On the topic of frequency dependent exponential decay matrices and Lie groups

Achille Aknin, Roland Badeau

STOCHASTIC REVERBERATION MODEL WITH A FREQUENCY DEPENDENT ATTENUATION

*Achille Aknin, Roland Badeau**

LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France
achille.aknin@telecom-paris.fr, roland.badeau@telecom-paris.fr

ABSTRACT

In various audio signal processing applications, such as source separation and dereverberation, accurate mathematical modeling of both source signals and room reverberation is needed to properly describe the audio data. In a previous paper, we introduced a stochastic room impulse response model based on the image source principle, and we proposed an expectation-maximization algorithm that was able to efficiently estimate the model parameters in various experimental settings. This paper aims to extend the model in order to account for the dependency of the exponential decay over frequency, due to the walls usually absorbing less energy at low frequencies than at high frequencies. Our experimental results show that this refinement of the model is able to generate realistic room impulse responses, that are perceptively very close to the original ones.

***Index Terms*—** Reverberation, room impulse response, probabilistic modeling, expectation-maximization algorithm, artificial reverberation.

1. INTRODUCTION

Audio signal processing applications such as source separation and dereverberation often involve the modeling of room impulse responses (RIRs): the observed signal $x(t)$ is usually modeled as the sum of convolution products of acoustic source signals $y_i(t)$ with the corresponding RIRs $h_i(t)$ (where index i refers to the different acoustic sources) possibly corrupted by additive noise $n(t)$. The joint estimation of h_i and y_i requires accurate mathematical models for both the source signals and the RIRs.

The modeling of source signals has been the main focus in numerous papers, using various approaches including autoregressive (AR) models [1], sinusoids plus noise models [2], Non-negative Matrix Factorization (NMF) [3, 4], heavy-tailed stochastic models [5], and Deep Neural Networks (DNN) [6, Chap. 7].

As for the RIR, one standard model [7, 8] is a Gaussian process with independent samples and exponentially decreasing variance: $\forall t > 0$,

$$\begin{aligned} h(t) &= e^{-at} b(t) \\ b(t) &\sim \mathcal{N}(0, \sigma^2). \end{aligned} \tag{1}$$

with $a, \sigma > 0$. This model is a good approximation of the late part of reverberation, but it does not accurately represent the early reflections. However these early reflections, where the energy is mostly concentrated, are perceptually very important. Other stochastic reverberation models include the use of a spatial covariance matrix [9], complex Gaussian latent variables [10], and a more general Student's t model [5].

*This work was partly supported by the research program ASCETE (ANR-19-CE48-0001) funded by ANR, the French State agency for research.

In a previous publication [11], we introduced a new reverberation model based on the theoretical work of [12, 13]. We present in this paper an extension of this model that accounts for the frequency dependence of the exponential decrease of the energy in any RIR. This frequency dependence is critical for realism, perceptive credibility and theoretical exactness, but it is difficult to implement and parametrize.

This paper is organised as follows: in Section 2 we recall the previous model and explain how it is extended, Section 3 introduces a parametric estimation algorithm for the new model and Section 4 presents some experimental results with both synthetic and real RIRs. Finally, we conclude in Section 5 with perspectives for future work.

Notation

- $\mathcal{N}(\mu, R)$: multivariate real Gaussian distribution of mean vector μ and covariance matrix R,
- M^T: transpose of matrix M,
- I: identity matrix,
- $\mathrm{Tr}(M)$: trace of matrix M,
- $\|.\|_2$: Euclidean vector norm,
- $\mathcal{F}_x(f)$: Fourier transform of signal x at frequency f,
- $\hat{X}$: discrete-time Fourier transform of signal x,
- $D_r(a,b) = \frac{1}{n}\sum_{i=0}^{n-1} \frac{|a(i)-b(i)|}{|a(i)|}$: relative distance between vectors a and b of size n.

