The University Of Sheffield.

Shrinkage Based Particle Filters for Tracking in Wireless Sensor Networks with Correlated Sparse Measurements

Aroland M'conie Jilui Kiring

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

THE UNIVERSITY OF SHEFFIELD
Department of Automatic Control and Systems Engineering
December 2017

Acknowledgements

My deepest gratitude and appreciation goes to my supervisor Prof. Lyudmila Mihaylova, without whom this thesis would not have been born, and without whom I would not grow as much as I have today. Thank you for your continuous guidance, support, and encouragement.

Thank you also to my second supervisor Dr. Inaki Esnaola for his patient guidance and useful critiques throughout my time working with him.

Thank you to Naveed Salman, Le Yang, and Chao Liu for the willingness to share their knowledge that has tremendously benefited me for my postgraduate experience in Sheffield.

Thank you to the Ministry of Higher Education, Malaysia (MoHE) and the University of Malaysia Sabah (UMS) for the financial support. Without the support, my journey in Sheffield would not have been possible.

Thank you to my amazing wife Farm Yan Yan for your endless love and understanding. Our first year in Sheffield brightened up by the birth of our daughter. Raising Hannah while both of us doing our Ph.D. is very challenging. Despite countless challenges, we never despair. We continue raising Hannah ourselves. I cherish every moment I spent with both of you. Watching Hannah grows happy and healthy made us a proud parent.

Thank you to my parents and parents-in-law for their love and support, to my brothers, especially Vincent for taking care our parents while I'm abroad doing my Ph.D., and to my sister.

Special thanks to those I have missed to mention, but who have contributed directly or indirectly towards the production of this thesis in one way or another.

Abstract

This thesis focuses on the development of mobile tracking approaches in wireless sensor networks (WSNs) with correlated and sparse measurements. In wireless networks, devices have the ability to transfer information over the network nodes via wireless signals. The strength of a wireless signal at a receiver is referred as the received signal strength (RSS) and many wireless technologies such as Wi-Fi, ZigBee, the Global Positioning Systems (GPS), and other Satellite systems provide the RSS measurements for signal transmission. Due to the availability of RSS measurements, various tracking approaches in WSNs were developed based on the RSS measurements.

Unfortunately, the feasibility of tracking using the RSS measurements is highly dependent on the connectivity of the wireless signals. The existing connectivity may be intermittently disrupted due to the low-battery status on the sensor node or temporarily sensor malfunction. In ad-hoc networks, the number of observation of the RSS measurements rapidly changing due to the movements of network nodes and mobile user. As a result, the tracking algorithms have limited data to perform state inference and this prevents accurate tracking. Furthermore, consecutive RSS measurements obtained from nearby sensor nodes exhibit spatio-temporal correlation, which provides extra information to be exploited. Exploiting the statistical information on the measurements noise covariance matrix increases the tracking accuracy. When the number of observations is relatively large, estimating the measurement noise covariance matrix is feasible. However, when they are relatively small, the covariance matrix estimation becomes ill-conditioned and non-invertible. In situations where the RSS measurements are corrupted by outliers, state inference can be misleading. Outliers can come from the sudden environmental disturbances, temporary sensor failures or even from the intrinsic noise of the sensor device. The outliers existence should be considered accordingly to avoid false and poor estimates.

This thesis proposes first a shrinkage-based particle filter for mobile tracking in WSNs. It estimates the correlation in the RSS measurement using the shrinkage estimator. The shrinkage estimator overcomes the problems of ill-conditioned and non-invertibility of the measurement noise covariance matrix. The estimated covariance matrix is then applied to the particle filter. Secondly, it develops a robust shrinkage based particle filter for the problem of outliers in the RSS measurements. The proposed algorithm provides a non-parametric shrinkage estimate and represents a multiple model particle filter. The performances of both proposed filters are demonstrated over challenging scenarios for mobile tracking.

Contents

Acronyms

AGPS	Assisted Global Positioning System
AOA	Angle of Arrival
APIT	Approximate Point in Triangle Test
DV-HOP	Distance Vector-Hop
EKF	Extended Kalman Filter
GPS	Global Positioning System
KF	Kalman Filter
MMSE	Minimum Mean Square Error
OS	Operating System
PCRLB	Posterior Cramer-Rao Lower Bound
PDF	Probability Density Function
PF	Particle Filter
PFIM	Posterior Fisher Information Matrix
RF	Radio Frequency
RMSE	Root Mean Square Error
RSPF	Robust Shrinkage based Particle Filter
RSS	Received Signal Strength
ShPF	Shrinkage based Particle Filter
SPF	Gaussian model based Particle Filter
SSMs	State-Space Models

TDOA	Time Difference of Arrival
TOA	Time of Arrival
TPF	Student's t model based Particle Filter
UKF	Unscented Kalman Filter
WSN	Wireless Sensor Network

List of Figures

List of Tables

List of Symbols

$(\ddot{x}_k, \ddot{y}_k)$ Acceleration

$(\dot{x}_k, \dot{y}_k)$ Velocity

(x, y) Coordinates

$\bar{\rho}$ Average correlation

β Path loss exponent

$\boldsymbol{\pi}$ Transition probabilities

$\boldsymbol{\pi}_s$ Noise model transition probabilities

$\mathbf{B}_u$ Command matrix

$\mathbf{B}_w$ Noise coefficient matrix

Γ Gamma function

γ Temporal correlation (chapter 5)

λ Shrinkage intensity or weight

$\boldsymbol{\Gamma}$ Correlation matrix

$\mathcal{G}$ Gamma distribution

$\mathcal{N}$ Normal distribution

$\mu_{i,0}$ Initial probabilities

$\nabla_{\mathbf{x}_k}$ First-order partial derivative operator

$\rho^{i,j}$	Spatial correlation
σ_v^2	Measurement noise variance
σ_w^2	Process noise variance
τ	Maneuvering time constant
A	State transition matrix
C	Covariance Matrix
H	Jacobian matrix of the expected measurements
I	Identity matrix
J	Posterior Fisher information matrix
m	System modes
P	State covariance matrix
Q	Process noise covariance matrix
R	Measurement noise covariance matrix
S	Shrinkage covariance matrix
T	Target matrix
u	Control input
v	Measurement noise
w	Process noise
x	State
z	Measurements
$\tilde{\rho}_{k,l}$	Temporal correlation (chapter 4)
d	Euclidean distance
D_c	Decorrelation distance

k	Time index
n	Number of sensors
N_p	Number of particles
N_{eff}	Effective sample size
N_{thresh}	Threshold sample size
P	Sample size or observations
$p(\cdot)$	Probability density function
s	Noise model
St	Student's t distribution
T	Discretization period
t_{window}	Sliding window time
v	Degrees of freedom (chapter 5)
w	Weight

Chapter 1

Introduction

1.1 Wireless Sensor Networks

A wireless sensor network (WSN) consists of a group of sensors that can sense, measure, and collect information about their surroundings (temperature, pressure, humidity, images, etc). Wireless sensors are small, equipped with processing and computing capabilities but their operation is suppress by having limited operating power. Sensors communicate wirelessly with one another to perform monitoring tasks in an autonomous manner and can react to events in a specific environment such as an explosion. These features enable sensors to operate on a broad range of applications such as air quality monitoring [1], [2], [3], [4], forest fire detection [5], [6], [7], [8], natural disaster prevention [9], [10], [11], traffic surveillance [12], [13], [14], structural monitoring [15], [16], [17], [18], healthcare [19], [20], [21], [22], military surveillance [23], [24], [25], etc. WSNs are also deployed in heavy industrial applications such as power distribution [26], waste-water treatment [27], and specialized factory automation.

Every sensor hardware is equipped with a module comprised of a power unit (batteries and/or solar cells), a sensing unit (sensors), a processing unit, and a transceiver unit (for communication in the network). The sensor devices can be either passive or active sensors. The passive sensors such as acoustic, humidity, and temperature sensors consumed lower power than the active sensors such as radar and sonar. The sensor lifetime is limited due to the dependency on a battery. In events such as detecting, processing, and transmitting raw or processed data, more power is consumed compared to the normal operation, thus reducing the lifetime of the sensors. Sensors are usually deployed in large quantities and it is challenging to collect data from WSNs if there is no connectivity between sensors due to a low battery life. Generally, sensors will send data to the central node, which then perform either computations or transmitting information to a base station. Therefore, to support the sensor operation, it is

important to have a continuous supply of power.

WSNs have the advantage over wired networks in term of implementation where they can be deployed in remote, harsh and hostile geographic locations however collecting data from the sensor networks is a challenging task. The communication between wireless sensors is usually in a form of radio, infrared, and optical signals. Managing a high number of sensors required standard operating and networking protocols. Some protocols may not well suit to the unique features and requirements of sensor networks. In ad hoc networks, the network topology may changes anytime after the sensor deployment and at some future time, additional sensors may be added in the network. To manage the networks, open standards and protocols are needed at the physical, link, network, and transport layers. The network protocols have to be cost-effective, supports sufficient transmission of data rates, has low power-consumption and guarantees security and reliability. Some examples of the applicable standard are the IEEE 802.15.4 and IEEE 802.11. The IEEE 802.11 supports 1- or 2-Mbps transmission in the 2.4-GHz band using either frequency-hopping spread spectrum or direct-sequence spread spectrum. To provide up to 54 Mbps in the 5-GHz band by using orthogonal frequency-division multiplexing encoding, the IEEE 802.11a is introduced. Other standards are such as the IEEE 802.11b, IEEE 802.11g and IEEE 802.16.

The development of a standard operating system (OS) for WSNs is also important as the OS-supported programming model has a significant impact on the application development. Some of the basic library in the OS includes networks protocols, distributed services, sensor drivers, and data acquisition tools. These features are optimized depending on a specific WSN application. The size of memory in the sensor hardware is usually small, therefore developing low-complexity algorithms are important when there is a need to install and run the algorithm on the tiny sensor nodes. Low-complexity algorithms consume less processing time and power when compared to high-complexity algorithms. The difference in the computation time and power consumption become evident when dealing with a large number of sensor data. Consequently, a trade-off between algorithm complexity, computation time, and power consumption has to be accepted. Besides that, sensors must deliver data within a determined time window so that data processing can be carried out as fast as possible. The transmission of data in indoor environments is still quite unpredictable compared to the outdoor environments. As more and more wireless systems are implemented, coexistence becomes a bigger issue. WSNs are application-oriented and depending on the type of application, WSNs present different challenges and constraints, which required specific designs and solutions. Ideally, the solutions should integrate well with the already established WSN systems in order to be practical. In this thesis, the focus will be towards tracking applications of the WSNs.

1.2 Tracking in Wireless Sensor Networks

Tracking is one of WSN's widely researched applications that deals with the state estimation of a moving object based on the observable measurements. The state represents the physical system of interest (position, velocity, etc.) and is normally described by means of a dynamic model. Since the state is not directly observable, an observation model is required for the state estimation. A widely used class of models that have been successfully applied in tracking is the state-space models (SSMs). The SSM consists of the state transition equation which describes the state behavior of the mobile user over time and the measurement equation which relates the observed measurement to the hidden state. The process of extracting the state information from the noisy measurements depends on the choice of motion models and filtering methods.

The challenges to track accurately in WSNs are highly influenced by factors such as the dynamic of sensor allocation, communication between nodes, computation, and target behavior. The sensor allocation determines the area of coverage and the probability that the sensor will detect an event of a moving object. For optimal coverage, a numerous number of static and/or mobile sensors were deployed to cover a wide range of area. Sensors have to be able to maintain a reliable communication between the nodes to enable successful data sharing in the network. Moreover, the ability of the sensor node to perform efficient computation by running a low-complexity algorithm can help to reduce power consumption while maintaining a satisfactory level of tracking accuracy. For the instant, to track a moving car accurately, its dynamic/motion is modeled by physical laws which describes how the state changes over time. Car driver actions such as steering movements and acceleration patterns were also considered. These are usually unknown and can be modeled as a random input with a certain probability distribution. Moreover, sensor measurements provide some information (distances, angles, etc) about the moving car in which the sensor data itself also contains measurement errors. Such errors can be accounted as another random input with a certain probability distribution. Hence, both the state and measurement transition processes have to account the effect of the random inputs. Therefore, the filtering process is carried out to extract information about the system state by filtering out noises (undesired information) in the state transition and measurement processes.

In the literature, tracking solutions in wireless systems is divided into two approaches: methods in which point coordinates are estimated using global positioning system (GPS) devices [28] and methods in which the coordinate and motion are estimated using SSM with filtering algorithms [29]. 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 [30]

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, SSM 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 [31], Gauss-Markov models, and Singer-type models [32], [33]. In [33], a dynamic mobility model that captures a wide range of vehicle maneuvering patterns is presented and employed for tracking in tactical weapons systems. Liu $et\ al.$ [34] use this mobility model to estimate the trajectory of mobile users in wireless asynchronous transfer mode (ATM) networks. The Kalman filter (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 $et\ al.$ [35] develop an extended Kalman filter (EKF), 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 [36]. Filtering methods based on the random sampling can also be applied in mobility tracking scenarios, such as the Ensemble Kalman filter (EnKF) [37] and the Unscented Kalman filter (UKF) [38], [39]. The Particle filter (PF) [40] is often used in non-linear and non-Gaussian filtering problems. In [41], a mobility model combined with a PF is developed for mobile tracking in cellular networks.

Tracking based on the SSM and filtering methods depend on the sensor measurements. Measurements or readings from sensors located approximately close to one another or in the same geographical area is spatially correlated. Besides that, when a consecutive measurement was obtained from the same sensors, it will also exhibit temporal correlation. The spatio-temporal correlation presents in the sensor measurements provides extra information that can be exploited to improve tracking accuracy. To exploit the correlation, a covariance matrix is estimated from the sensor measurements. A common approach to estimate the covariance matrix is by using a sample covariance estimator. However, the sample covariance matrix estimate is ill-conditioned, non-invertible, and introduces a large estimation error when the number of sensor measurement is sparse. In ad-hoc networks, signal connections between the user and sensors are spontaneously established when the user is within the communication range. The movement of user and sensors changes the network topology and results in a disruption of the existing connections. This reduces the number of measurements that are available in the network.

Moreover, the measurements obtained by sensors are also affected by random noises and these noises were usually described in term of probabilistic descriptions. In filtering problems, the state and measurement noises are always represented by distributions rather than single numbers. The measurement noise is often assumed to follow a Normal or Gaussian

distribution. The Gaussian distribution is statistically convenient and it is the most common distribution in nature. However, in practical applications, the distribution of measurement noise does not exactly match the Gaussian distribution. The measurement noise can take a non-Gaussian distribution when the data measurement is corrupted by the outliers. The outlier is an observation that does not fit with the expected pattern of distribution based on the data. Sensor measurements are difficult to interpret when the data is corrupted by the outliers. Outliers can come from sudden environmental disturbances, temporary sensor failures, and also because of the intrinsic noise of the sensor device itself. Therefore, there is an urgent need to develop filtering methods that work well beyond the assumption of Gaussian distribution of noise models.

The main objective of this thesis is to develop tracking approaches to estimate the state of a mobile user in the WSNs. The sensor measurements are spatio-temporally correlated and by exploiting the correlation present in the measurements, the tracking accuracy is expected to improve. However, exploiting the correlation in the measurements required the measurement noise covariance matrix information which is not available. Hence, the measurement noise covariance matrix has to be estimated. In situations where the number of observation is larger than the dimensionality of the signal model, the estimation is carried out using the sample estimator. However, when the number of observations is comparable or less than the dimensionality of the signal model, estimating the measurement noise covariance matrix becomes difficult due to the ill-posed problem of the covariance matrix. In this thesis, the number of observation is defined as the total number of snapshots or observations that can be obtained at one particular time instant. Meanwhile, the dimensional of the signal model is defined as the total number of sensors deployed in the network. Apart from that, the sensor measurements are also affected by the presence of outliers in the data measurements due to certain factors such as the sudden environmental disturbances. Outliers are defined as the values in the measurements that are numerically deviated from the expected Normal or Gaussian distribution of the sensor measurements. To account for the effect of outliers in the sensor measurements, the measurement noise is modeled in the form of Student's t distribution instead of Gaussian distribution. This method is analytically intractable and required some form of approximation.

In this thesis, two approaches are developed namely (1) a shrinkage-based particle filter, and (2) a robust shrinkage-based particle filter. A shrinkage-based particle filter is developed based on the idea of exploiting the correlation that presents in the sensor measurements. The developed filter overcomes the problems of covariance matrix estimation with the small or comparable number of observation when compared to the dimensionality of the signal model. On the other hand, a robust shrinkage-based particle filter is the optimization version of the first approach. The second filter is developed to make it more robust towards the effect of outliers in the sensor measurements. Throughout this thesis, vectors were denoted 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, $\mathcal{G}(\alpha, \beta)$ is a Gamma distribution with shape parameter α and rate parameter β, $\Gamma(\cdot)$ is the gamma function, $St(\boldsymbol{\mu}, \hat{\boldsymbol{\Sigma}}, v)$ denotes a multivariate Student's t distribution with mean $\boldsymbol{\mu}$, scale matrix $\hat{\boldsymbol{\Sigma}}$, and degrees of freedom parameter v, $[\mathbf{A}]_{ij}$ refers to the element at the i-th row and j-th column of matrix $\mathbf{A}$, $\|\mathbf{A}\|_F^2 = \mathrm{tr}(\mathbf{A}\mathbf{A}^T)/n$ to denote the scale Frobenius norm of $\mathbf{A}$ where $\mathbf{A}$ is a $n \times n$ matrix, and $\mathrm{tr}(\mathbf{A})$ to denote the trace of the matrix $\mathbf{A}$, $\mathbb{R}$ 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.

1.3 Challenges

Developing approaches for mobile tracking in WSNs introduces many challenges such as:

Exploiting correlated sensor measurements. In wireless networks where sensors are densely deployed, consecutive measurements obtained by sensors are spatio-temporally correlated. The spatio-temporal correlation is a feature that can be exploited for localization and tracking purposes. To exploit the correlation, the measurement noise covariance matrix has to be estimated. The estimated measurement noise covariance matrix has to remain positive definite and invertible otherwise any attempt to exploit the correlation in the sensor measurements would result in an ill-posed problem.

Dealing with sparse sensor measurements. The number of observations or snapshots of the sensor measurements at one particular time instant can change easily in ad-hoc networks due to the sensor and user movements or sensor malfunction. In this thesis, two types of network settings were considered in the simulation; (1) wireless networks comprises of mobile sensor nodes (Chapter 4) and (2) wireless networks comprises of static sensor nodes (Chapter 5). In the events, where the number of sensor measurement is sparse, due to the loss of signal connectivity in the network, estimating the measurement noise covariance matrix is detrimental. Consequently, an approach that is able to exploit the correlated data is not feasible. The sparsity in the measurements refers to situations where the total number of observations or snapshots is comparable or less than the dimensionality of the signal model. The number of connectivity in the networks can be increased by deploying more sensor nodes, however using a large number of sensors can be very expensive to implement and in addition to that, they can be a signal overlap and data redundancy problems.

Dealing with outliers in the sensor measurements. Sensor measurements are often difficult to interpret due to the presence of noise and outliers in the data measurements. The measurement noise is usually assumed to follow a Normal or Gaussian distribution. However, when outliers present, the assumption of Gaussian distribution is no longer hold. The measurement

noise can be non-Gaussianly distributed due to the presence of outliers. Outliers in the measurements are recorded when there is a temporary malfunction of sensor nodes. The sensor readings will numerically deviate from the expected Gaussian distribution of the data measurement. Non-Gaussian distribution is difficult to model by an analytic function, and even if it can be modeled, the function does not provide an easy mathematical solution. Therefore, making a state inference based on the noisy and corrupted sensor measurements is challenging. Most tracking methods operate with the assumption that the measurement noise follows a Gaussian distribution but in practice, this is not always the case. Developing a tracking method that works well with both the Gaussian and non-Gaussian probabilistic distributions is challenging because of it is analytically intractable.

1.4 Thesis Outline

The current chapter serves as an introduction to the thesis, discussing the motivations and challenges of the work presented in the thesis. The remainder of the thesis is organized into the following chapters:

Chapter 2 presents a background overview of various components related to localization and tracking methods. A classification and review of localization and tracking algorithms are presented.

Chapter 3 provides the necessary mathematical backgrounds required to understand the rest of the thesis and it has two main sections; the state space models and the filtering methods.

Chapter 4 proposes a tracking framework that works with correlated and sparse sensor measurements. The chapter begins with the framework overview, the state space models, the correlated shadowing noise model, the development of the proposed method, and the posterior Cramer Rao lower bound. This is followed by the validation of the proposed method using both simulation and experimental data.

Chapter 5 describes an extension to the work in Chapter 4, with an emphasis on the non-Gaussianly distributed of measurement noise. This chapter considers a scenario where static sensor nodes were deployed in a network, and provide measurements that are corrupted by the outliers. A robust shrinkage-based particle filter for tracking problem with outliers is proposed over a challenging tracking simulation with outliers.

Chapter 6 concludes the thesis by presenting the key contributions of the thesis, summary of the thesis findings, and also makes recommendations for future work.

1.5 Research Contributions

The main theme of this thesis revolves around the used of correlated and sparse wireless measurements for mobile tracking. Two approaches have been developed to exploit the correlated measurements that are (1) a shrinkage-based particle filter, and (2) a robust shrinkage-based particle filter. These filters are able to overcome the problems of covariance matrix estimation using sparse sensor measurements whereas the sparsity refers to situations where the total number of observations or snapshots is comparable or less than the dimensionality of the signal model. The first approach is validated using both simulation and experimental data. In the simulation settings, the sensor nodes are assumed to be mobile. When a target is detected in the networks, the sensor nodes moved towards the target. However, the mobile sensors were only able to move around their designated grid. On the other hand, the second approach is validated using simulation data where static sensor nodes were considered in the simulation network. It is developed to be more robust towards the presence of outliers in the data measurements. Hence, a non-parametric shrinkage estimator is derived under the non-Gaussian distribution of the measurement noise. The measurement noise is modeled in the form of the Student's t distribution instead of the Normal or Gaussian distribution in order to account for the effect of outliers in the sensor measurements. The main contributions of this thesis are details as follows and they are linked to the author's relevant publications listed in Section 1.6.

Chapter 4 - Shrinkage based particle filter for tracking with correlated sparse measurements [P1], [P2], [P3].

- A shrinkage-based particle filter is developed to exploit the correlation present in the RSS measurements and achieve accurate mobile user tracking.

- The developed filter combined the shrinkage estimator and the particle filter to jointly estimate the shadowing noise covariance matrix and the state of the mobile user.

- The shrinkage estimator is proposed to overcome the ill-posed problem of the covariance matrix estimation by the sample estimator when estimating with the small or comparable number of observation when compared to the dimensionality of the signal model.

- The developed filter is validated with both the simulation and experimental data whereby the simulated wireless network comprises of mobile sensor nodes.

Chapter 5 - Robust shrinkage based particle filter for tracking with non-Gaussian shadowing noise [P4].

- A robust shrinkage-based particle filter is developed to achieve accurate mobile user tracking in the events of outliers presence in the sensor measurements.

- The developed filter combined the non-parametric shrinkage estimator and the robust particle filter to jointly estimate the shadowing noise covariance matrix corrupted by outliers and the state of a mobile user.

- The non-parametric shrinkage estimator is derived to estimate the shadowing noise covariance matrix using the sensor measurements that are corrupted by the outliers.

- The developed filter is validated using a simulated wireless network comprises of static sensor nodes and is compared with a state of the art particle filter.

1.6 Publications

Parts of the material in this thesis have been published by the author in:

Peer Reviewed Journal Article

P1. A. Kiring, N. Salman, C. Liu, I. Esnaola, and L. Mihaylova, "Tracking with sparse and correlated measurements via a shrinkage-based particle filter," *IEEE Sensors Journal*, vol. 17, no. 10, pp. 3152-3164, May. 2017.

Peer Reviewed Conference Papers

P2. A. Kiring, N. Salman, C. Liu, I. Esnaola, and L. Mihaylova, "A shrinkage-based particle filter for tracking with correlated measurements," in *Proc. Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Oct. 2015, pp. 1-6.

P3. C. Liu, A. Kiring, N. Salman, I. Esnaola, and L. Mihaylova, "A Kriging algorithm for location fingerprinting based on received signal strength," in *Proc. Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Oct. 2015, pp. 1-6.

Work from the jointly authored publication in [P3] is not directly presented in the thesis. However, work contained in the publication that is directly attributable to me is towards the discussion of correlation data model and signal model.

Journal Publications Under Review

P4. A. Kiring, L. Yang, and L. Mihaylova, "Robust shrinkage based particle filter for tracking with non-Gaussian shadowing noise," Submitted to *IEEE Transactions on Vehicular Technology*, 2017.

Chapter 2

Background Review

In this chapter, the background overview of the localization and tracking methods are presented. The first part of this chapter will focus on classifying the localization algorithms, reviewed in Section 2.2 and the second part will focus on classifying the tracking algorithms, reviewed in Section 2.3. Finally, a summary is present in Section 2.4.

2.1 Overview

By definition, localization is a process of finding the exact coordinates of static nodes and tracking deals with the state estimation of a moving object based on the observable measurements. In this chapter, nodes refer to the static users or devices and objects refer to the moving users or devices. From an algorithmic perspective, localization is a one-time process for static users, whereas tracking is the continuous localization of the moving users over time. Almost all existing localization and tracking methods consist of two phases: distances/angles estimation and coordinates computation. At first, the distance and angles between two sensor nodes are estimated in the distance/angle estimation. Then, the coordinate of the unknown node is computed based on the distance and angle information of several reference nodes. The localization and tracking algorithms can be classified into few categories as follows:

2.2 Localization Algorithms in Wireless Sensor Networks

Localization algorithms in WSNs are divided into three categories; range based and range free algorithms [42], [43], cooperative and non-cooperative algorithms [44], and centralized

and non-centralized algorithms [44]. The range based and range free algorithms are differentiated according to the dependency of range measurements techniques. In the range-based technique, the algorithms have to first estimate the distance/angle between the unknown nodes and reference node before its can performed the coordinates computation. The reference node or also known as the anchor node refers to the sensor node with known coordinates in the network. On the other hand, the range-free technique is those algorithms that perform coordinates computational directly by using only the connectivity information between adjacent sensors. The next classification is cooperative and non-cooperative algorithms. Cooperative algorithms allow the exchange of information between all nodes in the network. Meanwhile, the non-cooperative algorithms only allow non-located nodes to exchange information with the anchor nodes. Finally, another classification subdivides the algorithms to centralized and non-centralized. In centralized algorithms, all computation is carried out in the central node. Conversely, in non-centralized algorithms, the computation is spread among the sensor nodes in the network. The central node refers to the sensor node that is located at the most central location of the cell network in relation to all the other sensor nodes.

2.2.1 Range-Based Localization

Range-based algorithms are those algorithms that have to first estimate the distance/angle between the unknown nodes and reference node before its can performed the coordinates computation. Range-based algorithms are much more accurate for localization than range-free algorithms but can required extra hardware for implementation and high power consumption [43]. Range-based algorithms estimate the distance or angle between the unknown and reference nodes using several measurement techniques such as:

Received Signal Strength-based. The received signal strength (RSS) is a power measurement present in a received signal. The RSS values are measured in dBm and they are usually very noisy due to shadow fading and multipath effects, which are depending on the environment. The RSS-based algorithms are the most common, cheapest, and simplest to implement since it does not require an additional hardware to implement. The strength of the signal attenuates and becomes weaker as the propagation distance increased. The signal strength is usually described by using the radio propagation path loss model. Consider the transmitter is transmitting a signal to the receiver. The received power is given by [45]

$$P_r = \frac{P_o}{d^\alpha}, \tag{2.1}$$

where P_o is the received power at a reference distance from the transmitter, and d is the distance between the transmitter and the receiver. It is shown that as the distance increases, the received power decreased proportionally to the distance. The path loss exponent is denoted by α and

it is characterized the propagation of the transmitted signal. Normally, for a free-space path, $\alpha = 2$ whereas for urban areas, $\alpha \in [2.7, 3.5]$ [46]. When expressed as log-normal variables, the power-distance relationships becomes

$$10\log_{10}P_r = 10\log_{10}P_o - 10\alpha\log_{10}d, \tag{2.2}$$

where $10\log_{10}P_r$ and $10\log_{10}P_o$ represent the transmitted and received power, respectively, at a reference distance of 1-meter. The term $10\alpha\log_{10}d$ indicates the power loss during the transmission. Another convenient way to express power-distance relationship is by using path loss rather than the signal power. Let the path loss in dB is given by

$$L = L_o + 10\alpha\log_{10}d + w, \tag{2.3}$$

where L is the total path loss between the transmitter and receiver, L_o is the path loss at a reference distance of 1-meter, and w is the path loss due to shadow fading and multi-path effects. The reference path loss is determined by $L_o = 10\log_{10}P_t - 10\log_{10}P_o$, where P_t is the transmitted signal power, and P_o is the received signal power at 1-meter distance.

The path loss model described in (2.3) is the most common model used for distance estimation. Hashemi [47] and Patwari [48] have collected real indoor RSS measurements to make comparison and analyses of the distance estimation based on the path loss model. It is shown that sufficient accuracy is achieved with the path loss model where the path loss exponent, shadow fading, and multipath effects are highly dependent on the environments. In [49], an analytical study of the path loss model had been carried out and the results showed that the distance estimation accuracy reduces when the RSS measurement is not stable and contain large power variations. Sucasas [50] investigated the applicability of the model in a dynamic network environment and the findings showed that the network dynamic does not affect the localization accuracy of the system if the nodes in the network are synchronized in broadcasting their location information to the central unit. Furthermore, the accuracy of the system is increased if more anchors are deployed in the network.

Besides using model-based, there is also a map-based RSS measurement. Localization using map-based techniques required the construction of the RSS map for the network environment. The constructed RSS map act as a comparison map to compared with the received signal measurements. Bahl and Padmanabhan [51] constructed a radio map consists of RSS measurements from an emanating beacon. The RSS measurements at various points in the network are identified and collected to represent the network environments. The coordinate of the mobile user is localized by mapping the received RSS with the constructed RSS map. It is showed that the RSS maps able to increase the localization accuracy by using multiple constructed maps. To maintain the feasibility of the method, the RSS maps have to be updated regularly when there are changes made in the network environment [52]. In general, the

RSS-based algorithms were chosen because of its low complexity and easy to implement for practical applications. The RSS-based algorithms offer sufficient localization accuracy for a broad range of wireless-based applications.

Timing-based. The timing-based approaches estimate the distance between the two nodes by timing the flight of a received signal. Some examples of the methods developed based on this approaches are the time of arrival (TOA) and the time difference of arrival (TDOA). In TOA systems, the time taken for the transmitted signal to reach the receiver is given by [53], [54],

$$\tau = dc, \tag{2.4}$$

where d is the distance between the two nodes, and c is the speed of light. This method requires a highly synchronized clock between a transmitter and a receiver, thus a dedicated hardware is needed. They have to be perfectly synchronized in order to achieve high accuracy estimation. The timing flight is highly dependent on the line-of-sight of the signal. In TDOA systems, the distance between the nodes is estimated by taking the time difference between the arrival times of a signal sent by a transmitter at two receivers [55], [56]. The clock at the two receivers needs to be synchronized as well.

Directionality-based. The directionality-based algorithms also refer as angle of arrival (AOA) algorithms estimate the angle of the sender by using the directionality of the received signal thus required the construction of antennas on the sensor nodes [57], [58]. For the instant, beamforming methods used the anisotropic antenna to find the direction of the received signal [55]. The antenna beam is rotated electronically or mechanically towards the direction of the received power. Then an angle will be computed based on the highest received power direction. When the transmitted signal consists of varying power, it can reduce the accuracy of the inferred angle. Thus, a non-rotating omnidirectional antenna is used together with the anisotropic antenna to normalize the received signal power at the receiver.

In addition, phase interferometry [55], another method based on AOA approach requires the set-up of an antenna array. The angle of the received signal is determined by examining the phase difference at each element of the array. Each antenna in the array is separated by a fixed distance, d and the phase difference of the received signal is given by

$$2\pi \frac{d\cos\theta}{\lambda}, \tag{2.5}$$

where θ is the direction of the transmitter and λ is the wavelength of the received signal. In [59], the anchor uses an array of four antennas and the performance is analyzed. The results reveal that heuristic weighting function is able to decrease the location errors in AOA measure-

ments localization. The main challenges of AOA measurements localization with WSNs are to maintain the acceptable accuracy while keeping the system simple and portable for practical applications.

2.2.2 Range-Free Localization

Range-free algorithms are those algorithms that can perform directly the coordinates computation by just using the connectivity information or the exchanged multi-hop routing information instead of the distance/angle information. Some common algorithms based on connectivity information are Centroid algorithm and Distance Vector-HOP (DV-HOP) algorithm. In Centroid algorithm [60], anchors broadcast their coordinates to all sensors in the network. The non-anchor nodes listen the transmitted signal for a certain period of time. After that, the non-anchor nodes computed their coordinates in the network by a centroid determination from all positions of the anchors. In DV-HOP algorithm [61], a reference node (anchor) broadcast their locations, counted by the number of hop to their neighbors. The nearest neighbors are indicated by a small number of hop and contrarily, a large number of hop shows that the node is located far from the anchors. Through this method, all nodes in the network will find a minimum possible number of hops to reach the anchor nodes. The minimum number of hops is converted into distance metrics and using estimation method, the coordinates of non-anchors are then computed. Other range-free algorithms including the approximate point in triangle test (APIT) [62] and the amorphous [63].

2.2.3 Cooperative and Non-Cooperative Localization

The cooperative localization algorithms are the methods that required the collaboration between all the nodes inside the network. There is no communication restriction amongst all the nodes. Therefore, the non-anchor node can cooperate with all the other nodes (either being anchor or non-anchor) in making the estimation of their locations. The cooperative localization is suitable to be used in a large-scale network as the accuracy and coverage depend on the density of nodes.

