Coherence Based Reflection Coefficient Estimation

Armin Goudarzi[1]
[1] DLR, 37073 Göttingen, Deutschland, Email: armin.goudarzi@dlr.de

Abstract
A new method to determine the acoustic reflection coefficient of surfaces in-situ is presented. It relies on a single noise source and a single microphone. The method is robust to low-end measurement equipment. It is derived from the blind source separation problem where the original but unknown signals are reconstructed from mixed, observable signals. To do so, an analytical solution to the mixing matrix is presented, based on the assumption of a monopole source and a single reflection. The reflection coefficient is derived from the short window coherence of the mixed signals at different times. The method is validated with simulated data.

Introduction
Knowledge of material properties is often necessary to assess acoustic conditions. State-of-the-art methods determine the reflection coefficient of absorber materials by measurements in impedance tubes [1] or reverberation rooms [2]. However, these methods require samples being taken which can not always be realised such as in outdoor environments. A method DIN EN 13472 exists [3], but the measurement setup needs to be calibrated carefully in a free-field room and is not robust to environment changes. A method proposed by Lanoye et al. [4] allows in-situ measurements of the reflection coefficient but relies on sound intensity probes, which are both expensive and sensitive. This paper presents a new in-situ approach based on calculating the coherence of a signal with a reflection time-shifted copy. It relies on a short window coherence Welch estimation and requires no knowledge of the used signal, the microphone's and speaker's frequency response or their calibration. Only the geometric distances between speaker, microphone and material need to be determined, making this method's experimental setup both fast and robust towards inaccuracies and allowing for the use if low-end measurement equipment.

Procedure
The procedure of calculating the reflection coefficient from the short window coherence is presented hereafter.

1. Generate a broadband noise signal (such as white or pink noise).
2. Place a noise source and a microphone above the reverberant surface and determine the distances between sound source, sensor and floor.
3. Measure the noise signal with a sampling frequency that will allow the detection of the reflection time in the autocorrelation later on.
4. Detect the time offset of the reflected sound via a distance-based travel time estimation or autocorrelation.
5. Generate a second signal from the measured signal which is shifted back in time by the time offset.
6. Calculate the coherence of these signals using a window length smaller than the time offset.
7. Calculate the reflection coefficient from the coherence using eq. 13.

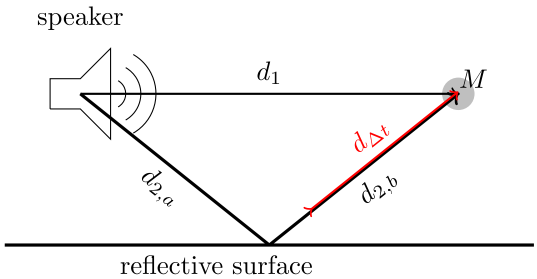

Figure 1: Experimental setup. d_1 is the direct sound travel distance and $d_2 = d_{2,a} + d_{2,b}$ is the total distance for the reflection. $d_{\Delta t} = d_2 - d_1$ is the travel distance difference between reflected and direct sound to the microphone M.

Theory
In figure 1 a schematic graphic of the problem consisting of a single source, sensor and reflection is shown. We can observe the microphone signal, we do not know the speakers original signal and we are interested in the properties of the reflective surface. The reflection coefficient will be derived from the blind source separation problem in the frequency domain. Thus, we focus on time-invariant problems. We denote the Fourier Transformation of the signal s_i with $S_i(f) = A_i(f)Q_i(f)$ with its frequency depended amplitude $A(f)$, its phase $Q(f)$ and its cross-spectral density with $S_{ij}(f) = S_i(f)S_j^*(f)$. The frequency dependency will be neglected in the following for better readability.

The blind source separation problem consists of true, uncorrelated signals S^t that are mixed with a mixing matrix $\boldsymbol{\lambda}$ and results in new, observable signals S.

$$S = \boldsymbol{\lambda} S^t \, . \tag{1}$$

The true signals amplitude A^t, its phase Q^t with $|Q| = 1$ and mixing coefficients $|\boldsymbol{\lambda}| \leq 1$ are unknown. To solve this problem we have to reconstruct the mixing matrix $\boldsymbol{\lambda}$ that includes the reflection coefficient. To do so, the problem is simplified by two assumptions. First, the mixing is symmetric, thus $\boldsymbol{\lambda}_{ij} = \boldsymbol{\lambda}_{ji}$. This assumption is valid in linear acoustics [5, p. 408]. Second, the sound wave can be modelled by a simple geometric monopole source [5, p. 70]. Thus, the mixing matrix can be derived analytically. Accounting for the decay of sound pressure and phase shift due to the propagation distance d we

