Theoretical Analysis of Euclidean Distance Matrix Completion for Ad Hoc Microphone Array Calibration

Mohammad J. Taghizadeh[a,b]

Email: mohammad.taghizadeh@idiap.ch

[a] *Idiap Research Institute, Martigny, Switzerland*
[b] *École Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

Abstract

We consider the problem of ad hoc microphone array calibration where the distance matrix consisted of all microphones pairwise distances have entries missing corresponding to distances greater than $d_{\max}$. Furthermore, the known entries are noisy modeled through additive independent random variables with strictly sub-Gaussian distribution, $\textsc{Sub}(c^2(d))$ with a bounded constant dependent on the distance d between the microphone pairs. In this report, we exploit matrix completion approach to recover the full distance matrix. We derive the theoretical guarantees of microphone calibration performance which demonstrates that the error of calibrating a network of N microphones using matrix completion decreases as $O(N^{-1/2})$.

Keywords: Ad hoc microphone array calibration, Matrix completion, Euclidean distance matrix, Missing pairwise distances, Microphone localization.

---

1. Introduction

Ad hoc microphone arrays consist of a set of sensor nodes spatially distributed over the acoustic field, in an ad hoc fashion. Processing of the data acquired with distributed sensors involves challenges attributed to the issues such as asynchronous sampling and unknown microphone positions. We address the problem of finding the sensor positions also referred to as *microphone calibration*. Finding the correct positioning of the microphones plays a key role in distant audio processing tasks such as source localization [1, 2, 3], high-quality acquisition for distant speech separation [4, 5, 6] and recognition [7, 8, 9]. Recent advances in mobile computing and communication technologies enable using cell phones, PDA's or tablets as a flexible acquisition set-up providing an ad hoc network of microphones. However, the unknown prior information on relative positions of the microphones is a key problem to achieve effective data processing.

The state-of-the-art techniques for microphone calibration often require information about distances between all microphones. Estimation of the pairwise distances becomes unreliable as the distances between the microphones are increased [10, 11, 12, 13]. Hence, the purpose of this report is to enable microphone calibration when some of the pairwise distances are missing. The matrix consisted of the squared pairwise distances has very low rank (explained in Section 2.1). The low-rank property has been investigated in the past years to devise efficient optimization schemes for matrix completion, i.e. recovering a low-rank matrix from randomly known entries. Candès et al. [14] showed that a small random fraction of the entries are sufficient to reconstruct a low-rank matrix *exactly.* Keshavan et al. proposed a matrix completion algorithm known as OptSpace and showed its optimality [15]. Furthermore, they proved that their algorithm is robust against noise [16]. Drineas et al. [11] exploited the low rank property to recover the distance matrix. However, they assume a nonzero probability of obtaining accurate distances for any pair of sensors regardless of their distance. This assumption severely restricts the applicability of their result for the microphone array calibration problem.

In this report, we build on our recent work on estimation of the microphones pairwise distances using the coherence model of a diffuse field [17]. This approach implies a local connectivity constraint as the pairwise distances of the further microphones can not be estimated. We construct a matrix of all the pairwise distances with missing entries corresponding to the unknown distances. We exploit the low-rank property of the square of this matrix to enable estimation of all the pairwise distances using matrix completion approach. The goal of this report is to provide the theoretical guarantees to bound the error for ad hoc microphone calibration considering the local connectivity of the noisy known entries. In Section 2, the mathematical basis and the model used for the calibration problem are described. The theoretical guarantees of calibration error using matrix completion are established in Section 3.

2. Problem Formulation

2.1. Distance Matrix

Consider a distance matrix $\boldsymbol{D}_{N\times N}$ consisting of the distances between N microphones constructed as

$$\boldsymbol{D} = \left[d_{ij}\right], \quad d_{ij} = \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|, \quad i, j \in \{1, \ldots, N\}\ , \tag{1}$$

where d_{ij} is the Euclidean distance between microphones i and j located at $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. Therefore, $\boldsymbol{D}$ is a symmetric matrix and it is often full rank.

