LINEAR LEAST SQUARES BASED ACOUSTIC SOURCE LOCALIZATION UTILIZING ENERGY MEASUREMENTS

D. Ampeliotis and K. Berberidis

Computer Engineering and Informatics Department and CTI/R&D
University of Patras, 26500, Rio-Patras, Greece
phone: + 30 2610 996973, fax: + 30 2610 996971,
emails: {ampeliot, berberid}@ceid.upatras.gr

ABSTRACT

Localization of an isotropic acoustic source using energy measurements from distributed sensors is considered. While most acoustic source localization algorithms require that distance estimates between the sensors and the source of interest are available, we propose a linear least squares criterion that does not make such an assumption. The new criterion can yield the location of the source and its transmit power in closed form. A weighted least squares cost function is also considered, and distributed implementation of the proposed estimators is studied. Numerical results indicate significant performance improvement as compared to a linear least squares based approach that utilizes energy ratios, and comparable performance to other estimators of higher computational complexity.

***Index Terms*—** Position measurement, Least squares methods, Distributed algorithms

1. INTRODUCTION

Localizing and tracking moving objects is an essential capability for a Wireless Sensor Network (WSN) [1]. On the other hand, sensor networks must operate using minimum resources: typical sensor nodes are battery powered and have limited processing ability. These constraints impose new challenges in algorithm development, and imply that power efficient, distributed and cooperative techniques should be employed.

Source localization methods fall mainly into two broad categories. The algorithms of the first category utilize Time Delay Of Arrival (TDOA) measurements, whereas the algorithms of the second category use Direction Of Arrival (DOA) measurements. DOA estimates are particularly useful for locating sources emitting narrowband signals [2], while TDOA measurements offer the increased capability of localizing sources emitting broadband signals [3]. However, both methods require high enough sampling rate and accurate synchronization. Recently [4], a new approach to source localization was proposed, that utilizes Received Signal Strength (RSS) measurements. In order to avoid the ambiguities due to the unknown power of the source, it was proposed to compute ratios of measurements taken at pairs of active sensors. Note that, a sensor node is characterized as active if its reading is greater than a predetermined threshold. In [5], maximum likelihood multiple-source localization based on RSS measurements was considered. In [6], a distributed "incremental subgradient" algorithm was prosposed to yield the source location estimate iteratively. In [7], the algorithm of [8] was modified so as to utilize RSS measurements instead of range difference measurements, however the source power was assumed known. Considering dense WSNs in [9], source location estimates that are robust to erroneous modelling of the energy decay function were derived by properly averaging the locations of active sensors. More recently, a distributed localization algorithm enjoying good convergence properties was proposed in [10]. In [11], a non-linear cost function for localization was proposed and it was proven that its gradient descent minimization is globally converging.

In this work, we show that distance estimation can properly be incorporated into a linear Least Squares (LS) cost function so that the source location and its unknown power can be estimated in closed form. A weighted LS cost function is also derived. Furthermore, distributed implementation of the estimators is considered.

This work is part of the 03ED910 research project, implemented within the framework of the "Reinforcement Programme of Human Research Manpower" (PENED) and co-financed by National and EC Funds

2. PROBLEM FORMULATION

The energy attenuation model of [4] is adopted. Let us consider N sensor nodes with known location vectors $\mathbf{r}_n \in \mathcal{R}^{p\times 1}$. An energy source is located at an unknown location $\mathbf{r} \in \mathcal{R}^{p\times 1}$. The RSS measurement y_n acquired by sensor node n is given by

$$y_n = g_n \frac{A}{||\mathbf{r}-\mathbf{r}_n||^2} + w_n, \qquad n = 1, 2, \dots, N \tag{1}$$

where g_n denotes the gain of node n and A denotes the power of the source as measured at 1 meter from it. The aim is to estimate the source location vector $\mathbf{r}$. Measurements y_n are corrupted by Additive White Gaussian Noise (AWGN) w_n with mean $\mu > 0$ and variance $\sigma^2 = 2\mu^2/M$, where M denotes the number of signal samples involved in the estimation of the power sample y_n [4]. In this work, for the sake of simplicity, we assume that the sensor nodes are well calibrated so that $g_n = 1$. Furthermore it is assumed that all sensors are identical so that the noise that corrupts the RSS estimates has the same statistics for all sensor nodes. However, the results presented here could be easily generalized for the case where each sensor node n has an estimate of its gain g_n and its measurement is corrupted by AWGN with mean μ_n, different for each sensor.

