

# TDOA Estimation by Mapped Steered Response Power Analysis Utilizing Higher-Order Moments

Xiao-Dong Zhai, Yuya Sugimoto, Shigeki Miyabe, Takeshi Yamada, and Shoji Makino

Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan.

E-mail: {zhaixiaodong, sugimoto}@mmlab.cs.tsukuba.ac.jp, {miyabe, maki}@ara.tsukuba.ac.jp, takeshi@cs.tsukuba.ac.jp

**Abstract**—In this paper, we propose a new estimation method for the time difference of arrival (TDOA) between two microphones with improved accuracy by exploiting higher-order moments. In the proposed method analyzes the steered response power (SRP) of the observed signals after nonlinearly mapped onto a higher-dimensional space. Since the mapping operation enhances the linear independence between different vectors by increasing the dimensionality of the observed signals, the TDOA analysis achieves higher resolution. The results of an experiment comparing the TDOA estimation performance of the proposed method with that of the conventional methods reveal the robustness of the proposed method against noise and reverberation.

## I. INTRODUCTION

In array signal processing, the time delay of arrival (TDOA) is used in various applications as prior information [1]. Also, TDOA estimation is used in direction of arrival (DOA) estimation, which is an important issue in audio scene analysis. In Particular, in the simultaneous localization problem of sound sources and microphones [2], accuracy of TDOA estimation is critical.

The general approach to estimating the TDOA is to analyze the peak of the interchannel correlation. The most representative method is the generalized cross-correlation (GCC) method [3], which sharpens the peak of the correlation function using a whitening filter. While many methods have been proposed for DOA estimation for a microphone array with three or more elements [4][5][6][7][8], the conventional GCC method is still the most reliable method for TDOA analysis for two microphones. Although the DOA estimation problem with two microphones is equivalent to TDOA estimation when the velocity and the distance between the microphones are given, high-resolution DOA estimation methods such as multiple signal classification (MUSIC) [5] and its alternatives [7][8] cannot achieve high accuracy because they assume that the number of microphones is much larger than the number of sources.

In this paper, we propose a new TDOA estimation method, referred to as *mapped steered response power (SRP)*, which improves the estimation accuracy. First, we map the observed signal onto a higher dimensional space using the nonlinear function we used in our previous work for the extension of MUSIC based on high-order moment analysis [8]. Next, in the mapped higher-dimensional space, we conduct SRP analysis [6], which is a multichannel extension of the GCC method. Owing to the improved linear independence resulting from the increase in the dimensionality, the SRP analysis after the nonlinear mapping achieves higher-resolution TDOA estimation. Through a simulation experiment to verify the effectiveness of our proposed method, we show that mapped SRP achieves higher TDOA estimation accuracy than the

conventional methods of GCC, MUSIC, and  $2q$ -MUSIC [7] in noisy and reverberant environments.

## II. STATEMENT OF PROBLEM

Here we formulate the signal model and the TDOA estimation problem. Suppose we have a target source whose amplitude is given by  $s(t)$ , where  $t$  denotes the continuous time. Our goal is to estimate the TDOA  $\Delta\tau$  of the target source  $s(t)$  between two microphones in a room where reverberation and noise exist. The observed signals are formulated as

$$x_1(t) = \{\alpha_1 \text{sinc}(t - \tau_1) + h_1(t)\} \otimes s(t) + n_1(t), \quad (1)$$

$$x_2(t) = \{\alpha_2 \text{sinc}(t - \tau_1 - \Delta\tau) + h_2(t)\} \otimes s(t) + n_2(t), \quad (2)$$

$$\Delta\tau \triangleq \tau_2 - \tau_1, \quad (3)$$

$$\text{sinc}(t) = \frac{\sin \pi t}{\pi t}, \quad (4)$$

where  $i$  denotes the index of the channel,  $x_i(t)$  denotes the observed signal at the  $i$ th channel,  $\alpha_i$  and  $\tau_i$  are the attenuation coefficient and the time of arrival of the direct wave propagating from the target source to the  $i$ th channel, respectively,  $h_i(t)$  is the impulse response of the indirect propagation path from the target source to the  $i$ th microphone,  $n_i(t)$  is the observation noise at the  $i$ th microphone, and  $\otimes$  is the convolution operator between two functions  $f(t)$  and  $g(t)$  given by

$$f(t) \otimes g(t) = \int_{-\infty}^{\infty} f(t') g(t - t') dt'. \quad (5)$$

We assume that the impulse responses are time-invariant and that the sound source  $s(t)$  the observation noise  $n_i(t)$  for  $i = 1, 2$  are stationary with zero mean. Also, we assume that  $s(t)$  and  $n_i(t)$  for  $i = 1, 2$  are statistically independent.