Let $\boldsymbol{X}_{N\times\zeta}$ denote the position matrix whose i^{th} row, $\boldsymbol{x}_i^T \in \mathbb{R}^{\zeta}$, is the position of microphone i in ζ-dimensional Euclidean coordinate where microphones are deployed and $.^T$ denotes the transpose operator. By squaring the elements of $\boldsymbol{D}$, we construct a matrix $\boldsymbol{M}_{N\times N}$ which can be written as

$$\boldsymbol{M} = \mathbf{1}_N\boldsymbol{\Lambda}^T + \boldsymbol{\Lambda}\mathbf{1}_N{}^T - 2\boldsymbol{X}\boldsymbol{X}^T\ , \tag{2}$$

where $\mathbf{1}_N \in \mathbb{R}^N$ is the all ones vector and $\boldsymbol{\Lambda} = (\boldsymbol{X} \circ \boldsymbol{X})\mathbf{1}_{\zeta}$ where $\circ$ denotes the Hadamard product. We observe that $\boldsymbol{M}$ is the sum of three matrices of rank 1, 1 and at most ζ respectively. Therefore, the rank of the squared distance matrix constructed of the elements $\boldsymbol{M}_{ij} = \left[d_{ij}^2\right]$ is at most $\zeta + 2$ [11]. For instance, if the microphones are located on a plane or shell of a sphere, $\boldsymbol{M}$ has rank 4 and if they are placed on a line or circle, the rank is exactly 3. Hence, there is significant dependency between the elements of $\boldsymbol{M}$ and exploiting this low-rank property is the core of the proposed method in this report. The maximum distance that can be computed by this method is assumed to be d_{max}. pairwise distances greater than d_{max} are missing implying a locality structure in the missing entries in the distance matrix $\boldsymbol{D}$ consisted of the pairwise distances. This locality constraint in distance estimation is a typical problem in ad hoc microphone arrays [18]. In addition, the computation algorithm can lead to deviation from the model resulting in unreliable estimates of the short distances causing random missing entries in $\boldsymbol{D}$. Furthermore, the known entries are noisy due to measurement inaccuracies and violation of diffuseness.

2.2. Objective

The noisy estimates of the pairwise distances is modeled as

$$\tilde{d}_{ij} = d_{ij} + w_{ij} \quad ; \quad \widetilde{\boldsymbol{D}} = \boldsymbol{D} + \boldsymbol{W} \,, \tag{3}$$

where w_{ij} is the measurement noise for distance d_{ij} and $\boldsymbol{W}$ is the corresponding measurement noise matrix. We introduce a noise matrix on the squared distance matrix as

$$\boldsymbol{Z} = \widetilde{\boldsymbol{M}} - \boldsymbol{M} = \widetilde{\boldsymbol{D}} \circ \widetilde{\boldsymbol{D}} - \boldsymbol{D} \circ \boldsymbol{D} \,, \tag{4}$$

where $\widetilde{\boldsymbol{M}}$ is the noisy squared distance matrix.

There are two kinds of missing entries. The first group is consisted of the structured missing entries corresponding to the distances greater than d_{max}. We denote this group by S defined as

$$S = \{(i, j) : \; d_{ij} \geq d_{max}\} \,, \tag{5}$$

where $d_{ij} = \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|$. These structured missing entries are denoted by a matrix

$$\boldsymbol{D}^s_{ij} = \begin{cases} \boldsymbol{D}_{ij} & \text{if } (i, j) \in S \\ 0 & \text{otherwise} \end{cases}$$