Meanwhile, non-cooperative localization is a communication-restricted algorithm. Non-anchor locates their position by established communication only with anchor nodes. Hence, the number of anchor nodes must be sufficient to guarantee the accuracy of this approach. The non-cooperative localization is suitable to be used on a small-scale network because they have better direct communication between the nodes with the anchor nodes in the network.

2.2.4 Centralized and Non-Centralized Localization

In centralized algorithms, one node is selected as a central unit and this unit will perform the data collection and computation. The centralized algorithms have high accuracy because the central unit has the entire network information. However, this increased the transmission traffic to the central unit and hence reduces the scalability of the algorithm. A node that located far from the central unit will consume more energy during communication with the central unit.

In non-centralized algorithms, computation is distributed among anchors and non-anchors nodes in the network and these nodes will act as the central unit. The central unit collects information about their neighbors and used this information to perform coordinates estimates. After that, the estimated coordinates by the central node will be shared with the rest of the nodes in the network. The localization accuracy obtained by this method is less accurate when compared to the centralized algorithms due to the error propagation. However, the non-centralized algorithms are much more feasible to be implemented for a large scale sensor networks. Figure 2.1 summarises the localization algorithms classification in WSNs.

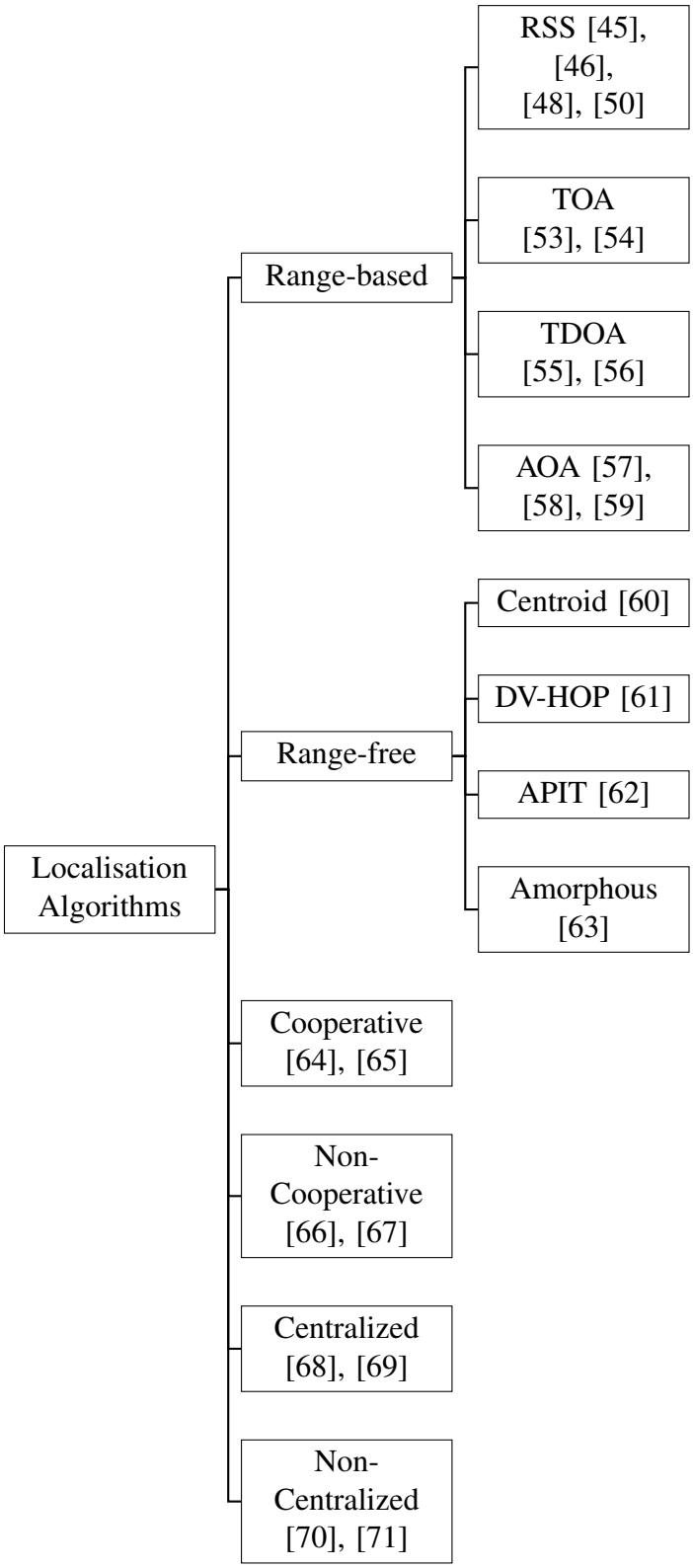

Figure 2.1: Localization algorithms classification in WSNs.

2.3 Tracking Algorithms in Wireless Sensor Networks

Mobile tracking in WSNs is generally carried out using the filtering methods such as the KF and its variants, the EKF and its variants, the UKF and its variants, and the PF and its variants. Unlike localization methods, the tracking methods have to consider all the possible motion scenarios. Tracking algorithms can be divided into few approaches such as [72], [73], [74]:

2.3.1 Cluster based Tracking

The cluster-based tracking is a method that divides the sensor nodes into a cluster head and several cluster nodes. The type of cluster can be static, dynamic, and spatio-temporal. In static clustering, clusters are formed statically thus it is not adaptable to dynamic scenarios. In a case where the cluster head failed, the entire nodes in the cluster will not be able to communicate with another cluster. In dynamic clustering, the clusters are dynamically constructed depending on different scenarios. When the nodes sense the mobile target, a cluster of nearby nodes is formed. Hence only nodes that belong to the cluster are activated. The nodes can also be member of different clusters at a different time. This reduced the energy consumption of the nodes and at the same time increased the estimation by using only the best possible nodes in the networks. Many clustering methods are formed based on the distance function but in spatio-temporal clustering, the construction of cluster based on distance measure is extended to the function of the position history of the mobile target. The spatio-temporal cluster based tracking has higher accuracy than the static and dynamical cluster based tracking.

2.3.2 Tree based Tracking

The tree-based tracking is a method that organized a sensor network into a tree-like structure where nodes that detect the mobile target are selected as the root node. When a mobile target detected in a network, the nearby node (root node) collect the sensed data, calculates the mobile motion, and send a warning message to the node closest to the next estimated state. The node that received this warning message will become the new root of the network that is responsible to perform the next state estimation. Unfortunately, when the mobile target moves at a high speed, it can pass several nodes without being detected. Therefore, to reduce missing coordinates, a recovery mechanism is required. However, using recovery mechanism frequently can lead to significant energy consumption. In the existing methods, an optimization is carried out revolves around the recovery mechanism such as query cost reduction, update cost reduction, coverage area, lifetime, etc.

2.3.3 Embedded Filter based Consensus

The cluster-based and tree-based tracking methods are based on the hierarchical architecture where the state estimation is performed at a central or cluster node. To reduce the state estimation dependency over the central or cluster node, the embedded filter based consensus is introduced. This method is based on peer-to-peer architecture, where the state estimation process is distributed among all sensors in the network. Since the estimation process involves many sensors, a set of rules is defined based on a specific consensus such as consensus on state, measurement or information. For a standard KF, consensus strategies can be applied to update step and/or prediction step without losing the filtering standard operation. In [75], a distributed KF is proposed for distributed sensor networks. The filter operates based on two consensuses; weighted measurements and inverse-covariance matrices. In [76], the problem of multi-target tracking for distributed sensor networks is presented. A multi-target tracking algorithm that utilized consensus over the estimated state and measurement information is proposed. Figures 2.2 summaries the tracking algorithms classification in WSNs.

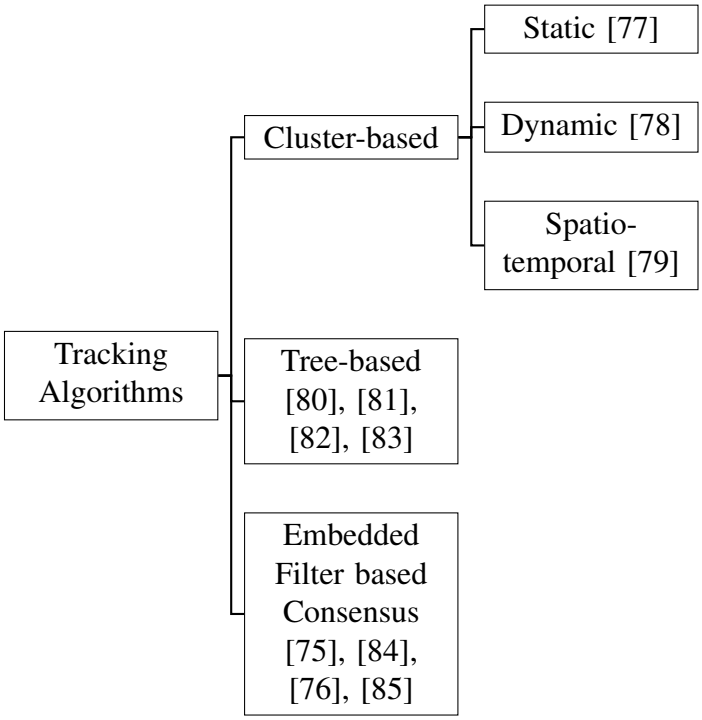

Figure 2.2: Tracking algorithms classification in WSNs.

2.4 Summary

In this chapter, a background review covering the localization and tracking methods in WSNs was presented. The chapter begins by identifying different type of measurements metrics that are available for both range-based and range-free methods. Range-based methods estimate the distance between nodes by the used of range information meanwhile range-free methods perform distance estimation by connectivity information. The estimated distance between the nodes can be highly erroneous thus filtering algorithms is adopted to reduce the error made during distance estimation. Apart from that, the localization methods can be cooperative and non-cooperative. In cooperative localization, all sensor nodes are able to communicate with one another but in non-cooperative localization, only anchor nodes are able to communicate with the rest of the nodes in the network. The coordinates computation by the filtering methods can also be performed in centralized and non-centralized manners.

The background review then shifted towards the tracking algorithms. The solution of mobile tracking can be mainly classified into schemes such as tree-based tracking, cluster-based tracking and embedded filter based consensus. The tree-based and cluster-based approaches belong to hierarchical networks. On the other hand, the embedded filter based consensus approaches belong to peer-to-peer networks. To maintain tracking performances for hierarchical networks, a selection of other sensor node is required when the central or cluster node is failed. In peer-to-peer networks, the most common approaches are the distributed KFs, which are based on the KF and their variants. A distributed tracking algorithm (embedded filter based consensus) has an advantage over the centralized distribution (tree-based and cluster-based) in term of scalability. Other approaches to tracking algorithms are such as UKF [86], adaptively robust UKF [87], sequential Markov chain Monte Carlo [88], etc. Some of the root causes of tracking errors are including the sensor node failure, communication failures, target behavior, energy consumption and tracking latency (computational time). In this thesis, a mobile tracking approach based on the WSNs will be developed to exploit the correlated and sparse wireless measurements.

Chapter 3

State Space Models and Filtering Methods

In this chapter, the theoretical concepts are presented for the problem of state estimation of a dynamic system using noisy measurements. The outline of this chapter is as follows: First, a general introduction to state space models is given in Section 3.1. Section 3.2 defines the Bayesian filtering. Section 3.3 and 3.4 introduce methods for linear and nonlinear filtering using the KF and its variants. Finally, Section 3.5 briefs the PF method which then becomes the framework for all following chapters.

3.1 State Space Models

In dynamical systems, the state represents the quantity of interest which captures relevant information about the system and typically changes over time. The way in which the state changes is described by mathematical equations as a function of time or at discrete-time instances. Typically, the state itself is not available instead they are manifested through a set of observed measurements. A widely used class of models that have been successfully utilized in tracking applications is the SSMs.

The SSM describes the behavior of the system state as a function of time using a series of first-order differential or difference equations [89]. It consists of the state transition and measurement equations. SSM can be a deterministic or stochastic model depending on the complexity of the simulation problems. A deterministic model is determined solely by the value of the parameters and the initial conditions. Meanwhile, the stochastic model has some

random components that characterized the system output. In mobile tracking, the main goal is to estimate the unknown state trajectory of the mobile users in a network. Often, the unknown states consist of position and velocity. In WSNs, sensors are deployed to collect signal strength transmitted by the mobile user. The collected signals are processed by the filtering methods to make inference about the mobile states. To account for the uncertainties in the system, the SSM is extended to include the process and the measurement noises. The SSM is assumed to be available in a probabilistic form thus estimation is formulated based on the probabilistic inference.

General Model

In general, the SSM is described as follows

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k), \tag{3.1a}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \tag{3.1b}$$

where $\mathbf{x}_k$ represents the state vector, $\mathbf{z}_k$ represents the measurement vector, and k represents the time instant. The function $\mathbf{f}_k(\cdot)$ relates the previous state and its process noise $\mathbf{w}_k$ to the upcoming state $\mathbf{x}_{k+1}$. Meanwhile, the function $\mathbf{h}_k(\cdot)$ maps the state $\mathbf{x}_k$ and its measurements noise $\mathbf{v}_k$ to the observed measurements of $\mathbf{z}_k$. Both functions $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ can be nonlinear functions and all quantities $\mathbf{x}_k$, $\mathbf{w}_k$, $\mathbf{z}_k$, and $\mathbf{v}_k$ are vectors with known dimensions. Moreover, the process and measurement noises are characterized in term of probabilistic descriptions.

Linear Model

The most common SSM is the linear model, given by

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{w}_k, \tag{3.2a}$$

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \tag{3.2b}$$

where $\mathbf{x}_k$ represents the state vector, $\mathbf{z}_k$ represents the measurement vector, and k represents the time instant. The state transition matrix and the measurement matrix are denoted by $\mathbf{F}_k$ and $\mathbf{H}_k$, respectively, and they are expressed by linear functions. The structure and dimension of the matrices are typically known. The process noise $\mathbf{w}_k$, and the measurement noise $\mathbf{v}_k$ are considered to be independent with known noise characteristics.

Maneuvering Model

Consider the following system

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{w}_k, \tag{3.3a}$$

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{3.3b}$$

where $\mathbf{F}_k$ is the state transition matrix, $\mathbf{G}_k$ is the command matrix, $\mathbf{H}_k$ is the measurement matrix, and k is the time instant. The matrices are typically known. The control input vector denoted by $\mathbf{u}_k$ is used to describe the uncertainties in the mobile user acceleration pattern or behavior. The control input is unknown and hence modeled as a random process. In [90], the random process can be in a form of white noise models, Markov process models, and semi-Markov jump process models. Some of the models that are derived based on the white noise models are including the constant velocity model, constant acceleration model, and polynomial model. In [89], a control input based on the white noise models is presented to model the user acceleration behavior. The maneuvering model is relatively simple and works well with less harsh maneuvering process. To deal with more harsh maneuvering and acceleration behavior of the mobile user, a maneuvering model with control input based on the Markov process models is presented in [29]. Finally, in [91], a maneuvering model with control input based on the semi-Markov jump process models is presented.

Correlated Process Noise (Singer) Model

A mobile user can move with a different acceleration levels at any time instants. The random acceleration levels can affected the degrees of evasive maneuvering taken by the mobile user. In [90], the mobile target acceleration levels are modeled by the control input. The control input assumed the target acceleration $\mathbf{a}(t)$ is correlated with autocorrelation function given by

$$R_{\mathbf{a}}(\tau) = \mathbb{E}[\mathbf{a}(t+\tau)\mathbf{a}(t)] = \sigma^2 e^{-\alpha|\tau|}, \tag{3.4}$$

where the maneuvering is parameterized by an acceleration variance σ^2 and a time constant $1/\alpha$. The process $\mathbf{a}(t)$ of continuous-time results from a linear system

$$\dot{\mathbf{a}}(t) = -\alpha \mathbf{a}(t) + \mathbf{w}(t), \tag{3.5}$$

where $\mathbf{w}(t)$ is a zero-mean white noise. Similarly, the process acceleration in the discrete-time equation is express in the form of

$$\mathbf{a}_{k+1} = \beta \mathbf{a}_k + \mathbf{w}_k^a, \tag{3.6}$$

where $\mathbf{w}_k^a$ is a zero-mean white noise with variance $\alpha^2(1-\beta^2)$ and $\beta = e^{-\alpha T}$. Augmenting the state with the target acceleration in one dimension of motion, the dynamical model for the

maneuvering mobile user of continuous-time becomes

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\alpha \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{w}(t), \tag{3.7}$$

and equivalently the maneuvering mobile user of discrete-time is given by

$$\mathbf{x}_{k+1} = \mathbf{F}_\alpha \mathbf{x}_k + \mathbf{w}_k = \begin{bmatrix} 1 & T & (\alpha T - 1 + e^{-\alpha T})/\alpha^2 \\ 0 & 1 & (1 - e^{-\alpha T})/\alpha \\ 0 & 0 & e^{-\alpha T} \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k. \tag{3.8}$$

where $\mathbf{F}_\alpha$ is the transition matrix, $\mathbf{x}_k$ is the mobile user state, and $\mathbf{w}_k$ is the process noise. The mobile user state is given by

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ \dot{x_k} \\ \ddot{x_k} \end{bmatrix}, \tag{3.9}$$

where x_k is the mobile user coordinate, $\dot{x_k}$ is the mobile user velocity, and $\ddot{x_k}$ is the mobile user acceleration at time instant k. The process noise is a zero-mean white noise sequence with covariance matrix $\mathbf{Q}_k$ whose elements q_{ij} having the form

$$\begin{aligned}
q_{11} &= \sigma^2(1 - e^{-2\alpha T} + 2\alpha T + \frac{2}{3}\alpha^3 T^3 - 2\alpha^2 T^2 - 4\alpha T e^{-\alpha T})/\alpha^4, \\
q_{12} &= \sigma^2(e^{-2\alpha T} + 1 - 2e^{-\alpha T} + 2\alpha T e^{-\alpha T} - 2\alpha T + \alpha^2 T^2)/\alpha^3, \\
q_{13} &= \sigma^2(1 - e^{-2\alpha T} - 2\alpha T e^{-\alpha T})/\alpha^2, \\
q_{22} &= \sigma^2(4e^{-\alpha T} - 3 - e^{-2\alpha T} + 2\alpha T)/\alpha^2, \\
q_{23} &= \sigma^2(e^{-2\alpha T} + 1 - 2e^{-\alpha T})/\alpha, \\
q_{33} &= \sigma^2(1 - e^{-2\alpha T})/\alpha,
\end{aligned} \tag{3.10}$$

where the maneuvering parameter α is chosen to indicate the type of maneuvering such as lazy turn, evasive turn, etc. The transition matrix can be expanded for more than one dimension of motion.

Semi-Markov Correlated Process Noise Model

This model is based on the modified Singer model but instead of having a zero-mean value, the model allows the target acceleration to have a random switching mean value thus a more realistic model of maneuvering was obtained [90]. Let

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{u}_k + \mathbf{w}_k, \tag{3.11a}$$

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \tag{3.11b}$$

where $\mathbf{F}_k$ is the state transition matrix, $\mathbf{G}_k$ is the command matrix, $\mathbf{H}_k$ is the measurement matrix, and k is the time instant. The matrices are typically known. The control input $\mathbf{u}_k$ is modeled as a finite-state semi-Markov jump process that takes values in a finite discrete set $\{u_1, \ldots, u_N\}$ where u refers to the target acceleration state and N refers to the total number of acceleration state. The correlated process noise is denoted by $\mathbf{w}_k$ and $\mathbf{v}_k$ is the measurement noise. The process noise covariance and the measurement noise covariance are assumed to be known. The possible values of $\mathbf{u}_k$ are quantized into N known levels with known transition probability given by

$$p_{ij} = \Pr\{u_{k+1} = u_j | u_k = u_i\}, \tag{3.12}$$

where $i, j = \{1, \ldots, N\}$. The transition probabilities tend to be sustained for a certain holding time thus its probability distribution is defined by

$$p_{ij}(\tau) = \Pr\{\tau_{ij} \leq \tau\}, \tag{3.13}$$

where $\tau_{ij} = t_k - t_{k-1}$ is the holding time, the time spend before the state transition occurred from u_i to u_j. In order to simplify the semi-Markov jump process, the holding time probability transition is described as having an exponential distribution so that a transition from maneuver state u_i becomes more likely as the time spent in u_i increases. With the assumption, the input process is no longer follow semi-Markov instead it is a Markov process [90].

The control input $\mathbf{u}_k$ in (3.3a) is introduced into the state space form to sufficiently captured the behavior of the system at all time. The state space model is assumed to be nonautonomous (time-varying) when the control input depends upon time explicitly. Conversely, if the control input does not depend upon time explicitly, the state space model is assumed to be autonomous (time-variant). In the correlated process noise (Singer) model and semi-Markov correlated process noise model, the control input described the different types of maneuver in which the acceleration switches between possible levels according to given transition probabilities and with given distributions governing the switching times for each level. Therefore, the control input of the models did not depend upon the time explicitly but rather on the finite discrete set of acceleration.

3.2 Bayesian Filtering

The Tracking Problem

Consider a general SSM given by

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}), \tag{3.14}$$

where $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is a target state vector, n_x is the dimension of the state vector, $\mathbb{R}$ is a set of real numbers, $k = \{0, \ldots, T\} \in \mathbb{N}$ is the time index with T is the final time, $\mathbf{f}_{k-1}$ is typically a non-linear function, and $\mathbf{w}_{k-1}$ is the random process noise. At each time step k, an observation of measurement vector $\mathbf{z}_k \in \mathbb{R}^{n_z}$ with dimension n_z is obtained, given by

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \tag{3.15}$$

where $\mathbf{h}_k$ is typically a non-linear function, and $\mathbf{v}_k$ is the measurement noise. A probabilistic description of (3.14) and (3.15) is given by

$$\mathbf{x}_k \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}), \tag{3.16}$$

$$\mathbf{z}_k \sim p(\mathbf{z}_k|\mathbf{x}_k), \tag{3.17}$$

where $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is the state transition density and $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function. The filtering methods seek to estimate the hidden mobile state $\mathbf{x}_k$ using all the observed measurements $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ up to time k.

To make inferences about $\mathbf{x}_k$ based on $\mathbf{z}_{1:k}$, a joint probability distribution for the two parameters is constructed $p(\mathbf{x}|\mathbf{z}_{1:k})$. Suppose that the joint probability distribution at the previous time step is available $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$, the state predictive distribution can be obtained via the Chapman-Kolmogorov equation, given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1}. \tag{3.18}$$

Next, by using Bayes rule, the filtering posterior state pdf is updated using

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \tag{3.19}$$

where $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$ is a normalization constant, and $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function. The recursive filtering operation of (3.18) and (3.19) form the basis of Bayesian filtering. A filtering method is said to be optimal, in Bayesian filtering, when it obtains a complete characterization of $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ in a recursive manner. Unfortunately, the optimal Bayesian solution is hardly achieved. It is only achievable when the problems are linear and Gaussian distributed. Otherwise, the filtering approaches have to resort to suboptimal methods.

3.3 The Kalman Filter

3.3.1 Background

The KF was invented by R. E. Kalman over five decades ago, and early use of the filter was in the navigation of the Apollo project going to the moon and back. Since then, many variants of nonlinear filters based on the KF have been developed to improve estimation accuracy. Today, KFs have been successfully applied in many applications [92], [93], [94], [95], [96], [97], [98], [99], [100], [101] and they are proven effective and reliable. KFs are relatively easy to design and implement but their performance is affected by several factors such as non-linearities of the model problems, ill-conditioning of the covariance matrix, and inaccurate models of the problem [102].

3.3.2 Basic Algorithm

Restate again the linear state space models (3.2) given by

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k, \tag{3.20}$$

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{3.21}$$

where $\mathbf{x}_k$ represents the state vector, $\mathbf{z}_k$ represents the measurement vector, k represents the time instant, $\mathbf{F}_k$ represents the state transition matrix, and $\mathbf{H}_k$ represents the measurement matrix. The transition matrix and measurement matrix are in a form of linear equations and usually are known. The system state $\mathbf{x}_{k+1}$ is a Gaussian and Markov process. The process noise is denoted by $\mathbf{w}_k$ and the measurement noise is denoted by $\mathbf{v}_k$. Here, the following assumptions are made for the noise processes.

Assumption 1. $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$ are independent processes.

Assumption 2. $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$ are uncorrelated, zero mean, Gaussian processes with known covariance matrices, given by

$$\mathrm{E}[\mathbf{w}_k \mathbf{w}_i^T] = \begin{cases} \mathbf{Q}_k, & \text{if } i = k \\ 0, & \text{if } i \neq j \end{cases}, \tag{3.22}$$

$$\mathrm{E}[\mathbf{v}_k \mathbf{v}_i^T] = \begin{cases} \mathbf{R}_k, & \text{if } i = k \\ 0, & \text{if } i \neq j \end{cases}, \tag{3.23}$$

$$\mathrm{E}[\mathbf{w}_k \mathbf{v}_i^T] = 0, \qquad \text{for all } k \text{ and } i\,. \tag{3.24}$$

where $\mathbf{Q}_k$ is the process noise covariance matrix, and $\mathbf{R}_k$ is the measurement noise covariance matrix.

Initial state description.
Suppose that at time $k = 0$, the state vector initial estimate $\hat{\mathbf{x}}_{0|0}$ and its covariance matrix $\mathbf{P}_{0|0}$ are available such that

$$\hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0], \tag{3.25}$$

$$\mathbf{P}_{0|0} = \mathbb{E}[(\hat{\mathbf{x}}_{0|0} - \mathbf{x}_0)(\hat{\mathbf{x}}_{0|0} - \mathbf{x}_0)^T]. \tag{3.26}$$

where $\mathbf{x}_0$ is a known mean.

Propagation of mean and covariances.
It is assumed that the state and measurement models are jointly Gaussian processes. Therefore, their probabilistic properties are determined by their means and covariances. At time k, the state and covariance estimates are given by

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k \hat{\mathbf{x}}_{k|k}, \tag{3.27}$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{Q}_k. \tag{3.28}$$

The estimated $\hat{\mathbf{x}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ are further improved using the observed sensor measurements $\mathbf{z}_{k+1}$ in accordance with the equations

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}), \tag{3.29}$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T \mathbf{S}_{k+1}^{-1} \mathbf{H}_{k+1} \mathbf{P}_{k+1|k}, \tag{3.30}$$

where $\hat{\mathbf{z}}_{k+1|k} = \mathbf{H}_k \hat{\mathbf{x}}_{k+1|k}$ is the estimated sensor measurements. The difference between the true $\mathbf{z}_{k+1}$ and estimated $\hat{\mathbf{z}}_{k+1|k}$ measurements is called innovation process. The Kalman gain, denoted by $\mathbf{K}_{k+1}$ is derived by minimizing the expectation of (3.26), leads to

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T \mathbf{S}_{k+1}^{-1}, \tag{3.31}$$

where

$$\mathbf{S}_{k+1} = \mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1}. \tag{3.32}$$

Finally, the covariance update in (3.30) can be represented in several different ways [89]. The KF is the optimal estimator, in the minimum mean square error (MMSE) sense, for linear systems with Gaussian random variables [103]. The KF only uses the propagation of mean and covariance in the update steps, which is summarized into two stages, time update $(3.27 - 3.28)$ and measurement update $(3.29 - 3.32)$. This makes KF fairly simple to design for practical applications and often provides good estimation accuracy.

3.4 Kalman Filters for Nonlinear Systems

The implementation of KF for nonlinear systems is important as most if not all real-world problems are nonlinear. However, the direct implementation of KF is not always possible thus required some form of approximation to the problems. The linearization process is the most common approach that was used to approximate the nonlinear functions in the system. Unfortunately, linearization process is only applicable if the error propagation can be approximated well by the linear function and the Jacobian matrices are exists [104]. One of the widely used KF-based estimators for the non-linear systems is the EKF.

3.4.1 Extended Kalman Filters

Consider the general form of SSM, given by

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}, \tag{3.33}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \tag{3.34}$$

where $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ are nonlinear functions, $\mathbf{w}_{k-1}$ is the process noise, and $\mathbf{v}_k$ is the measurement noise. The measurement noise is Gaussianly distributed with known covariance matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$. The nonlinear functions in (3.33) and (3.34) are approximated by the first order Taylor series. The EKF estimates the state and its covariance matrix as follows:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{f}_{k-1}(\mathbf{x}_{k-1|k-1}), \tag{3.35}$$

$$\mathbf{P}_{k|k-1} = \hat{\mathbf{F}}_{k-1}\mathbf{P}_{k-1|k-1}\hat{\mathbf{F}}_{k-1} + \mathbf{Q}_{k-1}, \tag{3.36}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1})), \tag{3.37}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\hat{\mathbf{H}}_k^T\mathbf{S}_k^{-1}\hat{\mathbf{H}}_k\mathbf{P}_{k|k-1}, \tag{3.38}$$

where

$$\mathbf{S}_k = \hat{\mathbf{H}}_k\mathbf{P}_{k|k-1}\hat{\mathbf{H}}_k^T + \mathbf{R}_k, \tag{3.39}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\hat{\mathbf{H}}_k^T\mathbf{S}_k^{-1}. \tag{3.40}$$

The matrices $\hat{\mathbf{F}}_{k-1}$ and $\hat{\mathbf{H}}$ are the Jacobians, obtained by linearizing the nonlinear functions of $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$, respectively, given by

$$\hat{\mathbf{F}}_{k-1} = [\nabla_{k-1}\mathbf{f}_{k-1}^T(\mathbf{x}_{k-1})]^T|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1|k-1}}, \tag{3.41}$$

$$\hat{\mathbf{H}}_k = [\nabla_k\mathbf{h}_k^T(\mathbf{x}_k)]^T|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}, \tag{3.42}$$

where ∇ is the first-order partial derivative operator. Note that equations (3.35) to (3.40) are similar to the KF equations of (3.27) to (3.32). The only difference is that the EKF features the Jacobian matrices $\hat{\mathbf{F}}_{k-1}$ and $\hat{\mathbf{H}}_k$ which are the linearized versions of the nonlinear functions of $\mathbf{f}_{k-1}(\cdot)$ and $\mathbf{h}_k(\cdot)$, respectively. The EKF is based on quadratic approximations to all nonlinearities in the state space equations. When the degree of non-linearities in (3.33) and (3.34) are severe, the tracking accuracy of EKF is reduced. To overcome the shortcoming, higher order of EKF is proposed where the system approximations require calculation of second-order (or higher order) Taylor series [105], [106]. By using the Taylor polynomial of a higher degree in the EKF operation, the system non-linearities are approximated better. Thus, the tracking accuracy achieved by the EKF is improved. Since the quadratic approximations in the state space models required calculation of higher order partial derivatives, the time taken to perform the computation is increased. Eventually, such higher order approximations increase the tracking accuracy but at the expense of increasing the computational time taken to perform the tracking process.

3.4.2 The Unscented Kalman Filters

The UKF is another popular variant of non-linear KF and it is developed to address the flaws of the EKF by representing the state using a probability distribution. By assuming the state is Gaussian distributed, a minimal set of sampling points, called sigma point is determined such that their mean and covariance are the same as the probability distribution. Each sigma points are propagated through the nonlinear system model, and the mean and covariance of the nonlinear transform points are calculated using a method called unscented transformation [39] which is then used to calculate the new state estimate. The UKF estimation accuracy is shown to outperform the EKF, and in term of computational complexity, the UKF and EKF required the same computational cost [102]. In [107], the summary of different UKF variants is presented. Both UKF and EKF are still limited to the assumption of Gaussian noise.

3.5 Particle Filters

3.5.1 Background

A PF is another nonlinear state estimator that is based on the principle of using a probability distribution. It uses sample points called particles to approximate any arbitrary distribution, so it is not limited to a Gaussian assumption. A significant number of particles are required to represent an arbitrary distribution of a system. PFs have existed since the 1950s but due

to the lack of computational power at that time, it makes PF unfavorable [108]. Since the introduction of a modern computing system, PFs have been applied to a wide range of fields [109], [110].

3.5.2 Basic Algorithm

Restate again the general SSM, given by

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k), \tag{3.43}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \tag{3.44}$$

where $\mathbf{x}_k$ is the state vector, $\mathbf{z}_k$ is the measurement vector, $\mathbf{w}_k$ is the process noise, and $\mathbf{v}_k$ is the measurement noise. The process and measurement noises admit arbitrary distributions are described to be independent of each other. The functions $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ are nonlinear and k is the time index.

Inference to Tracking.
To make inferences about the state vector $\mathbf{x}_k$ based on the observed measurement vector $\mathbf{z}_{1:k}$, the joint probability distribution is given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1}. \tag{3.45}$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is the state transition density, and $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ is the prior probability distribution. Solving the integral in (3.45) is intractable hence it is approximated by the used of the PF. Each particle weight is determined so that the sum of all weight is equal to one.

Particle approximation to the posterior.
The PF approximate the filtering distribution of $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ using a set of particles. Suppose that the joint probability distribution at the previous time step $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ is approximated by a set of particles $\{\mathbf{x}_k^{(i)}, i = 1, \ldots, N_p\}$, given by

$$p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(i)}), \tag{3.46}$$

where $\{w_k^{(i)}, i = 1, \ldots, N_p\}$ is the particles weights, N_p is the total number of particles, and $\delta(\cdot)$ is the Dirac delta function. Therefore, the integral in (3.45) is solved and becomes

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto p(\mathbf{z}_k|\mathbf{x}_k) \sum_{i=1}^{N_p} w_{k-1}^{(i)} p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}). \tag{3.47}$$

Next, the particle weights are determined based on the principle of importance sampling [111]. Suppose that the particles is taken from a distribution $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$, the weights are computed according to

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)})p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)}, \tag{3.48}$$

where $p(\mathbf{z}_k|\mathbf{x}_k^{(i)})$ is the likelihood function, $p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$ is the transition prior, and $q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$ is the proposal distribution. In [112], [113], [114], the proposal distribution is chosen such that $q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$, and the weight becomes

$$w_k^{(i)} \propto w_{k-1}^{(i)}\, p(\mathbf{z}_k|\mathbf{x}_k^{(i)}), \tag{3.49}$$

which then are normalized such that $\sum_{i=1}^{N_p} w_k^{(i)} = 1$. Finally, the posterior filtered density is given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx \sum_{i=1}^{N_p} w_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}). \tag{3.50}$$

The PF is affected by degeneracy phenomena [115], where the normalized weights tend to concentrate onto some particles only, after a certain number of recursive steps, leaving the other participle to be essentially degenerate. Therefore, resampling is carried out to reduce the effect of particle degeneracy.

