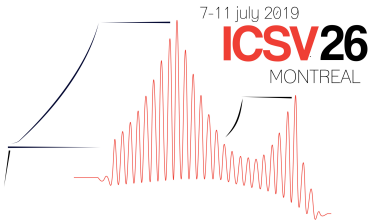
Refinement and implementation of a robust directional room impulse response denoising process, including applications to highly varied measurement databases

Pierre Massé, Thibaut Carpentier, Olivier Warusfel, Markus Noisternig

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



REFINEMENT AND IMPLEMENTATION OF A ROBUST DIRECTIONAL ROOM IMPULSE RESPONSE DENOISING PROCESS, INCLUDING APPLICATIONS TO HIGHLY VARIED MEASUREMENT DATABASES

Pierre Massé, Thibaut Carpentier, Olivier Warusfel and Markus Noisternig
Acoustic and Cognitive Spaces Group, IRCAM, CNRS, Sorbonne Université, UMR 9912 STMS, Paris, France
email: pierre.masse@ircam.fr

Recent developments in the measurement of directional room impulse responses (DRIR) by spherical microphone arrays (SMA) have led to their extensive use in sound spatialisation. Room reverberation effects can be reproduced in three-dimensional surround sound systems (e.g. Higher-Order Ambisonics) through multi-channel DRIR convolution. However, such measured impulse responses inevitably present a non-negligible noise floor, leading to a perceptible 'infinite reverberation effect'. Further, individual sensor noise and non-stationary measurement artefacts may additionally corrupt the deconvolved impulse response. This paper presents recent work regarding the implementation of state of the art DRIR analysis and denoising techniques and their application to extensive DRIR databases measured across a highly varied collection of spaces. We first review the basic energy decay relief (EDR) analysis and reverberation tail re-synthesis process, before presenting several novel refinements developed throughout the course of this implementation. Finally, an overview of the results obtained both globally and with respect to particular cases (complex architectural volumes, outdoor spaces, etc.) is included in order to further examine the capabilities of the denoising framework.

Keywords: DRIR, denoising, EDR, SMA, ambisonics

1. Introduction

The use of directional room impulse responses (DRIR) for recreating three-dimensional reverberation effects in spatialized sound installations has become a widespread practice in recent years, enabled by developments in both microphone and loudspeaker array technologies. Spherical microphone arrays (SMA), for example, have become a natural counterpart to the Higher-Order Ambisonics (HOA) format, based on a spherical harmonic (SH) representation of the sound field. DRIR can thus be measured using SMA, e.g. through the exponential sweep method (ESM), and encoded in HOA up to a maximum order dictated by the array's specific characteristics (number of transducers, configuration, etc.). However, such recordings are inevitably subject to measurement noise, which translates to a non-negligible noise floor in the impulse response, leading to a perceptible 'infinite reverberation effect'. Non-stationary noise and other artefacts may further corrupt the DRIR, rendering decay analysis and parameter extraction unstable.

In this paper, we present the robust implementation of a complete DRIR denoising framework and its application to a comprehensive measurement database covering a large variety of reverberant spaces.

This process is streamlined to exploit the unique advantages offered by a three-dimensional description of room reverberation. It involves three main steps: pre-deconvolution artefact suppression applied directly to the ESM measurement, DRIR energy decay relief (EDR) and diffuseness analysis, and finally reverberation tail re-synthesis. These are each detailed in turn, and then illustrated with specific examples obtained from the measured DRIR database, followed by a brief statistical review of the overall database results. The introduction is completed below by a summary of the theoretical bases underlying SMA measurement and SH analysis, and an overview of the existing litterature this current work draws from.

