Autoregressive Moving Average Modeling of Late Reverberation in the Frequency Domain

Simon LEGLAIVE, Roland BADEAU, Gaël RICHARD
LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, Paris, France

European Signal Processing Conference (EUSIPCO), 2016

Introduction

▷ Bayesian approaches for audio source separation and dereverberation need probabilistic priors over room responses.

▷ **Statistical room acoustics**: The frequency response of late reverberation is a centered, proper and Wide Sense Stationary (WSS) Complex Gaussian Random Process (CGRP)[1].

▷ **Objective**: Statistical characterization and parametric modeling of this random process.

▷ **Main idea**: Time domain dynamics of late reverberation (exponential decay) induces specific frequency correlations [1].

▷ **Contributions**:

- ◇ Theoretical expressions of the Power Spectral Density (PSD) and Autocovariance Function (ACVF) given some room parameters.
- ◇ Autoregressive Moving Average (ARMA) parametrization of the PSD and ACVF → new generative model.

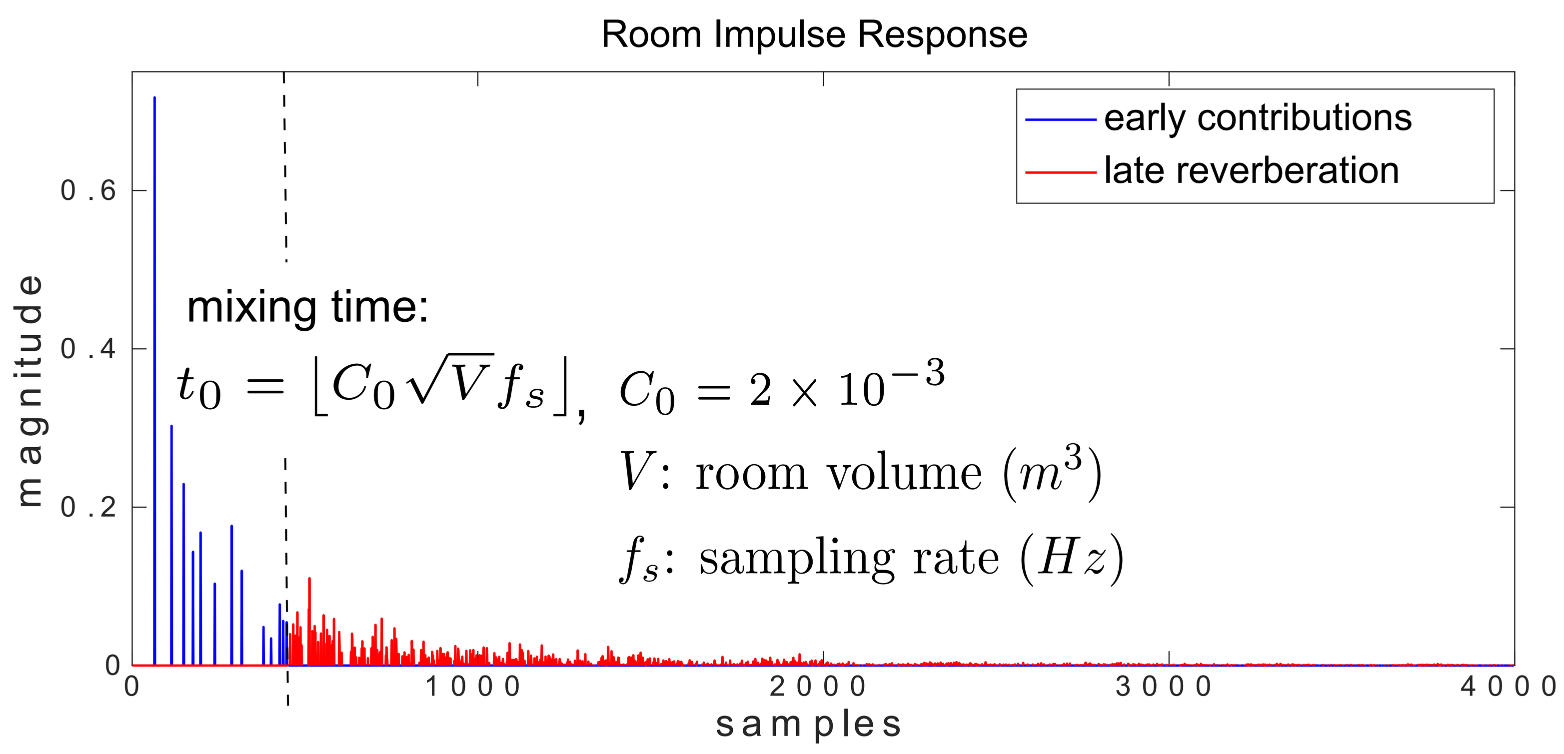
[1] Valid above Schroeder's frequency.