Resampling Methods.

Resampling step was introduced in [112] where particles are discarded or duplicated according to their weights, either at each k or when required. The degree of particle degeneracy can be computed by the effective sample size given by

$$N_{eff} = \frac{1}{\sum_{i=1}^{N_p} (w_k^{(i)})^2}. \tag{3.51}$$

By setting the threshold sample size value, resampling can be applied whenever N_{eff} falls below a threshold. Survey of resampling methods can be found in [115], [116], [117], [118]. The operation of the PF is illustrated in Algorithm 3.1.

Algorithm 3.1 The Particle Filter

(1) Input: Initialization
  for $k = 0$ **do**
    for $i = 1, \ldots, N_p$ **do**
      Samples: $\{\mathbf{x}_0^{(i)} \sim q(\mathbf{x}_0)\}$ where $q(\cdot)$ is the proposal important density.
      Set initial weights: $w_0^{(i)} = 1/N_p$.
    end for
  end for

  for $k = 1, \ldots, endtime$ **do**
    for $i = 1, \ldots, N_p$ **do**
      (2) Prediction Step
      Propagate the particles $\mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{w}_k)$ and the measurement noise $\mathbf{w}_k$, via a function $f_k(\cdot)$.

      (3) Measurement Update
      Calculate the measurement likelihood function:
      $\mathcal{L}\big(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}\big) = \big((2\pi)^n|\hat{\mathbf{R}}_k|\big)^{-\frac{1}{2}}\exp\big(-\frac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})\hat{\mathbf{R}}_k^{-1}(\mathbf{z}-\hat{\mathbf{z}})^T\big)$, where $\mathbf{v}_k \sim (0, \mathbf{R}_k)$.
      Update the weights: $w_k^{(i)} \propto w_{k-1}^{(i)}\mathcal{L}(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)})$.
    end for
    Normalize the weights: $\hat{w}_k^{(i)} = w_k^{(i)} / \sum_{i=1}^{N_p} w_k^{(i)}$.
    The posterior mean: $p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \sum_{i=1}^{N_p} \hat{w}_k^{(i)}\delta_{\hat{\mathbf{x}}_k^{(i)}}$.

    (4) Output: The estimated state: $\hat{\mathbf{x}}_k = p(\mathbf{x}_k|\mathbf{z}_{1:k})$.

    (5) Resampling Step
    Set the threshold sample size: $N_{thresh} = N_p/10$.
    Calculate the effective sample size: $N_{eff} = 1/\sum_{i=1}^{N_p}(\hat{w}_k^{(i)})^2$.
    if $N_{eff} < N_{thresh}$
      Applied resampling methods.
    end if
  end for

3.6 Summary

In this chapter, a theoretical review of the state space models and the filtering methods was presented. The mobile user state is described in a vector form consisting of the coordinates, velocity, and acceleration. On the other hand, the user motion model described the evolution of the mobile user state with respect to time. Since the mobile user state cannot be observed directly, the hidden state is manifested through a set of observed measurements given by the measurement model. The mathematical models of maneuvering target tracking are more practical for implementation since it facilitates more information about the user movements. The maneuver models relate the degree of maneuvering corresponding to the unknown acceleration level of the mobile user. The user acceleration is accounted in the maneuver model by using the control input variable. The control input is unknown and modeled as a random process in a form of white noise models, Markov process models, and semi-Markov jump process models.

The filtering methods make inference about the hidden state of a mobile user based on the observed measurements. The filtering methods can be divided into linear and nonlinear filters. The linear filters such as the KFs are implemented for tracking problems with linear state space models. Meanwhile, the nonlinear filters such as the EKFs, the UKFs, and the PFs are implemented for tracking problems with nonlinear state space models. Each filter has advantages and disadvantages to offer and usually, the selection of filters depending on the situation of tracking. Some of the approaches of using maneuvering model for target tracking include the EKF [35], the two-level hierarchical location prediction algorithm [34]; where the acceleration is modeled as semi-Markov processes, and the sequential Monte Carlo filters [40], [119]; where the acceleration is modeled as a Markov chain with a finite number of possible state. The key to successful target tracking depends on the feasibility of the state space models and the performance of the filtering methods. In this thesis, a mobile tracking approach based on the state space models will be developed to exploit the correlated and sparse wireless measurements in the WSNs.

Chapter 4

Shrinkage based Particle Filter for Tracking with Correlated Sparse Measurements

This chapter describes the development of a mobile tracking approach namely the shrinkage-based particle filter to estimate the dynamic state of a mobile user in the WSNs with correlated and sparse RSS measurements. The material in this chapter has been published in the papers [P1] and [P2].

4.1 Introduction

In wireless networks where sensors are densely deployed, measurements obtained by sensors are expected to be highly correlated in space and time domain. The readings from sensors located approximately close to one another or in the same geographical area will be spatially correlated. In addition to being spatially correlated, the sensor readings will also be temporally correlated due to the time correlation of the transmitted signal or oversampling of the received measurements. The spatio-temporal correlation is a feature that can be exploited for localization and tracking purposes. However, exploiting correlation is often makes it difficult to achieve tractable solutions. The information of the measurement noise covariance matrix is needed in order to exploit the correlation. The covariance matrix can be estimated from the received sensor measurements by using a sample estimator. The sample estimator provides a good estimate of the measurement noise covariance matrix but its performance highly depends

on the number of observations of the sensor measurements. In situations where the number of observation is larger than the dimensionality of the signal model, the sample estimator gives an accurate estimate. However, when the number of observations is comparable or less than the dimensionality of the signal model, the sample covariance matrix estimates become singular and ill-conditioned. The number of observation in the networks is defined as the total number of snapshots that can be obtained at one particular time instant. Meanwhile, the dimensional of the signal model is defined as the total number of sensor nodes that were deployed in the network.

A shrinkage-based particle filter is developed to improve the tracking accuracy of a mobile user by exploiting the spatio-temporal correlation present in the sensor measurements. The developed filter overcomes the ill-problem of measurement noise covariance matrix estimation with the small or comparable number of observation when compared to the dimensionality of the signal model. The main objectives of this chapter are: (i) to exploit the spatial-temporal correlation present in the RSS measurements; (ii) to estimate the measurement noise covariance matrix with the small or comparable number of observation when compared to the dimensionality of the signal model; and (iii) to develop a tracking framework that works with correlated and sparse RSS measurements. The outline of this chapter is as follows: Section 4.2 gives an overview of the proposed framework. Section 4.3 describes the mobile user model and the measurement model. Section 4.4 formulas the spatial and temporal correlation models for generating the correlated sensor measurement in the simulated data. In section 4.5, a shrinkage-based particle filter is developed to improve mobile tracking by exploiting the correlation in the sensor measurements. Section 4.6 reviews the posterior Cramer Rao lower bound for benchmarking purposes. Section 4.7 reports and discuss the results of the simulated and experimental data. Finally, Section 4.8 presents the summary.

4.2 Framework Overview

In this section, the scenario of a wireless network problem is defined and a general overview of the proposed tracking method is given based on which the problem framework is presented. This chapter consider a two-dimensional network consisting of n mobile sensor nodes that are uniformly deployed with known coordinates $(x_i, y_i),\ i \in \{1, \ldots, n\}$. The mobile sensors are uniformly deployed in a sense that they have the same range of radius between one another. This placement enables a full coverage of the monitored area by the sensor nodes. When a target (mobile user) is detected within the monitored area, the mobile sensors will move towards the target. However, the mobile sensors do not cross their designated grid in order to prevent collision with the target. The mobile user dynamic is modeled by the motion model given in Section 4.3.1. This motion model is chosen because it can represent well any ma-

neuvering patterns and has a time auto-correlation function. Once the target was detected, the mobile sensors will generate measurements in the form of RSS as described by the measurement model in Section 4.3.2. The RSS only provides range information between the sensors and the target. Therefore, to compute the coordinate of the target based on the information supplied by the RSS measurements, a filtering method is needed.

Let $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^n]^T$ be the vector of RSS measurements at time instant k and it is assumed that at each time instant k, there is a P number of snapshots of vector measurements were collected in the networks. The sensor measurements are spatially correlated due to the proximity distance between the sensor nodes and the target. Meanwhile, the temporal correlation is due to the total number of snapshots (oversampling) of the received measurements at the sensor node. Section 4.4.3 described the correlation models that are designed to generate spatio-temporally correlated measurements in the simulated data. A filtering method is developed in Section 4.5 to estimate the dynamic state of a mobile user $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ from the vector of RSS measurement. To exploit the correlated measurements, a measurement noise covariance matrix is estimated from the received sensor measurements at every time instant k using a data measurements consisting of a P number of snapshots of the RSS measurements. The number of snapshots in the networks can be varied as a result of sensors and user movements or sensor malfunctions. In situations where the number of P is larger than the number of n, the sample estimator gives an accurate estimate. However, when the number of P is comparable or less than the number of n, the sample covariance matrix estimates become singular and ill-conditioned.

4.3 State Space Models

4.3.1 Motion Model

Let the state vector at time instant k is defined by

$$\mathbf{x}_k = [x_k,\ \dot{x}_k,\ \ddot{x}_k,\ y_k,\ \dot{y}_k,\ \ddot{y}_k]^T, \tag{4.1}$$

where (x_k, y_k), $(\dot{x}_k, \dot{y}_k)$, and $(\ddot{x}_k, \ddot{y}_k)$ represent the user coordinates, velocity, and acceleration, respectively. The dynamics of the mobile user is described by a Singer model [34], [40] given by

$$\mathbf{x}_k = \mathbf{A}(T, \alpha)\mathbf{x}_{k-1} + \mathbf{B}_u(T)\mathbf{u}_k(\mathbf{m}_k) + \mathbf{B}_w(T)\mathbf{w}_k, \tag{4.2}$$

where T is the discretization period, $\alpha = 1/\tau$ is the reciprocal of the maneuvering time constant τ, and $\mathbf{u}_k = [u_{x,k},\ u_{y,k}]^T$ is the control input. A mobile user is moving with a changing speed and acceleration. They can move in the straight direction or take a right turn or/and left

turn at any time instant. The control input is responsible to capture these uncertainties in the mobile user behavior and it is usually unknown and modeled as a first-order Markov process which takes value from a set of models, called system modes $\mathbb{M} = \{\mathbf{m}_1, \ldots, \mathbf{m}_M\}$. Each of the modes is designed to represent the dynamics in the user acceleration behaviors. The transition probabilities are given by $\boldsymbol{\pi}_{ij} = P(\mathbf{u}_k = \mathbf{m}_j | \mathbf{u}_{k-1} = \mathbf{m}_i)$ for $i, j = 1, \ldots, M$ and the initial probabilities are given by $\mu_{i,0} = p(\mathbf{m} = \mathbf{m}_i)$ 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 (4.3), (4.4), and (4.5), respectively

$$\mathbf{A}(T, \alpha) = \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \tilde{\mathbf{A}} \end{bmatrix}, \quad \tilde{\mathbf{A}} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & \alpha \end{bmatrix}, \tag{4.3}$$

$$\mathbf{B}_u(T) = \begin{bmatrix} \tilde{\mathbf{B}}_u & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} & \tilde{\mathbf{B}}_u \end{bmatrix}, \quad \tilde{\mathbf{B}}_u = \begin{bmatrix} T^2/2 \\ T \\ 0 \end{bmatrix}, \tag{4.4}$$

$$\mathbf{B}_w(T) = \begin{bmatrix} \tilde{\mathbf{B}}_w & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} & \tilde{\mathbf{B}}_w \end{bmatrix}, \quad \tilde{\mathbf{B}}_w = \begin{bmatrix} T^2/2 \\ T \\ 1 \end{bmatrix}. \tag{4.5}$$

The process noise is denoted by $\mathbf{w}_k = [w_{x,k}, w_{y,k}]^T$ and it is modeled as a multivariate Gaussian random variable, with zero mean $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ and covariance matrix $\mathbf{Q}_k = \mathbb{E}[\mathbf{w}_k \mathbf{w}_k^T] = \sigma_w^2 \mathbf{I}$ where σ_w^2 is the process noise variance.

4.3.2 Measurement Model

The RSS is modeled as a function of distance using the path loss model [45] given by

$$z = z^0 + 10\beta \log_{10}(d) + v, \tag{4.6}$$

where z^0 is the signal power loss at a reference distance (usually 1-meter), z is the signal power loss at distance d, where d refers to the distance between the sensor node and the mobile user, and β is the path loss exponent which typically $\beta \in [2, 4]$ for urban area. The measurement noise is denoted by $v \sim \mathcal{N}(0, \sigma_v^2)$ and is accounted for the shadow fading. Shadow fading is caused by obstacles such as trees, buildings, and walls along the propagation path and it is found to be a zero mean, Gaussian random variable where σ_v^2 is the shadowing noise variance. The path loss z^0 is a constant and determined during system calibration and the total path loss $z = P_t - P_r$ can be determined by measuring the transmitted power P_t and the received power P_r. A minimum of three sensor measurements is required to perform accurate estimation via

a triangulation method. Therefore, (4.6) can be written in a vector form as

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \tag{4.7}$$

where $\mathbf{z}_k \in \mathbb{R}^n$ is the n sensor measurements at time instant k, i.e., $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^n]^T$ and $\mathbf{h}(\mathbf{x}_k) = [h(x_k^1), h(x_k^2), \ldots, h(x_k^n)]^T$ for $h(x_k^n) = z_k^0 + 10\beta \log_{10}(d_k^n)$. Meanwhile, the shadowing noise component is given by $\mathbf{v}_k = [v_k^1, v_k^2, \ldots, v_k^n]^T$ and having Gaussian distribution $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$, with covariance matrix $\mathbf{R}_k = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T] = \sigma_v^2 \mathbf{I}$ where σ_v^2 is the shadowing noise variance. The shadowing noise is assumed to be spatio-temporally correlated. The measurements only provide range information between the sensor nodes and the mobile user (target). Thus, a filtering method is needed to estimate $\mathbf{x}_k$ from measurements $\mathbf{z}_k \in \mathbb{R}^n$.

4.4 Correlated Shadowing Noise Models

In practice, the correlation present in the RSS measurements is unknown. The shadow fading component in the RSS measurement is spatially correlated due to the fact that the readings from sensors located approximately close to one another or in the same geographical area do not change abruptly. Apart from that, the consecutive measurements that were observed from the same sensors also exhibit temporal correlation. In the simulation, the correlation is induced in the sensor measurements by using the correlation models. The spatial and temporal correlation models are given as follows:

4.4.1 Spatial Correlation

The spatial correlation implies that the readings from sensors located approximately close to one another or in the same geographical area is correlated. Consider two sensors nodes, i-th and j-th received the signals from the mobile user at time instant k, the spatial correlation coefficient between the received signals by the two sensors is given by

$$\rho_k^{i,j} = \exp\left(- \frac{d_k^{i,j}}{D_c} \right), \tag{4.8}$$

where $d_k^{i,j}$ is the relative distance between the two sensors and D_c is the decorrelation distance which is assumed to be known [120]. Thus, the covariance between the measurements at the two sensors is given by

$$C_k^{i,j} = \rho_k^{i,j} \sigma_k^i \sigma_k^j, \tag{4.9}$$

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

4.4.2 Temporal Correlation

The temporal correlation implies that each consecutive observation of the sensor measurements is correlated in time domain. Consider a sensor node i-th received a signal from the same mobile user at two different time instants k and l, the temporal correlation coefficient [121] between the received signals at two time instants is given by

$$\tilde{\rho}^i_{k,l} = \exp\bigg(- \frac{d^i_{k,l}}{D_c} \ln 2 \bigg), \tag{4.10}$$

where $d^i_{k,l}$ is the distance traveled by the mobile user from the time instant k, to the time instant l, which is given by $d^i_{k,l} = \sqrt{(x^i_l - x^i_k)^2 + (y^i_l - y^i_k)^2}$ where (x^i_l, y^i_l) are the user coordinates at time instant l and (x^i_k, y^i_k) are the user coordinates at time instant k. Thus, the covariance between the received signals at two time instants is given by

$$\tilde{C}^i_{k,l} = \tilde{\rho}^i_{k,l} \tilde{\sigma}^i_k \tilde{\sigma}^i_l \ , \tag{4.11}$$

where $\tilde{\sigma}^i_k$ and $\tilde{\sigma}^i_l$ are the shadowing noise standard deviations at the time instant k and l, respectively.

4.4.3 Spatio-Temporal Correlation

Let $\mathbf{u}_k = [u^1_k, u^2_k, \ldots, u^n_k]^T$ be the uncorrelated shadowing noise from n sensor nodes at time instant k. The variable $\mathbf{v}_k = [v^1_k, v^2_k, \ldots, v^n_k]^T$ exhibits spatial correlation when

$$\mathbf{v}_k = \mathbf{C}_k \mathbf{u}_k, \tag{4.12}$$

where the covariance matrix $\mathbf{C}_k$ satisfied the spatial correlation matrix $\mathbf{\Gamma}_k$ at time instant k, such that

$$\mathbf{\Gamma}_k = \mathbf{C}_k \mathbf{C}^T_k, \tag{4.13}$$

where

$$\mathbf{\Gamma}_k = \begin{bmatrix} 1 & \rho^{1,2}_k & \ldots & \rho^{1,n}_k \\ \rho^{2,1}_k & 1 & \ldots & \rho^{2,n}_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n,1}_k & \rho^{n,2}_k & \ldots & 1 \end{bmatrix}. \tag{4.14}$$

The spatial correlation coefficient between the two sensor measurements (links) is given in (4.8). Thus, the covariance matrix that satisfied $\mathbf{\Gamma}_k$ is given by

$$\mathbf{C}_k = \begin{bmatrix} (\sigma_k^1)^2 & \rho_k^{1,2}\sigma_k^1\sigma_k^2 & \cdots & \rho_k^{1,n}\sigma_k^1\sigma_k^n \\ \rho_k^{2,1}\sigma_k^2\sigma_k^1 & (\sigma_k^2)^2 & \cdots & \rho_k^{2,n}\sigma_k^2\sigma_k^n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k^{n,1}\sigma_k^n\sigma_k^1 & \rho_k^{n,2}\sigma_k^n\sigma_k^2 & \cdots & (\sigma_k^n)^2 \end{bmatrix}, \tag{4.15}$$

where the matrix is formulated from (4.9). Both matrices determined the relationship and measured the dependency between two random variables. The correlation matrix $\mathbf{\Gamma}_k$ lies between $[-1, 1]$ and is the normalized version of the covariance matrix $\mathbf{C}_k$. At the next time instant $k+1$, the variable $\mathbf{v}_{k+1} = [v_{k+1}^1, v_{k+1}^2, \ldots, v_{k+1}^n]^T$ exhibits temporal correlation. The temporal correlation coefficient of the received measurements at two time instants is given in (4.10). Let the time instant $l = k+1$, the covariance matrix $\tilde{\mathbf{C}}_{k,l}$ that exhibits temporal correlation can be written as

$$\mathbf{C}_{k,l} = \begin{bmatrix} \tilde{\rho}_{k,l}^1\tilde{\sigma}_k^1\tilde{\sigma}_l^1 & 0 & \cdots & 0 \\ 0 & \tilde{\rho}_{k,l}^2\tilde{\sigma}_k^2\tilde{\sigma}_l^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\rho}_{k,l}^n\tilde{\sigma}_k^n\tilde{\sigma}_l^n \end{bmatrix}. \tag{4.16}$$

where the matrix is formulated from (4.11), k refers to the previous time instant, and l refers to the current time instant. To incorporate both spatio-temporal correlations in the shadowing noise, the covariance matrix in (4.12) has to contain both spatial and temporal coefficients. Therefore, the covariance matrix is constructed as given by

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_k, \mathbf{C}_{k,l}, \ldots \\ \mathbf{C}_{l,k}, \mathbf{C}_l, \ldots \\ \vdots \quad \vdots \quad \ddots \end{bmatrix}, \tag{4.17}$$

where the diagonal elements of the block covariance matrix are of the form of (5.6) which captured the spatial coefficient and the off-diagonal elements of the block covariance matrix are of the form of (4.16) which captured the temporal coefficient. The size of the covariance matrix in (4.17) depends on the measurements retained from all the previous time instants. When more measurement histories were retained, the strength of the temporal correlation is increased. However, this caused the dimensionality of the resulting block covariance matrix grows exponentially. The covariance matrix is required to be positive definite and invertible, otherwise, no solution exists. Moreover, the large dimensionality of the covariance matrix increases the computational cost in the filtering method. Thus, to limit the dimensionality of the resulting covariance matrix, a restriction on the history of previous vector measurements is imposed through a sliding window time.

4.5 Filter Development

4.5.1 Shadowing Noise Covariance Matrix

In order to exploit the correlation present in the RSS measurements, we need the knowledge of the shadowing noise covariance matrix. The estimation of the covariance matrix is not just useful in localization and tracking applications but also in many applications such as finances [122, 123], bioinformatics [124], engineering [125], and etc. A commonly used method to compute a covariance matrix is by using a sample covariance estimator. In general, the sample covariance estimator gives accurate estimates when the total number of parameters to estimate N is less than the total number of data available P. However, when $N \geq P$, the sample covariance matrix estimates becomes ill-conditioned and non-invertible. In the context of target tracking in WSNs, estimating the shadowing noise covariance matrix can become detrimental when the number of observed measurements is less than the dimensionality of the signal model.

To overcome these deficiencies, various methods have been proposed. One of the earliest studies in computing a covariance matrix is of Stein [126]. Stein found that the sample covariance matrix estimates from a zero mean multivariate normal distributions give inaccurate estimates when $N \geq P$, even though the sample estimator is unbiased and positive definite. The largest eigenvalues in the sample covariance matrix tend to be biased upward and the smallest eigenvalues tend to be biased downward. Stein introduces a shrinkage concept by shrinking the eigenvalues in the sample covariance matrix towards a structured estimate. However, Stein approach results in some of the eigenvalues become negative and not arranged in the correct order. Choosing the suitable structure estimate has become basis in the later development of the shrinkage approach.

Sharpe [127] introduces a single index covariance matrix estimator to estimate the covariance matrix of stock returns and Frost and Savarino [128] suggested that the sample covariance matrix is shrinking towards a structured matrix (also known as target matrix) which represents a prior knowledge of stock returns but they did not consider the correlation that might exist between the estimation error on the sample covariance matrix and the target matrix. Ledoit and Wolf [129] proposed an improved version of shrinkage covariance matrix by choosing the single index model to become the target matrix of their shrinkage estimator. The target matrix is combined linearly with the sample covariance matrix by means of shrinkage intensity (weight). The value of the weight is calculated by minimizing a risk function that employs the mean square error between the estimated and true covariance matrices. The estimator is well conditioned and invertible for problems of $N \geq P$, however, when operates in high dimensional matrix settings, the estimator become biased. In later works, Ledoit and Wolf [122, 130] incorporated different types of target matrix to further improve their shrinkage estimator.

Schaffer and Strimmer [124] extended the works of Ledoit and Wolf to the problem of inferring large-scale gene association networks. The authors proposed six target matrices and approaches to compute corresponding optimal shrinkage intensities. The formulated target matrices are well conditioned, positive definite, and approximate well the true covariance matrix. The variants of the target matrices guarantee a minimum mean square error between the estimated and true covariance matrices is achievable. Meanwhile, Fisher and Sun [131] also proposed shrinkage covariance matrix estimators with different variants of target matrices. The estimators are derived based on the assumption that the observations of the random vectors are drawn from the multivariate normal distribution with unknown mean and covariance matrix.

4.5.2 Sample Estimator

Let $\hat{\mathbf{C}}$ be the sample estimator of the true shadowing noise covariance matrix $\mathbf{C}$, given by

$$\hat{\mathbf{C}} = \frac{1}{P-1}\sum_{p=1}^{P}(\mathbf{z}_p - \bar{\mathbf{z}})(\mathbf{z}_p - \bar{\mathbf{z}})^T, \tag{4.18}$$

where $\mathbf{z}_p \in \mathbb{R}^n$ is the vector of sensor measurements that is sampled or observed $p = \{1, \ldots, P\}$ times, and $\bar{\mathbf{z}} = \frac{1}{P}\sum_{p=1}^{P}\mathbf{z}_p$ is the sample mean. The estimated sample covariance matrix in (4.18) is unbiased and provides accurate estimates in a case of $P \gg n$ where n refers to the total number of the sensor nodes. However, when there is only limited number of observation $P \leq n$, the sample covariance matrix estimate is ill-conditioned and not invertible. Therefore, the sample estimator is no longer considered as a good approximation of the true shadowing noise covariance matrix. An estimator for the shadowing noise covariance matrix is required to be both well-conditioned and invertible for a case of $P \leq n$.

4.5.3 Shrinkage Estimator

Let $\hat{\mathbf{S}}$ be the shrinkage estimator of $\mathbf{C}$, given by

$$\hat{\mathbf{S}} = \lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}}, \tag{4.19}$$

where $\hat{\mathbf{C}}$ is the sample estimates, $\mathbf{T}$ is the target matrix, and $\lambda \in [0, 1]$ is the shrinkage intensity (weight). Noticed that the sample estimate in (4.18) is improved by "shrinking the matrix coefficients to" the value supplied by the target matrix. 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 issues arise: (i) how should the target matrix be selected; and (ii) what value should be given to the weight. Schafer and Strimmer summarized six commonly used target matrices and associated estimators of the optimal weight. According to [120], the selection of target matrix should be driven by the data. Thus, in this chapter, two types of target matrices are selected to estimate the shadowing noise covariance matrix. The first target matrix is the diagonal, unit variance shrinkage target matrix given by

$$\mathbf{T}_1 = \mathbf{I}, \tag{4.20}$$

with optimal weight determined by

$$\hat{\lambda}_{\mathbf{T}_1} = \frac{\sum_{ij} \widehat{\mathrm{Var}}\Big(\Big[\hat{\mathbf{C}}\Big]_{ij}\Big)}{\sum_{i\neq j}\Big[\hat{\mathbf{C}}\Big]_{ij}^2 + \sum_i \Big(\Big[\hat{\mathbf{C}}\Big]_{ii} - 1\Big)^2}. \tag{4.21}$$

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

$$\Big[\mathbf{T}_2\Big]_{ij} = \begin{cases} \Big[\hat{\mathbf{C}}\Big]_{ii} & , \quad \text{for } i = j \\ \bar{\rho}\sqrt{\Big[\hat{\mathbf{C}}\Big]_{ii}\Big[\hat{\mathbf{C}}\Big]_{jj}}, & \text{for } i \neq j \end{cases}, \tag{4.22}$$

with optimal weight determined by

$$\hat{\lambda}_{\mathbf{T}_2} = \frac{\sum_{i\neq j} \widehat{\mathrm{Var}}\Big(\Big[\hat{\mathbf{C}}\Big]_{ij}\Big) - \bar{\rho} f_{ij}}{\sum_{i\neq j}\Big(\Big[\hat{\mathbf{C}}\Big]_{ij} - \bar{\rho}\sqrt{\Big[\hat{\mathbf{C}}\Big]_{ii}\Big[\hat{\mathbf{C}}\Big]_{jj}}\Big)^2}, \tag{4.23}$$

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

$$\bar{\rho} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{\Big[\hat{\mathbf{C}}\Big]_{ij}}{\Big[\hat{\mathbf{C}}\Big]_{ii}\Big[\hat{\mathbf{C}}\Big]_{jj}}, \tag{4.24}$$

and f_{ij} is calculated by

$$f_{ij} \quad = \quad \frac{1}{2}\Bigg\{\sqrt{\frac{\Big[\hat{\mathbf{C}}\Big]_{jj}}{\Big[\hat{\mathbf{C}}\Big]_{ii}}}\widehat{\mathrm{Cov}}\Big(\Big[\hat{\mathbf{C}}\Big]_{ii}, \Big[\hat{\mathbf{C}}\Big]_{ij}\Big) \quad + \quad \sqrt{\frac{\Big[\hat{\mathbf{C}}\Big]_{ii}}{\Big[\hat{\mathbf{C}}\Big]_{jj}}}\widehat{\mathrm{Cov}}\Big(\Big[\hat{\mathbf{C}}\Big]_{jj}, \Big[\hat{\mathbf{C}}\Big]_{ij}\Big)\Bigg\}. \tag{4.25}$$

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 A.

4.5.4 The Shrinkage-based Particle Filter

The shrinkage-based particle filter (ShPF) combined the shrinkage estimator and the particle filter to jointly estimate the shadowing noise covariance matrix and the state of the mobile user. The shrinkage estimator overcomes the ill-posed problem of the shadowing noise covariance matrix estimation by the sample estimator when estimating with the small or comparable number of observation when compared to the dimensionality of the signal model. The developed PF employed in the ShPF is adapted from [41] and become the basis for developing the ShPF algorithm. In [41], the PF is developed based on multiple models for the unknown user acceleration behaviors. It is operated with the assumption of fixed shadowing noise covariance matrix in the filter operations. The ShPF extended the filter operation by combining the PF with the shrinkage estimator to exploit the correlation in the sensor measurements. Instead of using a fixed shadowing noise covariance matrix, the ShPF has to estimate the shadowing noise at every time instant k. Consider again the motion and measurement models in Section (4.3), the state vector $\mathbf{x}_k$ is estimated based on the observed sensor measurements $\mathbf{z}_{1:k}$ up to time k. The joint probability distribution of the state and measurement is given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1}. \tag{4.26}$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is the state transition density, and $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ is the prior probability distribution. Solving the integral in (4.26) is intractable hence it is approximated by the PF. Suppose that the joint probability distribution at the previous time step $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ is approximated by a set of particles $\{\mathbf{x}_k^{(i)}, i = 1, \ldots, N_p\}$, given by

$$p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(i)}), \tag{4.27}$$

where $\{w_k^{(i)}, i = 1, \ldots, N_p\}$ is the particles weights, N_p is the total number of particles, and $\delta(\cdot)$ is the Dirac delta function. Therefore, the integral in (4.26) is solved and becomes

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto p(\mathbf{z}_k|\mathbf{x}_k) \sum_{i=1}^{N_p} w_{k-1}^{(i)} p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}). \tag{4.28}$$

Next, if $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$, the weights can be computed according to

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)})p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)}, \tag{4.29}$$

where $p(\mathbf{z}_k|\mathbf{x}_k^{(i)})$ is the likelihood function, $p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$ is the transition prior, and $q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$ is the proposal distribution. In [112], [113], [114], the proposal distribution is chosen such that

$q(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$, and the weight becomes

$$w_k^{(i)} \propto w_{k-1}^{(i)}\, p(\mathbf{z}_k|\mathbf{x}_k^{(i)}), \tag{4.30}$$

which then are normalized such that $\sum_{i=1}^{N_p} w_k^{(i)} = 1$. The likelihood function in (4.30) is computed by

$$p(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}) = \Big((2\pi)^n|\mathbf{S}_k|\Big)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})\mathbf{S}_k^{-1}(\mathbf{z}-\hat{\mathbf{z}})^T\Big), \tag{4.31}$$

where $\mathbf{z}$ and $\hat{\mathbf{z}}$ represent the actual and predicted RSS measurements, respectively, n is the number of sensor nodes, and $\mathbf{S}_k \in \mathbb{R}^{n\times n}$ is the shadowing noise covariance matrix at time instant k. Here, the shadowing noise covariance matrix is estimated by the shrinkage estimator in (4.19). When dealing with a large dimensionality of the covariance matrix, the computational time taken to compute the likelihood function will increase. Thus, to compute efficiently, the size of the estimated shadowing noise covariance matrix is limit by defining the size of the sliding window time, denoted as t_{window}. As a result, the computation of the likelihood function will involves a modified covariance matrix of size $n(t_{window}+1) \times n(t_{window}+1)$ [88]. Finally, the posterior filtered density is given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx \sum_{i=1}^{N_p} w_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}). \tag{4.32}$$

The PF is affected by degeneracy phenomena [115], where the normalized weights tend to concentrate onto some particles only, after a certain number of recursive steps, leaving the other participle to be essentially degenerate. Therefore, resampling is carried out to reduce the effect of particle degeneracy. In ShPF, the residual resampling algorithm [115], [132], is applied. In a nutshell, the operation of the ShPF is divided into three stages; (i) the prediction stage, (ii) the measurement update stage, and (iii) the resampling stage. At prediction stage, particles are propagated according to the motion model. At the measurement update stage, the particle weights are updated using the likelihood function of the measurements. Finally, at the resampling stage, the samples with low important weights were eliminated and samples with high importance weights were duplicated. Algorithm 4.1 describes the proposed ShPF for tracking problems with correlated and sparse RSS measurements.

