Proceedings Paper:

Salman, N, Mihaylova, L and Kemp, AH (2014) Localization of multiple nodes based on correlated measurements and shrinkage estimation. In: Proceedings of the 2014 Workshop on Sensor Data Fusion: Trends, Solutions, Applications, SDF 2014. 2014 Workshop on Sensor Data Fusion: Trends, Solutions, Applications, SDF 2014, 08-10 Oct 2014, Bonn. IEEE , 1 - 6. ISBN 9781479973873

https://doi.org/10.1109/SDF.2014.6954712

A Shrinkage Approach to Multiple Target Localization with Correlated Measurements

Naveed Salman*, *Student Member, IEEE,* Lyudmila Mihaylova*, *Senior Member, IEEE,* and A. H. Kemp**, *Member, IEEE*

***Abstract*—In this paper, we deal with the accurate estimation of the covariance matrix for the optimization of multiple target node (TN) localization using correlated received signal strength (RSS) measurements. Two location schemes i.e. the iterative generalized least squares (GLS) and the low complexity weighted linear least squares (WLLS) methods are investigated. For many applications the estimated covariance matrix needs to be positive definite and hence invertible. For an insufficient number of data points, the sample covariance matrix suffers from two drawbacks. Firstly, although it is unbiased, it consists of a large estimation error. Secondly, it is not positive definite. A shrinkage technique to estimate the covariance matrix has been proposed in the fields of finance and life sciences. In this paper, we introduce the shrinkage covariance matrix concept in the area of multiple TN localization in wireless networks with correlated measurements. An analytical expression of the multiple TN covariance matrix is derived for the WLLS method and the estimation is done via the shrinkage technique. Similarly, the shrinkage technique is employed for the covariance matrix estimation for the GLS algorithm. For a limited number of measurements, the shrinkage covariance technique improves the performance of both WLLS and GLS considerably.**

***Index Terms*—Localization, Received signal strength (RSS), Shrinkage estimation.**

I. INTRODUCTION

For a set of random variables, the covariance matrix is a symmetric matrix, whose diagonal elements represent the individual variances of the random variables while the off-diagonal elements are the corresponding cross covariances. When the variance of all random variables is set to unity, the covariance matrix is then known as the correlation matrix. The covariance matrices have applications in various fields. In finance, portfolio analysis is done via estimation of the covariance matrix of stock returns [1], [2]. In life sciences it is used in genes classification by finding the pairwise correlation between genes [3]. In most cases, the elements of the covariance matrix are unknown and need to be estimated. Due to its ease of use, the sample estimator is commonly used for covariance matrix estimation. Although the sample estimator is unbiased, it faces serious issues when the number of samples is equal or smaller than the number of variables. If the samples are equal or less than the number of variables, the sample covariance matrix will contain a large estimation error. Furthermore, the sample covariance matrix will not be positive definite, making it impractical to use. To elaborate on the previously given examples, for portfolio analysis, if the number of stocks is larger than the number of historical (e.g. monthly) stock returns then the corresponding sample covariance matrix will contain large errors and will always be singular. Similar problems are faced in gene classification research, if the number of genes under considerations exceeds the number of experiments performed.

The idea of shrinkage estimation for large covariance matrices was introduced rather recently in finance applications [2]. As mentioned before, the sample estimator is unbiased but poses large variance, the shrinkage idea is to introduce some structure to the sample covariance matrix. A target covariance matrix is chosen (which has a smaller number of variables) as a candidate for covariance matrix estimate. Due to the small number of parameters, the target covariance matrix will have a relatively small error variance although it will be biased. The idea is to choose an optimally compromised covariance matrix instead of the two extremes i.e. the large variance but unbiased sample covariance matrix and low variance but biased target covariance matrix.

In this paper, we apply the shrinkage theory for covariance matrix estimation to multiple node localization in wireless networks. Node localization in wireless networks is needed in many application including situation awareness, cattle/wild life monitoring and logistics [4]. A number of techniques have been developed for accurate positioning of wireless nodes. These may include the time of arrival (ToA) [5] and angle of arrival (AoA) methods [6]. Both ToA and AoA provide an accurate distance and hence location estimation. However, both techniques require additional hardware on board the target nodes (TNs). A prerequisite of ToA method is the availability of highly accurate and synchronized clocks on the TNs, while AoA requires an array of antennas [6], [7]. A low cost and simplistic technique to localize wireless nodes is the received signal strength (RSS) technique [8]. Considerable amount of research has been done to optimize the performance of RSS localization. Joint estimation of the path-loss exponent (PLE) and location is discussed in [9], [10], [11]. Cooperative RSS localization is studied in [12]. Tracking using RSS technique is investigated in [13]. Most of these studies assume individual links between TNs and sensor node (SN) to be independent of each other. This is an oversimplification as readings at the SN from TNs that are close to each other in a network introduces an element of spatial correlation. In this study we simulate our data using the widely accepted Gudmundson correlation model presented in [14]. The Gudmundson's model suggests that the correlation between two nodes is an exponentially

*Department of Automatic Control and Systems Engineering, University of Sheffield, U.K. (email: {n.salman, l.s.mihaylova}@sheffield.ac.uk)
**School of Electronic and Electrical Engineering, University of Leeds, U.K. (email: a.h.kemp@leeds.ac.uk)

decaying function of the distance between them. The authors of [15] show that if these correlations are taken into account they could enhance the location estimation performance. This proposition was done by deriving the Cramer-Rao bound for spatial correlation between nodes, however an algorithm capitalizing on such correlation was not presented.

In this paper, we optimize the performance of two multiple TN RSS location algorithms. Location estimation of TNs is a non-linear estimation problem. Due to the non-linear nature, the location estimation is generally performed using an iterative algorithm with a close initial estimate and Newton step update. However, location coordinates can also be estimated using a linear least squares (LLS) approach. In this paper we first develop an iterative generalized least squares (GLS) technique for multiple TN localization. Second, we also formulate a weighted LLS (WLLS) solution to the multiple TN location problem. A closed form expression is derived for the covariance matrix of multiple TN readings at the SNs. The covariance matrix expression depends on the variance and covariance values of these readings. In practical scenarios and in large networks these values are unknown and the number of snapshots is not large enough to provide an accurate sample estimate due to limited resources and packet lose. To counter this problem, we apply the shrinkage technique to estimate the variance and covariance elements required in the expression for the covariance matrix in the WLLS method. The shrinkage technique is also applied to estimate the covariance matrix for the GLS technique. Due to the lack of structure and for a limited number of snapshots the sample covariance matrix is non invertible rendering it impractical to use in the GLS method. The shrinkage covariance estimator on the other hand introduces a structure to the covariance matrix and hence it is invertible even for limited number of snapshots. It is shown via simulation that the shrinkage covariance matrix compared to the sample covariance matrix has a smaller estimation error. Consequently, the shrinkage covariance matrix, when used in the GLS and WLLS algorithm, results in superior performance. Especially in the case of GLS case, where the sample covariance estimator is non invertible for a limited number of snapshots, this problem is resolved with the shrinkage covariance matrix.

The rest of the paper is organized as follows. Section II introduces the signal model and the multiple TN location problem. Section III, develops the GLS and the WLLS estimator. In section IV we briefly discuss the shrinkage theory and covariance shrinkage estimator. Section V presents the simulation results which are followed by the conclusions. The derivation of the shrinkage estimator is provided in the Appendix.

II. System model

For future use, we define the following notations.

$\mathcal{R}^n$ is the set of n dimensional real numbers, $(.)^T$ is the transpose operator, $E(.)$ refers to the expectation operator, $[\mathbf{M}]_{ij}$ refers to the element at the i^{th} row and j^{th} column of matrix $\mathbf{M}$, $\mathbf{I}_N$ represents the N dimensional identity matrix. $\mathcal{N}(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2. $\|.\|_F^2$ represents the Frobenius norm.

A two dimensional (2-D) network is considered, consisting of N SNs with known locations $\boldsymbol{\phi}_{i,j} = [x_{i,j}, y_{i,j}]^T \left(\boldsymbol{\phi}_{i,j} \in \mathcal{R}^2\right)$ for $i,j = 1,...,N$. The network also consists of M TNs which have unknown coordinates $\boldsymbol{\theta}_{k,l} = [x_{k.l}, y_{k,l}]^T \left(\boldsymbol{\theta}_{k,l} \in \mathcal{R}^2\right)$ for $k,l = 1,\ldots,M$ that are to be estimated. The received power at the SNs due to random shadowing is log-normally distributed. This model is based on empirical results obtained in [16], [17]. It is assumed that the TNs transmit during preassigned epochs to avoid interference between different TN transmissions. Thus the distance d_{ik} between the k^{th} TN and the i^{th} AN is related to the path-loss at the i^{th} AN, $\mathscr{L}_{ik}$, and the PLE, α, as [18]

$$\mathscr{L}_{ik} = \mathscr{L}_0 + 10\alpha \log_{10} d_{ik} + w_{ik}, \tag{1}$$

where $\mathscr{L}_0$ is the path-loss at the reference distance d_0 ($d_0 < d_{ik}$, and is normally taken as 1 m) and w_{ik} is a zero-mean Gaussian random variable representing the multipath log-normal shadowing effect, i.e. $w_{ik} \sim \left(\mathcal{N}\left(0, \sigma_{ik}^2\right)\right)$. It is assumed that α is known via prior channel modeling or accurate estimation. The path-loss is calculated as

$$\mathscr{L}_{ik} = 10\log_{10} P_k - 10\log_{10} P_{ik}, \tag{2}$$

where P_k is the transmit power at the k^{th} TN and P_{ik} is the received power at the i^{th} AN. The distance d_{ik} is given by

$$d_{ik} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}. \tag{3}$$

The observed path-loss (in dB) from d_0 to d_{ik}, $z_{ik} = \mathscr{L}_{ik} - \mathscr{L}_0$, can be expressed as

$$z_{ik} = f_i(\boldsymbol{\theta}_k) + w_{ik}, \tag{4}$$

where $f_i(\boldsymbol{\theta}_k) = \gamma\alpha \ln d_{ik}$ and $\gamma = \frac{10}{\ln 10}$. (4) can be written in a matrix form as

$$\mathbf{Z} = \mathbf{F}(\boldsymbol{\theta}) + \mathbf{W}, \tag{5}$$

where the k^{th} column of $\mathbf{Z}$ and $\mathbf{W}$ represents the readings and associated noise elements at the SNs from the k^{th} TN respectively and $[\mathbf{F}(\boldsymbol{\theta})]_{ik} = f_i(\boldsymbol{\theta}_k)$. For convenience (5) can also be written in a 'roll out' form i.e. $\mathbf{z} = \text{vec}(\mathbf{Z})$, then for the k^{th} target, we have

$$\begin{bmatrix} z_{1k} \\ z_{2k} \\ \vdots \\ z_{Nk} \end{bmatrix} = \begin{bmatrix} f_1(\boldsymbol{\theta}_k) \\ f_2(\boldsymbol{\theta}_k) \\ \vdots \\ f_N(\boldsymbol{\theta}_k) \end{bmatrix} + \begin{bmatrix} w_{1k} \\ w_{2k} \\ \vdots \\ w_{Nk} \end{bmatrix}$$

or

$$\mathbf{z}_k = \mathbf{f}_k + \mathbf{w}_k. \tag{6}$$

Then we have $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{C}_{kk})$ where $\mathbf{C}_{kk} = \sigma_{ik}^2 \mathbf{I}_N$ and $E\left(\mathbf{w}_k (\mathbf{w}_l)^T\right) = \mathbf{C}_{kl}$. Here $\mathbf{C}_{kl}$ is the $N \times N$ covariance matrix between RSS readings at the SNs from the k^{th} and l^{th} TN and its elements are given by $\mathbf{I}_N \rho_{kl} \sigma_{ik} \sigma_{il}$. Here ρ_{kl} is the spatial correlation between the k^{th} and l^{th} TN. Its value depends on the relative distance between the TNs and can be modeled according to Gudmundson [14] as follows

$$\rho_{kl} = \exp\left(-\frac{d_{kl}}{d_c}\right), \tag{7}$$

where d_c is the 'decorrelation distance'. Field measurements in [23] suggest values for d_c for different environments. Eq. (6) when written for all TNs is given by

$$\mathbf{z} = \mathrm{vec}(\mathbf{Z}) = \left((\mathbf{z}_1)^T, (\mathbf{z}_2)^T, ..., (\mathbf{z}_M)^T\right)^T \tag{8}$$

or

$$\mathbf{z} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{w}, \tag{9}$$

where $\mathbf{w} \sim \mathcal{N}(0, \mathbf{C})$, where $\mathbf{C}$ encompasses correlations between all TNs;

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1M} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{M1} & \mathbf{C}_{M2} & \cdots & \mathbf{C}_{MM} \end{bmatrix}.$$

Two location algorithms to solve (9) are discussed in the next section.

III. Localization algorithms for multiple TNs

A. Generalized least squares algorithm

It is clear that (9) presents a non-linear estimation problem and a close form expression to solve (9) is not readily available. A solution to (9) is obtained by minimizing the following cost function [19]

$$\epsilon(\boldsymbol{\theta}) = (\mathbf{z} - \mathbf{f}(\boldsymbol{\theta}))^T \mathbf{C}^{-1} (\mathbf{z} - \mathbf{f}(\boldsymbol{\theta})), \tag{10}$$

where (10) can be minimized by first linearizing $\mathbf{f}(\boldsymbol{\theta})$ by the first order Taylor series expansion to a close initial estimate $\mathbf{f}(\boldsymbol{\theta}^a)$ i.e.

$$\mathbf{f}(\boldsymbol{\theta}) = \mathbf{f}(\boldsymbol{\theta}^a) + \mathbf{F}(\boldsymbol{\theta}^a)(\boldsymbol{\theta} - \boldsymbol{\theta}^a), \tag{11}$$

where

$$\mathbf{F}(\boldsymbol{\theta}^a) = \mathrm{diag}\left[\mathbf{F}(\boldsymbol{\theta}_1^a), \mathbf{F}(\boldsymbol{\theta}_2^a), ..., \mathbf{F}(\boldsymbol{\theta}_M^a)\right]$$

and $\mathbf{F}(\boldsymbol{\theta}_k^a)$ is the Jacobian matrix of the k^{th} TN and its elements are given by

$$\mathbf{F}(\boldsymbol{\theta}_k^a) = \begin{bmatrix} \frac{\partial f_1(\boldsymbol{\theta}_k)}{\partial x_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} & \frac{\partial f_1(\boldsymbol{\theta}_k)}{\partial y_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} \\ \frac{\partial f_2(\boldsymbol{\theta}_k)}{\partial x_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} & \frac{\partial f_2(\boldsymbol{\theta}_k)}{\partial y_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} \\ \vdots & \vdots \\ \frac{\partial f_N(\boldsymbol{\theta}_k)}{\partial x_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} & \frac{\partial f_N(\boldsymbol{\theta}_k)}{\partial y_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} \end{bmatrix}$$

for $\frac{\partial f_i(\boldsymbol{\theta}_k)}{\partial x_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} = \gamma\alpha(x_k^a - x_i)/d_{ik}^2$ and $\frac{\partial f_i(\boldsymbol{\theta}_k)}{\partial y_k}\Big|_{\boldsymbol{\theta}_k=\boldsymbol{\theta}_k^a} = \gamma\alpha(y_k^a - y_i)/d_{ik}^2$.