Thus, the noiseless recognized pairwise distance matrix is given by

$$\boldsymbol{D}^{\bar{s}} = \boldsymbol{D} - \boldsymbol{D}^{s} \ ,$$

and we obtain the known squared distance matrix as

$$\begin{aligned} \boldsymbol{M}^{s} &= \boldsymbol{D}^{s} \circ \boldsymbol{D}^{s} \\ \boldsymbol{M}^{\bar{s}} &= \boldsymbol{D}^{\bar{s}} \circ \boldsymbol{D}^{\bar{s}} = \boldsymbol{M} - \boldsymbol{M}^{s} \ . \end{aligned} \tag{6}$$

Considering the noise on the known entries, we obtain

$$\widetilde{\boldsymbol{M}}^{\bar{s}} = \boldsymbol{M}^{\bar{s}} + \boldsymbol{Z}^{\bar{s}} \ , \tag{7}$$

where $\boldsymbol{Z}^{\bar{s}}$ denotes the noise on the known entries in the squared distance matrix.

For modeling the random missing entries, we assume that each entry is sampled with probability p. Sampling can be introduced by a projection operator on an arbitrary matrix $\boldsymbol{Q}_{N\times N}$, given by

$$\Psi_E(\boldsymbol{Q})_{ij} = \begin{cases} \boldsymbol{Q}_{ij} & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $E \subseteq [N] \times [N]$ denotes the known entries after random erasing process and has cardinality $|E| \approx pN^2$. Therefore, the final recognized squared distance matrix is given by

$$\boldsymbol{M}^{E} = \Psi_E(\widetilde{\boldsymbol{M}}^{\bar{s}}) \ . \tag{9}$$

The goal of the matrix recovery algorithm is to find the missing entries and remove the noise, given matrix $\boldsymbol{M}^{E}$.

2.3. Noise Model

The level of noise in extracting the pairwise distances, w_{ij} in (3), increases as the distances grow (In many scenario in sensor localization we have same situation). We model this effect through

$$\boldsymbol{W} = \boldsymbol{\Upsilon} \circ \boldsymbol{D} \ ,$$

where the noise matrix $\boldsymbol{\Upsilon}_{N \times N}$ is consisted of i.i.d entries with sub-Gaussian distribution of a bounded constant ς^2, thus [16]

$$\mathbb{P}(|\boldsymbol{\Upsilon}_{ij}| \geq \beta) \leq 2\,\mathrm{e}^{-\frac{\beta^2}{2\varsigma^2}} \ . \tag{10}$$

Based on (7), $\boldsymbol{Z}^{\bar{s}}_{ij} = 2d^2_{ij}\boldsymbol{\Upsilon}_{ij} + d^2_{ij}\boldsymbol{\Upsilon}^2_{ij}$; thereby $\boldsymbol{Z}^{\bar{s}}_{ij}$ is also a sub-Gaussian random variable with a bounded constant $c(d_{ij}) = 2\varsigma d^2_{ij}$. The physical setup confines $|\boldsymbol{Z}^{\bar{s}}_{ij}| \leq 4a^2$ where a is the radius of the table.

2.4. Evaluation Measure

Extracting the absolute position of the microphones deployed in ζ dimensional space requires at least $\zeta + 1$ anchor points in addition to the distance matrix. Therefore, in a scenario that the only available information are pairwise distances, the evaluation measure must quantify the error in estimation of the *relative* position of the microphones thus robust to the rigid transformations (translation, rotation and reflection). Hence, we quantify the distance between the actual locations $\boldsymbol{X}$ and estimated locations $\hat{\boldsymbol{X}}$ as [19]

$$\begin{aligned} \mathrm{dist}(\boldsymbol{X}, \hat{\boldsymbol{X}}) =& \frac{1}{N} \left\| \boldsymbol{J}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{J} - \boldsymbol{J}\hat{\boldsymbol{X}}\hat{\boldsymbol{X}}^T\boldsymbol{J} \right\|_{\mathrm{F}} \ , \\ \boldsymbol{J} =& \ \mathbb{I}_N - (1/N)\mathbf{1}_N\mathbf{1}_N^T \end{aligned} \tag{11}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm and $\mathbb{I}_N$ is the $N \times N$ identity matrix. The distance measure stated in (11) is useful to compare the performance of different methods in terms of microphone array geometry estimation.