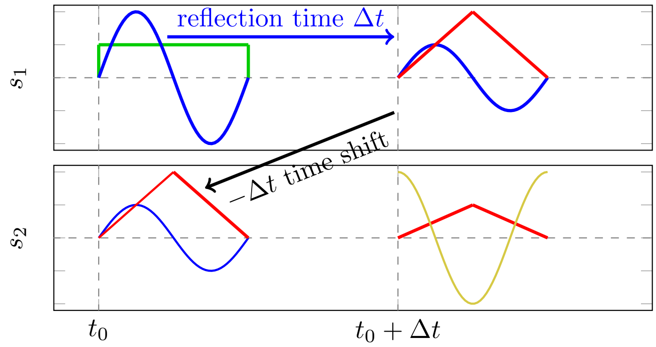

Figure 2: Schematic example of the signal s_1 in time domain which consists of random noise. The unique noise parts are displayed with a unique colour and shape, their reflections have a lower amplitude and appear after Δt. $s_2(t) = s_1(t + \Delta t)$, it is shifted back by the reflection time Δt.

model the monopole sound source with $\omega = 2\pi f$ and $k = \omega/c$

$$p(d,t) = A^t \frac{\exp(-j(kd+\omega t))}{d} = A^t Q^t \frac{Q_d}{d}\,. \quad (2)$$

Note that the sources observable phase consists of the source inherent phase Q^t (dependent on time and frequency) and its phase change due to the time and space distance Q_d (dependent on distance and wavenumber). The reflection of a signal is modelled via the complex reflection coefficient r [5, p. 179]

$$r = r_0 \exp(j\varphi)\,. \quad (3)$$

This neglects atmospheric dampening and non-linear effects. The phase shift due to the reflection will be denoted with Q_r. Given there exists a single source and a single reflective surface with a reflection coefficient r we derive the mixing matrix $\boldsymbol{\lambda}$ using the previous assumptions with

$$\boldsymbol{\lambda} = \begin{bmatrix} \frac{Q_1}{d_1} & \frac{Q_2 Q_r r_0}{d_2} \\ \frac{Q_2 Q_r r_0}{d_2} & \frac{Q_1}{d_1} \end{bmatrix}\,. \quad (4)$$

The observable signals are $S = \boldsymbol{\lambda} S^t$

$$\begin{aligned} S_1 &= \lambda_{11} S_1^t + \lambda_{12} S_2^t \\ S_2 &= \lambda_{21} S_1^t + \lambda_{22} S_2^t\,. \end{aligned} \quad (5)$$

At this point signal one and two are identical which makes sense, as the source signal is the same and the mixing matrix is symmetrical. Since we only have a single source, we rewrite these equations for a reference source time t_0. We relabel $A_i^t = A$ and $Q^t = Q_{t_0}$ at the source reference time t_0. The phase shift due to the different source emission times are denoted with $Q_{t_0 \pm \Delta t}$. The phase shift for the direct travel distance is Q_1, for the reflection we obtain $Q_2 = Q_1 Q_{\Delta t}$

$$\begin{aligned} S_1 &= \frac{A Q_{t_0} Q_1}{d_1} + \frac{r_0 Q_r Q_1 Q_{\Delta t} A Q_{t_0 - \Delta t}}{d_2} \\ S_2 &= \frac{A Q_{t_0} Q_1}{d_1} + \frac{r_0 Q_r Q_1 Q_{\Delta t} A Q_{t_0 - \Delta t}}{d_2}\,. \end{aligned} \quad (6)$$

The equation system is still underdetermined as we do not know r and A. To obtain an additional unique equation we shift S_2 in time so that the travel time difference between the direct sound and the reflected sound due to the propagation distance difference is compensated. Thus, the signals amplitudes and the mixing matrix are still identical but the true signals phase changes. The time shift and the observed signals are shown in figure 2. In the schematic graphic in s_1 at $t = t_0$ the blue sine represents the direct sound while the green square with a lower amplitude represents the reflected sound wave from $t = t_0 - \Delta t$. After the reflection time at $t = t_0 + \Delta t$ we detect the blue sine again with lower amplitude and some new direct sound, displayed as the red triangle. When we shift s_2 back by the reflection time Δt, so that $s_2(t) = s_1(t + \Delta t)$, its reflected part matches the direct part of s_1 while the direct sound part of s_2 is uncorrelated to s_1. Thus, in s_1 and s_2 at t_0 we have in total three independent signal parts, indicated by their unique source time phase. In frequency domain this time shift is expressed by $S_2(t_0) = S_1(t_0 + \Delta t)/Q_{\Delta t}$, thus

$$\begin{aligned} S_1(t_0) = &\frac{A Q_1 Q_{t_0}}{d_1} + \frac{r_0 Q_r Q_1 Q_{\Delta t} A Q_{t_0 - \Delta t}}{d_2} \\ S_2(t_0) = &\frac{A Q_1 Q_{t_0 + \Delta t}}{Q_{\Delta t} d_1} + \frac{r_0 Q_r Q_1 A Q_{t_0}}{d_2}\,. \end{aligned} \quad (7)$$

