

# APOCS: A RAPIDLY CONVERGENT SOURCE LOCALIZATION ALGORITHM FOR SENSOR NETWORKS

*Doron Blatt and Alfred O. Hero, III*

Department of Electrical Engineering and Computer Science,  
University of Michigan, Ann Arbor, MI  
{dblatt, hero}@eecs.umich.edu

## 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 hence local search methods might stagnate at a sub-optimal solution. In this paper, we treat the problem in its convex feasibility formulation. We propose the aggregated projection onto convex sets (APOCS) method, which, in contrast to the original POCS method, converges to a meaningful limit even when the problem is infeasible without requiring a diminishing step size. Simulation results show convergence to the global optimum with significantly faster convergence rates compared to the previous methods.

## 1. 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 by non-linear least squares via maximum likelihood, which is asymptotically optimal and offers a natural generalization to the case of multiple sources [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). The 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. It is shown that 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.

More recently, a kernel averaging method was proposed in [5], which avoids the global optimization problem associated with the nonlinear least squares problem. 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 (see [6] as well).

To solve the convex feasibility problem, the projection onto convex sets (POCS) method [7] (see also [8] Ch. 5) is applied. A closed form expression is given for the usually computationally demanding projection phase of POCS, which leads to a computationally efficient implementation.

A drawback of the standard form of POCS [7] is that for a finite number of samples the convex feasibility problem may become inconsistent, i.e., the intersection of the sets is empty, and the iterates reach a limit cycle instead of converging to a limit point. To overcome this weakness, the steered sequential projection method, which uses a decreas-

---

This work was supported in part by NIH/NCI grant 1P01 CA87634, by DARPA-MURI grant ARO DAAD 19-02-1-0262, and by NSF contract CCR-0325571.

ing sequence of relaxation parameters and converges to a meaningful limit, was proposed in [9]. A byproduct of the decreasing step size is a slowdown of convergence.

In this paper a variation of POCS, which does not require a diminishing step size for convergence, is proposed. The method is based on an aggregated projection phase, in which progress is made based on an average of the directions to the individual sets. It is shown that our 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 [10], a number of communication cycles across the network is sufficient for the implementation of the estimator.

Simulation results show global convergence of POCS in contrast to a local search method, with extremely fast convergence rates. For the inconsistent convex feasibility problem, a comparison between the steered sequential projection method [9] and our aggregated POCS (APOCS) is given.

## 2. 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 and computes an estimate of the source's energy. The energy estimates are modeled by

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

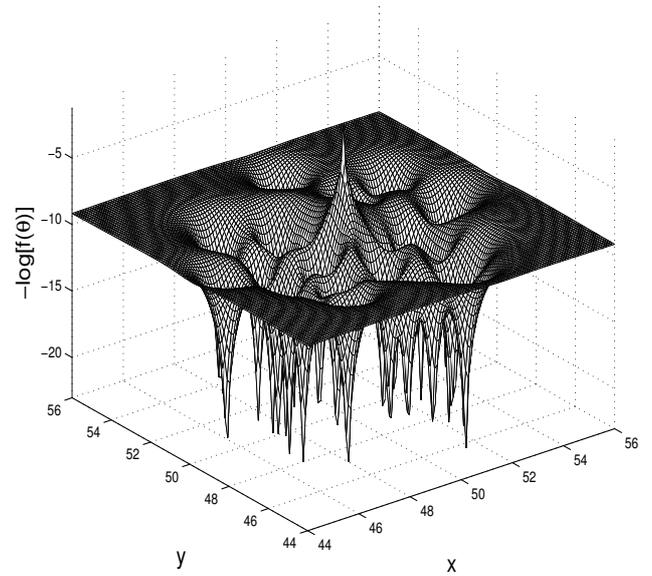
where  $y_l$  is the energy estimate at the  $l$ 'th sensor,  $A$  is the source's energy, and  $v_l$  is a zero-mean white Gaussian noise with known variance. This model has been validated through experiment for acoustic localization in [11].