2. FREQUENCY DEPENDENT REVERBERATION MODEL

2.1. Previous model

We recall in this section the model used in our previous article [11], as a starting point for the contribution of this paper. This model represents a RIR h of length L_h as $\forall u \in [\![0, L_h - 1]\!]$,

$$\begin{aligned} h(u) &= b(u) + w(u) \\ b(u) &= (G^{-1}E^{-1}\pi)(u) \end{aligned} \tag{2}$$

where:

- $w(u) \sim \mathcal{N}(0, \sigma^2)$ is white Gaussian noise corresponding to the measurement error of h,
- E is an $L_h \times L_h$ diagonal matrix of coefficients $E(u,u) = e^{au}, \forall u \in [\![0, L_h - 1]\!]$; $a > 0$ is a fixed exponential decrease parameter accounting for the absorption of the walls when sound is reflected on them,

- matrix G^{-1} implements an AR filter of order L_g that represents other convolutive effects such as the inner response of the microphone; more precisely, G is an $L_h \times L_h$ lower triangular Toeplitz matrix whose entries contain a Finite Impulse Response (FIR) filter g of length L_g such that $g(0) = 1$,
- π is a multivariate random vector containing i.i.d. random variables: $\forall v \in [\![0, L_h - 1]\!]$, $\pi(v) \sim \mathcal{N}(0, \lambda)$ where $\lambda > 0$ is the variance parameter.

This model is a good approximation for simple synthetic RIRs but has one main limitation in the case of real RIRs: the exponential decay encoded in matrix E^{-1} does not depend on the frequency of the input signal. Therefore this model leads to an estimation of the reverberation time (T_{60}) that does not depend on the frequency either: $T_{60} = \frac{3\ln(10)}{f_s a}$ (in seconds), where f_s is the sampling frequency of the modeled RIR. However, in most rooms, the energy is absorbed faster at high frequencies than at low frequencies.

2.2. Frequency dependent exponential decrease

In this section, we tackle the issue of implementing a frequency dependent exponential decrease instead of the simple diagonal matrix E of constant parameter a. One possible solution to this problem would be to partition the RIR h into several frequency sub-bands and process each sub-band i separately with one specific parameter a_i, but this approach has some limitations: we would need to choose the optimal frequency partition, which could be different depending on the RIR, the resulting model would involve many parameters, and sub-band modeling could create discontinuities at the edge between two successive sub-bands.

Instead, we define the set $\mathcal{P}$ of $L_h \times L_h$ matrices, where $P \in \mathcal{P}$ if and only if there exists a causal impulse response $p(v)$ such that:

$$P = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & p(0) & 0 & \dots & \vdots \\ 0 & p(1) & p^{*2}(0) & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & p(L_h-2) & p^{*2}(L_h-3) & & p^{*(L_h-1)}(0) \end{pmatrix}$$

where $p(0) \neq 0$ and p^{*j} is p convolved by itself j times (if $j = 0$, then $p^{*j} = \delta$ is the Kronecker delta function).

If we consider such a matrix P and extract the submatrix P_1 obtained by removing the first line and first column, we can see that $P_1 = TP'$, where T is the lower-triangular Toeplitz matrix corresponding to filter p, and P' is the $(L_h - 1) \times (L_h - 1)$ matrix structured as P and constructed from filter p. By induction this means that applying the matrix P at timestamp u is equivalent to applying a smaller matrix P'' at timestamp 0 and then filtering the result by p^{*u}.

This effect can be visualized in Fig. 1: we show the log-spectrogram of Px where $P \in \mathcal{P}$ and x is white Gaussian noise of variance 1, next to the frequency response of the filter p used in matrix P. We can see that at a given frequency f, the higher $\mathcal{F}_p(f)$, the slower the corresponding exponential decay effect.

Note that if p is just an impulse of amplitude e^a, then matrix P is the same as the previously used diagonal matrix E, which corresponds to the fact that $\mathcal{F}_p(f)$ is a constant: the effect of P does not depend on the frequency in this particular case.

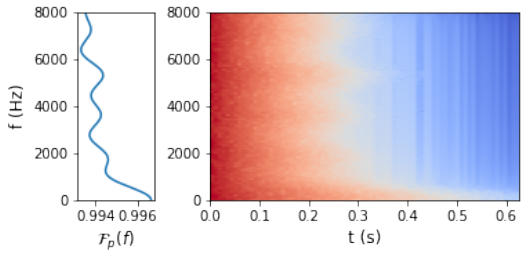


Figure 1: Side by side comparison of the frequency response of p (left) and the log-spectrogram of Px (right) where $P \in \mathcal{P}$ is the matrix corresponding to filter p and x is white Gaussian noise of variance 1.

