16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

ACOUSTIC SOURCE LOCALIZATION BY FUSING DISTRIBUTED MICROPHONE ARRAYS MEASUREMENTS

G. Prandi, G. Valenzise, M. Tagliasacchi, F. Antonacci, A. Sarti, S. Tubaro

Dipartimento di Elettronica e Informazione
Politecnico di Milano
P.zza Leonardo da Vinci, 32 20133 - Milano, Italy
{prandi, valenzise, tagliasa, antonacc, sarti, tubaro}@elet.polimi.it

ABSTRACT

This paper addresses the problem of localizing an acoustic source by optimally fusing the observations provided by distributed microphone arrays. We consider a two-step process. First, each array provides an estimate of the source position by measuring the TDOAs (time differences of arrival) between each pair of microphones of the array. Then, these estimates are optimally fused taking into account the geometry of the arrays, in order to provide the final estimate of the source location. We study the performance of this approach by providing a closed form expression of the Cramer-Rao bound for a few simple geometries.

1. INTRODUCTION

The problem of localizing a source by means of acoustic measurements collected by a microphone array has been thoroughly studied in the literature. A recent survey of the most relevant techniques can be found in [1]. A large class of these methods divides the problem in two steps: first, the time differences of arrivals (TDOA) between microphone pairs are estimated; then, the source location is obtained based on these measurements.

In some scenarios, especially when the total number of microphone sensors is large, it might be useful to split the localization problem among a set of microphone arrays. Each sub-array solves the localization problem independently, and the final estimate of the source location is provided by a central node that optimally fuses the estimates provided by the individual sub-arrays. The rationale behind splitting a large array into smaller sub-arrays lies mainly in the reduced cost in terms of computational complexity. It can be shown [2] that using all the possible TDOAs in a microphone array leads to better result in terms of localization precision. Thus, having M microphones on the whole, the computational cost due to the TDOA estimation of the *single* array strategy grows as $O(M^2)$, while the cost of distributing the computation among S sub-arrays, each with $N = M/S$ microphones, grows as $O(M^2/S)$. In addition, when the arrays are deployed as a sensor network, the cost of data communication between the microphones and the base station becomes also relevant, and it becomes more convenient to transmit the local estimate of the source position from each array instead of the whole acoustic signal from each sensor to the centralized processor.

In the literature, the problem of obtaining the source position by fusing multiple array estimates has been addressed, but mainly in the far-field scenario. In [3], the maximum likelihood estimation of the direction of arrival is studied, when a large array is partitioned into sub-arrays. In [4], a distributed processing scheme is proposed, which performs direction of arrival (DOA) estimation at the individual arrays and time-delay estimation between pairs of arrays. Therefore, partial communication between sub-arrays is enabled. To the authors' knowledge, the only previous work that addresses the localization problem in the near field scenario, by fusing multiple array observations, is [5]. Each sub-array computes a spatial likelihood function using a conventional SRP-PHAT technique. The spatial likelihood functions are blended together by properly weighting the different level of access that a microphone array has to different spatial positions. Nevertheless, the weighting is obtained experimentally, by probing the system in an off-line training phase. Also, no theoretical evidence is given about the performance of the proposed method. Recently, the relation between optimal microphone array geometries and the variance of location estimates has been formalized by Yang and Scheuing [2][6]. When the noise affecting TDOA measurements is Gaussian, the Cramer-Rao bound can be written in closed form, and it depends only on the noise covariance matrix and on the array geometry. In [2] necessary and sufficient conditions for the optimum array geometry which minimizes the bound are

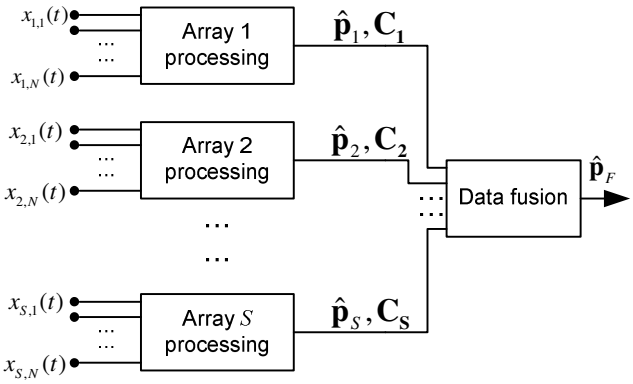


Fig. 1. Block diagram of the array fusion system.

provided, when TDOAs of all microphone pairs are collected. The work is extended in [6], where only a subset of TDOAs, all relative to a reference microphone, is adopted, such as in least squares based methods (e.g. spherical interpolation, spherical intersection, linear corrected least squares).

