Tracking With Sparse and Correlated Measurements via a Shrinkage-Based Particle Filter

Aroland Kiring, Naveed Salman, Chao Liu, Iñaki Esnaola, and Lyudmila Mihaylova, *Senior Member, IEEE*

***Abstract*— This paper presents a shrinkage-based particle filter method for tracking a mobile user in wireless networks. The proposed method estimates the shadowing noise covariance matrix using the shrinkage technique. The particle filter is designed with the estimated covariance matrix to improve the tracking performance. The shrinkage-based particle filter can be applied in a number of applications for navigation, tracking, and localization when the available sensor measurements are correlated and sparse. The performance of the shrinkage-based particle filter is compared with the posterior Cramer–Rao lower bound, which is also derived in this paper. The advantages of the proposed shrinkage-based particle filter approach are demonstrated via simulation and experimental results.**

***Index Terms*— Wireless sensor networks, tracking problems, received signal strength measurements, particle filter, covariance matrix, shrinkage estimation.**

I. Introduction

TRACKING a mobile user using the received signal strength (RSS) measurements is one of the many applications of wireless networks. The accuracy of tracking algorithms is affected by the quality of the received signal, the size of the obstructions encountered, and sudden changes in the speed of the mobile users. Various methods of target tracking based on the signal strength have been developed to address challenges such as improving the tracking accuracy [1]–[3] and saving the energy consumption of the deployed sensor nodes [4]. Techniques for tracking mobile users in wireless systems are divided into two groups: methods in which point coordinates are estimated using global positioning system (GPS) devices [5] and methods in which the coordinate and motion are estimated using an underlying mobility model with filtering algorithms [6].

GPS devices operate effectively in outdoor environments. However, when operates in indoor environments, or in areas where there is an obstructed line of sight to GPS satellites, e.g., in hills, high buildings, and dense forests, the GPS may not be able to establish a connection with the satellites. Sensors equipped with GPS capabilities require the installation of extra hardware making it costly to deploy. Assisted GPS (AGPS) devices [7] offer a better solution in situations where GPS devices have poor signal reception by establishing a communication with the satellite via cellular networks. Coordinates estimated using AGPS devices are faster but less accurate when compared to GPS devices.

Alternatively, mobility models with filtering algorithms can also be applied to estimate both the user coordinate and the motion. Various mobility models have been developed such as random walk based models [8], Gauss-Markov models, and Singer-type models [9], [10]. In [10], a dynamic mobility model that captures a wide range of vehicle maneuvering patterns is presented and employed for tracking in tactical weapons systems. Liu *et al*. [11] use this mobility model to estimate the trajectory of mobile users in wireless asynchronous transfer mode (ATM) networks by applying a modified Kalman filter (KF). However, the tracking accuracy is poor when there is a rapid change in the user acceleration.

The KF is the optimal estimator, in the minimum mean square error sense, for linear systems with observations corrupted by a Gaussian noise. However, when the KF is applied to a non-linear system, the estimator faces difficulties. Zaidi and Mark [12] develop an extended Kalman filter (EKF), which operates by first linearizing the state and/or measurement model before applying the standard KF. Unfortunately, the EKF produces unreliable estimates when the non-linearities in the system model and/or in the measurement model are severe. In this case, measurement-conversion techniques are used to solve the non-linear equations and improve the performance of the filter [13].

Filtering methods based on random sampling can also be applied in mobility tracking scenarios, such as the Ensemble Kalman filter (EnKF) and the Unscented Kalman filter (UKF). The EnKF performs a random sampling of the probability density function to represent the initial state estimate [14]. In contrast, the UKF relies on a deterministic choice of sampling points, called sigma points [15], [16]. The aforementioned methods assume that the process noise and the measurement noise are Gaussian distributed.

The Particle filter (PF) [17] is often used in non-linear and non-Gaussian filtering problems. The PF treats the random samples as particles, each with a corresponding weight. The PF updates the weights when a new measurement is received before it approximates the state of the mobile user. In [18] and [19], the performance of the PF is studied and the influences of the number of generated particles and the

Manuscript received June 17, 2016; revised November 11, 2016 and January 20, 2017; accepted January 30, 2017. Date of publication March 29, 2017; date of current version April 20, 2017. This work was supported in part by the Ministry of Education of Malaysia through Skim Latihan Akademik IPTA and in part by the U.K. Engineering and Physical Sciences Research Council through the Bayesian Tracking and Reasoning Over Time under Grant EP/K021516/1. The associate editor coordinating the review of this paper and approving it for publication was Dr. Shanhong Xia.

The authors are with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S10 2TN, U.K. (e-mail: amkiring1@sheffield.ac.uk; n.salman@sheffield.ac.uk; cliu47@sheffield.ac.uk; esnaola@sheffield.ac.uk; l.s.mihaylova@sheffield.ac.uk).

Digital Object Identifier 10.1109/JSEN.2017.2685684

resampling methods on its accuracy are analyzed. In [20], a mobility model combined with a PF is developed for mobile tracking in cellular networks.

The accuracy of the mobile user tracking method depends on the quality of the received signal. In ad hoc networks, connections between the user and the sensors are spontaneously established when the user is within the communication range. The mobility of sensors and user changes the network topology and results in a disruption of the communication. This limits the number of measurements that are available. The measurements from the nearby sensor nodes are often assumed to be spatially correlated. However, consecutive measurements also exhibit temporal correlation. For that reason, the tracking can be improved by exploiting the spatio-temporal correlation between the measurements. The shadowing noise covariance matrix depends on the distance between the sensor nodes and the mobile users. As a result, the movement of the mobile users induces changes on the covariance matrix.

To overcome these problems, we propose a shrinkage-based PF (ShPF) method for approximating the state of a mobile user based on sparse and correlated measurements. In particular, we adapt the Singer mobility model and the PF algorithm developed in [20] to the scenario of network with sparse and correlated measurements. The shrinkage technique is employed to overcome the limitation faced by the sample estimator. Then, the tracking accuracy of the ShPF is compared with the PF to validate the performance of the developed filter. The performance of the ShPF is demonstrated via simulation and experimental results. The posterior Cramer-Rao lower bound (PCRLB) is also derived and calculated for the simulated data.

The key contributions of this paper are: (i) the RSS based tracking framework is developed for correlated and sparse wireless sensor measurements; (ii) the framework employs a joint shrinkage technique and a PF algorithm for the estimation of the state of the mobile user; (iii) the ShPF performance is assessed over simulated data and real Wi-Fi data.

The rest of the paper is organized as follows. The problem formulation, user mobility model, sensor measurement model and measurement correlation model are given in Section II. Section III derives the shrinkage method. Section IV reviews the PCRLB. The proposed ShPF method is presented in Section V. Results and analyses of both simulated and real Wi-Fi data are given in Section VI. Finally, Section VII presents the conclusions.

II. Tracking System Model

A. Problem Formulation

Consider a two-dimensional (2-D) network consisting of n_s sensors that are uniformly distributed with known coordinates $(x_i, y_i),\ i \in \{1, \ldots, n_s\}$. These coordinates can be obtained using GPS devices, or by installing sensors at known points. The sensors measure the RSS of a mobile user, and all the measurements are collected and processed at a central unit to estimate the trajectory of the mobile user.

Notations: In the following, we denote vectors by lowercase boldface letters, matrices by uppercase boldcase letters, $(\cdot)^T$ is the transpose operator, $\mathbb{E}[\cdot]$ is the expectation operator, $\mathbf{I}$ denotes the identity matrix, $[\mathbf{A}]_{ij}$ refers to the element at the i-th row and j-th column of matrix $\mathbf{A}$, $\mathcal{N}(\mu, \sigma^2)$ represents a Gaussian distribution with mean μ and variance σ^2, $\|\cdot\|_F^2$ is the Frobenius norm, $|\cdot|$ is the matrix determinant, $\mathcal{L}(\cdot)$ is the likelihood function, $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^{m\times n}$ is the vector space of all m-by-n real matrices, and $\widehat{\mathrm{Cov}}(\cdot)$ is the covariance operator.

B. User Mobility Model

Let the state vector of a mobile user at time k be given by

$$\mathbf{x}_k = [x_k,\ \dot{x}_k,\ \ddot{x}_k,\ y_k,\ \dot{y}_k,\ \ddot{y}_k]^T, \tag{1}$$

where (x_k, y_k) represents the user coordinates, $(\dot{x}_k, \dot{y}_k)$ represents the user velocity, and $(\ddot{x}_k, \ddot{y}_k)$ represents the user acceleration, respectively. A Singer-type mobility model describes the evolution of the mobile user state (and also the speed and acceleration) with respect to time. The Singer model [11], [17] yields

$$\mathbf{x}_{k+1} = \mathbf{A}(T, \alpha)\mathbf{x}_k + \mathbf{B}_u(T)\mathbf{u}_{k+1} + \mathbf{B}_w(T)\mathbf{w}_{k+1}, \tag{2}$$

where T is the discretization period, $\alpha = 1/\tau$ is the reciprocal of the maneuvering time constant τ, and $\mathbf{u}_{k+1} = [u_{x,k+1},\ u_{y,k+1}]^T$ is the command process that changes the acceleration. Following [17] and [21], parameters $u_{x,k+1}$ and $u_{y,k+1}$ are modeled as a Markov chain and take values from a set of acceleration levels $\mathcal{M}_x$ and $\mathcal{M}_y$, where $\mathcal{M}_x$ and $\mathcal{M}_y$ are the set of all possible acceleration levels in the x and y directions, respectively. Consequently, $\mathbf{u}_{k+1}$ takes values from the set $\mathbb{M} = \mathcal{M}_x \times \mathcal{M}_y = \{\mathbf{m}_1, \ldots, \mathbf{m}_M\}$, with transition probabilities $\boldsymbol{\pi}_{ij} = P(\mathbf{u}_{k+1} = \mathbf{m}_j | \mathbf{u}_k = \mathbf{m}_i)$ and initial probabilities distribution $\mu_{i,0} = P\{\mathbf{m} = \mathbf{m}_i\}$ for modes $\mathbf{m}_i \in \mathbb{M}$ such that $\mu_{i,0} \geq 0$ and $\sum_{i=1}^{M} \mu_{i,0} = 1$. The matrix $\mathbf{A}(T, \alpha) \in \mathbb{R}^{6\times 6}$ is a state transition matrix, $\mathbf{B}_u(T) \in \mathbb{R}^{6\times 2}$ is the command matrix and $\mathbf{B}_w(T) \in \mathbb{R}^{6\times 2}$ is the noise coefficient matrix. These are given by (3), (4), and (5), respectively