To ensure proper convergence, we choose to keep the AR parametrization we already used for filter g and we apply the matrix P^{-1} (which is also in $\mathcal{P}$, because $\mathcal{P}$ is a Lie group[1]) instead of P itself. We limit the length of p to L_p. We also set $p(0) = 1$, while keeping the previous matrix E: E^{-1} will then encode the average exponential decrease over all frequencies, while P^{-1} will adjust the effect of E^{-1} depending on the frequency. This leads to the following final model:

$$\begin{aligned} h(u) &= b(u) + w(u), \\ b(u) &= (G^{-1}P^{-1}E^{-1}\pi)(u). \end{aligned} \tag{3}$$

The combined effects of P and E in the model result in a frequency dependent reverberation time $T_{60}(f)$ (in seconds):

$$T_{60}(f) = \frac{3\ln(10)}{f_s(a + \ln(|\mathcal{F}_p(f)|))}. \tag{4}$$

We can note that Eyring's formula [14], along with this equation, allows us to generate a realistic filter p by using the frequency dependent absorption parameters of a known material.

3. ESTIMATING THE PARAMETERS

The model defined in (3) includes the observed variable $h(u)$, the latent variable $b(u)$[2] and the parameters a, σ^2, λ, g (FIR filter of length L_g), and p (FIR filter of length L_p).

Since we only apply linear functions to π and w, the a priori distributions are Gaussian, thus we know that the a posteriori distribution of b given h is also Gaussian:

$$b \mid h, g, p, a, \lambda, \sigma^2 \sim \mathcal{N}(\mu, R). \tag{5}$$

Given this model, an Expectation-Maximization (EM) algorithm [15] can be used to jointly estimate the parameters g, p, a, λ and σ^2 and the latent variable $b(u)$ given an observed RIR $h(u)$ in the maximum likelihood sense. The algorithm alternates two steps:

- expectation step (E-step): computing the a posteriori distribution of b given h (in this case we only need its mean vector μ and covariance matrix R), given the current estimates of the parameters,

[1]The proof and some additional results are available on HAL: https://hal.archives-ouvertes.fr/hal-03298695/document.

[2]Note that we could equivalently define $w(u)$ or $\pi(u)$ as the latent variable.

- maximization step (M-step): maximizing (6) with respect to (w.r.t.) the parameters $\theta = (g, p, a, \lambda, \sigma^2)$ given the current estimate of the a posteriori distribution of b.

More specifically, the a posteriori expectation of the log-probability density function (PDF) of the joint distribution of observed and latent random variables is:

$$\begin{aligned}\mathcal{Q} =& \mathbb{E}_{\mathbb{P}(b|h,\theta)}\left[\ln \mathbb{P}(h, b \mid \theta)\right] \\ =& \mathbb{E}_{\mathbb{P}(b|h,\theta)}\left[\ln \mathbb{P}(b \mid \theta)\right] + \mathbb{E}_{\mathbb{P}(b|h,\theta)}\left[\ln \mathbb{P}(h \mid b, \theta)\right] \\ =& \tfrac{-L_h}{2}\ln(2\pi\lambda) - \tfrac{1}{2\lambda}\mathrm{Tr}(GPE\tilde{R}EP^TG^T) + \tfrac{L_h(L_h-1)}{2}a \\ & - \tfrac{L_h}{2}\ln(2\pi\sigma^2) - \tfrac{1}{2\sigma^2}\left[\|h-\mu\|_2^2 + \mathrm{Tr}(R)\right]\end{aligned} \tag{6}$$

with $\tilde{R} = R + \mu\mu^T$.

3.1. Expectation

For the expectation step, we can write

$$\ln(\mathbb{P}(h \mid \theta)) + \ln(\mathbb{P}(b \mid h, \theta)) = \ln(\mathbb{P}(b \mid \theta)) + \ln(\mathbb{P}(h \mid b, \theta))$$

and complete the square (i.e. identify linear and quadratic terms in b):

$$\frac{1}{2}(b-\mu)^T R^{-1}(b-\mu) = C - \frac{1}{2\lambda}\|GPEb\|_2^2 - \frac{1}{2\sigma^2}\|h-b\|_2^2$$

where C is a constant w.r.t. b. We then get:

$$\begin{aligned} R &= \lambda\sigma^2\left[\lambda I + \sigma^2 EP^TG^TGPE\right]^{-1}, \\ \mu &= \frac{Rh}{\sigma^2}. \end{aligned} \tag{7}$$

3.2. Maximization

As for the maximization step, the following updates directly maximize the a posteriori expectation of the joint log-PDF $\mathcal{Q}$ in (6).

Filter g: The optimal filter parameters are the unique solution of the following linear system: $\forall 0 < i < L_g$,

$$\sum_{j=0}^{L_g-1} g(j) \sum_{u=0}^{L_h-1} PE\tilde{R}EP^T(u-i, u-j) = 0. \tag{8}$$

Note that $g(0)$ is fixed to 1, so the solution to this equation is indeed unique (the system involves a positive definite matrix).