Using (11) in (10) and minimizing with respect to $\boldsymbol{\theta}$, so that the next estimation is given by

$$\boldsymbol{\theta}^{a+1} = \boldsymbol{\theta}^a + \varrho^a \tag{12}$$

where

$$\varrho^a = \left(\mathbf{F}(\boldsymbol{\theta}^a)^T \mathbf{C}^{-1} \mathbf{F}(\boldsymbol{\theta}^a)\right)^{-1} \mathbf{F}(\boldsymbol{\theta}^a)^T \mathbf{C}^{-1} (\mathbf{z} - \mathbf{f}(\boldsymbol{\theta}^a)). \tag{13}$$

Thus given an initial estimate $\boldsymbol{\theta}^1$, the generalized least squares (GLS) converges to the minima. The algorithm is stopped after a fixed number of iterations or when the value $\left|\epsilon\left(\boldsymbol{\theta}^{a+1}\right) - \epsilon(\boldsymbol{\theta}^a)\right|$ becomes smaller than a certain threshold. To avoid the complexity of the iterative procedure and a close initial estimate condition at the expanse of accuracy, a WLLS solution is presented in the next subsection.

B. Weighted linear least squares algorithm

A slight manipulation of (9) renders it to be linear. This technique was first proposed for ToA distance estimates in [20] and analyzed in [21]. For a single target RSS measurements the linearized model is discussed in [22]. Here we extend this method for the multiple TN case as follows. From (9) it can be readily shown that

$$E\left(\frac{1}{\beta_{ik}} \exp\left(\frac{2z_{ik}}{\gamma\alpha}\right)\right) = \hat{d}_{ik}^2, \tag{14}$$

where $\beta_{ik} = \exp\left(\frac{2\sigma_{ik}^2}{(\gamma\alpha)^2}\right)$. Similarly choosing a reference AN, it can be shown

$$E\left(\frac{1}{\beta_{rk}} \exp\left(\frac{2z_{rk}}{\gamma\alpha}\right)\right) = \hat{d}_{rk}^2, \tag{15}$$

where $\beta_{rk} = \exp\left(\frac{2\sigma_{rk}^2}{(\gamma\alpha)^2}\right)$. For linearization, the square of each distance equation is subtracted from the square of a reference distance equation. This results in a linear system which is represented in matrix form as[1]

$$\mathbf{b} = \mathbf{A}\boldsymbol{\theta} + \mathbf{v}, \tag{16}$$

where $\mathbf{b} = [\mathbf{b}_1, ..., \mathbf{b}_M]^T$ is the observation vector at the SNs from the TNs. The linearized observation from the k^{th} TN at the ANs is thus given as

$$\mathbf{b}_k = \begin{bmatrix} \delta_{rk} - \delta_{1k} - \kappa_{rk} + \kappa_1 \\ \delta_{rk} - \delta_{2k} - \kappa_{rk} + \kappa_2 \\ \vdots \\ \delta_{rk} - \delta_{N-1k} - \kappa_{rk} + \kappa_{N-1} \end{bmatrix}$$

for $\delta_{rk} = \frac{1}{\beta_{rk}} \exp\left(\frac{2z_{rk}}{\gamma\alpha}\right)$ and $\delta_{ik} = \frac{1}{\beta_{ik}} \exp\left(\frac{2z_{ik}}{\gamma\alpha}\right)$. While

$$\kappa_{rk} = x_{rk}^2 + y_{rk}^2 \text{ and } \kappa_i = x_i^2 + y_i^2$$

for $i \neq r$, $i = 1, ..., N-1$. Also (x_{rk}, y_{rk}) are the coordinates of the reference SN for the k^{th} TN. $\mathbf{A}$ is the $M(N-1) \times 2M$ data matrix for all TNs and is given by

$$\mathbf{A} = \mathrm{diag}\left(\mathbf{A}^1, \mathbf{A}^2 \cdots \mathbf{A}^M\right).$$

The diagonal elements of $\mathbf{A}$ are themselves data matrices for the individual TNs. Thus for the k^{th} TN, the data matrix is given by

[1]The inclusion of the constants β_{ik}, β_{rk} in (14) and (15) is essential for the LLS solution to be unbiased.

$$\mathbf{A}^k = 2\begin{bmatrix} x_1 - x_{rk} & y_1 - y_{rk} \\ x_2 - x_{rk} & y_2 - y_{rk} \\ \vdots & \vdots \\ x_{N-1} - x_{rk} & y_{N-1} - y_{rk} \end{bmatrix}.$$

$\mathbf{v}$ is the noise vector which has zero mean and a $M(N-1) \times M(N-1)$ covariance matrix $\mathbf{C_v}$ given by

$$\mathbf{C_v} = \begin{bmatrix} \mathbf{C_{v}}_{11} & \mathbf{C_{v}}_{12} & \cdots & \mathbf{C_{v}}_{1M} \\ \mathbf{C_{v}}_{21} & \mathbf{C_{v}}_{22} & \cdots & \mathbf{C_{v}}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C_{v}}_{M1} & \mathbf{C_{v}}_{M2} & \cdots & \mathbf{C_{v}}_{MM} \end{bmatrix}. \tag{17}$$

Expressions for the individual elements of $\mathbf{C_v}$ are given as follows.

The diagonal elements of $\mathbf{C_{v}}_{kk}$ are given by

$$[\mathbf{C_{v}}_{kk}]_{ii} = E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)^2\right]$$

$$= d_{ik}^4 \exp\left(\frac{4\sigma_{ik}^2}{(\gamma\alpha)^2}\right) - d_{ik}^4 + d_{rk}^4 \exp\left(\frac{4\sigma_{rk}^2}{(\gamma\alpha)^2}\right) - d_{rk}^4 \tag{18}$$

and the off-diagonal elements of $\mathbf{C_{v}}_{kk}$ are given by[2]

$$[\mathbf{C_{v}}_{kk}]_{ij} = E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rk} - \delta_{jk} - d_{rk}^2 + d_{jk}^2\right)\right]$$

$$= \left[d_{rk}^4 \exp\left(\frac{4\sigma_{rk}^2}{(\gamma\alpha)^2}\right) - d_{rk}^4\right]. \tag{19}$$

The diagonal elements of $\mathbf{C_{v}}_{kl}$ are given as

$$[\mathbf{C_{v}}_{kl}]_{ii} = E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rl} - \delta_{il} - d_{rl}^2 + d_{il}^2\right)\right]$$

$$= a1 + a2 - a3 - a4 \tag{20}$$

where

$$a1 = \frac{1}{\beta_{rk}}\frac{1}{\beta_{rl}} d_{rk}^2 d_{rl}^2 \exp\left(\frac{2\left(\sigma_{rk}^2 + \sigma_{rl}^2 + 2\rho_{kl}\sigma_{rk}\sigma_{rl}\right)}{(\gamma\alpha)^2}\right)$$

$$a2 = \frac{1}{\beta_{ik}}\frac{1}{\beta_{il}} d_{ik}^2 d_{il}^2 \exp\left(\frac{2\left(\sigma_{ik}^2 + \sigma_{il}^2 + 2\rho_{kl}\sigma_{ik}\sigma_{il}\right)}{(\gamma\alpha)^2}\right)$$

$$a3 = d_{rk}^2 d_{rl}^2 \quad \text{and} \quad a4 = d_{ik}^2 d_{il}^2.$$

Finally, the off diagonal elements of $\mathbf{C_{v}}_{kl}$ are given by

$$[\mathbf{C_{v}}_{kl}]_{ij} = E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rl} - \delta_{jl} - d_{rl}^2 + d_{jl}^2\right)\right]$$

$$= b1 - b2 \tag{21}$$

[2]Here the correlation between the measurements of the same TN at different ANs is not considered. This assumption is based upon two arguments. First, for efficient operation of the localization algorithm, the SNs are placed far apart hence retaining little correlation. Second, even if the SNs are placed close to each other, distance and hence correlation between them is known and can easily be incorporated in (19).

where

$$b1 = \frac{1}{\beta_{rk}}\frac{1}{\beta_{rl}} d_{rk}^2 d_{rl}^2 \exp\left(\frac{2\left(\sigma_{rk}^2 + \sigma_{rl}^2 + 2\rho_{kl}\sigma_{rk}\sigma_{rl}\right)}{(\gamma\alpha)^2}\right)$$

and

$$b2 = d_{rk}^2 d_{rl}^2.$$

The WLLS solution is obtained by minimizing the cost function

$$\varepsilon_{WLLS}(\boldsymbol{\theta}) = (\mathbf{b} - \mathbf{A}\boldsymbol{\theta})^T \mathbf{C_v}^{-1} (\mathbf{b} - \mathbf{A}\boldsymbol{\theta}), \tag{22}$$

The WLLS estimate is obtained as follows

$$\hat{\boldsymbol{\theta}}_{WLLS} = \mathbf{A}^{\ddagger}\mathbf{b}^{\ddagger}, \tag{23}$$

where $\mathbf{A}^{\ddagger} = \left[\mathbf{A}^T\mathbf{C_v}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^T$ and $\mathbf{b}^{\ddagger} = \mathbf{C_v}^{-1}\mathbf{b}$. Since the actual distances are not available, their estimated values are used in (23). The variance and correlation elements of $\mathbf{C_v}$ can be estimated accurately by the sample estimation of $\mathbf{C_v}$ for a large number of readings or snapshots. For a limited number of snapshots, a shrinkage estimator is proposed in section IV. The WLLS technique presented is relative low in complexity and does not require an initial estimate; however its performance is sub-optimal. This is because information from all the SNs is not utilized as the measurement from reference SN is used to linearize the system.

IV. Practical issues and shrinkage estimation

In practical scenarios the covariance matrix $\mathbf{C}$ is unknown and it needs to be estimated. For seamless operation of the iterative algorithm (12), $\mathbf{C}$ has to be positive definite and hence invertible. For a sample covariance matrix to be positive definite, the number of sample points need to be more than the number of variables. For multiple TN this poses a serious problem for the invertibility of $\mathbf{C}$ as the number readings need to exceed the number of unknowns which can not always be guaranteed in resource constraint networks. Furthermore, a sample covariance matrix estimate with limited number of sample points consists of large errors. Erroneous estimates of $\mathbf{C}$ could lead to performance degradation of the GLS algorithm. Similarly, the variance elements that are required to generate the covariance matrix $\mathbf{C_v}$ for the WLLS procedure need to be estimated. Here again, the sample estimator is not efficient for a limited number of snapshots. Thus, alternative techniques to estimate $\mathbf{C}$ and $\mathbf{C_v}$ needs investigation to insure minimum error and positive definiteness. In this paper we investigate the application of shrinkage estimation technique for covariance matrix estimation. The shrinkage estimator is introduced in the following subsection.

A. Shrinkage estimation of the covariance matrix

1) Shrinkage theory: In the context of multiple TN localization, we briefly explain the theory behind the shrinkage estimation as proposed by Ledoit and Wolf [2] for portfolio analysis. For a large dimension covariance matrix $\mathbf{C}$ of size

$NM \times NM$, let $\hat{\mathbf{C}}$ be the sample estimator of $\mathbf{C}$, then its elements are given by

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{1}{P-1}\sum_{p=1}^{P}\left(z_{ip} - \widehat{z}_i\right)\left(z_{jp} - \widehat{z}_j\right), \qquad (24)$$

where $i, j = 1, \ldots NM$ and $\widehat{z}_i$ and $\widehat{z}_j$ represent the sample means. The sample estimator in (24) is unbiased and easy to generate. Despite these advantages, for a small number of snapshots P, $\hat{\mathbf{C}}$ will exhibit a large error variance due to the large number of unknown parameters to be estimated. We can however introduce some structure to the sample covariance generally by reducing the number of free variables. Such structured estimates offer a relatively small error variance however they can be biased due to a misspecified structure. An optimal statistical solution is to reach an optimal tradeoff between the the unbiased but high variance estimate and the low variance but biased estimate. Thus the unbiased sample covariance is shrinked towards a biased structured covariance estimate.

To formalize the discussion and develop an expression for the shrinkage covariance estimate. We define $\mathbf{C}_0$ to be the true covariance matrix of the path-loss readings $\mathbf{z}$. Let $\hat{\mathbf{C}}$ be an unbiased estimate of $\mathbf{C}_0$. Let also the shrinkage target $\mathbf{T}$ be another biased structured estimate of $\mathbf{C}_0$ with relatively small number of parameters, due to which $\mathbf{T}$ will consist of a relatively small error variance. Thus instead of choosing between the two extremes, the shrinkage estimator $\mathbf{S}$ combines both estimates in a weighted fashion i.e.

$$\mathbf{S} = \lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}}, \qquad (25)$$

where λ is the shrinkage intensity and its value is between 0 and 1. It is noted that if $\lambda = 1$, then the shrinkage estimate is the shrinkage target and the unbiased covariance is given no weight. On the other hand, if $\lambda = 0$, then no shrinkage takes place and the unbiased estimator is chosen as the shrinkage estimate. Here two obvious questions arise, firstly, how to select the shrinkage target $\mathbf{T}$ and secondly, what value should be given to the shrinkage intensity λ. A number of shrinkage targets can be used in (25), e.g. [3] lists six target shrinkage intensities. In general, selection of $\mathbf{T}$ should be driven by the lower dimension structure in the data. Thus in the case of the covariance matrix $\mathbf{C}_0$ for multiple TNs, the natural shrinkage target is the constant correlation shrinkage target. For the constant correlation shrinkage target $[\mathbf{T}]_{ii} = \left[\hat{\mathbf{C}}\right]_{ii}$ and $[\mathbf{T}]_{ij} = \hat{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}$, where $\hat{\rho}$ is the average correlation of all the correlations between the network nodes i.e.

$$\hat{\rho} = \frac{1}{NM\,(NM-1)}\sum_{i=1}^{NM}\sum_{j\neq i}^{NM}\frac{\left[\hat{\mathbf{C}}\right]_{ij}}{\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}}. \qquad (26)$$

2) Optimal shrinkage intensity: In order to find the optimal shrinkage intensity $\hat{\lambda}$, the Frobenius norm of the difference between the shrinkage estimate and the true covariance matrix is minimized with respect to λ. This is achieved by minimizing the following cost function

$$L(\lambda) = E\left\|\lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}} - \mathbf{C}_0\right\|_F^2 \qquad (27)$$

or equivalently

$$L(\lambda) = E\left\{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left(\lambda\,[\mathbf{T}]_{ij} + (1-\lambda)\left[\hat{\mathbf{C}}\right]_{ij} - [\mathbf{C}_0]_{ij}\right)^2\right\} \qquad (28)$$

Expanding (28) leads to (29) given at the top of the next page.

