BLIND AMPLITUDE ESTIMATION OF EARLY ROOM REFLECTIONS USING ALTERNATING LEAST SQUARES

Tom Shlomo, Boaz Rafaely

School of Electrical and Computer Engineering
Ben-Gurion University of the Negev
Beer-Sheva 84105, Israel

ABSTRACT

Estimation of properties of early room reflections is an important task in audio signal processing, with applications in beamforming, source separation, room geometry inference, and spatial audio. While there are existing methods to blindly estimate the direction of arrival and of the early reflections, the blind estimation of reflections amplitudes remains an open problem. This work presents a preliminary attempt to blindly estimate reflections amplitudes. An iterative estimator is suggested, based on maximum likelihood and alternating least squares. We discuss some fundamental scaling ambiguities of the problem, and show connections between the proposed method and raking beamformers. A Simulation study demonstrates the effectiveness of the proposed method.

***Index Terms*—** Room reflections, room impulse response, rake beamforming, microphone array

1. INTRODUCTION

Estimation of the properties of early room reflections is useful for many tasks in signal processing, such as speech enhancement and source separation [1, 2], the analysis and synthesis of room acoustics [3], and spatial audio synthesis [4, 5].

Existing methods for the estimation of the parameters of early reflections can be categorized as blind and non-blind. Non-blind methods, operate on room impulse response signals, or alternatively assume the sound source signal is known, while blind methods operate on microphone signals directly, and assume no prior knowledge on the sound source signal. Most commonly, a single reflection is parameterized by its delay, direction of arrival (DOA), and amplitude [6].

Several blind methods for the estimation of the delays and DOAs of the early reflections have been published, mostly using spherical microphone arrays. These include methods that are based on cross correlation analysis and MVDR beamforming [7], and subspace methods such as MUSIC [8] and ESPRIT [9]. A survey comparing several such methods is available at [10]. The work of this paper is motivated by the recent introduction of the PHALCOR algorithm [11, 12], a blind method for the estimation of the delays and DOAs of the early reflections, that can accurately detect a large number of reflections. However, the applications described above, in particular speech enhancement and spatial audio synthesis, require also estimates of reflections amplitudes. While several non-blind methods to estimate reflections amplitudes exist [13, 14], to the best of our knowledge, this work presents the first attempt of blind estimation of reflections amplitudes.

The approach presented in this paper is based on a maximum likelihood (ML) estimator. As it requires solving a non-convex optimization problem, we suggest an iterative alternating maximization procedure to find a local maxima of the likelihood function. We show that each iteration requires solving a small least squares problem, therefore the name - alternating least squares (ALS) [15]. Furthermore, we draw some connections between the suggested method and raking beamformers [1]. A simulation study demonstrates the performance of the proposed method. An implementation of the proposed method, as well as source code to generate all the figures below, is available at `https://github.com/tomshlomo/als-amp-est`.

2. SYSTEM MODEL

This section presents the system model used in the paper. Consider a microphone array with Q microphones arranged in an arbitrary configuration. Consider also a sound field which is comprised of single source in a reverberant room, with a frequency domain signal denoted by $s(f)$. As the sound from the source propagates in the room, it is reflected from the room boundaries. Each reflections may be modeled as a different source, whose signal is a delayed and scaled copy of $s(f)$ [16]. Let $\mathbf{p}(f)$ denote the Q-vector of microphone signals at frequency f. Assuming the source and its reflections are in the far field, $\mathbf{p}(f)$ is described by the following model [17]:

$$\mathbf{p}(f) = \mathbf{V}(f)\mathbf{s}(f) + \mathbf{n}(f), \tag{1}$$

where $\mathbf{s}(f) \triangleq [s_0(f), \ldots, s_K(f)]^{\mathrm{T}}$ contains the frequency domain signals of the direct sound and the first K reflections,

This research was supported by Facebook Reality Labs.

$\mathbf{V}(f)$ is a $Q \times (K+1)$ steering matrix, describing the propagation between each source and the microphones [17], and $\mathbf{n}(f)$ includes noise and late reverberation terms, and is modeled using a circularly symmetric complex normal distribution, with a zero mean and covariance $\sigma^2 \mathbf{I}$.

The signals of the direct sound the first K reflections may be related to the source signal by:

$$s_k(f) = e^{-i2\pi f \tau_k} x_k s(f), \tag{2}$$

where τ_k and x_k are respectively the delay and amplitude of the k'th reflection (for $k \geq 1$) or the direct sound (for $k = 0$).