Variance parameter λ: The optimal parameter λ is:

$$\lambda = \frac{1}{L_h}\mathrm{Tr}\left(GPE\tilde{R}EP^TG^T\right). \tag{9}$$

Absorption parameter a: To update a, we use the fact that we can swap matrices in order to simplify the calculation: $GPE = E\tilde{G}\tilde{P}$ where $\tilde{g}(v) = g(v)e^{-av}$ for $0 \le v < L_g$ and $\tilde{p}(v) = p(v)e^{-av}$ for $0 \le v < L_p$.

Substituting the value of λ in the expression of $\mathcal{Q}$ in (6) and canceling the partial derivative w.r.t. a, we find:

$$\sum_{u=0}^{L_h-1} e^{2ua}\left(\tfrac{L_h-1}{2} - u\right)(\tilde{G}\tilde{P}\tilde{R}\tilde{P}^T\tilde{G}^T)(u,u) = 0. \tag{10}$$

This equation has no closed-form solution, but the solution is unique and we can use a dichotomy method to find the optimal value of a.

Frequency dependent exponential decrease filter p: To update p, we use the fact that we can also swap matrices $GP = P\tilde{G}$ with $\tilde{G} = P^{-1}GP$ and we maximize the log-probability $\mathcal{Q}$ w.r.t. p while considering that $\tilde{G}$ is constant w.r.t. p. Since we cannot find a closed-form solution that cancels the gradient w.r.t. vector p, we proceed to maximizing (6) using Newton's method [16], that involves computing the gradient vector and the Hessian matrix of $\mathcal{Q}$.

Computation of the gradient vector: The first order derivative of $\mathcal{Q}$ w.r.t. $p(i)$ for $0 < i < L_p$ is:

$$\frac{\partial\mathcal{Q}}{\partial p_i} = \frac{-1}{\lambda}\sum_{u=0}^{L_h-1} P^{\downarrow}D\tilde{G}E\tilde{R}E\tilde{G}^TP^T(u-i, u) \tag{11}$$

where D is the $L_h \times L_h$ diagonal matrix of coefficients $D(u,u) = u$, $P^{\downarrow}(u,v) = P(u-1, v-1)$ if $1 \le u, v < L_h$, and $P^{\downarrow}(u,v) = 0$ otherwise.

Computation of the Hessian matrix: The second order derivative of $\mathcal{Q}$ w.r.t. $p(i)$ and $p(j)$ for $0 < i, j < L_p$ is:

$$\begin{aligned}\frac{\partial^2\mathcal{Q}}{\partial p_j \partial p_i} =& \frac{-1}{\lambda}\sum_{u=0}^{L_h-1}\Big(P^{\downarrow\downarrow}D_2\tilde{G}E\tilde{R}E\tilde{G}^TP^T(u-i-j, u) \\ & + P^{\downarrow}D\tilde{G}E\tilde{R}E\tilde{G}^TD^TP^{\downarrow T}(u-i, u-j)\Big)\end{aligned} \tag{12}$$

where D_2 is the $L_h \times L_h$ diagonal matrix of coefficients $D(u,u) = (u-1)u$, $P^{\downarrow\downarrow}(u,v) = P(u-2, v-2)$ if $2 \le u, v < L_h$, and $P^{\downarrow\downarrow}(u,v) = 0$ otherwise.

White noise variance parameter σ^2:

The update of the white noise variance is:

$$\sigma^2 = \frac{1}{L_h}\left(\|h-\mu\|_2^2 + \mathrm{Tr}(R)\right). \tag{13}$$

3.3. Initialization of the parameters

An adequate initialization of the parameters is critical to make the EM algorithm converge faster (although our numerical simulations showed that the convergence to the optimal solution is not sensitive to initialization). Since we did not change the initialization part of the algorithm, the details can be found in [11]. The new parameter p can be initialized as $p(v) = 1$ if $v = 0$ and $p(v) = 0$ otherwise.

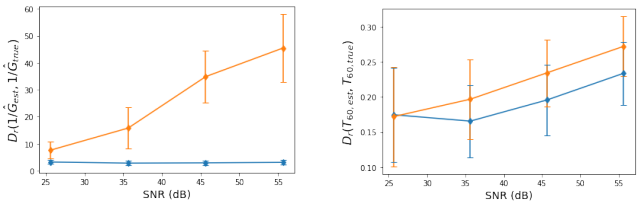
4. EXPERIMENTAL RESULTS

In this section, we present a few experimental results obtained by learning the model parameters on both synthetic and real RIRs. The experiments on synthetic RIRs allow us to compare the learned parameters to the true parameters, while accurately modeling real RIRs is a more challenging problem.

