ROBUST MICROPHONE PLACEMENT FOR SOURCE LOCALIZATION FROM NOISY DISTANCE MEASUREMENTS

Mohammad J. Taghizadeh^{1,2,3}, Saeid Haghighatshoar³, Afsaneh Asaei², Philip N. Garner², Hervé Bourlard^{2,3}

¹Huawei European Research Center, Munich, Germany

²Idiap Research Institute, Martigny, ³École Polytechnique Fédérale de Lausanne, Switzerland

ABSTRACT

We propose a novel algorithm to design an optimum array geometry for source localization inside an enclosure. We assume a square-law decay propagation model for the sound acquisition so that the additive noise on the measured source-microphone distances is proportional to the distances regardless of the noise distribution. We formulate the source localization as an instance of the "Generalized Trust Region Subproblem" (GTRS) whose solution gives the location of the source. We show that by suitable selection of the microphone locations, one can tremendously decrease the noise-sensitivity of the resulting solution. In particular, by minimizing the noise-sensitivity of the source location in terms of sensor positions, we find the optimal noise-robust array geometry for the enclosure. Simulation results are provided to show the efficiency of the proposed algorithm.

Index Terms – Robust microphone placement, Source localization, Generalized Trust Region Subproblem (GTRS).

1. INTRODUCTION

Microphone arrays are widely used to enable high-quality distant audio acquisition. They are an essential part of a plethora of sound technologies ranging from source localization and separation to distant speech recognition [1–3] and from sound field analysis and monitoring to virtual reality and surveillance [4, 5]. The optimum microphone array placement is a fundamental design problem that seeks the best spatial positions of the microphones such that a certain performance measure in terms of energy or cost efficiency, estimation, detection or identification accuracy is guaranteed. The focus of this paper is the optimum microphone array geometry for source localization based on noisy observations of the source-microphone distances.

The prior art often formulate the sensor placement problem for linear measurement models and the optimization procedures are derived for a scalar cost related to the mean squared error covariance matrix. It is also referred to as an optimal experimental design problem [6] in which a gird of sensors at all locations is hypothesized and the best subset of M sensors out of G possible locations is selected where M is typically known [7–9]. This formalism leads to a non-convex Boolean optimization problem which incurs a combinatorial search over all the $\binom{G}{M}$ possible combinations. In [10] a convex relaxation technique is presented for additive Gaussian linear models and the performance measures are independent of the unknown parameter. Alternative to the convex optimization, the se-

ments [11, 12] or solved using greedy algorithms [13] or heuristics. Non-linear measurement models are frequently encountered in

lection is achieved based on the coherence of the sensor measure-

applications like source localization and tracking. The error covariance matrix for non-linear models is not always available in closed form, and it often depends on the unknown parameter, hence, alternative approaches or performance measures are considered. A sensor selection algorithm for observations related to non-linear models is proposed in [14] within the Bayesian and sequential design. In [15], sensor selection for target tracking based on extended Kalman filtering is developed, in which a selection is performed by designing an appropriate gain matrix for a non-linear measurement model in additive Gaussian noise; the error covariance matrix is computed from the past state estimates so the solution is suboptimal. An alternative sensor selection framework is proposed in [16] where a sparse selection vector is designed such that a certain Cramér-Rao bound optimality on the estimates is guaranteed. This framework enables optimization over the number of microphones as a cardinality minimization problem such that a specified performance bound on localization error is obtained. The optimization procedure relies on the minimum eigen value of the Fisher information matrix and the optimal source localization can not always be achieved.

In this paper, we consider a minimax approach to design the microphone array without making any assumptions on the source location or statistics of the measurement noise. The main idea is to find a geometry for the array that gives the minimum estimation error for the source location when the source location and noise values are selected adversarily. We use an optimization approach based on Generalized Trust Region Subproblem (GTRS) to design a function whose minimum gives the minimax optimal geometry. We show for the rectangular enclosure, one can find the optimal solution efficiently.