## III. CONVENTIONAL GCC-BASED TDOA ESTIMATION

### A. TDOA Estimation Model in Time Domain

In this section, we describe the principle of TDOA estimation based on GCC analysis. For the signal model in Sect. II, the cross-correlation function  $C_{x_1 x_2}(\tau)$  between the channels can be formulated as follows:

$$C_{x_1 x_2}(\tau) = E_t [x_1(t) x_2(t + \tau)], \quad (6)$$

$$E_t [x(t)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt, \quad (7)$$

where  $E_t[\cdot]$  denotes the expectation of the argument in the time domain and  $T$  represents the length of the observation

interval  $0 \leq t < T$ . Furthermore, we assume that the source signal  $s(t)$  and noise  $n_i(t)$  are uncorrelated and that

$$C_{sn_i}(\tau) \approx 0. \quad (8)$$

According to Eqs. (7) and (8), the cross-correlation function  $C_{x_1x_2}(\tau)$  can be expanded as

$$C_{x_1x_2}(\tau) \approx \alpha_1\alpha_2 C_{ss}(\tau - \Delta\tau) + \{\alpha_1 h_2(-\tau) + \alpha_2 h_1(\tau) + h_1(\tau) \otimes h_2(-\tau)\} \otimes C_{ss}(\tau). \quad (9)$$

Assuming that the direct waves have sufficiently larger gain than the interference waves, we introduce the following relationship among  $\alpha_i$  and  $h_i$ :

$$\alpha_1^2 \approx \alpha_2^2 \gg \int_{-\infty}^{\infty} h_1^2(t) dt \approx \int_{-\infty}^{\infty} h_2^2(t) dt. \quad (10)$$

According to the assumptions of Eq. (10), the whole cross-correlation function  $C_{x_1x_2}(\tau)$  can be approximated as

$$C_{x_1x_2}(\tau) \approx \alpha_1\alpha_2 C_{ss}(\tau - \Delta\tau). \quad (11)$$

Moreover, if the source signal  $s(t)$  is white, the cross-correlation function  $C_{x_1x_2}(\tau)$  can be rewritten as:

$$C_{x_1x_2}(\tau) \approx \alpha_1\alpha_2 E[s^2(t)] \text{sinc}(\tau - \Delta\tau). \quad (12)$$

Under the assumptions given by Eqs. (10) and (12),  $\Delta\tau$  can be estimated by finding the maximum of  $C_{x_1x_2}(\tau)$ . However, when the above assumptions have an error owing to a reverberant and noisy recording environment and a colored signal, the peak of  $C_{x_1x_2}(\omega)$  becomes unclear. To ensure robustness against noise and reflected waves, the cross-correlation function  $C_{x_1x_2}$  is substituted for its whitened version  $\psi_{12}(\tau) \otimes C_{x_1x_2}(\tau)$  in the GCC method by applying the whitening filter  $\psi_{12}(\tau)$  to clarify the peak:

$$\Delta\tau \leftarrow \arg \max_{\tau} \psi_{12}(\tau) \otimes C_{x_1x_2}(\tau). \quad (13)$$

By designing an appropriate filter for the conditions, the TDOA estimation accuracy is greatly improved even in a noisy or reverberant environment.

### B. GCC Algorithm in Frequency Domain

Although we described the model of GCC-based TDOA estimation in the continuous time domain, the practical processing is usually conducted in the discrete time-frequency domain. First, we obtain the short-time Fourier transform (STFT)  $x(\omega, n)$  of the time domain observation  $x(t)$ , where  $-\pi \leq \omega < \pi$  denotes the angular frequency and  $n = 1, \dots, N$  is the time index of the  $N$  STFT frames. Since a delay of the length  $\tau$  is expressed by the phase rotation  $\exp(-j\omega\tau)$  in the time-frequency domain, the estimation of the TDOA  $\Delta\tau$  is equivalent to finding the filter  $\exp(-j\omega\Delta\tau)$  that maximizes the cross-spectrum  $G_{x_1x_2}(\omega)$ , which is the discrete Fourier

transform (DFT) of the cross-correlation function  $C_{x_1x_2}(t)$ :