1.1 Theoretical Background

Spherical microphone arrays enable the directional analysis of a given sound field by sampling over Q transducer positions on their surface. A natural representation for such three-dimensional signals is the spherical harmonic domain that serves as the basis of the spatial Fourier transform:

$$X_{l,m}(f,t) = \int_{\Omega \in S^2} x(f,\Omega,t) Y_{l,m}(\Omega) d\Omega, \tag{1}$$

where $\Omega = (\theta, \phi)$ is an angular point on the surface of a sphere with radius r, $x(f,\Omega,t)$ is the time-frequency domain representation of the sound field on a sphere with fixed radius $r = a$, and $Y_{l,m}(\Omega)$ are the spherical harmonics of order $l \in \mathbb{Z}^+$ and degree $m = [-l, l]$. Using a SMA, the integral in Eq. (1) is approximated by a weighted sum over the microphone positions; the specific weights are obtained by minimizing the error between this approximation and the ideal integral of Eq. (1) [1]. For this discrete transform to be exact, the maximum encoding order L and the array sampling configuration must lead to an encoding matrix with K non-vanishing singular values such that $K = (L+1)^2 \leq Q$ [2], thereby effectively limiting L for a given SMA. Finally, in order to project the ideal sound field as seen from *within* the sphere, a subsequent correction for the so-called holographic function (or mode strength) $b_l(f)$ of the SMA must be applied. Such is the case in the widespread HOA format, which uses the center of the sphere as a standardized reference point and defines the correcting filters accordingly [3].

1.2 Previous Work

The late diffuse reverberation tail of a room impulse response can be modelled by an exponentially decaying stochastic process [4], assuming sufficiently high echo density and modal overlap (time-frequency limits respectively corresponding to the so-called 'mixing time' and Schroeder frequency) [5]. These conditions lead to the description of an ideal diffuse wave field that can be synthesized in the form of a filtered Gaussian noise [6]. The exponential decay envelope of the diffuse reverberation tail is parameterized by a frequency-dependent decay coefficient (often represented as a $T_{60}(f)$ reverberation time) and an initial power spectrum $P_0(f)$; these parameters can be extracted by analysis of the EDR, a time-frequency extension of the Schroeder energy decay curve (EDC) [6]. A diffuse reverberation tail corrupted by non-decaying measurement noise can therefore be replaced by a synthesized zero-mean Gaussian noise with a frequency-dependent exponential decay envelope defined by the extracted parameters [7] [8]. The resulting signal-to-noise ratio (SNR) is thereby limited only by the quantization noise floor for the chosen audio format bit depth, $N_q = 20\log_{10}(2^{-d})$ dB, where d is the signal bit depth [9].

2. DRIR Denoising Process

The different parts of the proposed denoising framework are now presented in this section. The current work focuses more specifically on the artefact reduction and tail re-synthesis procedures (sections 2.1 and 2.2.3), and to a lesser extent on the EDR and diffuseness analysis steps (sections 2.2.1 and 2.2.2).

2.1 Raw Measurement Artefact Suppression

Impulse responses are often measured using the ESM method [10]. Measuring a repetition of several sweeps allows the multiple realizations to be averaged in order to increase the SNR, since the ensemble mean of any stationary sensor noise will tend to zero as the number of repetitions increases. However, momentary artefacts or non-stationary noise will on the contrary lower the SNR by accumulating over the averaged repetitions. This is especially troublesome in the case of impulsive artefacts which can generate audible 'inverse sweeps' upon deconvolution of the measurement.

In an attempt to minimize the influence of non-stationary measurement artefacts, the magnitude spectra of the individual sweep realizations are compared amongst each other in order to identify artefacts using non-negligible positive deviations from the mean magnitude spectrum as a discriminating criterion. This maximum allowed deviation is defined as $\xi(f,t) = \mu(f,t) + \alpha\sigma(f,t)$, where $\mu(f,t)$ is the mean magnitude spectrum, $\sigma(f,t)$ is the standard deviation over the available realizations, and α is an empirically-set deviation factor used as a control parameter. Artefact magnitude values identified as greater than $\xi(f,t)$ in each realization are then replaced with the corresponding mean over the remaining repetitions. This process is applied individually to the ESM measurement signals recorded by each SMA transducer; an example is presented below in section 4.1.

2.2 Reverberation Tail Analysis and Re-Synthesis

In this section, we first review the EDR analysis procedure used to extract the reverberation decay parameters, and then present a characterization of the DRIR's 'mixing time' using measures of the sound field's *diffuseness*, before showing that re-synthesizing the reverberation tail as a zero-mean Gaussian noise in the SH domain preserves the diffuse field's spatial properties.