2. PROBLEM STATEMENT

2.1. Signal Model

We consider a simple scenario for source localization in a rectangular-shaped room with M microphones whose positions are denoted by $x_i, i \in [M]$. In this paper, we consider a very simple case where M = 4. The results can be extended to more general cases. Let s be the location of the source in the room. To find the location of the source, each microphone estimates its distance from the source denoted by $d_i, i \in [M]$. We suppose $d_i = ||x_i - s||(1 + \eta_i)$ where η_i is the relative measurement noise. We do not assume any specific distribution for η_i except that the random variable $\eta_i \in [-\delta_i, \delta_i]$ where $\delta_i \in [0, 1)$ is a fixed given number showing the amount of

Afsaneh Asaei acknowledges the SNSF 200021-153507 grant on PHASER project.

noise in measurements of each microphone.

We briefly explain why this is a good model for the measurements. If we assume the square-law propagation model for the sound wave, it immediately results that the received signal power in each microphone is proportional to the inverse-square of its distance from the source. If we consider an algorithm to estimate the distance from the input signal, the estimation variance will be proportional to the squared-distance from the source. Therefore, we model the noisy distances as $d_i = ||x_i - s|| + \eta_i ||x_i - s||$, where η_i is a random variable so the resulting noise model $\eta_i ||x_i - s||$ is also a random variable with a variance proportional to distance.

2.2. Algorithm for Source Localization

To recover the position of the source, we consider the following quartic cost function:

$$g(z) = \sum_{i=1}^{M} (\|z - x_i\|^2 - d_i^2)^2.$$
(1)

The optimal source location is recovered by finding the global minima of the cost function. Adding an auxiliary variable γ and defining $y^T = (z^T, \gamma)$ and the following matrices:

$$A = \begin{pmatrix} -2x_1^T & 1\\ -2x_2^T & 1\\ \vdots & \vdots\\ -2x_M^T & 1 \end{pmatrix}, b = \begin{pmatrix} d_1^2 - \|x_1\|^2\\ \vdots\\ d_M^2 - \|x_M\|^2 \end{pmatrix},$$
(2)

the minimization problem can be written as follows

$$\min_{y} \left\{ \|Ay - b\|^{2} : y^{T}Ly = 0 \right\},$$
(3)

where $L = \text{diag}(\mathbb{1}_{\kappa \times 1}, -1)$ is a diagonal matrix and κ is the dimension of the ambient Euclidean space containing the microphones. This is a special instant of a quadratic optimization under a single quadratic constraint known as "Generalized Trust Region Subproblem" (GTRS) whose global minimum can be efficiently computed. Specifically, we have the following theorem: Specially, by [17] and [18], $y \in \mathbb{R}^{\kappa+1}$ is an optimal solution of (3) if and only if there is a $\lambda \in \mathbb{R}$ such that

$$(ATA + \lambda L)y = ATb, yTLy = 0,$$
(4)

$$A^T A + \lambda L \succeq 0. \tag{5}$$

This system of equations can be efficiently solved for y and λ which in particular gives z the optimal position of the source.

2.3. Noisy Measurements and Minimax Design

If there are measurement noises, the d_i parameters in the vector b in equation (2) will be the noisy distances, thus the equations (4) and (5) give an estimate of the source location. The estimation precision highly depends on the geometry of the microphone array (matrix A), the real location of the source s and measurement noises η_i . In some applications, one might have good estimates of the statistics of the noise or mobility pattern of the source inside the room specially if the source location is repeatedly estimated during time. In that case, one might design the sensor array based on these prior information.

In this paper, we consider a minimax approach to design the microphone array. More precisely, without making any assumptions on the initial source location or statistics of the noise, we use an optimization approach based on the equations (1), (4) and (5), to find the minimax-optimal configuration for the location of the microphones. The main idea is to find a geometry for the array that gives the minimum estimation error for the source location when the source location and noise values are selected in an adversary manner. We design a function whose minimum gives the minimax optimal geometry. Although optimizing this objective function might be difficult (requiring an exhaustive search) for a general room shape, in most cases the symmetry of the problem helps to find the optimal solution efficiently. In this section, we derive the results for a simple rectangular enclosure. Let us define the following function:

$$g(z; X, s, \eta) = \sum_{i=1}^{M} (\|z - x_i\|^2 - \|x_i - s\|^2 (1 + \eta_i)^2)^2, \quad (6)$$

where $X = [x_1, \ldots, x_M]$ is a matrix consisting of all microphone locations, $\eta = [\eta_1, \ldots, \eta_M]$ is the set of all noises and s is the real location of the source. Notice that $g(z; X, s, \eta)$ is the same function as equation (1) which should be minimized with respect to z to find the location of the source where the structure of the function g and thus the estimation quality of the source location depends on (X, s, η) . Also, as a design parameter we can choose X but (s, η) might be revealed adversarially as far as s is inside the enclosure and $|\eta_i| \leq \delta_i$. Defining u = z - s, one can write g as follows

$$g(u; X, s, \eta) = \sum_{i=1}^{M} (\|u\|^2 - 2(x_i - s)^T u - \|x_i - s\|^2 t(\eta_i))^2,$$
(7)

where $t(\eta_i) = 2\eta_i + \eta_i^2$. In practice, $|\eta_i| \ll 1$ and for simplicity one can assume that $t(\eta_i) \approx 2\eta_i$ where $\eta_i \in [-\delta_i, \delta_i]$. In order to find the source location, one should find the minimum of the function $g(u; X, s, \eta)$ with respect to u with the only difference that the optimization region is a rectangle centered at -s rather than 0. To simplify the design we assume that the optimization of the function $g(u; X, s, \eta)$ is done over all \mathbb{R}^2 . Notice that this is still a worst-case assumption because, for example, if in the unrestricted case, the minimum of the function occurs out of the room boundary, one can always find a better estimate of the source location inside the room. Let us denote by $\hat{u}(X, s, \eta)$ the minimum of the function $g(u; X, s, \eta)$ and let us define $e(X, s, \eta) = ||\hat{u}||^2$ to be the error (variance) of the estimation. It is immediate to check that if there is no noise $\eta = 0$, the minimum point will be $\hat{u} = 0$, i.e., e(X, s, 0) = 0 and the source location is exactly identified. For the minimax design that we consider, we are interested in e_{opt} = $\min_X \max_{s,\eta} e(X, s, \eta)$ where the maximization over η_i is done over the range $\eta_i \in [-\delta_i, \delta_i]$ and the outer minimization gives the minimax-optimal array geometry that we are interested in.

To further analyze the problem, we convert it into an instance of GTRS problem. Let us define $\gamma = \|u\|^2$, $y^T = (u^T, \gamma)$ and

$$A = \begin{pmatrix} -2(x_1 - s)^T & 1\\ -2(x_2 - s)^T & 1\\ \vdots & \vdots\\ -2(x_M - s)^T & 1 \end{pmatrix}, b = \begin{pmatrix} 2\eta_1 \|x_1 - s\|^2\\ \vdots\\ 2\eta_M \|x_M - s\|^2 \end{pmatrix}.$$
 (8)

Then, one can formulate finding the optimal \hat{u} as in equation (3) where the global minimum is given by equations (4) and (5).

Let us define the center of mass and the covariance of the array

geometry by

$$\mu = \frac{1}{M} \sum_{i=1}^{M} x_i, \ \Sigma = \frac{1}{M} \sum_{i=1}^{M} (x_i - \mu) (x_i - \mu)^T.$$
(9)

Choosing a coordinate system with axes parallel to the edges of the room with the origin at the center of the room, from the symmetry of the problem (room shape and symmetry of η), it results that the minimax optimal array geometry must be symmetric with respect to horizontal and vertical axes (in particular, $\mu = 0$). Therefore, there are two types of configurations that we should consider:

- C1: microphones are located at the vertices of a rectangle with edges parallel to the walls of the room.
- C2: microphones are on the vertices of a rhombus with diagonals parallel to the walls.