Algorithm 4.1 The Shrinkage-based Particle Filter (ShPF)

(1) Input: Initialization
  for $k = 0$ **do**
    for $i = 1, \ldots, N_p$ **do**
      Samples: $\{\mathbf{x}_0^{(i)} \sim q(\mathbf{x}_0)\}$ where $q(\cdot)$ is the proposal important density.
      Set initial weights: $w_0^{(i)} = 1/N_p$.
    end for
  end for

  for $k = 1, \ldots, endtime$ **do**
    (2) Shrinkage Estimation Step
    Estimates the shadowing noise covariance matrix by the shrinkage estimator:
    $\hat{\mathbf{S}} = \lambda \mathbf{T} + (1-\lambda)\hat{\mathbf{C}}$ for
    $\mathbf{T} = \begin{cases} \text{as given in } (4.20), & \text{for } \mathbf{T}_1 \\ \text{as given in } (4.22), & \text{for } \mathbf{T}_2 \end{cases}$, and $\hat{\lambda} = \begin{cases} \text{as given in } (4.21), & \text{for } \lambda_{\mathbf{T}_1} \\ \text{as given in } (4.23), & \text{for } \lambda_{\mathbf{T}_2} \end{cases}$.

    for $i = 1, \ldots, N_p$ **do**
      (2) Prediction Step
      Propagate the particles $\hat{\mathbf{x}}_k^{(i)} = \mathbf{A}(T,\alpha)\hat{\mathbf{x}}_{k-1}^{(i)} + \mathbf{B}_u(T)\mathbf{u}_k(\mathbf{m}_k)^{(i)} + \mathbf{B}_w(T)\mathbf{w}_k^{(i)}$
      with the process noise $\mathbf{w}_k^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$,

      (3) Measurement Update
      Calculate the measurement likelihood function:
      $p(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}) = \big((2\pi)^n|\mathbf{S}_k|\big)^{-\frac{1}{2}}\exp\big(-\frac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})\mathbf{S}_k^{-1}(\mathbf{z}-\hat{\mathbf{z}})^T\big)$, where $\mathbf{v}_k \sim (0, \mathbf{R}_k)$.
      Update the weights: $w_k^{(i)} \propto w_{k-1}^{(i)}\, p(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)})$.
    end for
    Normalize the weights: $\hat{w}_k^{(i)} = w_k^{(i)} / \sum_{i=1}^{N_p} w_k^{(i)}$.
    The posterior mean: $p(\mathbf{x}_k|\mathbf{z}_k) = \sum_{i=1}^{N_p} \hat{w}_k^{(i)} \delta_{\hat{\mathbf{x}}_k^{(i)}}$.

    (4) Output: The estimated state: $\hat{\mathbf{x}}_k = p(\mathbf{x}_k|\mathbf{z}_k)$.

    (5) Resampling Step
    Set the threshold sample size: $N_{thresh} = N_p/10$.
    Calculate the effective sample size: $N_{eff} = 1/\sum_{i=1}^{N_p}(\hat{w}_k^{(i)})^2$.
    if $N_{eff} < N_{thresh}$
      The residual resampling method is applied. The method discarded particles with low weights and replicated particles with high weights.
    end if
  end for

4.6 The Posterior Cramer-Rao Lower Bound

The posterior Cramer-Rao lower bound (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) [133]. The implementation of the PCRLB requires knowledge of the true state. However, the EKF and UKF based methods [134] can be applied to approximate the PCRLB for the state error estimates [135]. In this section, theoretical PCRLB is derived and implemented for the simulated trajectory which becomes the standard for benchmarking the performance of the ShPF.

Let $\hat{\mathbf{x}}_k$ be an unbiased estimate of the mobile user state $\mathbf{x}_k$, computed from a set of RSS 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

$$\mathbf{P}_k = \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] \geq \mathbf{J}_k^{-1}, \tag{4.33}$$

where $\mathbf{J}_k \in \mathbb{R}^{n_s \times n_s}$ is the PFIM given by

$$\mathbf{J}_k = \mathbb{E}[[\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)][\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)]^T], \tag{4.34}$$

or equivalently as

$$\mathbf{J}_k = -\mathbb{E}[\nabla_{\mathbf{X}_k}[\nabla_{\mathbf{X}_k} \log\, p(\mathbf{X}_k, \mathbf{Z}_k)]^T], \tag{4.35}$$

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 $\mathbf{Z}_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ is computed as [134]

$$p(\mathbf{X}_k, \mathbf{Z}_k) = p(\mathbf{x}_0) \prod_{i=1}^{k} p(\mathbf{z}_i|\mathbf{x}_i) \prod_{j=1}^{k} p(\mathbf{x}_j|\mathbf{x}_{j-1}), \tag{4.36}$$

where it is determined by the prior density function $p(\mathbf{x}_0)$ of the target initial state $\mathbf{x}_0$ and the conditional density functions of $p(\mathbf{z}_i|\mathbf{x}_i)$ and $p(\mathbf{x}_j|\mathbf{x}_{j-1})$, respectively. Tichavsky et al. [135] proposed a method of computing the PFIM recursively as

$$\mathbf{J}_{k+1} = \mathbf{D}_k^{22} - \mathbf{D}_k^{21}(\mathbf{J}_k + \mathbf{D}_k^{11})^{-1}\mathbf{D}_k^{12}, \tag{4.37}$$

where the terms in (4.37) are defined as

$$\mathbf{D}_k^{11} = -\mathbb{E}[\nabla_{\mathbf{x}_k}[\nabla_{\mathbf{x}_k}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})]^{\mathrm{T}}], \tag{4.38}$$

$$\mathbf{D}_k^{21} = -\mathbb{E}[\nabla_{\mathbf{x}_k}[\nabla_{\mathbf{x}_{k+1}}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})]^{\mathrm{T}}], \tag{4.39}$$

$$\mathbf{D}_k^{12} = -\mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_k}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})]^{\mathrm{T}}] = [\mathbf{D}_{\mathrm{k}}^{21}]^{\mathrm{T}}, \tag{4.40}$$

$$\begin{aligned}\mathbf{D}_k^{22} &= -\mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_{k+1}}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})]^{\mathrm{T}}] \\ &\quad - \mathbb{E}[\nabla_{\mathbf{x}_{k+1}}[\nabla_{\mathbf{x}_{k+1}}\log p(\mathbf{z}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}+1})]^{\mathrm{T}}].\end{aligned} \tag{4.41}$$

Approximated the PFIM.
Consider the nonlinear filtering problems based on the Gaussian assumptions, the log-probability density function of the state and measurement can be respectively formulated by

$$\nabla_{\mathbf{x}_k}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}}) = [\nabla_{\mathbf{x}_k}\mathbf{f}_k^T(\mathbf{x}_k)]\ \mathbf{Q}_k^{-1}\ [\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k)], \tag{4.42}$$

$$\nabla_{\mathbf{x}_k}\log p(\mathbf{z}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}+1}) = [\nabla_{\mathbf{x}_{k+1}}\mathbf{h}_{k+1}^T(\mathbf{x}_k)]\ \mathbf{R}_{k+1}^{-1}\ [\mathbf{z}_{k+1} - \mathbf{h}_{k+1}(\mathbf{x}_{k+1})], \tag{4.43}$$

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 $(4.38) - (4.41)$ are simplified as follows

$$\begin{aligned}\mathbf{D}_k^{11} &= \mathbb{E}[[\nabla_{\mathbf{x}_k}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})][\nabla_{\mathbf{x}_{\mathrm{k}}}\log p(\mathbf{x}_{\mathrm{k}+1}|\mathbf{x}_{\mathrm{k}})]^{\mathrm{T}}], \\ &= \mathbb{E}[[\nabla_{\mathbf{x}_k}\mathbf{f}_k^T(\mathbf{x}_k)]\ \mathbf{Q}_k^{-1}\ [\nabla_{\mathbf{x}_k}\mathbf{f}_k^T(\mathbf{x}_k)]^T], \\ &= \mathbb{E}[\mathbf{F}_k^T\mathbf{Q}_k^{-1}\mathbf{F}_k],\end{aligned} \tag{4.44}$$

$$\mathbf{D}_k^{12} = -\mathbb{E}[\mathbf{F}_k^T]\mathbf{Q}_k^{-1}, \tag{4.45}$$

$$\mathbf{D}_k^{21} = \mathbf{Q}_k^{-1} - \mathbb{E}[\mathbf{F}_k], \tag{4.46}$$

$$\begin{aligned}\mathbf{D}_k^{22} &= \mathbf{Q}_k^{-1} + \mathbb{E}[[\nabla_{\mathbf{x}_k}\mathbf{h}_k^T(\mathbf{x}_k)]\ \mathbf{R}_{k+1}^{-1}\ [\nabla_{\mathbf{x}_k}\mathbf{h}_k^T(\mathbf{x}_k)]^T], \\ &= \mathbf{Q}_k^{-1} + \mathbb{E}[\mathbf{H}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{H}_{k+1}],\end{aligned} \tag{4.47}$$

where $\mathbf{F}_k$ represents the state transition matrix and $\mathbf{H}_{k+1}$ is the Jacobian matrix evaluated at $\mathbf{x}_{k+1}$. Noted that all expressions involved in $(4.44) - (4.47)$ can be evaluated by using the mean and covariance of the state estimate instead of the true state.

The PCRLB for a Deterministic Trajectory.
Consider the case in which the motion model is generated deterministically. Hence, the true state is known and the process noise is zero. The expectation operator in $(4.44) - (4.47)$ can be dropped out. The recursive equation in (4.37) can be rewritten as

$$\mathbf{J}_{k+1} = \mathbf{Q}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1} - \mathbf{Q}_k^{-1} \mathbf{F}_k \Big(\mathbf{J}_k + \mathbf{F}_k^T \mathbf{Q}_k^{-1} \mathbf{F}_k \Big)^{-1} \mathbf{F}_k^T \mathbf{Q}_k^{-1}. \tag{4.48}$$

By applying the matrix inversion lemma, it yields

$$\mathbf{J}_{k+1} = \Big(\mathbf{Q}_k + \mathbf{F}_k \mathbf{J}_k^{-1} \mathbf{F}_k^T \Big)^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}. \tag{4.49}$$

Since $\mathbf{Q}_k = 0$, (4.49) becomes

$$\mathbf{J}_{k+1} = \Big[\mathbf{F}_k^{-1} \Big]^T \mathbf{J}_k \mathbf{F}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}. \tag{4.50}$$

4.6.1 Approximation based on the Extended Kalman Filter

Given the motion model in (4.2) and the measurement model in (4.7), the EKF is applied to estimate the state trajectory of a mobile user using the mean and covariance propagation. The EKF estimates the mean $\hat{\mathbf{x}}_k$ and covariance matrix $\mathbf{P}_k$ at time k recursively as follows [136], [137]. The prior state $\hat{\mathbf{x}}_{k|k-1}$ and covariance $\mathbf{P}_{k|k-1}$ estimate are given respectively by

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_u \mathbf{u}_k, \tag{4.51}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{B}_u \mathbf{Q}_k \mathbf{B}_u^T, \tag{4.52}$$

where $\mathbf{F}_k$ is the state transition matrix, $\mathbf{B}_u$ is the command matrix, and $\mathbf{u}_k$ is the control input. Then, the posterior state $\hat{\mathbf{x}}_{k|k}$ and covariance $\mathbf{P}_{k|k}$ estimate are computed by

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \tag{4.53}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T, \tag{4.54}$$

where

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k, \tag{4.55}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}. \tag{4.56}$$

The difference in the received and predicted RSS measurements $(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})$ is called the innovation process and this process improves the prior state estimates. The correction factor is denoted by the Kalman gain $\mathbf{K}_k$ and the uncertainty of the predicted output is denoted by $\mathbf{S}_k$. 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 Jacobian matrix component is obtained by taking the derivative of $\mathbf{h}(\mathbf{x}_k)$ with respect to the coordinates x_k and y_k and the matrix has a form

$$\mathbf{H}_k = \frac{10\beta}{\ln 10} \begin{bmatrix} \frac{\hat{x}-x_1}{(\hat{x}-x_1)^2+(\hat{y}-y_1)^2} & 0 & 0 & \frac{\hat{y}-y_1}{(\hat{x}-x_1)^2+(\hat{y}-y_1)^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\hat{x}-x_{n_s}}{(\hat{x}-x_{n_s})^2+(\hat{y}-y_{n_s})^2} & 0 & 0 & \frac{\hat{y}-y_{n_s}}{(\hat{x}-x_{n_s})^2+(\hat{y}-y_{n_s})^2} & 0 & 0 \end{bmatrix} . \tag{4.57}$$

After comparing (4.50) with (4.54), by replacing $\mathbf{J}_k$ by $\mathbf{P}_k^{-1}$ and by applying the matrix inversion lemma, the following expression is obtained

$$\mathbf{P}_{k+1}^{-1} = \left(\mathbf{F}_k \mathbf{P}_k \mathbf{F}_k\right)^{-1} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1} \, . \tag{4.58}$$

In (4.50), matrices $\mathbf{F}_k$ and $\mathbf{H}_{k+1}$ are evaluated at the true state but in (4.58), the matrices are evaluated at their estimated state. For a further derivation of the EKF concepts, the reader is referred to Chapter 3, Section 3.4 of the thesis.

4.7 Performance Evaluation

In this section, a simulation and experiment were conducted to measure how well the ShPF perform the task of tracking a mobile user based on the correlated and sparse RSS measurements. The ShPF is expected to track accurately by exploiting the correlation in the shadowing noise. The standard PF is compared against the proposed ShPF to validate the performance of the ShPF in terms of tracking accuracy. The accuracy of state estimates is assessed by the coordinates root mean square error (RMSE) given by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_k^i - x_k^i)^2 + (\hat{y}_k^i - y_k^i)^2} \tag{4.59}$$

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.

4.7.1 Simulation Results and Analysis

Table 4.1: Simulation Parameters for Tracking with Correlated and Sparse RSS Measurements

Comments	**Parameters**
Minimum speed of a mobile user	$v_{min} = 0.3\ ms^{-1}$
Maximum speed of a mobile user	$v_{max} = 5.4\ ms^{-1}$
Number of Particles	$N_p = 500$
Standard deviations of the process noise	$\sigma_w = 0.25\ ms^{-2}$
Standard deviations of the measurement noise	$\sigma_v = [0 - 4]$ dB
Path loss exponent	$\beta = 3$
Decorrelation distance	$D_c = 40$ m
Number of snapshots/observations of sensor measurements	$P = \{3, \ldots, 10\}$

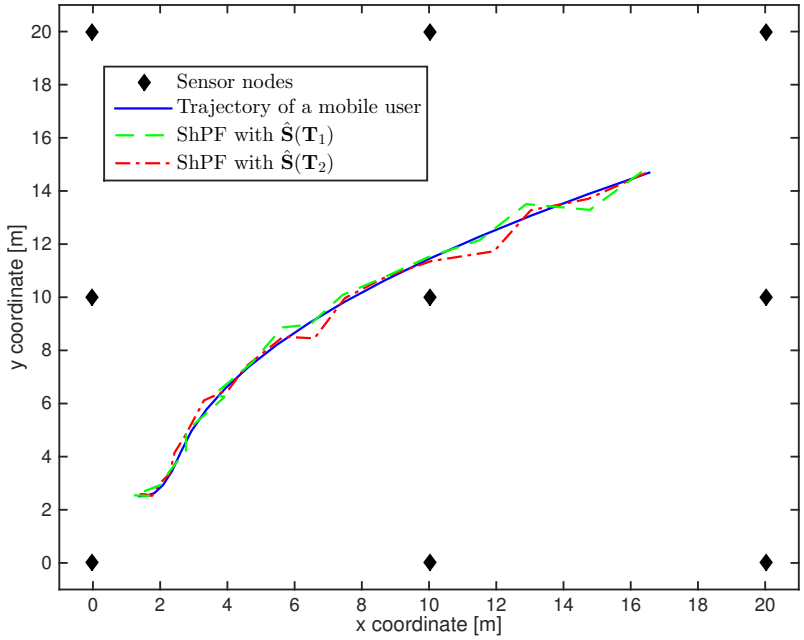

Figure 4.1: Coordinate of the sensor nodes, actual trajectory of the mobile user, and estimated trajectories by the ShPF from a single simulation run.

The simulated network comprises of nine mobile sensor nodes with a coverage radius of 5 meters was set-up as shown in Figure 4.1. The mobile sensors moved with varies speed, specified in the range of $(0.05 - 0.15)$ ms^{-1} and do not cross their designated grid in order to prevent collision with the target. Given the mobile user (target) initial state $\mathbf{x}_0 = [1.3,\ 0.02,\ 0.0,\ 2.5,\ 0.02,\ 0.0]^T$, the trajectory is generated by (4.2) where $T = 0.5$ s, $\alpha = 0.6$, and $k = \{1, \ldots, 10\}$. The control input $\mathbf{u}_k$ in (4.2) is modeled as a Markov process which takes values between the following modes $\mathbb{M} = \{[0.0\ 0.0]^T, [0.5\ 0.0]^T, [0.0\ 0.8]^T\}$ in units of ms^{-2},

with the transition probability matrix $\prod$ is defined by: $\left[\prod\right]_{ii} = 0.5$, for $i = 1, \ldots, 3$ (diagonal elements) and $\left[\prod\right]_{ij} = 0.25$, for $i, j = 1, \ldots, 3$ (off-diagonal elements). Meanwhile, the initial mode probabilities are $\mu_{1,0} = 0.8$ and $\mu_{i,0} = 0.1$ for $i = 2, 3$ such that the sum of elements in each row of the matrix is equal to one. Sensor measurements are generated according to (4.7) where the shadowing noise is Gaussian distributed and spatio-temporally correlated. It is assumed at time instant k, there is a P number of snapshots of sensor measurements are collected in the networks. During the snapshots or observations, the mobile user is assumed static. This assumption is made because the shadowing noise covariance matrix is estimated using the sensor measurements. If only a single observation of sensor measurement is available, the shrinkage method does not have enough data to compute the shadowing noise covariance matrix. At least, the minimum of a $P = 3$ number of snapshots or observations are needed by the shrinkage estimator to estimate the shadowing noise covariance matrix accurately. However, in practical applications, the numbers of snapshots (oversampling) of the received measurements are usually generated during the mobile user movement at different time instants. In this case, a concept of window time is considered to gather more snapshots of the measurements at a predefined size of window time. By using this concept, only the user state at the beginning and end of the window time is considered for the shadowing noise covariance matrix and mobile user state estimation. The size of the window time affects the strength of the temporal correlation between the received measurements.

Figure 4.2 shows the coordinates RMSE comparison of the PF and ShPF methods with its numerical representation is given in Table 4.2. The filtering methods were executed to estimate the mobile user state $\mathbf{x}_k$ at time instant k based on the P number of snapshots of sensor measurement $\mathbf{z}_k$. The simulation was repeated 100 times with random process noise and measurement noise for each run. Thus, an average of the coordinates RMSE is calculated. Table 4.1 shows the other parameter settings for the designed simulation. The PF is similar to the PF developed in [41] where a fixed shadowing noise covariance matrix was assumed in the filter operation. Here, the PF ($\mathbf{C} = \mathbf{I}$) is assumed to operate with the identity matrix because it does not know the correlation that exists between the shadowing noise. Meanwhile, the PF ($\mathbf{C} = \mathbf{C}_0$) is assumed to operate with the true shadowing noise covariance matrix. On the other hand, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) is the estimated shrinkage covariance matrix based on the target matrix in (4.20) and the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$) is the estimated shrinkage covariance matrix based on the target matrix in (4.22). It is shown that the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) and ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$) have a smaller coordinates RMSE when compared to the PF ($\mathbf{C} = \mathbf{I}$) for all number of observations or snapshots. This is because the ShPF has successfully exploited the correlated measurements by utilizing the estimated shadowing noise covariance matrix in the likelihood calculation of the filter operation. The PF ($\mathbf{C} = \mathbf{C}_0$) is the best estimator by exhibits the smallest coordinate RMSE for all number of observation. However, the assumption of using the true representation of the shadowing noise covariance matrix is too optimistic. In practice, it is impossible to know the true shadowing noise covariance matrix. The true shadowing

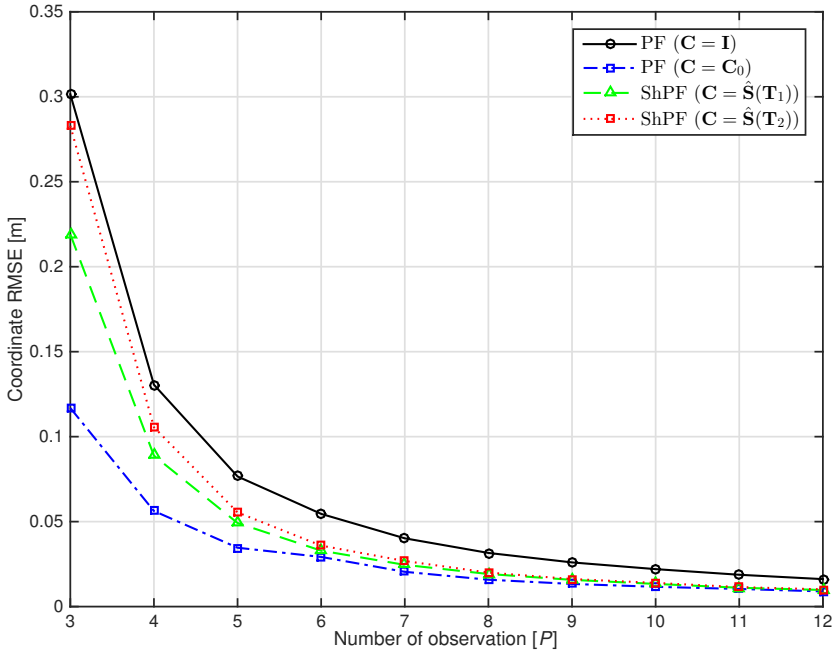

Figure 4.2: Coordinate RMSE comparison of the PF and ShPF using the simulated RSS measurements.

Table 4.2: Averaged coordinate RMSE of the PF and ShPF using the simulated RSS measurements.

Number of observation	PF ($\mathbf{C} = \mathbf{I}$)	PF ($\mathbf{C} = \mathbf{C}_0$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$)
3	0.3014	0.1168	0.2190	0.2829
4	0.1305	0.0562	0.0888	0.1055
5	0.0765	0.0346	0.0491	0.0554
6	0.0545	0.0292	0.0329	0.0358
7	0.0401	0.0204	0.0245	0.0268
8	0.0315	0.0158	0.0191	0.0200
9	0.0260	0.0133	0.0157	0.0162
10	0.0220	0.0116	0.0133	0.0139
11	0.0187	0.0103	0.0110	0.0114
12	0.0161	0.0089	0.0097	0.0101

noise covariance matrix can only be constructed in a designed simulation setup. For that reason, further analysis of the PF ($\mathbf{C} = \mathbf{C}_0$) is not carried out in the rest of the results. The coordinate RMSE of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) is smaller when compared to the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$), especially when estimating the shadowing noise covariance matrix with a small number of observations or snapshots, ($P \leq 5$). This showed that the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) gives a better estimate of the shadowing noise covariance matrix than the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$). In ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$), the shrinkage estimator estimates the shadowing noise covariance matrix

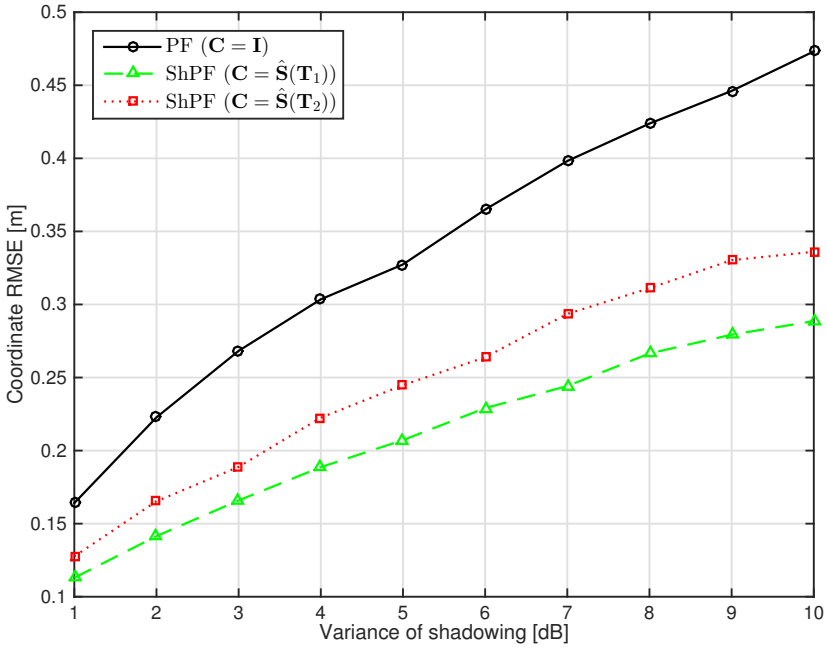

Figure 4.3: Coordinate RMSE comparison of the PF and ShPF for different values of shadowing variance in the RSS measurements.

Table 4.3: Averaged coordinate RMSE of the PF and ShPF for differrent values of shadowing variance in the RSS measurements.

Variance of shadowing	PF ($\mathbf{C} = \mathbf{I}$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$)
1	0.1640	0.1129	0.1276
2	0.2227	0.1414	0.1653
3	0.2682	0.1658	0.1888
4	0.3035	0.1886	0.2225
5	0.3274	0.2071	0.2448
6	0.3649	0.2290	0.2638
7	0.3981	0.2441	0.2935
8	0.4237	0.2664	0.3111
9	0.4461	0.2793	0.3303
10	0.4731	0.2886	0.3360

by assigning an equal weight terms towards both the matrix $\mathbf{T}_1$ and matrix $\hat{\mathbf{C}}$. By contrast, in ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$), the estimated shadowing noise covariance matrix is found to shrink more towards the matrix $\mathbf{T}_2$ and less towards the matrix $\hat{\mathbf{C}}$. Moreover, the optimal weight determined in ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) is calculated by taking all elements of the sample matrix $\hat{\mathbf{C}}$ whereas the optimal weight determined in ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$) is computed by taking only the off-diagonal elements of the sample matrix $\hat{\mathbf{C}}$. Intuitively, an optimal weight minimizes the risk function between the shrinkage estimator and the true covariance matrix.

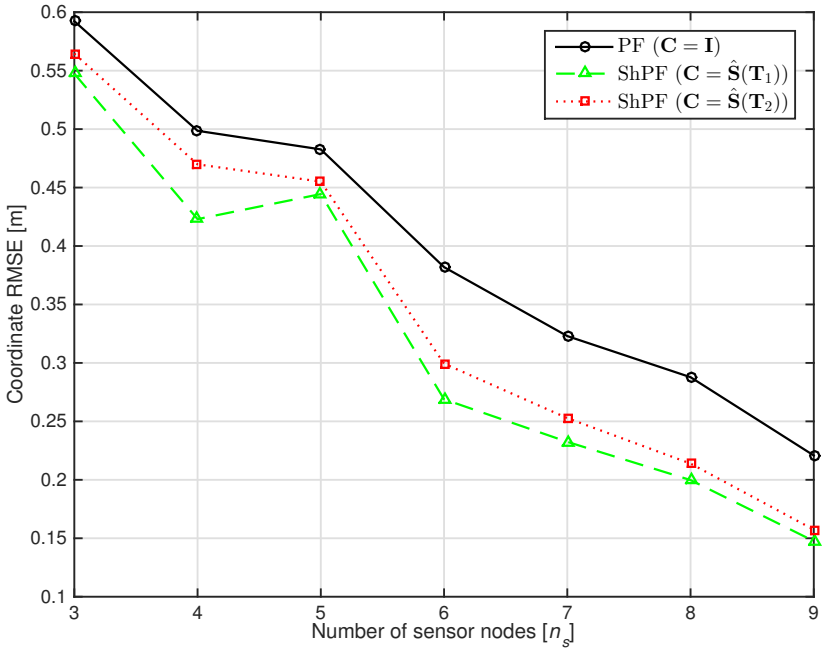

Figure 4.4: Coordinate RMSE comparison of the PF and ShPF for different numbers of sensor nodes in the network.

Table 4.4: Averaged coordinate RMSE of the PF and ShPF for differrent numbers of sensor nodes in the network.

Number of sensor nodes	PF ($\mathbf{C} = \mathbf{I}$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$)
3	0.5923	0.5474	0.5645
4	0.4985	0.4230	0.4698
5	0.4826	0.4444	0.4550
6	0.3819	0.2688	0.2994
7	0.3229	0.2324	0.2527
8	0.2877	0.1997	0.2139
9	0.2208	0.1476	0.1573

Figure 4.3 shows the coordinate RMSE comparison of the PF and ShPF for different values of shadowing variance in the RSS measurements with its numerical representation is given in Table 4.3. It is showed that when the value of the shadowing variance increases, the tracking accuracy decreases for both the PF and ShPF methods. The proposed ShPF gives a better tracking accuracy than the standard PF for all shadowing variance. The ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) has a lower coordinate RMSE when compared to that with the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_2)$).

Figure 4.4 shows the relationship between the number of sensor nodes and the tracking accuracy with its numerical representation is given in Table 4.4. The tracking accuracy for both PF and ShPF methods improves when the number of sensor nodes deployed increases.

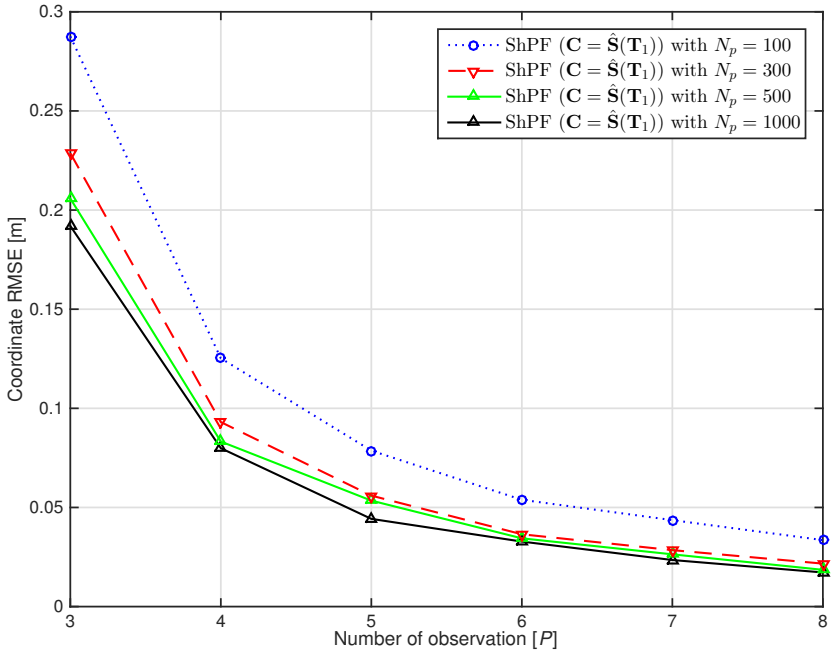

Figure 4.5: Coordinate RMSE comparison of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with $N_p = 100,\ 300,\ 500$ and 1000 particles.

Table 4.5: Averaged coordinate RMSE of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with different number of particles.

Number of observation	$N_p = 100$	$N_p = 300$	$N_p = 500$	$N_p = 1000$
3	0.2875	0.2287	0.2058	0.1919
4	0.1252	0.0930	0.0832	0.0798
5	0.0786	0.0559	0.0534	0.0442
6	0.0538	0.0364	0.0344	0.0327
7	0.0436	0.0284	0.0263	0.0234
8	0.0335	0.0216	0.0185	0.0172

As a result, more RSS measurements were collected and available for the estimation process by the filtering methods. The ShPF has a smaller coordinate RMSE when compared to the PF.

Figure 4.5 shows the coordinate RMSE of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with different number of particles with its numerical representation is given in Table 4.5. On average, when the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) operates with N_p = 100 particles, the tracking accuracy is reduced by almost 45% when compared with the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with N_p = 500 particles for ($P = 3$). However, when the number of particles is increased to N_p = 100 particles, the tracking accuracy increases but at the expense of increasing the computation time. The time taken by the

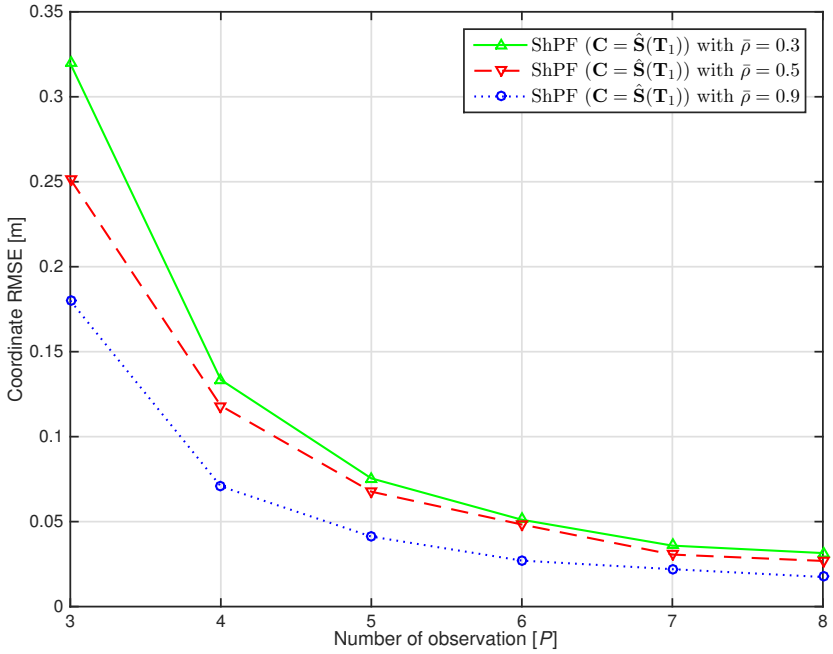

Figure 4.6: Coordinate RMSE comparison of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with the correlation between the RSS measurements is 0.3, 0.5 and 0.9.

Table 4.6: Averaged coordinate RMSE of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with different number of correlation between the RSS measurements.

