Autoregressive Moving Average Modeling of Late Reverberation in the Frequency Domain

Simon Leglaive, Roland Badeau, Gaël Richard

LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France.

Abstract—In this paper, the late part of a room response is modeled in the frequency domain as a complex Gaussian random process. The autocovariance function (ACVF) and power spectral density (PSD) are theoretically defined from the exponential decay of the late reverberation power. Furthermore we show that the ACVF and PSD are accurately parametrized by an autoregressive moving average (ARMA) model. This leads to a new generative model of late reverberation in the frequency domain. The ARMA parameters are easily estimated from the theoretical ACVF. The statistical characterization is consistent with empirical results on simulated and real data. This model could be used to incorporate priors in audio source separation and dereverberation.

Index Terms—Statistical room acoustics, late reverberation, Gaussian random process, autoregressive moving average model.

I. Introduction

When an audio source is recorded in a reverberant environment, the signal at the microphone is a filtered version of the source signal plus a possible additive noise. The filter involved in this process is called a Room Impulse Response (RIR) and represents the propagation from the source to the microphone. In the case of a single emitting source, estimating the source signal is called dereverberation. When several sources play simultaneously, the task is referred to as source separation. A large number of methods are based on probabilistic frameworks. One possible approach is to treat the acoustic parameters as being deterministic so that they are estimated in a maximum likelihood sense [1], [2]. Another way of doing is to incorporate priors on the acoustic parameters, leading to a maximum a posteriori estimation or a full Bayesian treatment [3, ch. 8]. In audio source separation, some recent approaches have shown that improved separation results can be obtained by taking probabilistic priors on RIRs into account. In [4] the authors proposed a prior motivated by statistical room acoustics in order to represent the spatial correlations between RIRs at different microphones. In [5], [6] an autoregressive prior in the frequency domain is proposed to model the direct path and the first echoes of the RIRs.

In statistical room acoustics, the late part of a Room Frequency Response (RFR) can be considered as a complex Gaussian random process. In [7] Schroeder expressed the theoretical autocovariance function (ACVF) of this process
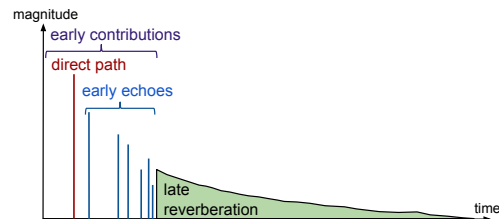
Fig. 1: Schematic illustration of a room impulse response.

from the time-domain exponential decay of the reverberation power. However Schroeder neglected the influence of the direct path and the early echoes, resulting in a mismatch between theoretical and empirical ACVF, as observed in [8]. In this paper, we consider the fact that the exponential decay is only valid for late reverberation. We thus obtain a more accurate match between the new expression of the theoretical ACVF and the experiments. Furthermore we show that an autoregressive moving average (ARMA) modeling of the late part of the RFR leads to an accurate parametrization of the ACVF and Power Spectral Density (PSD). Similar parametric modeling in the frequency domain has been applied to wireless channels in [9], but in a data-dependent way. From the ARMA model it will be possible to define a prior distribution on the late part of the RFR, which could be used for audio source separation in the same way as in [6].

We first present in section II the statistical late reverberation model and compare the theoretical results with experiments on simulated and real RIRs. Then we describe in section III the ARMA model for the late part of the RFR, where the parameters are easily estimated from the theoretical ACVF, without the need of any data. Experimental results show the validity of this ARMA model. We finally conclude in section IV on the proposed model invoking its suitability for audio source separation.