Also, after some simplification, equations (4) and (5) for this case can be written as follows:

$$\begin{pmatrix} 4M(\Sigma + ss^T) + \lambda I_{\kappa} & 2Ms\\ 2Ms^T & M - \lambda \end{pmatrix} y = w, \ y^T L y = 0, \quad (10)$$

where

$$w = \begin{pmatrix} -4\sum_{i=1}^{M} \eta_i \| x_i - s \|^2 (x_i - s) \\ 2\sum_{i=1}^{M} \eta_i \| x_i - s \|^2 \end{pmatrix}$$

and I_{κ} is the identity matrix of order κ where κ is the ambient dimension of the microhoness (We take $\kappa = 2$). There are still parameters (s, η) in these equation. We prove the following proposition which specifies the worst-case source location in the minimax design.

Proposition 1. Let \hat{X} be the minimax array configuration, i.e., $\max_{s,\eta} e(\hat{X}, s, \eta) = \min_X \max_{s,\eta} e(\hat{X}, s, \eta)$. Let $(\hat{s}, \hat{\eta})$ be the worst source location and noise parameter, i.e., $e(\hat{X}, \hat{s}, \hat{\eta}) = \max_{s,\eta} e(\hat{X}, s, \eta)$. Then \hat{s} must be on the vertices of the rectangular enclosure.

Proof. We just provide a sketch of the proof. The main idea is that for both types of configurations C1 and C2, one can increase $||x_i - s||^2$ by moving *s* closer to the vertices of the enclosure. As the noise scales proportional to the distance, this is equivalent to increasing the noise parameter δ_i which can potentially give a larger value in the minimax term $\max_{s,\eta} e(\hat{X}, s, \eta)$. This implies that the worst source location must be on the vertices.

Conjecture 1. In the minimax design, the worst case for the noise parameter η is when η_i is either $+\delta_i$ or $-\delta_i$.

Proposition 1 and Conjecture 1, completely specify the worst (s, η) parameter. To find the minimax optimal geometry, one only needs to do a simple optimization over all symmetric configurations of type C1 and C2 by simply solving equation (10) for y and λ . Notice that the last component of y is a positive number corresponding to the resulting estimation error $||u||^2$ and the minimax configuration is the one minimizing this component.

3. EXPERIMENTAL RESULTS

In this section, we conduct some experiments to demonstrate the proposed theories. The goal is to find the minimax optimal microphone array configuration which minimizes the localization error for the worst source location and the worst noise distribution. We consider a rectangular enclosure of dimension $6.6 \times 3.6 \text{ m}^2$. The number of microphones is M = 4.

3.1. Robust Microphone Array Configuration

As we explained in Section 2.3, in the minimax design, we assume a worst case scenario for the source mobility inside the enclosure and the distribution of noise. Based on Proposition 1, the worst-case source position is when it is located in one of the corners of the rectangular enclosure which has been depicted by hashed circles in Fig. 1. To find the minimax-optimal array geometry, we run a simple (one-dimensional) optimization expressed in (10) over all symmetric configurations of type C1 and C2.

Fig. 1 illustrates the six configurations (1)–(6) obtained for different noise levels on the source-microphone distances, i.e. $\delta_i = \{0.01, 0.02, 0.05, 0.1, 0.2, 0.3\}$ respectively. The resulting position of microphones for different noise levels are depicted with hexagonal shape and with different colors. Notice that between two types of array configurations (C1 and C2), the optimal one is always of type C1. The difference between the microphone positions at the green configurations (1) and (2) (corresponding to $\delta_i = 0.01, 0.02$) is less than 5 cm so they are not distinguishable in the picture. One can observe that as the level of noise increases, the microphones positioned on rectangles move away from the y-axis towards the corners. The exact positions are at 15, 36, 63, 114 and 183 cm distance from the y-axis corresponding to the different values of δ_i as stated above.



Fig. 1: Robust microphone configurations for source localization using four microphones: The numbers show the placement of the microphones (cm) along the x-axis with respect to the origin located at the room center. The worst-case source location is at the corners of the enclosure depicted by hashed circles. The configurations (1)–(6) correspond to $\delta_i = \{0.01, 0.02, 0.05, 0.1, 0.2, 0.3\}$; for example if $\delta_i = 0.1$ the purple configuration (4) is obtained by solving equation (10). We can see that larger noise levels lead to the microphone placements closer to the corners to achieve a robust design.