Finally, using $E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) = [\mathbf{C}_0]_{ij}$ in (30), the value of $\hat{\lambda}$ is obtained from the following expression

$$\hat{\lambda} = \frac{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[\mathrm{Var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) - \mathrm{Cov}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right]}{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[E\left([\mathbf{T}]_{ij} - \left[\hat{\mathbf{C}}\right]_{ij}\right)^2\right]}. \qquad (31)$$

Since the variance and covariances in (31) are also estimated, they are replaced with their sample estimates to obtain the optimal shrinkage intensity i.e.

$$\hat{\lambda} = \frac{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[\widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) - \widehat{\mathrm{Cov}}\left(\mathbf{T}_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right]}{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[E\left([\mathbf{T}]_{ij} - \left[\hat{\mathbf{C}}\right]_{ij}\right)^2\right]}. \qquad (32)$$

Expressions to obtain $\widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)$ and $\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ are derived in the Appendix.

The value of $\hat{\lambda} \in [0\ 1]$ in most cases. In the unlikely event if $\hat{\lambda}$ exceeds these limits, the following bound is imposed.

$$\hat{\hat{\lambda}} = \max\left(0, \min\left(1, \hat{\lambda}\right)\right).$$

Thus the shrinkage covariance matrix $\mathbf{S}$ obtained from (25) is used in the Newton step (13) to update the GLS algorithm (12). Similarly, the variance and covariance values obtained from $\mathbf{S}$ are used in the generation of the covariance matrix $\mathbf{C}_\mathbf{v}$ for the WLLS algorithm.

V. Simulation Results

In this section, we analyze the performance of the shrinkage covariance estimate $\mathbf{S}$ and its impact on the performance of WLLS and GLS algorithms. We consider a circular deployment of N SNs with radius R m. Within the network are randomly deployed M TNs. The network deployment is shown in Fig. 1. Each SN-TN link is given a random shadowing variance σ^2_{ijkl}. The PLE value is selected to be 3 ($\alpha = 3$). The GLS algorithm is operated for η number of iterations, while to show an average performance for different number of snapshots P, for every snapshot value, the simulation is run υ times independently.

Fig. 2 shows the error between the two covariance matrix estimates, i.e. sample covariance estimate $\hat{\mathbf{C}}$ and shrinkage covariance estimate $\mathbf{S}$. For this purpose, we base our evaluation on the Frobenius norm of the difference between the true

$$L(\lambda)=\sum_{i=1}^{NM}\sum_{j=1}^{NM}\lambda^2\mathrm{Var}\left([\mathbf{T}]_{ij}\right)+(1-\lambda)^2\,\mathrm{Var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)+2\lambda(1-\lambda)\mathrm{Cov}\left([\mathbf{T}]_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)+\left[\lambda E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)+E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)-[\mathbf{C}_0]_{ij}\right]^2 \tag{29}$$

Taking derivative with respect to λ and equating to 0 yields

$$\sum_{i=1}^{NM}\sum_{j=1}^{NM}\Bigg\{2\lambda\mathrm{var}\left([\mathbf{T}]_{ij}\right)-2(1-\lambda)\,\mathrm{var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)+2(1-2\lambda)\,\mathrm{Cov}\left([\mathbf{T}]_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)+$$
$$2\left[E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)\right]\left[\lambda E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)+E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)-[\mathbf{C}_0]_{ij}\right]\Bigg\}=0 \tag{30}$$

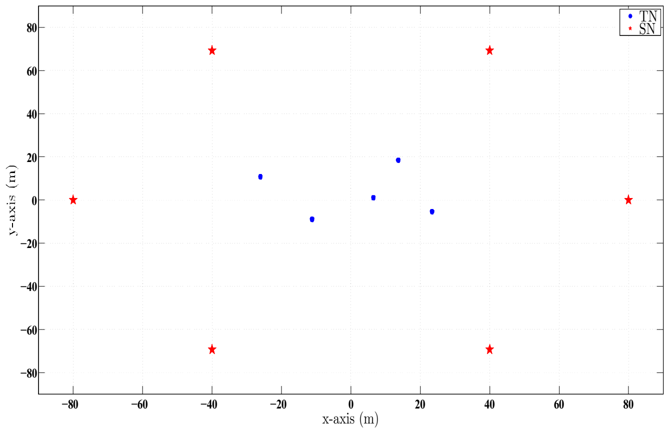

Figure 1. Network deployment

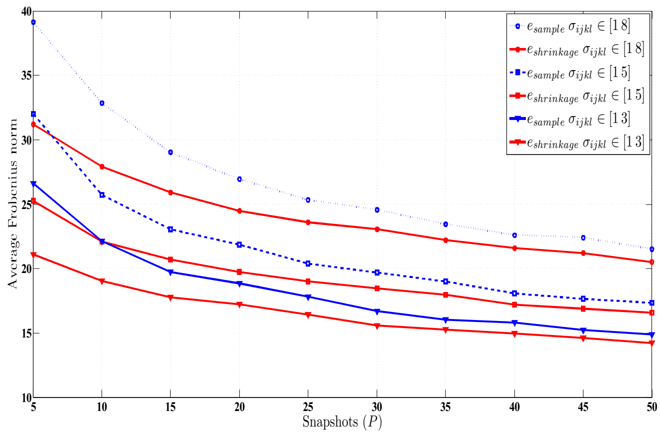

Figure 2. Error comparison of the estimated sample covariance matrix $\hat{\mathbf{C}}$ and shrinkage covariance matrix $\mathbf{S}$. $R = 80$ m, $N = 6$, $M = 5$, $\alpha = 3$, $d_c = 40\,\mathrm{m}$, $\upsilon = 100$.

and shrinkage covariance matrix i.e. $e_{shrinkage} = \|\mathbf{S} - \mathbf{C}_0\|_F^2$ and also between the true and sample covariance matrix i.e $e_{sample} = \left\|\hat{\mathbf{C}} - \mathbf{C}_0\right\|_F^2$. The comparison is done for random shadowing variance for each link between three different bounds i.e. each link is given a random σ_{ijkl}^2 value between [1 3], [1 5] and [1 8]. It is evident from Fig. 2 that $e_{shrinkage}$ is lower than than e_{sample} in all three cases. The difference in performance is more profound for smaller number of P. Also for larger σ_{ijkl}^2 bounds, larger error is shown by both estimators.

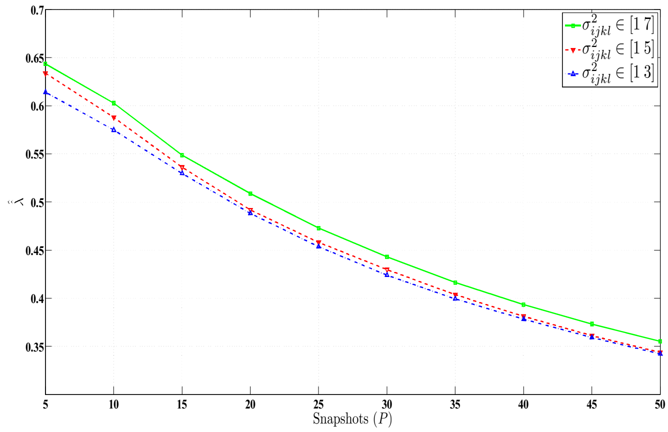

Figure 3. Average value for optimal shrinkage intensity $\hat{\hat{\lambda}}$ over $\upsilon = 100$ independent runs. $N = 6, M = 5, R = 80$ m, $\alpha = 3$, $\upsilon = 50, d_c = 40\,\mathrm{m}$,.

Fig. 3 compares the values of the optimal shrinkage intensity for different number of snapshots P and different sets of shadowing variance σ_{ijkl}^2. From Fig. 3, we deduce the following; i) the optimal shrinkage intensity is indeed between the two extremes i.e. $\hat{\hat{\lambda}} \in [0\ 1]$. ii) $\hat{\hat{\lambda}}$ decreases as the number of snapshots increases, this is expected as with more snapshots the sample estimator $\hat{\mathbf{C}}$ becomes a more accurate estimator of the covariance matrix $\mathbf{C}_0$ and hence a lower weight is given to the target shrinkage covariance $\mathbf{T}$ according to (25). iii) $\hat{\hat{\lambda}}$ is larger when the bound of shadowing variance σ_{ijkl}^2 is large, this too is supporting the shrinkage theory as with increased variance in the samples, less weight is given to the sample estimator $\hat{\mathbf{C}}$ for the same number of snapshots. The difference for the different sets of σ_{ijkl}^2 is however not significant.

Fig. 4, compares the average root mean squares error (RMSE) performance of the WLLS with sample covariance matrix WLLS_{sample}, WLLS with the shrinkage covariance matrix $\mathrm{WLLS}_{shrinkage}$ and the LLS estimator (i.e. when $\mathbf{C_v} = \mathbf{I}$) for different number of snapshots. The comparison is done for random shadowing variance between two sets. It is observed that the $\mathrm{WLLS}_{shrinkage}$ due to its relatively accurate

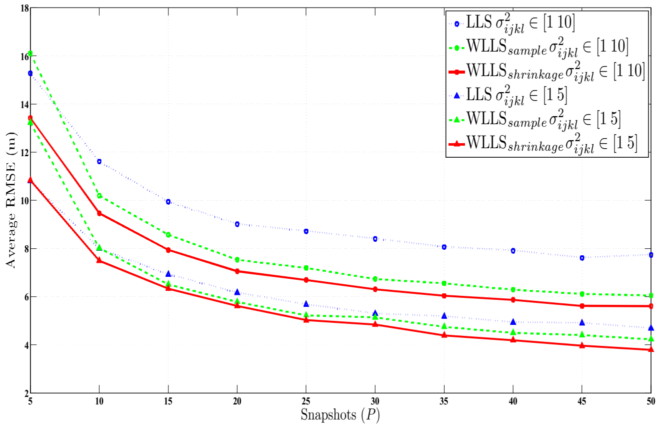

Figure 4. Average RMSE comparison between WLLS$_{shrinkage}$ and WLLS$_{sample}$. $N = 6$, $M = 5$, $R = 80$ m, $\alpha = 3$, $\upsilon = 50$, $d_c = 40$ m.

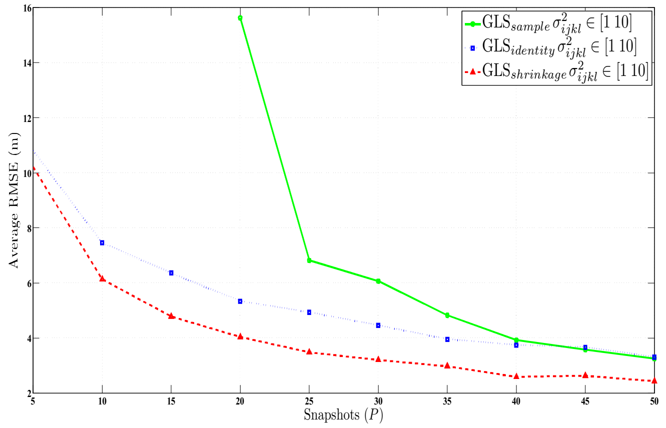

Figure 6. Average RMSE comparison between GLS$_{shrinkage}$ and GLS$_{sample}$. $N = 5$, $M = 6$, $R = 80$ m, $\alpha = 3$, $\eta = 5$, $\upsilon = 50$, $d_c = 40$ m, $\boldsymbol{\theta}^1_k = [0, 0]^T$.

estimation of the covariance matrix performs better than WLLS$_{sample}$, the performance difference is significant for a smaller number of snapshots. However even for larger values of P, WLLS$_{shrinkage}$ still performs superior to WLLS$_{sample}$.

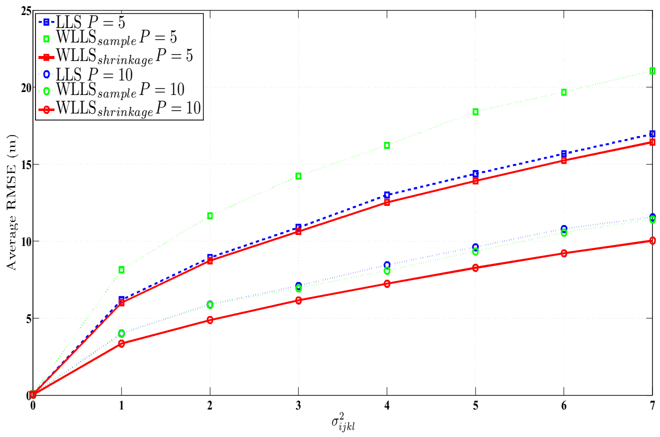

Figure 5. Average RMSE comparison between WLLS$_{shrinkage}$ and WLLS$_{sample}$. $N = 6$, $M = 5$, $R = 80$ m, $\alpha = 3$, $\upsilon = 50$, $d_c = 40$ m.