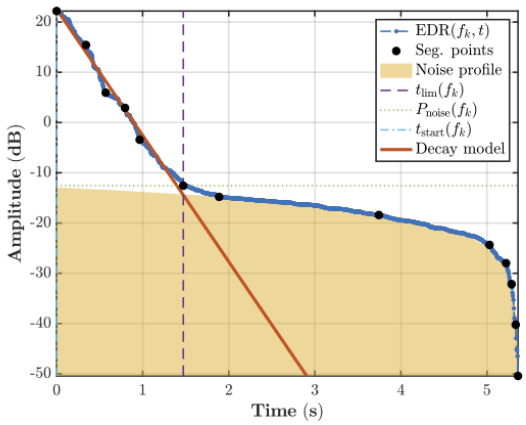
2.2.1 EDR Analysis

As proposed by Jot *et al.* [6], decay envelope parameters are extracted by analysing each EDR frequency bin independently. The basic strategy is to first identify the parts of the decay curve corrupted by non-decaying background noise, then estimate the reverberation time by linear regression over the unaffected sections, and finally 'extrapolate backwards' to find the initial power value (denoted P_0).

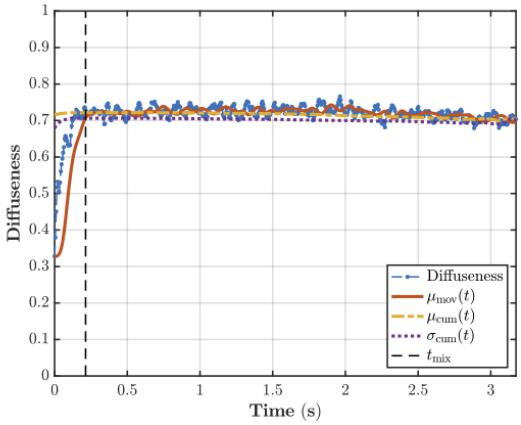
To this end, the decay curve is first segmented using an adaptive Ramer-Douglas-Peucker algorithm [11] in order to aid identification of the various sections; as a result, non-exponentially decaying early reflection regimes can be discarded by selecting an appropriate starting segment point (t_{start}). The noise floor limit point $\{P_{\text{noise}}, t_{\text{lim}}\}$ can then be found by fitting an ideal reverse-integrated noise profile to the segments (i.e. the reverse cumulative sum of an ideal constant-power noise envelope, see the yellow shaded area on Fig. 1a). An additional headroom of around 3 dB is defined above this noise profile in order to avoid the transition zone between the exponential decay and noise floor sections of the curve. The decay section of the curve, thus delimited by t_{start} and t_{lim}, can finally be modelled by linear regression in order to determine the T_{60} and P_0 values (see the solid orange line on Fig. 1a). The quality of this linear regression can be quantified with a *decay error* defined as the mean squared error (MSE) over the length of the decay section.

2.2.2 Diffuseness Analysis

As noted above, re-synthesizing the reverberation tail using a zero-mean Gaussian noise supposes that the sound field has become sufficiently diffuse for this model to be valid; in other words, one must be sure that the so-called 'mixing time' has been reached. In the case of omnidirectional or monophonic impulse responses, the mixing time has traditionally been estimated using statistical measures, such as the echo density [13] or kurtosis [14]. For three-dimensional DRIR, however, it seems judicious to exploit



(a) EDR analysis for a given frequency f_k of the zeroth-order (i.e. omnidirectional) SH component.



(b) Diffuseness analysis and mixing time identification using the SDR diffuseness measure [12].

Figure 1: Examples of energy decay (left) and diffuseness (right) analysis results obtained from a DRIR measured at the Kraftzentrale event venue in Duisburg, Germany, encoded to 4$^{\text{th}}$-order HOA.

the spatial characteristics on offer, and most notably measures of the sound field's diffuseness, which closely matches the conditions described for the reverberation model above. Diffuseness measures are therefore calculated on SH-domain DRIR signals according to Jarrett *et al.* [12], Epain and Jin [15], and Merimaa and Pulkki [16]; to obtain a simplified time-domain criterion, the frequency dimension is reduced by averaging these measures over the given SMA's optimal frequency range, defined with respect to its spatial aliasing limit $ka \simeq L$, where k is the wavenumber and a is the SMA radius [1].[1]