$$\mathbf{A}(T, \alpha) = \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \tilde{\mathbf{A}} \end{bmatrix}, \quad \tilde{\mathbf{A}} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & \alpha \end{bmatrix}, \tag{3}$$

$$\mathbf{B}_u(T) = \begin{bmatrix} \tilde{\mathbf{B}}_u & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} & \tilde{\mathbf{B}}_u \end{bmatrix}, \quad \tilde{\mathbf{B}}_u = \begin{bmatrix} T^2/2 \\ T \\ 0 \end{bmatrix}, \tag{4}$$

$$\mathbf{B}_w(T) = \begin{bmatrix} \tilde{\mathbf{B}}_w & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} & \tilde{\mathbf{B}}_w \end{bmatrix}, \quad \tilde{\mathbf{B}}_w = \begin{bmatrix} T^2/2 \\ T \\ 1 \end{bmatrix}, \tag{5}$$

and $\mathbf{w}_{k+1} = [w_{x,k+1}, w_{y,k+1}]^T$ is a multivariate Gaussian random variable, with zero mean and covariance matrix $\mathbf{Q} = \mathbb{E}[\mathbf{w}_{k+1}\mathbf{w}_{k+1}^T] = \sigma_w^2\mathbf{I}$ where σ_w is the standard deviation of the process noise.

C. Measurement Model

The measurement-based coordinate estimation can be related to RSS through the radio propagation path loss model [22]

$$z^i = z^0 + 10\beta \log_{10}(d^i) + v^i, \tag{6}$$

where z^0 is the signal power loss in dB at 1 m distance and z^i is the signal power loss at distance d^i, where i refers to the i-th sensor node. The parameter β is the path loss exponent which is typically $\beta \in [4, 8]$. Finally, $v^i \sim \mathcal{N}(0, (\sigma^i)^2)$ is a zero mean Gaussian random variable representing the shadowing noise where σ^i is the standard deviation of the shadowing noise at the i-th sensor node. The parameter z^0 can be calculated during system calibration and $z^i = P_t - P_r$ can be determined at the receiver node by measuring the received signal power P_r when the power of the transmitted signal P_t is known [23]. To enable accurate tracking, a minimum of three sensor measurements is needed. The measurement model in (6) can be written in vector form as

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \tag{7}$$

where $\mathbf{z}_k \in \mathbb{R}^{n_s}$ represents the measurements at n_s sensor nodes at time instant k, i.e., $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^{n_s}]^T$ and the elements of the vector $\mathbf{h}(\mathbf{x}_k)$ are given by $h(x_k^i) = z_k^0 + 10\beta \log_{10}(d_k^i)$ for $i = 1, 2, \ldots, n_s$. The vector $\mathbf{v}_k$ represents the shadowing noise, with covariance matrix $\mathbf{R}_k = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T]$, and is assumed to be correlated in space and time.

D. Correlated Data Model for Measurements Generation

The measurements described by (7) are assumed to be corrupted by additive noise correlated in both space and time. In practice, this correlation is unknown but in the simulations presented in Section VI, the spatial and temporal correlations of the measurements are produced by the Gudmundson model [23].

1) Spatial Dependence: The spatial correlation coefficient between the measurements at the i-th and the j-th sensor nodes, at time instant k, is given by

$$\rho_k^{i,j} = \exp\left(- \frac{d_k^{i,j}}{D_c} \right), \tag{8}$$

where $d_k^{i,j}$ is the relative distance between the two sensors and D_c is the decorrelation distance [24]. Thus, the covariance between the measurements at the two sensors is given by

$$C_k^{i,j} = \rho_k^{i,j} \sigma_k^i \sigma_k^j, \tag{9}$$

where σ_k^i and σ_k^j are the standard deviations of the shadowing noise at the i-th and the j-th sensor nodes, respectively.

2) Temporal Dependence: The temporal correlation between the RSS measured at time instants k and l by the i-th sensor, is given by

$$\tilde{\rho}_{k,l}^i = \exp\left(- \frac{d_{k,l}^i}{D_c} \right), \tag{10}$$

where $d_{k,l}^i$ is the distance traveled by the mobile user from the time instant k, to the time instant l, which is given by $d_{k,l}^i = \sqrt{(x_l^i - x_k^i)^2 + (y_l^i - y_k^i)^2}$ where (x_l^i, y_l^i) are the user coordinates at time instant l and (x_k^i, y_k^i) are the user coordinates at time instant k. Thus, the covariance between the RSS, measured at time instants k and l by the i-th sensor, is given by

$$\tilde{C}_{k,l}^i = \tilde{\rho}_{k,l}^i \tilde{\sigma}_k^i \tilde{\sigma}_l^i, \tag{11}$$

where $\tilde{\sigma}_k^i$ and $\tilde{\sigma}_l^i$ are the standard deviations of the shadowing noise at the time instant k and l, respectively.

3) Spatio-Temporal Dependence: A set of sensor measurements is collected from n_s sensors at different time instants. The measurements taken at time instant k are temporally correlated with the measurements taken at all previous time instants, l for $l \in \{1, \ldots, L\}$ given by

$$\mathbf{Z} = [\mathbf{z}_k^T, \ldots, \mathbf{z}_l^T]^T, \tag{12}$$

and the resulting block covariance matrix is given by

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{k,k} & \ldots & \mathbf{C}_{k,l} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{l,k} & \ldots & \mathbf{C}_{l,l} \end{bmatrix}, \tag{13}$$

where the diagonal elements of the block covariance matrix are of the form

$$\mathbf{C}_{k,k} = \begin{bmatrix} (\sigma_k^1)^2 & \rho_k^{1,2} \sigma_k^1 \sigma_k^2 & \cdots & \rho_k^{1,n_s} \sigma_k^1 \sigma_k^{n_s} \\ \rho_k^{2,1} \sigma_k^2 \sigma_k^1 & (\sigma_k^2)^2 & \cdots & \rho_k^{2,n_s} \sigma_k^2 \sigma_k^{n_s} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k^{n_s,1} \sigma_k^{n_s} \sigma_k^1 & \rho_k^{n_s,2} \sigma_k^{n_s} \sigma_k^2 & \cdots & (\sigma_k^{n_s})^2 \end{bmatrix}, \tag{14}$$

and the off-diagonal elements of the block covariance matrix are of the form

$$\mathbf{C}_{k,l} = \begin{bmatrix} \tilde{\rho}_{k,l}^1 \tilde{\sigma}_k^1 \tilde{\sigma}_l^1 & 0 & \cdots & 0 \\ 0 & \tilde{\rho}_{k,l}^2 \tilde{\sigma}_k^2 \tilde{\sigma}_l^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\rho}_{k,l}^{n_s} \tilde{\sigma}_k^{n_s} \tilde{\sigma}_l^{n_s} \end{bmatrix}. \tag{15}$$

The size of the observed measurement vector (12) depends on all the measurements from the previous time instants. As the number of acquired measurements increases, the size of the covariance matrix grows exponentially. Thus, to limit the dimensionality of the resulting covariance matrix, a restriction on the number of previous measurements is imposed through a sliding window time.

III. Shrinkage Method

A. Covariance Matrix of the Shadowing Noise

In practical scenarios, the covariance matrix of the shadowing noise is unknown. Therefore, the covariance matrix is estimated from the available sensor measurements. A common approach to estimate the covariance matrix is by using a sample covariance estimator. Given the set of RSS measurements $\{\mathbf{z}_p \in \mathbb{R}^{n_s}\}$, for $p = 1, \ldots, P$ where P is the number of observations, the sample covariance matrix estimate is given by

$$\hat{\mathbf{C}} = \frac{1}{P-1} \sum_{p=1}^{P} (\mathbf{z}_p - \bar{\mathbf{z}})(\mathbf{z}_p - \bar{\mathbf{z}})^T, \tag{16}$$

where $\bar{\mathbf{z}} = \frac{1}{P} \sum_{p=1}^{P} \mathbf{z}_p$ is the sample mean. The estimated sample covariance matrix in (16) is unbiased and provides accurate estimates when $P \gg n_s$. When $P \leq n_s$, the sample covariance matrix estimate is ill-conditioned, non-invertible, and introduces a large estimation error. This limitation can be overcome by using the shrinkage technique.

B. Shrinkage Estimation of the Covariance Matrix

The term shrinkage relates to the notion that an original estimate is improved through shrinking to the value supplied by additional information. A shrinkage covariance matrix estimate (improved estimate of the sample covariance matrix) is given by [25],

$$\hat{\mathbf{S}} = \lambda \mathbf{T} + (1-\lambda)\hat{\mathbf{C}}, \tag{17}$$

where the additional information introduced here is the target matrix, $\mathbf{T}$. The shrinkage intensity, $\lambda \in [0,1]$ controls the extent to which the improved estimate $\hat{\mathbf{S}}$ shrinks from the original estimate $\hat{\mathbf{C}}$ to the target matrix $\mathbf{T}$. Note that if $\lambda = 1$, the shrinkage estimate is equivalent to the target matrix and the sample covariance estimate is given no weight. On the other hand, if $\lambda = 0$, no shrinkage takes place and the sample covariance estimate dominates.