By combining Eqs. (1) and (2), we obtain the following model for $\mathbf{p}(f)$:

$$\mathbf{p}(f) = \mathbf{V}(f)\mathbf{D}(f)\mathbf{x}s(f) + \mathbf{n}(f), \tag{3}$$

where $\mathbf{D}(f) \triangleq \operatorname{diag}\left(e^{-i2\pi f \tau_0}, \ldots, e^{-i2\pi f \tau_K}\right)$ and $\mathbf{x} \triangleq [x_0, \ldots, x_K]$.

3. PROBLEM STATEMENT

The goal of this paper work is to estimate the vector of reflections amplitudes $\mathbf{x}$ from the observed, noisy microphone signals $\mathcal{P} \triangleq (\mathbf{p}(f_1), \ldots, \mathbf{p}(f_J))$, where $f_1, \ldots, f_J$ are frequencies covering a pre-defined band. Note that the reflections amplitudes are assumed frequency independent within the band. In practice, the amplitudes may vary slowly in frequency. It will therefore be assumed that the bandwidth is small enough such that the amplitudes are approximately constant for the entire frequency band. The processing below may be applied to different frequency bands to obtain amplitudes estimates that are varying in frequency.

As we consider the case of blind estimation, besides the vector of amplitudes $\mathbf{x}$, the source signal $s(f)$ is also assumed unknown, as well as it's delay τ_0. The DOAs of the direct sound and the first K reflections, as well as the delay between the first K reflections and the direct sound, are assumed known. They can be obtained by applying any existing method to estimate them, as detailed in the introduction section. Note that while knowledge of the DOAs fully determines the steering matrix $\mathbf{V}(f)$, knowledge of the delays between the direct sound and the reflections determines $\mathbf{D}(f)$ only up to a phase, which depends on the unknown delay τ_0:

$$\mathbf{D}(f) = e^{-i2\pi f \tau_0} \tilde{\mathbf{D}}(f), \tag{4}$$

where $\tilde{\mathbf{D}}(f)$ is the known matrix:

$$\tilde{\mathbf{D}}(f) \triangleq \operatorname{diag}\left(1, e^{-i2\pi f \tilde{\tau}_1}, \ldots, e^{-i2\pi f \tilde{\tau}_K}\right), \tag{5}$$

and $\tilde{\tau}_k \triangleq \tau_k - \tau_0$.

An important limitation of blind amplitude estimation is that it is fundamentally *scale ambiguous*. Let c be an arbitrary non-zero scalar. We can always re-write the model (3), as follows:

$$\mathbf{p}(f) = \mathbf{V}(f)\mathbf{D}(f)\left(\mathbf{x}c\right)\left(s(f)/c\right) + \mathbf{n}(f). \tag{6}$$

As evident, the same observations $\mathbf{p}(f)$ that are obtained by a given $\mathbf{x}$ and $s(f)$, could be just as well explained by $\mathbf{x}c$ and $s(f)/c$. Note that most applications are insensitive to such ambiguities, and often only the relative amplitudes of the reflections are important.

4. ALGORITHM DESCRIPTION

In this section, we present our algorithm for estimating the reflections amplitudes. We begin be deriving the ML estimator of $\mathbf{x}$ from $\mathcal{P}$, by modeling $\mathbf{x}$, $\{s(f_j)\}_{j=1}^{J}$ and τ_0 as deterministic unknowns:

$$\hat{\mathbf{x}}_{\text{ML}} \triangleq \underset{\mathbf{x} \in \mathbb{R}^{K+1}}{\arg\min} \left\{ \min_{\substack{s(f_1), \ldots, s(f_J) \in \mathbb{C}, \\ \tau_0 \geq 0}} -\mathcal{L}(\boldsymbol{\theta} | \mathcal{P}) \right\}, \tag{7}$$

where $\boldsymbol{\theta} \triangleq (\mathbf{x}, s(f_1), \ldots, s(f_J), \tau_0)$ is the vector of unknown parameters, and $\mathcal{L}$ is the log-likelihood function of $\boldsymbol{\theta}$ given $\mathcal{P}$. By assuming $\mathbf{n}(f_j)$ is uncorrelated with $\mathbf{n}(f_i)$ for $i \neq j$, and ignoring constant terms, we can simplify the inner maximization in Eq. (7) as follows:

$$\begin{aligned}
&\min_{\substack{s(f_1), \ldots, s(f_J) \in \mathbb{C}, \\ \tau_0 \geq 0}} -\mathcal{L}(\boldsymbol{\theta} | \mathcal{P}) \\
&= \min_{\substack{s(f_1), \ldots, s(f_J) \in \mathbb{C}, \\ \tau_0 \geq 0}} \sum_{j=1}^{J} \left\| \mathbf{A}(f_j)\mathbf{x}e^{-i2\pi f_j \tau_0} s(f_j) - \mathbf{p}(f_j) \right\|^2 \\
&= \min_{\tilde{s}(f_1), \ldots, \tilde{s}(f_J) \in \mathbb{C}} \sum_{j=1}^{J} \left\| \mathbf{A}(f_j)\mathbf{x}\tilde{s}(f_j) - \mathbf{p}(f_j) \right\|^2
\end{aligned} \tag{8}$$

where $\mathbf{A}(f) \triangleq \mathbf{V}(f)\tilde{\mathbf{D}}(f)$, and $\|\cdot\|$ is the Euclidean norm. The last equality in (8) was made by introducing the variables $\tilde{s}(f) = e^{-i2\pi f \tau_0} s(f)$. By combining Eqs. (7) and (8), we obtain the following expression for the ML estimator:

$$\begin{aligned}
\hat{\mathbf{x}}_{\text{ML}} = \underset{\mathbf{x} \in \mathbb{R}^{K+1}}{\arg\min} \Bigg\{ \\
\min_{\tilde{s}(f_1), \ldots, \tilde{s}(f_J) \in \mathbb{C}} \sum_{j=1}^{J} \left\| \mathbf{A}(f_j)\mathbf{x}\tilde{s}(f_j) - \mathbf{p}(f_j) \right\|^2 \Bigg\}
\end{aligned} \tag{9}$$

As the ML estimator for $\mathbf{x}$ involves solving a bi-linear least squares problem, which is a non convex optimization problem, we resort to a sub-optimal solution that finds only a local maximum of the likelihood function. We show next that when fixing $\mathbf{x}$, the inner optimization in Eq. (9) is solved analytically. Similarly, when fixing $\tilde{s}(f_1), \ldots, \tilde{s}(f_J)$, the outer optimization (over $\mathbf{x}$) is also solved analytically.

4.1. Solving for $\tilde{s}(f)$ given $\hat{\mathbf{x}}$

When $\hat{\mathbf{x}}$ is fixed, the inner optimization problem in Eq. (9) becomes separable, and can be solved for each frequency independently using scalar least squares:

$$\tilde{s}(f) \triangleq \underset{\tilde{s}\in\mathbb{C}}{\arg\min} \left\|\mathbf{A}(f)\hat{\mathbf{x}}\tilde{s} - \mathbf{p}(f)\right\|^2 = \frac{1}{\left\|\mathbf{A}(f)\hat{\mathbf{x}}\right\|^2} \left(\mathbf{A}(f)\hat{\mathbf{x}}\right)^{\mathrm{H}} \mathbf{p}(f) \tag{10}$$

A more intuitive interpretation for Eq. (10) can be obtained by rewriting it as:

$$\tilde{s}(f) = \frac{1}{\left\|\mathbf{A}(f)\hat{\mathbf{x}}\right\|^2} \sum_{k=0}^{K} x_k^* e^{i2\pi f\tilde{\tau}_k} \mathbf{v}_k(f)^{\mathrm{H}}\mathbf{p}(f), \tag{11}$$

where $\mathbf{v}_k$ is the steering vector of the k'th reflection (for $k \geq 1$) or the direct sound (for $k = 0$). The term $\mathbf{v}_k(f)^{\mathrm{H}}\mathbf{p}(f)$ is the output of a delay-and-sum beamformer directed towards to k'th source. The term $e^{i2\pi f\tilde{\tau}_k}$ represents a delay of the beamfomer output according to the delay of the reflection, which coherently aligns the summed signals (this processing is known as raking [1]). In this interpretation, the amplitudes $\hat{\mathbf{x}}$ determine the weight of each reflection in the summation, giving more weight to strong reflections. Finally, the term $1/\left\|\mathbf{A}(f)\hat{\mathbf{x}}\right\|^2$ can be considered as an equalization filter, ensuring that the overall beamformer is distortionless.