The mixing time is then estimated by identifying the onset of the maximum diffuseness plateau, i.e. the point at which the sound field has reached sufficient echo and modal densities for the stochastic reverberation model to be valid. As shown on Fig. 1b for the signal-to-diffuse ratio (SDR) diffuseness measure [12], the mixing time identification algorithm makes use of a reverse-cumulative mean $\mu_{\text{cum}}(t)$ and standard deviation $\sigma_{\text{cum}}(t)$ in conjunction with a forward-moving average $\mu_{\text{mov}}(t)$ in order to determine t_{mix} as the first point t_n such that $\mu_{\text{mov}}(t_n) \leq \mu_{\text{cum}}(t_n) - \sigma_{\text{cum}}(t_n)$.

2.2.3 Diffuse Reverberation Tail Synthesis

The signal synthesized in order to replace the non-decaying noise floor must imperatively retain the spatial properties of a diffuse field, most notably in terms of spatial coherence. This can be guaranteed by performing the reverberation tail re-synthesis directly in the SH domain. Jarrett *et al.* [12] show that a diffuse SMA signal of the form

$$X_{l,m}^{\text{diff}}(f,t) = \sqrt{P_{\text{diff}}(f,t)}\, b_l(f) \int_{\Omega \in S^2} \Phi(f,t,\Omega) Y_{l,m}(\Omega) d\Omega, \tag{2}$$

where $P_{\text{diff}}(f,t)$ is the diffuse field power envelope, $\Phi(f,t,\Omega)$ is the independent plane wave phase such that $|\Phi(f,t,\Omega)| = 1 \; \forall \; k,t,\Omega$ and $E\left[\Phi(f,t,\Omega)\Phi^*(f,t,\Omega')\right] = \delta_{\Omega,\Omega'}$, and $b_l(f)$ are the aforementioned array holographic functions (with δ representing the Kronecker delta and $E\left[.\right]$ mathematical expectation), leads to a characteristic spatial coherence value of $\gamma_{l,m;l',m'}^{\text{diff}}(f,t) = 0 \; \forall \; (l,m) \neq (l',m')$.

[1]The lower limit of this range is chosen to include the maximum amount of octaves possible between the SMA's spatial aliasing limit and typical values of the Schroeder frequency in mixing rooms, i.e. around 100∼200 Hz for large halls. In the case of the mh acoustics Eigenmike® used in this work, the spatial aliasing limit calculates to $f_{\text{alias}} \simeq 5.2$ KHz, so we can define a 5-octave range stretching down to 162.5 Hz.

Subsequently, synthesizing a zero-mean Gaussian noise of power $P_{\text{diff}}(f,t)$ and random phase $\Phi_{l,m}(f,t)$ per SH component gives a cross-power spectral density (PSD) of

$$\hat{\Psi}^{\text{diff}}_{l,m;l'm'}(f,t) = E\left[\hat{X}^{\text{diff}}_{l,m}(f,t)\hat{X}^{\text{diff}*}_{l',m'}(f,t)\right] = P_{\text{diff}}(f,t)\delta_{l,m;l'm'}, \tag{3}$$

and thus the same characteristic diffuse field spatial coherence value:

$$\hat{\gamma}^{\text{diff}}_{l,m;l'm'}(f,t) = \frac{\hat{\Psi}^{\text{diff}}_{l,m;l'm'}(f,t)}{\sqrt{\hat{\Psi}^{\text{diff}}_{l,m;lm}(f,t)}\sqrt{\hat{\Psi}^{\text{diff}}_{l,'m';l'm'}(f,t)}} = 0 \;\forall\; (l,m) \neq (l',m'). \tag{4}$$

Furthermore, the covariance properties described by Epain and Jin [15] for a diffuse field in the SH domain can be similarly shown to hold for such component-wise synthesis.

The tail re-synthesis denoising process can therefore be summarized as extending the reverberation tail on each SH component of a measured DRIR using a zero-mean Gaussian noise decaying according to a rate $T_{60}^{(l,m)}(f)$ and an initial power $P_0^{(l,m)}(f)$, starting at the noise limit time-frequency point $t_{\text{lim}}^{(l,m)}(f)$ and thereby replacing the non-decaying noise floor.