2.5. Matrix Completion

We recall our problem of having N microphones distributed on a space of dimension ζ. Hence, the squared distance matrix $\boldsymbol{M}$ has rank $\eta = \zeta + 2$, but it is only partially known. The objective is to recover $\boldsymbol{M}_{N\times N}$ of rank $\eta \ll N$ from a sampling of its entries without having to ascertain all the N^2 entries, or collect N^2 or more measurements about $\boldsymbol{M}$. The approach proposed through *matrix completion* relies on the fact that a low-rank data matrix carries much less information than its ambient dimension implies. Intuitively, as the matrix $\boldsymbol{M}$ has $(2N-\eta)\eta$ degrees of freedom[1], we need to know at least ηN of the row entries as well as ηN of the column entries reduced by η^2 number of the repeated values to recover the entire elements of $\boldsymbol{M}$.

Given $\boldsymbol{M}^E$ defined in (9), the matrix completion recovers an estimate of the distance matrix $\hat{\boldsymbol{M}}$ through the following optimization

$$\begin{array}{lll} \text{Minimize} & \operatorname{rank}(\hat{\boldsymbol{M}}) & \\ \text{subject to} & \hat{\boldsymbol{M}}_{ij} = \boldsymbol{M}_{ij}\,, & (i,j) \in E \end{array} \tag{12}$$

We use the procedure of OptSpace proposed by Keshavan et al. [16] for estimating a matrix given the desired rank η.

3. Theoretical Guarantees for Microphone Calibration

We denote the smallest singular value of the squared distance matrix by $\sigma_\eta(\boldsymbol{M})$. Based on the following theorem we guarantee that there is an upper bound on the calibration error which decreases by the number of microphones.

Theorem 1. There exist constants C_1 and C_2, such that the output $\hat{\boldsymbol{X}}$ satisfies

$$\operatorname{dist}(\boldsymbol{X}, \hat{\boldsymbol{X}}) \leq C_1 \frac{a^2}{p} + C_2\, \varsigma \frac{d_{max}^2}{\sqrt{pN}} \tag{13}$$

with probability greater than $1 - N^{-3}$, provided that the right-hand side is less than $\sigma_\eta(\boldsymbol{M})/N$.

[1] The degrees of freedom can be estimated by counting the parameters in the singular value decomposition (the number of degrees of freedom associated with the description of the singular values and of the left and right singular vectors). When the rank is small, this is considerably smaller than N^2 [20].

3.1. *Proof of Theorem 1*

The squared distance matrix $\boldsymbol{M} \in \mathbb{R}^{N\times N}$ with rank$-\eta$, singular values $\sigma_k(\boldsymbol{M})$, $k \in [\eta]$ and singular value decomposition $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^T$ is (μ_1, μ_2)-*incoherent* if the following conditions hold.

$\mathcal{A}_1$. For all $i \in [N]$: $\sum_{k=1}^{\eta} \boldsymbol{U}_{ik}^2 \le \eta\,\mu_1$.

$\mathcal{A}_2$. For all $i, j \in [N]$: $\left|\sum_{k=1}^{\eta} \boldsymbol{U}_{ik}(\sigma_k(\boldsymbol{M})/\sigma_1(\boldsymbol{M}))\boldsymbol{U}_{jk}\right| \le \sqrt{\eta}\,\mu_2$.

where without loss of generality, $\boldsymbol{U}^T\boldsymbol{U} = N\mathbb{I}$.