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 sensor strength is treated in section 4.

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

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

The fact that the objective function is a sum of  $L$  components was exploited in the implementation of the distributed



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

incremental gradient method in [2] and [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 (2) is presented in Fig. 1. The details of the simulation that generated this figure are given in Sec. 5. 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 (2). It is easily seen that the function

$$f_l(\theta) = \left[ y_l - \frac{A}{\|r_l - \theta\|^2} \right]^2 \quad (3)$$

obtains its minimum on the circle

$$C_l = \left\{ \theta \in \mathbb{R}^2 : \|\theta - r_l\| = \sqrt{A/y_l} \right\}. \quad (4)$$

Let  $D_l$  be the disk defined by

$$D_l = \left\{ \theta \in \mathbb{R}^2 : \|\theta - r_l\| \leq \sqrt{A/y_l} \right\}. \quad (5)$$

Then our estimation problem is solved by finding a point in the intersection of the sets  $D_l$ ,  $l = 1, \dots, L$ . That is,

$$\hat{\theta} \in D = \bigcap_{l=1}^L D_l \subset \mathbb{R}^2. \quad (6)$$

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  $D_l, l = 1, \dots, L$ ,

$$\hat{\theta} = \arg \min_{\theta \in \mathbb{R}^2} \sum_{l=1}^L \|\theta - \mathcal{P}_{D_l}(\theta)\|^2 \quad (7)$$

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,

$$\mathcal{P}_S(x) = \arg \min_{y \in S} \|x - y\| \quad (8)$$

where  $\|\cdot\|$  is the Euclidean norm. Observe that (7) includes (6) as a special case when a minimum value of zero is attainable.

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 can be shown geometrically (see [6]) that when the number of samples  $n$  increases to infinity, or in the absence of measurement noise, the convex feasibility problem (7) has a unique solution at the true source's location, denoted by  $\theta^*$ , if and only if  $\theta^*$  lies in  $\mathcal{H}$ .

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  $D_l$  (5), or, when a decreasing sequence of relaxation parameters are used [9], the method converges to the optimal solution. As an alternative to the diminishing step size approach of [9], the APOCS method is proposed in the next section.

### 3. AGGREGATED POCS AND ITS DISTRIBUTED IMPLEMENTATION

Before describing APOCS, the POCS method [7, 8] is reviewed.

#### POCS:

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

$$\theta^{k+1} = \theta^k + \lambda^k [\mathcal{P}_{D_{\kappa(k)}}(\theta^k) - \theta^k] \quad (9)$$

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 (8).

An important property of POCS when applied to the source localization problem is that the projection operator has a closed form expression. Clearly, if  $\|\theta - r_l\| \leq \sqrt{A/y_l}$  then  $\theta \in D_l$  and  $\mathcal{P}_{D_l}(\theta) = \theta$ , otherwise,

$$\mathcal{P}_{D_l}(\theta^k) = r_l + [\alpha \cos(\phi), \alpha \sin(\phi)]^T \quad (10)$$

where  $\alpha = \sqrt{A/y_l}$ , and  $\phi = \text{atan}(\theta^k(2) - r_l(2), \theta^k(1) - r_l(1))$ , where  $\text{atan}(\cdot, \cdot)$  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.

As mentioned earlier, when the convex feasibility problem is inconsistent, POCS converges to a limit cycle, unless  $\lambda^k$  is decreased to zero [9]. The byproduct of this variation, which is called the steered sequential projection method in [9], is slowdown of convergence. As an alternative we propose the APOCS method.

#### APOCS:

1. Initialization:  $\theta^0$  is arbitrary,  $d^0 = [0, 0]^T$ .
2. Phase one: For  $1 \leq k \leq L$ ,

$$\begin{aligned} d^k &= d^{k-1} + [\mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1}) - \theta^{k-1}] \\ \theta^k &= \mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1}) \end{aligned} \quad (11)$$

where  $\mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1})$  is defined in (8).