Number of observation	$\bar{\rho} = 0.3$	$\bar{\rho} = 0.5$	$\bar{\rho} = 0.9$
3	0.3195	0.2515	0.1800
4	0.1334	0.1181	0.0707
5	0.0754	0.0676	0.0411
6	0.0511	0.0482	0.0270
7	0.0358	0.0306	0.0220
8	0.0314	0.0268	0.0174

ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) to complete a single run in the simulation with N_p = 500 particles is 1.37 seconds. The simulation process is executed by means of a desktop computer with an Intel core 3.3 GHz processor, 4 GB RAM, and 465 GB hard drive.

Figure 4.6 displays the relationship between the coordinate RMSE and the measurement correlation with its numerical representation is given in Table 4.6. The RSS measurement is simulated with different correlation coefficients where the spatial coefficient in (4.8) and the temporal coefficient in (4.10) are set to 0.3, 0.5 and 0.9, respectively. When the measurements are highly correlated, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) displays the lowest coordinate RMSE. Contrary, the tracking accuracy is reduced when the measurements are weakly correlated.

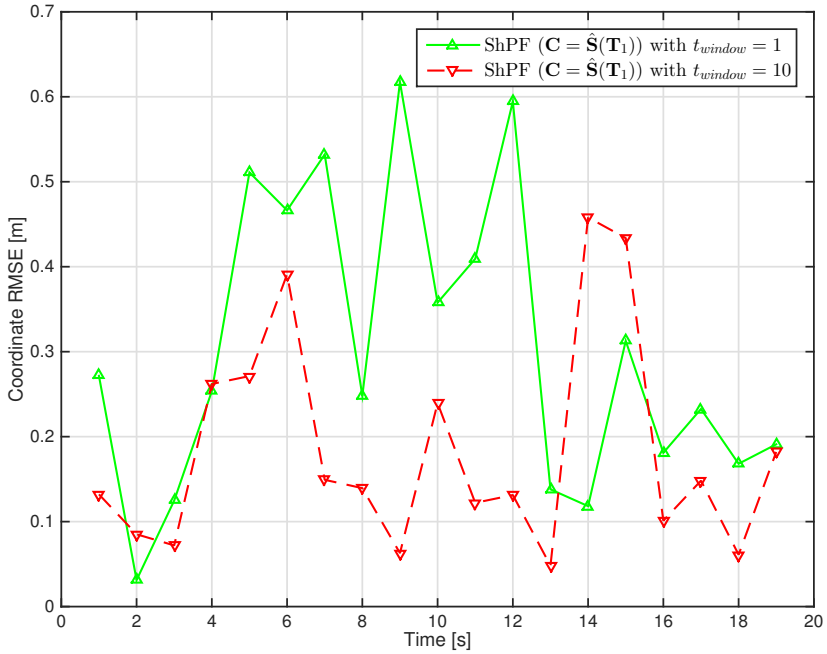

Figure 4.7: Coordinate RMSE comparison of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with window sizes $t_{window} = 1$ and $t_{window} = 10$.

Table 4.7: Averaged coordinate RMSE of the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)) with different window sizes.

Time	$t_{window} = 1$	$t_{window} = 10$
1	0.2716	0.1321
2	0.0312	0.0851
3	0.1248	0.0720
4	0.2535	0.2620
5	0.5110	0.2713
6	0.4663	0.3911
7	0.5321	0.1492
8	0.2480	0.1394
9	0.6173	0.0629
10	0.3577	0.2395
11	0.4093	0.1221
12	0.5945	0.1313
13	0.1380	0.0481
14	0.1182	0.4580
15	0.3128	0.4337
16	0.1810	0.1011
17	0.2326	0.1482
18	0.1681	0.0611
19	0.1910	0.1820

Figure 4.7 analyzed the effects of temporal correlation with its numerical representation is given in Table 4.7. The ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 1$ is compared with the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 10$. 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. The ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 10$ shows a smaller coordinate RMSE than the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 1$. This showed that when more measurement history was retained, better information could be extracted from the shadowing noise covariance matrix. However, by doing so, the dimensionality of the covariance matrix grows exponentially and consequently the time taken to compute the likelihood function increases. The ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 10$ takes 5.97 seconds to complete a single simulation run. On the other hand, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 1$ only takes 2.33 seconds to complete. The sliding window time is set depends on the environment. In urban environments, the correlation coefficient varies more than in suburban and rural environments. Thus, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 10$ is useful. However, in suburban and rural environments, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) with $t_{window} = 1$ is considered to be efficient. The size of the sliding window imposes a trade-off between the tracking accuracy and the computational complexity of the ShPF algorithm.

Finally, Figure 4.8 shows a comparison of the coordinate RMSE of the PF, ShPF, and PCRLB at every time instant k with its numerical representation is given in Table 4.8. The coordinate RMSE of the PF and the ShPF are measured using (4.59). Meanwhile, the PCRLB is calculated using

$$RMSE_{PCRLB} = \sqrt{(\mathbf{P}_k(1,1) + \mathbf{P}_k(4,4))}, \tag{4.60}$$

where $\mathbf{P}_k$ is the covariance matrix of the EKF with the Jacobian evaluated at the true state vector $\mathbf{x}_k$ at time instant k. Here, the PCRLB is introduced as to provide a fundamental lower bound on the variance of an unbiased estimator. The PCRLB sets a useful benchmark which characterizes the lowest possible variance that can be achieved by any estimator. The implementation of the PCRLB requires knowledge of the true state which is not possible in practice. The true state of the mobile user is only available through designed simulation setup. It is shown that the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) provides a smaller coordinate RMSE than the PF ($\mathbf{C} = \mathbf{I}$) for most of the time. This shows that the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) has successfully exploited the correlation in the RSS measurements. On the other hand, the PCRLB sets the lowest theoretical accuracy that can be achieved by any estimator, which is not always possible due to the requirements of the true state. For that reason, the standard PF is considered as the baseline method in this case.

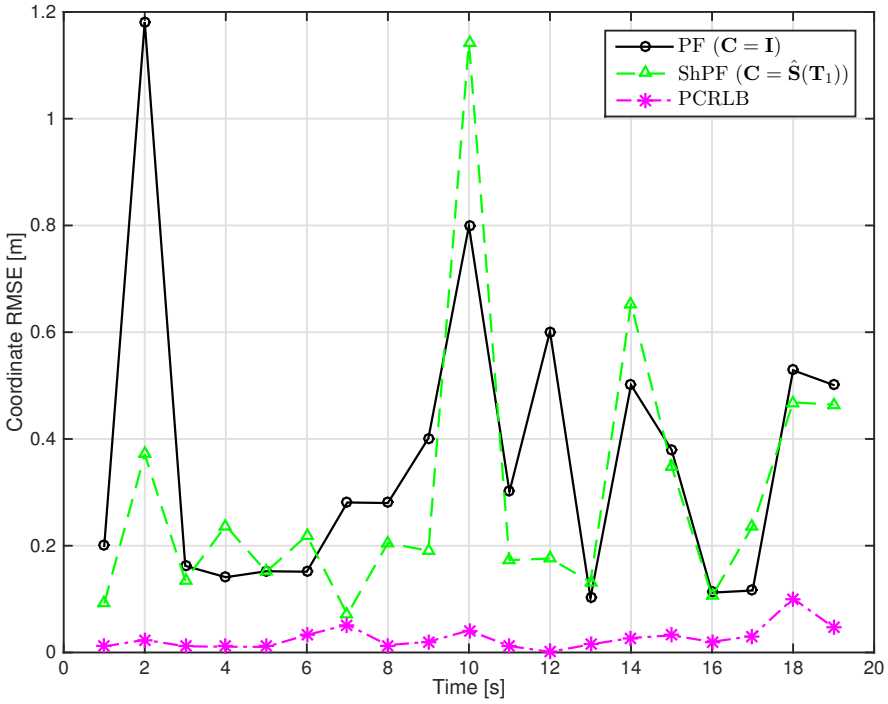

Figure 4.8: Coordinate RMSE comparison of the PF, ShPF, and PCRLB.

Table 4.8: Averaged coordinate RMSE of the PF, ShPF, and PCRLB.

Time	PF ($\mathbf{C} = \mathbf{I}$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)	PCRLB
1	0.1995	0.0917	0.0114
2	1.1810	0.3712	0.0237
3	0.1623	0.1352	0.0115
4	0.1410	0.2375	0.0108
5	0.1522	0.1512	0.0105
6	0.1513	0.2190	0.0327
7	0.2810	0.0711	0.0521
8	0.2800	0.2051	0.0137
9	0.4013	0.1906	0.0201
10	0.7989	1.1423	0.0411
11	0.3012	0.1731	0.0121
12	0.6014	0.1760	0.0011
13	0.1021	0.1321	0.0151
14	0.5011	0.6539	0.0271
15	0.3801	0.3497	0.0321
16	0.1121	0.1053	0.0199
17	0.1161	0.2367	0.0310
18	0.5287	0.4685	0.1010
19	0.5011	0.4643	0.0471

4.7.2 Experimental Results and Analysis

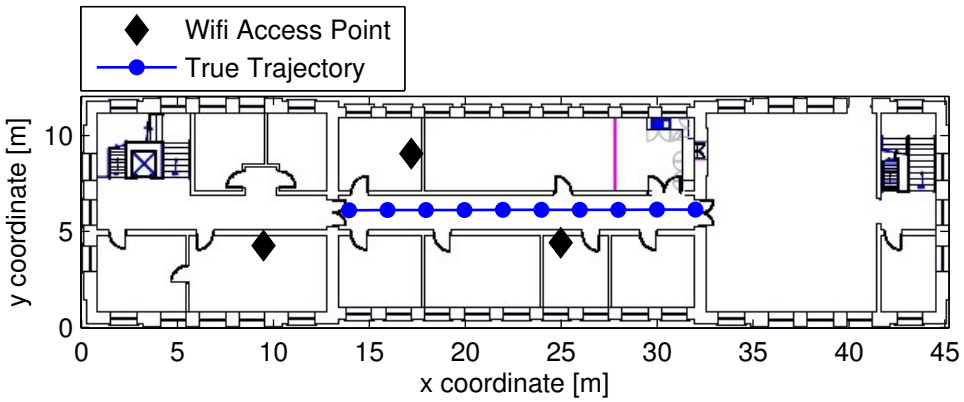

Figure 4.9: Experimental setup.

In this section, the PF and ShPF are validated with the real experimental data. Wi-Fi networks, with Wi-Fi signal strengths, were collected from the D floor of the Amy Johnson Building, at the University of Sheffield, United Kingdom. In the experimental set-up, a user carries a mobile smartphone moving from one end of the corridor to the other end. Three Wi-Fi access points were identified on 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 Linköping University and can be downloaded for free on Google Play [138].

Figure 4.9 displayed the true trajectory of a mobile user and the coordinates of the Wi-Fi access points 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 represent 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 signal noises at each point coordinate in the trajectory. The aim of the PF and ShPF were to estimate the coordinates of the mobile user using the collected Wi-Fi signal strengths. Again, the tracking accuracy is assessed by the coordinate RMSE to validate the performance of the PF and ShPF.

In the motion model, the following parameters were initialized. The user initial state is set to $\mathbf{x}_0 = [12,\ 2.00,\ 0,\ 6,\ 0.01,\ 0.0]^T$ and the trajectory is simulated by (4.2) where $T = 1.0$ s, $\alpha = 0.6$, and $k = \{1, \ldots, 10\}$. The state variance is set to $\sigma_w^2 = 1$ and the control input is set to $\mathbf{u}_k = [0.0\ 0.0]^T$ in units of ms^{-2} with the assumption of the mobile user is moving in a straight direction with a constant velocity in a range of $v = (0.32 - 0.37)$ ms^{-1} and no

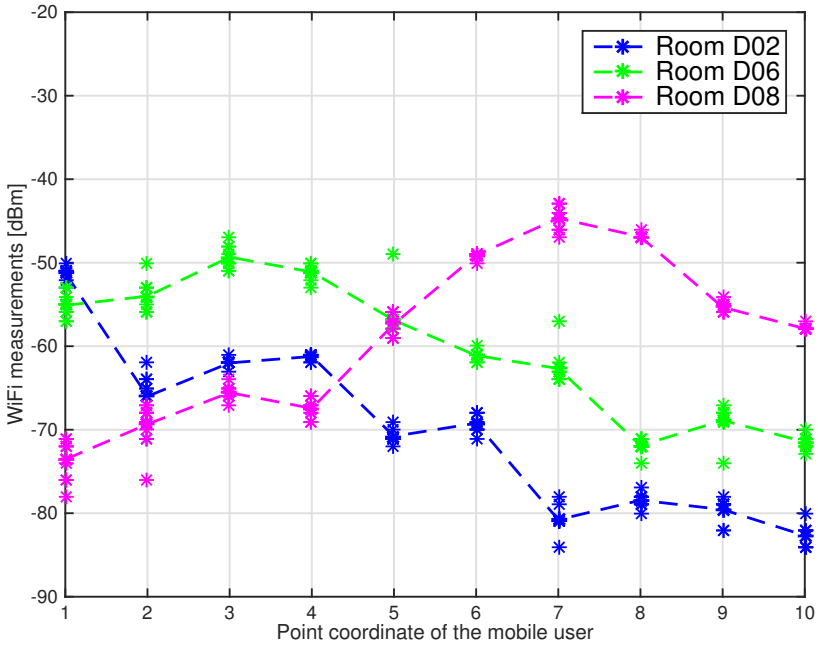

Figure 4.10: The recorded Wi-Fi measurements from the testbed taken at Room D02, D06, and D08 of the Amy Johnson Building.

Table 4.9: Averaged Wi-Fi measurements taken at the Amy Johnson Buildings.

Point coordinate of the mobile user	Room D02	Room D06	Room D08
1	-51.22	-55.11	-73.55
2	-66.00	-54.00	-69.33
3	-62.00	-49.33	-65.55
4	-61.22	-51.11	-67.44
5	-70.77	-56.88	-57.33
6	-69.22	-61.11	-49.11
7	-80.77	-62.66	-44.66
8	-78.44	-71.88	-47.00
9	-79.55	-68.88	-55.33
10	-82.66	-71.44	-57.88

acceleration takes place. Meanwhile, in the measurement model, the parameters were initialized as follows. The signal power loss at a 1-meter distance is set to $z^0 = -36$ dBm and the path loss exponent is set to $\beta = 3$. The actual value of β is usually varied depending on the movements of people on the floor, the location of walls, doors, chairs and even by the construction materials used in the building.

Figure 4.10 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 floor with its

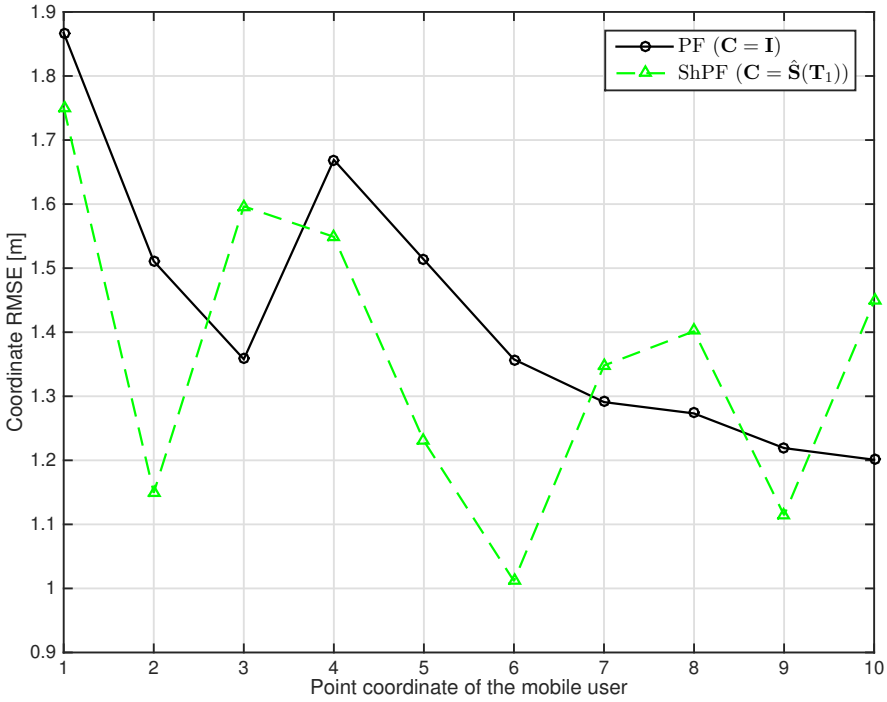

Figure 4.11: Coordinate RMSE comparison of the PF and ShPF using the Wi-Fi measurements.

Table 4.10: Coordinate RMSE of the PF and ShPF using the Wi-Fi measurements.

Point coordinate of the mobile user	PF ($\mathbf{C} = \mathbf{I}$)	ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$)
1	1.8673	1.7510
2	1.5109	1.1493
3	1.3578	1.5965
4	1.6690	1.5489
5	1.5132	1.2303
6	1.3576	1.0115
7	1.2909	1.3480
8	1.2732	1.4015
9	1.2191	1.1143
10	1.2009	1.4482

numerical representation is given in Table 4.9. These signals are received at ten different point coordinates in the trajectory. At each point coordinate, there is a maximum of $(P = 8)$ observations available to be processed by the PF and ShPF with their mean values indicated by the plotted lines. The variations in the observations are due to the shadowing and 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.

Finally, Figure 4.11 compares the coordinate RMSE of the PF ($\mathbf{C} = \mathbf{I}$) and ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) at all point coordinates with its numerical representation is given in Table 4.10. Both filtering algorithms are assumed to know the initial coordinate of the mobile user and used the collected Wi-Fi measurements for estimation process. The PF ($\mathbf{C} = \mathbf{I}$) is deployed by assuming the filter does not know the correlation between the Wi-Fi measurements. On the other hand, the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) uses the shrinkage estimator to exploit the correlation that might exists between the Wi-Fi measurements. It shows that the ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) outperforms the tracking performance of the PF ($\mathbf{C} = \mathbf{I}$) for most of the coordinate estimates. For some coordinates, the tracking accuracy of ShPF ($\mathbf{C} = \hat{\mathbf{S}}(\mathbf{T}_1)$) increases by almost 26% when compared with the PF ($\mathbf{C} = \mathbf{I}$). However, for some point coordinates, the improvement is not significant because the correlation between the Wi-Fi measurements is less than 0.6. Finally, it cannot be ruled out that other factors such as uncertainties in determining the exact coordinates of mobile user and Wi-Fi access points in the building floor might also contribute to the coordinate RMSE readings.

4.8 Summary

In this chapter, the challenging problem of tracking a mobile user using the correlated and sparse RSS measurement was considered. In wireless networks where sensors are densely deployed, measurements obtained by sensors exhibit spatial and temporal correlations. Exploiting the correlation provides extra information to enhance the performance of tracking. However, it would result in an estimation of a covariance matrix that can be problematic when the number of observations is comparable or less than the dimensionality of the signal model. This can happen due to the sensor and user movements or sensor malfunctions in the networks. The number of observation in the networks is defined as the total number of snapshots that can be obtained at one particular time instant. Meanwhile, the dimensional of the signal model is defined as the total number of sensor nodes that were deployed in the network.

A shrinkage-based particle filter (ShPF) was proposed for mobile tracking with correlated and sparse RSS measurements. The method has successfully exploited the correlated measurements by estimating the shadowing noise covariance matrix through the shrinkage estimator. When more historical measurements were retained, the dimensionality of the shadowing noise covariance matrix increases and caused the computational time to increase as well. Thus, to compute efficiently and preserve positive definiteness and invertibility of the covariance matrix, a sliding window time is imposed. The PCRLB is also calculated for simulated data to compare it with the coordinate RMSE of the PF and ShPF. The performance of the ShPF has been validated with the simulation and experimental data. Results from both data showed that the ShPF works best when the RSS measurements are highly correlated and sparse.

Chapter 5

Robust Shrinkage based Particle Filter for Tracking with Non-Gaussian Shadowing Noise

In the previous Chapter 4, the ShPF method had been developed to estimate the dynamic state of a mobile user in the WSNs with correlated and sparse RSS measurements. The developed filter able to track accurately by exploiting the correlated measurements. In this chapter, the performance of the ShPF will be extending to the problems of outliers in the correlated RSS measurements. This chapter describes the development of a robust shrinkage based particle filter to estimate the dynamic state of a mobile user with correlated RSS measurements that are corrupted by outliers. The material in this chapter has been submitted in a paper [P4].

5.1 Introduction

The state space approach to modeling dynamic systems consists of two parts; the state transition model and the measurement model. The system state cannot be observed directly, instead, they are manifested through a set of observed measurements. In WSNs, sensors are deployed to provide indirect observations about the system state. Sensors such as GPS receivers, radars, visual sensing systems, and robotic sensing systems often provide measurements that are not easily interpretable due to the presence of outliers. The outlier is an observation that does not fit with the expected pattern of distribution (usually assumed to be Gaussian) based on the available data. Outliers can come from sudden environmental disturbances, temporary sensor

failures or from the intrinsic noise of the sensor device. The presence of outliers frequently in the measurements can have serious consequences. For examples, the presence of outliers in the dataset caused an inaccurate estimation of the static and moving objects [139], [140], unreliable face recognition [141], imbalance robot controller [142], and damaging the system identification in [143].

The traditional approach to deal with outliers is by using the rejection process [144]. The data is first assumed to follow a Gaussian distribution and any observations that located outside the distribution are treated as residuals. The rejection process minimizes the number of residuals however it creates bias in the parameter estimates. The KF is the optimal estimator for linear Gaussian systems, giving the minimum mean squared error [145]. The performance of KF deteriorates when the measurement is corrupted by outliers. To overcome the effect of outliers, robust KFs were proposed [146], [147]. In [148], [149] a robust KF based on a weighted least-squares approach is presented. Each observation is given a weight which indicates the probability of its contribution towards the estimated state at each time step. The method required the used of heuristics function and manual parameter tuning. A similar approach is adopted in [150] where the weight of observations and the system dynamics are adjusted by using a variational Expectation-Maximization framework [151]. The approach does not require manual parameter tuning however implicit assumptions about the noise distribution were made.

Apart from that, approaches based on resampling techniques [152] and numerical integration [153] were also developed to deal with outliers in the measurements. These methods are complex and required more time to compute thus they are less favorable for real-time applications. Another popular approach is based on the non-parametric methods [154] which used a probability of distribution to represent the state of interest. The non-parametric methods have an advantage in overcoming the constraints of analytical intractable of non-linear models however they are computationally expensive since they required a large number of particles to approximate a state distribution. In [86], the UKF is proposed as the proposal distribution for a PF. The UKF proposal distribution allows the samples in the prior distribution to move to regions of a high likelihood that lies in the tails of the distribution. By doing so, the PF tracks more accurately for the problems of non-Gaussian probability distributions.

Alternatively, the probabilistic description of the state noise and/or measurement noise can be modeled using a Student's t distribution instead of a Gaussian distribution to account for the effects of outliers in the system models. In [155], the measurement noise is modeled to follow a Student's t distribution and the hidden state is approximated based on a structured variational approach [156]. The method has been successfully tested on both simulated and real data and has outperformed the filtering method developed in [150]. The structural variational approach also adopted by [157] to develop a robust filter and smoother for filtering with non-linear state space models. The method provides a good trade-off between accuracy and computational

efficiency but it is less flexible when compared to [158]. In [159], an approach based on the Student's t noise model of both the state and measurement noises is presented. The method is computational inexpensive since it is developed based on the KF framework. However, the values of the degrees of freedom have to be approximated at every time steps to prevent the Student's t filter from converged to the standard KF. The Student's t filter showed promising results in challenging tracking scenarios.

In [160], the measurement noise is modeled using a skewed (asymmetric) and heavy-tailed distributions to account the presence of outliers in the measurements. A skew-t Bayesian filter and smoother were then developed to estimate the hidden state based on the skewed and heavy-tailed distribution of the measurement noise. The method has the lowest root mean square error when compared to the KF. However, as the number of estimated parameters increased, more time is required in the computation process as compared to [157]. In [161], a robust PF is presented which uses two type of Student's t noise models with different degrees of freedom to model the measurement noise. The robust PF is developed based on the Bayesian model averaging approach [162]. The approach overcomes the uncertainty in model selection by employing a number of different noise models instead of a single noise model. The robust PF operates with the assumption of a fixed measurement noise covariance matrix. The method was tested on simulated data and outperformed the standard PF. The robust PF is a recent method which is developed based on an established benchmark method of a Bayesian model averaging approach. Motivated by the work, the robust PF is extended in this chapter by introducing the non-parametric shrinkage estimator.

Known as the robust shrinkage-based particle filter (RSPF), the filter is developed by combining the non-parametric shrinkage estimator with the robust PF to jointly estimate the shadowing noise covariance matrix that is corrupted by the outliers and the state of a mobile user. Therefore, instead of using a fixed and predefined value of the measurement noise covariance matrix as in [161], the RSPF estimates the shadowing noise covariance matrix at every time instants. By exploiting the statistical information on the estimated shadowing noise covariance matrix, the accuracy of the mobile tracking is increased. The main objectives of this chapter are: (i) to address the presence of outliers in the RSS measurements using the Student's t noise model; (ii) to formulate the non-parametric shrinkage estimator for robust estimation of the shadowing noise covariance matrix using the RSS measurements with outliers; and (iii) to develop a tracking framework that works with correlated RSS measurements that are corrupted by outliers. The outline of this chapter is as follows: Section 5.2 give an overview of the proposed framework. Section 5.3 describes the state space models. Section 5.4 formulas the spatial and temporal correlation models. In section 5.5, the presence of outlier in the measurement is addressed and the robust filtering method is developed. Section 5.6 validates the performance of the developed filter and finally, Section 5.7 presents the summary.

5.2 Framework Overview

In this section, the scenario of a wireless network problem is defined and a general overview of the proposed tracking method is given based on which the problem framework is presented. This chapter considers a two dimensional network consisting of n static sensor nodes that are randomly deployed with known coordinates $(x_i, y_i),\ i \in \{1, \ldots, n\}$. The static sensor nodes have an unequal range of radius between one another and were deployed to give full coverage of the mobile user movement. The mobile user dynamic is modeled by the Singer model which can represent well any maneuvering patterns taken by the mobile user and has a time auto-correlation function. When a target (mobile user) is detected within the monitored area by the sensor nodes, the sensors will generate measurements in the form of signal strength. The received signal strength (RSS) measurement at the sensor nodes is modeled as a function of distance using the path loss model. The RSS measurements only provide range information between the sensors and the target. Therefore, to compute the coordinate of the mobile user in the wireless network using the RSS measurements, a filtering method is needed.

Let $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^n]^T$ be the vector of RSS measurements at time instant k and it is assumed that at each time instant k, there is a P number of snapshots of vector measurements were collected in the networks. The RSS measurements are spatio-temporally correlated and corrupted by outliers where an outlier is the sensor readings that do not fit with the expected pattern of a Gaussian distribution. To account for the effect of outliers in the RSS measurements, the measurement noise is modeled using a Student's t distribution instead of a Gaussian distribution. The Student's t distribution has much heavier tails than the Gaussian distribution. The spatial correlation in the RSS measurements is due to the proximity distance between the sensor nodes and the target and the temporal correlation is due to the total number of snapshots (oversampling) of the received measurements at the sensor node. Section 5.4 described the correlation models that are designed to generate spatio-temporally correlated measurements. The objective of our framework is to estimate the state trajectory of a mobile user $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ using the vector of correlated RSS measurements that are corrupted by outliers. A filtering method is developed in Section 5.5 to estimate the dynamic state of a mobile user based on the received sensor measurements. To exploit the correlated measurements, a non-parametric shrinkage estimator is derived. The non-parametric shrinkage estimator estimates the shadowing noise covariance matrix at every time instant k using a data measurements consisting of a P number of snapshots of the RSS measurements. The number of snapshots in the networks can be varied as a result of sensors and user movements or sensor malfunctions. In situations where the number of snapshots P is larger than the number of sensor nodes n, the sample estimator gives an accurate estimate. However, when the number of P is comparable or less than the number of n, the sample covariance matrix estimates become singular and ill-conditioned.

5.3 State Space Models

5.3.1 Motion Model

Let the state vector at time instant k is defined by

$$\mathbf{x}_k = [x_k,\ \dot{x}_k,\ \ddot{x}_k,\ y_k,\ \dot{y}_k,\ \ddot{y}_k]^T, \tag{5.1}$$

where (x_k, y_k), $(\dot{x}_k, \dot{y}_k)$, and $(\ddot{x}_k, \ddot{y}_k)$ represent the user coordinates, velocity, and acceleration, respectively. The dynamics of the mobile user is described by a Singer model and it is chosen because the model can represent well any maneuvering patterns of the mobile user and has a time auto-correlation function. The process noise of the motion model $\mathbf{w}_k \in \mathbb{R}^{n_x}$ admits a zero mean multivariate Gaussian distribution, $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ with covariance matrix $\mathbf{Q}_k = \mathbb{E}[\mathbf{w}_k \mathbf{w}_k^T] = \sigma_w^2 \mathbf{I}$ where σ_w^2 is the process noise variance. Further detail of the Singer model is given in Chapter 4, Section 4.3.1.

5.3.2 Measurement Model

The RSS measurements at the sensor nodes is modeled as a function of distance using the path loss model where the shadowing component of the measurement noise model is assumed to follow a zero mean multivariate Gaussian distribution, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$ with covariance matrix $\mathbf{R}_k = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T] = \sigma_v^2 \mathbf{I}$ where σ_v^2 is the shadowing noise variance. The RSS only provides range information between the sensors and the mobile user. Therefore, to compute the coordinate of the mobile user in the wireless network using the RSS measurements, a filtering method is needed. Further detail of the path loss model is given in Chapter 4, Section 4.3.2.

5.4 Correlated Shadowing Noise Models

The RSS measurements at the sensor nodes are assumed to be correlated in space and time. The correlation comes from the shadowing noise component of the measurement model. Shadowing noise is a result of attenuation signals due to obstacles such as trees, buildings, walls, etc. Let $\mathbf{u}_k = [u_k^1, u_k^2, \ldots, u_k^n]^T$ be the uncorrelated shadowing noise observed by n sensor nodes at time instant k. The variable $\mathbf{v}_k = [v_k^1, v_k^2, \ldots, v_k^n]^T$ exhibits spatial correlation when

$$\mathbf{v}_k = \mathbf{C}_k \mathbf{u}_k, \tag{5.2}$$

where the covariance matrix $\mathbf{C}_k$ satisfied the correlation matrix $\mathbf{\Gamma}_k$ at time instant k, such that

$$\mathbf{\Gamma}_k = \mathbf{C}_k \mathbf{C}_k^T, \tag{5.3}$$

where

$$\mathbf{\Gamma}_k = \begin{bmatrix} 1 & \rho_k^{1,2} & \dots & \rho_k^{1,n} \\ \rho_k^{2,1} & 1 & \dots & \rho_k^{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k^{n,1} & \rho_k^{n,2} & \dots & 1 \end{bmatrix}. \tag{5.4}$$

The spatial correlation coefficient between two sensor nodes (links) is given by [163]

$$\rho_k = \begin{cases} 0.8 - \frac{|\theta|}{150}, & \text{if} \quad \theta \leq 60^\circ \\ 0.6 \quad , & \text{if} \quad \theta > 60^\circ \end{cases}, \tag{5.5}$$

where θ is the angle of arrival between two links. Thus, the covariance matrix that satisfied the correlation matrix is given by

$$\mathbf{C}_k = \begin{bmatrix} (\sigma_k^1)^2 & \rho_k^{1,2}\sigma_k^1\sigma_k^2 & \cdots & \rho_k^{1,n}\sigma_k^1\sigma_k^n \\ \rho_k^{2,1}\sigma_k^2\sigma_k^1 & (\sigma_k^2)^2 & \cdots & \rho_k^{2,n}\sigma_k^2\sigma_k^n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k^{n,1}\sigma_k^n\sigma_k^1 & \rho_k^{n,2}\sigma_k^n\sigma_k^2 & \cdots & (\sigma_k^n)^2 \end{bmatrix}, \tag{5.6}$$

where the variance of the correlated random variables are the same as the variance of the uncorrelated random variables, that is $\sigma_{\mathbf{v}}^2 = \sigma_{\mathbf{u}}^2$. The correlation matrix $\mathbf{\Gamma}_k$ lies between $[-1, 1]$ and is the normalized version of the covariance matrix $\mathbf{C}_k$. At the next time instant $k+1$, the shadowing noise exhibit temporal correlation, given by

$$\begin{bmatrix} v_{k+1}^1 \\ v_{k+1}^2 \\ \vdots \\ v_{k+1}^n \end{bmatrix} = \gamma_{k+1}(\Delta x) \begin{bmatrix} v_k^1 \\ v_k^2 \\ \vdots \\ v_k^n \end{bmatrix} + \begin{bmatrix} u_{k+1}^1 \\ u_{k+1}^2 \\ \vdots \\ u_{k+1}^n \end{bmatrix}, \tag{5.7}$$

where $\gamma_{k+1}(\Delta x)$ is the temporal correlation coefficient [121], given by

$$\gamma_{k+1}(\Delta x) = e^{-\frac{|\Delta x|}{D_c}\ln 2}, \tag{5.8}$$

and Δx is the change in distance by the mobile user from the time instant k to the time instant $k+1$, and D_c is the decorrelation distance.