For both experimental setups, we empirically fixed the hyper-parameters $L_g = 21$ and $L_p = 10$ (in practice, the method is not very sensitive to these parameters' values).

4.1. Synthetic RIR

The synthetic RIRs we used were generated by Roomsimove [17]. Roomsimove is a MATLAB toolbox that simulates a parallelepipedic room, and allows us to specify the dimensions of the room, the frequency dependent absorption of the walls, the filter g, and the position of the source and microphone in the room. We manually added the noise $w(u)$ in order to account for the measurement error. We used the following parameters in Roomsimove:



(a) $D_r(1/\hat{G}_{est}, 1/\hat{G}_{true})$ with the estimated parameters from the new model (blue) and the estimation from the previous model (orange).

(b) $D_r(T_{60,est}, T_{60,true})$ with the estimated parameters from the new model (blue) and the estimation from the previous model (orange).

Figure 2: *Mean and standard deviation over 20 experiments of the estimation of the AR filter of coefficients g (left) and the frequency dependent T_{60} (right) to the true parameters for different SNRs.*

- $g_{true}^{*(-1)}$ is a combination of a high-pass filter with low cut-off frequency at 20 Hz, that is implemented as a recursive filter of order (2,2) (default settings of Roomsimove), and a FIR filter approaching the frequency response of an Audio-Technica ATM650 microphone[3],
- the room is of size $2 \times 3 \times 4$ (in meters),
- the sampling frequency is $f_s = 16000$ Hz,
- the sources and microphones in the room are omnidirectional,
- the frequency dependent $T_{60,true}$ can be computed from the frequency dependent absorption values of the walls and room dimensions, which are inputs of Roomsimove, through Eyring's formula [14].

Fig. 2 shows the mean and standard deviation over 20 experiments of the relative distance D_r, as defined in the notation paragraph of Section 1, between the estimation and the true value of the AR filter of coefficients g and the frequency dependent T_{60} for different Signal to Noise Ratios (SNR). We also included, as baselines, the same comparisons with the estimations of our previous model [11] on the same RIR, i.e. without the inclusion of matrix P: the corresponding T_{60} does not depend on frequency f.

We can see that, although the estimation is more difficult in the case of low SNR values, it is generally improved for both the parameter g and the frequency dependent T_{60}.

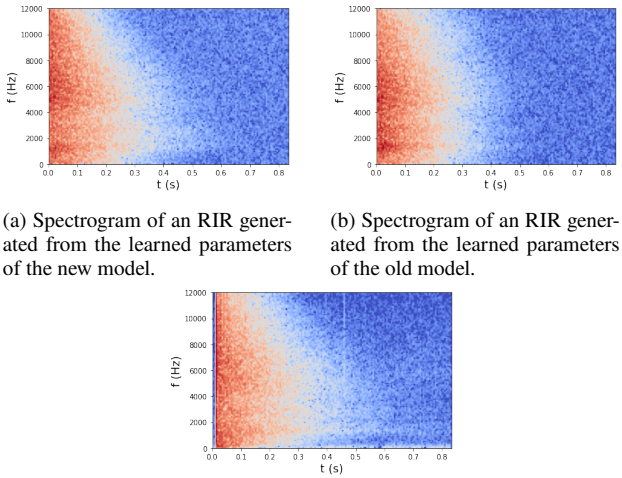
4.2. Real RIRs

The experiments on real RIRs were conducted on RIRs from the Mardy [18] dataset. This dataset was chosen for its relatively short reverberation time, allowing for faster execution of the EM algorithm.

The main drawback of the proposed algorithm being its computational time, we had to extract the first 8000 samples of the RIRs to learn the parameters, with a sampling frequency of $f_s = 24000$ Hz, losing some information in the process. Even with this trimming of the RIRs, one iteration takes up to 2 minutes on an AMD Ryzen 5 3600 3.6 GHz processor and we needed 200 iterations to reach a good estimation of the parameters.

With the estimated parameters, we then generated a synthetic RIR according to the stochastic model. Fig. 3 shows three spectrograms: in 3c, the spectrogram of an RIR from the Mardy database, in 3a, the spectrogram of an RIR generated from the new model parameters learned on the real RIR, and in 3b, the spectrogram of

[3]Based on http://recordinghacks.com/microphones/Audio-Technica/ATM650.