II. Statistical late reverberation model

An RIR has a well defined structure. As represented in Fig. 1 we distinguish three regions: firstly the direct path between the source and the microphone, then the early echoes, coming from the first reflections on the walls of the room, and finally the late reverberation where the sound is considered to be diffuse. Diffusion means that the energy density is uniformly distributed across the room and the energy at one point is uniformly distributed over all directions [10]. The direct path

and early echoes are usually considered as being deterministic, while the characterization of late reverberation involves the use of statistical room acoustics. In this paper, we group the direct path and the early echoes under the term "early contributions". Let $h(t) = h_e(t) + h_l(t)$ denote the whole RIR, and $h_e(t)$, $h_l(t)$ the early and late parts respectively. The transition instant between early reflections and late reverberation is referred to as the mixing time and can be defined according to the volume V of the room in m^3 [11], [12]:

$$t_0 = \lfloor C_0 \sqrt{V} f_s \rceil \text{ samples,} \tag{1}$$

with C_0 a normalization constant, $\lfloor \cdot \rceil$ the floor function and f_s the sampling rate in Hz. Different values of C_0 have been proposed in the literature [12], we choose $C_0 = 2 \times 10^{-3}$. The RFR is defined as $H(k) = H_e(k) + H_l(k)$ where for $k = 0, ..., T-1$:

$$H_{(\cdot)}(k) = \mathcal{F}_T\{h_{(\cdot)}(t)\} = \sum_{t=0}^{T-1} h_{(\cdot)}(t) e^{-j2\pi kt/T}. \tag{2}$$

Equation (2) is the T-point Discrete Fourier Transform (DFT) of signal $h_{(\cdot)}(t)$. We also define the inverse DFT by:

$$h_{(\cdot)}(t) = \mathcal{F}_T^{-1}\{H_{(\cdot)}(k)\} = \frac{1}{T} \sum_{k=0}^{T-1} H_{(\cdot)}(k) e^{j2\pi kt/T}. \tag{3}$$

A. Late reverberation model

The theory of statistical room acoustics for diffuse sound fields can be applied in the frequency domain if (a) the dimensions of the room are large compared to the wavelength of the source signal, (b) the average spacing of the room resonance frequencies are smaller than one third of the modes' bandwidth and (c) the source and microphone are located at least at half a wavelength from the walls [7], [8], [13]. Condition (b) is fulfilled above Schroeder's frequency:

$$f_{sch} = C_1 \sqrt{\frac{T_{60}}{V}} \text{ Hz,} \tag{4}$$

with C_1 a normalization constant approximately equal to 2000. T_{60} is the reverberation time in seconds defined as the time it takes for the sound energy to decrease by 60 dB after extinction of the source. We assume in the following that these conditions are fulfilled. It is well known that for the diffuse part of an RIR the power decays exponentially [11]. Following this observation we define the Power Temporal Profile (PTP) of the RIR for $t = 0, ..., T-1$:

$$\bar{h}_l(t) = \mathbb{E}[|h_l(t)|^2] = P_0^2 e^{-2t/\tau} \mathbb{1}_{t \geq t_0}(t), \tag{5}$$

where $\mathbb{1}_{t \in \mathcal{T}}(t)$ is the indicator function which equals 1 if $t \in \mathcal{T}$, 0 otherwise, P_0^2 is a constant related to the total power of late reverberation and τ is linked to the reverberation time T_{60} through:

$$\tau = \frac{T_{60} f_s}{3 \ln(10)} \text{ samples.} \tag{6}$$

It is important to mention that different realizations of the room responses can be interpreted as different observations at several source and microphone positions in the room. Expectations have thus to be understood as spatial averaging.

We now describe some important statistical results about the late part of the RFR. According to the theory of statistical room acoustics, the signal received by a microphone in a reverberant field combines the contributions of a large number of modes of the room with randomly distributed amplitudes and phases. Assuming that the modes are independent and identically distributed, the central limit theorem implies that the RFR is a random process whose real and imaginary parts are independent Gaussian processes with the same variance. Assuming a centered and wide sense stationary (WSS) process we define the ACVF $\gamma(m)$ and the PSD $\phi(t)$ of the RFR by:

$$\gamma(m) = \mathbb{E}[H_l(k) H_l(k-m)^*]; \tag{7}$$

$$\phi(t) = \frac{1}{T} \mathbb{E}[|\mathcal{F}_T\{H_l(k)\}|^2], \tag{8}$$

where $(\cdot)^*$ denotes complex conjugate. As a centered proper and WSS complex Gaussian random process the RFR is thus fully characterized by its ACVF or its PSD. We have to mention that as we work in discrete time and frequency, all signals are T-periodic. Strictly speaking the RFR is thus a T-periodic WSS random process. Moreover $\phi(t)$ has to be understood as a discretized PSD function. One can refer to [9] for a review of some properties of periodic random processes.