$$\Delta\tau \leftarrow \arg \max_{\tau} J_{\text{GCC}}(\tau), \quad (14)$$

$$J_{\text{GCC}}(\tau) = \int_{-\pi}^{\pi} \Psi_{12}(\omega) G_{x_1x_2}(\omega) \exp(-j\omega\tau) d\omega, \quad (15)$$

$$G_{x_1x_2}(\omega) = E_n[x_1^*(\omega, n) x_2(\omega, n)], \quad (16)$$

$$E_n[x(\omega, n)] = \frac{1}{N} \sum_{n=1}^N x(\omega, n), \quad (17)$$

where  $[\cdot]^*$  denotes the complex conjugate,  $E_n[\cdot]$  denotes the expectation of the argument in frequency domain, and  $\Psi_{12}(\omega)$  is the frequency-domain whitening filter corresponding to the DFT of  $\psi_{12}(t)$ . The design of  $\Psi_{12}(\omega)$  will be described in Sect. III-D.

### C. Multichannel extension by SRP

SRP analysis is an extension of the GCC method for estimating multiple TDOAs of all the pairs of three channels or more simultaneously. Although this paper focuses on the TDOA estimation of a two-channel pair, we review the SRP process because the methodology is utilized in the formulation of the proposed method.

Without loss of generality, the estimation of the TDOAs of  $M$  channels can be formulated as the estimation of the TDOAs  $\Delta\tau_{1j}$  ( $j = 1, \dots, M$ ) between the  $j$ th channel and the first channel regarded as the reference channel as follows:

$$\Delta\tau_{1j} = \tau_j - \tau_1. \quad (18)$$

Note that the TDOA  $\Delta\tau_{11}$  is obviously zero. We define the following vector  $\boldsymbol{\tau}$  consisting of all TDOAs  $\Delta\tau_{11}, \dots, \Delta\tau_{1M}$ :

$$\boldsymbol{\tau} = [\Delta\tau_{11}, \dots, \Delta\tau_{1M}]^T, \quad (19)$$

where  $[\cdot]^T$  denotes the transpose. Then, the correct estimation of the vector  $\boldsymbol{\tau}$  ought to maximize the cross-correlation functions of all channel pairs. Hence, the SRP method estimates the vector  $\boldsymbol{\tau}$  by maximizing the summation of the GCC criteria of all channel pairs.

$$\boldsymbol{\tau} \leftarrow \arg \max_{\boldsymbol{\tau}'} J_{\text{SRP}}(\boldsymbol{\tau}') \quad \text{subject to } \tau'_1 = 0, \quad (20)$$

$$J_{\text{SRP}}(\boldsymbol{\tau}') = \sum_{i=1}^M \sum_{j=1}^M \int_{-\pi}^{\pi} \Psi_{ij}(\omega) G_{x_i x_j}(\omega) \exp(j\omega(\tau'_i - \tau'_j)) d\omega, \quad (21)$$

$$\boldsymbol{\tau}' = [\tau'_1, \dots, \tau'_M]^T. \quad (22)$$

Also, the TDOA estimation that maximizes an SRP  $J_{\text{SRP}}(\boldsymbol{\tau})$  can be expressed by a quadratic form of the generalized correlation matrix  $\boldsymbol{\Psi}(\omega) \oplus \boldsymbol{R}(\omega)$  and the steering vector  $\mathbf{b}(\omega; \boldsymbol{\tau}')$  as follows:

$$J_{\text{SRP}}(\boldsymbol{\tau}') = \mathbf{b}(\omega; \boldsymbol{\tau}')^H (\boldsymbol{\Psi}(\omega) \oplus \boldsymbol{R}(\omega)) \mathbf{b}(\omega; \boldsymbol{\tau}'), \quad (23)$$

$$\mathbf{b}(\omega; \boldsymbol{\tau}') = \frac{1}{\sqrt{M}} [\exp(-j\omega\tau'_1), \dots, \exp(-j\omega\tau'_M)]^T, \quad (24)$$

$$\boldsymbol{\Psi} = [\Psi_{ij}(\omega)]_{ij}, \quad (25)$$

$$\begin{aligned} \boldsymbol{R}(\omega) &= E_n[\mathbf{x}(\omega, n) \mathbf{x}^H(\omega, n)] \\ &= [G_{x_i x_j}(\omega)]_{ij}, \end{aligned} \quad (26)$$

$$\mathbf{x}(\omega, n) = [x_1(\omega, n), \dots, x_M(\omega, n)]^T, \quad (27)$$

