# Energy Based Sensor Network Source Localization via Projection onto Convex Sets (POCS)

Doron Blatt and Alfred O. Hero, III

***Abstract*— This paper addresses the problem of locating an acoustic source using a sensor network in a distributed manner, i.e., without transmitting the full data set to a central point for processing. This problem has been traditionally addressed through the maximum likelihood framework or nonlinear least squares. These methods, even though asymptotically optimal under certain conditions, pose a difficult global optimization problem. It is shown that the associated objective function may have multiple local optima and saddle points and hence any local search method might stagnate at a sub-optimal solution. In this paper, we formulate the problem as a convex feasibility problem and apply a distributed version of the projection onto convex sets (POCS) method. We give a closed form expression for the projection phase, which usually constitutes the heaviest computational aspect of POCS. Conditions are given under which, when the number of samples increases to infinity or in the absence of measurement noise, the convex feasibility problem has a unique solution at the true source location. In general, the method converges to a limit point or a limit cycle in the neighborhood of the true location. Simulation results show convergence to the global optimum with extremely fast convergence rates compared to the previous methods.**

## I. Introduction

The problem of locating a source that emits acoustic waves using a wireless network of acoustic sensors has been addressed by several authors (see [1] and references therein). This problem has been traditionally solved though maximum likelihood, which is equivalent to nonlinear least squares estimation when the observation noise is modeled as a white Gaussian process. The maximum likelihood estimator is asymptotically optimal, it can be applied to both the cases of known and unknown source power, and offers a natural generalization to the multiple sources case [1]. However, there are two major drawbacks to the method of [1]: (a) it requires the transmission of a certain statistic from each node in the network to a central point for processing, and (b) the solution of a global optimization problem is required for the derivation of the estimator.

Rabbat and Nowak [2], [3] proposed a distributed implementation of the incremental gradient algorithm to solve the nonlinear least squares problem in a distributed manner, i.e., without the need to transmit the data to a central point for processing. The advantage of in-network computation relative to the fusion center approach in terms of communication bandwidth and energy consumption is well documented in the literature (see e.g. [4] and references therein). As in [4] our premise is that as the network becomes denser, it is cheaper to perform several communication cycles across the network than to transmit the data from each sensor to a central point.

A drawback of the method in [2], [3], or any other local search method, is that it is sensitive to local optima and saddle points. As will be shown below, the objective function associated with this problem is indeed multi-modal and may have a number of local optima and saddle points. Therefore, while a single communication cycle requires less energy and bandwidth than transmitting the data to a central point, solving a global optimization problem may require a large number of cycles, rendering the distributed implementation impractical.

In this paper the problem is formulated as a convex feasibility problem instead of nonlinear least squares. Necessary and sufficient conditions are given under which, when the number of samples increases to infinity or in the absence of measurement noise, the convex feasibility problem has a unique solution at the true source location.

To solve the convex feasibility problem, we propose the projection onto convex sets (POCS) method [5] (see also [6] Ch. 5). It is shown that this method can be implemented in a distributed manner, i.e., each sensor performs the bulk of its computations based on its own data and it is not required that the full data set be sent to a central point for processing. As in Nowak's distributed EM algorithm [7], a number of communication cycles across the network is sufficient for the implementation of the estimator. A closed form expression is given for the usually computationally demanding projection phase of POCS, which leads to a computationally efficient implementation. For a finite number of samples, it is shown that convergence to a point or a limit cycle in the vicinity of the true source position occurs. Simulation results show global convergence of the proposed method in contrast to a local search method, with extremely fast convergence rates.

## II. Problem Formulation

The energy attenuation model of [1] is adopted. Consider a sensor network composed of $L$ sensors distributed at known spatial locations, denoted $r_l$, $l = 1, \dots, L$, where $r_l \in \mathbb{R}^2$. Generalization to $\mathbb{R}^3$ is straightforward but is not explored here. A stationary acoustic source is located at an unknown location $\theta^* \in \mathbb{R}^2$. Each sensor collects $n$ noisy measurements of the acoustic signal transmitted by the source. Neglecting the propagation time from the source to the sensors, the received signal is modeled by

$$x_l(t) = \frac{a(t)}{||r_l - \theta^*||} + w_l(t), \quad t = 1, \dots, n, \quad l = 1, \dots, L$$

where $a(t)$ is the intensity of the source signal measured 1 m from the source, and $w_l(t)$ is a zero-mean white Gaussian noise with known variance $\sigma^2$, which is independent of $a(t)$. The estimation of the source location is based on the source energy estimates at each of the sensors

$$z_l = \frac{1}{n}\sum_{t=1}^{n} x_l^2(t) = \frac{A}{||r_l - \theta^*||^2} + \frac{2}{n}\sum_{t=1}^{n}\frac{a(t)w_l(t)}{||r_l - \theta^*||} + v_l$$

where

$$A = 1/n \sum_{t=1}^{n} a_l^2(t) \tag{1}$$

and $v_l = 1/n \sum_{t=1}^{n} w_l^2(t)$. Neglecting the cross term due to the independence assumption, invoking the central limit theorem to model $v_l$, and subtracting the assumed known noise variance $\sigma^2$, we arrive at the energy attenuation model of [1], which was validated through an experiment in [8],

$$y_l = \frac{A}{||r_l - \theta^*||^2} + v_l, \quad l = 1, \dots, L \tag{2}$$

where $v_l$ is a zero-mean white Gaussian noise with variance $2\sigma^4/n$.