3. Phase two: For  $k > L$ ,

$$\begin{aligned} d^k &= d^{k-1} + [\mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1}) - \theta^{k-1}] \\ &\quad - [\mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1-L}) - \theta^{k-1-L}] \\ \theta^k &= \theta^{k-1} + \frac{\mu}{L} d^k \end{aligned} \quad (12)$$

where  $\kappa(k) = k \bmod L$ , and  $\mu$  is a constant step size chosen small enough to ensure convergence.

The APOCS iteration can also be write explicitly as

$$\theta^k = \theta^{k-1} + \frac{\mu}{L} \sum_{j=1}^L [\mathcal{P}_{D_{\kappa(k-j+1)}}(\theta^{k-j}) - \theta^{k-j}]. \quad (13)$$

Hence, at each iteration, progress is made in a direction which is the average of the  $L$  previous projections. When a new projection direction to a set is computed, it replaces the old projection direction to that set, which was computed  $L$  iterations before. When far from the eventual limit the averaged projection directions is an approximation to the direction towards to intersection of the sets. As the method approaches the eventual limit it converges to a fixed point of the algorithm at the point in which the average projection directions vanishes. This point is precisely the minimizer of (7). This is analogous to the behavior of the aggregated incremental gradient method for minimizing the sum of differentiable functions [12].

Note that all the information required for the computation of (11) and (13) is available at sensor  $\kappa(k)$  and hence a distributed implementation is possible. Following [10], assume without loss of generality that the indices  $l = 1, \dots, L$  correspond to a cycle through the network. Phase 1 is performed as follows. Let sensor 1 be initiated with a pre-specified initial value  $\theta^0$ . Sensor 1 generates  $d^1$  and  $\theta^1$  through (11) and transmits them to sensor 2. Upon receiving  $d^{k-1}$  and  $\theta^{k-1}$  from sensor  $k-1$ , sensor  $k$  calculates  $d^k$  and  $\theta^k$  according to (11) and transmits them to sensor  $k+1$ . Phase two is carried out in a similar manner. Upon receiving  $d^{k-1}$  and  $\theta^{k-1}$  from sensor  $\kappa(k-1)$ , sensor  $\kappa(k)$  calculates  $d^k$  and  $\theta^k$  according to (13) and transmits them to sensor  $\kappa(k+1)$ . Note that  $[\mathcal{P}_{D_{\kappa(k)}}(\theta^{k-1-L}) - \theta^{k-1-L}]$  was obtained at sensor  $\kappa(k)$  after the previous cycle.

#### 4. THE CASE OF UNKNOWN SOURCE POWER

When the source is not collaborating with the network, the signal power  $A$  is unknown. To eliminate the dependency of the optimization problem on  $A$ , an energy ratios based source localization method was proposed in [11] (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. 3.

Li and Hu [11] showed that for the noise free problem, 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:

$$\varphi_{lk} = \sqrt{y_l/y_k} = \frac{\|\theta - r_l\|}{\|\theta - r_k\|}. \quad (14)$$

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  $\varphi_{lk} = 1$ , (14) 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

$$J(\theta) = \sum_{l_1=1}^{L_1} (\|\theta - c_{l_1}\| - \zeta_{l_1})^2 + \sum_{l_2=1}^{L_2} (\theta^T v_{l_2} - \tau_{l_2})^2 \quad (15)$$

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 (14) as a convex feasibility

problem, assume without loss of generality that  $\varphi_{lk} \leq 1$  (otherwise replace it with  $\varphi_{kl}$ ). Define the discs

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

and the hyperplanes

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

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

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

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 (15) with a single hyperplane and then solve the resulting linear least squares problem. This approach can also be converted to a convex feasibility problem.