For a (μ_1, μ_2)-incoherent matrix $\boldsymbol{M}$, (14) is correct with probability greater than $1 - N^{-3}$; cf. [16]-Theorem 1.2.

$$\frac{1}{N}\|\boldsymbol{M} - \hat{\boldsymbol{M}}\|_{\mathrm{F}} \le \frac{C_1' \,\|\Psi_E(\boldsymbol{M}^s)\|_2 + C_2' \,\left\|\Psi_E(\boldsymbol{Z}^{\bar{s}})\right\|_2}{p\,N} \;, \tag{14}$$

provided that

$$|E| \ge C_1' N \kappa_\eta^2(\boldsymbol{M}) \max\left\{\mu_1 \eta \log N\,;\, \mu_1^2\eta^2\kappa_\eta^4(\boldsymbol{M})\,;\, \mu_2^2\eta^2\kappa_\eta^4(\boldsymbol{M})\right\} \;, \tag{15}$$

and

$$\frac{C_1' \,\|\Psi_E(\boldsymbol{M}^s)\|_2 + C_2' \,\left\|\Psi_E(\boldsymbol{Z}^{\bar{s}})\right\|_2}{p\,N} \le \sigma_\eta(\boldsymbol{M})/N \;, \tag{16}$$

where the condition number $\kappa_\eta(\boldsymbol{M}) = \sigma_1(\boldsymbol{M})/\sigma_\eta(\boldsymbol{M})$.

To prove Theorem 1, we show the correctness of the upper bound stated in (13) based on the following Theorems 2 and 3.

Theorem 2. There exists a constant C_1'', such that with probability greater than $1 - N^{-3}$,

$$\|\Psi_E(\boldsymbol{M}^s)\|_2 \le C_1''\, a^2\, N \;. \tag{17}$$

The proof of this theorem is explained in section 3.2.

Theorem 3. There exists a constant C_2'', such that with probability greater than $1 - N^{-3}$,

$$\left\|\Psi_E(\boldsymbol{Z}^{\bar{s}})\right\| \le C_2''\varsigma d_{\max}^2 \sqrt{pN} \;. \tag{18}$$

The proof of this theorem is explained in section 3.3.

On the other hand, the following condition holds for any arbitrary network of microphones [21]

$$\mathrm{dist}(\boldsymbol{X}, \hat{\boldsymbol{X}}) \leq \frac{1}{N} \|\boldsymbol{M} - \hat{\boldsymbol{M}}\|_{\mathrm{F}} \ . \tag{19}$$

Therefore, based on Theorem 2, Theorem 3 and the relations (14) and (19), the upper bound stated in (13) is correct where $C_1 = C_1' C_1''$ and $C_2 = C_2' C_2''$. A Journal paper has been submitted which shows the correctness of the conditions (15) and (16) along with the (μ_1, μ_2)-incoherency of $\boldsymbol{M}$. That finishes the proof of Theorem 1.

■

3.2. Proof of Theorem 2

The goal is to find the bound of the norm of the squared distance matrix with missing entries according to structures indicated by E and S. Based on (5) and (8), we define matrix $\mathcal{E}$ as

$$\mathcal{E}_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \cap S \\ 0 & \text{otherwise} \end{cases}$$

Both E and S are symmetric matrices, hence $\mathcal{E}$ is also symmetric. Due to the physical setup, we know that $\Psi_E(\boldsymbol{M})_{ij} \leq 4a^2$ for all $i, j \in [N]$ and from the norm definition we have

$$\|\Psi_E(\boldsymbol{M}^s)\|_2 \quad \leq \quad 4a^2 \max_{\|\mathbf{h}\|=\|\boldsymbol{\hbar}\|=1} \sum_{i,j} |h_i|\,|\hbar_j|\,\mathcal{E}_{ij} = 4a^2 \|\mathcal{E}\|_2 \ ,$$

where $\mathbf{h} = [h_1, h_2, ..., h_N]^T$ and $\boldsymbol{\hbar} = [\hbar_1, \hbar_2, ..., \hbar_N]^T$ are right and left eigenvectors of matrix $\mathcal{E}$. In order to bound $\|\mathcal{E}\|_2$, we first define a binomial random variable vector $\boldsymbol{\nu} = [\nu_1, \nu_2, ..., \nu_N]^T$ where