5.5 Filter Development

5.5.1 Student's t Distributions

The readings of the RSS measurements at the sensor nodes are easily corrupted by the presence of outliers. An outlier is an observation that does not fit with the expected pattern of a Gaussian distribution. Therefore, to account for the effect of outliers in the sensor measurements, the shadowing noise component of the measurement model admits a Student's t distribution instead of a Gaussian distribution. Let $\mathbf{v}_k \in \mathbb{R}^n$ be a zero-mean multivariate Gaussian distribution, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$ with shadowing noise covariance matrix $\mathbf{R}_k$. The probability density function is given by

$$p(\mathbf{v}_k) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\sqrt{|\mathbf{R}_k|}} \exp(-\frac{1}{2}(\mathbf{v}_k - \bar{\mathbf{v}}_k)\mathbf{R}_k^{-1}(\mathbf{v}_k - \bar{\mathbf{v}}_k)^T), \tag{5.9}$$

where $\bar{\mathbf{v}}_k$ is the mean value. Let $\mathbf{y}_k \in \mathbb{R}^n$ be another random vector that have a Gamma distribution, $\mathbf{y}_k \sim \mathcal{G}(\frac{v}{2}, \frac{v}{2})$. Then

$$\mathbf{z}_k = \bar{\mathbf{z}}_k + \frac{\mathbf{v}_k}{\sqrt{\mathbf{y}_k}}, \tag{5.10}$$

follows a multivariate Student's t distributions $\mathbf{z}_k \sim St(\boldsymbol{\mu}, \hat{\boldsymbol{\Sigma}}, v)$ with probability density function given by [164]

$$p(\mathbf{z}_k) = \frac{\Gamma(\frac{v+n}{2})}{\Gamma(\frac{v}{2})} \frac{1}{(v\pi)^{\frac{n}{2}}} \frac{1}{\sqrt{|\mathbf{R}_k|}} \left(1 + \frac{(\mathbf{z}_k - \bar{\mathbf{z}}_k)^T \mathbf{R}_k^{-1}(\mathbf{z}_k - \bar{\mathbf{z}}_k)}{v}\right)^{-\frac{n+v}{2}}, \tag{5.11}$$

where $\bar{\mathbf{z}}_k$ is the mean, $\mathbf{R}_k$ is the scale matrix, and v is the degrees of freedom. The peakness of (5.11) changes by varying the value of v. As $v \to \infty$, the joint probability density function becomes multivariate normal distribution. The covariance of a Student's t random variable is given by $\frac{v}{v-2}\mathbf{R}_k$ for $v > 2$. The Student's t distribution has much heavier tails than the Gaussian distribution and are able to generate outliers for more accurate modeling of the measurement noise.

5.5.2 Non-Parametric Shrinkage Estimator

To exploit the correlation in the RSS measurements, computing the shadowing noise covariance matrix is essential. The non-parametric shrinkage estimator is derived to estimate the shadowing noise covariance matrix using the correlated sensor measurements that are corrupted by outliers. Let $\hat{\mathbf{S}}$ be the shrinkage estimator of the true shadowing noise covariance

matrix, given by

$$\hat{\mathbf{S}} = \lambda \mathbf{T} + (1-\lambda)\hat{\mathbf{C}} \tag{5.12}$$

where $\hat{\mathbf{C}}$ is the sample estimates, $\mathbf{T}$ is the target matrix, and $\lambda \in [1, 0]$ is the shrinkage intensity (weight). The estimated sample covariance matrix $\hat{\mathbf{C}}$ in (5.12) is highly sensitive to the effect of outliers in the measurements. To account for the outliers, an unbiased shrinkage estimator is proposed under non-parametric setups. Suppose that the respective target matrix is chosen

$$\mathbf{T} = \begin{cases} \left[\hat{\mathbf{C}}\right]_{ii}, & \text{if} \quad i = j \\ 0 \quad , & \text{if} \quad i \neq j \end{cases} . \tag{5.13}$$

Under a multivariate Gaussian distribution, the optimal weight is approximated by [124]

$$\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]_{ij}\right)}{\sum_{i \neq j} \left[\hat{\mathbf{C}}^2\right]_{ij}}, \tag{5.14}$$

where the derivation of $\widehat{\mathrm{Var}}\left(\left[\hat{\mathbf{C}}\right]\right)$ and $\hat{\mathbf{C}}$ is given in Appendix A. In contrast, under non-parametric setups, the optimal weight is approximated by

$$\hat{\lambda} = \frac{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \frac{\mathrm{tr}[\hat{\mathbf{C}}\mathbf{T}]}{n} - \hat{a}_1 + \hat{b}_1}{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \hat{c}_1}, \tag{5.15}$$

where

$$\hat{a}_1 = \frac{(P-1)}{P(P-2)(P-3)}\Big((P-1)(P-2)\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} + n\hat{c}_1 - \frac{PQ}{n}\Big), \tag{5.16}$$

$$\hat{b}_1 = \frac{P-1}{(P+1)n}\mathrm{tr}[\mathbf{T}^2] - \frac{1}{P+1}\hat{b}_2, \tag{5.17}$$

$$\hat{b}_2 = \frac{-1}{(P-2)(P-3)n} \times \Big(2(P-1)^2\mathrm{tr}[\hat{\mathbf{C}}^2] + (P-1)^2(\mathrm{tr}[\hat{\mathbf{C}}])^2 - P(P+1)Q\Big), \tag{5.18}$$

$$\hat{c}_1 = \frac{P-1}{P(P-2)(P-3)n^2}\Big(2\mathrm{tr}[\hat{\mathbf{C}}] + (P^2-3P+1)(\mathrm{tr}[\hat{\mathbf{C}}])^2 - PQ\Big), \tag{5.19}$$

$$\tag{5.20}$$

and

$$Q := \frac{1}{P-1}\sum_{p=1}^{P}\Big((\mathbf{z}_p - \bar{\mathbf{z}})^T(\mathbf{z}_p - \bar{\mathbf{z}})\Big)^2, \tag{5.21}$$

is proposed by [165] as another statistics to solve the estimation of $\hat{a}_1$. The parameter n refers to the number of sensor nodes and P refers to the total number of snapshots of the RSS measurements. Further derivation of (5.15) is given in the Appendix B.

5.5.3 Multiple Models Particle Filter

In target tracking, the maneuvering patterns of the mobile user are not known. The maneuvering patterns are assumed to follow a few possible patterns of operation based on some relevant information such as velocity, driving situations, driving manners, routes, and etc. In [136], the maneuvering patterns are grouped into a few possible patterns known as the modes (or regimes) of operation. Let, the state dynamic and measurement model are given by

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{m}_k, \mathbf{w}_{k-1}), \tag{5.22}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{m}_k, \mathbf{v}_k), \tag{5.23}$$

where $\mathbf{x}_k$ is the state vector, $\mathbf{w}_{k-1}$ is the process noise, $\mathbf{z}_k$ is the measurement, $\mathbf{v}_k$ is the measurement noise, and $\mathbf{m}_k$ is the modes (or regimes) that is in effect during the transition period of the mobile user from the time instant $k-1$ to the time instant k. The mode is commonly modeled by a time-homogeneous m-state first-order Markov chain with transitional probabilities given by

$$\boldsymbol{\pi}_{ij} = p(\mathbf{m}_k = j | \mathbf{m}_{k-1} = i) \quad (i, j \in \mathbb{M}), \tag{5.24}$$

where $M = \{1, 2, \ldots, m\}$ for M is the total number of modes. Thus, the transition probability matrix $\boldsymbol{\pi}_{ij}$ is a matrix of size $(M \times M)$ with each element of $i, j \in \mathbb{M}$ satisfying

$$\boldsymbol{\pi}_{ij} \geq 0 \quad \text{and} \quad \sum_{j=1}^{s} \boldsymbol{\pi}_{ij} = 1. \tag{5.25}$$

The initial probabilities of the modes are given by

$$\mu_i = p(\mathbf{m}_1 = i), \tag{5.26}$$

for $i \in \mathbb{M}$, such that

$$\mu_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{s} \mu_i = 1. \tag{5.27}$$

With the modes of operation, the state vector becomes an augmented hybrid state vector $\mathbf{y}_k = [\mathbf{x}_k^T, \mathbf{m}_k]^T$ with the recursive solution for the hybrid state estimation is given by

Prediction:

$$p(\mathbf{x}_k, \mathbf{m}_k = j | \mathbf{Z}_{k-1}) = \sum_i \boldsymbol{\pi}_{ij} \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{m}_k = j)\, p(\mathbf{x}_{k-1}, \mathbf{m}_{k-1} = i | \mathbf{Z}_{k-1})\, d\mathbf{x}_{k-1}, \tag{5.28}$$

Update:

$$p(\mathbf{x}_k, \mathbf{m}_k = j | \mathbf{Z}_k) = \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}_k = j)\, p(\mathbf{x}_k, \mathbf{m}_k = j | \mathbf{Z}_{k-1})}{\sum_j \int p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}_k = j)\, p(\mathbf{x}_k, \mathbf{m}_k = j | \mathbf{Z}_{k-1}) d\mathbf{x}_k} \tag{5.29}$$

where $p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{m}_k = j)$ is the likelihood function of the mode $\mathbf{m}_k = j$, $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{m}_k = j)$ is the state transition density of the mode $\mathbf{m}_k = j$, and $p(\mathbf{x}_{k-1}, \mathbf{m}_{k-1} = i|\mathbf{Z}_{k-1})$ is the prior probability distribution of the mode $\mathbf{m}_{k-1} = i$ based on the observed measurements $\mathbf{Z}_{k-1}$ up to time $k-1$. The solution defined by (5.28) and (5.29) are quite similar with (3.46) and (3.50) of the particle filters where the only difference being that the state vector is now $\mathbf{y}_k = [\mathbf{x}_k^T, \mathbf{m}_k]^T$ instead of $\mathbf{y}_k = [\mathbf{x}_k^T]^T$.

5.5.4 Bayesian Model Averaging

The Bayesian model averaging approach [162] provides solutions to the model uncertainty problems in selecting the best model for use at a given time. The shadowing component of the measurement model given in Section 5.3.2 can admit either a Gaussian distribution or a Student's t distribution depending on whether an outlier is a presence or not in the measurement data. If no outliers, the filtering method can assume that the measurement noise is Gaussian distributed. Otherwise, when the outliers presence, it is best for the filtering method to assume a Student's t distribution instead of a Gaussian distribution. However, choosing which a noise model to follow at a given time is heavily depends on the knowledge of the outlier in the sensor measurements. Therefore, if the outlier is the quantity of interest which is not readily available, then its posterior distribution given the measurement data D, is given by

$$p(o|D) = \sum_{s=1}^{S} p(o|M_s, D)\, p(M_s|D), \tag{5.30}$$

where o is the outlier, M_s is the noise model, and $s = \{1, \ldots, S\}$ is the number of model; for S is the total model. The posterior distribution in (5.30) gives an average of the posterior distributions under the consideration of all the noise models. Furthermore, the posterior probability for each model M_s is given by

$$p(M_s|D) = \frac{p(D|M_s)\, p(M_s)}{\sum_{l=1}^{S} p(D|M_l)\, p(M_l)}, \tag{5.31}$$

where

$$p(D|M_s) = \int p(D|\theta_s, M_s)\, p(\theta_s|M_s)\, d\theta_s \tag{5.32}$$

is the integrated likelihood of the model M_s, θ_s is the vector parameters of the model M_s, and $p(M_s)$ is the prior probability that M_s is the true model. By averaging all the possible distributions of measurement noise using the Bayesian model averaging approach, it provides a better average prediction over the used of any single distribution of noise model.

5.5.5 The Robust Shrinkage Particle Filter

The measurement noise is often assumed to follow a Gaussian distribution with a fixed and predefined value of the shadowing noise covariance matrix. When the sensor measurements are corrupted by outliers, the assumption of normal distribution is no longer hold. Moreover, the elements of the shadowing noise covariance matrix are changing frequently due to the movement of the mobile user in the networks. Thus, the assumption of a fixed shadowing noise covariance matrix is only held true when dealing with a static user. Hence, a robust shrinkage particle filter (RSPF) is developed to overcome the problems. Here two issues are identified and tackled: (i) how to address the model noise uncertainty due to outliers; and (ii) how to approximate the shadowing noise covariance matrix due to outliers. The first issue is addressed by using multiple models particle filter and Bayesian model averaging approaches as described in Section 5.5.3 and Section 5.5.4. The second issue is addressed by using the non-parametric shrinkage estimator as described in Section 5.5.2.

Let $\mathcal{S} = \{s_1, s_2\}$ denotes the multiple noise models where s_1 is based on the multivariate Gaussian distribution and s_2 is based on the multivariate Student's t distribution. Let also $\mathcal{H}_k = s$ denotes the event that the s-th noise model is chosen at time instant k. Suppose that at time instant $k-1$, the following distribution is obtained

$$\begin{aligned} p(\mathbf{x}_{k-1}|\mathbf{z}_{k-1}) &= \sum_{s=1}^{S} p(\mathbf{x}_{s,k-1}|\mathcal{H}_{k-1} = s, \mathbf{z}_{k-1})\; p(\mathcal{H}_{k-1} = s|\mathbf{z}_{k-1}), \\ &= \sum_{s=1}^{S} p_s(\mathbf{x}_{s,k-1}|\mathbf{z}_{k-1})\; p(\mathcal{H}_{k-1} = s|\mathbf{z}_{k-1}), \end{aligned} \tag{5.33}$$

where (5.33) is approximated by a set of weighted particles $\{\mathbf{x}_{k-1}^{(i)}, w_{k-1}^{(i)}\}$, for $i = 1, \ldots, N_p$, and a noise model transition probability $\boldsymbol{\pi}_{s,k-1}$, for $s = \{1, 2\}$, given by

$$p(\mathbf{x}_{k-1}|\mathbf{z}_{k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(i)}) \sum_{s=1}^{S} \boldsymbol{\pi}_{s,k-1}. \tag{5.34}$$

At time instant k, the particles are drawn from a proposal distribution $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$, and under the hypothesis $\mathcal{H}_k = s$, the weights can be recursively computed according to

$$w_{s,k}^{(i)} \propto w_{k-1}^{(i)} \frac{p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}) p(\hat{\mathbf{x}}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{q(\hat{\mathbf{x}}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)}, \tag{5.35}$$

where $p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)})$ denotes the likelihood function associated to the s-th noise model, given by

$$p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}) = \begin{cases} \text{eq. (5.9)} & \text{for Gaussian distribution} \\ \text{eq. (5.11)} & \text{for Student's } t \text{ distribution} \end{cases}$$

It is shown that the computation of the likelihood functions required information about the shadowing noise covariance matrix, which is not available. Therefore, an approximation of the shadowing noise covariance matrix is required to exploit the correlation in the measurements. Suppose that the target matrix (5.13) is chosen, the optimal weight under non-parametric setups is given in (5.15). Next, by choosing the proposal important density $q(\hat{\mathbf{x}}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = p(\hat{\mathbf{x}}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$, the weights simplifies to

$$w_{s,k}^{(i)} \propto w_{k-1}^{(i)} p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}), \tag{5.36}$$

and normalized by

$$\hat{w}_{s,k}^{(i)} = w_{s,k}^{(i)} / \sum_{i=1}^{N_p} w_{s,k}^{(i)}. \tag{5.37}$$

On the other hand, the transition probability of the s-th noise model at time step $k-1$ is given by

$$\boldsymbol{\pi}_{s,k-1} = \frac{\boldsymbol{\pi}_{s,k-1}^{\varphi}}{\sum_{s=1}^{S} \boldsymbol{\pi}_{s,k-1}^{\varphi}}, \tag{5.38}$$

where $\varphi \in [0, 1]$ denotes the forgetting factor [166]. By employing Bayes' rule, the transition probability of the s-th noise model is recursively computed by

$$\boldsymbol{\pi}_{s,k} = \frac{\boldsymbol{\pi}_{s,k-1}^{\varphi}\, p_s(\mathbf{z}_k|\mathbf{z}_{k-1})}{\sum_{s=1}^{S} \boldsymbol{\pi}_{s,k-1}^{\varphi}\, p_s(\mathbf{z}_k|\mathbf{z}_{k-1})}, \tag{5.39}$$

where $p_s(\mathbf{z}_k|\mathbf{z}_{k-1})$ is the marginal likelihood of the s-th noise model, defined by

$$p_s(\mathbf{z}_k|\mathbf{z}_{k-1}) = \int p_s(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{k-1}) d\mathbf{x}_k. \tag{5.40}$$

Since $p(\mathbf{x}_k|\mathbf{z}_{k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)} \delta_{\hat{\mathbf{x}}_k^{(i)}}$, the integral in (5.40) is approximated by

$$p_s(\mathbf{z}_k|\mathbf{z}_{k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)}\, p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}). \tag{5.41}$$

Finally, the posterior filtering density at time step k is formulated as

$$p(\mathbf{x}_k|\mathbf{z}_k) = \sum_{s=1}^{S} p_s(\mathbf{x}_{s,k}|\mathbf{z}_k)\ \boldsymbol{\pi}_{s,k}, \tag{5.42}$$

$$= \sum_{s=1}^{S} w_{s,k}^{(i)} \delta_{\hat{\mathbf{x}}_k^{(i)}}\ \boldsymbol{\pi}_{s,k}. \tag{5.43}$$

Algorithm 5.1 summarises the RSPF.

Algorithm 5.1 The Robust Shrinkage Particle Filter (RSPF)

(1) Input: Initialization
for $k = 0$ **do**
  for $i = 1, \ldots, N_p$ **do**
    Samples: $\{\mathbf{x}_0^{(i)} \sim q(\mathbf{x}_0), \mathbf{m}_0^{(i)} \sim P_0(\mathbf{m})\}$ where $P_0(\mathbf{m})$ are the initial mode probabilities for $\mathbf{m}_0$. Set initial weights: $w_0^{(i)} = 1/N_p$.
  end for
  Set initial noise models transition probabilities, $\boldsymbol{\pi}_{s,0}$ for $s = \{1, 2\}$.
end for

for $k = 1, \ldots, endtime$ **do**
  (2) Shrinkage Estimation Step
  Shadowing noise covariance estimates: $\hat{\mathbf{S}}_k = (1 - \hat{\lambda}_k)\hat{\mathbf{C}}_k + \hat{\lambda}_k \mathbf{T}_k$, for
  $\mathbf{T}_k = \begin{cases} [\hat{\mathbf{C}}]_{ii}, & \text{if} \quad i = j \\ 0\,, & \text{if} \quad i \neq j \end{cases}$, and $\hat{\lambda}_k =$ as given in (5.15).

  for $i = 1, \ldots, N_p$ **do**
    (3) Prediction State
    Propagate the particles $\hat{\mathbf{x}}_k^{(i)}$ according to the motion model.

    (4) Measurement Update
    Calculate the likelihood functions: $p_s\big(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}\big) = \begin{cases} \text{as given in } (5.9), & \text{for } s_1. \\ \text{as given in } (5.11), & \text{for } s_2. \end{cases}$
    Calculate the marginal likelihood of noise models:
    $p_s(\mathbf{z}_k|\mathbf{z}_{k-1}) \approx \sum_{i=1}^{N_p} w_{k-1}^{(i)}\, p_s\big(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)}\big)$.
    Update the weights: $w_{s,k}^{(i)} \propto w_{k-1}^{(i)} p_s(\mathbf{z}_k|\hat{\mathbf{x}}_k^{(i)})$.
  end for
  Normalize the weights: $\hat{w}_{s,k}^{(i)} = w_{s,k}^{(i)} / \sum_{i=1}^{N_p} w_{s,k}^{(i)}$.

  (5) Model Averaging Step
  Specify noise models transition probabilities: $\boldsymbol{\pi}_{s,k-1} = \frac{\boldsymbol{\pi}_{s,k-1}^{\varphi}}{\sum_{s=1}^{S} \boldsymbol{\pi}_{s,k-1}^{\varphi}}$, and $\varphi \in [0, 1]$.
  Update noise models transition probabilities: $\boldsymbol{\pi}_{s,k} = \frac{\boldsymbol{\pi}_{s,k-1}\; p_s(\mathbf{z}_k|\mathbf{z}_{k-1})}{\sum_{s=1}^{S} \boldsymbol{\pi}_{s,k-1}\; p_s(\mathbf{z}_k|\mathbf{z}_{k-1})}$.
  The posterior mean: $p(\mathbf{x}_k|\mathbf{z}_k) = \sum_{s=1}^{S} \hat{w}_{s,k}^{(i)} \delta_{\hat{\mathbf{x}}^{(i)}}\; \boldsymbol{\pi}_{s,k}$.

  (6) Output: State estimates: $\hat{\mathbf{x}}_k$

  (7) Resampling Step
  Set the threshold sample size: $N_{thresh} = N_p/10$.
  Calculate the effective sample size:
  $N_{eff} = 1/\sum_{i=1}^{N_p} (\hat{w}_k^{(i)})^2$ where $\hat{w}_k^{(i)} = \sum_{s=1}^{S} \hat{w}_{s,k}^{(i)}\; \boldsymbol{\pi}_{s,k}$.
  if $N_{eff} < N_{thresh}$
    The residual resampling method is applied.
  end if
end for

5.6 Performance Validation

Table 5.1: Simulation Parameters for Tracking with Non-Gaussian Shadowing Noise

Comments	**Parameters**
Minimum speed of a mobile user	$v_{min} = 3.53$ ms^{-1}
Maximum speed of a mobile user	$v_{max} = 17.67$ ms^{-1}
Number of Particles	$N_p = 500$
Standard deviation of the process noise	$\sigma_w = 0.25$ ms^{-2}
Standard deviation of the shadowing noise	$\sigma_v = [1 - 4]$ dB
Path loss exponent	$\beta = 3$
Decorrelation distance	$D_c = 10$ m
Degree of freedom	$v = 3$
Forgetting factor	$\varphi = 0.8$
Initial probabilities of noise models	$\boldsymbol{\pi}_{s,0} = 0.5$
Number of snapshot/observed measurements	$P = 3$
Window time	$t_{window} = 2$

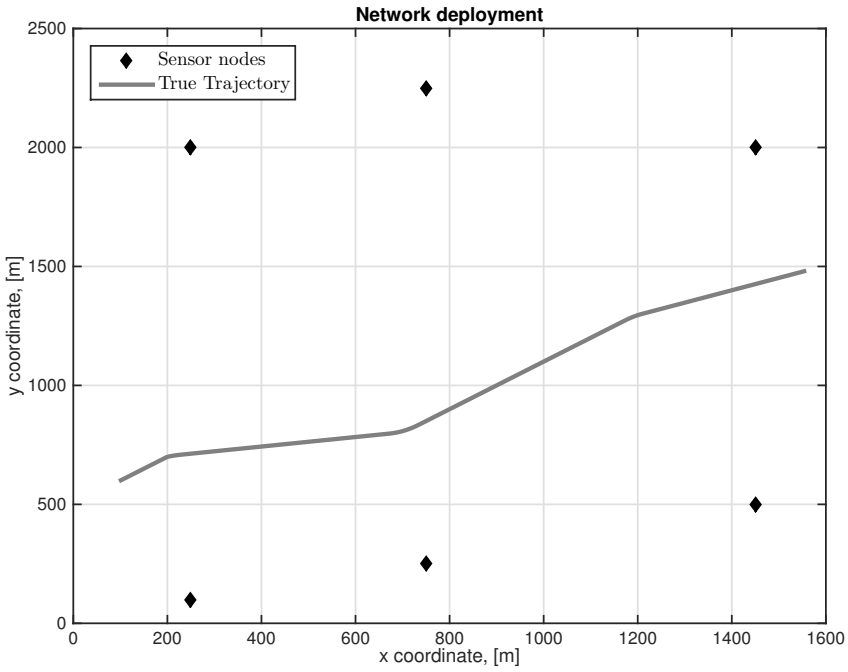

Figure 5.1: Network deployment.

Given the initial state $\mathbf{x}_0 = [100,\ 2.5,\ 0.0,\ 600,\ 2.5,\ 0.0]^T$, the mobile user trajectory in Figure 5.1 is generated according to the motion model in (4.2) where $T = 0.5$ s, $\alpha = 0.6$, and $k = \{1, \ldots, 150\}$. The control input $\mathbf{u}_k$ in (4.2) is modeled as a Markov process which takes values between the following modes $\mathbb{M} = \{[0.0\ 0.0]^T, [2.0\ 0.0]^T, [0.0\ 2.0]^T, [0.0\ -$

$2.0]^T, [-2.0\ 0.0]^T\}$ in units of ms^{-2}, with the transition probability matrix $\prod$ is defined by: $\left[\prod\right]_{ii} = 0.7$, for $i = 1, \ldots, 5$ (diagonal elements) and $\left[\prod\right]_{ij} = 0.075$, for $i, j = 1, \ldots, 5$ (off-diagonal elements). Meanwhile, the initial mode probabilities are $\mu_{1,0} = 0.8$ and $\mu_{i,0} = 0.05$ for $i = 2, \ldots, 5$ such that the sum of elements in each row of the matrix is equal to one. Sensor measurements are generated according to measurement model in (4.7) where the shadowing noise is spatio-temporally correlated and corrupted by outliers. When the outliers presence, the shadowing noise is drawn from a Student's t distribution $\mathbf{v}_k \sim \mathrm{St}(0, \hat{\mathbf{R}}_k, v)$ and the shadowing noise is set to be heavily tailed with $v = 3$. Otherwise, if no outliers, the shadowing noise is drawn from a zero-mean Gaussian distribution $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$. It is assumed at time instant k, there is a P number of snapshots of sensor measurements are collected in the networks. This assumption is made because the shadowing noise covariance matrix is estimated using the sensor measurements. If only a single observation of sensor measurement is available, the non-parametric shrinkage estimator does not have enough data to compute the shadowing noise covariance matrix. At least, the minimum of a $P = 3$ number of snapshots or observations are needed by the non-parametric shrinkage estimator to estimate the shadowing noise covariance matrix accurately. To generate a multiple number of snapshots (oversampling) of the received sensor measurements, a concept of time window is considered. By using this concept, only the user state at the beginning and end of the window time is considered for the shadowing noise covariance matrix and mobile user state estimation. The size of the window time affects the strength of the temporal correlation between the received measurements. Moreover, the estimated speed and acceleration are also imposed so that $\hat{v} \in [0,\ 20]$ ms^{-1} and $\hat{a} \in [-2,\ 2]$ ms^{-2}, respectively. The particles $\mathbf{x}_0^{(i)}$ is drawn from an initial distribution $\mathbf{x}_0^{(i)} \sim \mathcal{N}(\mathbf{x}_0, \mathbf{P})$ where $\mathbf{P} = \mathrm{diag}\{x_0^2,\ \dot{x}_0^2,\ \ddot{x}_0^2,\ y_0^2,\ \dot{y}_0^2,\ \ddot{y}_0^2\}$. Table 5.1 shows the other parameter settings for the designed simulation.

Case 1 *Tracking with the true shadowing noise covariance matrix.* First, considered the case where the true value of the shadowing noise covariance matrix $\mathbf{R}_k$ is available during the state estimation. The shadowing noise component of the measurement is drawn from a zero-mean multivariate Gaussian distribution at all time except at time steps $k = \{50, \ldots, 100\}$, where the shadowing noise component is drawn from a multivariate Student's t distributions to indicate that outlier is present in the sensor readings. Three different filtering methods were investigated for performance comparison, namely the PF, TPF, and RPF, respectively. These filtering methods are detailed as follows:

- PF - A Gaussian model based Particle filter.
- TPF - A Student's t model based Particle filter $(v = 3)$.
- RPF - A robust Particle filter.

The PF details are given in [41] but instead of using a fixed shadowing noise covariance

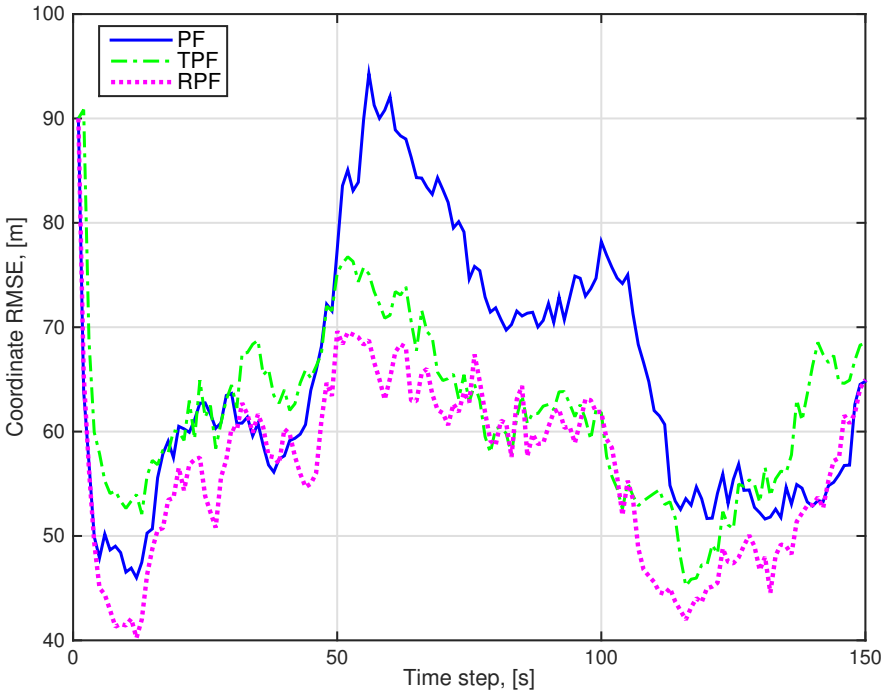

Figure 5.2: Coordinate RMSE comparison of the PF, TPF, and RPF as a function of time.

matrix at every time instant k, the PF, in this case, is operated with the true shadowing noise covariance matrix. The PF was a well-established method for state estimation and thus was selected for comparison purposes. The TPF is basically a PF which only employs the Student's t model instead of the Gaussian model. On the contrary, the RPF is a combination of the multiple models particle filter and the Bayesian model averaging methods which employed both the Gaussian and Student's t models. The RPF is assumed to operate with the true shadowing noise covariance matrix as with the PF and TPF. Here, all filtering methods are aware of the true shadowing noise covariance matrix of the measurement noise. However, they did not know when the measurement would be corrupted by the outliers. The PF operates by assuming no outliers presence in the measurements at all time and conversely, the TPF assumed that the outliers always present in the measurements at all time. On the other hand, the RPF operates by taking consideration that the outliers may be present at any time instants. The performance of the different filters is summarized in Figure 5.2. It is shown that the RPF exhibits the lowest coordinates RMSE better than the PF and TPF at most of the time. At time instants, $k = \{1, \ldots, 49\}$ and $k = \{101, \ldots, 150\}$, where the sensor measurements were not corrupted by the outliers, the PF displayed a more accurate state estimates when compared to the TPF. Conversely, when the sensor measurements were corrupted by the outliers that are at time instants $k = \{50, \ldots, 100\}$, the TPF showed a better coordinate estimated than the PF. The coordinate RMSE is rapidly increased in the event of measurement with outliers and when no outliers present, the coordinate RMSE is decreased. This is evident that outliers have dis-

torted the accuracy of the state estimation by the filtering methods. The filtering method with a Gaussian model operates the best when there is no outliers presence in the measurements. However, when there is an outlier, the filtering method with a Student's t model is shown to perform well than with the Gaussian model. Instead of choosing between the two noise models, the RPF used both models to operate and gives a better state estimate when compared to the PF and TPF.

Case 2 *Tracking with the estimated shadowing noise covariance matrix.* Next, considered the case where the true value of the shadowing noise covariance matrix of the measurement noise, $\mathbf{R}_k$ presented in Section 5.5.1 is not available. Hence, the shadowing noise covariance matrix has to be estimated using the sensor measurements. Let the shadowing noise covariance matrix is estimated by using the non-parametric shrinkage estimator $\hat{\mathbf{S}}_k = \mathbf{R}_k$ given in Section 5.5.2. The setting for the simulation outliers was the same as in Case 1. Here, an extension to the filtering methods presented in Case 1 is proposed and investigated, namely, the SPF, TSPF, and RSPF, respectively. These filtering methods are detailed as follows:

- SPF - A Gaussian model based shrinkage Particle filter.
- TSPF - A Student's t model based shrinkage Particle filter $(v = 3)$.
- RSPF - A robust shrinkage Particle filter.

The SPF details are given in Chapter 4, Section 4.5.4 of the thesis but instead of using the diagonal, unit variance target matrix and the constant correlation shrinkage target matrix, the SPF, in this case, is operated with the diagonal, unequal variance target matrix [124]. The TSPF is basically a SPF which only employs the Student's t model instead of the Gaussian model. On the contrary, the RSPF is developed using the multiple models particle filter, Bayesian model averaging, and non-parametric shrinkage estimator which is details in Algorithm 5.1. The RSPF jointly estimate the shadowing noise covariance matrix of the measurements that are corrupted by the outliers as well as the state of a mobile user. In this case, all filtering methods had to estimate the shadowing noise covariance matrix of the measurements in order to exploit the correlation. They were also set to be blind to the occurrence of outliers in the measurements. The SPF operates with the assumption of no outliers present in the measurement thus it only employed a Gaussian-based model. Conversely, the TSPF operates with the assumption of outliers are always present in the measurements thus only employed the Student's t-based model. On the other hand, the RSPF operates by taking into consideration that outliers can exist at any time instants, therefore combined both the Gaussian and Student's t models via a Bayesian model averaging approach. Since the true shadowing noise covariance matrix is not available, all the filtering methods used the non-parametric shrinkage estimator to estimate the shadowing noise covariance matrix in its operation. Figure 5.3 shows the coordinate RMSE comparison of the different filters. It is shown that the RSPF displayed the most accurate

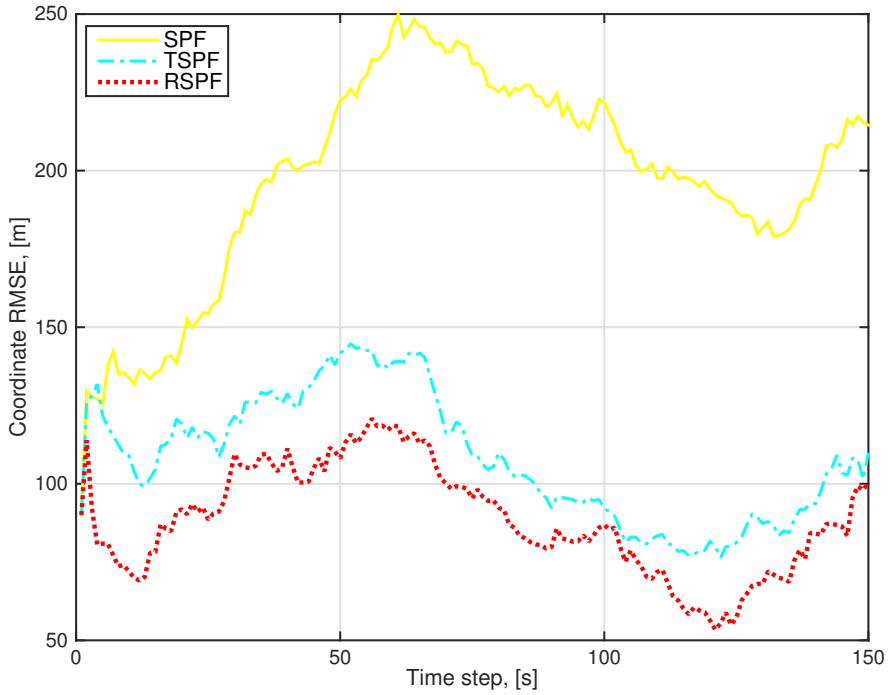

Figure 5.3: Coordinate RMSE comparison of the SPF, TSPF, and RSPF as a function of time.

coordinate estimates by exhibits the lowest coordinate RMSE at all time, followed by the TSPF and finally the SPF. Since the true shadowing noise covariance matrix of the measurement is not available and has to be estimated, the SPF, TSPF, and RSPF are exhibited higher coordinate RMSE when compared to the PF, TPF, and RPF as displayed in Case 1.