#### 5. 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 (1). The source is located at  $\theta^* = [50, 50]^T$  and emits a signal with energy  $A$  set to 100. Following the approach of [13, 3], in our simulation 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  $y_l$  (1) are above 5 participate. For example, in the realization presented in Figs. 1-4,  $\bar{L} = 31$  sensors detected the source and entered the estimation phase.

In Fig. 2, 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 (2). 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.

In contrast to this shortcoming of the SD local search method, the standard POCS method [7] converges to the vicinity of the global optimum regardless of the initial point. In Fig. 3 the paths taken by the POCS method are presented.

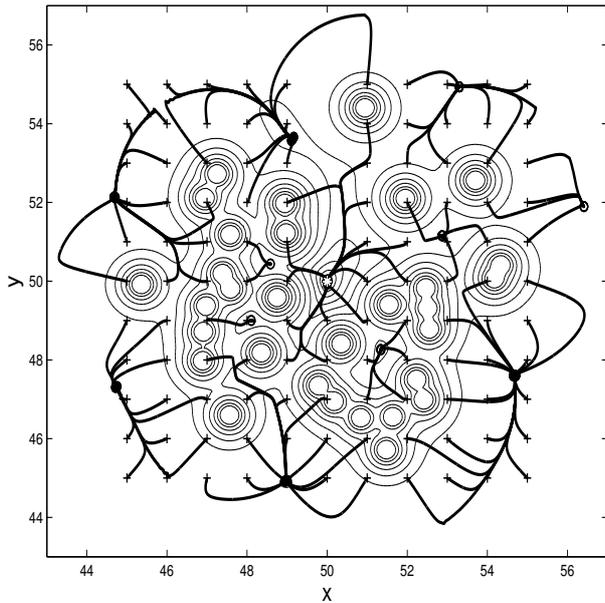


Fig. 2. Paths taken by the steepest descent method.

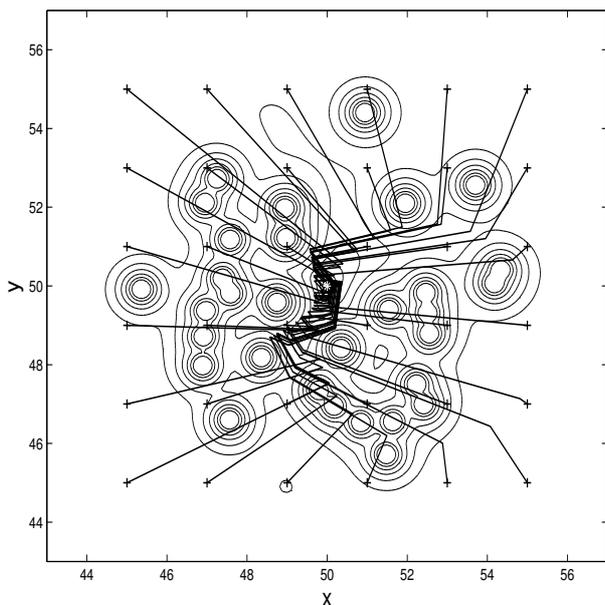


Fig. 3. Paths taken by the POCS method.

It is seen that the convergence is extremely fast; after as few as three sub iterations (11), the sequence reaches the vicinity of the global optimum.

A comparison between the steered sequential projection method [9] and our APOCS method is presented in Fig. 4. To make the comparison fair, both methods performed phase 1 of APOCS. Otherwise, the steered sequential projection method may slow down too early and not