This paper analyzes the performance of a two-step localization process in a 2-dimensional space based on TDOA measurements. Each sub-array provides an estimate of the source location to a central processing node. Unlike [4], the scheme we study in this paper does not involve any data communication between sub-arrays. The central processing node performs optimal fusion, based on the incoming estimates and the knowledge of the array geometry. The Cramer-Rao bound is derived for a few simple array geometries for which the problem is mathematically tractable.

2. FUSION OF LOCALIZATION MEASUREMENTS

Consider the system illustrated in Figure 1, which consists of S distributed microphone arrays, each with N microphones. The source is located in $\mathbf{p} = [p_1, p_2]^T$. We assume that each array provides an unbiased estimate of the location of the acoustic source, i.e. $\hat{\mathbf{p}}_s$. Possibly, an estimate of the covariance matrix $\mathbf{C}_s = E[(\hat{\mathbf{p}}_s - \mathbf{p})(\hat{\mathbf{p}}_s - \mathbf{p})^T]$ is also given. As we shall see later, $\mathbf{C}_s$ depends on the actual estimator adopted, and it is typically related to the array and source geometry and to the measurement noise. The data fusion module receives $\hat{\mathbf{p}}_s$ and $\mathbf{C}_s$, $s = 1, \ldots, S$, and it computes the final estimate of the source, i.e. $\hat{\mathbf{p}}_F$.

The data model adopted by the fusion module is summarized by the following equation:

$$\hat{\mathbf{p}}_s = \mathbf{p} + \mathbf{w}_s, \tag{1}$$

where $\mathbf{w}_s \in \mathbb{R}^2$ denotes the position error in the estimate provided by array s, which is directly related to the error in the estimated TDOAs at each microphone. We postulate that no anomalies occur in the TDOA estimation process, i.e. there are no spurious peaks in the autocorrelation function due, for example, to reverberation [7]. If we assume that the SNR is high and no reverberation occurs, then the errors on the TDOA estimates are relatively small, and $\mathbf{w}_s$ can be modeled as a zero mean, Gaussian distributed random process with covariance matrix $E[\mathbf{w}_s \mathbf{w}_s^T] = \mathbf{C}_s$ (i.e. $\mathbf{w}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_s)$). Thus, the likelihood function can be written as

$$L(\mathbf{p}; \hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_S) = \prod_{s=1}^{S} \frac{1}{2\pi|\mathbf{C}_s|} e^{-\frac{1}{2}(\hat{\mathbf{p}}_s - \mathbf{p})^T \mathbf{C}_s^{-1} (\hat{\mathbf{p}}_s - \mathbf{p})}. \tag{2}$$

We assume that the data fusion module adopts a maximum likelihood estimator (MLE) to compute $\hat{\mathbf{p}}_F$. For linear data models and Gaussian noise as in (1), the MLE is unbiased and efficient, and its variance attains the Cramer-Rao lower bound (CRLB) also for a finite number of measurements S. If the noise pdf is unknown, the MLE still represents a meaningful estimator, as it is the best linear unbiased estimator (BLUE) [8].

In order to find the MLE, we need to compute the first order derivative of the log-likelihood function with respect to $\mathbf{p}$, and set it equal to zero, i.e.:

$$\begin{aligned} \frac{\partial \log L(\mathbf{p}; \hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_S)}{\partial \mathbf{p}} &= \frac{\partial}{\partial \mathbf{p}} \Bigg[-\log \sum_{s=1}^{S} 2\pi|\mathbf{C}_s| - \ldots \\ &\quad \ldots - \frac{1}{2} \sum_{s=1}^{S} (\hat{\mathbf{p}}_s - \mathbf{p})^T \mathbf{C}_s^{-1} (\hat{\mathbf{p}}_s - \mathbf{p}) \Bigg] \\ &= \sum_{i=1}^{S} \mathbf{C}_s^{-1} (\hat{\mathbf{p}}_s - \mathbf{p}) = 0. \end{aligned} \tag{3}$$