3. Measured DRIR Database

In order to assess the overall denoising procedure described above, it was applied to an extensive DRIR database measured over a highly varied collection of spaces, ranging from traditional theatres and concert halls to more unconventional and experimental environments such as forested valleys and castle ruins. The database totals 255 measurements over 19 different locations, all performed with the 32-capsule mh acoustics Eigenmike SMA and a d&b E12 loudspeaker.

4. Results

We now present some specific results obtained from the measurement database described above as examples chosen to highlight the main aspects of the denoising process, before offering a more global statistical overview of the database results.

4.1 Raw Measurement Artefact Suppression

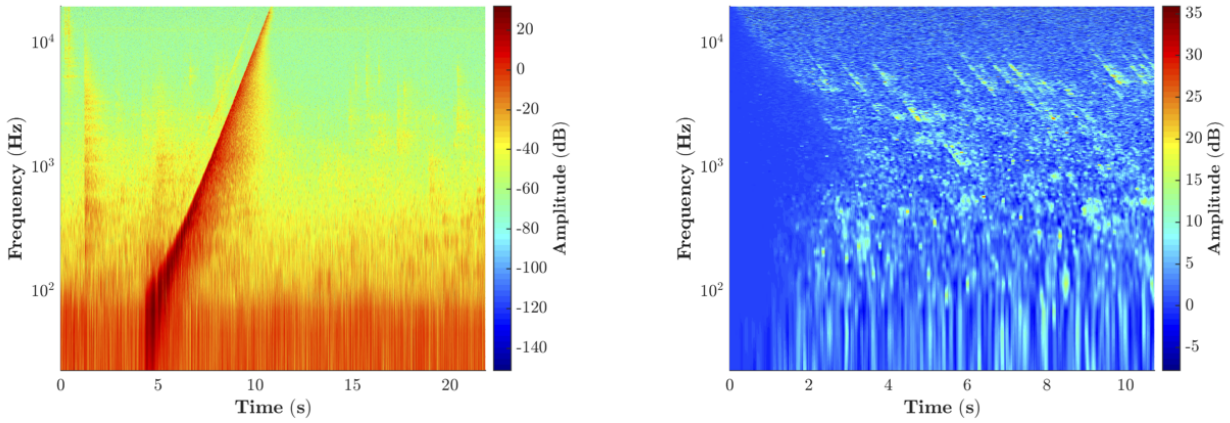
The example used to illustrate the artefact reduction procedure is a DRIR measurement performed in the Christuskirche, the German Protestant Church in Paris. This particular measurement presented several important impulsive artefacts, especially in the first of the four sweep repetitions (see Fig. 2a).

By applying the artefact suppression technique described in section 2.1 to this measurement, the overall SNR is found to be increased by approx. 2 dB with respect to the original deconvolved IR; more importantly, however, Fig. 2b shows that non-stationary artefacts with magnitudes over 15 dB are removed from the noise floor. This is crucial in regularizing the EDR so that the noise profile approaches the ideal fit used to identify the $\{P_{\text{noise}}, t_{\text{lim}}\}$ point as in section 2.2.1 (see Fig. 1a).

Finally, to further quantify artefact reduction, we define an *artefact-to-total-energy ratio*, i.e. the ratio of the total energy in a given time frame identified as 'outlying', as described in section 2.1, to the total signal energy in that frame:

$$\chi_q(t) = \frac{\sum_{k=1}^{K}|\tilde{x}(f_k,\Omega_q,t)|^2}{\sum_{k=1}^{K}|x(f_k,\Omega_q,t)|^2},\; \tilde{x}(f_k,\Omega_q,t) = \begin{cases} x(f_k,\Omega_q,t), |x(f_k,\Omega_q,t)| > \xi_q(f_k,t) \\ 0 \end{cases}. \tag{5}$$

For the current example, this averaged to $\bar{\chi}_q = 27.4\%$ over the four repetitions, with a peak of around 35% over the heavily corrupted first realization, shown in figure 2a.