We first assume that $A$ is known. This assumption is valid when an additional sensor is added to an already deployed network and the new sensor transmits an acoustic signal with known power to enable the network to estimate its location. The case of unknown source power is treated in Sec. IV.

The material in this paper will be presented in part at the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing.

This research was partially supported by DARPA-MURI grant ARO DAAD 19-02-1-0262.

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122 USA, Tel: 734-647-2045, Fax: 734-763-8041, (email: dblatt@umich.edu; hero@umich.edu).

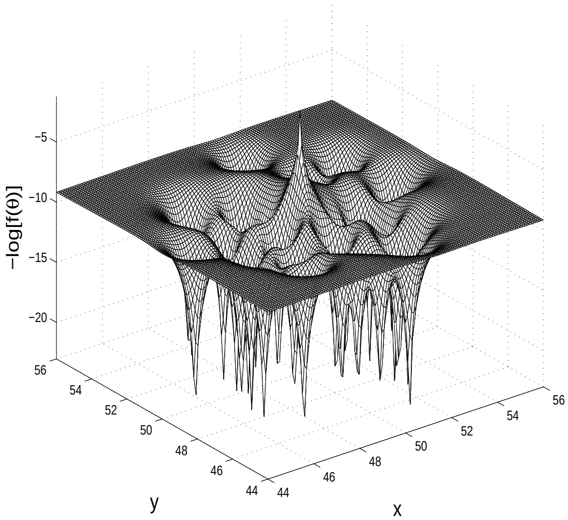


Fig. 1. The negative log of the nonlinear least squares objective function.

The maximum likelihood estimator (MLE) [1] is found by solving the nonlinear least squares problem

$$\widehat{\theta}_{ML} = \arg\min_{\theta \in \mathbb{R}^2} \sum_{l=1}^{L} \left[ y_l - \frac{A}{||r_l - \theta||^2} \right]^2 . \tag{3}$$

The fact that the objective function is a sum of $L$ components was exploited in the implementation of the distributed incremental gradient method in [2], [3]. However, since the objective function has multiple local optima and saddle points, the incremental gradient method may stagnate at one of these sub-optimal solutions instead of converging to the optimal one. A realization of the negative log of the objective function in (3) is presented in Fig. 1. The details of the simulation that generated this figure are given in Sec. V. It can be seen that the objective function has many local optima and saddle points and that the global optimum is peaked.

An alternative formulation of the problem of estimating the source's location is the following. Consider the $l$ summands in the objective function (3). It is easily seen that the function

$$f_l(\theta) = \left[ y_l - \frac{A}{||r_l - \theta||^2} \right]^2 \tag{4}$$

obtains its minimum on the circle

$$C_l = \left\{ \theta \in \mathbb{R}^2 : ||\theta - r_l|| = \sqrt{A/y_l} \right\} . \tag{5}$$

Let $D_l$ be the disk defined by

$$D_l = \left\{ \theta \in \mathbb{R}^2 : ||\theta - r_l|| \leq \sqrt{A/y_l} \right\} . \tag{6}$$

We propose to solve the source localization problem by letting the estimator be a point in the intersection of the sets $D_l$, $l = 1, \ldots, L$, that is,

$$\widehat{\theta} \in D = \bigcap_{l=1}^{L} D_l \subset \mathbb{R}^2 . \tag{7}$$

Note that due to observation noise the intersection $D$ might be empty. In this case, our estimator is any point that minimizes the sum of distances to the sets $D_l$, $l = 1, \ldots, L$, that is,

$$\widehat{\theta} = \arg\min_{\theta \in \mathbb{R}^2} \sum_{l=1}^{L} ||\theta - \mathcal{P}_{D_l}(\theta)||^2 \tag{8}$$

where for a set $S \subseteq \mathbb{R}^2$ and a point $x \in \mathbb{R}^2$, $\mathcal{P}_S(x)$ is the orthogonal projection of $x$ onto $S$, that is,

$$\mathcal{P}_S(x) = \arg\min_{y \in S} ||x - y|| \tag{9}$$

where $|| \cdot ||$ is the Euclidean norm. Observe that (8) includes (7) as a special case when a minimum value of zero is attainable, and note that in general $\widehat{\theta} \neq \widehat{\theta}_{ML}$. Since the sets $D_l$ are convex, both the consistent and inconsistent convex feasibility problems, (7) and (8), respectively, can be solved via the POCS method to be described below.