(a) Spectrogram of an RIR generated from the learned parameters of the new model.

(b) Spectrogram of an RIR generated from the learned parameters of the old model.

(c) Spectrogram of the true RIR.

Figure 3: Three spectrograms in decibels (dB) compared: the true RIR (bottom) and both generated RIRs with the new (left) and old (right) model.

an RIR generated from the previous model [11] parameters, not accounting for the frequency dependence of the exponential decay.

We can see that the frequency dependence of the exponential decay permits us to describe the true RIR more accurately, and that the new model manages to create a perceptually credible synthetic RIR that is very close to the original one. Indeed, the informal listening tests[4] that we conducted, by filtering an anechoic recording from the TSP dataset [19] with the true and estimated RIRs, confirmed this observation: as non-professional sound engineers, we were able to hear that the new model generated an RIR with much more perceptive quality than the previous model, and which is very close to the true RIR.

5. CONCLUSION

This paper extended previous works on stochastic reverberation modeling, and investigated the pertinence of introducing a frequency dependent exponential attenuation in order to better represent real RIRs. The implementation of an EM algorithm showed that this model is able to represent real RIRs with enough accuracy to generate perceptually credible synthetic RIRs, close to the original ones.

Our main future objective is to reduce the computational complexity of the estimation algorithm, which in turn would enhance the estimation itself since we would be able to process longer RIRs and perform more iterations of the algorithm. To this end, one possible approach would be to investigate the mathematical properties of the Lie group $\mathcal{P}$ and its corresponding Lie algebra, and to exploit a more efficient parametrization of matrix P.

Other directions to be explored include: conducting more experiments once the computational cost of the estimating algorithm will have been reduced; extending this monophonic model to a spatial model (i.e. modeling the statistical dependencies between sensors); and estimating the reverberation model from reverberated audio signals instead of the clean RIRs, for example in the context of dereverberation or source separation.

[4]The code used and a few audio examples are available on the GitHub repository https://github.com/Aknin/P_matrix_model

6. REFERENCES

[1] A. Hyvärinen, "A unifying model for blind separation of independent sources," *Signal Processing*, vol. 85, no. 7, pp. 1419–1427, 2005.

[2] X. Serra *et al.*, "Musical sound modeling with sinusoids plus noise," *Musical signal processing*, pp. 91–122, 1997.

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

[4] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 5, pp. 971–982, 2013.

[5] S. Leglaive, R. Badeau, and G. Richard, "Student's t source and mixing models for multichannel audio source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 6, pp. 1154–1168, 2018.

[6] J.-T. Chien, *Source separation and machine learning*. Academic Press, 2018.

[7] M. R. Schroeder, "Frequency-correlation functions of frequency responses in rooms," *The Journal of the Acoustical Society of America*, vol. 34, no. 12, pp. 1819–1823, 1962.

[8] J. A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 3, no. 2, pp. 13–28, 1979.

[9] N. Q. Duong, E. Vincent, and R. Gribonval, "Underdetermined reverberant audio source separation using a full-rank spatial covariance model," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 7, pp. 1830–1840, 2010.

[10] L. Girin and R. Badeau, "On the use of latent mixing filters in audio source separation," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2017, pp. 225–235.

[11] A. Aknin, T. Dupré, and R. Badeau, "Evaluation of a stochastic reverberation model based on the image source principle," in *International Conference on Digital Audio Effects*, 2020.

[12] R. Badeau, "Unified stochastic reverberation modeling," in *Proc. of 26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, Sept. 2018.

[13] ——, "Common mathematical framework for stochastic reverberation models," *The Journal of the Acoustical Society of America*, 2019, Special issue on room acoustics modeling and auralization.

[14] G. Millington, "A modified formula for reverberation," *The Journal of the Acoustical society of America*, vol. 4, no. 1A, pp. 69–82, 1932.

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[16] C. T. Kelley, *Solving nonlinear equations with Newton's method*. SIAM, 2003.

[17] E. Vincent and D. R. Campbell, "Roomsimove toolbox," 2008, GNU Public License https://members.loria.fr/EVincent/software-and-data.

[18] J. Y. Wen, N. D. Gaubitch, E. A. Habets, T. Myatt, and P. A. Naylor, "Evaluation of speech dereverberation algorithms using the MARDY database," in *Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, 2006.

[19] P. Kabal, "TSP speech database," McGill University, Montréal, Québec, Canada, Tech. Rep., 2002.