To understand the effect of outliers in the sensor measurements, Figure 5.4 is plotted by combining the results obtained in Figure 5.2 and Figure 5.3, respectively. The presence of outliers in the measurements affected the tracking accuracy in two ways; (1) false estimation of the shadowing noise covariance matrix, and (2) false estimation of the mobile state. When the filtering methods (PF, TPF, and RPF) operated with the true shadowing noise covariance matrix of the measurement, the outliers only affected the state estimation of the mobile user. The PF, TPF, and RPF are still able to exploit the correlated measurement using the statistical information provided by the true measurement noise covariance matrix. Conversely, when the true shadowing noise covariance matrix of the measurement is not available and has to be estimated in order to exploit the correlated measurements, the SPF, TSPF, and RSPF were developed. The SPF, TSPF, and RSPF have to estimate the mobile user state accurately using the corrupted measurements. The utilization of the non-parametric shrinkage estimator by the SPF, TSPF, and RSPF in their operations provided a way to estimate the shadowing noise covariance matrix that is more robust towards the effect of outliers in the measurements. By comparing the two groups of filtering method, the RPF and RSPF indicated to be more accurate compared to the others. For that reason, only RPF and RSPF estimators are further

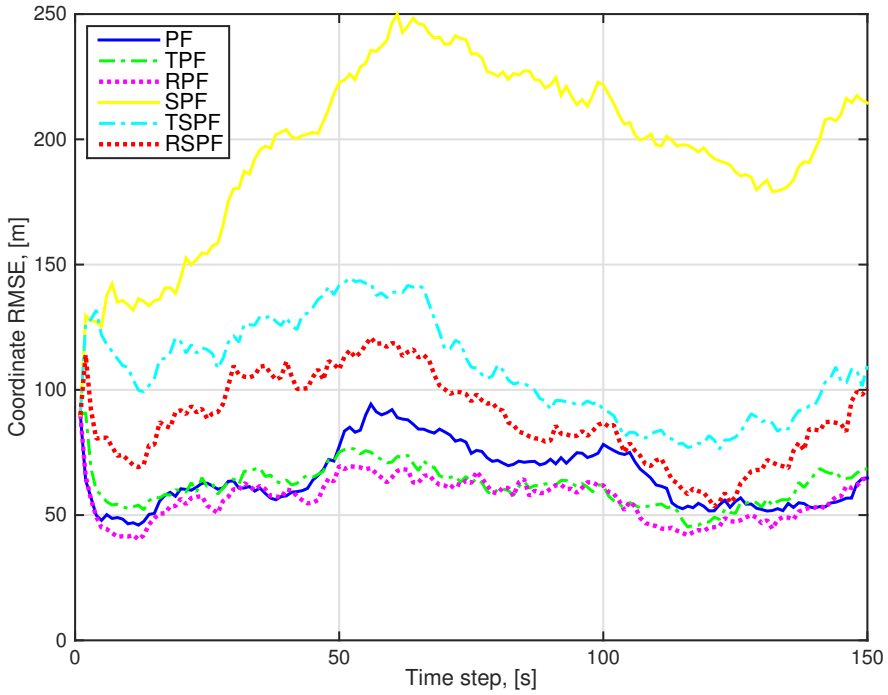

Figure 5.4: Coordinate RMSE comparison of the PF, TPF, RPF, SPF, TSPF, and RSPF as a function of time.

analysis. However, a significant difference of coordinate RMSE is displayed when comparing the RPF and the RSPF. In RPF, the outliers only distorted the state estimation of the mobile user. Meanwhile, in RSPF the outliers have distorted both the state estimation and the shadowing noise covariance matrix estimation. In practice, it is impossible to know the true shadowing noise covariance matrix of the measurements except through a designed simulation setup. Hence, the assumption of using the true representation of the shadowing noise covariance matrix is too optimistic. The RSPF is more practical for implementation when compared to the RPF and for that reason, the analysis of RPF is excluded and further analysis of the RSPF is carried out.

Effects of temporal correlation on the RSPF tracking accuracy. Figure 5.5 displays the effects of temporal correlation with different window sizes of the RSPF. In RSPF with $t_{window} = 4$, only four measurements are held per second while in RSPF with $t_{window} = 10$, there are ten measurements are held per second. The setting for the simulation outliers was the same as in Case 1. The RSPF $(t_{window} = 10)$ showed a better coordinate estimation than the RSPF $(t_{window} = 4)$ because of more measurement history was retained for the shadowing noise covariance matrix estimation process. The strength of the temporal correlation increases when more measurement history was retained thus a better statistical information can be extracted from the measurements by the RSPF. However, when operated with a high number of snapshots/observations, the dimensionality of the estimated shadowing noise covariance

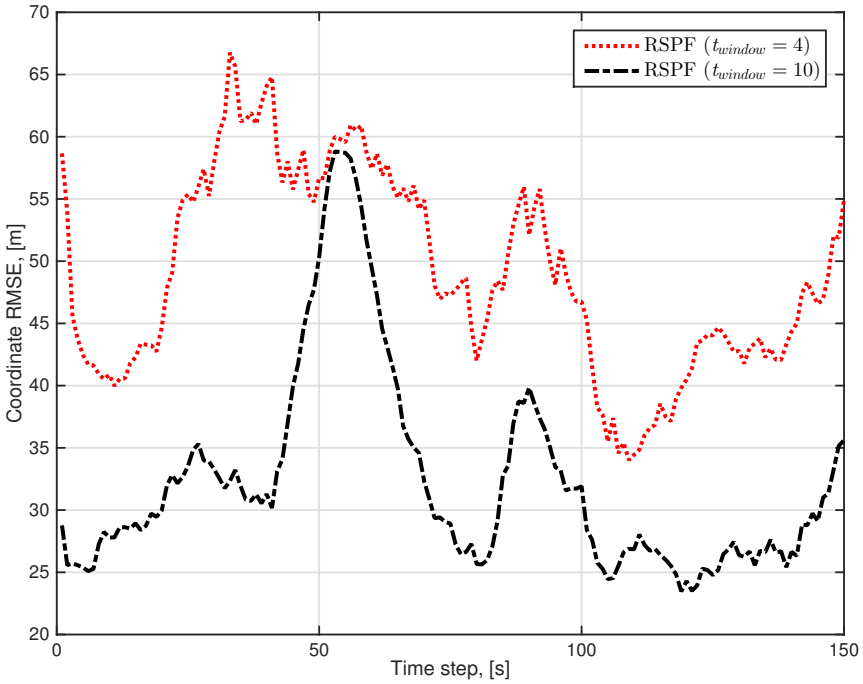

Figure 5.5: Coordinate RMSE comparison of the RSPF with window sizes $t_{window} = 4$ and $t_{window} = 10$.

matrix grows exponentially and thus increases the execution time of the RSPF.

Effects of forgetting factor on the RSPF noise models posterior probabilities. Next, the effects of the forgetting factor on the RSPF noise models posterior probabilities are evaluated. The value of the forgetting factor is set to $\varphi = 0.9$ and $\varphi = 0.1$, respectively and the average posterior probability of each measurement noise model was recorded at every time instants. To understand the effects of the forgetting factor, the RPF noise models posterior probabilities also plotted and compared with the RSPF noise models posterior probabilities. Figures 5.6 and 5.7 show the average posterior probability of noise models by the RPF for $\varphi = 0.9$ and $\varphi = 0.1$, respectively and Figures 5.8 and 5.9 show the average posterior probability of noise models by the RSPF for $\varphi = 0.9$ and $\varphi = 0.1$, respectively. The setting for the simulation outliers was the same as in Case 1. Intuitively, when no outliers present in the measurements that is at time instants, $k = \{1, \ldots, 49\}$ and $k = \{101, \ldots, 150\}$, the Gaussian noise model will dominate the Student's t noise model and vice verse when outliers present in the measurements that is at time instants $k = \{50, \ldots, 100\}$. This is true with the RPF of $\varphi = 0.1$ but when the forgetting factor is increases to $\varphi = 0.9$, the posterior probabilities of the Student's t noise model will then dominate the posterior probabilities of the Gaussian noise model at all time instants.

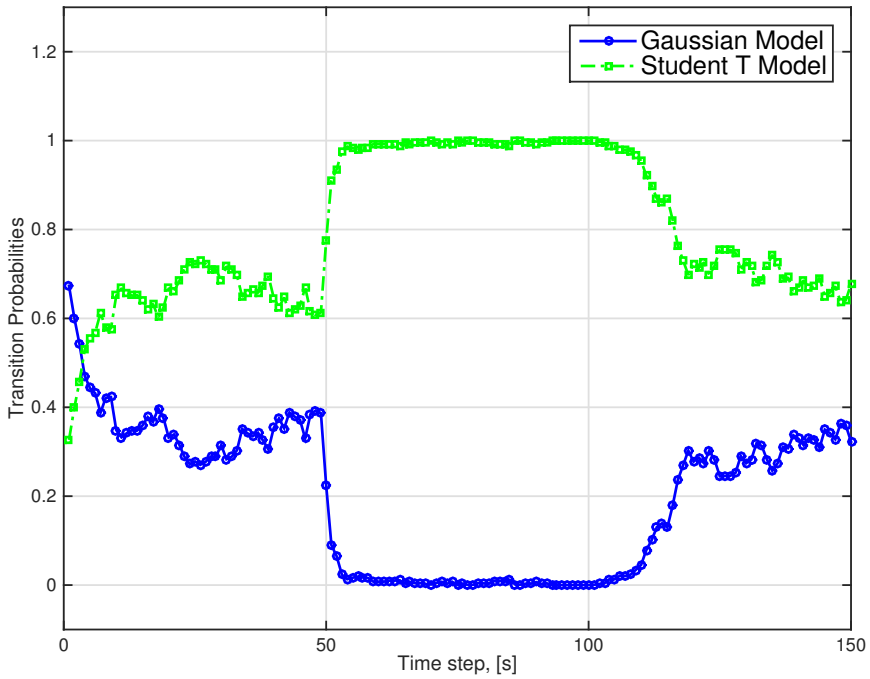

Figure 5.6: Average posterior probability of noise models by RPF for $\varphi = 0.9$.

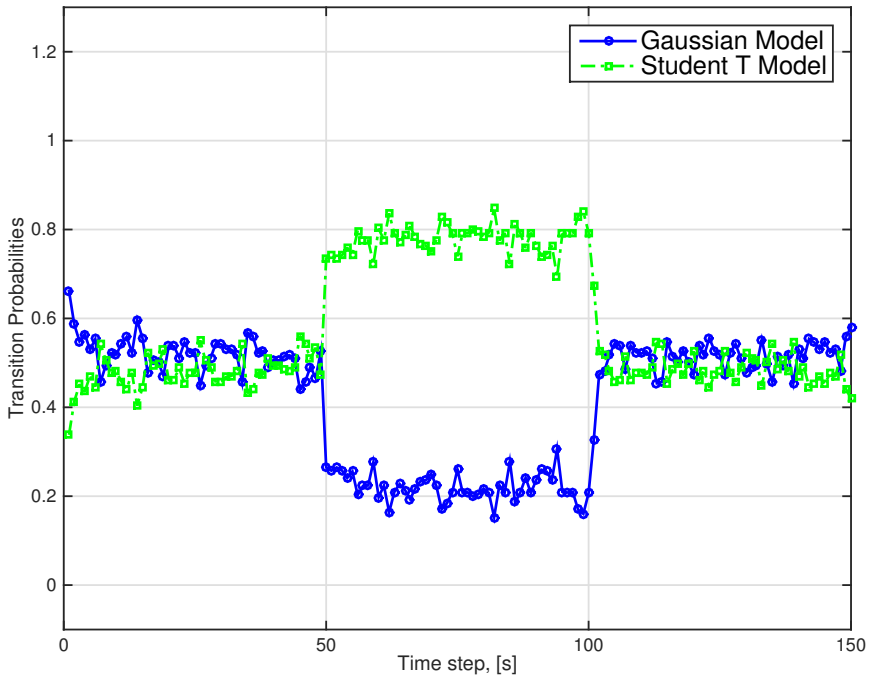

Figure 5.7: Average posterior probability of noise models by RPF for $\varphi = 0.1$.

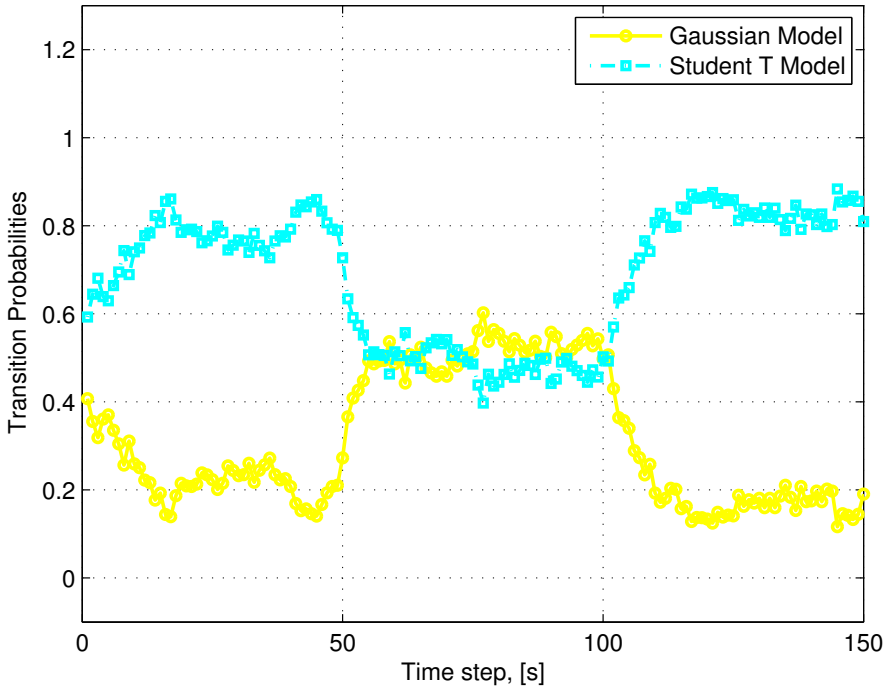

Figure 5.8: Average posterior probability of noise models by RSPF for $\varphi = 0.9$.

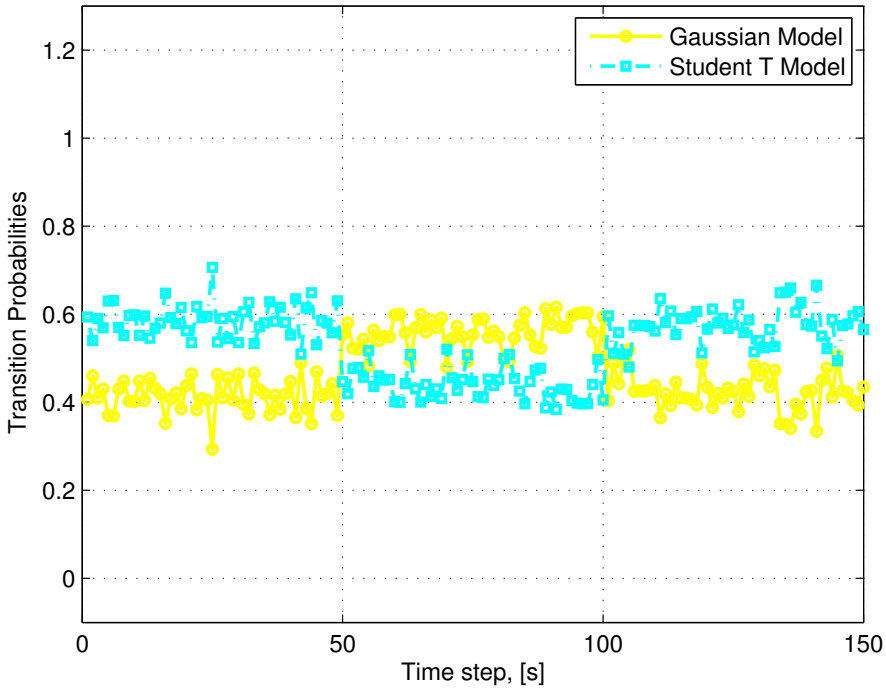

Figure 5.9: Average posterior probability of noise models by RSPF for $\varphi = 0.1$.

Contrary, in the RSPF posterior probability of noise models, when no outliers present in the measurements, the Student's t noise model will always dominates the Gaussian noise model as shown in Figures 5.8 and 5.9 at time instants $k = \{1, \ldots, 49\}$ and $k = \{101, \ldots, 150\}$. Comparing the two Figures at time instants $k = \{1, \ldots, 49\}$ and $k = \{101, \ldots, 150\}$, when the value of forgetting factor in the RSPF was decreased from $\varphi = 0.9$ to $\varphi = 0.1$, the gap of the transition probabilities between the two noise models reduced significantly. This indicated that when RSPF is operated with $\varphi = 0.9$, the transition probabilities between one noise model to the other is less affected by the distribution of the shadowing noise thus the filter can hold the previous information for much longer. However, when the RSPF is operated with $\varphi = 0.1$, the transition probabilities are highly affected by the distribution of the shadowing noise. Hence, the filter will update the noise distribution information frequently.

Furthermore, when there is outlier present in the measurements, the posterior probability of the Gaussian noise model and the Student's t noise model were equally distributed between them as shown in Figures 5.8 and 5.9 at time instants $k = \{50, \ldots, 100\}$. In RSPF, the Gaussian noise model is given higher transition probability when compared to the Student's t noise model. In contrast, the Gaussian noise model is given less transition probability than the Student's t noise model by the RPF as shown in Figures 5.6 and 5.7 at time instants $k = \{50, \ldots, 100\}$. The posterior probability/transition probability of noise models depends on the shadowing noise covariance matrix of the measurements. When estimated shadowing noise covariance matrix was deployed instead of the true shadowing noise covariance matrix, the transition probabilities are adjusted according to the statistical information of the covariance matrix. In the events of outliers present in the measurements, the interpretation of the statistical information of the covariance matrix can be misleading. Thus, choosing which noise models to operate at a given time instants are heavily depended on the knowledge of outliers in the measurements.

Effects of correlation in the shadowing noise. Finally, the effects of correlation in the shadowing noise of the measurements are evaluated. Let the true value of the shadowing noise covariance matrix is available during the state estimation. Then, consider the setting for the simulation outliers was the same as in Case 1. Figure 5.10 shows the coordinate RMSE of the PF, TPF, and RPF with correlated shadowing noise and Figure 5.11 with uncorrelated shadowing noise. When there is a correlation between measurements, all the filtering methods are converging more slowly between one another. This is indicated by comparing the coordinate RMSE of the filters in Figures 5.10 and 5.11 at time instants $t > 100$ which is during the transition period from outliers present in the measurements to no outliers present in the measurements. On the other hand, when there is no correlation between measurements, the filtering methods are converging more quickly between one another.

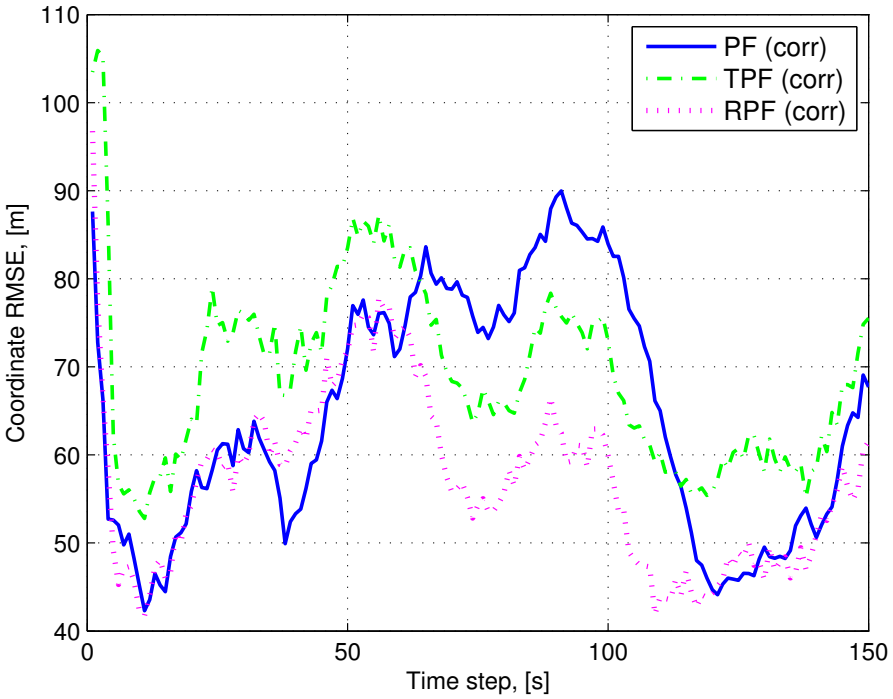

Figure 5.10: Coordinate RMSE comparison of the filters with correlated shadowing noise.

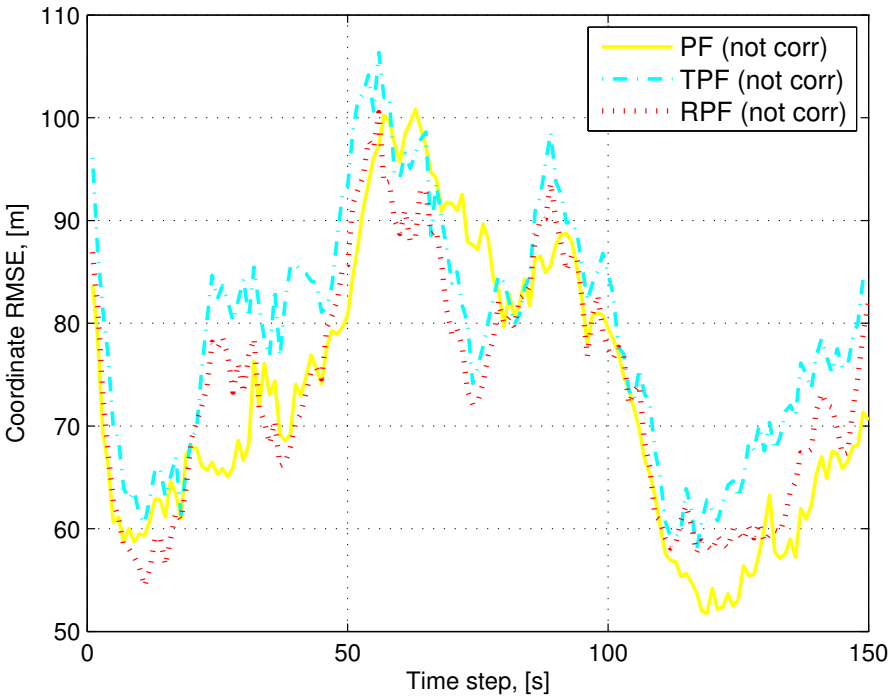

Figure 5.11: Coordinate RMSE comparison of the filters with uncorrelated shadowing noise.

5.7 Summary

In this chapter, the challenging problem of tracking a mobile user with correlated RSS measurements that are corrupted by outliers was considered. A heavy-tailed Student's t noise model is presented to simulate the RSS measurements with outliers. To exploit the correlated measurement, a non-parametric shrinkage estimator is derived to estimate the shadowing noise covariance matrix of the corrupted measurements. A robust shrinkage-based particle filter (RSPF) was proposed to jointly estimate the shadowing noise covariance matrix and the state of the mobile user. The RSPF is developed by combining the multiple models particle filter, Bayesian model averaging, and the non-parametric shrinkage estimator. The performance of the RSPF is validated using the simulated data. The simulation results showed that the RSPF is robust against the presence of outliers in the RSS measurements and capable to track accurately the mobile user in the wireless networks.

However, the validation of the RSPF is not extended using the experimental data. Outliers in the sensor readings can come from sudden environmental disturbances, temporary sensor failures or from the intrinsic noise of the sensor device itself. In practice, it is impossible to design an experimental setup that can guarantee outliers present in the sensor readings at a given time instant. The sensor readings are affected by multipath and shadow fading effects caused by obstacles along the signal propagation path which are usually assumed to follow a Gaussian distribution instead of a Student's t distribution. Unless there is knowledge of the outlier in the measurements, the assumption of using the Student's t distribution will not be truly justified. However in the future, it is applicable for the RSPF to be tested using the experimental data if the data are carefully designed to reflect the effects of the outliers present in the sensor measurements.

Chapter 6

Conclusion

6.1 Thesis Contributions

The aim of this thesis was to developed mobile tracking approach in WSNs with the correlated and sparse measurements. The key contributions of this thesis are details as follows:

In Chapter 4, a ShPF for mobile tracking in WSNs was presented. The shadowing component of the RSS measurements is assumed to be spatio-temporally correlated. Thus, a correlation model is designed to generate correlated RSS in the simulated data. In practice, the correlation present in the measurement is unknown. Modeling realistic correlated measurements using correlation models may not be always feasible. In particular, when both spatial and temporal correlation is incorporate in the correlation model, they may not be invertible and well-conditioned. Next, a shrinkage estimator is proposed to estimate the shadowing noise covariance matrix of the RSS measurements. By exploiting the statistical information on the shadowing noise covariance matrix, the correlation present in the measurement is successfully extracted. The estimated shadowing noise covariance matrix is well-conditioned and invertible even when the number of snapshots/observations of the RSS measurement is comparable or less than the dimensionality of the signal model. After that, the estimated shadowing noise covariance matrix is applied to the particle filter (PF). This information enhanced the statistical inference of the PF likelihood function. Thus, high tracking accuracy is achieved. Results form both simulated and experimental data showed that the ShPF able to exploit the correlated and sparse RSS measurement with high accuracy tracking.

In Chapter 5, a robust shrinkage-based particle filter (RSPF) for the problem of outliers in the RSS measurements is presented. The shadowing component of the RSS measurement is assumed to be spatio-temporally correlated. Hence, a correlation model is designed to generate the correlated RSS measurements. When the outliers presence, the shadowing noise follows

the Student's t distributions. Otherwise, if no outliers, the shadowing noise is drawn from the Gaussian distributions. A Student's t noise model is proposed to indicate that the outliers are present in the sensor readings. The magnitude of correlation in the RSS measurements is unintentionally changed due to the presence of outliers. Since exploiting correlation in the measurement increases the tracking accuracy, a non-parametric shrinkage estimator is derived to robustly estimate the shadowing noise covariance matrix of the corrupted measurements. The non-parametric shrinkage estimator is combined with the multiple models particle filter to jointly estimate the shadowing noise covariance matrix and the state of the mobile user. To address the state estimation error due to outliers, the Bayesian model averaging approach is adopted. The approach combined both the Gaussian and heavier-tailed Student's t noise models and dynamically adjusted the effects of each model in its operation. Results from the simulated data showed that the RSPF is robust against the presence of outliers in the measurements and capable to track accurately the mobile user using the corrupted RSS measurements.

6.2 Direction for Future Work

Mobile tracking with correlated and sparse measurements is still a challenging task to be explored. Some directions for future works based on the findings in this thesis are as follows:

Exploring new shadowing noise correlation model.

In Chapter 4 and Chapter 5, the correlation models were developed to generate spatio-temporal correlation in the shadowing noise of the RSS measurements. One disadvantage of using the current shadowing correlation models is that they are designed based on the assumption that the shadowing noise follows a lognormal distribution. The lognormal distribution is a probability distribution of random variables whose logarithm has a Gaussian distribution. For the random variables whose logarithm has a non-Gaussian distribution, the random variables are said to follow a non-lognormal distribution. Future research could seek to design new correlation models in a case of non-lognormal distribution of the shadowing noise.

Improving correlation extraction from the measurements.

In Chapter 4, the accuracy of mobile tracking is shown to improve by exploiting the statistical information on the shadowing noise covariance matrix. The shrinkage-based particle filter used the shrinkage estimator to estimate the shadowing noise covariance matrix in its operation. The shrinkage estimator is developed by combining the sample covariance matrix with the diagonal unit variance shrinkage target matrix or the constant correlation shrinkage target matrix. Each target matrices are different in term of its structure as explained in [124]. The

shrinkage estimator operates by shrinking the value of the coefficients of the sample covariance matrix to the value supplied by the target matrix. Future work could focus to design new and flexible structure of target matrix using a best-suited representation of the correlation problems in the wireless network settings.

Non-parametric shrinkage estimator.

In Chapter 5, the non-parametric shrinkage estimator is derived and proposed to robustly estimate the shadowing noise covariance matrix of the corrupted RSS measurements by the outliers. The optimal shrinkage intensity (weight) of the non-parametric shrinkage estimator is derived with the assumption of using a diagonal target matrix structure (target D in [124]) in its operation. Future work could consider deriving the weight of the non-parametric shrinkage estimators by using a much more flexible structure of the target matrix.

Extending to multiple mobile users tracking.

The proposed tracking algorithms that are presented in Chapter 4 and Chapter 5 of the thesis are applied to the problem of single mobile user tracking in wireless sensor networks. Future work could consider solutions for tracking multiple mobile users using the correlated and sparse wireless measurements.