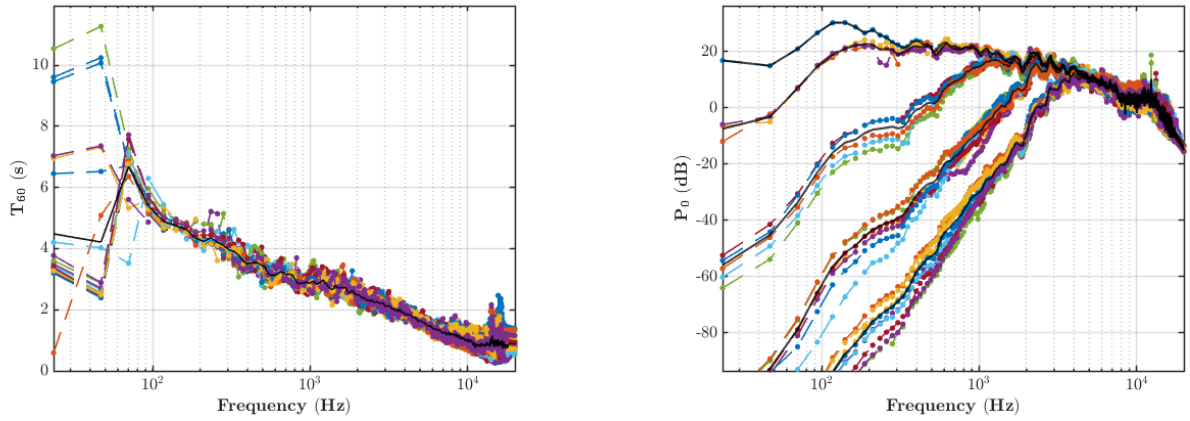
(a) Raw ESM measurement.

(b) PSD of difference in deconvolved IR.

Figure 2: Example of ESM measurement with non-stationary noise artefacts (a), and PSD of difference between deconvolved IRs before/after artefact suppression; time-averaging period $T_{\mathrm{PSD}} = 85.3$ ms (b).

4.2 Reverberation Tail Analysis and Re-Synthesis

As mentioned above, the EDR analysis is first conducted on each SH component individually. However, the encoding filters used to correct for the Eigenmike's holographic function dramatically cut the lower frequencies of the higher orders [3], and as a result the EDR analysis can become unreliable due to the reduced dynamic range. Therefore, both the $T_{60}^{(l,m)}(f)$, $P_0^{(l,m)}(f)$, and $\{P_{\mathrm{noise}}, t_{\mathrm{lim}}\}_{l,m}(f)$ extraction and the subsequent reverberation tail re-synthesis are performed before applying the corrective filters.

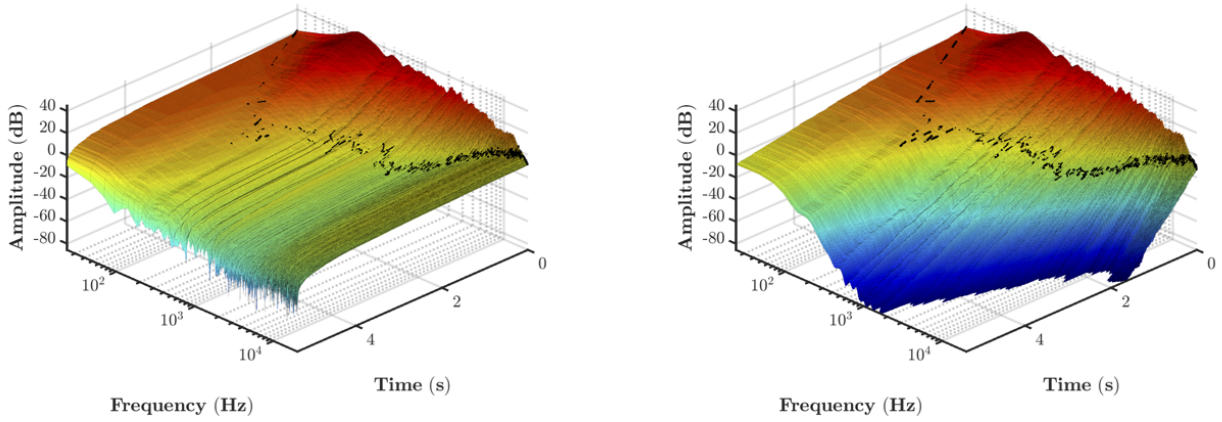


(a) $T_{60}^{(l,m)}(f)$.