Here two questions arise. Firstly, how should the shrinkage target matrix $\mathbf{T}$ be selected? Secondly, what value should be given to the shrinkage intensity λ? Schafer and Strimmer propose six target matrices and approaches to compute corresponding optimal shrinkage intensities in [26]. According to [24], the selection of target matrix $\mathbf{T}$ should be driven by the data. Thus in the case of the shadowing noise covariance estimation for path loss measurements, two types of target matrices are selected. The first target matrix is the diagonal, unit variance shrinkage target matrix given by

$$\mathbf{T}_1 = \mathbf{I}, \tag{18}$$

with optimal shrinkage intensity determined by

$$\hat{\lambda}_{\mathbf{T}_1} = \frac{\sum_{ij} \widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)}{\sum_{i\neq j}\left[\hat{\mathbf{C}}\right]_{ij}^2 + \sum_i \left(\left[\hat{\mathbf{C}}\right]_{ii} - 1\right)^2}. \tag{19}$$

The second target matrix is the constant correlation shrinkage target covariance matrix given by

$$\left[\mathbf{T}_2\right]_{ij} = \begin{cases} \left[\hat{\mathbf{C}}\right]_{ii}, & \text{for } i = j \\ \bar{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}, & \text{for } i \neq j \end{cases}, \tag{20}$$

with optimal shrinkage intensity determined by

$$\hat{\lambda}_{\mathbf{T}_2} = \frac{\sum_{i\neq j} \widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right) - \bar{\rho} f_{ij}}{\sum_{i\neq j}\left(\left[\hat{\mathbf{C}}\right]_{ij} - \bar{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}\right)^2}, \tag{21}$$

where $\bar{\rho}$ is the average correlation of the off-diagonal elements in the sample covariance matrix estimate computed as

$$\bar{\rho} = \frac{1}{n_s(n_s-1)} \sum_{i=1}^{n_s}\sum_{j\neq i}^{n_s} \frac{\left[\hat{\mathbf{C}}\right]_{ij}}{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}, \tag{22}$$

and the parameter f_{ij} is given by

$$f_{ij} = \frac{1}{2}\left\{ \sqrt{\frac{\left[\hat{\mathbf{C}}\right]_{jj}}{\left[\hat{\mathbf{C}}\right]_{ii}}} \widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) + \sqrt{\frac{\left[\hat{\mathbf{C}}\right]_{ii}}{\left[\hat{\mathbf{C}}\right]_{jj}}} \widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right\}. \tag{23}$$

The shrinkage intensity is optimal when the value minimizes the risk function in (60). The expressions of $\mathbf{T}_1$, $\hat{\lambda}_{\mathbf{T}_1}$, $\mathbf{T}_2$, and $\hat{\lambda}_{\mathbf{T}_2}$ are derived in the appendix.

IV. The Posterior Cramer-Rao Lower Bound

The PCRLB provides a lower bound on the mean square error obtained with any non-linear filter and is equivalent to the inverse of the posterior Fisher information matrix (PFIM) [27]. The implementation of the PCRLB requires knowledge of the true state. However, an EKF and UKF based method [28] can be applied to compute the PCRLB formulation in [29] using the mean and the covariance of the online state estimates.

A. The Extended Kalman Filter

The state vector $\mathbf{x}_k$ at time k is estimated based on the n_s received sensor measurements, $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^{n_s}]^T$. Consider the state model given in (2) and the measurement model in (7) where the parameters $\mathbf{Q}_k$ and $\mathbf{R}_k$ represent the process noise covariance matrix and the measurement noise covariance matrix, respectively. The non-linear function $\mathbf{h}(\mathbf{x}_k)$ in (7) relates the state to the measurements. By using the EKF method, the estimate of the state $\hat{\mathbf{x}}_k$ and the state covariance $\mathbf{P}_k$ at time k are calculated recursively as follows [30], [31].

The predicted state vector $\hat{\mathbf{x}}_{k|k-1}$ and the predicted covariance matrix $\mathbf{P}_{k|k-1}$ are given respectively by

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_u \mathbf{u}_k, \tag{24}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1}\mathbf{F}_k^T + \mathbf{B}_u \mathbf{Q}_k \mathbf{B}_u^T, \tag{25}$$

where $\mathbf{F}_k$ represents the state transition matrix, $\mathbf{B}_u$ denotes the command matrix, and $\mathbf{u}_k$ is the command process. Then, the posterior state vector $\hat{\mathbf{x}}_{k|k}$ and the posterior covariance matrix $\mathbf{P}_{k|k}$ are computed using

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \tag{26}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T, \tag{27}$$

where

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k, \tag{28}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T \mathbf{S}_k^{-1}. \tag{29}$$

The difference between the measurement $\mathbf{z}_k$ and the predicted measurement $\hat{\mathbf{z}}_{k|k-1}$ is called innovation process and the process improves the prior state estimate. The Kalman gain $\mathbf{K}_k$ is the correction factor and $\mathbf{S}_k$ represents the uncertainty of the predicted output. The parameter $\mathbf{H}_k$ represents the Jacobian matrix of the expected measurements $\mathbf{h}(\mathbf{x}_k)$ of a mobile user from all sensors. The components of the Jacobian matrix are

$$\mathbf{H}_k = \frac{10\beta}{\ln 10}\begin{bmatrix} \frac{\hat{x}-x_1}{(\hat{x}-x_1)^2+(\hat{y}-y_1)^2} & 0 & 0 & \frac{\hat{y}-y_1}{(\hat{x}-x_1)^2+(\hat{y}-y_1)^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\hat{x}-x_{n_s}}{(\hat{x}-x_{n_s})^2+(\hat{y}-y_{n_s})^2} & 0 & 0 & \frac{\hat{y}-y_{n_s}}{(\hat{x}-x_{n_s})^2+(\hat{y}-y_{n_s})^2} & 0 & 0 \end{bmatrix}. \tag{30}$$

obtained by taking the derivative of $\mathbf{h}(\mathbf{x}_k)$ with respect to the coordinates x_k and y_k and this matrix has the form shown in (30) at the top of the page.

B. Recursive Computation of Filtering Information Matrix

Let $\hat{\mathbf{x}}_k$ be an unbiased estimate of the state vector $\mathbf{x}_k$, based on a set of sensor measurements that are collected up to time k, i.e., $\mathbf{Z}_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$. The error covariance matrix of $\hat{\mathbf{x}}_k$ is lower bounded by

$$\mathbf{P}_k = \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] \geq \mathbf{J}_k^{-1}, \tag{31}$$

where $\mathbf{J}_k \in \mathbb{R}^{n_s \times n_s}$ is the Fisher information matrix (FIM). The FIM is given by

$$\mathbf{J}_k = \mathbb{E}[[\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)][\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)]^T], \tag{32}$$

or equivalently as

$$\mathbf{J}_k = -\mathbb{E}[\nabla_{\mathbf{X}_k}[\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)]^T], \tag{33}$$

where $\nabla_{\mathbf{X}_k}$ is the first-order partial derivative operator with respect to $\mathbf{X}_k$. The joint probability distribution of $\mathbf{X}_k = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k\}$ and $\mathbf{Z}_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ computed as [28]

$$p(\mathbf{X}_k, \mathbf{Z}_k) = p(\mathbf{x}_0) \prod_{i=1}^{k} p(\mathbf{z}_i|\mathbf{x}_i) \prod_{j=1}^{k} p(\mathbf{x}_j|\mathbf{x}_{j-1}) \tag{34}$$

is determined by the prior density function $p(\mathbf{x}_0)$ of the target initial state $\mathbf{x}_0$ and the conditional density functions $p(\mathbf{z}_i|\mathbf{x}_i)$ and $p(\mathbf{x}_j|\mathbf{x}_{j-1})$, respectively. Tichavsky *et al.* [29] propose a method of computing the FIM recursively as

$$\mathbf{J}_{k+1} = \mathbf{D}_k^{22} - \mathbf{D}_k^{21}(\mathbf{J}_k + \mathbf{D}_k^{11})^{-1}\mathbf{D}_k^{12}, \tag{35}$$

where the terms in (35) are defined as

$$\mathbf{D}_k^{11} = -\mathbb{E}[\nabla_{\mathbf{x}_k}[\nabla_{\mathbf{x}_k} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)]^{\mathrm{T}}], \tag{36}$$

$$\mathbf{D}_k^{21} = -\mathbb{E}[\nabla_{\mathbf{x}_k}[\nabla_{\mathbf{x}_{k+1}} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)]^{\mathrm{T}}], \tag{37}$$

$$\mathbf{D}_k^{12} = -\mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_k} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)]^{\mathrm{T}}] = [\mathbf{D}_k^{21}]^{\mathrm{T}}, \tag{38}$$

$$\mathbf{D}_k^{22} = -\mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_{k+1}} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)]^{\mathrm{T}}] - \mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_{k+1}} \log\, p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})]^{\mathrm{T}}], \tag{39}$$

and the state and measurement density functions satisfy the following:

$$\nabla_{\mathbf{x}_k} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k) = [\nabla_{\mathbf{x}_k} \mathbf{f}_k^T(\mathbf{x}_k)]\ \mathbf{Q}_k^{-1}\ [\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k)], \tag{40}$$

$$\nabla_{\mathbf{x}_k} \log\, p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) = [\nabla_{\mathbf{x}_{k+1}} \mathbf{h}_{k+1}^T(\mathbf{x}_k)]\ \mathbf{R}_{k+1}^{-1} \times [\mathbf{z}_{k+1} - \mathbf{h}_{k+1}(\mathbf{x}_{k+1})], \tag{41}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_{k+1}$ are the process noise covariance matrix and the measurement noise covariance matrix, respectively. By assuming that the covariance matrices are invertible, the matrices defined in (36) − (39) are simplified as follows

$$\mathbf{D}_k^{11} = \mathbb{E}[[\nabla_{\mathbf{x}_k} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)][\nabla_{\mathbf{x}_k} \log\, p(\mathbf{x}_{k+1}|\mathbf{x}_k)]^{\mathrm{T}}], = \mathbb{E}[[\nabla_{\mathbf{x}_k} \mathbf{f}_k^T(\mathbf{x}_k)]\ \mathbf{Q}_k^{-1}\ [\nabla_{\mathbf{x}_k} \mathbf{f}_k^T(\mathbf{x}_k)]^T], = \mathbb{E}[\mathbf{F}_k^T \mathbf{Q}_k^{-1} \mathbf{F}_k], \tag{42}$$

$$\mathbf{D}_k^{12} = -\mathbb{E}[\mathbf{F}_k^T]\mathbf{Q}_k^{-1}, \tag{43}$$

$$\mathbf{D}_k^{21} = \mathbf{Q}_k^{-1} - \mathbb{E}[\mathbf{F}_k], \tag{44}$$

$$\mathbf{D}_k^{22} = \mathbf{Q}_k^{-1} + \mathbb{E}[[\nabla_{\mathbf{x}_k} \mathbf{h}_k^T(\mathbf{x}_k)]\ \mathbf{R}_{k+1}^{-1}\ [\nabla_{\mathbf{x}_k} \mathbf{h}_k^T(\mathbf{x}_k)]^T], = \mathbf{Q}_k^{-1} + \mathbb{E}[\mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}], \tag{45}$$

where $\mathbf{F}_k$ represents the state transition matrix and $\mathbf{H}_{k+1}$ is the Jacobian matrix evaluated at $\mathbf{x}_{k+1}$.