4.2. Solving for $\mathbf{x}$ given $\tilde{s}(f)$

When $\tilde{s}(f_1), \ldots, \tilde{s}(f_J)$ are fixed, the outer optimization in Eq. (9) can be rewritten as:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathbb{R}^{K+1}}{\arg\min} \left\|\tilde{\mathbf{A}}\mathbf{x} - \tilde{\mathbf{p}}\right\|, \tag{12}$$

where $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{p}}$ are defined block-wise as follows:

$$\tilde{\mathbf{A}} \triangleq \left[\tilde{s}(f_1)\mathbf{A}(f_1)^{\mathrm{T}} \quad \cdots \quad \tilde{s}(f_J)\mathbf{A}(f_J)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{13}$$

$$\tilde{\mathbf{p}} \triangleq \left[\mathbf{p}(f_1)^{\mathrm{T}} \quad \cdots \quad \mathbf{p}(f_J)^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{14}$$

Finally, the solution to Eq. (12) is given by:

$$\hat{\mathbf{x}} = \begin{bmatrix}\Re(\tilde{\mathbf{A}})\\ \Im(\tilde{\mathbf{A}})\end{bmatrix}^{\dagger} \begin{bmatrix}\Re(\tilde{\mathbf{p}})\\ \Im(\tilde{\mathbf{p}})\end{bmatrix}, \tag{15}$$

where $\Re$ and $\Im$ return the real and imaginary part respectively, and operate entry-wise, and $(\cdot)^{\dagger}$ is the Moore-Penrose pseudo-inverse. Note that $\mathbf{x}$ is assumed real [6], which can be justified for rooms with hard walls in which the attenuation of reflections relative to the direct sound is dominated by the radial spreading of energy, rather than wall absorption. Nevertheless, $\mathbf{x}$ can be assumed complex instead, in which case the solution can be obtained using standard least squares, i.e $\hat{\mathbf{x}} = \tilde{\mathbf{A}}^{\dagger}\tilde{\mathbf{p}}$.

Algorithm 1: The ALS algorithm for blind amplitude estimation of early room reflections.

Input: Microphones signals $\mathbf{p}(f_1), \ldots, \mathbf{p}(f_J)$, Steering matrices $\mathbf{V}(f_1), \ldots, \mathbf{V}(f_J)$, delays of the first K reflections $\tilde{\tau}_1, \ldots, \tilde{\tau}_K$.
Output: $\hat{\mathbf{x}}$, the amplitude of the direct sound and the first K reflections.

```
x̂ ← [1, 0, ..., 0]^T
repeat
    for j = 1, ..., J do
        Update s̃(f_j) according to Eq. (10)
    Update x̂ according to Eq. (15)
until Convergence
```

Eq. (15) may be well-posed even if $K + 1 > Q$, that is, even if the number of reflections is larger than the number of microphones. This somewhat surprising result, is attributed to the fact that Eq. (12) considers multiple frequencies simultaneously. Nevertheless, note that the computational complexity of Eq. (15) is $\mathcal{O}(K^2J)$, that is, quadratic in the (typically small) number of reflections, and only linear in the number of frequencies.

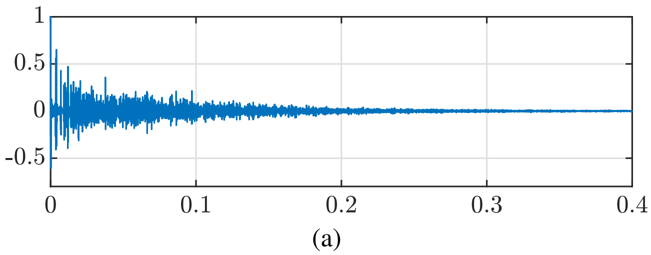
4.3. Alternating Least Squares

Given the analsys above, we propose to solve Eq. (9) by alternating between Eqs. (10) and (12) until convergence. Since each step requires solving a least squares problem, we refer to this method as ALS. Note that ALS is used in other domains where optimization problems involving bilinear functions arise, e.g collaborative filtering with low rank models [15].

The convergence of the algorithm is guaranteed, since at each step the cost function is decreased. To initialize the algorithm, we suggest using simply $\mathbf{x} = [1, \hat{0}, \ldots, 0]^{\mathrm{T}}$, which we empirically found to give good performance. Furthermore, this initialization can be theoretically justified in two ways. First, it is often the case that the amplitude of the direct sound is significantly larger than the reflections. Second, as can be seen from Eq. (11), the first estimate of $\tilde{s}$ that results from this initialization, is equivalent to the one obtained by steering a delay and sum beamformer towards the DOA of the direct sound. Algorithm 1 presents a pseudo-code for the proposed method.

