Aveiro, Portugal, 09 2018. [Online]. Available: http://dafx2018.web.ua.pt/papers/DAFx2018_paper_23.pdf

[11] J. Stautner and M. Puckette, "Designing multi-channel reverberators," *Computer Music Journal*, vol. 6, no. 1, pp. 52–65, 1982.

[12] D. Rocchesso and J. O. Smith, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 1, pp. 51–63, Jan 1997.

[13] F. Zotter, M. Frank, M. Kronlachner, and J. Choi, "Efficient phantom source widening and diffuseness in ambisonics," in *EAA Symposium on Auralization and Ambisonics*, Berlin, 2014.

[14] M. Frank and F. Zotter, "Spatial impression and directional resolution in the reproduction of reverberation," in *Fortschritte der Akustik - DEGA*, Aachen, 2016.

[15] ——, "Exploring the perceptual sweet area in ambisonics," in *Audio Engineering Society Convention 142*, Berlin, 2017. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18604

[16] Plug-in suite by iem. Software Package. Institute of Electronic Music and Acoustics. [Online]. Available: https://plugins.iem.at/

[17] B. Moore and B. Glasberg, "A revision of zwicker's loudness model," *Acta Acustica united with Acustica*, vol. 82, pp. 335–345, 03 1996.

[18] S. Choisel and F. Wickelmaier, "Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preferenc," *The Journal of the Acoustical Society of America*, vol. 121, pp. 388–400, 02 2007.

[19] Opendata archive. Archive. Institute of Electronic Music and Acoustics. [Online]. Available: https://opendata.iem.at/