The robust configurations obtained in this section do not exploit any prior knowledge on the source position or the noise distribution. In the next section, we assume that both the source mobility and the noise distribution are known and we find the average-optimal configuration using an exhaustive search over all possible microphones placements to find the configuration corresponding to the minimum average source localization error where the average is taken over the source location and the noise statistics.

3.2. Comparison with an Average-optimal Array Geometry

In the minimax-optimal design, the philosophy is to guard against the worst source location and measurement noise. This is a reasonable assumption if one does not have any prior knowledge about the mobility pattern of the source or the statistics of the noise. In some cases, it might be possible to know both the mobility of the source and the statistics of the noise, thus it would be possible to find an average-optimal array geometry where the average is taken over

Array Configuration	Relative noise standard deviation					
	0.01	0.02	0.05	0.1	0.2	0.3
Average-optimum	3.59	7.16	17.9	35.67	73.93	115.4
Robust	4.1	8.56	19.6	37.4	77.2	122.84
Compact	35.33	83.98	170.6	200.2	209.34	215.89
Random	10.44	14.8	35.76	61.57	111.7	173.7
Corner	6.04	12.24	30.1	60.8	98.74	152.33

Table 1: Localization error (cm) using different microphone placements at various Guassian noise with a relative standard deviation δ_i .

the distribution of the source and measurement noise. It will be interesting to know how the minimax-optimal design compares with this average-optimal design. In this section, we assume that the source is uniformly distributed inside the enclosure and the measurement noise is Gaussian. To find the average-optimum array configuration, the area of the room is discretized into a grid of 600 (uniform) cells. All of the $\binom{600}{4}$ array configurations are considered while the source is uniformly randomly sampled inside the room to quantify the average localization errors. The simulated noise which is added to the real distances between source and microphones, is $\eta_i ||x_i - s||$ where η_i is a random variable, which follows zero-mean Gaussian with different standard deviations $\delta_i = \{0.01, 0.02, 0.05, 0.1, 0.2, 0.3\}.$ For each source and array configuration, 50 realizations of the noise are considered and the average localization error for each configuration is quantified. The resulting average-optimum microphone placement is depicted in Fig. 2 where for each value of δ_i the configurations (1)-(6) are obtained accordingly.



Fig. 2: Average-optimum microphone configurations for source localization using four microphones. The numbers show the placement of the microphones (cm) with respect to the origin located at the room center. Number (1)-(6) corresponding to $\delta_i = \{0.01, 0.02, 0.05, 0.1, 0.2, 0.3\}$ accordingly. One can see that larger noise levels lead to larger apertures.

We can see that as the level of noise increases, two of the microphones (black ones) remain at fixed positions on the walls while the other two microphones positioned at the middle move towards the walls to minimize the mean (expected) source localization error. The exact placement of the middle microphones with respect to the origin of the room coordinates (room center) is indicated at the pictures; we can see that the positions move from 60 cm to 162 cm as the noise level is increased.

The first two rows of Table 1 compare the performance of the average-optimal and minimax design for this scenario. It is seen that minimax-optimal design performs very well (comparable with the optimal one).

3.3. Source Localization Performance

In this section, we evaluate the performance of source localization using different microphone array configurations. Table 1 lists the localization error for different microphone array design at different Gaussian noise levels. Each number is obtained by averaging over 600 arbitrary positions of the source where the noisy distances are given from 50 realizations. The first row corresponds to the average-optimum placement. The second row indicates the error for the robust configuration obtained through the proposed algorithm. We can see that although the robust configuration is different than the average-optimum one, the expected localization performance is very close to the optimal value at all noise levels. Hence, the proposed algorithm enables an efficient microphone array design to achieve robust source localization. Further empirical observations show that the localization error using the robust configuration is 40% less than the average-optimum configuration if the source is located at the enclosure corners (the worst-case scenario). It may be noted the the average-optimum configuration is obtained under the assumptions that the noise distribution is known (Guassian) and the source mobility is uniform. If these assumptions are violated, it leads to the degradation of the performance obtained from the average-optimum configuration. On the other hand, the robust configuration is achieved without any assumption on the noise distribution and source mobility. Hence, the performance can be generalized to other setups.