Therefore, the MLE of the source position is given by

$$\hat{\mathbf{p}}_F = \left(\sum_{s=1}^{S} \mathbf{C}_s^{-1} \right)^{-1} \cdot \left(\sum_{s=1}^{S} \mathbf{C}_s^{-1} \hat{\mathbf{p}}_s \right). \tag{4}$$

The value $\hat{\mathbf{p}}_F$ is a stationary point of the likelihood function (2). It is a point of maximum, since the second order derivative is negative definite. In fact:

$$\frac{\partial}{\partial \mathbf{p}} \left[\frac{\partial \log L(\mathbf{p}; \hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_S)}{\partial \mathbf{p}} \right] = -\sum_{s=1}^{S} \mathbf{C}_s^{-1}. \tag{5}$$

The CRLB is found as the element (i,i) of the inverse of the Fisher information matrix, i.e.

$$\mathrm{var}(\hat{p}_i) \geq [\mathbf{I}_F^{-1}(\mathbf{p})]_{(i,i)}, \tag{6}$$

where

$$\begin{aligned} \mathbf{I}_F(\mathbf{p}) &= -E \left\{ \frac{\partial}{\partial \mathbf{p}} \left[\frac{\partial \log L(\mathbf{p})}{\partial \mathbf{p}} \right] \right\} \\ &= \sum_{s=1}^{S} \mathbf{C}_s^{-1}. \end{aligned} \tag{7}$$

We notice that, if the MLE is used to perform data fusion, the CRLB of the final estimate $\hat{\mathbf{p}}_F$ depends on the covariance matrices $\mathbf{C}_s$, $s = 1, \ldots, S$ of the individual arrays. In the following section, we introduce an expression for $\mathbf{C}_s$, when the acoustic source localization is performed by means of TDOA measurements.

3. SOURCE LOCALIZATION FROM TDOA MEASUREMENTS

Each microphone array consists of N sensors at locations $\mathbf{q}_n$, $n = 1, \ldots, N$. If we consider the TDOA measurement noise to be zero-mean, Gaussian and i.i.d. with variance σ^2, the Fisher information matrix (FIM) is given by [2]:

$$I(\mathbf{p}) = \frac{\mathbf{G}\mathbf{G}^T}{(\sigma v)^2} \tag{8}$$

where v denotes the speed of sound and

$$\mathbf{G} = [\ldots, \mathbf{g}_{ij}, \ldots]_{(i,j) \in \{(i,j) | 1 \leq j < i \leq N\}} \tag{9}$$

$$\mathbf{g}_{ij} = \mathbf{g}_i - \mathbf{g}_j \tag{10}$$

$$\mathbf{g}_i = \frac{\mathbf{p}-\mathbf{q}_i}{\|\mathbf{p}-\mathbf{q}_i\|}, \quad i = 1,\ldots,N \tag{11}$$

i.e. $\mathbf{g}_i$ is the unit norm vector pointing from the ith microphone to the source.

The Fisher information matrix can be rewritten as [2]

$$I(\mathbf{p}) = \frac{N\mathbf{g}\mathbf{g}^T - (\mathbf{g}\mathbf{1})(\mathbf{g}\mathbf{1})^T}{(\sigma v)^2} \tag{12}$$

where $\mathbf{g} = [\mathbf{g}_1,\ldots,\mathbf{g}_N]$ and $\mathbf{1} = [1,1,\ldots,1]^T$.

4. A SIMPLE DISTRIBUTED ARRAY GEOMETRY

In order to make the following discussion mathematically tractable, we consider a simple geometry with S arrays. Each array consists of N microphones located along a circle centered in the source position, i.e. $\mathbf{p} = \mathbf{0}$. For the first array, i.e. $s = 1$, the microphone positions are given by

$$\mathbf{q}_n = R[\cos(\alpha_n), \sin(\alpha_n)]^T, \quad n = 1,\ldots,N, \tag{13}$$

where R is the radius of the circle, and

$$\alpha_n = \frac{2\pi n}{K} - \pi\left(\frac{N-1}{K}\right). \tag{14}$$

The value of $K \geq N$ determines the angular aperture of the array (see Figures 2 and 3). Large values of K correspond to small angular distances between microphones. When $K = N$, microphones are uniformly spaced over the whole circle, while $K \gg N$ means that the array is located on a small circular sector.