where,  $[\cdot]^H$  denotes the complex conjugate transpose,  $\oplus$  denotes the Hadamard product,  $\Psi_{ij}(\omega)$  is the GCC filter for the  $i$ th channel pair, and  $[\cdot]_{ij}$  denotes a matrix consisting of the argument as its  $(i, j)$  entry. Note that TDOA estimation using the SRP method when  $M = 2$  is equivalent to that using the GCC method because

$$J_{\text{SRP}}([\tau'_1, \tau'_2]^T) = 2J_{\text{GCC}}(\tau'_2 - \tau'_1) + \sum_{i=1}^2 \int_{-\pi}^{\pi} \Psi_{ii}(\omega) G_{x_i x_i}(\omega) d\omega \propto J_{\text{GCC}}(\tau'_2 - \tau'_1). \quad (28)$$

#### D. Design of GCC Filter

The GCC and SRP methods sharpen the peaks of the cross-correlation function through the whitening of the observation using whitening filters  $\Psi_{ij}(\omega)$  appropriately designed for the surrounding environment. While several filters have been proposed on the basis of different criteria, we introduce the GCC-SCOT and GCC-PHAT filters as examples.

- GCC-SCOT filter

$$\Psi_{ij}(\omega) = \frac{1}{\sqrt{G_{x_i x_i}(\omega) G_{x_j x_j}(\omega)}}. \quad (29)$$

This filter is known to be robust against noise.

- GCC-PHAT filter

$$\Psi_{ij}(\omega) = \frac{1}{|G_{x_i x_j}(\omega)|}. \quad (30)$$

This filter has good robustness against reflections.

### IV. PROPOSED METHOD

In this section, we propose a new extension of GCC-based TDOA estimation, mapped SRP, by using higher-dimensional mapping to improve the resolution of TDOA analysis. The TDOA analysis of mapped SRP is based on a search for a steering vector, similarly in the SRP method. First, both the two-dimensional steering and observed signal vectors are mapped nonlinearly onto a higher-dimensional space. Next, SRP analysis is conducted in the higher-dimensional space to find the peak of the cross-correlation.

#### A. TDOA Estimation Algorithm of Mapped SRP

Here we consider a function  $\phi: \mathbb{C}^2 \rightarrow \mathbb{C}^M$  that maps the elements in the two-dimensional space of the observed signals onto another space with dimensionality  $M > 2$ . We assume that the map  $\phi$  satisfies the following three conditions.

- 1) The magnitude relation of the norm is retained:

$$\|\mathbf{x}\| \geq \|\mathbf{y}\| \rightarrow \|\phi(\mathbf{x})\| \geq \|\phi(\mathbf{y})\|, \quad (31)$$

where  $\|\cdot\|$  denotes the norm.

- 2) The origin remains intact:

$$\mathbf{x} = \mathbf{0} \rightarrow \phi(\mathbf{x}) = \mathbf{0}. \quad (32)$$

- 3) The orthogonality between vectors is preserved:

$$\mathbf{x}^H \mathbf{y} = 0 \rightarrow \phi^H(\mathbf{x}) \phi(\mathbf{y}) = 0. \quad (33)$$

The correlation matrix  $\mathbf{R}^\phi(\omega)$  of the mapped observed signal  $\phi(\mathbf{x}(\omega, n))$  is obtained as

$$\mathbf{R}^\phi(\omega) = E_n [\phi(\mathbf{x}(\omega, n)) \phi^H(\mathbf{x}(\omega, n))]. \quad (34)$$

We also map the steering vector  $\mathbf{b}(\omega; \tau)$  given by

$$\mathbf{b}(\omega; \tau) = [1, \exp(-j\omega\tau)], \quad (35)$$

and the TDOA  $\Delta\tau$  is estimated as the delay  $\tau$  that maximizes the SRP score  $J^\phi(\tau)$  in the mapped  $M$ -dimensional space:

$$\Delta\tau \leftarrow \arg \max_{\tau} J^\phi(\tau), \quad (36)$$

$$J^\phi(\tau) = \int_{-\pi}^{\pi} J^\phi(\omega; \tau) d\omega, \quad (37)$$

$$J^\phi(\omega; \tau) = \phi(\mathbf{b}(\omega; \tau))^H (\Psi^\phi(\omega) \oplus \mathbf{R}^\phi(\omega)) \phi(\mathbf{b}(\omega, \tau)), \quad (38)$$

$$\Psi^\phi(\omega) = [\Psi_{ij}^\phi(\omega)]_{ij}, \quad (39)$$

where  $\Psi_{ij}^\phi$  is the whitening filter of the  $i$ th and  $j$ th dimensions of the mapped observed signal vector  $\phi(\mathbf{x}(\omega, n))$ , whose design is described later.