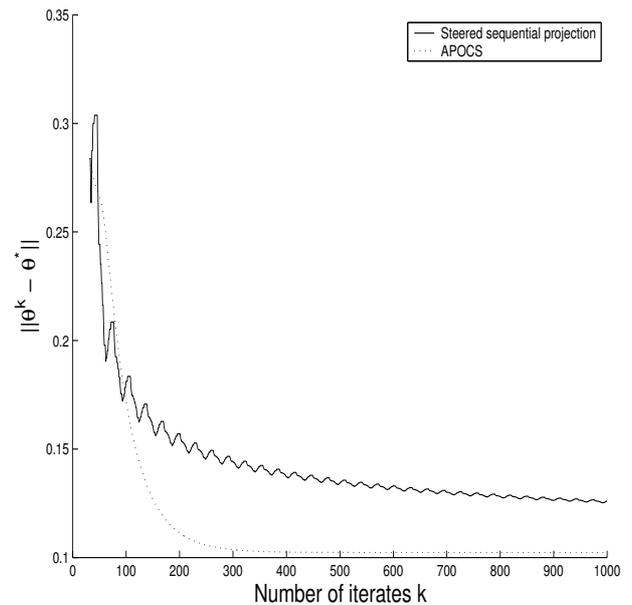


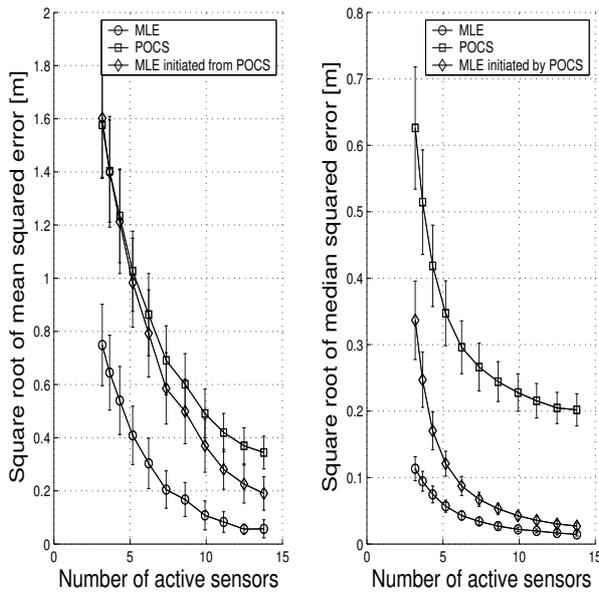
Fig. 4. Steered sequential projection method vs. APOCS.

reach the limit point in reasonable time. The APOCS method is less sensitive to the initial point as it uses a constant step size  $\mu$ . It can be seen that APOCS outperforms the steered sequential projection method in terms of the convergence rate.

The performance of POCS in terms of mean squared estimation error was also evaluated. As a benchmark, the performance of POCS was compared to the performance of the global MLE [1] and an estimator which is obtained by performing a local search on the ML objective function (2) initiated at the POCS estimator. The performance of the three estimators, as the number of sensors was increased from 100 to 2100 in 200 increments, was evaluated through 20000 Monte Carlo iterations. In Fig. 5 the square root of the mean squared error and the median squared error of the three methods are presented with  $\pm\sigma$  confidence intervals 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, so as to avoid the more ambitious task of generating the data multiple times.

## 6. CONCLUSIONS

The problem of distributed acoustic source localization using a wireless sensor network was formulated as a convex feasibility problem and solved via the APOCS method, a convergent variation of POCS. The solution has global convergence properties with fast convergence rates.



**Fig. 5.** Local performance: POCS vs. MLE, mean (left) and median (right).

## 7. 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] M. G. Rabbat, R. D. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, March 2005, pp. 1057 – 1060.
- [6] D. Blatt and A. O. Hero, "Sensor network source localization via projection onto convex sets (POCS)," in *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, USA, Mar. 2005.
- [7] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 1–24, 1967.
- [8] Y. Censor and S. A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*. New York: Oxford University Press, 1997.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] D. Blatt, A. Hero, and H. Gauchman, "A convergent incremental gradient method with constant step size," *under revision for: SIAM J. Optim.*, 2004.
- [13] 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: Springer-Verlag, New York, April 2003, pp. 285–300.