C. The PCRLB for a Deterministic Trajectory

For models (2) and (7), obtaining a closed-form solution to the FIM is non-trivial. Therefore, we consider the case in which the trajectory of the mobile user is generated in a deterministic way. Hence, the process noise is zero and the expectation operator in (42) − (45) can be dropped out. The recursive equation in (35) can be rewritten as

$$\mathbf{J}_{k+1} = \mathbf{Q}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1} - \mathbf{Q}_k^{-1} \mathbf{F}_k \Big(\mathbf{J}_k + \mathbf{F}_k^T \mathbf{Q}_k^{-1} \mathbf{F}_k\Big)^{-1} \times \mathbf{F}_k^T \mathbf{Q}_k^{-1}. \tag{46}$$

By applying the matrix inversion lemma, it yields

$$\mathbf{J}_{k+1} = \Big(\mathbf{Q}_k + \mathbf{F}_k \mathbf{J}_k^{-1} \mathbf{F}_k^T\Big)^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}. \tag{47}$$

Since $\mathbf{Q}_k = 0$, (47) becomes

$$\mathbf{J}_{k+1} = \Big[\mathbf{F}_k^{-1}\Big]^T \mathbf{J}_k \mathbf{F}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}. \tag{48}$$

After comparing (48) with the EKF covariance matrix computed in (27), by replacing $\mathbf{J}_k$ by $\mathbf{P}_k^{-1}$ and by applying the matrix inversion lemma, the following expression is obtained:

$$\mathbf{P}_{k+1}^{-1} = \Big(\mathbf{F}_k \mathbf{P}_k \mathbf{F}_k\Big)^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}. \tag{49}$$

In (49), matrices $\mathbf{F}_k$ and $\mathbf{H}_{k+1}$ are evaluated at the true state of the state vector but in the EKF equations, the matrices are evaluated at their estimated state vector.

V. Particle Filtering With Shrinkage for Dealing With Correlated Measurements

The PF estimates the state of the mobile user $\mathbf{x}_k$ recursively from the conditional probability density function $p(\mathbf{x}_k|\mathbf{Z}_{1:k})$ and a set of sensor measurements, $\mathbf{Z}_{1:k} = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ up to

time k via the Chapman-Kolmogorov equation and Bayes' rule given by

$$p(\mathbf{x}_k|\mathbf{Z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{Z}_{1:k-1})d\mathbf{x}_{k-1}, \quad (50)$$

$$p(\mathbf{x}_k|\mathbf{Z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{Z}_{1:k-1})}, \quad (51)$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function and $p(\mathbf{z}_k|\mathbf{Z}_{1:k-1})$ is the normalizing constant. Solving the integral in (50) is intractable in general and hence, it is approximated with the sequential important sampling PF [32].

The posterior density function is approximated by a set of particles $\{\hat{\mathbf{x}}_k^{(i)},\ i = 1, \ldots, N_p\}$ with its associated weights $\{W_k^{(i)},\ i = 1, \ldots, N_p\}$ where N_p is the total number of particles, given by

$$\hat{p}(\mathbf{x}_k|\mathbf{Z}_{1:k}) = \sum_{i=1}^{N_p} \widehat{W}_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}), \quad (52)$$

where $\delta(\cdot)$ is the Dirac delta measure and

$$W_k^{(i)} \propto W_{k-1}^{(i)} \mathcal{L}(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}), \quad (53)$$

is the weight of the i-th particle and is normalized such that

$$\widehat{W}_k^{(i)} = \frac{W_k^{(i)}}{\sum_{i=1}^{N_p} W_k^{(i)}}. \quad (54)$$

The PF works based on three stages: the prediction stage, the measurement stage and the resampling stage. During the prediction stage, each particle transition state is propagated according to the user mobility model. In the measurement stage, the weight of each particle is evaluated using the likelihood function. Finally, the resampling stage replaces one set of particles and their weights with another set. The resampling step is essential to avoid particle degeneracy [18] and in this paper, the residual resampling algorithm [19], [33], is applied.

A. Likelihood Function of the Shrinkage Particle Filter

The likelihood function of the ShPF in (53) is given by

$$\mathcal{L}\Big(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}\Big) = \Big((2\pi)^{n_s}|\mathbf{S}_k|\Big)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})\mathbf{S}_k^{-1}(\mathbf{z}-\hat{\mathbf{z}})^T\Big), \quad (55)$$

where $\mathbf{z}$ and $\hat{\mathbf{z}}$ represent the actual and the predicted RSS measurements respectively, n_s is the number of sensors, and $\mathbf{S}_k \in \mathbb{R}^{n_s \times n_s}$ is the shadowing noise covariance matrix at time instant k. The covariance matrix in the likelihood function is estimated by the shrinkage estimator in (17) and is expected to improve the tracking performance of the PF with respect to the tracking accuracy. By defining the size of the sliding window time, denoted as t_{window}, the computation of the likelihood in (55) involves a modified covariance matrix of size $n_s(t_{window}+1) \times n_s(t_{window}+1)$ [34].

B. A Particle Filter with Shrinkage Algorithm

The ShPF is developed based on the PF model presented in [20]. Algorithm 1 describes the proposed ShPF.

Algorithm 1: Shrinkage-Based PF

(1) Initialization
for $k = 0$ and $i = 1, \ldots, N_p$ **do**
  Samples from initial estimate $\hat{\mathbf{x}}_0^{(i)} \sim p(\hat{\mathbf{x}}_0)$.
  Assign initial important weights $W_0^{(i)} = 1/N_p$.
end for
for $k = 1, \ldots, endtime$ **do**
  (2) Shrinkage Covariance Estimator
  Estimate the shadowing noise covariance matrix using the shrinkage estimator $\hat{\mathbf{S}} = \lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}}$.
  for $i = 1, \ldots, N_p$ **do**
    (3) Prediction Step
    Propagate the samples $\hat{\mathbf{x}}_k^{(i)} = \mathbf{A}(T,\alpha)\hat{\mathbf{x}}_{k-1}^{(i)} + \mathbf{B}_u(T)\mathbf{u}_k^{(i)} + \mathbf{B}_w(T)\mathbf{w}_k^{(i)}$ with noise realizations $\mathbf{w}_k^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$.
    (4) Measurement Update
    Compute the weights $W_k^{(i)} \propto W_{k-1}^{(i)} \mathcal{L}(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)})$.
    The likelihood function is calculated using

$$\mathcal{L}\Big(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}\Big) = \Big((2\pi)^{n_s}|\hat{\mathbf{S}}_k|\Big)^{-\frac{1}{2}} \times \exp\Big(-\frac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})\hat{\mathbf{S}}_k^{-1}(\mathbf{z}-\hat{\mathbf{z}})^T\Big).$$

  end for
  Normalize the weights $\widehat{W}_k^{(i)} = W_k^{(i)} / \sum_{i=1}^{N_p} W_k^{(i)}$, $i = 1, \ldots, N_p$.
  (5) Output Estimate
  The estimated state is $\hat{\mathbf{x}}_k = \sum_{i=1}^{N_p} \widehat{W}_k^{(i)} \hat{\mathbf{x}}_k^{(i)}$, $i = 1, \ldots, N_p$.
  (6) Resampling Step
  Set the threshold sample size $N_{thresh} = N_p/10$.
  Calculate the effective sample size $N_{eff} = 1/\sum_{i=1}^{N_p} (\widehat{W}_k^{(i)})^2$.
  Resampling **if** $N_{eff} < N_{thresh}$ **then**
    Multiply/suppress particles with high/low importance weights in order to obtain N_p new random particles approximately distributed according to the posterior state distribution. The residual resampling algorithm is applied [19], [33].
  end if
end for

VI. Performance Evaluation

The tracking accuracy of the proposed ShPF is compared with the PF without the shrinkage over simulated data and real Wi-Fi data. The root mean square error (RMSE) is used as the performance metric for assessing the accuracy of the state estimates. The coordinate RMSE is given by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\hat{x}_k^i - x_k^i)^2 + (\hat{y}_k^i - y_k^i)^2} \quad (56)$$

where $\{\hat{x}_k, \hat{y}_k\}$ is the estimated trajectory and $\{x_k, y_k\}$ is the actual trajectory, collected up to time k, and N is the number of simulation runs.

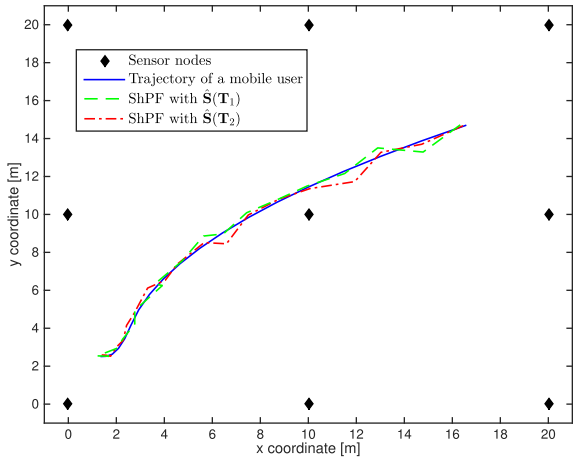

Fig. 1. Coordinate of the sensor nodes, actual trajectory of the mobile user, and estimated trajectories by the ShPF from a single realization.