Before describing POCS, we give necessary and sufficient conditions for the consistency of the estimator (8). Denote by $\mathcal{H}$ the convex hull of the sensors' spatial locations, i.e.,

$$\mathcal{H} = \left\{ x \in \mathbb{R}^2 : x = \sum_{l=1}^{L} \alpha_l r_l ,\ \alpha_l \geq 0,\ \sum_{l=1}^{L} \alpha_l = 1 \right\} .$$

It is possible to show geometrically (see Fig. 2) that when the number of samples $n$ increases to infinity, or in the absence of measurement noise, the convex feasibility problem (8) has a unique solution at the true source's location, denoted by $\theta^*$, if and only if $\theta^*$ lies in $\mathcal{H}$, that is,

$$\bigcap_{l=1}^{L} \left\{ \theta \in \mathbb{R}^2 : ||r_l - \theta|| \leq ||r_l - \theta^*|| \right\} = \{\theta^*\}$$
$$\text{if and only if} \quad \theta^* \in \mathcal{H}$$

where $L \geq 2$. As seen in Fig. 2 (bottom), when the source lies outside $\mathcal{H}$, even in the asymptotic case there is no unique solution to (8). Rather, there is a continuous set of points (the shaded area) that minimize the objective function. In this situation, our formalization is not appropriate for estimating the source location.

In the general case of finite number of samples and finite signal to noise ratio, one of two cases can occur: (a) $D \neq \emptyset$, and (b) $D = \emptyset$. In the former, the POCS method is guaranteed to converge to a point in $D$. In the latter, the POCS method converges to a limit cycle in the vicinity of the point that minimizes the sum of distances to the sets $D_l$ (6), or, when a certain sequence of relaxation parameters are used, the method converges to the optimal solution.

## III. Distributed Implementation of POCS

The POCS method [5], [6] is given by the following algorithm.

1) Initialization: $\theta^0$ is arbitrary.
2) Iterative step: For all $k \geq 0$,

$$\theta^{k+1} = \theta^k + \lambda^k \left[ \mathcal{P}_{D_{\kappa(k)}}(\theta^k) - \theta^k \right] \tag{10}$$

where $\{\lambda^k\}_{k \geq 1}$ is a sequence of relaxation parameters satisfying for all $k$, $\epsilon_1 \leq \lambda^k \leq 2 - \epsilon_2$ for some $\epsilon_1, \epsilon_2 > 0$, $\kappa(k) = k \bmod L$, and $\mathcal{P}_S(x)$ is defined in (9).

Usually the projection operator is the most computationally demanding element of POCS. In our application, however, a closed form expression is available for (10). Clearly, if $||\theta - r_l|| \leq \sqrt{A/y_l}$ then $\theta \in D_l$ and $\mathcal{P}_{D_l}(\theta) = \theta$, otherwise,

$$\mathcal{P}_{D_l}(\theta^k) = r_l + [\alpha \cos(\phi),\ \alpha \sin(\phi)]^T \tag{11}$$

where $\alpha = \sqrt{A/y_l}$, and $\phi = \mathrm{atan}(\theta^k(2) - r_l(2), \theta^k(1) - r_l(1))$, where $\mathrm{atan}(\cdot, \cdot)$ is the four quadrant inverse tangent function, and for a vector $x \in \mathbb{R}^2$, $x(1)$ and $x(2)$ denote its first and second coordinates, respectively.

When $||\theta^k - r_{\kappa(k)}|| > \sqrt{A/y_{\kappa(k)}}$, the vector $\left[\theta^k - \mathcal{P}_{D_{\kappa(k)}}(\theta^k)\right]$ points in the same direction as $\nabla f_{\kappa(k)}(\theta^k)$, the gradient of $f_{\kappa(k)}$ at the point $\theta^k$. Hence, the POCS method is closely related to the

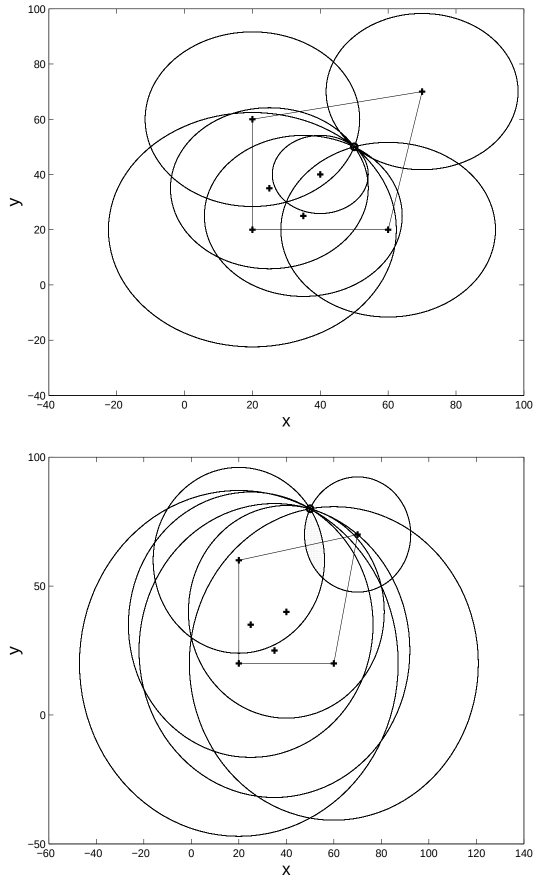


Fig. 2. Source, denoted by a black dot, is located inside (top) and outside of (bottom) the convex hull $\mathcal{H}$ of the sensors' locations, denoted by crosses.