(b) $P_0^{(l,m)}(f)$.

Figure 3: Component-wise reverberation times ((a), left), and initial power spectra ((b), right). Black lines show the mean over all components for the $T_{60}^{(l,m)}(f)$, and per SH order for the $P_0^{(l,m)}(f)$.

Figure 3 shows the reverberation time ($T_{60}^{(l,m)}$) and initial power spectrum ($P_0^{(l,m)}$) results obtained from analyzing the Kraftzentrale DRIR SH component-wise. The mixing time is determined as described in section 2.2.2 and Fig. 1b for each of the three mentioned diffuseness measures, and then averaged to a single $\bar{t}_{\mathrm{mix}}$ value (here $\bar{t}_{\mathrm{mix}} = 221$ ms). In order to confirm the validity of the diffuse field hypothesis used to justify the reverberation tail synthesis method described in section 2.2.3, this mixing time is compared to a $\bar{t}_{\mathrm{lim}}^{(0,0)}$ value obtained by averaging over the audible frequency range of the omnidirectional (zeroth-order, $Y_{0,0}(\Omega)$) component analysis results (here $\bar{t}_{\mathrm{lim}}^{(0,0)} = 1.29$ s). The effect of the denoising process is illustrated in Fig. 4 with the EDR of the original (Fig. 4a, left) and denoised (Fig. 4b, right) omnidirectional components.



(a) Original (after artefact reduction).

(b) After reverberation tail re-synthesis.

Figure 4: The effect of the reverberation tail re-synthesis denoising process as seen on the EDR of the Kraftzentrale DRIR's omnidirectional component. The superimposed black line shows the identified noise floor limit $\{P_{\text{noise}}, t_{\text{lim}}\}_{0,0}(f_k)$ from which point onwards the reverberation tail is synthesized.

The diffuseness profiles of the denoised DRIR are found to approach those of the original (seen in figure 1b for the SDR diffuseness measure [12]), confirming the approach taken in section 2.2.3. To quantify the effectiveness of the denoising process, we define the *continuity error* as the percent error between the original $T_{60}^{(l,m)}(f_k)$ used to parameterize the synthesized reverberation tail and a final $\widehat{T}_{60}^{(l,m)}(f_k)$ obtained by linear regression over 6 dB around the $t_{\text{lim}}^{(l,m)}(f_k)$ point on the denoised $\text{EDR}_{l,m}(f_k, t)$ curve. This error calculated to an average (over both SH components and frequency) of 9.39% in this case, whereas the decay error was 0.944 dB/time frame.

4.3 Application to Measured DRIR Database

Out of the 255 DRIR in the measurement database described in section 3, 226 were denoised using the proposed method. The remaining 29 were either found to be unsuitable for the denoising process presented here, with $\bar{t}_{\text{mix}} > \bar{t}_{\text{lim}}^{(0,0)}$ violating the diffuse field hypothesis, or EDR analysis results leading to a decay error $\epsilon_{\text{dec}}^{(l,m)}(f_k) > 6$ dB per time frame in over half of the frequency bins in the audible range, for any SH component. The latter was used as an emperically-determined limit in order to catch DRIR with such elevated analysis errors that they may not be considered as describing exponentially decaying reverberation (as is assumed by the analysis method described in section 2.2.1). Of the denoised DRIRs, 81.0% had an overall average continuity error of less than 10% (i.e. averaged over both frequency and SH components). Additionally, 67.3% had an overall average decay error of less than 3 dB per time frame.

5. Conclusion

This paper presented a comprehensive DRIR denoising procedure, as well as its subsequent application to an extensive measurement database covering a large variety of reverberant spaces. The three essential aspects of the denoising framework were presented in detail and each step was illustrated by an example taken from the DRIR database results. Artefact suppression on ESM measurements was found to not only increase the final impulse response SNR but also regularize the EDR analysis, and the reverberation tail re-synthesis technique was shown to provide an arbitrary target SNR or dynamic range (limited only by the signal bit depth) with limited continuity error, whilst maintaining the diffuse reverberant field's spatial properties, as characterized by its diffuseness.