From the condition of true, uncorrelated signals we deduct that during signal averaging for the Welch estimation $\mathbb{E}[\ldots]$ [6] of the cross power spectral densities that

$$|\mathbb{E}\left[Q_i Q_j^*\right]| = \begin{cases} 0, & \text{for } i \neq j \\ 1, & \text{for } i = j \end{cases}\,. \quad (8)$$

In our case we know that for Welch block size smaller than Δt the signal parts are only correlated when they have the same source time phase, thus

$$|\mathbb{E}\left[Q_{t_0} Q_{t_0 \pm n\Delta t}^*\right]| = \begin{cases} 0, & \text{for } n \in \mathbb{Z} \setminus \{0\} \\ 1, & \text{for } n = 0 \end{cases}\,. \quad (9)$$

Calculating the cross power spectral density we find for $S_{ii} = S_{11} = S_{22}$

$$\begin{aligned} S_{11} = &\mathbb{E}\Bigg[\frac{A^2 |Q_1|^2 |Q_{t_0}|^2}{|Q_{\Delta t}|^2 d_1^2} + \\ &\frac{A^2 r_0^2 |Q_1|^2 |Q_r|^2 |Q_{t_0 - \Delta t}|^2 |Q_{\Delta t}|^2}{d_2^2} + \underbrace{\ldots}_{=0}\Bigg] \\ = &\mathbb{E}\left[A^2 \left(\frac{1}{d_1^2} + \frac{r_0^2}{d_2^2}\right)\right]\,. \end{aligned} \quad (10)$$

For $S_{ij} = S_{ji}^*$ we find

$$\begin{aligned} S_{12} = &\mathbb{E}\left[\frac{A^2 r_0 |Q_1|^2 Q_r^*}{d_1 d_2} + \underbrace{\ldots}_{=0}\right] \\ = &\mathbb{E}\left[\frac{A^2 r_0 Q_r^*}{d_1 d_2}\right]\,. \end{aligned} \quad (11)$$

At this point we already estimated the phase of the reflection coefficient with $\angle r = \angle S_{ji}$ but we still need to find A and r_0. To do so we use these relations and calculate the magnitude squared coherence $\gamma_{ij}^2 = \frac{|S_{ij}|^2}{S_{ii}S_{jj}}$ with $d_0 = d_1/d_2$, $d_0 \geq 0$ and $A \geq 0$

$$\gamma_{12}^2 = \frac{|r_0 Q_r d_0|^2}{(1 + r_0^2 d_0^2)^2} = \frac{|r_0 d_0|^2}{(1 + r_0^2 d_0^2)^2} \,. \tag{12}$$

Since the mixing is $0 \leq r_0 \leq 1$ we can directly calculate the reflection coefficient from the coherence with the quadratic formula

$$r_0 = -\frac{\sqrt{1 - 4\gamma^2} - 1}{2\gamma d_0} \tag{13}$$

in the interval $0 \leq \gamma^2 \leq \frac{1}{4}$. Note, we assumed that the signals are only correlated within the time shift window. This means that for Fourier analysis we can only use block sizes that are smaller than the time shift. However, zero-padding is possible to achieve a reasonable frequency resolution. Eq. 13 is independent of the original sound source amplitude A^t. This means that neither the amplitude of the signal nor the frequency response of the microphone or loudspeaker has to be known. The only requirement is that a sufficient signal is detected with an omnidirectional microphone.