incremental gradient method [9, p. 109], which was implemented in [2] and [3] to solve (3). The incremental gradient method generates a sequence $\{\theta^k\}_{k\geq 0}$ according to

$$\theta^{k+1} = \theta^k - \mu^k \nabla f_{\kappa(k)}(\theta^k) \tag{12}$$

where $\mu^k$ is a positive step size, possibly depending on $k$. The difference between the two methods is that when $||\theta^k - r_{\kappa(k)}|| \leq \sqrt{A/y_{\kappa(k)}}$, the POCS iteration is $\theta^{k+1} = \theta^k$, whereas the incremental gradient iteration is a step in the direction that points from $r_{\kappa(k)}$ to $\theta^k$, i.e., a step towards the circle $C_{\kappa(k)}$. Therefore, the incremental gradient method (12) is more closely related to Kaczmarz's algorithm [6] (also known as the algebraic reconstruction technique) than to POCS. In particular, if the step size $\mu^k$ is determined by a line search in the direction $\nabla f_{\kappa(k)}(\theta^k)$, then the iterates are identical to those obtained when applying Kaczmarz's algorithm to the problem for finding the intersection of the circles $\{C_l\}_{l=1}^L$. In contrast to the global convergence property of POCS, Kaczmarz's algorithm is known to converge only locally when applied to nonlinear problems such as finding the intersection of the circles $\{C_l\}_{l=1}^L$.

The relaxation parameters $\lambda^k$ (10) play an important role in the convergence of the method. At the first phase of the implementation of the POCS method, the relaxation parameters are set to 1. As the method progresses, a convergence criterion is repeatedly checked. If convergence to a single point is detected, e.g., by verifying that $\sum_{l=1}^L ||\theta^{k-l} - \theta^{k-l-1}||$ is smaller than a threshold, it is concluded that $D \neq \emptyset$, and the final estimate $\widehat{\theta}$ is set to the limit point. If convergence to a limit cycle is detected, i.e., each sensor converges to a different value, it is concluded that $D = \emptyset$ and the method enters phase two. At phase two the relaxation parameters are decreased at a rate of $1/k$. In [10], it is shown that this relaxation sequence leads to convergence to the point $x$ that minimizes the sum of squared distances to the sets $D_l$, that is, to $\widehat{\theta}$ defined in (8). It should be noted that if the transition to phase two occurs prematurely, this convergence result still holds. The effect will be a slowdown of convergence. A sub-optimal but computationally cheaper alternative to phase two is to approximate $\widehat{\theta}$ by the arithmetic mean of the points in the limit cycle. This simple approach was used in the simulation reported in Sec. V. Due to its global convergence properties, the estimate resulting from the POCS method could also be used to trigger a local search for the nonlinear least squares estimator such as the one in [2]. However, we cannot guarantee that this initial point falls within the attraction region of the global maximum of the likelihood function.

Note that all the information required for the computation of (11) (or (10)) is available at sensor $l$ and hence a distributed implementation is possible. Following [7], assume without loss of generality that the indices $l = 1, \ldots, L$ correspond to a cycle through the network. Let sensor 1 be initiated with a pre-specified initial value $\theta^0$. Sensor 1 generates $\theta^1$ through (10) and transmits $\theta^1$ to sensor 2. Upon receiving $\theta^k$ from sensor $\kappa(k)$, sensor $\kappa(k+1)$ calculates $\theta^{k+1}$ and transmits it to sensor $\kappa(k+2)$. The information cycle continues until the detection of convergence to either a limit point or a limit cycle. The convergence detection criteria can be easily implemented in a distributed manner as well. Phase two can be implemented in a similar way.

### A. Communication Bandwidth and Energy Consumption

Consider first the centralized approach, in which each sensor transmits its coordinates $r_l \in \mathbb{R}^2$ (unless these are known a priori) and its energy estimate $y_l$ (2) to a fusion center. Assume that the network is distributed over the cube $[0,1]^2$. For a fixed quantization level of the spatial coordinates and energy estimates, conveying the information from the sensors to a fusion center requires the transmission of $O(L)$ bits over a distance of $O(1)$ per bit [4].

In our decentralized implementation of the POCS method and in the decentralized implementation of the incremental gradient method [2], [3], information is conveyed in communication cycles as described above. In every communication hop, a sensor transmits the current source location estimate to the next sensor in the cycle. The spatial coordinates of the sensors do not need to be shared. For a fixed quantization level of the source location estimate, performing a single communication cycle across the network requires the transmission of $O(L)$ bits over a distance of $O(\sqrt{\log^2 L/L})$ per bit [4]. Finally, the total number of bits is obtained by multiplying the number of bits per cycle by the expected number of cycles.