In Fig. 5, we compare the average RMSE of the location estimate of WLLS$_{sample}$, WLLS$_{shrinkage}$ and LLS corresponding to increasing value of shadowing variance σ^2_{ijkl}. In this case, σ^2_{ijkl} is same for all SN-TN links i.e. $\sigma^2_{ijkl} = \sigma^2 \,\forall\, i, j, k, l$. The plots are for two fixed snapshot values i.e $P = 5$ and $P = 10$. It is again noted that the WLLS$_{shrinkage}$ performs superior to both WLLS$_{sample}$ and LLS. It is also interesting to note that the for $P = 5$, WLLS$_{sample}$ performs even worse than the LLS, this is also evident from Fig. 4 for $P = 5$. The reason being the large error in the sample estimate $\hat{\mathbf{C}}_{\mathbf{v}}$ with a small number of P.

Fig. 6 compares the performance of the GLS algorithm with the sample covariance matrix GLS$_{sample}$, GLS with shrinkage covariance matrix GLS$_{shrinkage}$ and the GLS estimator with identity covariance matrix GLS$_{identity}$. The average RMSE of all the TNs is compared with the number of snapshots P. The center of the network is selected as the initial estimate for all TNs i.e. $\boldsymbol{\theta}^1_k = [0, 0]^T \forall\, k$. The GLS algorithm is iterated $\eta = 5$ times. Due to the lack of structure in the sample covariance matrix $\hat{\mathbf{C}}$, it is non-invertible for $P < 20$ and hence GLS$_{sample}$ produces no result for these values of P. $P = 20$ is by no means a minimum value for the invertibility of $\hat{\mathbf{C}}$, in fact the minimum number of snapshots will increase with the increase in the number of TNs. For $P > 20$ the GLS$_{sample}$ shows unacceptable error and performs worse than GLS$_{identity}$. On the other hand, the shrinkage covariance matrix $\mathbf{S}$ is both invertible and consists of a small error variance, consequently GLS$_{shrinkage}$ performs better than GLS$_{sample}$ and GLS$_{identity}$.

VI. Conclusions

In this paper, we focus on the optimized localization of multiple TNs with correlated measurements. We have shown that the correlation between readings from different TNs at the SNs can improve the estimation accuracy. Two RSS based localization algorithms were investigated. A closed form expression is derived for the covariance matrix for multiple TN WLLS algorithm. Furthermore, the covariance matrix is estimated with a limited number of snapshots and the shrinkage estimator. The shrinkage estimator is also used to estimate the covariance matrix for the GLS algorithm. To evaluate the performance of the proposed methods, correlation between multiple TN readings at the same SN is based on the the Gudmundson model. It is shown via simulation results that the proposed WLLS and GLS algorithms with the shrinkage covariance matrix perform superior to the same algorithms using the sample covariance matrix. Especially for a small number of snapshots the sample covariance matrix consists of a large error (and is non invertible in case of GLS). This results in degraded performance and in the worst case, the performance is inferior to the case where the identity matrix is used for the covariance matrix. The shrinkage covariance matrix on the other hand, guarantees a relatively small estimation error and it is always invertible. This results in superior location estimation performance in both WLLS and GLS.

ACKNOWLEDGMENT

We acknowledge the support from the UK Engineering and Physical Sciences Research Council (EPSRC) for the support via the Bayesian Tracking and Reasoning over Time (BTaRoT) grant EP/K021516/1.

APPENDIX

Following Schafer and Strimmer [3] and Kwan [24] whose study focussed on gene classification and security returns respectively, we derive optimal shrinkage intensity in the context of multitarget localization. We define the following: For P snapshots of path-loss measurements z_i, the sample mean is given by $\widehat{z}_i = P^{-1}\sum_{p=1}^{P} z_{ip}$. Let

$$v_{ijp} = (z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j) \tag{33}$$

for $i, j = 1, ..., NM$ and $p = 1, ..., P$. Also let $\widehat{v}_{ij} = P^{-1}\sum_{p=1}^{P} v_{ijp}$. Then the sample covariance is given by

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{P}{P-1}\widehat{v}_{ij} = \frac{1}{P-1}\sum_{p=1}^{P}(z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j). \tag{34}$$

For ease of understanding, v_{ijp} should be viewed as a random variable. The unbiased variance of individual elements of $\hat{\mathbf{C}}$ is given by

$$\widehat{\mathrm{Var}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}\right\} = \frac{P^2}{(P-1)^2}\mathrm{Var}\left(\widehat{v}_{ij}\right). \tag{35}$$

Using the identity to find the variance of the mean of a random variable, we have

$$\widehat{\mathrm{Var}}\left(\widehat{v}_{ij}\right) = \frac{1}{P}\mathrm{var}\left(v_{ij}\right) \tag{36}$$

or

$$\widehat{\mathrm{Var}}\left(\widehat{v}_{ij}\right) = \frac{1}{P}\left[\frac{1}{P-1}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)^2\right]. \tag{37}$$

Thus the variance of the sample covariance matrix is given by

$$\widehat{\mathrm{Var}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}\right\} = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)^2. \tag{38}$$

On the other hand, the covariance elements are given as

$$\widehat{\mathrm{Cov}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{kl}\right\} = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)\left(v_{klp} - \widehat{v}_{kl}\right). \tag{39}$$

An expression for $\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ is slightly more complicated, here we use the so called 'delta method'. For this purpose, we consider the individual elements of $\hat{\mathbf{C}}$ as variables and expand $[\mathbf{T}]_{ij} = \hat{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}$ via Taylor series around the corresponding point estimates $\left[\bar{\mathbf{C}}\right]_{ii}$, $\left[\bar{\mathbf{C}}\right]_{jj}$ and $\left[\bar{\mathbf{C}}\right]_{ij}$ provided by the samples. We have

$$[\mathbf{T}]_{ij} = \bar{\rho}\sqrt{\left[\bar{\mathbf{C}}\right]_{ii}\left[\bar{\mathbf{C}}\right]_{jj}} + \frac{\bar{\rho}}{2}\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{jj}}{\left[\bar{\mathbf{C}}\right]_{ii}}}\left([\mathbf{C}]_{ii} - \left[\bar{\mathbf{C}}\right]_{ii}\right) + \frac{\bar{\rho}}{2}\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{ii}}{\left[\bar{\mathbf{C}}\right]_{jj}}}\left([\mathbf{C}]_{jj} - \left[\bar{\mathbf{C}}\right]_{jj}\right), \tag{40}$$

where $\bar{\rho}$ is obtained using (26) for $\bar{\mathbf{C}}_{ii}$ $\bar{\mathbf{C}}_{jj}$ and $\bar{\mathbf{C}}_{ij}$. Then from the definition of the covariance matrix, we have

$$\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = E\left[\left([\mathbf{T}]_{ij} - E\left([\mathbf{T}]_{ij}\right)\right)\left([\mathbf{C}]_{ij} - E\left([\mathbf{C}]_{ij}\right)\right)\right] \tag{41}$$

using (40) in (41) leads to

$$\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{\bar{\rho}}{2}\left(\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{jj}}{\left[\bar{\mathbf{C}}\right]_{ii}}}\mathrm{Cov}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) + \sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{ii}}{\left[\bar{\mathbf{C}}\right]_{jj}}}\mathrm{Cov}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right).$$

Finally using (33) and (39), it is straightforward to define $\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ and $\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ as

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left\{\left[(z_{ip} - \widehat{z}_i)^2 - \widehat{v}_{ii}\right]\left[(z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j) - \widehat{v}_{ij}\right]\right\} \tag{42}$$

and similarly

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left\{\left[(z_{jp} - \widehat{z}_j)^2 - \widehat{v}_{jj}\right]\left[(z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j) - \widehat{v}_{ij}\right]\right\}, \tag{43}$$

which completes the derivation.

REFERENCES

[1] O. Ledoit, and M. Wolf, "Honey, I Shrunk the Sample Covariance Matrix", *Journal of Portfolio Management*, 110-119, 2004.

[2] O. Ledoit, and M. Wolf, "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection", *Journal of Empirical Finance*, 10, 603-621, 2003.

[3] J. Schafer, K. Strimmer, "A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics," *Statistical Applications in Genetics and Molecular Biology*, Vol. 4, no. 1, 2005.

[4] N. Patwari, J. N. Ash, S. Kyperountas, A. O. H. III, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[5] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surveys and Tutorials*, vol. 11, no. 3, pp. 107–124, Aug. 2009.

[6] B.D. Van Veen and K.M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[7] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample ML techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S.S. Haykin, J. Litva, and T. Shepherd, Eds. New York: Springer- Verlag, 1993, pp. 99–151.

[8] N. Patwari, A. O. Hero, III, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2137-2148, Aug. 2003.

[9] N. Salman, M. Ghogho, and A. H. Kemp, “On the joint estimation of the RSS-based location and path-loss exponent,” *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 34-37, Feb. 2012.

[10] N. Salman, A. H. Kemp and M. Ghogho, “Low Complexity Joint Estimation of Location and Path-Loss Exponent”, *IEEE Wireless Commun Lett.,* vol. 1, no. 4, pp. 364–367, Aug. 2012.

[11] A. Bel, J.L, Vicario, G. Seco-Granados, Localization Algorithm with On-line Path Loss Estimation and Node Selection, Sensors, 11, pp. 6905-6925, Jul. 2011.

[12] R. Ouyang, A.-S. Wong, and C.-T. Lea, “Received signal strength-based wireless localization via semidefinite programming: Noncooperative and cooperative schemes,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1307– 1318, Mar. 2010.

[13] L. Mihaylova, D. Angelova, D. R. Bull, N. Canagarajah, "Localization of Mobile Nodes in Wireless Networks with Correlated in Time Measurement Noise," I*EEE Transactions on Mobile Computing*, vol.10, no.1, pp.44,53, Jan. 2011.

[14] M. Gudmundson, “Correlation model for shadow fading in mobile radio systems,” IEE Electronic Letters, 27(23):2145-2146, 7 Nov, 1991.

[15] N. Patwari and P. Agrawal, “Effects of correlated shadowing: Connectivity, localization, and RF tomography,” in *Proc. IEEE/ACM Int. Conf. Information Processing in Sensor Networks (IPSN)*, 2008, pp. 82–93.

[16] H. Hashemi, “The indoor radio propagation channel,” *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, July 1993.

[17] T.S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[18] K. Pahlavan and A. Levesque, *Wireless Information Networks*. New York: John Wiley & Sons, Inc., 1995.

[19] G. A. F. Seber, C. J. Wild, *Nonlinear regression*. John Wiley & Sons Inc, 2003.

[20] J. J. Caffery, “A new approach to the geometry of TOA location,” in Proc. *IEEE Veh. Technol. Conf. (VTC)*, vol. 4, Boston, MA, Sep. 2000, pp. 1943–1949.

[21] I. Guvenc, S. Gezici, Z. Sahinoglu, “Fundamental limits and improved algorithms for linear least-squares wireless position estimation,”*Wirel. Commun. Mob. Comput.* (2010).

[22] N. Salman, M. Ghogho, A. H. Kemp, “Optimized Low Complexity Sensor Node Positioning in Wireless Sensor Networks,” *IEEE Sensors Journal*, vol.14, no.1, pp.39-46, Jan. 2014.

[23] D. Baum, J. Hansen, and J. Salo, “An interim channel model for beyond-3G systems: extending the 3GPP spatial channel model (SCM),” in *Proc. IEEE VTC*, May/Jun. 2005, vol. 5, pp. 3132–3136.

[24] C. C. Y. Kwan, “Estimation error in the average correlation of security returns and shrinkage estimation of covariance and correlation matrices,” *Elsevier Finance Research Letters*, 5, 236-244, 2008.

A Shrinkage Approach to Multiple Target Localization with Correlated Measurements

Naveed Salman*, *Student Member, IEEE,* Lyudmila Mihaylova*, *Senior Member, IEEE,* and A. H. Kemp**, *Member, IEEE*

Abstract

In this paper, we deal with the accurate estimation of the covariance matrix for the optimization of multiple target node (TN) localization using correlated received signal strength (RSS) measurements. Two location schemes i.e. the iterative generalized least squares (GLS) and the low complexity weighted linear least squares (WLLS) methods are investigated. For many applications the estimated covariance matrix needs to be positive definite and hence invertible. For an insufficient number of data points, the sample covariance matrix suffers from two drawbacks. Firstly, although it is unbiased, it consists of a large estimation error. Secondly, it is not positive definite. A shrinkage technique to estimate the covariance matrix has been proposed in the fields of finance and life sciences. In this paper, we introduce the shrinkage covariance matrix concept in the area of multiple TN localization in wireless networks with correlated measurements. An analytical expression of the multiple TN covariance matrix is derived for the WLLS method and the estimation is done via the shrinkage technique. Similarly, the shrinkage technique is employed for the covariance matrix estimation for the GLS algorithm. For a limited number of measurements, the shrinkage covariance technique improves the performance of both WLLS and GLS considerably.

Index Terms

Localization, Received signal strength (RSS), Shrinkage estimation.

I. INTRODUCTION

For a set of random variables, the covariance matrix is a symmetric matrix, whose diagonal elements represent the individual variances of the random variables while the off-diagonal elements are the corresponding cross covariances. When the variance of all random variables is set to unity, the covariance matrix is then known as the correlation matrix. The covariance matrices have applications in various fields. In finance, portfolio analysis is done via estimation of the covariance matrix of stock returns [1],

*Department of Automatic Control and Systems Engineering, University of Sheffield, U.K. (email: {n.salman, l.s.mihaylova}@sheffield.ac.uk)

**School of Electronic and Electrical Engineering, University of Leeds, U.K. (email: a.h.kemp@leeds.ac.uk)

[2]. In life sciences it is used in genes classification by finding the pairwise correlation between genes [3]. In most cases, the elements of the covariance matrix are unknown and need to be estimated. Due to its ease of use, the sample estimator is commonly used for covariance matrix estimation. Although the sample estimator is unbiased, it faces serious issues when the number of samples is equal or smaller than the number of variables. If the samples are equal or less than the number of variables, the sample covariance matrix will contain a large estimation error. Furthermore, the sample covariance matrix will not be positive definite, making it impractical to use. To elaborate on the previously given examples, for portfolio analysis, if the number of stocks is larger than the number of historical (e.g. monthly) stock returns then the corresponding sample covariance matrix will contain large errors and will always be singular. Similar problems are faced in gene classification research, if the number of genes under considerations exceeds the number of experiments performed.