5. SIMULATION STUDY

In this section, we demonstrate the performance of the proposed algorithm with simulated data. An acoustic scene, that consists of a speaker and a rigid spherical microphone array in a shoe box room, was simulated using the image method [16]. The speech signal is a one second long female speech sam-



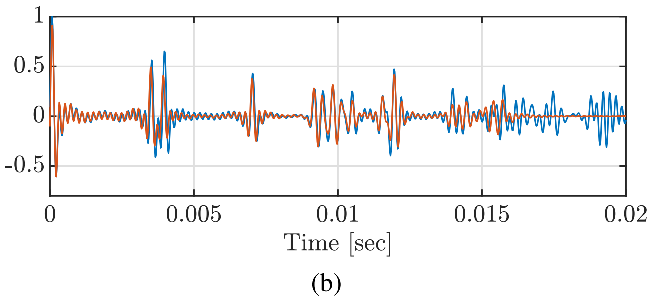


Fig. 1. (a) RIR, band pass filtered between 0.5 and 5 kHz. (b) Zoomed view to the early part of the RIR (blue), and the RIR reconstructed using the estimated amplitudes (red).

ple from the TSP Speech Database [18]. The array has 32 microphones, and a radius of 4.2 cm (similar to the Eigenmike [19]). The microphone signals were sampled at 48 kHz. Sensor noise was added, such that the direct sound to noise ratio is 30dB. The room dimensions are $7 \times 5 \times 3$ m, and its reverberation time is $T_{60} = 0.8$ sec. The center of the array and the speaker are positioned at $(1.7, 2.2, 1.8)$ m and $(4.8, 2.5, 1.1)$ m, respectively, such that the direct to reverberation ratio is -12.8dB. The room impulse response (RIR) between the speaker and a single microphone is shown in Fig. 1(a).

We apply our method on various frequency bands of width 1kHz, and center frequency varies between 1kHz and 10kHz, denoted by f_c. To test the sensitivity of the proposed method, the DOA and delay of the early reflections are perturbed with an additive zero mean Gaussian noise, with standard deviations denoted by σ_Ω and σ_τ, respectively. The number of early reflections K is set to 20, which correspond to approximately the first 15ms of the RIR.

Given the scaling ambiguity discussed in section 3, we use the following scale invariant error measure, which we refer to as scale invariant mean squared error (SIMSE):

$$\text{SIMSE}\left(\hat{\mathbf{x}}, \mathbf{x}\right) \triangleq \frac{\min_{c \in \mathbb{R}} \left\|\hat{\mathbf{x}}c - \mathbf{x}\right\|^2}{\left\|\mathbf{x}\right\|^2} = 1 - \left|\frac{\mathbf{x}^{\mathrm{T}}\hat{\mathbf{x}}}{\left\|\mathbf{x}\right\| \left\|\hat{\mathbf{x}}\right\|}\right|^2$$

Fig. 1(b) shows the early part of the RIR, as well as its reconstruction using the estimated amplitudes and the noisy DOAs and delays. In this example, σ_Ω and σ_τ were set to 10° and 10µs respectively. The figure shows that the estimated RIR follows the true response fairly accurately.

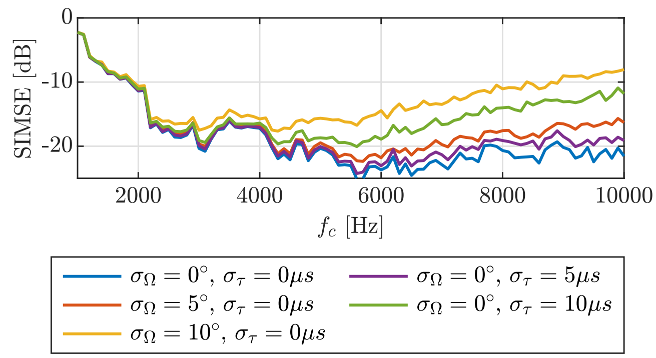


Fig. 2. SIMSE vs. frequency, for $K = 20$ reflections, and different levels of DOAs and delay perturbations σ_Ω, σ_τ.