The TDOA estimations using SRP and mapped SRP, given by Eqs. (20) and (36), respectively, are conducted by searching for a steering vector having higher linear dependence with the true steering vector. The linear independence between two different vectors is increased by the map  $\phi$  owing to the increase in the dimensionality. Therefore, the peak of the SRP becomes clearer and the resolution of TDOA estimation is improved.

#### B. Mapping for Analysis of 2dth-Order Moments

As the function  $\phi$  satisfying the conditions given by Eqs. (31)–(33), we employ the mapping  $\phi_d: \mathbb{C}^2 \rightarrow \mathbb{C}^{2^d}$  used in the MUSIC-based DOA analysis in [8]:

$$\phi_d(\mathbf{x}(\omega, n)) = \left[ \prod_{l=1}^d x_{c_{1l}}^{l\otimes}, \dots, \prod_{l=1}^d x_{c_{2^d l}}^{l\otimes} \right]^T, \quad (40)$$

$$x^{l\otimes} = \begin{cases} x & (\text{if } l \text{ is odd}) \\ x^* & (\text{otherwise}), \end{cases} \quad (41)$$

where  $d$  is the degree of the mapping,  $c_{kl}$  is an index specifying the entry number of vector  $\mathbf{x}$ , which is the  $(k, l)$  entry of the  $2^d \times d$  index matrix  $\mathbf{C}$  with  $d$  repeated permutations of 1 and 2 as its row vectors:

$$\mathbf{C} \triangleq [c_{kl}]_{kl} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & \dots & 2 \end{bmatrix}. \quad (42)$$

As discussed in [8], the entries of the correlation matrix  $\mathbf{R}^\phi(\omega)$  in Eq. (34) correspond to 2dth-order moments. This characteristic reveals a trade-off between the resolution and the robustness when choosing the degree  $d$ . While the mapping  $\phi_d$  with a large value of  $d$  improves the resolution of TDOA estimation owing to the increase in the dimensionality, given by the order of  $2^d$ , it loses the robustness against outliers because of the large variance of the high-order moment estimation. Thus, the degree  $d$  has to be chosen carefully. In the following experiments, we fixed  $d$  to 2 so that robust TDOA estimation could be obtained with a short observation.

Note that the computational complexity of the proposed method increases on the order of  $2^{2d}$ -fold to obtain the mapped correlation matrix compared with the conventional GCC.

TABLE I  
EXPERIMENTAL CONDITIONS.

Room size	Room 1: 3.4 × 4 × 2.7 m Room 2: 5.5 × 9 × 2.6 m
Microphone spacing	0.2 m
Reverberation time ( $T_{60}$ )	Room.1: 0.3 s, Room.2: 0.8 s
Noise type	Diffuse pink noise
SNR	0, -5, -10 dB
Frame length	512 samples
Frame shift length	128 samples
Frequency band	0.2–4 kHz
Audio signal	18 male and 14 female utterances
Sound source	Loudspeaker 1 m from array
Source directions	Seven azimuths of $-90^\circ, -60^\circ, \dots, 90^\circ$
Observation length	1 s

### C. Filters used in Mapped SRP

Here we describe the filter used in mapped SRP. Their characteristics are also similar to those for GCC.

- Mapped SRP-SCOT filter

$$\Psi'_{ij}(\omega) = \frac{1}{\sqrt{E_n \left[ \left| \Pi_{l=1}^d x_{c_{il}}^{l\otimes} \right|^2 \right] E_n \left[ \left| \Pi_{l=1}^d x_{c_{jl}}^{l\otimes} \right|^2 \right]}}. \quad (43)$$

- Mapped SRP-PHAT filter

$$\Psi'_{ij}(\omega) = \frac{1}{\left| E_n \left[ \left( \Pi_{l=1}^d x_{c_{il}}^{l\otimes} \right) \left( \Pi_{l=1}^d x_{c_{jl}}^{l\otimes} \right)^* \right] \right|}. \quad (44)$$

## V. EXPERIMENT

To confirm the effectiveness of mapped SRP, we conducted an experiment to evaluate the TDOA estimation accuracy. As conventional methods for comparison, we implemented GCC-SCOT and GCC-PHAT, which are discussed in Sect. III, MUSIC, and 2q-MUSIC ( $q = 2$ ). Note that we excluded the result for GCC-ML, a well-known whitening filter used in the GCC method because it did not perform well in this experiment owing to its weakness against noise and reverberation as pointed out in [6].