From equation (8) we can express the PSD of the process $\{H_l(k)\}_k$ according to the PTP defined in equation (5) (see appendix A):

$$\phi(t) = T \bar{h}_l(T - t). \tag{9}$$

The PTP is T-periodic so that $\bar{h}_l(T) = \bar{h}_l(0)$. Applying the Wiener-Khinchin theorem from this expression of the PSD we obtain the theoretical ACVF:

$$\gamma(m) = \mathcal{F}_T^{-1}\{\phi(t)\} = P_0^2 e^{-2T/\tau} \frac{1 - e^{(j2\pi m/T + 2/\tau)(T - t_0 + 1)}}{1 - e^{j2\pi m/T + 2/\tau}}. \tag{10}$$

P_0^2 is related to the variance $\sigma_{rev}^2 := \gamma(0)$ of the RFR by:

$$P_0^2 = \sigma_{rev}^2 e^{2T/\tau} \frac{1 - e^{2/\tau}}{1 - e^{2(T - t_0 + 1)/\tau}}. \tag{11}$$

Moreover it can be shown that (see appendix B and [14]):

$$\sigma_{rev}^2 = \frac{1 - \alpha}{\pi \alpha S}, \tag{12}$$

with α the average absorption coefficient (without dimension) and S the total wall area in m^2. The absorption coefficient can be computed using Norris-Eyring's formula [3, p. 24]:

$$\alpha = 1 - e^{-24 \ln(10)V/(cST_{60})}, \tag{13}$$

with c the speed of sound in m.s^{-1}.

B. Experimental validation

As already mentioned before, the authors in [8] compared Schroeder's theoretical ACVF [7] with empirical ACVFs computed from simulated and recorded RIRs. The results we present in this section have been obtained using an experimental procedure close to the one presented in [8]. From a
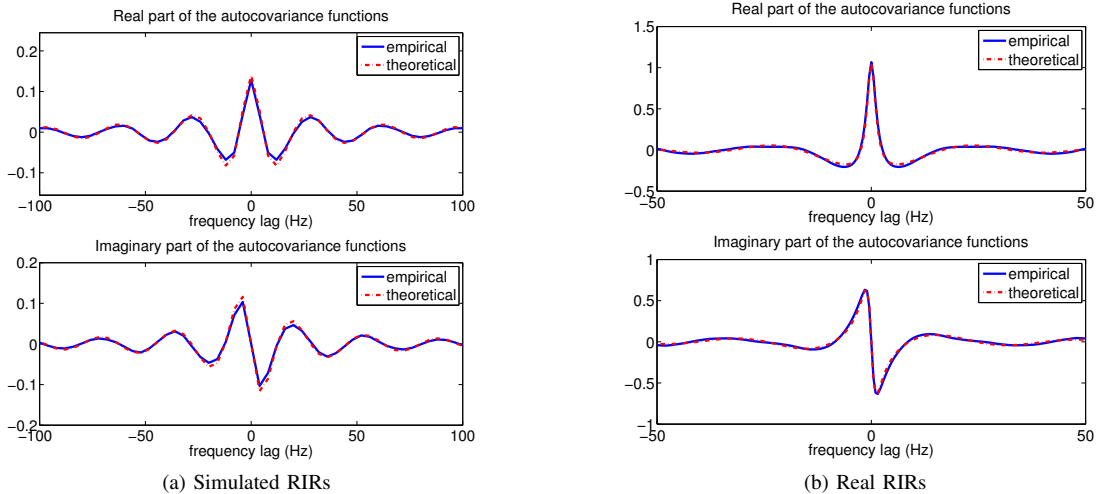
(a) Simulated RIRs

(b) Real RIRs

Fig. 2: Theoretical and empirical ACVFs.

given set of RIRs (simulated or recorded) at various locations in the room we first remove the early contributions by setting to zero the samples for $t < t_0$. For each realization $\tilde{h}_l^r(t)$ (one RIR for a particular location of source and microphone), we compute the RFR $\tilde{H}_l^r(k)$ and only keep the K samples in the interval $[f_{sch}, f_s/2]$ Hz. The empirical ACVF for one realization is obtained using the following estimator for $m = -K+1, ..., K-1$:

$$\tilde{\gamma}_r(m) = \frac{1}{K} \sum_{k=0}^{K-1-|m|} \tilde{H}_l^r(k)\tilde{H}_l^r(k-m)^*. \qquad (14)$$