The third row of Table 1 presents the error if a compact circular microphone array of diameter 20 cm is used at the center of the room. We can see that using a compact microphone for localization leads to huge error which is up to 8 times bigger than the localization accuracy achieved using the robust design. This error increases quickly by increasing the noise on the distances. In addition, we evaluate the localization error if the microphones are positioned randomly. For this experiment, we choose 20 random setups and compute the average localization error for 50 realizations of the noisy distances. We can see that the localization error is about two times more than the robust configuration. Finally, the last row shows the localization error when the microphones are positioned at the corner of the room.

4. CONCLUSION AND FUTURE WORK

In this paper, we proposed a minimax design for a microphone array consisting of four microphones in a rectangular enclosure. We assumed a square-law decay propagation model for the sound and designed the array for the worst source location and statistics of the measurement noise. We proposed an efficient algorithm to identify the robust microphone array configuration to minimize the worstcase source localization error. We showed that this robust configuration yields the performance very close to the average-optimum design. The robust placement was also shown to achieve substantial improvement over the compact, ad hoc and heuristic microphone array configurations. As an extension, one can consider a more complicated signal model for the source consisting of reflections from the boundaries which can be characterized using the image-source model of multipath propagation.

References

- H. T. Do, "Robust cross-correlation-based methods for soundsource localization and separation using a large-aperture microphone array," Ph.D. dissertation, Brown University, 2011.
- [2] A. Asaei, M. Golbabaee, H. Bourlard, and V. Cevher, "Structured sparsity models for reverberant speech separation," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 22, no. 3, pp. 620–633, 2014.
- [3] A. Asaei, "Model-based sparse component analysis for multiparty distant speech recognition," Ph.D. dissertation, École Polytechnique Fédéral de Lausanne (EPFL), 2013.
- [4] T. Hiekkanen, T. Lempiainen, M. Mattila, V. Veijanen, and V. Pulkki, "Reproduction of virtual reality with multichannel microphone techniques," in *Proceeding of 122nd AES Convention*, 2007.
- [5] G. Valenzise, L. Gerosa, M. Tagliasacchi, E. Antonacci, and A. Sarti, "Scream and gunshot detection and localization for audio-surveillance systems," in *Advanced Video and Signal Based Surveillance*, 2007. AVSS 2007. IEEE Conference on, 2007, pp. 21–26.
- [6] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [7] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 54–69, 2005.
- [8] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements," *Signal Processing Magazine*, *IEEE*, vol. 22, no. 4, pp. 41–53, 2005.

- [9] V. Kekatos, G. B. Giannakis, and B. Wollenberg, "Optimal placement of phasor measurement units via convex relaxation," *IEEE Trans. on Power Systems*, vol. 27, no. 3, pp. 1521–1530, 2012.
- [10] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, 2009.
- [11] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," 2013.
- [12] A. Carmi and P. Gurfil, "Sensor selection via compressed sensing," *Automatica*, vol. 49, no. 11, pp. 3304–3314, 2013.
- [13] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Decision and Control* (*CDC*), 2010 49th IEEE Conference on. IEEE, 2010, pp. 2572–2577.
- [14] D. M. T. I. Ford and C. P. Kitsos, "Recent advances in nonlinear experimental design," *Technometrics*, vol. 31, pp. 49–60, 1989.
- [15] E. Masazade, M. Fardad, and P. K. Varshney, "Sparsitypromoting extended kalman filtering for target tracking in wireless sensor networks," *Signal Processing Letters, IEEE*, vol. 19, no. 12, pp. 845–848, 2012.
- [16] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *arXiv preprint arXiv*:1310.5251, 2013.
- [17] J. J. Moré, "Generalizations of the trust region problem," Optimization Methods and Software, vol. 2, pp. 189–209, 1993.
- [18] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *Signal Processing*, *IEEE Transactions on*, vol. 56, no. 5, pp. 1770–1778, 2008.