The idea of shrinkage estimation for large covariance matrices was introduced rather recently in finance applications [2]. As mentioned before, the sample estimator is unbiased but poses large variance, the shrinkage idea is to introduce some structure to the sample covariance matrix. A target covariance matrix is chosen (which has a smaller number of variables) as a candidate for covariance matrix estimate. Due to the small number of parameters, the target covariance matrix will have a relatively small error variance although it will be biased. The idea is to choose an optimally compromised covariance matrix instead of the two extremes i.e. the large variance but unbiased sample covariance matrix and low variance but biased target covariance matrix.

In this paper, we apply the shrinkage theory for covariance matrix estimation to multiple node localization in wireless networks. Node localization in wireless networks is needed in many application including situation awareness, cattle/wild life monitoring and logistics [4]. A number of techniques have been developed for accurate positioning of wireless nodes. These may include the time of arrival (ToA) [5] and angle of arrival (AoA) methods [6]. Both ToA and AoA provide an accurate distance and hence location estimation. However, both techniques require additional hardware on board the target nodes (TNs). A prerequisite of ToA method is the availability of highly accurate and synchronized clocks on the TNs, while AoA requires an array of antennas [6], [7]. A low cost and simplistic technique to localize wireless nodes is the received signal strength (RSS) technique [8]. Considerable amount of research has been done to optimize the performance of RSS localization. Joint estimation of the path-loss exponent (PLE) and location is discussed in [9], [10], [11]. Cooperative RSS localization is studied in [12]. Tracking using

RSS technique is investigated in [13]. Most of these studies assume individual links between TNs and sensor node (SN) to be independent of each other. This is an oversimplification as readings at the SN from TNs that are close to each other in a network introduces an element of spatial correlation. In this study we simulate our data using the widely accepted Gudmundson correlation model presented in [14]. The Gudmundson's model suggests that the correlation between two nodes is an exponentially decaying function of the distance between them. The authors of [15] show that if these correlations are taken into account they could enhance the location estimation performance. This proposition was done by deriving the Cramer-Rao bound for spatial correlation between nodes, however an algorithm capitalizing on such correlation was not presented.

In this paper, we optimize the performance of two multiple TN RSS location algorithms. Location estimation of TNs is a non-linear estimation problem. Due to the non-linear nature, the location estimation is generally performed using an iterative algorithm with a close initial estimate and Newton step update. However, location coordinates can also be estimated using a linear least squares (LLS) approach. In this paper we first develop an iterative generalized least squares (GLS) technique for multiple TN localization. Second, we also formulate a weighted LLS (WLLS) solution to the multiple TN location problem. A closed form expression is derived for the covariance matrix of multiple TN readings at the SNs. The covariance matrix expression depends on the variance and covariance values of these readings. In practical scenarios and in large networks these values are unknown and the number of snapshots is not large enough to provide an accurate sample estimate due to limited resources and packet lose. To counter this problem, we apply the shrinkage technique to estimate the variance and covariance elements required in the expression for the covariance matrix in the WLLS method. The shrinkage technique is also applied to estimate the covariance matrix for the GLS technique. Due to the lack of structure and for a limited number of snapshots the sample covariance matrix is non invertible rendering it impractical to use in the GLS method. The shrinkage covariance estimator on the other hand introduces a structure to the covariance matrix and hence it is invertible even for limited number of snapshots. It is shown via simulation that the shrinkage covariance matrix compared to the sample covariance matrix has a smaller estimation error. Consequently, the shrinkage covariance matrix, when used in the GLS and WLLS algorithm, results in superior performance. Especially in the case of GLS case, where the sample covariance estimator is non invertible for a limited number of snapshots, this problem is resolved with the shrinkage covariance matrix.

The rest of the paper is organized as follows. Section II introduces the signal model and the multiple TN location problem. Section III, develops the GLS and the WLLS estimator. In section IV we briefly discuss the shrinkage theory and covariance shrinkage estimator. Section V presents the simulation results which are followed by the conclusions. The derivation of the shrinkage estimator is provided in the Appendix.

II. System model

For future use, we define the following notations.

$\mathcal{R}^n$ is the set of n dimensional real numbers, $(.)^T$ is the transpose operator, $E(.)$ refers to the expectation operator, $[\mathbf{M}]_{ij}$ refers to the element at the i^{th} row and j^{th} column of matrix $\mathbf{M}$, $\mathbf{I}_N$ represents the N dimensional identity matrix. $\mathcal{N}(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2. $\|.\|_F^2$ represents the Frobenius norm.

A two dimensional (2-D) network is considered, consisting of N SNs with known locations $\boldsymbol{\phi}_{i,j} = [x_{i,j}, y_{i,j}]^T (\boldsymbol{\phi}_{i,j} \in \mathcal{R}^2)$ for $i, j = 1, ..., N$. The network also consists of M TNs which have unknown coordinates $\boldsymbol{\theta}_{k,l} = [x_{k.l}, y_{k,l}]^T$ $(\boldsymbol{\theta}_{k,l} \in \mathcal{R}^2)$ for $k, l = 1, \ldots, M$ that are to be estimated. The received power at the SNs due to random shadowing is log-normally distributed. This model is based on empirical results obtained in [16], [17]. It is assumed that the TNs transmit during preassigned epochs to avoid interference between different TN transmissions. Thus the distance d_{ik} between the k^{th} TN and the i^{th} AN is related to the path-loss at the i^{th} AN, $\mathscr{L}_{ik}$, and the PLE, α, as [18]

$$\mathscr{L}_{ik} = \mathscr{L}_0 + 10\alpha \log_{10} d_{ik} + w_{ik}, \tag{1}$$

where $\mathscr{L}_0$ is the path-loss at the reference distance d_0 ($d_0 < d_{ik}$, and is normally taken as 1 m) and w_{ik} is a zero-mean Gaussian random variable representing the multipath log-normal shadowing effect, i.e. $w_{ik} \sim (\mathcal{N}(0, \sigma_{ik}^2))$. It is assumed that α is known via prior channel modeling or accurate estimation. The path-loss is calculated as

$$\mathscr{L}_{ik} = 10 \log_{10} P_k - 10 \log_{10} P_{ik}, \tag{2}$$

where P_k is the transmit power at the k^{th} TN and P_{ik} is the received power at the i^{th} AN. The distance d_{ik} is given by

$$d_{ik} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}. \tag{3}$$

The observed path-loss (in dB) from d_0 to d_{ik}, $z_{ik} = \mathscr{L}_{ik} - \mathscr{L}_0$, can be expressed as

$$z_{ik} = f_i(\boldsymbol{\theta}_k) + w_{ik}, \tag{4}$$

where $f_i(\boldsymbol{\theta}_k) = \gamma\alpha \ln d_{ik}$ and $\gamma = \frac{10}{\ln 10}$. (4) can be written in a matrix form as

$$\mathbf{Z} = \mathbf{F}(\boldsymbol{\theta}) + \mathbf{W}, \tag{5}$$

where the k^{th} column of $\mathbf{Z}$ and $\mathbf{W}$ represents the readings and associated noise elements at the SNs from the k^{th} TN respectively and $[\mathbf{F}(\boldsymbol{\theta})]_{ik} = f_i(\boldsymbol{\theta}_k)$. For convenience (5) can also be written in a 'roll out' form i.e. $\mathbf{z} = \mathrm{vec}(\mathbf{Z})$, then for the k^{th} target, we have

$$\begin{bmatrix} z_{1k} \\ z_{2k} \\ \vdots \\ z_{Nk} \end{bmatrix} = \begin{bmatrix} f_1(\boldsymbol{\theta}_k) \\ f_2(\boldsymbol{\theta}_k) \\ \vdots \\ f_N(\boldsymbol{\theta}_k) \end{bmatrix} + \begin{bmatrix} w_{1k} \\ w_{2k} \\ \vdots \\ w_{Nk} \end{bmatrix}$$

or

$$\mathbf{z}_k = \mathbf{f}_k + \mathbf{w}_k. \tag{6}$$

Then we have $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{C}_{kk})$ where $\mathbf{C}_{kk} = \sigma_{ik}^2 \mathbf{I}_N$ and $E\left(\mathbf{w}_k (\mathbf{w}_l)^T\right) = \mathbf{C}_{kl}$. Here $\mathbf{C}_{kl}$ is the $N \times N$ covariance matrix between RSS readings at the SNs from the k^{th} and l^{th} TN and its elements are given by $\mathbf{I}_N \rho_{kl} \sigma_{ik} \sigma_{il}$. Here ρ_{kl} is the spatial correlation between the k^{th} and l^{th} TN. Its value depends on the relative distance between the TNs and can be modeled according to Gudmundson [14] as follows

$$\rho_{kl} = \exp\left(-\frac{d_{kl}}{d_c}\right), \tag{7}$$

where d_c is the 'decorrelation distance'. Field measurements in [23] suggest values for d_c for different environments. Eq. (6) when written for all TNs is given by

$$\mathbf{z} = \mathrm{vec}(\mathbf{Z}) = \left((\mathbf{z}_1)^T, (\mathbf{z}_2)^T, ..., (\mathbf{z}_M)^T\right)^T \tag{8}$$

or

$$\mathbf{z} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{w}, \tag{9}$$

where $\mathbf{w} \sim \mathcal{N}(0, \mathbf{C})$, where $\mathbf{C}$ encompasses correlations between all TNs;

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1M} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{M1} & \mathbf{C}_{M2} & \cdots & \mathbf{C}_{MM} \end{bmatrix}.$$

Two location algorithms to solve (9) are discussed in the next section.

III. Localization algorithms for multiple TNs

A. Generalized least squares algorithm

It is clear that (9) presents a non-linear estimation problem and a close form expression to solve (9) is not readily available. A solution to (9) is obtained by minimizing the following cost function [19]

$$\epsilon(\boldsymbol{\theta}) = (\mathbf{z} - \mathbf{f}(\boldsymbol{\theta}))^T \mathbf{C}^{-1} (\mathbf{z} - \mathbf{f}(\boldsymbol{\theta})), \tag{10}$$

where (10) can be minimized by first linearizing $\mathbf{f}(\boldsymbol{\theta})$ by the first order Taylor series expansion to a close initial estimate $\mathbf{f}(\boldsymbol{\theta}^a)$ i.e.

$$\mathbf{f}(\boldsymbol{\theta}) = \mathbf{f}(\boldsymbol{\theta}^a) + \mathbf{F}(\boldsymbol{\theta}^a)(\boldsymbol{\theta} - \boldsymbol{\theta}^a), \tag{11}$$

where

$$\mathbf{F}(\boldsymbol{\theta}^a) = \mathrm{diag}\left[\mathbf{F}(\boldsymbol{\theta}_1^a), \mathbf{F}(\boldsymbol{\theta}_2^a), ..., \mathbf{F}(\boldsymbol{\theta}_M^a)\right]$$

and $\mathbf{F}(\boldsymbol{\theta}_k^a)$ is the Jacobian matrix of the k^{th} TN and its elements are given by

$$\mathbf{F}(\boldsymbol{\theta}_k^a) = \begin{bmatrix} \left.\frac{\partial f_1(\boldsymbol{\theta}_k)}{\partial x_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} & \left.\frac{\partial f_1(\boldsymbol{\theta}_k)}{\partial y_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} \\ \left.\frac{\partial f_2(\boldsymbol{\theta}_k)}{\partial x_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} & \left.\frac{\partial f_2(\boldsymbol{\theta}_k)}{\partial y_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} \\ \vdots & \vdots \\ \left.\frac{\partial f_N(\boldsymbol{\theta}_k)}{\partial x_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} & \left.\frac{\partial f_N(\boldsymbol{\theta}_k)}{\partial y_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} \end{bmatrix}$$

for $\left.\frac{\partial f_i(\boldsymbol{\theta}_k)}{\partial x_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} = {}^{\gamma\alpha\left(x_k^a - x_i\right)}\!/_{d_{ik}^2}$ and $\left.\frac{\partial f_i(\boldsymbol{\theta}_k)}{\partial y_k}\right|_{\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^a} = {}^{\gamma\alpha\left(y_k^a - y_i\right)}\!/_{d_{ik}^2}$.

Using (11) in (10) and minimizing with respect to $\boldsymbol{\theta}$, so that the next estimation is given by

$$\boldsymbol{\theta}^{a+1} = \boldsymbol{\theta}^{a} + \varrho^{a} \tag{12}$$

where

$$\varrho^{a} = \left(\mathbf{F}\left(\boldsymbol{\theta}^{a}\right)^{T}\mathbf{C}^{-1}\mathbf{F}\left(\boldsymbol{\theta}^{a}\right)\right)^{-1}\mathbf{F}\left(\boldsymbol{\theta}^{a}\right)^{T}\mathbf{C}^{-1}\left(\mathbf{z}-\mathbf{f}\left(\boldsymbol{\theta}^{a}\right)\right). \tag{13}$$

Thus given an initial estimate $\boldsymbol{\theta}^{1}$, the generalized least squares (GLS) converges to the minima. The algorithm is stopped after a fixed number of iterations or when the value $\left|\epsilon\left(\boldsymbol{\theta}^{a+1}\right)-\epsilon\left(\boldsymbol{\theta}^{a}\right)\right|$ becomes smaller than a certain threshold. To avoid the complexity of the iterative procedure and a close initial estimate condition at the expanse of accuracy, a WLLS solution is presented in the next subsection.