Room Response Decomposition

For $t, k = 0, ..., T-1$:

$$\underbrace{h(t) = h_e(t) + h_l(t)}_{\text{Room impulse response (RIR)}} \quad \underset{\mathcal{F}_T^{-1}}{\overset{\mathcal{F}_T}{\rightleftarrows}} \quad \underbrace{H(k) = H_e(k) + H_l(k)}_{\text{Room frequency response (RFR)}}$$

with $\mathcal{F}_T$ the discrete Fourier transform operator.



Statistical Room Acoustics

▷ Late reverberation power decays exponentially, we thus define the **Power Temporal Profile (PTP)** by:

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

P_0^2: related to the total power of late reverberation; $\tau = \frac{T_{60} f_s}{3 \ln(10)}$ samples; $\mathbb{E}[\cdot]$: expectation = spatial averaging; $\mathbb{1}_{(\cdot)}(t)$: indicator function.

▷ $\{H_l(k)\}_k$ is a T-periodic, centered and proper CGRP with

- ◇ **Autocovariance Function**:

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

- ◇ **Power Spectral Density**:

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

▷ **Variance** $\sigma_{rev}^2 := \gamma(0)$ of $\{H_l(k)\}_k$:

$$\sigma_{rev}^2 \propto (1-\alpha)/(\pi \alpha S).$$

α: average absorption coefficient; S: total wall area (m^2).

Theoretical Statistical Characterization

We can show that the PSD is related to the PTP by:

$$\phi(t) = T \bar{h}_l(T-t). \quad (1)$$

It follows from the Wiener-Khinchin theorem:

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

$$\text{with } P_0^2 = \sigma_{rev}^2 e^{2T/\tau} \frac{1 - e^{2/\tau}}{1 - e^{2(T-t_0+1)/\tau}}. \quad (3)$$

Experimental validation: Monte-Carlo simulation for computing empirical ACVFs from simulated and measured room responses.

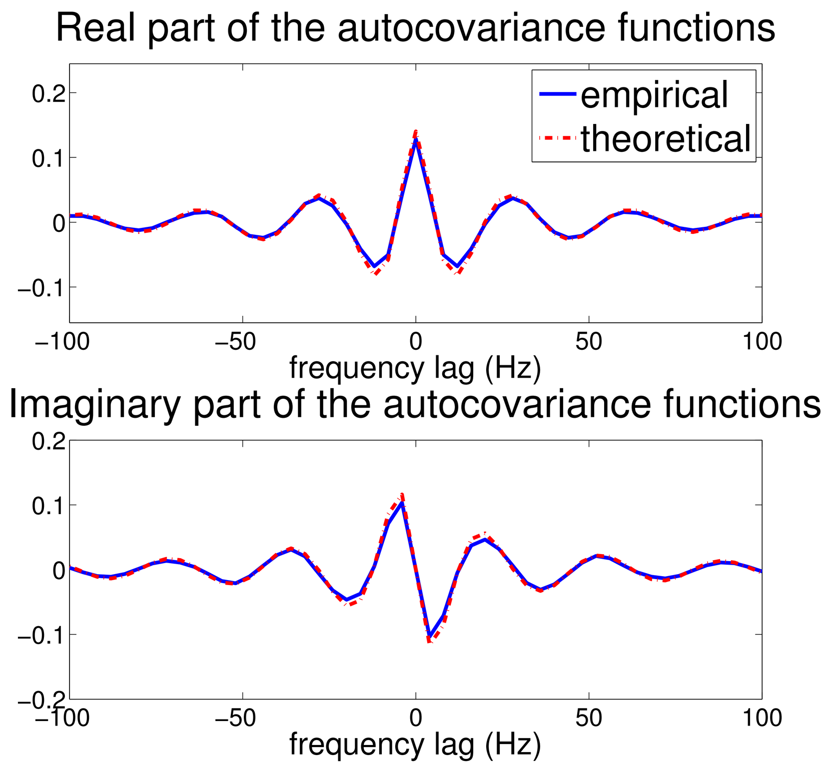


Figure 1: From 196 simulated RIRs, $T_{60} = 0.25$ s, room: $10 \times 6.6 \times 3.3$ m.