### A. Experimental Conditions

We employed convolutive mixtures of clean speech with impulse responses recorded in two different rooms with reverberation times of 0.3 s and 0.8 s. In order to simulate noisy observation, diffused pink noises [9] with three different SNRs of 0, -5, -10 dB were superimposed on the observed signals. Table I shows the other experimental conditions.

For the evaluation, we employed the root mean squared error (RMSE) between the estimated TDOA and the true source TDOA.

### B. Experimental Results

Figure 1 shows the experimental results under different conditions. In this TDOA estimation scenario of two-microphone pair, We can see no clear superiority of the MUSIC-based methods, which usually performs better when more channels are available. When the noise and reverberation are relatively weak, all the methods perform very good and TDOA estimation accuracy of the proposed method is similar to those of the conventional GCC methods. However, when the noise or the reverberation becomes stronger, the TDOA estimation

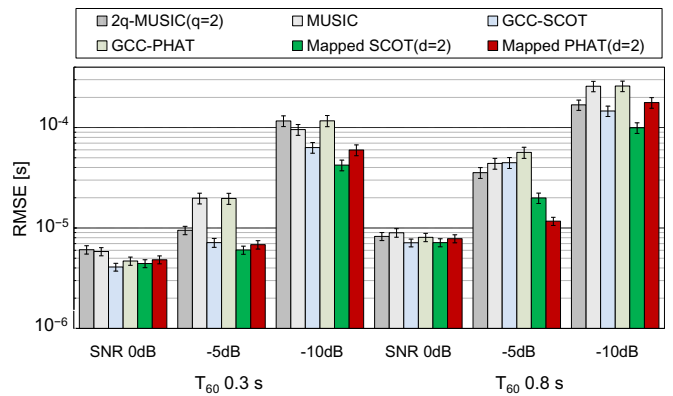


Fig. 1. TDOA estimation accuracy.

accuracy of the proposed method is better than those of the conventional methods. As a result, we find that our proposed mapped SRP performs better than the conventional methods when the noise and reverberation are heavy.

## VI. CONCLUSIONS

In this paper, we proposed a new TDOA estimation method called mapped SRP, which conducts SRP analysis of an observation nonlinearly mapped onto a higher dimensional space. The increased dimensionality improves the linear independence between different steering vectors, and the TDOA estimation by SRP analysis achieves improved resolution. The covariance of the mapped observation analyzed in mapped SRP corresponds to higher-order moments. We confirmed that the SRP with fourth order moment analysis achieves higher TDOA estimation accuracy than the conventional methods under in the case of large noise or reverberation.

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 23240023.

## REFERENCES

- [1] M. Brandstein and W. Darren, Eds., *Microphone Arrays*, Springer Verlag, Berlin, 2001.
- [2] K. Hasegawa, N. Ono, S. Miyabe, and S. Sagayama, "Blind estimation of locations and time offsets for distributed recording devices," *Proc. LVA/ICA*, pp.57–64, 2010.
- [3] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust. Speech Signal Process.*, vol.24, pp.320–327, 1976.
- [4] H. F. Silverman, Y. Yu, J. M. Sachar, and W. R. Patterson, "Performance of real-time source-location estimators for a large-aperture microphone array," *IEEE Trans. Speech Audio Process.*, vol.4, no.13, pp.593–606, 2005.
- [5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol.34, no.3, pp.276–280, 1986.
- [6] M. S. Brandstein, H. F. Silverman, "A robust method for speech signal time-delay estimation in reverberant rooms," *Proc. ICASSP*, vol.1, pp.375–378, 1997.
- [7] P. Chevalier, A. Ferreol, and L. Albera, "High resolution direction finding from higher order statistics: The 2q-music algorithm," *IEEE Trans. Signal Process.*, vol.54, no.8, pp.2986–2997, 2006.
- [8] Y. Sugimoto, S. Miyabe, T. Yamada, S. Makino, and B.-H. Juang, "Underdetermined DOA estimation by the nonlinear MUSIC exploiting higher-order moments," *Proc. IWAENC*, 2012.
- [9] I. A. McCowan and H. Bourslard, "Microphone array post-filter based on noise field coherence," *IEEE Trans. Speech Audio Process.*, vol.11, no.6, pp.709–716, 2003.