B. Weighted linear least squares algorithm

A slight manipulation of (9) renders it to be linear. This technique was first proposed for ToA distance estimates in [20] and analyzed in [21]. For a single target RSS measurements the linearized model is discussed in [22]. Here we extend this method for the multiple TN case as follows. From (9) it can be readily shown that

$$E\left(\frac{1}{\beta_{ik}}\exp\left(\frac{2z_{ik}}{\gamma\alpha}\right)\right) = \hat{d}_{ik}^{2}, \tag{14}$$

where $\beta_{ik} = \exp\left(\frac{2\sigma_{ik}^{2}}{(\gamma\alpha)^{2}}\right)$. Similarly choosing a reference AN, it can be shown

$$E\left(\frac{1}{\beta_{rk}}\exp\left(\frac{2z_{rk}}{\gamma\alpha}\right)\right) = \hat{d}_{rk}^{2}, \tag{15}$$

where $\beta_{rk} = \exp\left(\frac{2\sigma_{rk}^{2}}{(\gamma\alpha)^{2}}\right)$. For linearization, the square of each distance equation is subtracted from the square of a reference distance equation. This results in a linear system which is represented in matrix form as[1]

$$\mathbf{b} = \mathbf{A}\boldsymbol{\theta} + \mathbf{v}, \tag{16}$$

where $\mathbf{b} = \left[\mathbf{b}_{1}, ..., \mathbf{b}_{M}\right]^{T}$ is the observation vector at the SNs from the TNs. The linearized observation from the k^{th} TN at the ANs is thus given as

[1]The inclusion of the constants β_{ik}, β_{rk} in (14) and (15) is essential for the LLS solution to be unbiased.

$$\mathbf{b}_k = \begin{bmatrix} \delta_{rk} - \delta_{1k} - \kappa_{rk} + \kappa_1 \\ \delta_{rk} - \delta_{2k} - \kappa_{rk} + \kappa_2 \\ \vdots \\ \delta_{rk} - \delta_{N-1k} - \kappa_{rk} + \kappa_{N-1} \end{bmatrix}$$

for $\delta_{rk} = \frac{1}{\beta_{rk}} \exp\left(\frac{2z_{rk}}{\gamma\alpha}\right)$ and $\delta_{ik} = \frac{1}{\beta_{ik}} \exp\left(\frac{2z_{ik}}{\gamma\alpha}\right)$. While

$$\kappa_{rk} = x_{rk}^2 + y_{rk}^2 \text{ and } \kappa_i = x_i^2 + y_i^2$$

for $i \neq r,\ i = 1, ..., N-1$. Also (x_{rk}, y_{rk}) are the coordinates of the reference SN for the k^{th} TN. $\mathbf{A}$ is the $M(N-1) \times 2M$ data matrix for all TNs and is given by

$$\mathbf{A} = \operatorname{diag}\left(\mathbf{A}^1, \mathbf{A}^2 \cdots \mathbf{A}^M\right).$$

The diagonal elements of $\mathbf{A}$ are themselves data matrices for the individual TNs. Thus for the k^{th} TN, the data matrix is given by

$$\mathbf{A}^k = 2 \begin{bmatrix} x_1 - x_{rk} & y_1 - y_{rk} \\ x_2 - x_{rk} & y_2 - y_{rk} \\ \vdots & \vdots \\ x_{N-1} - x_{rk} & y_{N-1} - y_{rk} \end{bmatrix}.$$

$\mathbf{v}$ is the noise vector which has zero mean and a $M(N-1) \times M(N-1)$ covariance matrix $\mathbf{C_v}$ given by

$$\mathbf{C_v} = \begin{bmatrix} \mathbf{C_v}_{11} & \mathbf{C_v}_{12} & \cdots & \mathbf{C_v}_{1M} \\ \mathbf{C_v}_{21} & \mathbf{C_v}_{22} & \cdots & \mathbf{C_v}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C_v}_{M1} & \mathbf{C_v}_{M2} & \cdots & \mathbf{C_v}_{MM} \end{bmatrix}. \tag{17}$$

Expressions for the individual elements of $\mathbf{C_v}$ are given as follows.

The diagonal elements of $\mathbf{C_v}_{kk}$ are given by

$$
\begin{aligned}
\left[\mathbf{C}_{\mathbf{v}kk}\right]_{ii} &= E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)^2\right] \\
&= d_{ik}^4 \exp\left(\frac{4\sigma_{ik}^2}{(\gamma\alpha)^2}\right) - d_{ik}^4 + d_{rk}^4 \exp\left(\frac{4\sigma_{rk}^2}{(\gamma\alpha)^2}\right) - d_{rk}^4
\end{aligned} \tag{18}
$$

and the off-diagonal elements of $\mathbf{C}_{\mathbf{v}kk}$ are given by[2]

$$
\begin{aligned}
\left[\mathbf{C}_{\mathbf{v}kk}\right]_{ij} &= E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rk} - \delta_{jk} - d_{rk}^2 + d_{jk}^2\right)\right] \\
&= \left[d_{rk}^4 \exp\left(\frac{4\sigma_{rk}^2}{(\gamma\alpha)^2}\right) - d_{rk}^4\right].
\end{aligned} \tag{19}
$$

The diagonal elements of $\mathbf{C}_{\mathbf{v}kl}$ are given as

$$
\begin{aligned}
\left[\mathbf{C}_{\mathbf{v}kl}\right]_{ii} &= E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rl} - \delta_{il} - d_{rl}^2 + d_{il}^2\right)\right] \\
&= a1 + a2 - a3 - a4
\end{aligned} \tag{20}
$$

where

$$
a1 = \frac{1}{\beta_{rk}}\frac{1}{\beta_{rl}} d_{rk}^2 d_{rl}^2 \exp\left(\frac{2\left(\sigma_{rk}^2 + \sigma_{rl}^2 + 2\rho_{kl}\sigma_{rk}\sigma_{rl}\right)}{(\gamma\alpha)^2}\right)
$$

$$
a2 = \frac{1}{\beta_{ik}}\frac{1}{\beta_{il}} d_{ik}^2 d_{il}^2 \exp\left(\frac{2\left(\sigma_{ik}^2 + \sigma_{il}^2 + 2\rho_{kl}\sigma_{ik}\sigma_{il}\right)}{(\gamma\alpha)^2}\right)
$$

$$
a3 = d_{rk}^2 d_{rl}^2 \quad \text{and} \quad a4 = d_{ik}^2 d_{il}^2.
$$

Finally, the off diagonal elements of $\mathbf{C}_{\mathbf{v}kl}$ are given by

[2]Here the correlation between the measurements of the same TN at different ANs is not considered. This assumption is based upon two arguments. First, for efficient operation of the localization algorithm, the SNs are placed far apart hence retaining little correlation. Second, even if the SNs are placed close to each other, distance and hence correlation between them is known and can easily be incorporated in (19).

$$
\begin{aligned}
[\mathbf{C_v}_{kl}]_{ij} &= E\left[\left(\delta_{rk} - \delta_{ik} - d_{rk}^2 + d_{ik}^2\right)\left(\delta_{rl} - \delta_{jl} - d_{rl}^2 + d_{jl}^2\right)\right] \\
&= b1 - b2
\end{aligned}
\tag{21}
$$

where

$$b1 = \frac{1}{\beta_{rk}}\frac{1}{\beta_{rl}} d_{rk}^2 d_{rl}^2 \exp\left(\frac{2\left(\sigma_{rk}^2 + \sigma_{rl}^2 + 2\rho_{kl}\sigma_{rk}\sigma_{rl}\right)}{\left(\gamma\alpha\right)^2}\right)$$

and

$$b2 = d_{rk}^2 d_{rl}^2.$$

The WLLS solution is obtained by minimizing the cost function

$$\varepsilon_{WLLS}\left(\boldsymbol{\theta}\right) = \left(\mathbf{b} - \mathbf{A}\boldsymbol{\theta}\right)^T \mathbf{C_v}^{-1}\left(\mathbf{b} - \mathbf{A}\boldsymbol{\theta}\right), \tag{22}$$

The WLLS estimate is obtained as follows

$$\hat{\boldsymbol{\theta}}_{WLLS} = \mathbf{A}^{\ddagger}\mathbf{b}^{\ddagger}, \tag{23}$$

where $\mathbf{A}^{\ddagger} = \left[\mathbf{A}^T\mathbf{C_v}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^T$ and $\mathbf{b}^{\ddagger} = \mathbf{C_v}^{-1}\mathbf{b}$. Since the actual distances are not available, their estimated values are used in (23). The variance and correlation elements of $\mathbf{C_v}$ can be estimated accurately by the sample estimation of $\mathbf{C_v}$ for a large number of readings or snapshots. For a limited number of snapshots, a shrinkage estimator is proposed in section IV. The WLLS technique presented is relative low in complexity and does not require an initial estimate; however its performance is sub-optimal. This is because information from all the SNs is not utilized as the measurement from reference SN is used to linearize the system.

IV. Practical issues and shrinkage estimation

In practical scenarios the covariance matrix $\mathbf{C}$ is unknown and it needs to be estimated. For seamless operation of the iterative algorithm (12), $\mathbf{C}$ has to be positive definite and hence invertible. For a sample covariance matrix to be positive definite, the number of sample points need to be more than the number of variables. For multiple TN this poses a serious problem for the invertibility of $\mathbf{C}$ as the number readings need to exceed the number of unknowns which can not always be guaranteed in resource constraint networks. Furthermore, a sample covariance matrix estimate with limited number of sample points consists of large errors. Erroneous estimates of $\mathbf{C}$ could lead to performance degradation of the GLS algorithm. Similarly, the variance elements that are required to generate the covariance matrix $\mathbf{C_v}$ for the WLLS procedure need to be estimated. Here again, the sample estimator is not efficient for a limited number of snapshots. Thus, alternative techniques to estimate $\mathbf{C}$ and $\mathbf{C_v}$ needs investigation to insure minimum error and positive definiteness. In this paper we investigate the application of shrinkage estimation technique for covariance matrix estimation. The shrinkage estimator is introduced in the following subsection.

A. Shrinkage estimation of the covariance matrix

1) Shrinkage theory: In the context of multiple TN localization, we briefly explain the theory behind the shrinkage estimation as proposed by Ledoit and Wolf [2] for portfolio analysis. For a large dimension covariance matrix $\mathbf{C}$ of size $NM \times NM$, let $\hat{\mathbf{C}}$ be the sample estimator of $\mathbf{C}$, then its elements are given by

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{1}{P-1}\sum_{p=1}^{P}\left(z_{ip} - \widehat{z}_i\right)\left(z_{jp} - \widehat{z}_j\right), \tag{24}$$

where $i, j = 1, \ldots NM$ and $\widehat{z}_i$ and $\widehat{z}_j$ represent the sample means. The sample estimator in (24) is unbiased and easy to generate. Despite these advantages, for a small number of snapshots P, $\hat{\mathbf{C}}$ will exhibit a large error variance due to the large number of unknown parameters to be estimated. We can however introduce some structure to the sample covariance generally by reducing the number of free variables. Such structured estimates offer a relatively small error variance however they can be biased due to a misspecified structure. An optimal statistical solution is to reach an optimal tradeoff between the the unbiased but high variance estimate and the low variance but biased estimate. Thus the unbiased sample covariance is shrinked towards a biased structured covariance estimate.

To formalize the discussion and develop an expression for the shrinkage covariance estimate. We define $\mathbf{C}_0$ to be the true covariance matrix of the path-loss readings $\mathbf{z}$. Let $\hat{\mathbf{C}}$ be an unbiased estimate of $\mathbf{C}_0$. Let also the shrinkage target $\mathbf{T}$ be another biased structured estimate of $\mathbf{C}_0$ with relatively small number of parameters, due to which $\mathbf{T}$ will consist of a relatively small error variance. Thus instead of choosing between the two extremes, the shrinkage estimator $\mathbf{S}$ combines both estimates in a weighted fashion i.e.

$$\mathbf{S} = \lambda \mathbf{T} + (1-\lambda)\,\hat{\mathbf{C}}, \tag{25}$$

where λ is the shrinkage intensity and its value is between 0 and 1. It is noted that if $\lambda = 1$, then the shrinkage estimate is the shrinkage target and the unbiased covariance is given no weight. On the other hand, if $\lambda = 0$, then no shrinkage takes place and the unbiased estimator is chosen as the shrinkage estimate. Here two obvious questions arise, firstly, how to select the shrinkage target $\mathbf{T}$ and secondly, what value should be given to the shrinkage intensity λ. A number of shrinkage targets can be used in (25), e.g. [3] lists six target shrinkage intensities. In general, selection of $\mathbf{T}$ should be driven by the lower dimension structure in the data. Thus in the case of the covariance matrix $\mathbf{C}_0$ for multiple TNs, the natural shrinkage target is the constant correlation shrinkage target. For the constant correlation shrinkage target $[\mathbf{T}]_{ii} = \left[\hat{\mathbf{C}}\right]_{ii}$ and $[\mathbf{T}]_{ij} = \hat{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}$, where $\hat{\rho}$ is the average correlation of all the correlations between the network nodes i.e.

$$\hat{\rho} = \frac{1}{NM\,(NM-1)} \sum_{i=1}^{NM}\sum_{j\neq i}^{NM} \frac{\left[\hat{\mathbf{C}}\right]_{ij}}{\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}}. \tag{26}$$

2) Optimal shrinkage intensity: In order to find the optimal shrinkage intensity $\hat{\lambda}$, the Frobenius norm of the difference between the shrinkage estimate and the true covariance matrix is minimized with respect to λ. This is achieved by minimizing the following cost function

$$L(\lambda) = E\left\|\lambda \mathbf{T} + (1-\lambda)\,\hat{\mathbf{C}} - \mathbf{C}_0\right\|_F^2 \tag{27}$$

or equivalently

$$L(\lambda) = E\left\{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left(\lambda\,[\mathbf{T}]_{ij} + (1-\lambda)\left[\hat{\mathbf{C}}\right]_{ij} - [\mathbf{C}_0]_{ij}\right)^2\right\} \tag{28}$$

Expanding (28) leads to (29).

Finally, using $E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) = [\mathbf{C}_0]_{ij}$ in (30), the value of $\hat{\lambda}$ is obtained from the following expression

$$\hat{\lambda} = \frac{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[\mathrm{Var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) - \mathrm{Cov}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right]}{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[E\left([\mathbf{T}]_{ij} - \left[\hat{\mathbf{C}}\right]_{ij}\right)^2\right]}. \tag{31}$$

$$L(\lambda)=\sum_{i=1}^{NM}\sum_{j=1}^{NM}\lambda^2\mathrm{Var}\left([\mathbf{T}]_{ij}\right)+(1-\lambda)^2\,\mathrm{Var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)+2\lambda(1-\lambda)\mathrm{Cov}\left([\mathbf{T}]_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)$$
$$+\left[\lambda E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)+E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)-[\mathbf{C}_0]_{ij}\right]^2 \quad (29)$$

Taking derivative with respect to λ and equating to 0 yields

$$\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left\{2\lambda\mathrm{var}\left([\mathbf{T}]_{ij}\right)-2(1-\lambda)\,\mathrm{var}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)+2(1-2\lambda)\,\mathrm{Cov}\left([\mathbf{T}]_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)+\right.$$
$$\left.2\left[E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)\right]\left[\lambda E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)+E\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)-[\mathbf{C}_0]_{ij}\right]\right\}=0 \quad (30)$$

Since the variance and covariances in (31) are also estimated, they are replaced with their sample estimates to obtain the optimal shrinkage intensity i.e.

$$\hat{\lambda}=\frac{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[\widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)-\widehat{\mathrm{Cov}}\left(\mathbf{T}_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)\right]}{\sum_{i=1}^{NM}\sum_{j=1}^{NM}\left[E\left([\mathbf{T}]_{ij}-\left[\hat{\mathbf{C}}\right]_{ij}\right)^2\right]}. \quad (32)$$

Expressions to obtain $\widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)$ and $\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij},\left[\hat{\mathbf{C}}\right]_{ij}\right)$ are derived in the Appendix.