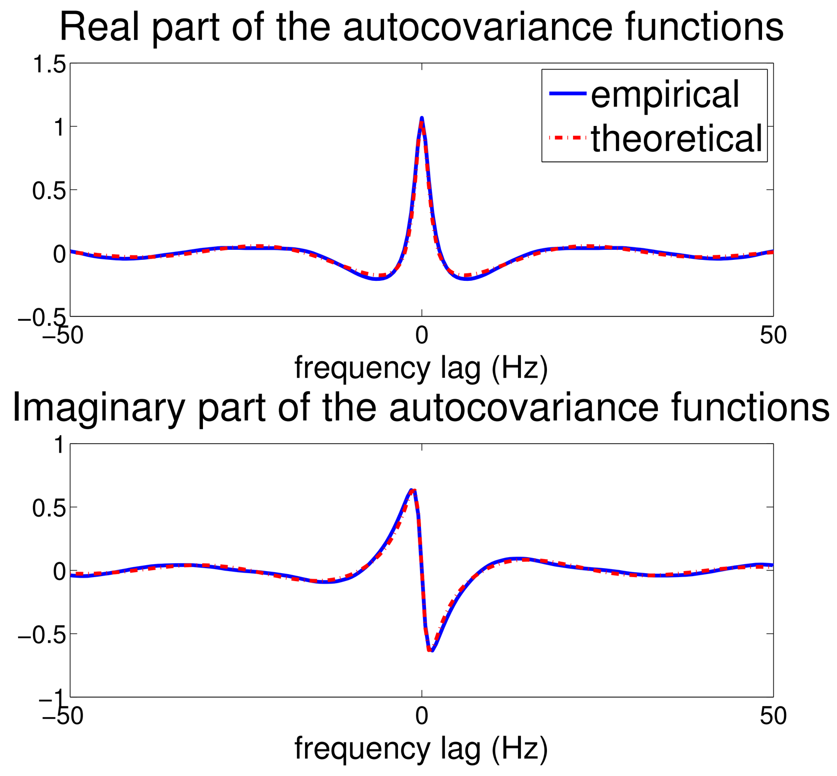


Figure 2: From 130 measured RIRs, $T_{60} = 1.8$ s, room: $7.5 \times 9.9 \times 3.5$ m.

ARMA parametrization

Objective: Describe the statistical properties with a parametric model.

▷ We assume that $\{H_l(k)\}_k$ follows an ARMA(P, Q) model.

▷ The ARMA parameters can be estimated from the theoretical ACVF, without the need of any data.

▷ The ARMA model is generative.