Provided that the distances $\rho_n = ||\mathbf{r}-\mathbf{r}_n||$ between each sensor node n and the source of interest have been estimated as $\hat{\rho}_n$, an estimate of the source location is given by (see [1] pp. 31-32, also equation (2.2) of [11]):

$$\tilde{\mathbf{r}} = \arg\min_{r} ||\mathbf{P}\mathbf{r} - \mathbf{b}||^2 \tag{2}$$

where $\mathbf{P}$ is a $N(N-1)/2 \times p$ matrix and $\mathbf{b}$ is a $N(N-1)/2 \times 1$ vector, whose k-th rows are given by

$$\mathbf{P}_k^T = 2(\mathbf{r}_j^T - \mathbf{r}_i^T), \qquad b_k = ||\mathbf{r}_j||^2 - ||\mathbf{r}_i||^2 + \hat{\rho}_i^2 - \hat{\rho}_j^2 \tag{3}$$

for $i = 1, \dots, N-1,\ j = i+1, \dots, N$ and k corresponds to an arbitrary ordering of all the $\binom{N}{2}$ pairs of sensor nodes $\{i, j\}$. Although there are $\binom{N}{2}$ linear equations in (2), it can easily be verified that only $N-1$ equations are linearly independent.

Unfortunately, the aforementioned LS criterion cannot be directly used for localization using RSS measurements, since it requires the distances of the sensors from the source, that are not available. Of course, if the power A was known, such distance estimates could be obtained using $\hat{\rho}_n \approx \sqrt{A/(y_n - \mu)}$. However, the method that will be presented in the next Section relies directly on the energy measurements and does not need a previous estimate of A.

3. THE NEW LOCALIZATION ALGORITHM

3.1. Alternative formulation of the LS problem

The k-th equation in (2) is

$$2(\mathbf{r}_j^T - \mathbf{r}_i^T)\mathbf{r} = ||r_j||^2 - ||r_i||^2 + \hat{\rho}_i^2 - \hat{\rho}_j^2 \ . \qquad (4)$$

Let us define $z_n = y_n - w_n$ and $\hat{z}_n = y_n - \mu$ with z_n denoting the noiseless measurements (deterministic) and $\hat{z}_n$ the noisy mean-corrected measurements, respectively. We add the term

$$\frac{A}{\hat{z}_j} - \frac{A}{\hat{z}_i} \qquad (5)$$

on both sides of (4), and replace the distance estimates $\hat{\rho}_n^2$ by the correct squared distances $\rho_n^2 = A/z_n$ to get

$$2(\mathbf{r}_j^T - \mathbf{r}_i^T)\mathbf{r} + \left(\frac{1}{\hat{z}_j} - \frac{1}{\hat{z}_i}\right) A = ||\mathbf{r}_j||^2 - ||\mathbf{r}_i||^2 + e_{i,j} \qquad (6)$$

where

$$e_{i,j} = A\left(\frac{w_i - \mu}{\hat{z}_i z_i}\right) - A\left(\frac{w_j - \mu}{\hat{z}_j z_j}\right) \ . \qquad (7)$$

From (7), we can observe that if the random variables $\hat{z}_i$ and $\hat{z}_j$ take large values (i.e. high SNR) then $e_{i,j}$ will behave as a noise term. Furthermore, since the expected values of the numerators in (7) are zero, we regard $e_{i,j}$ as a zero mean noise term.

We can express the $\binom{N}{2}$ equations in (6) in matrix form as

$$\mathbf{P}'\mathbf{r}' = \mathbf{b}' + \mathbf{e} \qquad (8)$$

where the k-th row of matrix $\mathbf{P}'$ is given by

$$\mathbf{P}_k'^T = \left[2(\mathbf{r}_j^T - \mathbf{r}_i^T) \qquad \left(\frac{1}{\hat{z}_j} - \frac{1}{\hat{z}_i}\right)\right] , \qquad (9)$$

the k-th elements of vectors $\mathbf{b}'$ and $\mathbf{e}$ are respectively

$$b_k' = ||\mathbf{r}_j||^2 - ||\mathbf{r}_i||^2 , \quad e_k = e_{i,j} \qquad (10)$$

and $\mathbf{r}'^T = [\mathbf{r}^T \quad A]$. Thus, minimization of $||\mathbf{P}'\mathbf{r}' - \mathbf{b}'||^2$ gives [12]

$$\tilde{\mathbf{r}}' = (\mathbf{P}'^T\mathbf{P}')^{-1}\mathbf{P}'^T\mathbf{b}' \qquad (11)$$

to yield both the source location vector $\mathbf{r}$ and its power A in closed form. In the following, we will refer to the above estimator as the Linear Least Squares (LLS) estimator. As pointed out by anonymous reviewers, an estimator similar to (11) was reported in [13]. However, in that work, no expression for $e_{i,j}$ required for computing accurate weights, as discussed in the next section, was given.