From the above analysis it can be seen that the communication burden grows linearly with the number of sensors in the centralized approach and sub-linearly ($\sqrt{L}\log L$) in the decentralized approach. To compare the two approaches for a fixed number of sensors $L$, however, one should factor in the number of cycles. As shown in Sec. V the implementation of the incremental gradient method [2], [3] may require hundreds of cycles to find the global maximum of the non-linear objective function (3) and hence has advantages with respect to the centralized approach only in dense networks. In contrast, it is shown that for the same scenario the distributed POCS implementation requires as few as 4 cycles to achieve convergence, and hence, leads to a reduction in energy and bandwidth requirements as compared to the centralized approach for sparser networks.

## IV. The Case of Unknown Source Power

When the source is not collaborating with the network, the signal power $A$ (1) is unknown. To eliminate the dependency of the optimization problem on $A$, an energy ratios based source localization method was proposed in [8] (see [1] as well). In this section it is shown that it is also possible to represent the estimation of the source location based on the energy ratios as a convex feasibility problem and hence solve it in a distributed manner as described in Sec. III.

Considering the noise free problem, Li and Hu [8] showed that the ratio between the energy readings at two sensors, $y_l$ and $y_k$, defines a circle or a hyperplane on which the source may lie:

$$\varphi_{lk} = \sqrt{y_l/y_k} = \frac{||\theta - r_l||}{||\theta - r_k||}. \tag{13}$$

When $\varphi_{lk} \neq 1$, the resulting circle is given by

$$\{\theta : ||\theta - c_{lk}||^2 = \zeta_{lk}^2\}$$

where $c_{lk} = (r_l - \varphi_{lk}^2 r_k)/(1-\varphi_{lk}^2)$, and $\zeta_{lk} = \varphi_{lk}||r_l - r_k||^2/(1-\varphi_{lk}^2)$. When $\varphi_{lk} = 1$, (13) defines the hyperplane

$$\{\theta : \theta^T v_{lk} = \tau_{lk}\}$$

where $v_{lk} = r_l - r_k$, and $\tau_{lk} = (||r_l||^2 - ||r_k||^2)/2$. In the presence of observation noise, given a set of $L_1 + L_2$ ratios, the location of the source is estimated by minimizing the cost function

$$J(\theta) = \sum_{l_1=1}^{L_1} (||\theta - c_{l_1}|| - \zeta_{l_1})^2 + \sum_{l_2=1}^{L_2} (\theta^T v_{l_2} - \tau_{l_2})^2 \tag{14}$$

where $L_1$ and $L_2$ are the number of circles and hyperplanes, respectively. In [1], this estimator is called the energy-ratio nonlinear least squares.

To formulate the problem of estimating the source location from the energy ratios (13) as a convex feasibility problem, assume without loss of generality that $\varphi_{lk} \leq 1$ (otherwise replace it with $\varphi_{kl}$). Define the discs

$$\widetilde{D}_{l_1} = \{\theta : ||\theta - c_{l_1}||^2 \leq \zeta_{l_1}^2\}$$

and the hyperplanes

$$\widetilde{h}_{l_2} = \{\theta : \theta^T v_{l_2} = \tau_{l_2}\}.$$

Hence, the POCS method can be implemented in a distributed manner to estimate the source location by finding a point in the intersection of the convex sets

$$\left(\bigcap_{l_1=1}^{L_1} \widetilde{D}_{l_1}\right) \bigcap \left(\bigcap_{l_2=1}^{L_2} \widetilde{h}_{l_2}\right).$$

Note that the projection onto a hyperplane has a closed form expression as well. To optimize the energy consumption, the energy ratios should be selected based on geographical vicinity.

Li and Hu also proposed to replace every two circles in (14) with a single hyperplane and then solve the resulting linear least squares problem. This approach can also be converted to a convex feasibility problem.

## V. Simulation Results

This section presents a simulation of a sensor network of $L = 5000$ nodes, distributed randomly in a $100\text{m} \times 100\text{m}$ field. At each sensor a measurement of the acoustic source energy was generated according to (2). The source is located at $\theta^* = [50, 50]^T$ and emits a signal with $A$ (1) set to $100$. The energy measurement noise variance is $2\sigma^4/n = 1$. Following [11], [3], not all sensors participate in the estimation task. At an acquisition phase, each sensor decides whether or not a source is present using a simple threshold test. Only those sensors whose energy estimates $y_l$ (2) are above $5$ participate. The threshold $5$ corresponds to an average SNR greater than 7dB at the active sensors and it was chosen to balance the number of active sensors and their associated SNR levels. The performance of the algorithm was insensitive to small changes in the threshold. When this threshold is used the average number of active sensors is $32$ with standard deviation of $5$. In the realization presented here, $\overline{L} = 31$ sensors detected the source and entered the estimation phase.