Influence of background noise

Outdoor measurements are often disturbed by background noise $\hat{s}$. Thus, its influence on the proposed method is examined. Assume a noise source in the farfield which is uncorrelated to our primary noise source with the normalised amplitude $\hat{A}/d_1$. The normalisation by the distance of the original noise source makes a comparison easier. We also assume no reflections within the Fourier window. This will result in an underestimation of the coherence since the auto power S_{ii} increases while the cross power spectrum S_{ij} remains the same. Finally, this leads to an underestimation of the reflection coefficient. We can quantify the underestimation using

$$S_{ii} = \mathbb{E}\left[A^2\left(\frac{1}{d_1^2} + \frac{r_0^2}{d_2^2}\right) + \frac{\hat{A}^2}{d_1^2}\right] \tag{14}$$

$$S_{ij} = \mathbb{E}\left[\frac{A^2 r_0 Q_r^*}{d_1 d_2}\right] . \tag{15}$$

This leads to

$$\gamma^2 = \frac{|r_0 d_0|^2}{\left(1 + r_0^2 d_0^2 + \frac{\hat{A}^2}{A^2}\right)^2} \,. \tag{16}$$

Even small ratios of $\hat{A}^2/A^2$ have a big impact on the estimation of r_0. When the background noise source is quasi-stationary we can prevent this underestimation by first measuring the background noise $\hat{s}$ exclusively and subtracting its PSD $\hat{S}_{11}$ from S_{11} before calculating the coherence γ_{mod}^2.

$$\gamma_{\mathrm{mod}}^2 = \frac{|S_{ij}|^2}{(S_{ii} - \hat{S}_{ii})^2} \tag{17}$$

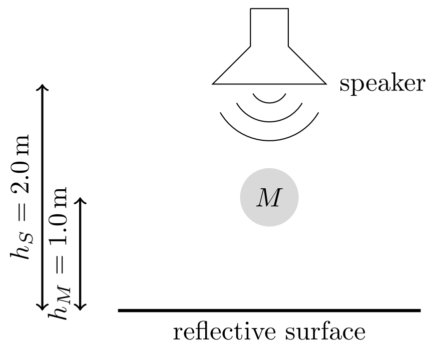

Figure 3: Experimental setup to determine r_0 of a reflective surface with the microphone M placed at $h_M = 1.0$ m and the speaker placed at $h_S = 2.0$ m.

Results

An analytical example is given to illustrate the procedure. A virtual speaker is placed at $h_S = 2$ m, a virtual microphone is placed at $h_M = 1$ m above a reflective surface, see figure 3. We calculate $d_1 = (h_S - h_M) = 1.0$ m and $d_2 = (h_S + h_M) = 3.0$ m, $d_0 = 1/3$. The reflection coefficient r_0 is modelled using a Bessel bandpass filter of order=1, $f_l = 200$ Hz, $f_h = 1500$ Hz. A signal s^t is generated using white noise with fs = 8192 Hz, $t = 100$ s. Then the signal is lowered in amplitude and shifted in time by λ_{11} to generate s_1^t. To generate the reflected signal s_1^r the Bessel bandpass filter is applied and the signal is shifted in time and lowered in amplitude by λ_{12}. The relative time offset between the signals is $\Delta t = (h_S + h_M)/c \approx 0.0059$ s with $c = 340\,\mathrm{m\,s^{-1}}$ and the relative amplitude (excluding the Bessel filter) is $A_1^t/A_1^r = 1/(h_S + h_M) = 1/3$. The reflection time corresponds to approximately 23 samples. Thus, a Welch block size of nperseg = 22 is used. This results in a very low frequency resolution ($\Delta f \approx 372$ Hz) so a zero padding with nfft = 256 is used and results in a frequency resolution of $\Delta f = 64$ Hz. To calculate the Cross Spectral Density, the second signal $s_2(t) = s_1(t + \Delta t)$ is generated.

In figure 4 the results of the method are shown. The real reflection coefficient is the bandpass filter response, shown in black. Since we know S_{11}^t we can calculate the practical optimum for r_0 that includes the Welch estimation error. Using eq. 10 and the Welch estimation $\mathbb{E}[S_{11}^t]$ for the true signal at the microphone position $S_1^t = \lambda_{11} S^t$ with

$$S_{11}^t = \mathbb{E}\left[\frac{A^2}{d_1^2}\right] \tag{18}$$

$$S_{11} = \mathbb{E}\left[\frac{A^2}{d_1^2}\left(1 + r_0^2 d_0^2\right)\right] , \tag{19}$$

we calculate

$$r_0 = \sqrt{\left(\frac{S_{11}}{S_{11}^t} - 1\right) d_0^{-2}} \,. \tag{20}$$

The result is shown in blue. The result of the method proposed in this paper is shown in red. The red dots indi-