3.2. Weighted least squares

In the previous, we treated all the error terms of the LS cost function as equally important. However, each pair of nodes $\{i, j\}$ should have a different weight on the respective error term of the LS function. For example, pairs of nodes with high SNRs should have a greater influence on the cost function than pairs with smaller SNRs. Thus, in this section we derive proper weights and introduce a modification of the aforementioned linear LS problem.

Our scope is to compute a weighting matrix $\mathbf{W}$ and solve the LS problem $\mathbf{WP}'\mathbf{r} \sim \mathbf{Wb}'$. The optimal selection for matrix $\mathbf{C} = \mathbf{W}^T\mathbf{W}$ is the inverse of the noise covariance matrix [12]

$$\mathbf{C} = E[\mathbf{ee}^T]^{-1} \ . \qquad (12)$$

It can be shown that matrix $E[\mathbf{ee}^T]$ is a sparse matrix having $O(N^2)$ non-zero elements out of $\binom{N}{2}^2$ total elements. The diagonal elements of matrix $E[\mathbf{ee}^T]$ at the k-th row and k-th column ($k = 1, 2, \dots, \binom{N}{2}$) will be given by

$$\begin{aligned} E[e_{i,j}^2] &= A^2 E\left[\left(\frac{w_i - \mu}{\hat{z}_i z_i}\right)^2\right] + A^2 E\left[\left(\frac{w_j - \mu}{\hat{z}_j z_j}\right)^2\right] \\ &- 2A^2 E\left[\frac{w_i - \mu}{\hat{z}_i z_i}\right] E\left[\frac{w_j - \mu}{\hat{z}_j z_j}\right] \ . \end{aligned} \qquad (13)$$

Now, assuming that both terms of $e_{i,j}$ are zero mean and using the first-order approximation for the variance of the ratio of two random variables R_1 and R_2 given by [14]:

$$V\left[\frac{R_1}{R_2}\right] \approx V[R_2]\frac{E^2[R_1]}{E^4[R_2]} + \frac{V[R_1]}{E^2[R_2]} - 2Cov[R_1, R_2]\frac{E[R_1]}{E[R_2]}$$

we have that

$$E[e_{i,j}^2] \approx \frac{A^2}{z_i^4}\sigma^2 + \frac{A^2}{z_j^4}\sigma^2 \ . \qquad (14)$$

Since z_n denotes the unobserved noise-free measurements, we replace z_n by $\hat{z}_n$. Note that $\hat{z}_n = y_n - \mu$ is the best estimator of z_n, since we have only one measurement y_n. Furthermore, as the solution of the weighted LS problem does not change due to multiplication of the weighting matrix by a scalar, we finally have

$$E[e_{i,j}^2] \propto d_k = \frac{1}{\hat{z}_i^4} + \frac{1}{\hat{z}_j^4} \ . \qquad (15)$$

Using similar reasoning, we can find the following expressions for the non-zero non-diagonal elements of $E[\mathbf{ee}^T]$

$$E[e_{i,j}e_{j,l}] \propto -\frac{1}{\hat{z}_j^4} \quad \text{and} \quad E[e_{i,j}e_{i,l}] \propto \frac{1}{\hat{z}_i^4} \ .$$

However, keeping in mind the restrictive computation and communication abilities of wireless sensor networks, we will depart from the above optimal selection of the weighting matrix. In particular, we will assume that $E[\mathbf{ee}^T]$ is a diagonal matrix. This approach, although not optimal, has the advantage that it can be implemented in a distributed fashion. Furthermore, numerical experiments verified that small performance degradation occurs by performing this approximation. Thus, in place of matrix $\mathbf{C}$ we use the diagonal matrix

$$\hat{\mathbf{C}} = diag[1/d_1 \quad 1/d_2 \cdots 1/d_{\binom{N}{2}}] \qquad (16)$$

and in place of the associated matrix $\mathbf{W}$ we use

$$\hat{\mathbf{W}} = diag\left[\frac{1}{\sqrt{d_1}} \quad \frac{1}{\sqrt{d_2}} \cdots \frac{1}{\sqrt{d_{\binom{N}{2}}}}\right] \ . \qquad (17)$$