Following this Monte Carlo simulation the final empirical ACVF is obtained by averaging over the whole set of realizations:

$$\hat{\gamma}(m) = \frac{1}{R} \sum_{r=1}^{R} \tilde{\gamma}_r(m). \qquad (15)$$

We will compare this empirical ACVF to the theoretical one defined by equations (10)-(13).

During our experiments we first observed a mismatch between the data and the theoretical variance given in equation (12). Although widely used in the literature this sole expression is rarely compared with experiments. As mentioned in [15, ch. 7] the expression of the reverberant mean square pressure that leads to equation (12) (see appendix B) may be inaccurate depending on the room dimensions and the source to microphone distance. Moreover this formula is obtained by considering that the reverberation field is established after the first reflection of the direct path on a wall, which disagrees with the definition of late reverberation we gave before. This could explain the mismatch. However an empirically accurate correction consists in multiplying σ_{rev}^2 as defined in equation (12) by a constant $C_{rev} = 75$. In the following this correction will always be applied. Moreover we can mention that the variance is only a scale factor in the ACVF and PSD. The correlations between the samples of a random process can

also be defined according to the normalized autocorrelation function which does not depend on the variance of the process.

We simulated 196 RIRs from the image source method using the Roomsimove toolbox[1]. The room was a 10 x 6.6 x 3 m shoebox with reverberation time $T_{60} = 250$ ms. The source position remained fixed while the microphone was moved over 196 positions. The results for simulated RIRs are presented in Fig. 2a. We observe a good fit between the theoretical expression in equation (10) and the empirical ACVF.

For the experiment on real RIRs we used 130 responses from the C4DM (Center for Digital Music) database [16]. The RIRs were recorded in a room of size 7.5 x 9 x 3.5 m with reverberation time $T_{60} = 1.8$ s. The source position was again fixed while the microphone was moved. Once again we observe a good fit between theory and experiments in Fig. 2b.

For both simulated and real RIRs, the theoretical expression of the ACVF we proposed in equation (10) better agree with experimental results than the one Schroeder proposed in [7] and that was further investigated in [8].

III. AUTOREGRESSIVE MOVING AVERAGE MODEL

We have seen in the previous section that the statistical properties of the late part of an RFR are fully summarized by the ACVF defined in equation (10). This ACVF only depends on the reverberation time, the volume of the room and the total wall area. In this section we will show that an ARMA model in the frequency domain is able to accurately fit the theoretical ACVF and PSD we defined in section II-A.

A. ARMA model and estimation

We define the following ARMA(P, Q) model for the late part of the RFR $\{H_l(k)\}_k$:

$$\Phi(L)H_l(k) = \Theta(L)\epsilon(k), \qquad (16)$$

---

[1]http://www.irisa.fr/metiss/members/evincent/Roomsimove.zip

where $\Phi(L) = \sum_{p=0}^{P} \varphi_p L^p$, $\Theta(L) = \sum_{q=0}^{Q} \theta_q L^q$ with $\varphi_0 = \theta_0 = 1$ and L is the lag operator, i.e. $LH_l(k) = H_l(k-1)$. $\epsilon(k)$ is a proper complex white Gaussian noise of variance σ_ϵ^2 for $k \in [0, ..., T-1]$ and is extended by T-periodicity elsewhere. $H_l(k)$ is the output of a linear filter with transfer function $\Psi(z^{-1}) = \Theta(z^{-1})/\Phi(z^{-1})$. The polynomials $\Theta(z^{-1})$ and $\Phi(z^{-1})$ are assumed to have their zeros within the unit circle so that $\Psi(z^{-1})$ is the transfer function of an invertible causal and stable filter. From this ARMA model we can write the following parametric representations of the PSD $\phi(t)$ and ACVF $\gamma(m)$:

$$\phi(t) = \sigma_\epsilon^2 \frac{\left| \mathcal{F}_T\left\{ \{\theta_q\}_{q=0,...,Q} \right\} \right|^2}{\left| \mathcal{F}_T\left\{ \{\varphi_p\}_{p=0,...,P} \right\} \right|^2}, \tag{17}$$

$$\sum_{p=0}^{P} \varphi_p \gamma(m-p) = \begin{cases} \sigma_\epsilon^2 \sum_{q=\underline{m}}^{Q} \theta_q \psi_{q-\underline{m}}^* & \text{if } 0 \le \underline{m} \le Q \\ 0 & \text{if } \underline{m} > Q, \end{cases} \tag{18}$$

where $\Psi(z^{-1}) = \sum_{k=0}^{+\infty} \psi_k z^{-k}$ and $\underline{m} = m\,(\mathrm{mod}\ T)$ such that $\underline{m} \in [0, ..., T-1]$. For equation (18) to be valid we have to assume that there is no overlap of the T-periodic ACVF $\gamma(m)$, i.e. $\psi_k = 0$ for $k > \lfloor T/2 \rfloor$.

The ARMA parameters are estimated from the sole knowledge of the theoretical ACVF given by (10), without need of any data. The procedure is fully described in [17, ch. 2] and we recall it here briefly. The AR parameters $\{\varphi_p\}_{p=1}^{P}$ are first estimated by solving the *modified Yule-Walker equations* defined from equation (18) for $m = Q+1, ..., Q+P$. From the estimated AR parameters we then define the ACVF $\gamma'(m) = \Phi^*(L^{-1})\Phi(L)\gamma(m)$ where $\Phi^*(L^{-1}) = \sum_{p=0}^{P} \varphi_p^* L^{-p}$. This is the ACVF of an MA process whose coefficients and noise variance are $\{\theta_q\}_{q=1}^{Q}$ and σ_ϵ^2 respectively, as introduced in equation (16). We approximate this $\mathrm{MA}(Q)$ process by a high-order $\mathrm{AR}(L)$ with $L = 10Q$. From $\gamma'(m)$ we compute the parameters of this high-order AR model by solving the standard Yule-Walker equations. We obtain the estimate of σ_ϵ^2 and we denote by $\{\varphi_p'\}_{p=1}^{L}$ the resulting AR coefficients. Finally, to obtain the MA parameters $\{\theta_q\}_{q=1}^{Q}$, we fit an $\mathrm{AR}(Q)$ model on the sequence $\{1, \varphi_1', ..., \varphi_L'\}$ again by solving the Yule-Walker equations.

B. Experimental validation

We applied this procedure to compute the ARMA parameters from the theoretical ACVF defined by equation (10). We used the same room parameters and reverberation time (250 ms) as presented in section II-B for the simulated RIRs. We empirically chose an AR order $P = 7$ while the MA order is set as $Q = 2$ by observing the support of the ACVF $\gamma'(m)$ defined in the previous section. Once the ARMA parameters are estimated, the PSD can be computed using equation (17), and the ACVF is easily obtained by inverse DFT of the PSD. The results are shown in Fig. 3. We observe a very good fit of the ARMA model on both theoretical ACVF and PSD. The PSDs are represented in decibels so we can observe that the exponential decay is well verified.
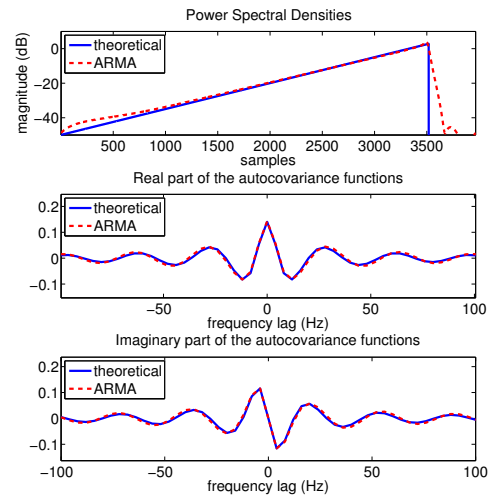