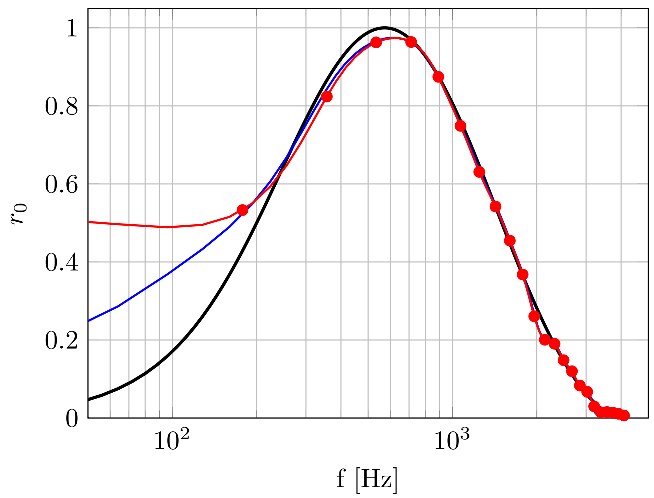

Figure 4: The true reflection coefficient (the bandpass filter response, black) and the practical optimum that could be determined due to the Welch estimation error (blue) using eq. 20. The estimated degree of reflection using eq. 13 is shown in red. The red dots indicate the frequency bins that result from no zero padding.

cate the frequency bins that result from no zero padding.

Discussion

The proposed method to determine the reflection coefficient relies on a single sound source and sensor. The sensor and source must be omnidirectional to ensure that the direct and reflected sound waves are altered by the same microphone transfer function. Since the absolute amplitude cancels out in the equations the sensor and sound source can have any frequency response and allow for low-end measurement equipment. To deploy the proposed method the sensor signal must be copied and shifted by the reflection time. The reflection time can be determined by a geometric calculation or via the autocorrelation. If the autocorrelation method is used notice that by definition we will shift the signal by the reflection time delay plus the phase shift of the absorption material. This can be a problem for thick absorption materials, especially when the phase shift is frequency dependent. Otherwise, the phase of the reflection coefficient will be zero. To obtain the correct phase, the geometric travel time for the reflected sound wave should be used.

When placing the sound source and the sensor we have two make a trade-off. On the one hand, we want to maximize the reflection time, which allows us to use larger block sizes for the FFT. On the other hand, we want to minimize the travel time difference as the intensity of the reflected sound wave decreases with d^2. This results in an extremely low coherence. Due to the non-linear relation in equation 13 between the coherence and the reflection coefficient errors in the estimation will be magnified. This is especially noticeable at low frequencies compared to the block size. In the given example the reflection coefficient was reasonably well estimated above $f \geq 250\,\mathrm{Hz}$. This is between the first two frequency bins that would be achieved without zero-padding for the Welch estimation. We can obtain a higher frequency resolution by using zero padding, but the quality of the results can not be improved due to the Welch estimation error. The estimation in the low-frequency region can not be improved as we would need to increase the delay between the original and reflected signal which on the other hand increases the coherence estimation error due to the decay in amplitude of the reflected signal.

The coherence estimation rapidly drops in the presence of a background noise source which results in an underestimation of the reflection coefficient. For quasi-stationary background noise sources we can eliminate this effect by subtracting the exclusive background noise PSD from S_{11} when calculating the coherence.

Conclusion

This paper proposes a method to determine a material's reflection coefficient in-situ using a single microphone and a sound source. It is straight forward in its application and does not rely on specialist hardware. The frequency response of both devices and the employed broadband signal does not need to be known. Assuming a monopole source and a single reflection the reflection coefficient is derived from the signal's time-shifted auto-coherence with FFT block lengths shorter than the reflection time. This results in a low frequency resolution. Placing the sensor further apart from the material increases the reflection time and frequency resolution but results in a decrease of coherence estimation accuracy. The influence of quasi-stationary background noise can be eliminated but a high signal-to-noise ratio of the set-up is recommended.

References

[1] ISO 10534-2:1998 - Determination of sound absorption coefficient and impedance in impedance tubes - part 2: Transfer-function method, Nov. 1998.

[2] ISO 354:2003 - Measurement of sound absorption in a reverberation room, May 2003.

[3] DIN ISO 13472-1:2004 - Measurement of sound absorption properties of road surfaces in situ - part 1: Extended surface method, Apr. 2004.

[4] R. Lanoye, G. Vermeir, W. Lauriks, R. Kruse, and V. Mellert. Measuring the free field acoustic impedance and absorption coefficient of sound absorbing materials with a combined particle velocity-pressure sensor. *The Journal of the Acoustical Society of America*, 119(5):2826–2831, May 2006.

[5] M. Möser. *Engineering Acoustics*. Springer Berlin Heidelberg, 2009.

[6] P. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, June 1967.