$$\nu_i = \sum_{j \in [N]} |\mathcal{E}_{ij}| \ . \tag{20}$$

Based on the Gershgorin circle theorem we have $\|\mathcal{E}\|_2 \leq \|\boldsymbol{\nu}\|_\infty$. Each entry in matrix $\mathcal{E}$ is one with probability $p\,q$ where q is the probability that the entry is included in structured missing entries or

$$q = \mathbb{P}\{|\boldsymbol{x}_i - \boldsymbol{x}_j| \geq d_{\max}\} .$$

Hence, we have

$$\mathbb{E}[\nu_i] = Npq \, , \tag{21}$$

For bounding $\mathbb{E}[\nu_i]$, it is necessary to bound q. The lowest probability of missing distances corresponds to the case that the microphone location with respect to the edge of the circular table has a distance more than $d_{\max}$ and the highest probability corresponds to the case that the microphone is located right at the edge of the table. The maximum of $d_{\max}$ is a. We denote the upper bound and lower bound with $q_{\max}(a, d_{\max})$ and $q_{\min}(a, d_{\max})$ respectively, therefore

$$q_{\min}(a, d_{\max}) \leq q \leq q_{\max}(a, d_{\max}) \, . \tag{22}$$

$q_{\min}(a, d_{\max}) = \max\{1 - (\frac{d_{\max}}{a})^2, 0\}$ and

$$q_{\max} = 1 - \frac{2\gamma}{\pi} + \frac{1}{2\pi}\sin 4\gamma + \frac{2\xi^2}{\pi}[2\gamma + \sin 2\gamma] - 2\xi^2 \, , \tag{23}$$

where $\xi = d_{\max}/2a$ and $\gamma = \sin^{-1}\xi$. Based on (21) and (22) we have

$$Npq_{\min}(a, d_{\max}) \leq \mathbb{E}[\nu_i] \leq Npq_{\max}(a, d_{\max}) \, . \tag{24}$$

By applying the Chernoff bound to ν_i we have

$$\mathbb{P}(\nu_i > (1+\epsilon)\mathbb{E}[\nu_i]) \leq 2^{-(1+\epsilon)\mathbb{E}[\nu_i]} \, ,$$

where ϵ is an arbitrary positive constant. Therefore, based on (24) we have

$$\mathbb{P}(\nu_i > (1+\epsilon)Np\,q_{\max}) \leq 2^{-(1+\epsilon)Np\,q_{\min}} \, . \tag{25}$$

By applying the union bound we have

$$\mathbb{P}(\max_{i\in[N]} \nu_i > (1+\epsilon)Np\, q_{\max}) \leq 2^{-(1+\epsilon)Np\, q_{\min}+\log_2 N} \,. \tag{26}$$

We assume that $q_{\min}$ grows as $O(\frac{\log_2 N}{N})$; this assumption indicates that the ratio of the structured missing entries with respect to N decreases as N grows or in other words, $d_{\max}$ increases as the size of the network N grows. Therefore, we have

$$\mathbb{P}(\max_{i\in[N]} \nu_i > (1+\epsilon)Np\, q_{\max}) \leq N^{-\theta} \,, \tag{27}$$

where the positive parameter $\theta = (1+\epsilon)p - 1$; by choosing $\epsilon \geq 4/p - 1$, with probability greater than $1 - N^{-3}$, we have

$$||\Psi_E(\boldsymbol{M}^s)||_2 \leq 4a^2 \max_{i\in[N]} \nu_i \,,$$

and based on (27)

$$||\Psi_E(\boldsymbol{M}^s)||_2 \leq 4a^2(1+\theta)q_{\max}N \,.$$

The value for $q_{\max}$ for large N based on (23) is $1/3 + \sqrt{3}/(2\pi)$. Anyway q_{max} is probability measure and is less than one. Therefore, we achieve