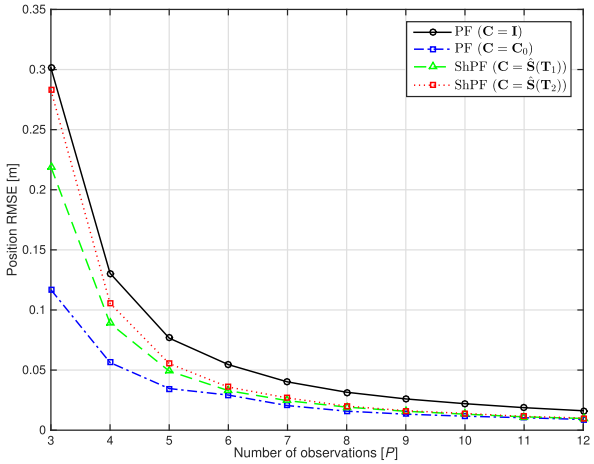

Fig. 2. Coordinate RMSE comparison of the PF and ShPF using the simulated RSS measurements.

A. Simulation Results

The simulated network area contains nine sensor nodes ($n_s = 9$) with a coverage radius of 5 meters, as shown in Figure 1. In order to maintain coverage and reduce the tracking error, all sensor nodes are able to move towards the mobile user but do not cross their designated grid. The speed of the sensor nodes varies within the specified range of $(0.05 - 0.15)$ ms^{-1}. The mobile user can move to any part of the network with varying speed and acceleration. The simulated trajectory of the mobile user is generated deterministically according to (2) and, with this trajectory, the sensor measurements are randomly generated according to (7) with different noise realizations for each simulation run. The command processes $u_{x,k+1}$ and $u_{y,k+1}$ in the filters is assumed to be a Markov chain, taking values between the following acceleration levels

$$\mathcal{M}_x = \{0.0,\ 0.5,\ 0.0\}, \tag{57}$$

$$\mathcal{M}_y = \{0.0,\ 0.0,\ 0.8\}, \tag{58}$$

in units of ms^{-2}. Initial mode probabilities are $\mu_{1,0} = 0.8$ and $\mu_{i,0} = 0.1$ for $i = 2, 3$. The transition probability matrix $\prod$ has the following diagonal elements ($[\prod]_{ii} = 0.5$, for $i = 1, 2, 3$) and the off-diagonal elements are ($[\prod]_{ij} = 0.25$, for $i, j = 1, 2, 3$). The sum of elements in each row of the matrix is equal to one. The sensor measurements are assumed to be correlated according to (13) with $t_{window} = 1$ is used. The number of simulation runs is 100 and an average of the RMSE is calculated. The ShPF takes 1.37 seconds to complete a single run in the simulation. The performance validation is carried out by means of a desktop computer with an Intel core 3.3 GHz processor, 4 GB RAM, and 465 GB hardrive. The simulation parameters of the ShPF and the respective Singer model are given in Table I. Figure 1 presents the actual and estimated trajectories of the mobile user over a single simulation run. The true initial state of the simulated trajectory is set to $\mathbf{x}_0 = [1.3,\ 0.02,\ 0,\ 2.5,\ 0.02,\ 0]^T$.

Figure 2 shows the RMSE comparison of the PF and ShPF algorithms. The PF operates with shadowing noise covariance matrix $\mathbf{C} = \mathbf{I}$ and the ShPF operates with shadowing noise covariance matrix given by (18) with $\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$ and by (20) with $\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$, respectively. Let $\mathbf{C}_0$ be the true shadowing noise covariance matrix, $\hat{\mathbf{S}}(\mathbf{T}_1)$ is the estimated shrinkage covariance matrix based on the target matrix $\mathbf{T}_1$, and $\hat{\mathbf{S}}(\mathbf{T}_2)$ is the estimated shrinkage covariance matrix based on the target matrix $\mathbf{T}_2$. It is shown that for all number of observations, the ShPF has a smaller RMSE when compared to the PF. This is because the elements of the shadowing noise covariance matrix $\mathbf{C}$, employed in the ShPF is estimated by capturing the correlation that present between the measurements while the PF does not. The tracking accuracy improves significantly especially when the ShPF deals with a small number of observations, i.e. $P < 8$. For a larger number of observations, ($P > 8$) there is no obvious difference in the RMSE of the coordinate estimate between the ShPF and the PF. The ShPF with the target matrix $\mathbf{T}_1$ achieves a lower RMSE value compared to that with the target matrix $\mathbf{T}_2$, especially when it operates with a limited number of observations ($P \leq 5$). This is because when the target matrix $\mathbf{T}_1$ is employed in (17), the shrinkage covariance matrix estimate $\hat{\mathbf{S}}$, shrinks with equally weighted terms toward both $\mathbf{T}_1$ and $\hat{\mathbf{C}}$. However, when the target matrix $\mathbf{T}_2$ is employed in (17), the shrinkage covariance matrix estimate $\hat{\mathbf{S}}$, shrinks more towards $\mathbf{T}_2$ and less towards $\hat{\mathbf{C}}$. The optimal shrinkage intensity in (19) is calculated by taking all elements of $\hat{\mathbf{C}}$. On the other hand, the optimal shrinkage intensity in (21) is calculated by taking only the off-diagonal elements of $\hat{\mathbf{C}}$.

TABLE I
SIMULATION PARAMETERS FOR TRACKING OF A MOBILE USER

Discretisation time step T	0.5 s
Minimum speed of a mobile user V_{min}	0.3 ms^{-1}
Maximum speed of a mobile user V_{max}	5.4 ms^{-1}
Reciprocal maneuver time constant α	0.6
Number of Particles N_P	500
Standard deviations of the noise σ_k^i, σ_k^j in (9)	[0 – 4] dB
Standard deviations of the noise $\tilde{\sigma}_k^i, \tilde{\sigma}_l^i$ in (11)	[0 – 4] dB
Path loss exponent β	3
Decorrelation distance D_c	40 m

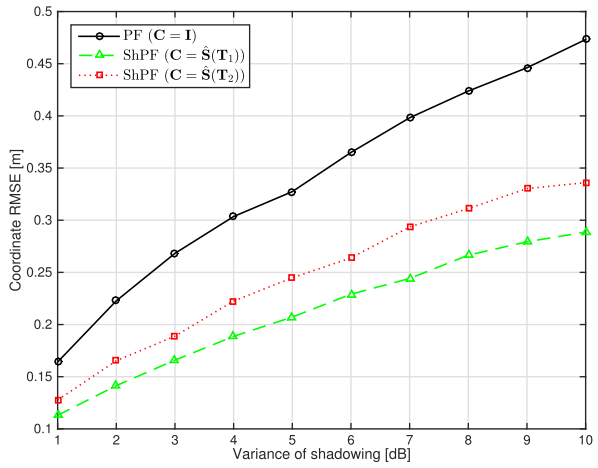

Fig. 3. Coordinate RMSE comparison of the PF and ShPF for different values of shadowing variance used in the RSS measurements.

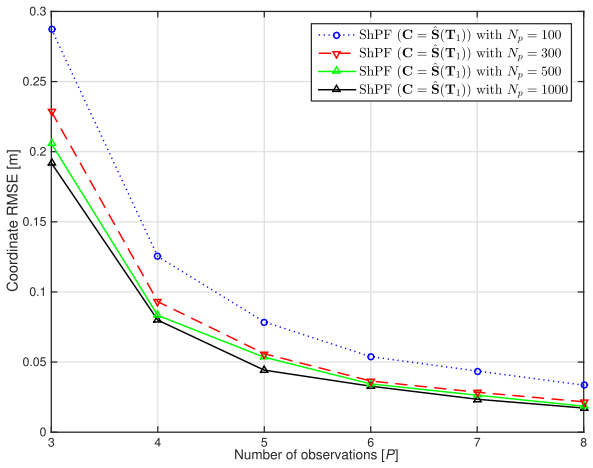

Fig. 5. Coordinate RMSE of the ShPF with $N_p = 100,\ 300,\ 500$ and 1000 particles.

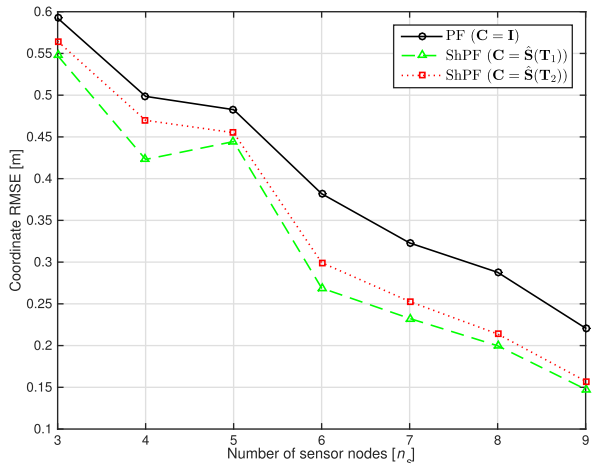

Fig. 4. Coordinate RMSE comparison of the PF and ShPF for different numbers of sensor nodes (anchors) in the network.

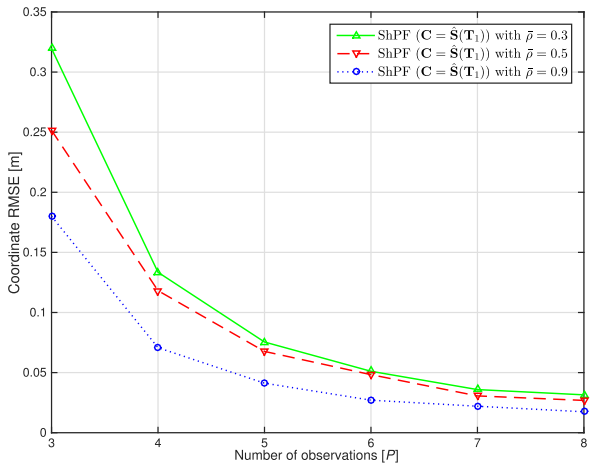

Fig. 6. Coordinate RMSE of the ShPF with the correlation between the RSS measurements is 0.3, 0.5 and 0.9.