A realization of the objective function associated with the MLE (3) is shown in Fig. 1. To optimize the viewing angle of this figure, the negative log of the objective function is presented. Hence, the optimum point is the global maximum rather than the minimum, which appears close to the true location of the source. The objective function has multiple local optima and saddle points, which impose difficulties on any local search method. In Fig. 3, the paths taken by the steepest descent (SD) method initiated from multiple points on a grid are presented on top of the contour plot of the nonlinear objective function (3). The SD method could also be implemented in a distributed manner, e.g., distributed Fisher scoring [12]. The initial points are depicted by crosses, followed by a line which follows the path taken by the algorithm, and ends at the convergence points depicted by circles. It is seen that only when the method is initiated close to the global optimum at the center of the plot, does convergence to the global optimum occur. The method mostly stagnates at local optima or saddle points.

The paths taken by the incremental gradient method of [2], [3] are presented in Fig. 4. Since the gradient $\nabla f_l(\theta)$ diverges at the sensor location $r_l$, a small step size $\mu$ (12) is required to achieve convergence, e.g., in the implementation presented in Fig. 4 $\mu = 10^{-4}$. As in Fig. 3, the initial points are depicted by crosses, followed by a line which follows the path taken by the algorithm, and ends at the convergence points depicted by circles. The crosses that are not followed by a line correspond to initial points that lead to divergence. Each path corresponds to hundreds of communication cycles and it is seen that only two of the initial points result in convergence to the global optimum at the center of the figure.

To combat the high variability of the gradient $\nabla f_l(\theta)$ over the optimization domain, Rabbat and Nowak suggested to normalize the descent direction in (12) [13]. The modified algorithm is given by

$$\theta^{k+1} = \theta^k - \mu \frac{\nabla f_{\kappa(k)}(\theta^k)}{||\nabla f_{\kappa(k)}(\theta^k)||} \tag{15}$$

where $\mu$ is a constant step size. In Fig. 5 the paths taken by this algorithm for $\mu = 0.1$ are given. It is seen that this variation of the incremental gradient method is not sensitive to local maxima. However, even in the absence of noise, a convergence proof is not available for this variation. In the simulation, convergence to a single point or a limit cycle was not detected after the first $200$ communication cycles. An example in which the method (15) does not converge to the global maximum is given in Fig. 6. In this simulation there are four sensors operating in a noise free environment and a source is located at $[50, 50]$. The paths initiated from the top four crosses do not converge to the global optimum at $[50, 50]$.

In contrast to the shortcoming of the local search methods, the proposed POCS method converges to the vicinity of the global optimum regardless of the initial point. In Fig. 7 the paths taken by the POCS method are presented. The order of the sensors in the information cycle described in Sec. III was selected randomly. Convergence to a limit or a limit cycle is declared if at sensor $1$

$$||\theta^{kL+1} - \theta^{(k-1)L+1}|| < 10^{-3}. \tag{16}$$

Once convergence is detected, the final estimate is the average of

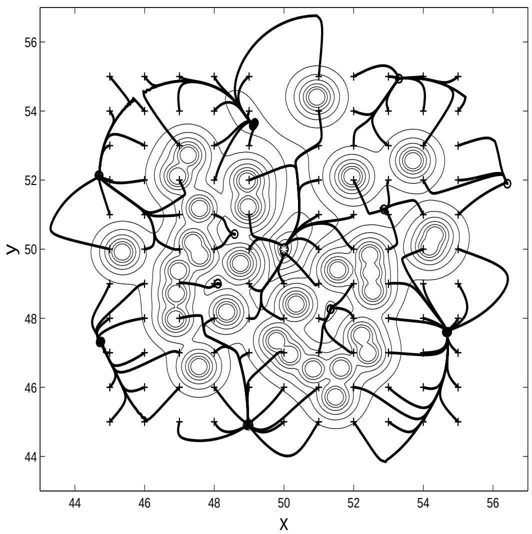


Fig. 3. Paths taken by the steepest descent method.

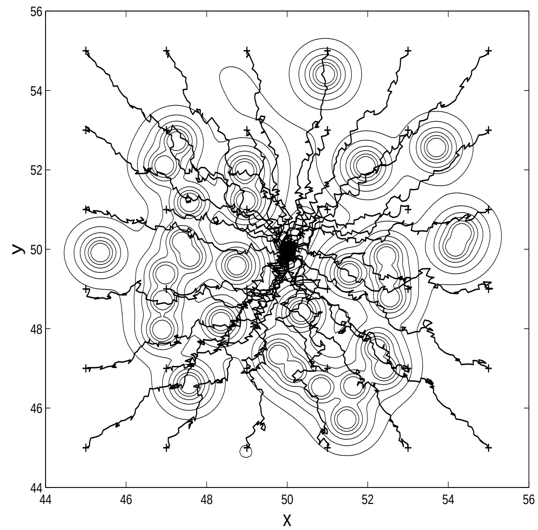


Fig. 5. Paths taken by the normalized IG method.