$$||\Psi_E(\boldsymbol{M}^s)||_2 \leq C_1'' a^2 \, N \,.$$

■

3.3. Proof of Theorem 3

Based on the noise model described in Section 2.3, $\boldsymbol{Z}^{\bar{s}}_{ij}$ is obtained as

$$\boldsymbol{Z}^{\bar{s}}_{ij} = d^2_{ij}\boldsymbol{\Upsilon}_{ij}\left(2 + \boldsymbol{\Upsilon}_{ij}\right) \approx 2d^2_{ij}\boldsymbol{\Upsilon}_{ij}, \tag{28}$$

where $d_{ij} \leq d_{\max}$ and based on concentration inequality for 1-Lipschitz function $\|.\|$ on random variables $\Psi_E(\mathbf{Z}^{\bar{s}})$ with zero mean and sub-Gaussian tail with parameter $4\varsigma^2 d_{\max}^4$(10), (28),

$$\mathbb{P}\left(\left| \left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\| - \mathbb{E}\left(\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\|\right) \right| > t \right) \leq \exp\left(\frac{-t^2}{8\,\varsigma^2 d_{\max}^4}\right). \tag{29}$$

By setting $t = 2d_{\max}^2 \sqrt{6\varsigma^2 \log N}$ we have

$$\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\| \leq \mathbb{E}\left(\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\|\right) + 2d_{\max}^2 \sqrt{6\varsigma^2 \log N} \tag{30}$$

with probability bigger than $1 - N^{-3}$. So we need to extract bound for expectation of $\Psi_E(\mathbf{Z}^{\bar{s}})$ that has symmetric random enties. By using Theorem 1.1 from [22],

$$\mathbb{E}\left(\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\|\right) \leq C_4\,\mathbb{E}\left(\max_{j\in[N]} \left\|\Psi_E(\mathbf{Z}^{\bar{s}}_{.j})\right\|\right) \tag{31}$$

Furthermore by using union bound and with apply Chernoff bound on the sum of independent random variables [16]

$$\mathbb{E}\left(\max_{j\in[N]} \left\|\Psi_E(\mathbf{Z}^{\bar{s}}_{.j})\right\|^2\right) \leq C_5 d_{max}^4\,\varsigma^2 pN \tag{32}$$

Since

$$\mathbb{E}\left(\max_{j\in[N]} \left\|\Psi_E(\mathbf{Z}^{\bar{s}}_{.j})\right\|\right) \leq \sqrt{\mathbb{E}\left(\max_{j\in[N]} \left\|\Psi_E(\mathbf{Z}^{\bar{s}}_{.j})\right\|^2\right)} \tag{33}$$

Base on relations (31), (32) and (33)

$$\mathbb{E}\left(\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\|\right) \leq C_6 d_{\max}^2 \varsigma \sqrt{pN} \tag{34}$$

By using (34) and (30) for $pN \gg \log N$ we have

$$\left\|\Psi_E(\mathbf{Z}^{\bar{s}})\right\| \leq C_2'' d_{\max}^2 \varsigma \sqrt{pN}$$

$\blacksquare$

4. Conclusions

Rigorous analysis on the calibration of ad hoc microphone arrays are provided when the pairwise distances are only partially observed in a noisy setting. We exploited the matrix completion algorithm to recover the entire distance matrix and established theoretical guarantees for the error of microphone calibration. This analysis elucidates that the calibration error decreases as the local connectivity and the size of the network grows. The proposed algorithm and theoretical guarantees are applicable to the general ad hoc sensor array scenarios.

Acknowledgment

This work was supported by the Swiss National Science Foundation under the National Center of Competence in Research (NCCR) on "Interactive Multi-modal Information Management" (IM2).

References

[1] A. Asaei, M. J. Taghizadeh, M. Bahrololum, and M. Ghanbari, "Verified speaker localization utilizing voicing level in split-bands," *Signal Processing*, vol. 89(6), 2009.