In Figure 3, the RMSEs of the coordinate estimate between the PF and the ShPF are compared using different values of shadowing variance in the RSS measurements. Note that when the value of the shadowing variance increases, the tracking accuracy decreases for all the approaches. However, the proposed ShPF gives a better estimate than the PF without the shrinkage in tracking the mobile user. The ShPF with the target matrix $\mathbf{T}_1$, has a lower tracking error when compared to that with the target matrix $\mathbf{T}_2$.

Figure 4 shows the RMSE of the coordinate estimate for a different number of sensor nodes for all the approaches. The RMSE of the coordinate estimation improves when the number of sensor nodes increases as a result of having more data available in the estimation process. The ShPF with the target matrix $\mathbf{T}_1$ has a slightly smaller coordinate RMSE when compared to the case with the target matrix $\mathbf{T}_2$.

The trade-off between the accuracy and the number of particles used in the ShPF is presented in Figure 5. On average, when the ShPF operates with $N_p = 100$ particles, the accuracy of tracking is reduced by almost 45% when compared with the ShPF with $N_p = 500$ particles for small number of observations ($P = 3$). When the ShPF operates with 1000 particles, the tracking accuracy increases but at the expense of increasing the computation time.

The effect of the correlation between the RSS measurements in terms of coordinate RMSE is observed in Figure 6. The spatial correlation in (8) and the temporal correlation in (10) are set to 0.3, 0.5 and 0.9, respectively. For the highest correlation value between measurements, the ShPF displays the lowest RMSE for the coordinate estimates of the mobile user. However, when the correlation between the measurements is weak, the tracking accuracy of the ShPF is also reduced. The correlation between the measurements contributes to the optimal shrinkage intensity calculated in (19) for the ShPF with the target matrix $\mathbf{T}_1$.

In Figure 7, the ShPF with $t_{window} = 10$ gives a better tracking performance than the ShPF with $t_{window} = 1$. In the simulation, the measurements are generated every 0.01 seconds. Thus, every second approximately one hundred measurements are acquired. When $t_{window} = 1$, only ten measurements are held per second, while for $t_{window} = 10$, the number of measurements increases to one hundred. As a result,

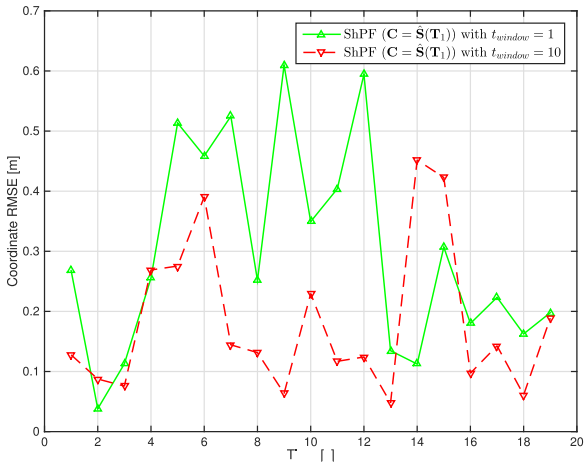

Fig. 7. Coordinate RMSE comparison of the ShPF with window sizes $t_{window} = 1$ and $t_{window} = 10$.

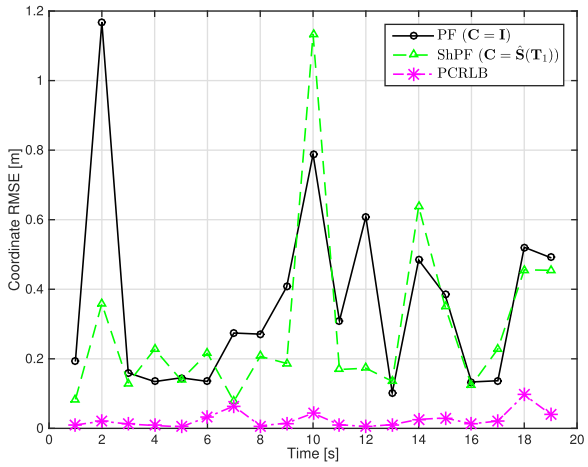

Fig. 8. Coordinate RMSE comparison of the PF, ShPF, and PCRLB.

the estimation of the temporal correlation by the filtering algorithm is improved by using more history of the past target positions. However, this increases the computation time of the tracking algorithm due to the increased complexity in evaluating the likelihood function by using the inversion of large covariance matrix. The ShPF with $t_{window} = 10$ takes 5.97 seconds to complete a single run. Meanwhile, the ShPF with $t_{window} = 1$ only takes 2.33 seconds to complete. Thus, the sliding window time needs to be chosen in a way that the size of the measurements vector is not large. The size of the sliding window time also depends on the environment. In urban environments the correlation coefficient varies more than in suburban and rural environments. To preserve the statistical structure in noisy scenarios a sliding window is used to encompass several measurements. The size of the sliding window imposes a trade-off between the tracking accuracy and the computational complexity of the tracking algorithm.

Finally, Figure 8 shows a comparison of the coordinate RMSE of the PF, ShPF, and PCRLB. The coordinate RMSE of the PF and the ShPF are measured using (56). Meanwhile,

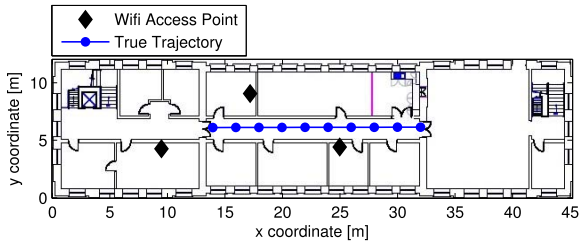

Fig. 9. Experimental setup.

the PCRLB is calculated using

$$RMSE_{PCRLB} = \sqrt{(\mathbf{P}_k(1,1) + \mathbf{P}_k(4,4))}, \tag{59}$$

where $\mathbf{P}_k$ is the covariance matrix of the EKF with the Jacobian evaluated at the estimated state vector $\hat{\mathbf{x}}_k$ at time instant k. It is shown that the ShPF provides a smaller RMSE than the PF without the shrinkage most of the time. This is because the proposed ShPF method operates with a better estimates of the shadowing noise covariance matrix in its operation than the PF, which does not consider the correlation between the measurements. On the other hand, the PCRLB sets a lower bound on the estimated covariance and provides a useful benchmark against the tracking performance of the standard PF and the ShPF. For that reason, the standard PF is considered as the baseline method in this case.

B. Experimental Results

The performance of the ShPF has been investigated in a Wi-Fi network, with real Wi-Fi signal strengths collected from the D floor of the Amy Johnson Building, at the University of Sheffield, United Kingdom. For this experiment, a user carries a mobile smartphone moving from one end of the corridor to the other end. There are three Wi-Fi access points available in the floor where the user moves. A Xiaomi mobile smartphone running an Android 4.4.2 operating system is installed with the Sensor Fusion App and it is used as a receiver to collect the transmitted Wi-Fi signals from all three access points. The App is developed by LinkÃping University and can be downloaded for free on Google Play [35]. Figure 9 presents the coordinates of the Wi-Fi access points and the true trajectory of a mobile user that are superimposed on the layout of the building floor. The size of the building floor area is 414.74 square meters with the black diamonds representing the coordinates of the Wi-Fi access points. A total of ten point coordinates have been identified in the user trajectory for data collection, denoted by blue circles, and each point coordinate is separated by 2 meters. As the user moves, the Wi-Fi signals are collected with their corresponding measurement noise at each point coordinate in the trajectory. The PF and ShPF algorithms estimate the coordinate of the mobile user using the collected Wi-Fi signal strengths. The RMSE is again used as a performance metric to compare the ShPF and the PF methods.

Figure 10 plots the transmitted Wi-Fi measurements from three Wi-Fi access points located in three different rooms: room D02, D06, and D08 of the building. These signals are received at ten different point coordinates in the trajectory.

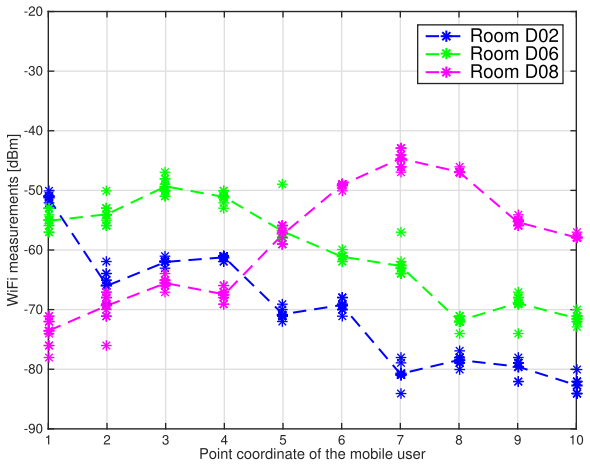

Fig. 10. The recorded Wi-Fi measurements from the testbed taken at Room D02, D06, and D08 of the Amy Johnson Building.

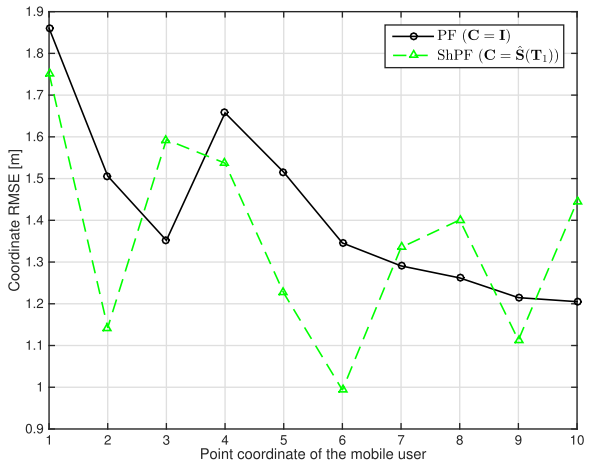

Fig. 11. Coordinate RMSE comparison of the PF and ShPF using the Wi-Fi measurements.

At each point coordinate, there are a maximum of $P = 8$ observations available to be processed by the tracking algorithms with their mean values indicated by the plotted lines. The variations in the observations are due to the multipath fading effects. The Wi-Fi signal strength increases when the receiver approaches the transmitter (Wi-Fi access points) and the signal strength decreases when the distance between the transmitter and receiver increases.