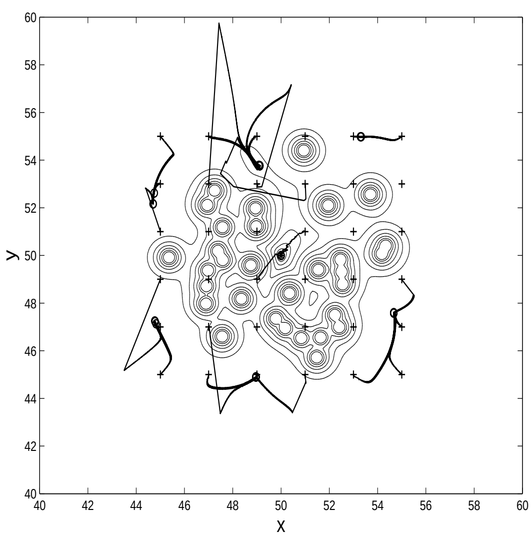


Fig. 4. Paths taken by the IG method.

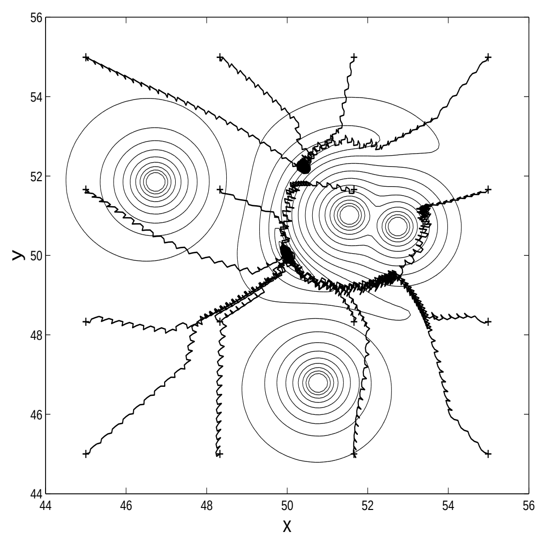


Fig. 6. The normalized IG method does not always converge globally.

the sensors' estimates $1/L \sum_{l=1}^{L} \theta^{(k-1)L+l}$, which can be easily computed in a distributed manner [2] through a single communication cycle. A better illustration of the method is presented in Fig. 8, in which four representative paths are superimposed on top of the convex sets (discs) (6). At each iteration the sequence generated by the algorithm is projected onto a different disc, unless it is already inside it. It is seen that the convergence is extremely fast; after as few as three sub-iterations (10), the sequence reaches the vicinity of the global optimum. In part of the sub-iterations (10), little or no progress is made if the previous iterate is close to or inside the corresponding disc, respectively, but three communication cycles were sufficient to satisfy the convergence criterion (16) regardless of the initial point. Adding the communication cycle required for the computing the average we conclude that it requires $4$ communication cycle to implement the distributed POCS method in this scenario; a significant reduction in energy and bandwidth requirements compared to the incremental gradient method implementation [2], [3] or its modification (15).

The performance of POCS in terms of estimation errors was also evaluated. As a benchmark, the performance of POCS was compared to the performance of the MLE [1]. The MLE was found by performing a grid search over the field area followed by a local search initiated at the highest maximum. Note that due to the peakedness of the global maximum (see Fig. 1), a fine grid search is required. Therefore, a grid search based optimization is operational only in the centralized approach when all sensor readings are available at a central location. In our implementation the grid search resolution was set to $0.1\mathrm{m}\times 0.1\mathrm{m}$. Also presented is the performance of an estimator which is obtained by performing a local search on the ML objective function (3) initiated at the POCS estimator. The performance of the three estimators was evaluated through $20000$ Monte Carlo iterations, whereby the number of sensors in the field was increased from $100$ to $2100$ in $200$ increments. In Fig. 9 the square root of the mean squared error and the median squared error of the three methods are presented with $\pm\sigma$ confidence interval as a function of the average number of sensors that entered the estimation phase. The standard deviations of the mean and median estimators were estimated from $1000$ bootstrap data samples.

As shown in Fig. 9, the POCS method is more sensitive to changes in the number of active sensors then the MLE in terms of estimation

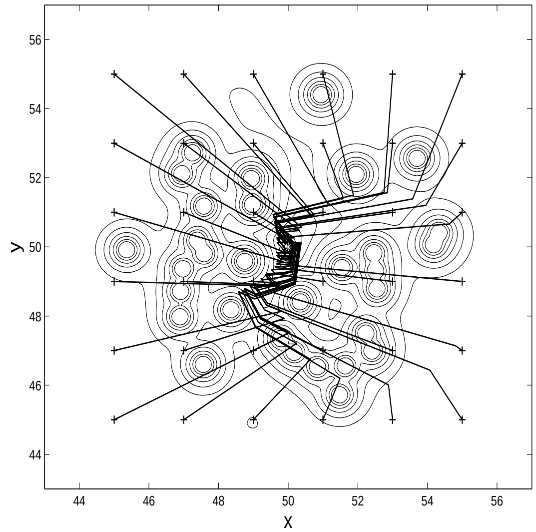


Fig. 7. Paths taken by the POCS method.