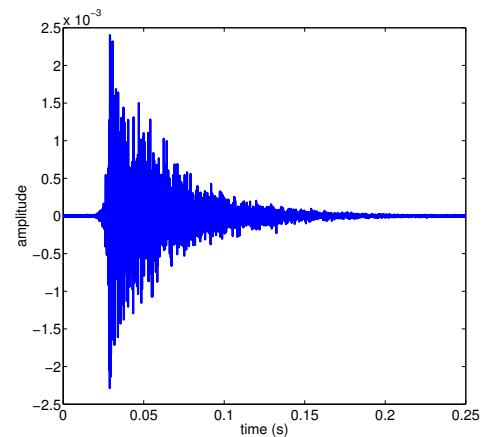Fig. 3: ARMA(7,2) modeling of the PSD and ACVF.

Fig. 4: Synthesized late reverberation from ARMA(7,2) filtering in frequency domain of a complex white Gaussian noise.

As a final way of assessing the validity of the whole model (theoretical ACVF, PSD and ARMA parametrization), we synthesize an artificial late reverberation RIR. Using the estimated parameters, the late RFR is obtained by ARMA filtering in the frequency domain of a complex white Gaussian noise. To obtain the late part of the RFR we only keep the samples within the frequency range $[f_{sch}, f_s/2]$ Hz. The RFR values below Schroeder's frequency are fixed to zero, but we have to mention that the results are similar by extending the model to $[0, f_s/2]$ Hz. The time domain late RIR represented in Fig. 4 is finally obtained using the hermitian symmetry of the RFR and the inverse DFT. We can clearly observe the exponential decay of the late reverberation starting from the mixing time, as described by the PTP in equation (5). The estimated reverberation time of this RIR obtained by using Schroeder's backward integration method [18] is 280 ms which is close to the targeted value of 250 ms.

IV. CONCLUSIONS

In this paper, we refined the expression of Schroeder's ACVF [7] for the late part of an RFR by considering that late

reverberation only appears after early contributions. We experimentally showed the accuracy of the obtained ACVF. We also proposed a new parametric model of the ACVF and PSD, leading to a generative ARMA model of late reverberation in the frequency domain. From this ARMA representation in equation (16) it is possible to define a prior distribution on the late part of an RFR. In future work this prior will be applied to source separation to constrain the estimation of the mixing filters, following the same procedure as in [6]. Refinement of the model includes the investigation of a more accurate and justified expression of the variance σ_{rev}^2. This model could also be extended for frequency dependent reverberation times, by using different ARMA models for different frequency bands.

APPENDIX A

Proof of equation (9): From straightforward calculations one can show that:

$$\mathcal{F}_T\{h_l^*(T-1-t)\} = e^{j2\pi k/T}H_l^*(k). \tag{19}$$

Applying the inverse DFT, noted $\mathcal{F}_T^{-1}\{\cdot\}$, to equation (19) and using the circular convolution and shift theorems we have:

$$h_l^*(T-1-t) = \delta(t+1) \circledast \mathcal{F}_T^{-1}\{H_l^*(k)\}, \tag{20}$$

where $\circledast$ denotes the circular convolution, and $\delta(t)$ the Dirac delta function. Using the fact that $\mathcal{F}_T^{-1}\{H_l^*(k)\} = \frac{1}{T}\mathcal{F}_T\{H_l(k)\}^*$ in equation (20) we obtain:

$$\mathcal{F}_T\{H_l(k)\} = Th_l(T-t). \tag{21}$$

From equations (21) and (8) we finally obtain (9).

APPENDIX B

Proof of equation (12): Let $f = kf_s/T$. The relation between the RFR $H_l(f)$ and the complex sound pressure (CSP) $P_l(f)$ due to late reverberation at some point in a room for a punctual source is given by [19, p. 311], [14]:

$$P_l(f) = -j2\pi f\rho_0 Q(f)H_l(f), \tag{22}$$

with ρ_0 the air density in kg.m^{-3} and $Q(f)$ the volume velocity of the source in m^3.s^{-1}. Using the previous equality we can link the variances of the CSP and of the RFR:

$$\sigma_{rev}^2 = \mathbb{E}[|H_l(f)|^2] = \frac{\mathbb{E}[|P_l(f)|^2]}{(2\pi f)^2\rho_0^2|Q(f)|^2}, \tag{23}$$

where $|Q(f)|^2$ is expressed as [19, p. 311], [14]:

$$|Q(f)|^2 = \frac{W_s c}{\pi f^2\rho_0}. \tag{24}$$

c is the speed of sound in m.s^{-1} and W_s is the power output of the sound source in Watts.