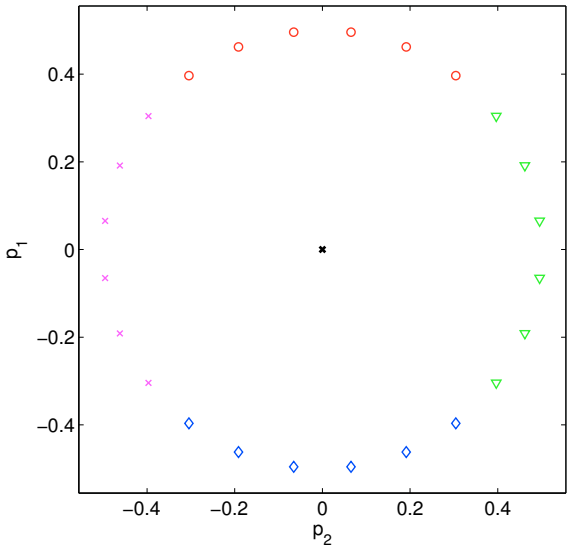
Also, we postulate that each array adopts an unbiased and efficient estimator, which attains the CRLB. Actually, such an estimator does not exist [2], since the MLE estimator for source location [1] is only *asymptotically* efficient, i.e. $\mathbf{C}_s > \mathrm{I}_s^{-1}$. Nevertheless, as shown in [9], the variance of the MLE is very close to the theoretical CRLB when just a small number of microphones (e.g. $N \geq 6$ sensors) is used for each array. Therefore, in the following we consider the sensor fusion performance in terms of the CRLB, which can be considered a tight bound on the estimator variance, provided that the number of microphones for each array is not too small, or equivalently, that the number of arrays S is not too large (see Figures 2-3).

For the problem at hand, $\mathbf{g}_n = -\mathbf{q}_n/R$, and the Fisher information matrix for the first array is given by

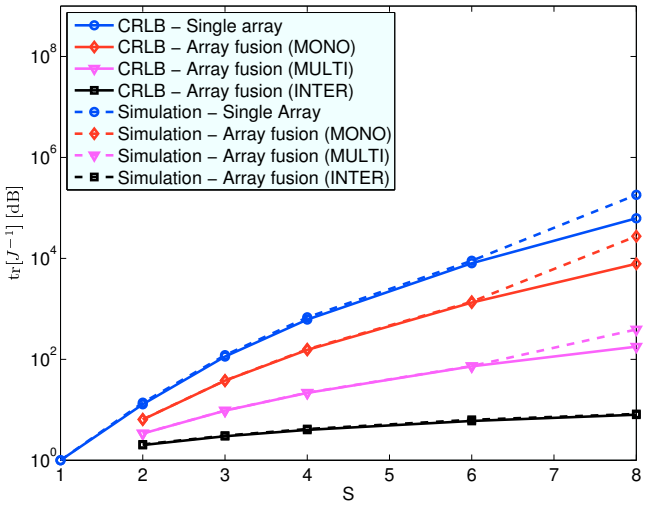
$$\mathbf{I}_1 = \frac{\mathbf{G_1}\mathbf{G_1}^T}{(\sigma_1 v)^2} = \frac{1}{(\sigma_1 v)^2}\begin{pmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_2 \end{pmatrix} \tag{15}$$

where

$$\begin{aligned} \frac{1}{\lambda_1} &= N\sum_{n=1}^{N}\cos^2(\alpha_n) - \left[\sum_{n=1}^{N}\cos(\alpha_n)\right]^2, \\ \frac{1}{\lambda_2} &= N\sum_{n=1}^{N}\sin^2(\alpha_n) \end{aligned} \tag{16}$$



(a) Multi-array geometry.



(b) Cramer-Rao Bound

Fig. 2. In the top part of the figure, multi-array geometry with $M = 24$ microphones, $S = 4$ arrays, and $K = SN = M$, i.e. each array is distributed over $1/4$ of the circle. In the bottom part, the trace of the inverse FIM for different kinds of array geometries, when the available number of microphones $M = 24$ is distributed across S arrays and $K = M$ ($K = N$ for interleaved configuration). The values of $\mathrm{tr}\{\mathbf{I}_F^{-1}\}$ are normalized by the trace of the inverse FIM of a single array with $S = 1$.