Fig. 2 shows the estimation performance, as a function of the band central frequency f_c, and consider several values of σ_Ω and σ_τ. For each value of f_c, σ_Ω and σ_τ, we repeat the simulation 50 times, and calculate the average SIMSE. As evident, when $\sigma_\Omega = \sigma_\tau = 0$, the performance degrades as frequency decreases. This is attributed to the poor spatial resolution in the low frequencies: as the frequency decreases, the wavelength compared to the array size increases, and therefore the spatial response of the different reflections become too similar to separate. When σ_Ω increases, from about 4kHz, the performance starts to degrade. This is since when the wavelength of the array is small compared to the array size, small perturbations in the DOAs can significantly change the steering matrix $\mathbf{V}(f)$. The effect of delay errors also increases with frequency. An explanation can be given by considering the raking beamformer interpretation of the $\tilde{s}$ update, shown in Eq. (11). When the error in $\tilde{\tau}_k$ is significant compared to $1/f$ (the wave period), the aligned beamformers outputs are not coherently summed, and performance is therefore degraded.

6. CONCLUSIONS

In this work, a novel method to blindly estimate the amplitudes of early reflections is proposed. The method is based on maximum likelihood estimation and ALS, and is related to rake beamforming. A simulation study showed that the proposed method is able to accurately estimate the amplitudes of a large number of reflections, under reverberant and noisy conditions. The results show that while the performance in the low and middle frequency bands is insensitive to DOA and delay errors, at the high frequencies their effect is more significant. While existing methods for blind DOA estimation may be accurate enough, achieving the required delay accuracy is more of a challenge, and requires further research. Future work will apply the proposed method on experimental data, and study its possible use for spatial audio synthesis applications.

7. REFERENCES

[1] Ivan Dokmanić, Robin Scheibler, and Martin Vetterli, "Raking the cocktail party," *IEEE journal of selected topics in signal processing*, vol. 9, no. 5, pp. 825–836, 2015.

[2] Emmanuel Vincent, Nancy Bertin, Rémi Gribonval, and Frédéric Bimbot, "From blind to guided audio source separation: How models and side information can improve the separation of sound," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 107–115, 2014.

[3] Juha Merimaa and Ville Pulkki, "Spatial impulse response rendering i: Analysis and synthesis," *Journal of the Audio Engineering Society*, vol. 53, no. 12, pp. 1115–1127, 2005.

[4] Ville Pulkki, Symeon Delikaris-Manias, and Archontis Politis, *Parametric time-frequency domain spatial audio*, Wiley Online Library, 2018.

[5] Luca Remaggi, Philip Jackson, and Philip Coleman, "Estimation of room reflection parameters for a reverberant spatial audio object," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[6] Heinrich Kuttruff, *Room acoustics*, CRC Press, 2016.

[7] Edwin Mabande, Konrad Kowalczyk, Haohai Sun, and Walter Kellermann, "Room geometry inference based on spherical microphone array eigenbeam processing," *The Journal of the Acoustical Society of America*, vol. 134, no. 4, pp. 2773–2789, 2013.

[8] Dima Khaykin and Boaz Rafaely, "Coherent signals direction-of-arrival estimation using a spherical microphone array: Frequency smoothing approach," in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 2009, pp. 221–224.

[9] Byeongho Jo and Jung-Woo Choi, "Robust localization of early reflections in a room using semi real-valued eb-esprit with three recurrence relations and laplacian constraint," in *International Commission for Acoustics (ICA)*. International Commission for Acoustics (ICA), 2019.

[10] Haohai Sun, Edwin Mabande, Konrad Kowalczyk, and Walter Kellermann, "Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing," *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2828–2840, 2012.

[11] Tom Shlomo, "Blind localization of early room reflections using phase aligned spatial correlation," M.S. thesis, School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel, October 2020.

[12] Tom Shlomo and Boaz Rafaely, "Blind localization of early room reflections from reverberant speech using phase aligned spatial correlation," *The 179th Meeting of the Acoustical Society of America*, (a) (accepted for publication).

[13] Sakari Tervo, "Single snapshot detection and estimation of reflections from room impulse responses in the spherical harmonic domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 12, pp. 2466–2480, 2016.

[14] Dejan Markovic, Christian Hofmann, Fabio Antonacci, Konrad Kowalczyk, Augusto Sarti, and Walter Kellermann, "Reflection coefficient estimation by pseudospectrum matching," in *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*. VDE, 2012, pp. 1–4.

[15] Moritz Hardt, "Understanding alternating minimization for matrix completion," in *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*. IEEE, 2014, pp. 651–660.

[16] Jont B Allen and David A Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[17] Harry L Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*, John Wiley & Sons, 2004.

[18] Peter Kabal, "TSP speech database," *McGill University, Database Version*, vol. 1, no. 0, pp. 09–02, 2002.

[19] MH Acoustics, "Em32 eigenmike microphone array release notes (v17. 0)," *25 Summit Ave, Summit, NJ 07901, USA*, 2013.