Figure 11 compares the RMSE of the coordinate estimates that are achieved by the PF and ShPF at all point coordinates in the user trajectory. The PF is operates with the identity covariance matrix by assuming the filter does not know the correlation between the measurements. On the other hand, the ShPF operates with the shrinkage covariance matrix to exploit the correlation between the measurements. Both tracking methods are assumed to know the starting coordinate of the mobile user in the trajectory. The filtering algorithms have to infer the next coordinate of the mobile user using the prior information and the Wi-Fi measurements. The path loss exponent is set to $\beta = 3$ and the signal power loss at 1 m distance is $z^0 = -36$ dBm. In indoor environments, the path loss exponent is affected by the layout of the building floor such as the location of walls, doors, chairs and even the construction materials in the building. It shows that the ShPF outperforms the tracking performance of the PF without the shrinkage for most of the coordinate estimates. For some coordinates, the ShPF tracking accuracy increases by 26% when compared with the PF without the shrinkage. However, for some point coordinates, the improvement is not significant because the correlation between the Wi-Fi measurements is less than 0.6. The ShPF works best when the Wi-Fi measurements are highly correlated. Finally, it cannot be ruled out that part of the error in tracking is also caused by the uncertainty in determining the exact coordinate of the Wi-Fi access points in the building.

VII. Conclusions

This paper presents a new approach to target tracking in a wireless sensor network by combining the particle filter and the shrinkage estimation technique. The shrinkage technique is shown to improve the estimate of the shadowing noise covariance matrix. The shadowing noise covariance matrix is used in a particle filter for recursive updates of the likelihood. The proposed shrinkage-based particle filter works effectively with correlated and limited measurements. In dense wireless sensor networks, the measurements obtained by sensors are spatio-temporally correlated. Exploiting this spatio-temporal correlation improves the performance of target detection and tracking. However, the high dimension of the covariance matrix may lead to challenges from computational point of view. Simulation and experimental results have shown that the proposed method improves the tracking accuracy when compared to the commonly used particle filter methods. The posterior Cramer Rao lower bound is also calculated for simulated data to compare it with the root mean square error of the shrinkage-based particle filter method.

Appendix

Consider the problem of estimating the true covariance matrix $\mathbf{C}$ by and an estimator $\hat{\mathbf{S}}$, where λ minimizes the risk function

$$R(\lambda) = \mathbb{E}\|\hat{\mathbf{S}} - \mathbf{C}\|_F^2. \tag{60}$$

This implies that

$$R(\lambda) = \sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{S}}_p) + [\mathbb{E}(\hat{\mathbf{S}}_p) - \mathbf{C}_p]^2 \tag{61}$$

$$= \sum_{p=1}^{P} \mathrm{Var}(\lambda \mathbf{T}_p + (1-\lambda)\hat{\mathbf{C}}_p) + [\mathbb{E}(\lambda \mathbf{T}_p + (1-\lambda)\hat{\mathbf{C}}_p) - \mathbf{C}_p]^2 \tag{62}$$

$$= \sum_{p=1}^{P} \lambda^2 \mathrm{Var}(\mathbf{T}_p) + (1-\lambda)^2 \mathrm{Var}(\hat{\mathbf{C}}_p) + 2\lambda(1-\lambda) \times \mathrm{Cov}(\hat{\mathbf{C}}_p, \mathbf{T}_p) + [\lambda \mathbb{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p) + \mathrm{Bias}(\hat{\mathbf{C}})]^2. \tag{63}$$

Taking the derivative with respect to λ and setting equal to zero yields

$$R'(\lambda) = \sum_{p=1}^{P} 2\lambda \mathrm{Var}(\mathbf{T}_p) + 2(1-\lambda)\mathrm{Var}(\hat{\mathbf{C}}_p) + 2(1-2\lambda) \times \mathrm{Cov}(\hat{\mathbf{C}}_p, \mathbf{T}_p) + 2[\mathbb{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p)] \times [\lambda \mathbb{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p) + \mathrm{Bias}(\hat{\mathbf{C}})], \quad (64)$$

which leads the optimal shrinkage intensity given by

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{C}}_p) - \mathrm{Cov}(\mathbf{T}_p, \hat{\mathbf{C}}_p) - \mathrm{Bias}(\hat{\mathbf{C}}_p)\mathbb{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathbb{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}. \quad (65)$$

If $\hat{\mathbf{C}}$ is an unbiased estimator of $\mathbf{C}$, then the optimal shrinkage intensity becomes

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{C}}_p) - \mathrm{Cov}(\mathbf{T}_p, \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathbb{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}, \quad (66)$$

and if $\widehat{\mathrm{Var}}(\hat{\mathbf{C}}_p)$ and $\widehat{\mathrm{Cov}}(\mathbf{T}_p, \hat{\mathbf{C}}_p)$ are unbiased estimators of $\mathrm{Var}(\hat{\mathbf{C}}_p)$ and $\mathrm{Cov}(\mathbf{T}_p, \hat{\mathbf{C}}_p)$, then the optimal shrinkage intensity is given by

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \widehat{\mathrm{Var}}(\hat{\mathbf{C}}_p) - \widehat{\mathrm{Cov}}(\mathbf{T}_p, \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathbb{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}. \quad (67)$$

To obtain the expressions of $\widehat{\mathrm{Var}}(\hat{\mathbf{C}})$ and $\widehat{\mathrm{Cov}}(\mathbf{T}, \hat{\mathbf{C}})$, we defined the following: For P observations of sensor measurements z_i, the sample mean is given by $\bar{z}_i = P^{-1}\sum_{p=1}^{P} z_{ip}$ for $p = 1, \ldots, P$. Let

$$v_{ijp} = (z_{ip} - \bar{z}_i)(z_{jp} - \bar{z}_j), \quad (68)$$

$i, j = 1, \ldots, n_s$ and $p = 1, \ldots, P$ be random variables with a sample mean given by $\bar{v}_{ij} = P^{-1}\sum_{p=1}^{P} v_{ijp}$. Then, the sample covariance matrix estimate is given by

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{1}{P-1}\sum_{p=1}^{P}(z_{ip} - \bar{z})(z_{jp} - \bar{z}). \quad (69)$$

The unbiased variance of individual elements of $\hat{\mathbf{C}}$ is given by [36]

$$\widehat{\mathrm{Var}}\left[\hat{\mathbf{C}}\right]_{ij} = \frac{P^2}{(P-1)^2}\widehat{\mathrm{Var}}(\bar{v}_{ij}) \quad (70)$$

$$= \frac{P^2}{(P-1)^2}\left[\frac{1}{P}\widehat{\mathrm{Var}}(v_{ij})\right] \quad (71)$$

$$= \frac{P^2}{(P-1)^2}\left[\frac{1}{P}\left[\frac{1}{P-1}\sum_{p=1}^{P}(v_{ijp} - \bar{v}_{ij})^2\right]\right] \quad (72)$$

$$= \frac{P}{(P-1)^3}\sum_{p=1}^{P}(v_{ijp} - \bar{v}_{ij})^2. \quad (73)$$

Let v_{kl} be another random variable where the sample mean of the variable is $\bar{v}_{kl}$, and the covariance elements are obtained as [36]

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{kl}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}(v_{ijp} - \bar{v}_{ij})(v_{klp} - \bar{v}_{kl}). \quad (74)$$

To determine $\hat{\lambda}$ in (67) requires an expression for $\widehat{\mathrm{Cov}}(\mathbf{T}_p, \hat{\mathbf{C}}_p)$. Consider $\left[\mathbf{T}\right]_{ij} = \bar{\rho}\sqrt{\left[\hat{\mathbf{C}}\right]_{ii}\left[\hat{\mathbf{C}}\right]_{jj}}$, for $i \neq j$ and let $\left[\bar{\mathbf{C}}\right]_{ii}$, $\left[\bar{\mathbf{C}}\right]_{jj}$ and $\left[\bar{\mathbf{C}}\right]_{ij}$ be the point estimates respectively of $\left[\mathbf{C}\right]_{ii}$, $\left[\mathbf{C}\right]_{jj}$ and $\left[\mathbf{C}\right]_{ij}$, respectively. Then, $\left[\mathbf{T}\right]_{ij}$ is expanded via Taylor series which results in

$$\left[\mathbf{T}\right]_{ij} = \bar{\rho}\sqrt{\left[\bar{\mathbf{C}}\right]_{ii}\left[\bar{\mathbf{C}}\right]_{jj}} + \frac{\bar{\rho}}{2}\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{jj}}{\left[\bar{\mathbf{C}}\right]_{ii}}}\left(\left[\mathbf{C}\right]_{ii} - \left[\bar{\mathbf{C}}\right]_{ii}\right) + \frac{\bar{\rho}}{2}\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{ii}}{\left[\bar{\mathbf{C}}\right]_{jj}}}\left(\left[\mathbf{C}\right]_{jj} - \left[\bar{\mathbf{C}}\right]_{jj}\right), \quad (75)$$

where $\bar{\rho}$ is the average correlation of all the correlation values in the sample and is obtained using (22) for $\left[\bar{\mathbf{C}}\right]_{ii}$, $\left[\bar{\mathbf{C}}\right]_{jj}$ and $\left[\bar{\mathbf{C}}\right]_{ij}$. Based on the definition of the covariance matrix, this implies that

$$\widehat{\mathrm{Cov}}\left(\left[\mathbf{T}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \mathbb{E}\left[\left(\left[\mathbf{T}\right]_{ij} - \mathbb{E}\left[\left[\mathbf{T}\right]_{ij}\right]\right) \times \left(\left[\mathbf{C}\right]_{ij} - \mathbb{E}\left[\left[\mathbf{C}\right]_{ij}\right]\right)\right], \quad (76)$$

and using (75) and (76) yields

$$\widehat{\mathrm{Cov}}\left(\left[\mathbf{T}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{\bar{\rho}}{2}\left\{\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{jj}}{\left[\bar{\mathbf{C}}\right]_{ii}}}\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) + \sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{ii}}{\left[\bar{\mathbf{C}}\right]_{jj}}}\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right)\right\}. \quad (77)$$

Finally, using (68) and (74) the covariance elements are expressed as

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left[(z_{ip} - \bar{z}_i)^2 - \bar{v}_{ii}\right] \times \left[(z_{ip} - \bar{z}_i)(z_{jp} - \bar{z}_j) - \bar{v}_{ij}\right], \quad (78)$$

and similarly

$$\widehat{\mathrm{Cov}}\left(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\right) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\left[(z_{jp} - \bar{z}_j)^2 - \bar{v}_{jj}\right] \times \left[(z_{ip} - \bar{z}_i)(z_{jp} - \bar{z}_j) - \bar{v}_{ij}\right]. \quad (79)$$

which completes the derivation.