We now have to express the reverberant field mean square pressure $\mathbb{E}[|P_l(f)|^2]$. We denote by α the average absorption coefficient (without dimension) and S the total wall area in m^2. In a reverberant field the intensity is constant everywhere and is related to the mean square pressure by [19, p. 581]:

$$I_{rev} = \frac{\mathbb{E}[|P_l(f)|^2]}{4\rho_0 c}. \tag{25}$$

For a diffuse sound field in steady-state conditions, the reverberant power supplied by the source must be equal to the power absorbed by the walls [15, ch. 7]. The reverberant power can be defined as the sound power from the source that remains after the first reflection, $(1-\alpha)W_s$, and the absorbed power is αSI_{rev}. From this condition we have:

$$(1-\alpha)W_s = \alpha S\frac{\mathbb{E}[|P_l(f)|^2]}{4\rho_0 c}, \tag{26}$$

which leads to the reverberant field mean square pressure:

$$\mathbb{E}[|P_l(f)|^2] = \frac{4\rho_0 c(1-\alpha)W_s}{\alpha S}. \tag{27}$$

Substituting equations (27) and (24) into (23) we obtain the result given by equation (12).

REFERENCES

[1] B. Schwartz, S. Gannot, and E. A. Habets, "Multi-microphone speech dereverberation using expectation-maximization and kalman smoothing," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Marakech, Morocco, 2013, pp. 1–5.

[2] A. Ozerov and C. Févotte, "Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation," *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 18, no. 3, pp. 550–563, 2010.

[3] P. A. Naylor and N. D. Gaubitch, *Speech dereverberation*. Springer Science & Business Media, 2010.

[4] N. Q. Duong, E. Vincent, and R. Gribonval, "Spatial location priors for gaussian model based reverberant audio source separation," *EURASIP J. Adv. Signal Process.*, vol. 2013, 2013.

[5] S. Leglaive, R. Badeau, and G. Richard, "A priori probabiliste anéchoïque pour la séparation sous-déterminée de sources sonores en milieu réverbérant," in *Colloque GRETSI*, Lyon, France, 2015.

[6] ——, "Multichannel audio source separation with probabilistic reverberation modeling," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, 2015, p. 5.

[7] M. R. Schroeder, "Frequency-correlation functions of frequency responses in rooms," *J. Acoust. Soc. Am.*, vol. 34, no. 12, pp. 1819–1823, 1962.

[8] T. Gustafsson, B. D. Rao, and M. Trivedi, "Source localization in reverberant environments: Modeling and statistical analysis," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 791–803, 2003.

[9] G. Gu, X. Gao, J. He, and M. Naraghi-Pour, "Parametric modeling of wideband and ultrawideband channels in frequency domain," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1600–1612, 2007.

[10] T. Schultz, "Diffusion in reverberation rooms," *J. Sound Vibration*, vol. 16, no. 1, pp. 17–28, 1971.

[11] J.-M. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Audio Engin. Soc. Conv. 103*, 1997.

[12] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses," in *Audio Engin. Soc. Conv. 128*, 2010.

[13] E. A. Habets, "Speech dereverberation using statistical reverberation models," in *Speech Dereverberation*. Springer, 2010, pp. 57–93.

[14] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment: Robustness results," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 3, pp. 311–319, 2000.

[15] D. A. Bies and C. H. Hansen, *Engineering noise control: theory and practice*. CRC press, 2009.

[16] R. Stewart and M. B. Sandler, "Database of omnidirectional and b-format room impulse responses." in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, Texas, 2010, pp. 165–168.

[17] S. Kay, "Spectral estimation," *Advanced topics in signal processing*, pp. 58–122, 1988.

[18] M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Am.*, vol. 37, no. 3, pp. 409–412, 1965.

[19] P. M. Morse and K. U. Ingard, *Theoretical acoustics*. Princeton university press, 1968.