Therefore

$$\mathbf{C}_1 > \sigma^2 v^2 \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{17}$$

We assume that the microphone locations for the other arrays (i.e. $s = 2,\ldots,S$) are obtained by rotating the first array by an angle equal to α_s. By defining different values for α_s, we can obtain the following distributed geometries for a set of S microphone arrays:

- *mono-sector array*: $\alpha_s = 0, s = 2,\ldots,S$
- *multi-sector array*:
  - $S = 2$, $\alpha_2 = \pi/2$[1]
  - $S > 2$, $\alpha_s = \dfrac{2\pi(s-1)}{S}, s = 2,\ldots,S$
- *interleaved array*: $K = N$, $\alpha_s =$ any

If we assume that the source is isotropic and the gains of the microphones are all equal, $\sigma_1 = \sigma_2 = \ldots = \sigma_S = \sigma$.

In the following, we compare the three aforementioned geometries in terms of localization performance. To this end, as in [2], the metrics adopted is the trace of the inverse of the FIM, i.e. $\mathrm{tr}\{\mathbf{I}_F^{-1}\}$.

4.1. Mono-sector array

In this scenario, we have a superposition of S co-located arrays. Therefore, $\mathbf{C}_s = \mathbf{C}_1$, $s = 2,\ldots,S$ and

$$\begin{aligned}\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{MONO}} &= \frac{\mathrm{tr}\{\mathbf{C}_1\}}{S} \\ &> \frac{\mathrm{tr}\{\mathbf{I}_1^{-1}\}}{S} \\ &= \frac{\sigma^2 v^2(\lambda_1+\lambda_2)}{S}. \end{aligned} \tag{18}$$

An expression for $\mathrm{tr}\{\mathbf{I}_1^{-1}\}$ was found in [6], and it is equal to

$$\mathrm{tr}\{\mathbf{I}_1^{-1}\} = \frac{4\sigma^2 v^2}{N^2}\frac{1-\rho_2^2}{(1-\rho_1)(1+\rho_1-2\rho_2^2)}, \tag{19}$$

with

$$\rho_1 = \frac{\sin(2\pi N/K)}{N\sin(2\pi/K)}, \quad \rho_2 = \frac{\sin(\pi N/K)}{N\sin(\pi/K)}, \tag{20}$$

A more compact approximation of the previous expression can be obtained as follows. From [2] we have:

$$\mathrm{tr}\{\mathbf{I}_1^{-1}\} = \mathrm{tr}\{(\mathbf{G}\mathbf{G}^T)^{-1}\} \geq \frac{4\sigma^2 v^2}{\mathrm{tr}\{\mathbf{G}\mathbf{G}^T\}} \tag{21}$$

Although the equality holds only for arrays uniformly distributed along the circle, a good approximation is given by:

$$\mathrm{tr}\{\mathbf{I}_1^{-1}\} \sim \frac{4\sigma^2 v^2}{\mathrm{tr}\{\mathbf{G}\mathbf{G}^T\}}S^{3/2} = \frac{4\sigma^2 v^2 S^{3/2}}{N^2 - \left|\frac{\sin(\pi N/K)}{\sin(\pi/K)}\right|^2} \tag{22}$$

Therefore, we obtain

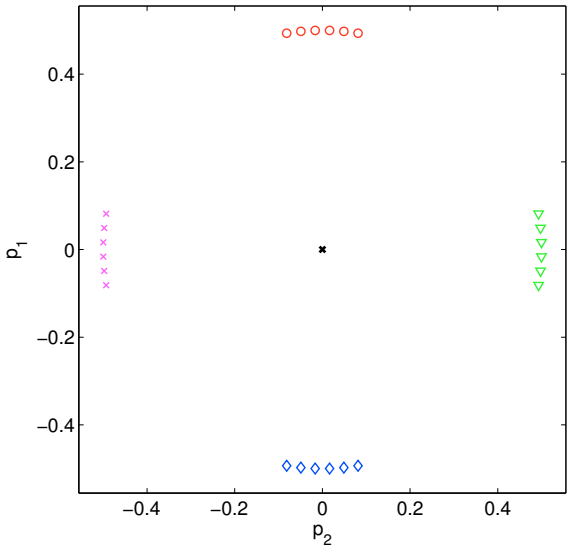
$$\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{MONO}} > \frac{4\sigma^2 v^2 S^{1/2}}{N^2 - \left|\frac{\sin(\pi N/K)}{\sin(\pi/K)}\right|^2} \tag{23}$$