The value of $\hat{\lambda}\in[0\ 1]$ in most cases. In the unlikely event if $\hat{\lambda}$ exceeds these limits, the following bound is imposed.

$$\hat{\hat{\lambda}}=\max\left(0,\min\left(1,\hat{\lambda}\right)\right).$$

Thus the shrinkage covariance matrix $\mathbf{S}$ obtained from (25) is used in the Newton step (13) to update the GLS algorithm (12). Similarly, the variance and covariance values obtained from $\mathbf{S}$ are used in the generation of the covariance matrix $\mathbf{C_v}$ for the WLLS algorithm.

V. Simulation results

In this section, we analyze the performance of the shrinkage covariance estimate $\mathbf{S}$ and its impact on the performance of WLLS and GLS algorithms. We consider a circular deployment of N SNs with radius R m. Within the network are randomly deployed M TNs. The network deployment is shown in Fig. 1. Each SN-TN link is given a random shadowing variance σ_{ijkl}^2. The PLE value is selected to be 3 ($\alpha=3$).

The GLS algorithm is operated for η number of iterations, while to show an average performance for different number of snapshots P, for every snapshot value, the simulation is run υ times independently.

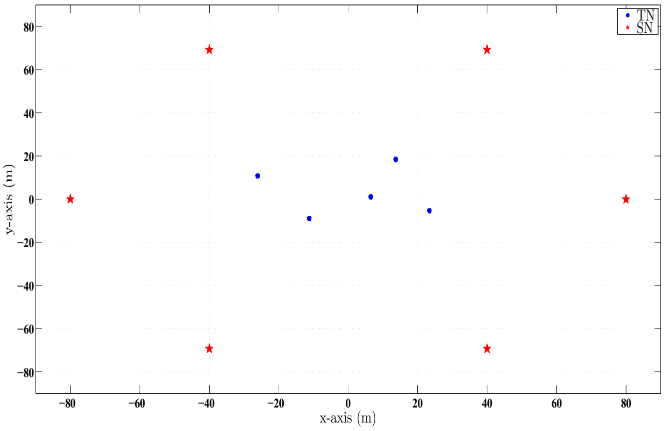

Figure 1. Network deployment

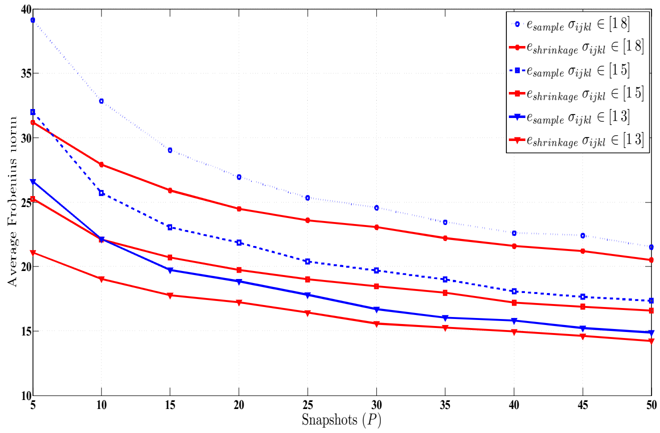

Figure 2. Error comparison of the estimated sample covariance matrix $\hat{\mathbf{C}}$ and shrinkage covariance matrix $\mathbf{S}$. $R = 80$ m, $N = 6$, $M = 5$, $\alpha = 3$, $d_c = 40$ m, $\upsilon = 100$.

Fig. 2 shows the error between the two covariance matrix estimates, i.e. sample covariance estimate $\hat{\mathbf{C}}$ and shrinkage covariance estimate $\mathbf{S}$. For this purpose, we base our evaluation on the Frobenius norm of the difference between the true and shrinkage covariance matrix i.e. $e_{shrinkage} = \|\mathbf{S} - \mathbf{C}_0\|_F^2$ and also between the true and sample covariance matrix i.e $e_{sample} = \left\|\hat{\mathbf{C}} - \mathbf{C}_0\right\|_F^2$. The comparison is done for random shadowing variance for each link between three different bounds i.e. each link is given a random σ_{ijkl}^2 value between [1 3], [1 5] and [1 8]. It is evident from Fig. 2 that $e_{shrinkage}$ is lower than than e_{sample} in all three cases. The difference in performance is more profound for smaller number of P. Also for larger σ_{ijkl}^2 bounds, larger error is shown by both estimators.

Fig. 3 compares the values of the optimal shrinkage intensity for different number of snapshots P

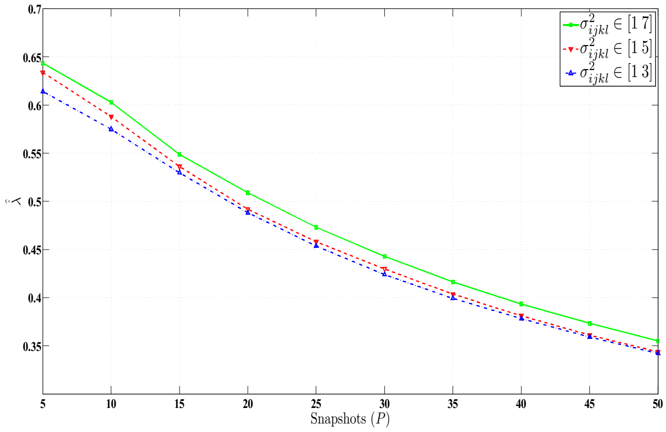

Figure 3. Average value for optimal shrinkage intensity $\hat{\hat{\lambda}}$ over $\upsilon = 100$ independent runs. $N = 6$, $M = 5$, $R = 80$ m, $\alpha = 3$, $\upsilon = 50$, $d_c = 40$ m,.

and different sets of shadowing variance σ^2_{ijkl}. From Fig. 3, we deduce the following; i) the optimal shrinkage intensity is indeed between the two extremes i.e. $\hat{\hat{\lambda}} \in [0\ 1]$. ii) $\hat{\hat{\lambda}}$ decreases as the number of snapshots increases, this is expected as with more snapshots the sample estimator $\hat{\mathbf{C}}$ becomes a more accurate estimator of the covariance matrix $\mathbf{C}_0$ and hence a lower weight is given to the target shrinkage covariance $\mathbf{T}$ according to (25). iii) $\hat{\hat{\lambda}}$ is larger when the bound of shadowing variance σ^2_{ijkl} is large, this too is supporting the shrinkage theory as with increased variance in the samples, less weight is given to the sample estimator $\hat{\mathbf{C}}$ for the same number of snapshots. The difference for the different sets of σ^2_{ijkl} is however not significant.

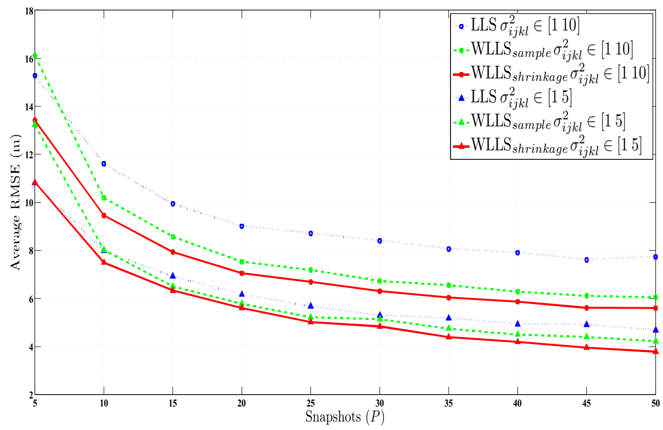

Figure 4. Average RMSE comparison between $\text{WLLS}_{shrinkage}$ and WLLS_{sample}. $N = 6, M = 5, R = 80$ m, $\alpha = 3,\ \upsilon = 50, d_c = 40$m.

Fig. 4, compares the average root mean squares error (RMSE) performance of the WLLS with sample covariance matrix WLLS_{sample}, WLLS with the shrinkage covariance matrix $\text{WLLS}_{shrinkage}$ and the LLS estimator (i.e. when $\mathbf{C}_\mathbf{v} = \mathbf{I}$) for different number of snapshots. The comparison is done for random shadowing variance between two sets. It is observed that the $\mathrm{WLLS}_{shrinkage}$ due to its relatively accurate estimation of the covariance matrix performs better than WLLS_{sample}, the performance difference is significant for a smaller number of snapshots. However even for larger values of P, $\mathrm{WLLS}_{shrinkage}$ still performs superior to WLLS_{sample}.

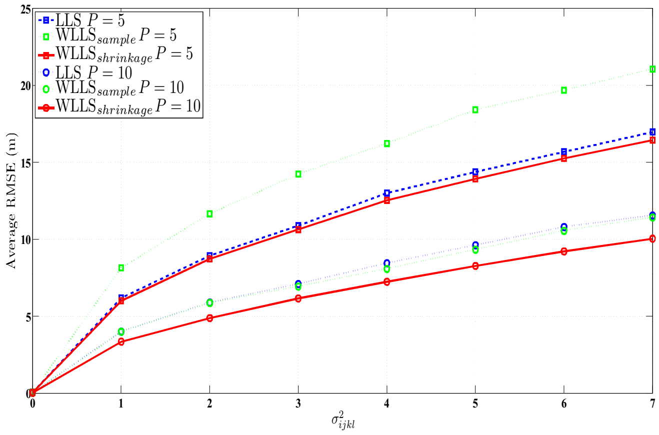

Figure 5. Average RMSE comparison between $\mathrm{WLLS}_{shrinkage}$ and WLLS_{sample}. $N = 6, M = 5, R = 80$ m, $\alpha = 3,\ \upsilon = 50, d_c = 40\,\mathrm{m}$.

In Fig. 5, we compare the average RMSE of the location estimate of WLLS_{sample}, $\mathrm{WLLS}_{shrinkage}$ and LLS corresponding to increasing value of shadowing variance σ^2_{ijkl}. In this case, σ^2_{ijkl} is same for all SN-TN links i.e. $\sigma^2_{ijkl} = \sigma^2\, \forall\, i, j, k, l$. The plots are for two fixed snapshot values i.e $P = 5$ and $P = 10$. It is again noted that the $\mathrm{WLLS}_{shrinkage}$ performs superior to both WLLS_{sample} and LLS. It is also interesting to note that the for $P = 5$, WLLS_{sample} performs even worse than the LLS, this is also evident from Fig. 4 for $P = 5$. The reason being the large error in the sample estimate $\hat{\mathbf{C}}_{\mathbf{v}}$ with a small number of P.

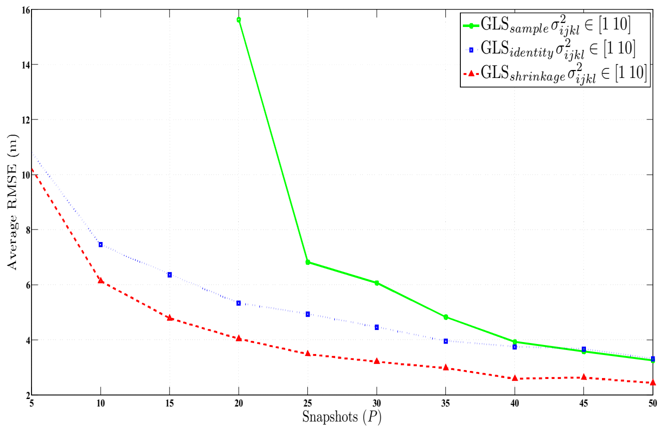

Figure 6. Average RMSE comparison between $\mathrm{GLS}_{shrinkage}$ and GLS_{sample}. $N = 5$, $M = 6$, $R = 80$ m, $\alpha = 3, \eta = 5,\ \upsilon = 50, d_c = 40\,\mathrm{m}$, $\boldsymbol{\theta}^1_k = [0, 0]^T$.

Fig. 6 compares the performance of the GLS algorithm with the sample covariance matrix GLS_{sample}, GLS with shrinkage covariance matrix $\mathrm{GLS}_{shrinkage}$ and the GLS estimator with identity covariance matrix $\mathrm{GLS}_{identity}$. The average RMSE of all the TNs is compared with the number of snapshots P. The center of the network is selected as the initial estimate for all TNs i.e. $\boldsymbol{\theta}_k^1 = [0, 0]^T \forall\, k$. The GLS algorithm is iterated $\eta = 5$ times. Due to the lack of structure in the sample covariance matrix $\hat{\mathbf{C}}$, it is non-invertible for $P < 20$ and hence GLS_{sample} produces no result for these values of P. $P = 20$ is by no means a minimum value for the invertibility of $\hat{\mathbf{C}}$, in fact the minimum number of snapshots will increase with the increase in the number of TNs. For $P > 20$ the GLS_{sample} shows unacceptable error and performs worse than $\mathrm{GLS}_{identity}$. On the other hand, the shrinkage covariance matrix $\mathbf{S}$ is both invertible and consists of a small error variance, consequently $\mathrm{GLS}_{shrinkage}$ performs better than GLS_{sample} and $\mathrm{GLS}_{identity}$.

VI. Conclusions

In this paper, we focus on the optimized localization of multiple TNs with correlated measurements. We have shown that the correlation between readings from different TNs at the SNs can improve the estimation accuracy. Two RSS based localization algorithms were investigated. A closed form expression is derived for the covariance matrix for multiple TN WLLS algorithm. Furthermore, the covariance matrix is estimated with a limited number of snapshots and the shrinkage estimator. The shrinkage estimator is also used to estimate the covariance matrix for the GLS algorithm. To evaluate the performance of the proposed methods, correlation between multiple TN readings at the same SN is based on the the Gudmundson model. It is shown via simulation results that the proposed WLLS and GLS algorithms with the shrinkage covariance matrix perform superior to the same algorithms using the sample covariance matrix. Especially for a small number of snapshots the sample covariance matrix consists of a large error (and is non invertible in case of GLS). This results in degraded performance and in the worst case, the performance is inferior to the case where the identity matrix is used for the covariance matrix. The shrinkage covariance matrix on the other hand, guarantees a relatively small estimation error and it is always invertible. This results in superior location estimation performance in both WLLS and GLS.