Input from node i: θ_{t-1}, $\mathbf{R}_{t-1}$, $\mathbf{r}_i$, $\hat{z}_i$

Set the forgetting factor λ

$\mathbf{p}_t = \left[2(\mathbf{r}_j^T - \mathbf{r}_i^T) \quad \left(\frac{1}{\hat{z}_j} - \frac{1}{\hat{z}_i} \right) \right]^T$

$b_t = ||\mathbf{r}_j||^2 - ||\mathbf{r}_i||^2$

$w_t = \begin{cases} 1 & \text{For the LLS estimator} \\ 1/\sqrt{d_k} & \text{For the DWLLS estimator} \end{cases}$

$\mathbf{v}_t = w_t \mathbf{R}_{t-1} \mathbf{p}_t$

$\mathbf{k}_t = \frac{1}{\lambda + w_t \mathbf{p}_t^T \mathbf{v}_t} \mathbf{v}_t$

$\xi_t = w_t b_t - w_t \theta_{t-1}^T \mathbf{p}_t$

$\theta_t = \theta_{t-1} + \mathbf{k}_t \xi_t$

$\mathbf{R}_t = \lambda^{-1} \mathbf{R}_{t-1} - \lambda^{-1} w_t \mathbf{k}_t \mathbf{p}_t^T \mathbf{R}_{t-1}$

Select next node l

Output to node l: θ_t, $\mathbf{R}_t$, $\mathbf{r}_j$, $\hat{z}_j$

Table 1. Distributed source localization using the RLS algorithm. Node j receives data at time t from node i, updates the estimates, and forwards its data to node l

Finally, the solution of the weighted LS problem can be expressed as

$$\tilde{\mathbf{r}}'_w = (\mathbf{P}'^T \hat{\mathbf{C}} \mathbf{P}')^{-1} \mathbf{P}'^T \hat{\mathbf{C}} \mathbf{b}' . \quad (18)$$

In the following, we will refer to the above estimator as the Diagonally Weighted Linear Least Squares (DWLLS) estimator. Also, we will refer to the estimator that does not make the diagonal matrix approximation as the Weighted Linear Least Squares estimator (WLLS).

4. DISTRIBUTED IMPLEMENTATION

Linear LS cost functions are particularly attractive since well-known adaptive algorithms (i.e. Least Mean Square (LMS), Recursive Least Squares (RLS) [15]) exist for minimization and/or tracking of the solution in a scenario where the source moves and/or its power varies with time. It is quite easy to derive such adaptive estimators from (11) and (18). Furthermore, by observing the k-th linear equation in the aforementioned cost functions, we note that it is dependent only on information available to nodes i and j. Thus, distributed minimization is possible.

Let us consider for example that all active nodes in the network are organized in a circle, and that node i has an estimate θ_{t-1} of $\tilde{\mathbf{r}}'$ (or $\tilde{\mathbf{r}}'_w$) at some discrete time instant $t-1$. Then, node i forwards its estimate to the next active node j. Table 1 describes the steps taken by node j at time t, following the RLS algorithm. Node j will in turn forward its data to the next node l, and so on.

The communication cost of the algorithm in Table 1 consists in transmitting the current estimate θ_t, the *inverse correlation matrix* $\mathbf{R}_t$, as well as the location vector $\mathbf{r}_j$ and measurement $\hat{z}_j$ in every iteration of the distributed algorithm. For static sensors, the location vectors need to be transmitted only once. Similarly, measurements $\hat{z}_j$ need to be transmitted only if they change (i.e. moving source and/or new samples are acquired). Thus, the overall communication cost for the static case after K iterations will be to transmit $Np + N + K((p+1) + (p+1)^2)$ real numbers, where p is the number of dimensions of the deployment field and N is the number of active nodes. Using the LMS approach, we save the cost of transmitting the inverse correlation matrix, thus reducing the communication cost to $Np + N + K(p+1)$ real numbers. However, this will reduce convergence speed.