4.2. Multi-sector array

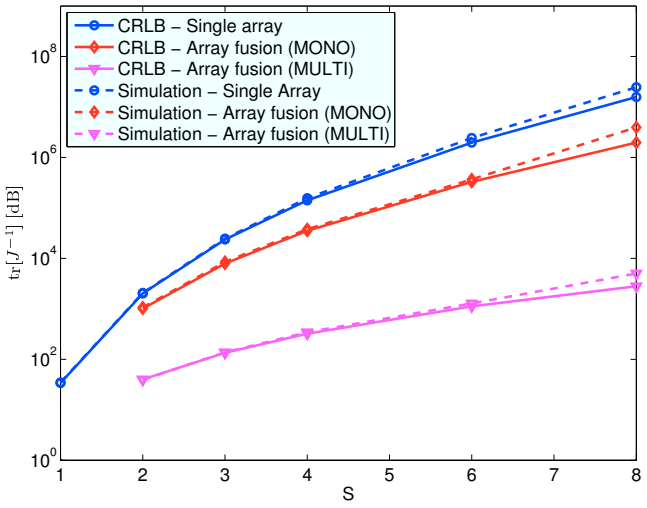
In this scenario each array spans a sector, and there are S distinct sectors uniformly spaced along the circle. Therefore, we can write

$$\mathbf{C}_s = \mathbf{T}_s \mathbf{C}_1 \mathbf{T}_s^T \tag{24}$$

[1] It can be shown that $\alpha_2 = \pi$ is not optimal, while $\alpha_2 = \pi/2$ is.



(a) Multi-array geometry.



(b) Cramer-Rao Bound

Fig. 3. In the top part of the figure, multi-array geometry with $M = 24$ microphones, $S = 4$ arrays, and $K = 4SN = 4M$, i.e. each array is distributed over $1/16$ of the circle. In the bottom part, the Cramer-Rao bound for different kinds of array geometries, when the available number of microphones $M = 24$ is distributed across S arrays and $K = 4M$. The values of $\mathrm{tr}\{\mathbf{I}_F^{-1}\}$ are normalized by the same value of Figure 2 for comparison.

where $\mathbf{T}_s$ is an orthonormal rotation matrix given by

$$\mathbf{T}_s = \begin{bmatrix} \cos(\alpha_s) & -\sin(\alpha_s) \\ \sin(\alpha_s) & \cos(\alpha_s) \end{bmatrix} \tag{25}$$

Setting

$$a = \sum_{s=1}^{S} \cos^2(\alpha_s), \quad b = \sum_{s=1}^{S} \sin^2(\alpha_s),$$
$$c = \sum_{s=1}^{S} \sin(\alpha_s)\cos(\alpha_s) \tag{26}$$

and combining (7) and (24), we get

$$\begin{aligned}
\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{MULTI}} &= \left(\sum_{s=1}^{S} \mathbf{C}_s^{-1}\right)^{-1} = \left(\sum_{s=1}^{S} \mathbf{T}_s \mathbf{C}_1^{-1} \mathbf{T}_s^T\right)^{-1} \\
&> \sigma^2\nu^2 \begin{bmatrix} \frac{a}{\lambda_1}+\frac{b}{\lambda_2} & \frac{c}{\lambda_1}+\frac{c}{\lambda_2} \\ \frac{c}{\lambda_1}+\frac{c}{\lambda_2} & \frac{b}{\lambda_1}+\frac{a}{\lambda_2} \end{bmatrix}^{-1} \\
&= \frac{2\sigma^2\nu^2}{S} \begin{bmatrix} \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2} & 0 \\ 0 & \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2} \end{bmatrix} \\
&= \frac{4\sigma^2\nu^2}{\mathrm{tr}\{\mathbf{G}\mathbf{G}^T\}} \cdot S^{-1} \\
&= \frac{4\sigma^2\nu^2 S^{-1}}{N^2 - \left|\frac{\sin(\pi N/K)}{\sin(\pi/K)}\right|^2}
\end{aligned} \tag{27}$$