REFERENCES

[1] X. Wang, M. Fu, and H. Zhang, "Target tracking in wireless sensor networks based on the combination of KF and MLE using distance measurements," *IEEE Trans. Mobile Comput.*, vol. 11, no. 4, pp. 567–576, Apr. 2012.

[2] S. Mahfouz, F. Mourad-Chehade, P. Honeine, J. Farah, and H. Snoussi, "Target tracking using machine learning and Kalman filter in wireless sensor networks," *IEEE Sensors J.*, vol. 14, no. 10, pp. 3715–3725, Oct. 2014.

[3] K. Radnosrati, F. Gunnarsson, and F. Gustafsson, "New trends in radio network positioning," in *Proc. 18th Int. Conf. Inf. Fusion*, Jul. 2015, pp. 492–498.

[4] S. Cho, L. Han, B. Joo, and S. Han, "P-LEACH: An efficient cluster-based technique to track mobile sinks in wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 10, no. 9, Sep. 2014.

[5] X. Wang, A. K. S. Wong, and Y. Kong, "Mobility tracking using GPS, Wi-Fi and cell ID," in *Proc. Int. Conf. Inf. Netw.*, Feb. 2012, pp. 171–176.

[6] Z. R. Zaidi and B. L. Mark, "Real-time mobility tracking algorithms for cellular networks based on Kalman filtering," *IEEE Trans. Mobile Comput.*, vol. 4, no. 2, pp. 195–208, Mar. 2005.

[7] G. M. Djuknic and R. E. Richton, "Geolocation and assisted GPS," *Computer*, vol. 34, no. 2, pp. 123–125, Feb. 2001.

[8] I. F. Akyildiz, Y.-B. Lin, W.-R. Lai, and R.-J. Chen, "A new random walk model for PCS networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1254–1260, Jul. 2000.

[9] B. L. Mark and Z. R. Zaidi, "Robust mobility tracking for cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 1, Jun. 2002, pp. 445–449.

[10] R. A. Singer, "Estimating optimal tracking filter performance for manned maneuvering targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 6, no. 4, pp. 473–483, Jul. 1970.

[11] T. Liu, P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 6, pp. 922–936, Aug. 1998.

[12] Z. R. Zaidi and B. L. Mark, "A mobility tracking model for wireless ad hoc networks," in *Proc. Conf. IEEE Wireless Commun. Netw. (WCNC)*, vol. 3, Mar. 2003, pp. 1790–1795.

[13] Z. Zhao, T. X. R. Li, and V. P. Jilkov, "Best linear unbiased filtering with nonlinear measurements for target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 4, pp. 1324–1336, Oct. 2004.

[14] N. K. Singh, S. Bhaumik, and S. Bhattacharya, "Tracking of ballistic target on re-entry using ensemble Kalman filter," in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2012, pp. 508–513.

[15] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering With Matlab Exercises*. Hoboken, NJ, USA: Wiley, 2012.

[16] E. A. Wan and R. V. D. Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proc. IEEE Adapt. Syst. Signal Process., Commun., Control Symp.*, Jul. 2000, pp. 153–158.

[17] Z. Yang and X. Wang, "Joint mobility tracking and handoff in cellular networks via sequential Monte Carlo filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 269–281, Jan. 2003.

[18] T. Li, M. Bolic, and P. M. Djuric, "Resampling methods for particle filtering: Classification, implementation, and strategies," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 70–86, May 2015.

[19] R. Douc and O. Cappe, "Comparison of resampling schemes for particle filtering," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, Sep. 2005, pp. 64–69.

[20] L. Mihaylova, D. Angelova, S. Honary, D. R. Bull, C. N. Canagarajah, and B. Ristic, "Mobility tracking in cellular networks using particle filtering," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3589–3599, Oct. 2007.

[21] L. Mihaylova, D. Angelova, D. Bull, and N. Canagarajah, "Localization of mobile nodes in wireless networks with correlated in time measurement noise," *IEEE Trans. Mobile Comput.*, vol. 10, no. 1, pp. 44–53, Jan. 2011.

[22] K. Pahlavan and A. H. Levesque, *Wireless Information Networking*. New York, NY, USA: Wiley, 2005.

[23] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[24] N. Salman, L. Mihaylova, and A. H. Kemp, "Localization of multiple nodes based on correlated measurements and shrinkage estimation," in *Proc. Sensor Data Fusion, Trends, Solutions, Appl. (SDF)*, Oct. 2014, pp. 1–6.

[25] O. Ledoit and M. Wolf, "Honey, I shrunk the sample covariance matrix," *J. Portfolio Manage.*, vol. 30, no. 4, pp. 110–119, 2003.

[26] J. Schäfer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statist. Appl. Genet. Molecular Biol.*, vol. 4, no. 1, p. 32, Nov. 2005.

[27] A. Tulsyan, B. Huang, R. B. Gopaluni, and J. F. Forbes, "A particle filter approach to approximate posterior Cramer-Rao lower bound: The case of hidden states," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2478–2495, Oct. 2013.

[28] M. Lei, B. J. van Wyk, and Y. Qi, "Online estimation of the approximate posterior Cramer-Rao lower bound for discrete-time nonlinear filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 37–57, Jan. 2011.

[29] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[30] B. Ristic, S. Arulampalam, and N. C. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications Partical Filters for Tracking Applications*. Norwood, MA, USA: Artech House, 2004.

[31] R. Khan, S. U. Khan, S. Khan, and M. U. A. Khan, "Localization performance evaluation of extended Kalman filter in wireless sensors network," *Procedia Comput. Sci.*, vol. 32, pp. 117–124, 2014.

[32] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[33] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Stat. Assoc.*, vol. 93, no. 443, pp. 1032–1044, 1998.

[34] R. Lamberti, F. Septier, N. Salman, and L. Mihaylova, "Sequential Markov Chain Monte Carlo for multi-target tracking with correlated RSS measurements," in *Proc. IEEE 10th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2015, pp. 1–6.

[35] Linkoping University. *Sensor Fusion App*, accessed on Jan. 7, 2016. [Online]. Available: https://play.google.com/store/apps

[36] C. C. Y. Kwan, "Estimation error in the average correlation of security returns and shrinkage estimation of covariance and correlation matrices," *Finance Res. Lett.*, vol. 5, no. 4, pp. 236–244, Dec. 2008.

Aroland Kiring received the B.Eng. degree in electronic and computer engineering from University Malaysia Sarawak, Malaysia, in 2006, and the M.Sc. degree in electronic, communication and computer engineering from the University of Nottingham, U.K., in 2009. He is currently pursuing the Ph.D. degree with the Department of Automatic and Control Systems, University of Sheffield, U.K. His current research interests include received signal strength-based target tracking and algorithms optimization for target tracking in wireless networks. He is a Graduate Member of the Board of Engineers, Malaysia, and the Technological Association of Malaysia.

Naveed Salman received the bachelor's (Hons.) degree in electrical and electronics engineering from the NWFP University of Engineering and Technology, Peshawar, Pakistan, and the master's and Ph.D. degrees from the University of Leeds, U.K., in 2009 and 2014, respectively. He was a Research Associate with the Department of Automatic Control and Systems Engineering, University of Sheffield, from 2014 to 2016. He is currently working in collaboration with Nestle UK Ltd., on the Innovate U.K. Project Effective Milk Processing With Variable Composition at the National Centre of Excellence for Food Engineering, Sheffield Hallam University. He is the author of a number of journal and conference papers. He was a recipient of the 2012 GW Carter Best Paper Award from Leeds University. He serves as a Reviewer for several international journals and conferences, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE COMMUNICATIONS LETTERS.

Chao Liu received the degree in automation engineering from Xiamen University, China, in 2010, and the M.Sc. degree in advanced engineering robotics from the University of Bristol, U.K., in 2013. He is currently pursuing the Ph.D. degree with the Department of Automatic Control and Systems Engineering, University of Sheffield.

Iñaki Esnaola received the M.S. degree in electrical engineering from the University of Navarra, Spain, in 2006, and the Ph.D. degree in electrical engineering from the University of Delaware, Newark, DE, in 2011. From 2010 to 2011, he was a Research Intern with Bell Laboratories, Alcatel-Lucent, Holmdel, NJ, and from 2011 to 2013, he was a Post-Doctoral Research Associate at Princeton University, Princeton, NJ. He is currently a Lecturer with the Department of Automatic Control and Systems Engineering, University of Sheffield, and a Visiting Research Collaborator at the Department of Electrical Engineering, Princeton University. His research interests include information theory and communication theory with an emphasis on the application to electricity grid problems.

Lyudmila Mihaylova (M'98–SM'08) is a Professor of Signal Processing and Control with the Department of Automatic Control and Systems Engineering, University of Sheffield, U.K. Her research is in the areas of machine learning and autonomous systems with various applications, such as navigation, surveillance, and sensor network systems. She has given a number of talks and tutorials, including the plenary talks for the CEIT-2016, Tunisia, and the IEEE Sensor Data Fusion 2015, Germany, invited talks at the University of California, Los Angeles, IPAMI Traffic Workshop 2015, USA, and the IET ICWMMN 2013, Beijing, China. She was elected in 2016 as President of the International Society of Information Fusion (ISIF). She is on the Board of Directors of ISIF. She was the General Co-Chair of IET Data Fusion and Target Tracking 2012 and 2014 Conferences, a Program Co-Chair for the 19th International Conference on Information Fusion, Heidelberg, Germany, 2016, and the Academic Chair of Fusion 2010 Conference. He is an Associate Editor of the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS and the Elsevier *Signal Processing Journal.*