Organizing the active nodes into a circle, might prove to be a challenging task. In particular, since not all sensor pairs contain the

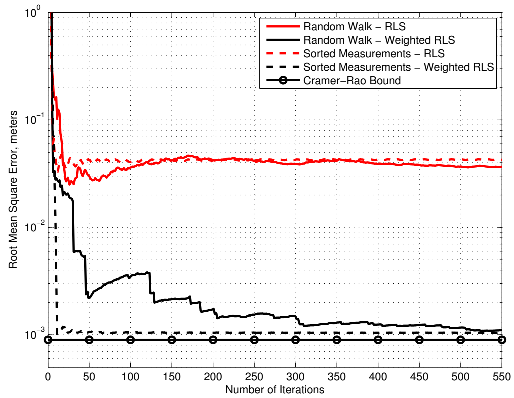

Fig. 1. RMS Localization Error as a function of the Iterations

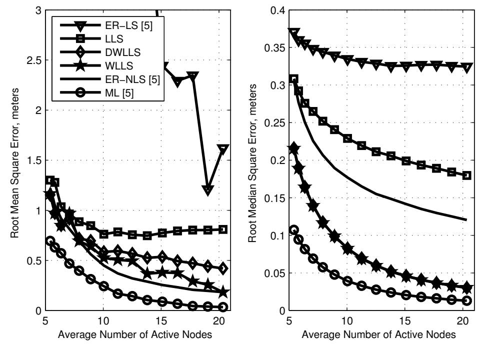

Fig. 2. Localization Error as a function of $\bar{N}$

same "quality" of information, we have that the selection of the circle is critical to the performance of the algorithm. Thus, a proper cost function should be defined for this task in conjunction with a distributed algorithm for its minimization. In order to avoid this "overhead" computation and communication, we propose to forward estimates from node to node in a random walk fashion. Thus, node j selects the next node l among its neighbors uniformly at random. Besides the computation and communication savings, this approach will use all the $\binom{N}{2}$ pairs of nodes, rather than using only N pairs along a circle. The expected time in which a random walk traverses all the edges of a complete graph of N nodes is studied in [16]. Simulation results presented in the next section, demonstrate that the random walk based approach achieves minimization of the cost function, at the cost of increasing the time required for convergence.

5. NUMERICAL RESULTS

In order to assess the convergence speed of the proposed algorithms we simulated a sensor network in which 1500 nodes were uniformly developed over a $100m \times 100m$ field. A signal source with $A = 100$ was located at $\mathbf{r} = [50\ 50]^T$. The noise variance was $\sigma^2 = 1$. $N = 11$ active nodes cooperated for the estimation based on their distance from the source. Figure 1 presents the Root Mean Squared (RMS) error of the location estimate as a function of the iterations performed. A total of $50N$ iterations were conducted and the results are the average of 1000 different noise realizations. We examined

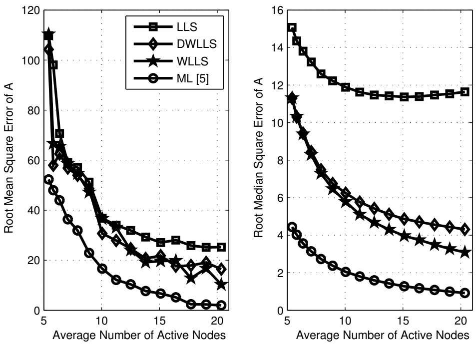

Fig. 3. Error for A estimation as a function of $\bar{N}$

the random walk based RLS versions of the LLS and DWLLS estimators, the RLS versions of LLS and DWLLS where the nodes have been organized in a circle in order of decreasing measurement, and the Cramer-Rao lower bound [5]. All algorithms start from an initial location at $[100\ 100]^T$ while the initial value of A was set equal to zero. The inverse correlation matrix for all RLS algorithms was initialized to $100\mathbf{I}$ while the *forgetting factor* λ was 1. From Figure 1, we note that organization of nodes in a circle of decreasing measurement gives very fast convergence speed, as expected, and a single circle (11 hops) is adequate for convergence. In practice however, one should also take into account the time required to achieve this organization of the nodes. On the other hand, the random walk based algorithms converge slower but require no organization of the nodes. Interestingly, the DWLLS estimator in this setting achieves a performance close to the Cramer-Rao bound.