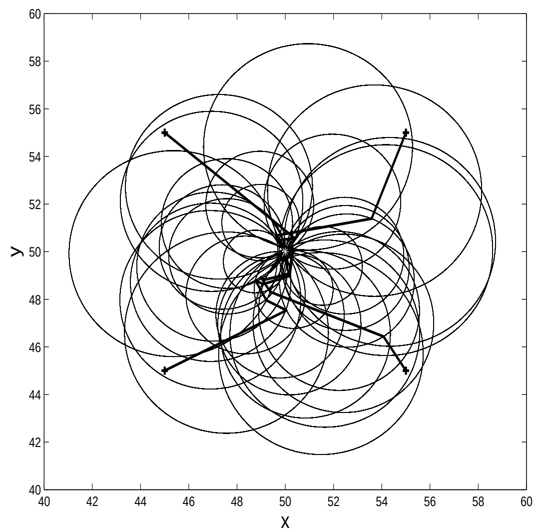


Fig. 8. Paths taken by the POCS method superimposed on the convex sets.

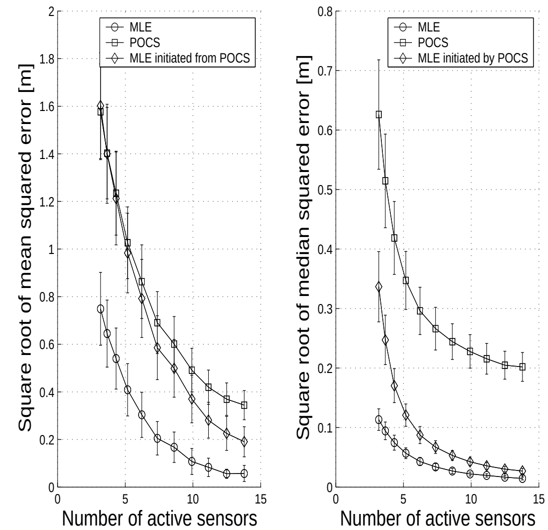


Fig. 9. Local performance: POCS vs. MLE, mean (left) and median (right).

errors. We have observed the same kind of behavior when the number of sensors was fixed and the SNR was varied. This equivalence is expected since the number of active sensors is linked to the SNR level through the detection threshold. In contrast to the sensitivity of POCS to the number of active sensors or to the SNR level, when the intersection of the disks is in the vicinity of the source location, we always observed extremely fast convergence rates of POCS similar to those shown in Fig. 7.

## VI. Conclusions

The problem of distributed acoustic source localization using a wireless sensor network was formulated as a convex feasibility problem and solved via the POCS method. The solution has global convergence properties with fast convergence rates. Finally we note that this concept can be applied to other problems in which the objective function depends on the parameters through terms of the form $||\theta - c_l||$, where $c_l,\ l = 1, \ldots, L$ are data dependent terms. In particular, this concept can be easily generalized to the three dimensional case. The effect of quantization and channel noise are worthy of additional study.

## VII. Acknowledgement

The authors would like to thank Prof. R. Nowak and M. Rabbat for sharing with us their code for implementing the incremental gradient method and its modification (15).

## References

[1] X. Sheng and Y. H. Hu, “Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks,” *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 44–53, Jan. 2005.

[2] M. G. Rabbat and R. D. Nowak, “Decentralized source localization and tracking,” in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, May 2004, pp. 921 – 924.

[3] ——, “Distributed optimization in sensor networks,” in *Proceedings of the Third International Symposium on Information Processing in Sensor Networks*. Berkeley, California: ACM Press, New York, April 2004, pp. 20–27.

[4] ——, “Quantized incremental algorithms for distributed optimization,” *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 798 – 808, Apr. 2005.

[5] L. G. Gubin, B. T. Polyak, and E. V. Raik, “The method of projections for finding the commmon point of convex sets,” *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 1–24, 1967.

[6] Y. Censor and S. A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*. New York: Oxford University Press, 1997.

[7] R. D. Nowak, “Distributed EM algorithms for density estimation and clustering in sensor networks,” *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2245 – 2253, Aug. 2003.

[8] D. Li and Y. H. Hu, “Energy-based collaborative source localization using acoustic microsensor array,” *EURASIP Journal on Applied Signal Processing*, no. 4, pp. 321–337, 2003.

[9] D. P. Bertsekas, *Nonlinear programming: second edition*. Belmont, MA: Athena Scientific, 1999.

[10] Y. Censor, A. R. D. Pierro, and M. Zaknoon, “Steered sequential projections for the inconsistent convex feasibility problem,” *Nonlinear analysis: theory, methods, and application, series A*, vol. 59, pp. 385–405, 2004.

[11] X. Sheng and Y. H. Hu, “Energy based acoustic source localization,” in *Information Processing in Sensor Networks, Second International Workshop, IPSN 2003*, ser. Lecture Notes in Computer Science, Z. Feng and G. Leonidas, Eds., vol. 2634. Palo Alto, California: Sringer-Verlag, New York, April 2003, pp. 285–300.

[12] D. Blatt and A. Hero, “Distributed maximum likelihood for sensor networks,” in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, May 2004, pp. 929 – 932.

[13] M. G. Rabbat and R. D. Nowak, *Personal Communication*, 2005.