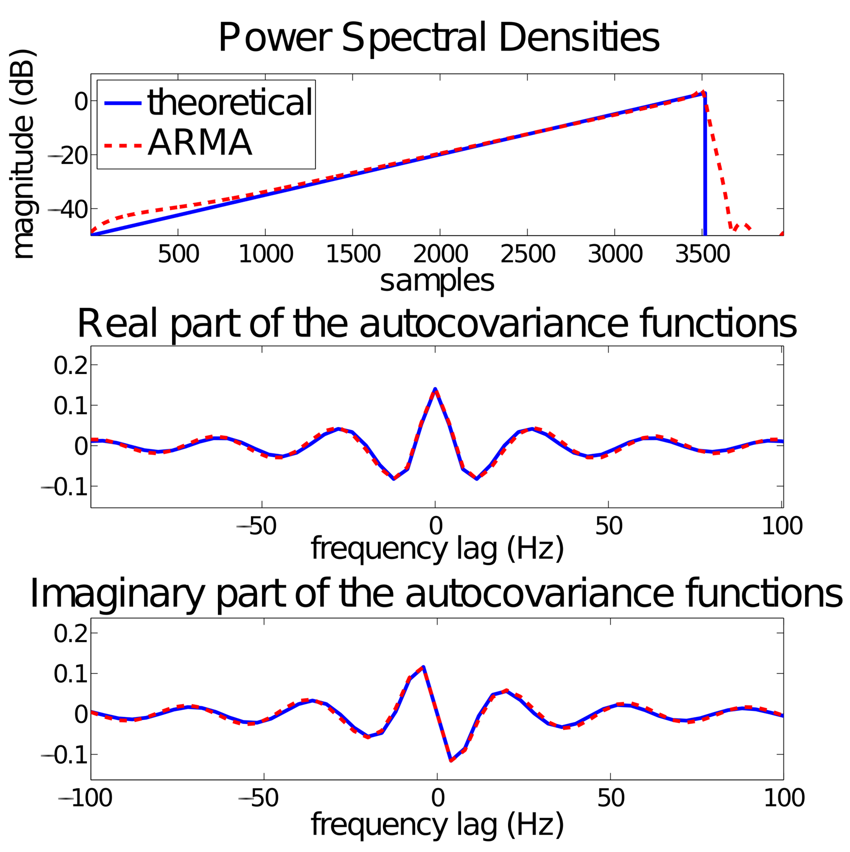


Figure 3: ARMA(7,2) parametrization.

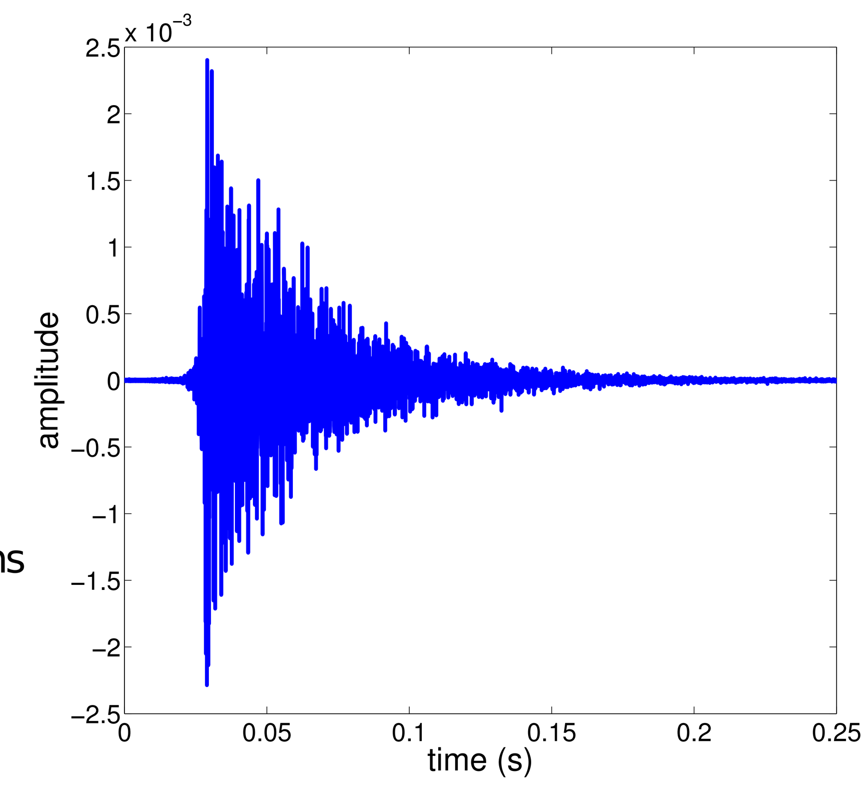


Figure 4: Synthesized late RIR from the ARMA(7,2) model.

Conclusions

▷ Theoretical statistical characterization and generative ARMA model of late reverberation in the frequency domain.

▷ Suitable model for defining a prior over late reverberation in audio source separation and dereverberation.

▷ Future work: Non-stationary ARMA model for frequency dependent reverberation times.

Reference: [1] M. R. Schroeder, "Frequency-correlation functions of frequency responses in rooms," JASA, 1962.