In another simulation, the performance of the proposed estimators was compared to the the ER-LS, ER-NLS and Maximum Likelihood (ML) estimators of [5]. The number of nodes was increased from 300 to 3100 in 200 increments. Active nodes were selected as those with measurement higher than a threshold equal to $5 + \mu$, i.e. only sensors whose SNR ($(y_n - \mu)/\sigma^2$) is greater than 7dB take part in the estimation procedure. For ML estimation, two multiresolution iterations were performed, using a $10m \times 10m$ square grid centered at the actual source location for the first iteration, and a $5m \times 5m$ grid for the second iteration. Each grid used 81×81 search points logarithmically distributed around their center. Also, optimization for the ER-NLS approach was conducted by applying the method described in [10]. Figures 2 and 3 demonstrate the error (RMS and root median square) for localization and power estimation respectively, as a function of the average number of active nodes $\bar{N}$. The results are the average of 50000 Monte Carlo runs. Network realizations resulting in less than 5 active nodes were not permitted.

Figure 2 demonstrates that the proposed estimators offer better localization accuracy than the ER-LS, and comparable performance to the ER-NLS approach. In particular, one can observe that the ER-NLS approach (optimized using POCS [10]) gives smaller RMS error than the DWLLS, while the DWLLS approach achieves better root median square error. We noticed that this phenomenon is due to some "ill-conditioned" network topologies that may be realized. These topologies result in very big location errors for LS-based algorithms and dominate RMS error calculation. In order to overcome this, we have applied robust methods based on the Huber norm [17]. Numerical results not included here have shown the effectiveness of such robust methods in tackling this phenomenon.

Figure 3 demonstrates that the proposed estimators provide estimates for the power of the acoustic source that get close to the estimates provided by the Maximum Likelihood estimator, as the network density increases. Note that the ER-LS and ER-NLS estimators of [5], do not provide an estimate of A and thus they are not included in Figure 3.

6. CONCLUSIONS

In this work, we derived novel linear least squares criteria for joint estimation of the location and power of an acoustic source. The proposed criteria utilize received signal strength measurements and are particularly suited to low-cost networks of distributed sensors. Weighted least squares were also considered and distributed implementation was studied. Numerical results verified that the proposed estimators offer good performance-complexity trade-offs as compared to existing estimators. Future work will focus on the application of optimization methods based on robust statistics in order to explore the potential performance improvement of the proposed criteria.

7. REFERENCES

[1] F. Zhao and L. Guibas, *Wireless Sensor Networks: An Information Processing Approach*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.

[2] S. Haykin, *Array signal processing*, Englewood Cliffs, NJ, Prentice-Hall, Inc., 1985.

[3] N. Owsley and G. Swope, "Time delay estimation in a sensor array," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 29, no. 3, pp. 519–523, 1981.

[4] D. Li and Y.-H. Hu, "Energy-based collaborative source localization using acoustic microsensor array," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 1, pp. 321–337, 2003.

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

[6] M. G. Rabbat and R. D. Nowak, "Decentralized source localization and tracking," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2004, pp. 921–924.

[7] T. Pham, B. M. Sadler, and H. Papadopoulos, "Energy-based source localization via ad-hoc acoustic sensor network," *Statistical Signal Processing, 2003 IEEE Workshop on*, pp. 387–390, 28 Sept.-1 Oct. 2003.

[8] B. Friedlander, "A passive localization algorithm and its accurancy analysis," *Oceanic Engineering, IEEE Journal of*, vol. 12, no. 1, pp. 234–245, Jan 1987.

[9] M. G. Rabbat, R. D. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2005, pp. 1057–1060.

[10] D. Blatt and A. O. Hero, "Energy-based sensor network source localization via projection onto convex sets," *Signal Processing, IEEE Transactions on*, vol. 54, no. 9, pp. 3614–3619, September 2006.

[11] B. Fidan, S. Dasgupta, and B. D. O. Anderson, "Conditions for guaranteed convergence in sensor and source localization," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2007, pp. II–1081 – II–1084.

[12] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1999.

[13] T. Pham and H.C. Papadopoulos, "Distributed tracking in ad-hoc sensor networks," *Statistical Signal Processing, 2005 IEEE/SP 13th Workshop on*, pp. 1226–1231, 17-20 July 2005.

[14] M. Kendall and A. Stuart, *The Advanced Theory of Statistics. Vol. 2: Inference and Relationship*, Number 2. C. Griffin, London, 1979.

[15] S. Haykin, *Adaptive Filter Theory (4th Edition)*, Prentice Hall, September 2001.

[16] D. Zuckerman, "On the time to traverse all edges of a graph," *Information Processing Letters*, vol. 38, no. 6, pp. 335–337, 1991.

[17] Peter J. Huber, "Robust regression: Asymptotics, conjectures and monte carlo," *The Annals of Statistics*, vol. 1, no. 5, pp. 799–821, 1973.