4.3. Interleaved array

In this scenario, each array has uniformly spaced microphones along the circle. The covariance matrix $\mathbf{C}_1$ is given by

$$\mathbf{C}_1 = \sigma^2\nu^2 \begin{bmatrix} 2/N^2 & 0 \\ 0 & 2/N^2 \end{bmatrix} \tag{28}$$

Since $\mathbf{C}_s = \mathbf{C}_1$, $s = 2, \ldots, S$,

$$\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{INTER}} > \frac{\mathrm{tr}\{\mathbf{C}_1\}}{S} = \frac{4\sigma^2\nu^2}{SN^2} \tag{29}$$

5. SIMULATION RESULTS

It is interesting to compare the three geometries described in Section 4 in terms of localization accuracy, for the same total number of microphones M. Figure 2(b) shows the Cramer-Rao bound when $M = 24$, $N = M/S$ and $K = N$, normalized by $4\sigma^2\nu^2/M^2$, i.e. the optimal accuracy attained when the M microphones are uniformly distributed around the source to form a single array. This normalized value is referred to in the literature [6] as *geometric dilution of precision*, since it accounts for the loss induced by the array geometry. The plot shows the Cramer-Rao bound for the mono-sector, multi-sector and interleaved geometries, for a different number of sub-arrays. The plot also shows as a reference the Cramer-Rao bound for a single array of $N = M/S$ microphones. As an example, Figure 2(a) shows the multi-sector scenario when $S = 4$. We notice that the mono-sector geometry performs worse. By comparing (23) and (27), we observe that $\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{MONO}}/\mathrm{tr}\{\mathbf{I}_F^{-1}\}_{\mathrm{MULTI}} \sim S^{3/2}$, i.e. the gain of adopting a distributed array geometry increases with the number of sub-arrays. This is due to the fact that the geometric dilution of precision of the individual arrays is partially compensated by the fusion process only in the multi-sector scenario. The interleaved scenario achieves the best performance, since the individual arrays do not suffer from any geometric dilution of precision.

Figure 3(b) shows similar results, but when $K = 4N$, i.e. the angular aperture of the individual sub-arrays is 4 times smaller than before. Figure 3(a) shows an example when $S = 4$. Comparing Figure 3(b) with Figure 2(b), we notice that, for the same number of sub-arrays S, the localization accuracy decreases, due to the higher geometric dilution of precision of the individual arrays.

6. CONCLUSIONS

In this paper we investigated the Cramer-Rao bound of the localization accuracy, when the estimates provided by distributed sub-arrays are fused together in a central node. Future investigation will address the accuracy-resources tradeoffs that arise when computational and/or communication power is constrained.

7. REFERENCES

[1] Y. Huang, J. Benesty, and G. W. Elko, *Audio Signal Processing for the Next-Generation Multimedia Communication Systems*, chapter 9, Source Localization, Kluwer Academic Publishers, 2004.

[2] B. Yang and J. Scheuing, "Cramer-rao bound and optimum sensor array for source localization from time differences of arrival," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, March 2005.

[3] P. Stoica, A. Nehorai, and T. Soderstrom, "Decentralized array processing using the MODE algorithm," *Circuits, Systems and Signal Processing*, vol. 14, pp. 17–38, 1995.

[4] R. J. Kosick and B. M. Sadler, "Source localization with distributed sensor arrays and partial spatial coherence," *IEEE Trans. Signal Process.*, vol. 52, pp. 601–616, 2004.

[5] P. Aarabi, "The fusion of distributed microphone arrays for sound localization," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 338–347, 2003.

[6] B. Yang and J. Scheuing, "A theoretical analysis of 2d sensor arrays for TDOA based localization," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006.

[7] J. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 6, pp. 998–1003, 1982.

[8] S.M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.

[9] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, "Acoustic source localization in distributed sensor networks," *Conference Record of the 38th Asilomar Conference on Signals, Systems and Computers*, vol. 2, 2004.