Acknowledgment

We acknowledge the support from the UK Engineering and Physical Sciences Research Council (EPSRC) for the support via the Bayesian Tracking and Reasoning over Time (BTaRoT) grant EP/K021516/1.

APPENDIX

Following Schafer and Strimmer [3] and Kwan [24] whose study focussed on gene classification and security returns respectively, we derive optimal shrinkage intensity in the context of multitarget localization. We define the following: For P snapshots of path-loss measurements z_i, the sample mean is given by $\widehat{z}_i = P^{-1}\sum_{p=1}^{P} z_{ip}$. Let

$$v_{ijp} = \left(z_{ip} - \widehat{z}_i\right)\left(z_{jp} - \widehat{z}_j\right) \tag{33}$$

for $i, j = 1, ..., NM$ and $p = 1, ..., P$. Also let $\widehat{v}_{ij} = P^{-1}\sum_{p=1}^{P} v_{ijp}$. Then the sample covariance is given by

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{P}{P-1}\widehat{v}_{ij} = \frac{1}{P-1}\sum_{p=1}^{P}\left(z_{ip} - \widehat{z}_i\right)\left(z_{jp} - \widehat{z}_j\right). \tag{34}$$

For ease of understanding, v_{ijp} should be viewed as a random variable. The unbiased variance of individual elements of $\hat{\mathbf{C}}$ is given by

$$\widehat{\mathrm{Var}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}\right\} = \frac{P^2}{\left(P-1\right)^2}\mathrm{Var}\left(\widehat{v}_{ij}\right). \tag{35}$$

Using the identity to find the variance of the mean of a random variable, we have

$$\widehat{\mathrm{Var}}\left(\widehat{v}_{ij}\right) = \frac{1}{P}\mathrm{var}\left(v_{ij}\right) \tag{36}$$

or

$$\widehat{\mathrm{Var}}\left(\widehat{v}_{ij}\right) = \frac{1}{P}\left[\frac{1}{P-1}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)^2\right]. \tag{37}$$

Thus the variance of the sample covariance matrix is given by

$$\widehat{\mathrm{Var}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}\right\} = \frac{P}{\left(P-1\right)^3}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)^2. \tag{38}$$

On the other hand, the covariance elements are given as

$$\widehat{\mathrm{Cov}}\left\{\left[\hat{\mathbf{C}}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{kl}\right\} = \frac{P}{\left(P-1\right)^3}\sum_{p=1}^{P}\left(v_{ijp} - \widehat{v}_{ij}\right)\left(v_{klp} - \widehat{v}_{kl}\right). \tag{39}$$

An expression for $\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ is slightly more complicated, here we use the so called 'delta method'. For this purpose, we consider the individual elements of $\hat{\mathbf{C}}$ as variables and expand $[\mathbf{T}]_{ij} =$

$\hat{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}$ via Taylor series around the corresponding point estimates $\left[\bar{\mathbf{C}}\right]_{ii}$, $\left[\bar{\mathbf{C}}\right]_{jj}$ and $\left[\bar{\mathbf{C}}\right]_{ij}$ provided by the samples. We have

$$[\mathbf{T}]_{ij} = \bar{\rho}\sqrt{[\bar{\mathbf{C}}]_{ii}\,[\bar{\mathbf{C}}]_{jj}} + \frac{\bar{\rho}}{2}\sqrt{\frac{[\bar{\mathbf{C}}]_{jj}}{[\bar{\mathbf{C}}]_{ii}}}\left([\mathbf{C}]_{ii} - [\bar{\mathbf{C}}]_{ii}\right) + \frac{\bar{\rho}}{2}\sqrt{\frac{[\bar{\mathbf{C}}]_{ii}}{[\bar{\mathbf{C}}]_{jj}}}\left([\mathbf{C}]_{jj} - [\bar{\mathbf{C}}]_{jj}\right), \tag{40}$$

where $\bar{\rho}$ is obtained using (26) for $\bar{\mathbf{C}}_{ii}$ $\bar{\mathbf{C}}_{jj}$ and $\bar{\mathbf{C}}_{ij}$.

Then from the definition of the covariance matrix, we have

$$\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = E\left[\left([\mathbf{T}]_{ij} - E\left([\mathbf{T}]_{ij}\right)\right)\left([\mathbf{C}]_{ij} - E\left([\mathbf{C}]_{ij}\right)\right)\right] \tag{41}$$

using (40) in (41) leads to

$$\widehat{\mathrm{Cov}}\left([\mathbf{T}]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{\bar{\rho}}{2}\left(\sqrt{\frac{[\bar{\mathbf{C}}]_{jj}}{[\bar{\mathbf{C}}]_{ii}}}\mathrm{Cov}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) + \sqrt{\frac{[\bar{\mathbf{C}}]_{ii}}{[\bar{\mathbf{C}}]_{jj}}}\mathrm{Cov}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right).$$

Finally using (33) and (39), it is straightforward to define $\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ and $\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)$ as

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left\{\left[(z_{ip} - \widehat{z}_i)^2 - \widehat{v}_{ii}\right]\left[(z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j) - \widehat{v}_{ij}\right]\right\} \tag{42}$$

and similarly

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left\{\left[(z_{jp} - \widehat{z}_j)^2 - \widehat{v}_{jj}\right]\left[(z_{ip} - \widehat{z}_i)(z_{jp} - \widehat{z}_j) - \widehat{v}_{ij}\right]\right\}, \tag{43}$$

which completes the derivation.

REFERENCES

[1] O. Ledoit, and M. Wolf, "Honey, I Shrunk the Sample Covariance Matrix", *Journal of Portfolio Management*, 110-119, 2004.

[2] O. Ledoit, and M. Wolf, "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection", *Journal of Empirical Finance*, 10, 603-621, 2003.

[3] J. Schafer, K. Strimmer, “A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics,” *Statistical Applications in Genetics and Molecular Biology*, Vol. 4, no. 1, 2005.

[4] N. Patwari, J. N. Ash, S. Kyperountas, A. O. H. III, R. L. Moses, and N. S. Correal, “Locating the nodes: cooperative localization in wireless sensor networks,” *IEEE Signal Processing Mag*., vol. 22, no. 4, pp. 54–69, Jul. 2005.

[5] I. Guvenc and C.-C. Chong, “A survey on TOA based wireless localization and NLOS mitigation techniques,” *IEEE Commun. Surveys and Tutorials*, vol. 11, no. 3, pp. 107–124, Aug. 2009.

[6] B.D. Van Veen and K.M. Buckley, “Beamforming: A versatile approach to spatial filtering,” *IEEE ASSP Mag*., vol. 5, no. 2, pp. 4–24, Apr. 1988.

[7] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, “Exact and large sample ML techniques for parameter estimation and detection in array processing,” in *Radar Array Processing*, S.S. Haykin, J. Litva, and T. Shepherd, Eds. New York: Springer- Verlag, 1993, pp. 99–151.

[8] N. Patwari, A. O. Hero, III, M. Perkins, N. S. Correal, and R. J. O’Dea, “Relative location estimation in wireless sensor networks,” *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2137-2148, Aug. 2003.

[9] N. Salman, M. Ghogho, and A. H. Kemp, “On the joint estimation of the RSS-based location and path-loss exponent,” *IEEE Wireless Commun. Lett*., vol. 1, no. 1, pp. 34-37, Feb. 2012.

[10] N. Salman, A. H. Kemp and M. Ghogho, “Low Complexity Joint Estimation of Location and Path-Loss Exponent”, *IEEE Wireless Commun Lett.,* vol. 1, no. 4, pp. 364–367, Aug. 2012.

[11] A. Bel, J.L, Vicario, G. Seco-Granados, Localization Algorithm with On-line Path Loss Estimation and Node Selection, Sensors, 11, pp. 6905-6925, Jul. 2011.

[12] R. Ouyang, A.-S. Wong, and C.-T. Lea, “Received signal strength-based wireless localization via semidefinite programming: Noncooperative and cooperative schemes,” *IEEE Trans. Veh. Technol*., vol. 59, no. 3, pp. 1307– 1318, Mar. 2010.

[13] L. Mihaylova, D. Angelova, D. R. Bull, N. Canagarajah, "Localization of Mobile Nodes in Wireless Networks with Correlated in Time Measurement Noise," I*EEE Transactions on Mobile Computing*, vol.10, no.1, pp.44,53, Jan. 2011.

[14] M. Gudmundson, “Correlation model for shadow fading in mobile radio systems,” IEE Electronic Letters, 27(23):2145-2146, 7 Nov, 1991.

[15] N. Patwari and P. Agrawal, “Effects of correlated shadowing: Connectivity, localization, and RF tomography,” in *Proc. IEEE/ACM Int. Conf. Information Processing in Sensor Networks (IPSN)*, 2008, pp. 82–93.

[16] H. Hashemi, “The indoor radio propagation channel,” *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, July 1993.

[17] T.S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[18] K. Pahlavan and A. Levesque, *Wireless Information Networks*. New York: John Wiley & Sons, Inc., 1995.

[19] G. A. F. Seber, C. J. Wild, *Nonlinear regression*. John Wiley & Sons Inc, 2003.

[20] J. J. Caffery, “A new approach to the geometry of TOA location,” in Proc. *IEEE Veh. Technol. Conf. (VTC)*, vol. 4, Boston, MA, Sep. 2000, pp. 1943–1949.

[21] I. Guvenc, S. Gezici, Z. Sahinoglu, “Fundamental limits and improved algorithms for linear least-squares wireless position estimation,”*Wirel. Commun. Mob. Comput.* (2010).

[22] N. Salman, M. Ghogho, A. H. Kemp, “Optimized Low Complexity Sensor Node Positioning in Wireless Sensor Networks,” *IEEE Sensors Journal*, vol.14, no.1, pp.39-46, Jan. 2014.

[23] D. Baum, J. Hansen, and J. Salo, “An interim channel model for beyond- 3G systems: extending the 3GPP spatial channel model (SCM),” in *Proc. IEEE VTC*, May/Jun. 2005, vol. 5, pp. 3132–3136.

[24] C. C. Y. Kwan, "Estimation error in the average correlation of security returns and shrinkage estimation of covariance and correlation matrices," *Elsevier Finance Research Letters*, 5, 236-244, 2008.

IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE. Please read the form carefully and keep a copy for your files.

TITLE OF PAPER/ARTICLE/REPORT, INCLUDING ALL CONTENT IN ANY FORM, FORMAT, OR MEDIA (hereinafter, "the Work"):
A Shrinkage Approach to Multiple Target Localization with correlated Measurements.

COMPLETE LIST OF AUTHORS:
Naveed Salman, Lyudmila Mihaylova, Andrew H. Kemp

IEEE PUBLICATION TITLE (Journal, Magazine, Conference, Book):

COPYRIGHT TRANSFER

1. The undersigned hereby assigns to The Institute of Electrical and Electronics Engineers, Incorporated (the "IEEE") all rights under copyright that may exist in and to: (a) the above Work, including any revised or expanded derivative works submitted to the IEEE by the undersigned based on the Work; and (b) any associated written or multimedia components or other enhancements accompanying the Work.

CONSENT AND RELEASE

2. In the event the undersigned makes a presentation based upon the Work at a conference hosted or sponsored in whole or in part by the IEEE, the undersigned, in consideration for his/her participation in the conference, hereby grants the IEEE the unlimited, worldwide, irrevocable permission to use, distribute, publish, license, exhibit, record, digitize, broadcast, reproduce and archive, in any format or medium, whether now known or hereafter developed: (a) his/her presentation and comments at the conference; (b) any written materials or multimedia files used in connection with his/her presentation; and (c) any recorded interviews of him/her (collectively, the "Presentation"). The permission granted includes the transcription and reproduction of the Presentation for inclusion in products sold or distributed by IEEE and live or recorded broadcast of the Presentation during or after the conference.

3. In connection with the permission granted in Section 2, the undersigned hereby grants IEEE the unlimited, worldwide, irrevocable right to use his/her name, picture, likeness, voice and biographical information as part of the advertisement, distribution and sale of products incorporating the Work or Presentation, and releases IEEE from any claim based on right of privacy or publicity.

4. The undersigned hereby warrants that the Work and Presentation (collectively, the "Materials") are original and that he/she is the author of the Materials. To the extent the Materials incorporate text passages, figures, data or other material from the works of others, the undersigned has obtained any necessary permissions. Where necessary, the undersigned has obtained all third party permissions and consents to grant the license above and has provided copies of such permissions and consents to IEEE.

☐ Please check this box if you do not wish to have video/audio recordings made of your conference presentation

See reverse side for Retained Rights/Terms and Conditions, and Author Responsibilities.

GENERAL TERMS

- The undersigned represents that he/she has the power and authority to make and execute this assignment.
- The undersigned agrees to indemnify and hold harmless the IEEE from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.
- In the event the above work is not accepted and published by the IEEE or is withdrawn by the author(s) before acceptance by the IEEE, the foregoing copyright transfer shall become null and void and all materials embodying the Work submitted to the IEEE will be destroyed.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.

(1)______________________________ 15/08/2014.
Author/Authorized Agent for Joint Authors **Date**

U.S. GOVERNMENT EMPLOYEE CERTIFICATION (WHERE APPLICABLE)

This will certify that all authors of the Work are U.S. government employees and prepared the Work on a subject within the scope of their official duties. As such, the Work is not subject to U.S. copyright protection.

(2)______________________________ ______________________
Authorized Signature **Date**

(Authors who are U.S. government employees should also sign signature line (1) above to enable the IEEE to claim and protect its copyright in international jurisdictions.)

CROWN COPYRIGHT CERTIFICATION (WHERE APPLICABLE)

This will certify that all authors of the Work are employees of the British or British Commonwealth Government and prepared the Work in connection with their official duties. As such, the Work is subject to Crown Copyright and is not assigned to the IEEE as set forth in the first sentence of the Copyright Transfer Section above. The undersigned acknowledges, however, that the IEEE has the right to publish, distribute and reprint the Work in all forms and media.

(3)______________________________ ______________________
Authorized Signature **Date**

(Authors who are British or British Commonwealth Government employees should also sign line (1) above to indicate their acceptance of all terms other than the copyright transfer.)

rev. 020711