[2] M. J. Taghizadeh, P. N. Garner, H. Bourlard, H. R. Abutalebi, and A. Asaei, "An integrated framework for multi-channel multi-source localization and voice activity detection," in *IEEE Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, 2011.

[3] A. Asaei, H. Bourlard, M. J. Taghizadeh, and V. Cevher, "Model-based sparse component analysis for reverberant speech localization," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2014.

[4] A. Asaei, M. Davies, H. Bourlard, and V. Cevher, "Computational methods for structured sparse component analysis of convolutive speech mixtures," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012.

[5] A. Mahmoodzadeh, H. Abutalebi, H. Soltanian-Zadeh, and H. Sheikhzadeh, "Single channel speech separation with a frame-based pitch range estimation method in modulation frequency," in *IEEE 5th International Symposium on Telecommunications (IST)*, 2010.

[6] A. Asaei, M. Golbabaee, H. Bourlard, and V. Cevher, "Structured sparsity models for reverberant speech separation," *IEEE/ACM Transactions on Acoustics, Speech, and Signal Processing*, vol. 22(3), 2014.

[7] C.-T. Do, M. J. Taghizadeh, and P. N. Garner, "Combining cepstral normalization and cohlear implant-like speech processing for microphone array-based speech recognition," in *IEEE workshop on Spoken Language Technology (SLT)*, 2012.

[8] A. Saeb and F. Razzazi, "A fast compressive sensing approach for phoneme classification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012.

[9] A. Asaei, H. Bourlard, and P. N. Garner, "Sparse component analysis for speech recognition in mullti-speaker environment," in *Proceedings of INTERSPEECH*, 2010.

[10] I. McCowan, M. Lincoln, and I. Himawan, "Microphone array shape calibration in diffuse noise fields," *IEEE Transactions on Audio,Speech and Language Processing*, vol. 16(3), 2008.

[11] P. Drineas, M. Javed, M. Magdon-Ismail, G. Pandurangant, R. Virrankoski, and A. Savvides, "Distance matrix reconstruction from incomplete distance information for sensor network localization," *Sensor and Ad Hoc Communications and Networks*, vol. 2, 2006.

[12] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from connectivity in sensor networks," *IEEE Transactions on Parallel Distribute Systems*, vol. 15(11), 2004.

[13] A. Asaei, B. Raj, H. Bourlard, and V. Cevher, "Structured sparse coding for microphone array position calibration," in *Proceeding of 5th ISCA workshop on Statistical and Perceptual Audition, SAPA-SCALE Conference,*, 2012.

[14] E. J. Candes and Y. Plan, "Matrix completion with noise," *IEEE Signal Processing Magazine*, vol. 98(6), 2010.

[15] R. H. Keshavan and S. Oh, "OptSpace: A gradient descent algorithm on the grassman manifold for matrix completion," 2009, `arXiv:0910.5260`.

[16] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from noisy entries," *Journal of Machine Learning Research*, vol. 11, 2010.

[17] M. J. Taghizadeh, R. Parhizkar, P. N. Garner, and H. Bourlard, "Euclidean distance matrix completion for ad-hoc microphone array calibration," in *IEEE 18th International Conference in Digital Signal Processing*, 2013.

[18] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from mere connectivity," in *MobiHoc*, 2003, pp. 201–212.

[19] I. Borg and P. J. Groenen, "Modern multidimensional scaling theory and applications," in *Springer*, 2005.

[20] E. Candes and B. Recht, "Exact matrix completion via convex optimization," *Magazine Communications of the ACM*, vol. 55, 2012.

[21] R. Parhizkar, "Euclidean distance matrices: Properties, algorithms and applications," Ph.D. dissertation, École polytechnique fédérale de Lausanne EPFL, 2014.

[22] Y. Seginer, "The expected norm of random matrices," *Combinatorics, Probability and Computing*, vol. 9, 2000.