Several developments are currently underway to further improve the presented method, including the

detection and analysis of DRIR with multi-slope decays. Additionally, further study of the spatial properties of background noise is being explored in order to deal with localized noise sources. More generally, the question of denoising highly heterogeneous spaces that may not become diffuse before the DRIR reaches the noise floor (e.g. large outdoor spaces such as plazas or squares) is of great interest within the context of pursuing reverberation models that may extend beyond traditionally 'mixing' volumes.

ACKNOWLEDGEMENTS

The authors would like to thank Augustin Muller and Pedro Garcia-Velasquez for measuring the great majority of the DRIR database used in this work, as well as Franck Zagala for providing the foundations of the EDR analysis and SH-domain denoising algorithms.

REFERENCES

1. Rafaely, B. Analysis and Design of Spherical Microphone Arrays, *IEEE Trans on Speech and Audio Proc*, **13** (1), 135–143, (2005).
2. Noisternig, M., Zotter, F. and Katz, B. F. G., (2011), Reconstructing Sound Source Directivity in Virtual Acoustic Environments. *Principles and Applications of Spatial Hearing*, pp. 357–373, World Scientific Publishing.
3. Daniel, J. and Moreau, S. Further Study of Sound Field Coding with Higher Order Ambisonics, *Proc of 116th Audio Eng Soc Conv*, Berlin, Germany, (2004).
4. Schroeder, M. R. Natural-Sounding Artificial Reverberation, *J Audio Eng Soc*, **10** (3), 219–223, (1962).
5. Polack, J.-D., *La transmission de l'énergie sonore dans les salles*, Ph.D. thesis, Université du Maine, (1988).
6. Jot, J.-M., Cerveau, L. and Warusfel, O. Analysis and Synthesis of Room Reverberation Based on a Statistical Time-Frequency Model, *Proc of 103rd Audio Eng Soc Conv*, New York, U.S.A., (1997).
7. Noisternig, M., Carpentier, T., Szpruch, T. and Warusfel, O. Denoising of Directional Room Impulse Responses Measured with Spherical Microphone Arrays, *Proc of 40th Annual German Congress on Acoustics (DAGA)*, Oldenburg, Germany, pp. 600–601, (2014).
8. Guski, M. and Vorländer, M. Comparison of Noise Compensation Methods for Room Acoustic Impulse Response Evaluations, *Acta Acustica United with Acustica*, **100** (2), 320–327, (2014).
9. Carpentier, T., Szpruch, T., Noisternig, M. and Warusfel, O. Parametric Control of Convolution-Based Room Simulators, *Proc of 2013 Int Symp on Room Acoustics*, Toronto, Canada, (2013).
10. Farina, A. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique, *Proc of 108th Audio Eng Soc Conv*, Paris, France, (2000).
11. Prasad, D. K., Leung, M. K., Quek, C. and Cho, S. Y. A Novel Framework for Making Dominant Point Detection Methods Non-Parametric, *Image and Vision Computing*, **30** (12), 843–859, (2012).
12. Jarrett, D. P., Thiergart, O., Habets, E. A. P. and Naylor, P. A. Coherence-Based Diffuseness Estimation in the Spherical Harmonic Domain, *Proc of IEEE Conv of Electrical & Electronics Eng*, Eilat, Israel, (2012).
13. Abel, J. S. and Huang, P. A Simple, Robust Measure of Reverberation Echo Density, *Proc of 121st Audio Eng Soc Conv*, San Francisco, U.S.A., (2006).
14. Defrance, G. and Polack, J.-D. Measuring the Mixing Time in Auditoria, *Acoustics'08*, Paris, France, (2008).
15. Epain, N. and Jin, C. T. Spherical Harmonic Signal Covariance and Sound Field Diffuseness, *IEEE/ACM Trans on Audio, Speech, and Language Proc*, **24** (10), 1796–1807, (2016).
16. Merimaa, J. and Pulkki, V. Spatial Impulse Response Rendering I: Analysis and Synthesis, *J Audio Eng Soc*, **53** (12), 1115–1127, (2005).