Appendix A: The Shrinkage Derivation for Tracking with Correlated Sparse Measurements

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

$$\begin{aligned} R(\mathbf{C}, \hat{\mathbf{S}}) &= \min_{\lambda} \ \mathrm{E}[\|\hat{\mathbf{S}} - \mathbf{C}\|^2], \\ &= \min_{\lambda} \ \mathrm{E}[\|\hat{\mathbf{S}} - \mathrm{E}[\hat{\mathbf{S}}]\|^2] + \|\mathrm{E}[\hat{\mathbf{S}}] - \mathbf{C}\|^2, \end{aligned} \tag{6.1}$$

which is composed of a variance and squared bias. If $\hat{\mathbf{S}} = \lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}}$, this implies that

$$\begin{aligned} R(\lambda) &= \sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{S}}_p) + [\mathrm{E}(\hat{\mathbf{S}}_p) - \mathbf{C}_p]^2, \\ &= \sum_{p=1}^{P} \mathrm{Var}(\lambda\mathbf{T}_p + (1-\lambda)\hat{\mathbf{C}}_p) + [\mathrm{E}(\lambda\mathbf{T}_p + (1-\lambda)\hat{\mathbf{C}}_p) - \mathbf{C}_p]^2, \\ &= \sum_{p=1}^{P} \lambda^2\mathrm{Var}(\mathbf{T}_p) + (1-\lambda)^2\mathrm{Var}(\hat{\mathbf{C}}_p) + 2\lambda(1-\lambda)\mathrm{Cov}(\hat{\mathbf{C}}_p, \mathbf{T}_p) + [\lambda\mathrm{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p) + \mathrm{Bias}(\hat{\mathbf{C}})]^2. \end{aligned} \tag{6.2}$$

Take derivative with respect to λ and set equal to zero yields

$$R'(\lambda) = \sum_{p=1}^{P} 2\lambda \mathrm{Var}(\mathbf{T}_p) + 2(1-\lambda)\mathrm{Var}(\hat{\mathbf{C}}_p) + 2(1-2\lambda)\mathrm{Cov}(\hat{\mathbf{C}}_p, \mathbf{T}_p) \\ + 2[\mathrm{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p)][\lambda \mathrm{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p) + \mathrm{Bias}(\hat{\mathbf{C}})], \tag{6.3}$$

which leads to the optimal shrinkage intensity given by

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{C}}_p) - \mathrm{Cov}(\mathbf{T}_p, \hat{\mathbf{C}}_p) - \mathrm{Bias}(\hat{\mathbf{C}}_p)\mathrm{E}(\mathbf{T}_p - \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathrm{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}. \tag{6.4}$$

If $\hat{\mathbf{C}}$ is unbiased estimator of $\mathbf{C}$, then $\hat{\lambda}$ becomes

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \mathrm{Var}(\hat{\mathbf{C}}_p) - \mathrm{Cov}(\mathbf{T}_p, \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathrm{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}, \tag{6.5}$$

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

$$\hat{\lambda} = \frac{\sum_{p=1}^{P} \widehat{\mathrm{Var}}(\hat{\mathbf{C}}_p) - \widehat{\mathrm{Cov}}(\mathbf{T}_p, \hat{\mathbf{C}}_p)}{\sum_{p=1}^{P} \mathrm{E}[(\mathbf{T}_p - \hat{\mathbf{C}}_p)^2]}. \tag{6.6}$$

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

$$v_{ijp} = (z_{ip} - \bar{z}_i)(z_{jp} - \bar{z}_j), \tag{6.7}$$

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

$$\left[\hat{\mathbf{C}}\right]_{ij} = \frac{1}{P-1} \sum_{p=1}^{P} (z_{ip} - \bar{z})(z_{jp} - \bar{z}). \tag{6.8}$$

The unbiased variance of individual elements of $\hat{\mathbf{C}}$ is given by [123]

$$\widehat{\mathrm{Var}}\Big[\hat{\mathbf{C}}\Big]_{ij} = \frac{P^2}{(P-1)^2}\widehat{\mathrm{Var}}(\bar{v}_{ij}) \tag{6.9}$$

$$= \frac{P^2}{(P-1)^2}\left[\frac{1}{P}\widehat{\mathrm{Var}}(v_{ij})\right] \tag{6.10}$$

$$= \frac{P^2}{(P-1)^2}\left[\frac{1}{P}\left[\frac{1}{P-1}\sum_{p=1}^{P}(v_{ijp}-\bar{v}_{ij})^2\right]\right] \tag{6.11}$$

$$= \frac{P}{(P-1)^3}\sum_{p=1}^{P}(v_{ijp}-\bar{v}_{ij})^2. \tag{6.12}$$

Next, let v_{kl} be another random variable where the sample mean of the variable is $\bar{v}_{kl}$. The covariance elements are obtained as [123]

$$\widehat{\mathrm{Cov}}\Big(\Big[\hat{\mathbf{C}}\Big]_{ij},\Big[\hat{\mathbf{C}}\Big]_{kl}\Big) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}(v_{ijp}-\bar{v}_{ij})(v_{klp}-\bar{v}_{kl}). \tag{6.13}$$

To determine $\hat{\lambda}$ in (6.6) requires an expression for $\widehat{\mathrm{Cov}}(\mathbf{T}_p, \hat{\mathbf{C}}_p)$. Now, consider a target matrix of $\Big[\mathbf{T}\Big]_{ij} = \bar{\rho}\sqrt{\Big[\hat{\mathbf{C}}\Big]_{ii}\Big[\hat{\mathbf{C}}\Big]_{jj}}$, for $i \neq j$ and let $\Big[\bar{\mathbf{C}}\Big]_{ii}$, $\Big[\bar{\mathbf{C}}\Big]_{jj}$ and $\Big[\bar{\mathbf{C}}\Big]_{ij}$ be the point estimates of $\Big[\mathbf{C}\Big]_{ii}$, $\Big[\mathbf{C}\Big]_{jj}$ and $\Big[\mathbf{C}\Big]_{ij}$, respectively. The point estimate here is refer to the elements of the sample covariance which is also considered as a random variable. Then, $\Big[\mathbf{T}\Big]_{ij}$ is expands via Taylor series such that

$$\Big[\mathbf{T}\Big]_{ij} = \bar{\rho}\sqrt{\Big[\bar{\mathbf{C}}\Big]_{ii}\Big[\bar{\mathbf{C}}\Big]_{jj}} + \frac{\bar{\rho}}{2}\sqrt{\frac{\Big[\bar{\mathbf{C}}\Big]_{jj}}{\Big[\bar{\mathbf{C}}\Big]_{ii}}}\Big(\Big[\mathbf{C}\Big]_{ii}-\Big[\bar{\mathbf{C}}\Big]_{ii}\Big) + \frac{\bar{\rho}}{2}\sqrt{\frac{\Big[\bar{\mathbf{C}}\Big]_{ii}}{\Big[\bar{\mathbf{C}}\Big]_{jj}}}\Big(\Big[\mathbf{C}\Big]_{jj}-\Big[\bar{\mathbf{C}}\Big]_{jj}\Big), \tag{6.14}$$

where $\bar{\rho}$ is the average correlation of the off-diagonal elements in the sample covariance and is obtained using

$$\bar{\rho} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{\Big[\hat{\mathbf{C}}\Big]_{ij}}{\sqrt{\Big[\hat{\mathbf{C}}\Big]_{ii}\Big[\hat{\mathbf{C}}\Big]_{jj}}}, \tag{6.15}$$

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

$$\widehat{\mathrm{Cov}}\Big(\left[\mathbf{T}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\Big) = \mathbb{E}\Big[\Big(\left[\mathbf{T}\right]_{ij} - \mathbb{E}\Big[\left[\mathbf{T}\right]_{ij}\Big]\Big)\Big(\left[\mathbf{C}\right]_{ij} - \mathbb{E}\Big[\left[\mathbf{C}\right]_{ij}\Big]\Big)\Big], \tag{6.16}$$

and using (6.14) and (6.16) yields

$$\widehat{\mathrm{Cov}}\Big(\left[\mathbf{T}\right]_{ij}, \left[\hat{\mathbf{C}}\right]_{ij}\Big) = \frac{\bar{\rho}}{2}\Bigg\{\sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{jj}}{\left[\bar{\mathbf{C}}\right]_{ii}}}\widehat{\mathrm{Cov}}\Big(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\Big) + \sqrt{\frac{\left[\bar{\mathbf{C}}\right]_{ii}}{\left[\bar{\mathbf{C}}\right]_{jj}}}\widehat{\mathrm{Cov}}\Big(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\Big)\Bigg\}. \tag{6.17}$$

Finally, using (6.7) and (6.13) the covariance elements are expressed as

$$\widehat{\mathrm{Cov}}\Big(\left[\hat{\mathbf{C}}\right]_{ii}, \left[\hat{\mathbf{C}}\right]_{ij}\Big) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\Big[(z_{ip}-\bar{z}_i)^2 - \bar{v}_{ii}\Big]\Big[(z_{ip}-\bar{z}_i)(z_{jp}-\bar{z}_j) - \bar{v}_{ij}\Big], \tag{6.18}$$

and similarly

$$\widehat{\mathrm{Cov}}\Big(\left[\hat{\mathbf{C}}\right]_{jj}, \left[\hat{\mathbf{C}}\right]_{ij}\Big) = \frac{P}{(P-1)^3}\sum_{p=1}^{P}\Big[(z_{jp}-\bar{z}_j)^2 - \bar{v}_{jj}\Big]\Big[(z_{ip}-\bar{z}_i)(z_{jp}-\bar{z}_j) - \bar{v}_{ij}\Big]. \tag{6.19}$$

which completes the derivation.

Target matrix should reflect the true characteristics of the quantity that being estimated. In general, target matrix decreases the mean square errors of the sample covariance matrix. Six types of target matrix structures are summarized in Table 6.1 along with its optimal weight given by [124].

Table 6.1: Six commonly used shrinkage targets for the covariance matrix and its optimal shrinkage intensity.

Target A: "diagonal, unit variance" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} 1, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j \end{cases}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij}) + \sum_i \widehat{\text{Var}}(\hat{\mathbf{C}}_{ii})}{\sum_{i \neq j} \hat{\mathbf{C}}_{ij}^2 + \sum_i (\hat{\mathbf{C}}_{ii} - 1)^2}$	**Target B:** "diagonal, common variance" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} v = \text{avg}(\hat{\mathbf{C}}_{ii}), & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j \end{cases}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij}) + \sum_i \widehat{\text{Var}}(\hat{\mathbf{C}}_{ii})}{\sum_{i \neq j} \hat{\mathbf{C}}_{ij}^2 + \sum_i (\hat{\mathbf{C}}_{ii} - v)^2}$
Target C: "common (co)variance" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} v = \text{avg}(\hat{\mathbf{C}}_{ii}), & \text{if} \quad i = j \\ c = \text{avg}(\hat{\mathbf{C}}_{ij}), & \text{if} \quad i \neq j \end{cases}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij}) + \sum_i \widehat{\text{Var}}(\hat{\mathbf{C}}_{ii})}{\sum_{i \neq j} (\hat{\mathbf{C}}_{ij} - c)^2 + \sum_i (\hat{\mathbf{C}}_{ii} - v)^2}$	**Target D:** "diagonal, unequal variance" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} \hat{\mathbf{C}}_{ii}, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j \end{cases}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij})}{\sum_{i \neq j} \hat{\mathbf{C}}_{ij}^2}$
Target E: "perfect positive correlation" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} \hat{\mathbf{C}}_{ii}, & \text{if} \quad i = j \\ \sqrt{\hat{\mathbf{C}}_{ii} \hat{\mathbf{C}}_{jj}}, & \text{if} \quad i \neq j \end{cases}$ $f_{ij} = \frac{1}{2} \left\{ \sqrt{\frac{\hat{\mathbf{C}}_{jj}}{\hat{\mathbf{C}}_{ii}}} \widehat{\text{Cov}}(\hat{\mathbf{C}}_{ii}, \hat{\mathbf{C}}_{ij}) + \sqrt{\frac{\hat{\mathbf{C}}_{ii}}{\hat{\mathbf{C}}_{jj}}} \widehat{\text{Cov}}(\hat{\mathbf{C}}_{jj}, \hat{\mathbf{C}}_{ij}) \right\}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij}) - f_{ij}}{\sum_{i \neq j} (\hat{\mathbf{C}}_{ij} - \sqrt{\hat{\mathbf{C}}_{ii} \hat{\mathbf{C}}_{jj}})^2}$	**Target F:** "constant correlation" $\Big[\mathbf{T}\Big]_{ij} = \begin{cases} \hat{\mathbf{C}}_{ii}, & \text{if} \quad i = j \\ \bar{\rho} \sqrt{\hat{\mathbf{C}}_{ii} \hat{\mathbf{C}}_{jj}}, & \text{if} \quad i \neq j \end{cases}$ $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(\hat{\mathbf{C}}_{ij}) - \bar{\rho} f_{ij}}{\sum_{i \neq j} (\hat{\mathbf{C}}_{ij} - \bar{\rho} \sqrt{\hat{\mathbf{C}}_{ii} \hat{\mathbf{C}}_{jj}})^2}$

Appendix B: The Non-Parametric Shrinkage Derivation for Tracking with Non-Gaussian Shadowing Noise

Consider the problem of estimating the true covariance matrix $\mathbf{C}_0$ by a shrinkage estimator $\hat{\mathbf{S}} = \lambda\mathbf{T} + (1-\lambda)\hat{\mathbf{C}}$, where $\hat{\mathbf{C}}$ is the estimated sample covariance matrix, $\mathbf{T}$ is the target matrix, and $\hat{\lambda} \in [0,\ 1]$ is the shrinkage intensity (weight). The sample covariance estimates is given by

$$\hat{\mathbf{C}} = \frac{1}{P-1}\sum_{p=1}^{P}(\mathbf{z}_p - \bar{\mathbf{z}}_p)(\mathbf{z}_p - \bar{\mathbf{z}}_p)^T, \tag{6.20}$$

where $\mathbf{z}_p \in \mathbb{R}^n$ is the RSS measurements, $\bar{\mathbf{z}} = \frac{1}{P}\sum_{p=1}^{P}\mathbf{z}_p$ is the sample mean, and $p = \{1, \ldots, P\}$ is the number of observed measurements at one time instant. Suppose that the respective target matrix is chosen

$$\mathbf{T} = \begin{cases} \left[\hat{\mathbf{C}}\right]_{ii}, & \text{if} \quad i = j \\ 0 \quad , & \text{if} \quad i \neq j \end{cases}, \tag{6.21}$$

then the optimal weight is determined by minimizing the Frobenius norm of difference between $\hat{\mathbf{S}}$ and $\mathbf{C}_0$ with respect to λ given by

$$\begin{aligned} R(\mathbf{C}_0, \hat{\mathbf{S}}) &= \min_{\lambda}\ \mathrm{E}[\|\hat{\mathbf{S}} - \mathbf{C}_0\|_F^2] \\ &= \min_{\lambda}\ \mathrm{E}[\|(1-\lambda)\hat{\mathbf{C}} + \lambda\mathbf{T} - \mathbf{C}_0\|_F^2], \end{aligned} \tag{6.22}$$

and by expanding (6.22) gives

$$R(\lambda) = \lambda^2 \mathbb{E}\big[\big\|\mathbf{C}_0 - \mathbf{T}\big\|_F^2\big] + (1-\lambda)^2 \mathbb{E}\big[\big\|\hat{\mathbf{C}} - \mathbf{C}_0\big\|_F^2\big] + 2\lambda(1-\lambda)\mathbb{E}\big[\langle \mathbf{C}_0 - \mathbf{T}, \hat{\mathbf{C}} - \mathbf{C}_0\rangle\big]. \tag{6.23}$$

Taking derivative with respect to λ and equating to zero yields

$$2\lambda \mathbb{E}\big[\big\|\mathbf{C}_0 - \mathbf{T}\big\|_F^2\big] - 2(1-\lambda)\mathbb{E}\big[\big\|\hat{\mathbf{C}} - \mathbf{C}_0\big\|_F^2\big] + 2(1-2\lambda)\mathbb{E}\big[\langle \mathbf{C}_0 - \mathbf{T}, \hat{\mathbf{C}} - \mathbf{C}_0\rangle\big] \;=\; 0. \tag{6.24}$$

Next the followings were defined

$$\begin{aligned} \alpha^2 &= \mathbb{E}\big[\big\|\mathbf{C}_0 - \mathbf{T}\big\|_F^2\big], \\ \beta^2 &= \mathbb{E}\big[\big\|\hat{\mathbf{C}} - \mathbf{C}_0\big\|_F^2\big], \\ \gamma^2 &= \mathbb{E}\big[\langle \mathbf{C}_0 - \mathbf{T}, \hat{\mathbf{C}} - \mathbf{C}_0\rangle\big], \end{aligned} \tag{6.25}$$

with decomposition [130]

$$\begin{aligned} \delta^2 &= \mathbb{E}\big[\big\|\hat{\mathbf{C}} - \mathbf{T}\big\|_F^2\big] \\ &= \mathbb{E}\big[\big\|\mathbf{C}_0 - \mathbf{T}\big\|_F^2\big] + \mathbb{E}\big[\big\|\hat{\mathbf{C}} - \mathbf{C}_0\big\|_F^2\big] + 2\mathbb{E}\big[\langle \mathbf{C}_0 - \mathbf{T}, \hat{\mathbf{C}} - \mathbf{C}_0\rangle\big] \\ &= \alpha^2 + \beta^2 + 2\gamma^2. \end{aligned} \tag{6.26}$$

By taking (6.25) and (6.26), equation (6.24) can be rewritten in the form of

$$2\lambda\alpha^2 - 2(1-\lambda)\beta^2 + 2\gamma^2 - 4\lambda\gamma^2 = 0, \tag{6.27}$$

and by utilizing $\mathbb{E}\big[\langle \mathbf{T} - \mathbf{C}_0, \hat{\mathbf{C}} - \mathbf{C}_0\rangle\big] = -\gamma^2$ yields

$$2\lambda\alpha^2 - 2(1-\lambda)\beta^2 - 2\gamma^2 + 4\lambda\gamma^2 = 0, \tag{6.28}$$

where

$$\begin{aligned}\lambda &= \frac{\beta^2+\gamma^2}{\delta^2},\\ &= \frac{\mathbb{E}\big[\big\|\hat{\mathbf{C}}-\mathbf{C}_0\big\|_F^2\big]+\mathbb{E}\big[\langle \mathbf{C}_0-\mathbf{T},\hat{\mathbf{C}}-\mathbf{C}_0\rangle\big]}{\mathbb{E}\big[\big\|\hat{\mathbf{C}}-\mathbf{T}\big\|_F^2\big]}.\end{aligned} \tag{6.29}$$

Given the results in [131], the equation can be further simplified as follows

$$\begin{aligned}\lambda &= \frac{\mathbb{E}\big[\mathrm{tr}[\hat{\mathbf{C}}^2]\big]-\mathbb{E}\big[\mathrm{tr}[\hat{\mathbf{C}}\mathbf{T}]\big]-\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0\hat{\mathbf{C}}]\big]+\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0\mathbf{T}]\big]}{\mathbb{E}\big[\big\|\hat{\mathbf{C}}-\mathbf{T}\big\|_F^2\big]},\\ &= \frac{\mathbb{E}\big[\mathrm{tr}[\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{T})]\big]-\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0(\hat{\mathbf{C}}-\mathbf{T})]\big]}{\mathbb{E}\big[\mathrm{tr}[(\hat{\mathbf{C}}-\mathbf{T})^2]\big]}.\end{aligned} \tag{6.30}$$

Let $\mathrm{tr}[\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{T})]$ and $\mathrm{tr}[(\hat{\mathbf{C}}-\mathbf{T})^2]$ are the exact unbiased estimators of $\mathbb{E}\big[\mathrm{tr}[\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{T})]\big]$ and $\mathbb{E}\big[\mathrm{tr}[(\hat{\mathbf{C}}-\mathbf{T})^2]\big]$, respectively, then

$$\lambda = \frac{\mathrm{tr}[\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{T})]-\hat{\mathrm{G}}}{\mathrm{tr}[(\hat{\mathbf{C}}-\mathbf{T})^2]}, \tag{6.31}$$

for $\hat{\mathrm{G}}=\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0(\hat{\mathbf{C}}-\mathbf{T})]\big]$. Noted that $\mathbb{E}\big[\hat{\mathbf{C}}\big]=\mathbf{C}_0$, the estimator of $\hat{\mathrm{G}}$ can be further rewritten as

$$\begin{aligned}\hat{\mathrm{G}} &= \mathrm{tr}[\mathbf{C}_0^2]-\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0\mathbf{T}]\big]\\ &= \hat{a}-\hat{b},\end{aligned} \tag{6.32}$$

for $\hat{a}=\mathrm{tr}[\mathbf{C}_0^2]$ and $\hat{b}=\mathbb{E}\big[\mathrm{tr}[\mathbf{C}_0\mathbf{T}]\big]$. Under non-normality, the exact unbiased estimator of $\hat{a}$ is formulated as [165], [167]

$$\hat{a}_1 = \frac{(P-1)}{P(P-2)(P-3)}\Big((P-1)(P-2)\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n}+n\hat{c}_1-\frac{PQ}{n}\Big), \tag{6.33}$$

where

$$\hat{c}_1 = \frac{P-1}{P(P-2)(P-3)n^2}\Big(2\mathrm{tr}[\hat{\mathbf{C}}]+(P^2-3P+1)(\mathrm{tr}[\hat{\mathbf{C}}])^2-PQ\Big), \tag{6.34}$$

and

$$Q := \frac{1}{P-1}\sum_{p=1}^{P}\Big((\mathbf{z}_p - \bar{\mathbf{z}})^T(\mathbf{z}_p - \bar{\mathbf{z}})\Big)^2. \tag{6.35}$$

The equation (6.35) is proposed by [165] as another statistics to solve the estimation of $\hat{a}_1$. Meanwhile, for the estimation of $\hat{b}$, the exact unbiased estimator under non-normality is formulated as [168]

$$\hat{b}_1 = \frac{P-1}{(P+1)n}\mathrm{tr}[\mathbf{T}^2] - \frac{1}{P+1}\hat{b}_2, \tag{6.36}$$

where

$$\hat{b}_2 \;=\; \frac{-1}{(P-2)(P-3)n}\Big(2(P\,-\,1)^2\mathrm{tr}[\hat{\mathbf{C}}^2] + (P\,-\,1)^2(\mathrm{tr}[\hat{\mathbf{C}}])^2 \,-\, P(P\,+\,1)Q\Big). \tag{6.37}$$

Finally, by replacing estimators $\hat{a} = \hat{a}_1$ and $\hat{b} = \hat{b}_1$, respectively, in (6.31) the shrinkage weight under non-normality yields

$$\begin{aligned}\hat{\lambda} &= \frac{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \frac{\mathrm{tr}[\hat{\mathbf{C}}\mathbf{T}]}{n} - \hat{a}_1 + \hat{b}_1}{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \frac{\mathrm{tr}[\mathbf{T}^2]}{n}}, \\ &= \frac{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \frac{\mathrm{tr}[\hat{\mathbf{C}}\mathbf{T}]}{n} - \hat{a}_1 + \hat{b}_1}{\frac{\mathrm{tr}[\hat{\mathbf{C}}^2]}{n} - \hat{c}_1},\end{aligned} \tag{6.38}$$

which completes the derivation.

References

[1] C. H. Wang, Y. K. Huang, X. Y. Zheng, T. S. Lin, C. L. Chuang, and J. A. Jiang, "A self sustainable air quality monitoring system using WSN," in *Proc. of the 5th IEEE Int. Conf. on Service-Oriented Computing and Applicat.*, Dec. 2012, pp. 1–6.

[2] S. Bhattacharya, S. Sridevi, and R. Pitchiah, "Indoor air quality monitoring using wireless sensor network," in *Proc. of the 6th Int. Conf. on Sensing Technology*, Dec. 2012, pp. 422–427.

[3] A. Kadri, E. Yaacoub, M. Mushtaha, and A. Abu-Dayya, "Wireless sensor network for real-time air pollution monitoring," in *Proc. of the 1st Int. Conf. on Commun., Signal Process., and their Applicat.*, Feb. 2013, pp. 1–5.

[4] M. Pavani and P. T. Rao, "Real time pollution monitoring using wireless sensor networks," in *Proc. IEEE 7th Annual Inform. Technology, Electronics and Mobile Communication Conf.*, Oct. 2016, pp. 1–6.

[5] J. Zhang, W. Li, Z. Yin, S. Liu, and X. Guo, "Forest fire detection system based on wireless sensor network," in *Proc. of the 4th IEEE Conf. on Industrial Electronics and Applicat.*, May. 2009, pp. 520–523.

[6] P. Bolourchi and S. Uysal, "Forest fire detection in wireless sensor network using fuzzy logic," in *Proc. of the 5th Int. Conf. on Computational Intell., Commun. Syst. and Networks*, Jun. 2013, pp. 83–87.

[7] L. Guang-Hui, Z. Jun, and W. Zhi, "Research on forest fire detection based on wireless

sensor network," in *Proc. of the 6th World Congress on Intell. Control and Automation*, vol. 1, 2006, pp. 275–279.

[8] M. Hefeeda and M. Bagheri, "Wireless sensor networks for early detection of forest fires," in *Proc. IEEE Int. Conf. on Mobile Adhoc and Sensor Syst.*, Oct. 2007, pp. 1–6.

[9] I. Benkhelifa, N. Nouali-Taboudjemat, and S. Moussaoui, "Disaster management projects using wireless sensor networks: An overview," in *Proc. of the 28th Int. Conf. on Advanced Inform. Networking and Applicat. Workshops*, May. 2014, pp. 605–610.

[10] A. Khan and L. Jenkins, "Undersea wireless sensor network for ocean pollution prevention," in *Proc. of the 3rd Int. Conf. on Commun. Syst. Software and Middleware and Workshops.*, Jan. 2008, pp. 2–8.

[11] D. Chen, Z. Liu, L. Wang, M. Dou, J. Chen, and H. Li, "Natural disaster monitoring with wireless sensor networks: A case study of data-intensive applications upon low-cost scalable systems," *Mobile Netw. Appl.*, vol. 18, no. 5, pp. 651–663, Oct. 2013.

[12] A. Pascale, M. Nicoli, F. Deflorio, B. D. Chiara, and U. Spagnolini, "Wireless sensor networks for traffic management and road safety," *IET Intell. Transport Syst.*, vol. 6, no. 1, pp. 67–77, Mar. 2012.

[13] E. H. Ng, S. L. Tan, and J. G. Guzman, "Road traffic monitoring using a wireless vehicle sensor network," in *Proc. Int. Symp.on Intell. Signal Process. and Commun. Syst.*, Feb. 2009, pp. 1–4.

[14] W. Balid, H. Tafish, and H. H. Refai, "Development of portable wireless sensor network system for real-time traffic surveillance," in *Proc. IEEE 18th Int. Conf. on Intell. Transportation Syst.*, Sept. 2015, pp. 1630–1637.

[15] J. Paek, K. Chintalapudi, R. Govindan, J. Caffrey, and S. Masri, "A wireless sensor network for structural health monitoring: performance and experience," in *Proc. of the 2nd IEEE Workshop on Embedded Networked Sensors*, May. 2005, pp. 1–10.

[16] N. Kurata, S. Saruwatari, and H. Morikawa, "Ubiquitous structural monitoring using

wireless sensor networks," in *Int. Symp. on Intell. Signal Process. and Commun.*, Dec. 2006, pp. 99–102.

[17] M. Reyer, S. Hurlebaus, J. Mander, and O. E. Ozbulut, "Design of a wireless sensor network for structural health monitoring of bridges," in *Proc. of the 5th Int. Conf. on Sensing Technology*, Nov. 2011, pp. 515–520.

[18] A. Ledeczi, T. Hay, P. Volgyesi, D. R. Hay, A. Nadas, and S. Jayaraman, "Wireless acoustic emission sensor network for structural monitoring," *IEEE Sensors J.*, vol. 9, no. 11, pp. 1370–1377, Nov. 2009.

[19] N. Srivastava, "Challenges of next-generation wireless sensor networks and its impact on society," *J. of Telecommun.*, vol. 1, pp. 128–133, Feb. 2010.

[20] H. Huo, Y. Xu, H. Yan, S. Mubeen, and H. Zhang, "An elderly health care system using wireless sensor networks at home," in *Proc. of the 3rd Int. Conf.on Sensor Technologies and Applicat.*, Jun. 2009, pp. 158–163.

[21] J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, A. Terzis, and M. Welsh, "Wireless sensor networks for healthcare," *Proc. of the IEEE*, vol. 98, no. 11, pp. 1947–1960, Nov. 2010.

[22] C. R. Baker, K. Armijo, S. Belka, M. Benhabib, V. Bhargava, N. Burkhart, A. D. Minassians, G. Dervisoglu, L. Gutnik, M. B. Haick, C. Ho, M. Koplow, J. Mangold, S. Robinson, M. Rosa, M. Schwartz, C. Sims, H. Stoffregen, A. Waterbury, E. S. Leland, T. Pering, and P. K. Wright, "Wireless sensor networks for home health care," in *Proc. of the 21st Int. Conf. on Advanced Inform. Networking and Applicat. Workshops.*, vol. 2, May. 2007, pp. 832–837.

[23] E. Onur, C. Ersoy, H. Delic, and L. Akarun, "Surveillance wireless sensor networks: Deployment quality analysis," *IEEE Network*, vol. 21, no. 6, pp. 48–53, Nov. 2007.

[24] D. E. Burgner and L. A. Wahsheh, "Security of wireless sensor networks," in *Proc. of the 8th Int. Conf. on Inform. Technology: New Generations*, Apr. 2011, pp. 315–320.

[25] S. H. Lee, S. Lee, H. Song, and H. S. Lee, "Wireless sensor network design for tactical

military applications: Remote large-scale environments," in *Proc. of the 28th IEEE Conf. on Military Commun.*, 2009, pp. 911–917.

[26] G. Tuna, V. C. Gungor, and K. Gulez, "Wireless sensor networks for smart grid applications: A case study on link reliability and node lifetime evaluations in power distribution systems," *Int. J. of Distributed Sensor Networks*, vol. 9, no. 2, Jan. 2013.

[27] Y. Chen, C. M. Twigg, O. A. Sadik, and S. Tong, "A self-powered adaptive wireless sensor network for wastewater treatment plants," in *Proc. IEEE Int. Conf. on Pervasive Computing and Commun. Workshops*, Mar. 2011, pp. 356–359.

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

[29] Z. Zaidi and B. 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.

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

[31] 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.

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

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

[34] 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.

[35] Z. R. Zaidi and B. L. Mark, "A mobility tracking model for wireless ad hoc networks,"

in *Proc. Conf. IEEE Wireless Commun. and Netw. (WCNC)*, vol. 3, Mar. 2003, pp. 1790–1795.

[36] 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.

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

[38] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises*. John Wiley and Sons, Inc, 2012.

[39] E. A. Wan and R. V. D. Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proc. IEEE Adaptive Syst. for Signal Process., Commun., and Control Symp. (Cat. No.00EX373)*, 2000, pp. 153–158.

[40] 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.

[41] 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.

[42] A. Mesmoudi, M. Feham, and N. Labraoui, "Wireless sensor networks localization algorithms: a comprehensive survey," *Int. J. of Computer Netw. Commun.*, vol. 5, no. 6, pp. 45–64, Nov. 2013.

[43] G. Han, H. Xu, T. Q. Duong, J. Jiang, and T. Hara, "Localization algorithms of wireless sensor networks: a survey," *Telecommun. Syst.*, vol. 52, no. 4, pp. 2419–2436, Apr. 2013.

[44] U. Nazir, N. Shahid, M. A. Arshad, and S. H. Raza, "Classification of localization

algorithms for wireless sensor network: A survey," in *Proc. Int. Conf. on Open Source Syst. and Technologies*, Dec. 2012, pp. 1–5.

[45] K. Pahlavan and A. H. Levesque, *Wireless Inform. Networks*. John Wiley & Sons, 2005.

[46] J. Miranda, R. Abrishambaf, T. Gomes, P. Gonçalves, J. Cabral, A. Tavares, and J. Monteiro, "Path loss exponent analysis in wireless sensor networks: Experimental evaluation," in *Proc. of the 11th IEEE Int. Conf. on Industrial Informatics*, July. 2013, pp. 54–58.

[47] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[48] N. Patwari, R. O'Dea, and Y. Wang, "Relative location in wireless networks," in *Proc. of the 53rd IEEE Vehicular Technology*, vol. 2, May. 2001, pp. 1149–1153.

[49] A. W. A.J. Coulson and R. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 494–502, Apr. 1998.

[50] V. Sucasas, S. Hadzic, H. Marques, J. Rodriguez, and R. Tafazolli, "Performance evaluation of RSS based localization systems in mobile environments," in *Proc. of the 17th IEEE Int. Workshop on Comput. Aided Modeling and Design of Commun. Links and Netw.*, Sept. 2012, pp. 26–30.

[51] P. Bahl, V. N. Padmanabhan, and A. Balachandran, "Enhancements to the radar user location and tracking system," Microsoft Research, Tech. Rep., 2000.

[52] C. Steiner, F. Althaus, F. Trösch, and A. Wittneben, "Ultra-wideband geo-regioning: A novel clustering and localization technique," *EURASIP J. on Advances in Signal Process., Special Issue on Signal Process. for Location Estimation and Tracking in Wireless Environments*, Nov. 2007.

[53] R. Li and Z. Fang, "LLA: a new high precision mobile node localization algorithm based on toa," *J. of Commun.*, vol. 8, no. 5, pp. 604–611, Aug. 2010.

[54] K. Yu, Y. Guo, and M. Hedley, "Toa-based distributed localisation with unknown internal delays and clock frequency offsets in wireless sensor networks," *IET Signal Process.*, vol. 3, no. 2, pp. 106–118, Mar. 2009.

[55] G. Mao, B. Fidan, and B. D. Anderson, "Wireless sensor network localization techniques," *Computer Networks*, vol. 51, no. 10, pp. 2529 – 2553, 2007.

[56] J. Sallai, G. Balogh, M. Maróti, Ákos Lédeczi, and B. Kusy, "Acoustic ranging in resource-constrained sensor networks," in *Proc. of ICWN*, 2004, p. 04.

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

[58] R. Peng and M. Sichitiu, "Angle of arrival localization for wireless sensor networks," in *IEEE Sensor and Ad Hoc Commun. and Netw.*, vol. 1, Sept. 2006, pp. 374–382.

[59] P. Kułakowski, J. Vales-Alonso, E. Egea-López, W. Ludwin, and J. García-Haro, "Angle-of-arrival localization based on antenna arrays for wireless sensor networks," *Comput. and Elect. Eng.*, vol. 36, no. 6, pp. 1181 – 1186, 2010.

[60] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *IEEE Pers. Commun.*, vol. 7, no. 5, pp. 28–34, Oct. 2000.

[61] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proc. of the 9th Int. Conf. on Mobile Computing and Networking*, 2003, pp. 81–95.

[62] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proc. of the 9th Annual Int. Conf. on Mobile Computing and Networking*, 2003, pp. 81–95.

[63] R. Nagpal, H. Shrobe, and J. Bachrach, "Organizing a global coordinate system from local information on an ad hoc sensor network," in *Proc. of the 2nd Int. Conf. on Inform. Process. in Sensor Networks*, 2003, pp. 333–348.

[64] M. Jin, S. Xia, H. Wu, and X. Gu, "Scalable and fully distributed localization with mere connectivity," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 3164–3172.

[65] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from mere connectivity," in *Proc. of the 4th ACM Int. Symp. on Mobile Ad Hoc Networking & Computing*, 2003, pp. 201–212.

[66] R. C. Luo, O. Chen, and S. H. Pan, "Mobile user localization in wireless sensor network using grey prediction method," in *Proc. of the 31st Annual Conf. of IEEE Industrial Electron. Society*, Nov. 2005.

[67] P. M. Djuric, M. Vemula, M. F. Bugallo, and J. Miguez, "Non-cooperative localization of binary sensors," in *Proc. IEEE/SP 13th Workshop on Stat. Signal Process.*, July. 2005, pp. 1244–1249.

[68] A. A. Kannan, G. Mao, and B. Vucetic, "Simulated annealing based wireless sensor network localization with flip ambiguity mitigation," in *Proc. IEEE 63rd Vehicular Technology Conf.*, vol. 2, May. 2006, pp. 1022–1026.

[69] C. Alippi and G. Vanini, "A rssi-based and calibrated centralized localization technique for wireless sensor networks," in *Proc. of the 4th IEEE Int. Conf. on Pervasive Computing and Commun. Workshops*, Mar. 2006.

[70] A. A. Ahmed, H. Shi, and Y. Shang, "Sharp: a new approach to relative localization in wireless sensor networks," in *Proc. of the 25th IEEE Int. Conf.on Distributed Computing Syst. Workshops*, Jun. 2005, pp. 892–898.

[71] R. Huang, G. V. Zaruba, and M. Huber, "Complexity and error propagation of localization using interferometric ranging," in *Proc. IEEE Int. Conf. on Commun.*, Jun. 2007, pp. 3063–3069.

[72] K. Ramya, K. P. Kumar, and V. S. Rao, "A survey on target tracking techniques in wireless sensor networks," *J. of Computer Science & Engineering Survey*, vol. 3, no. 4, pp. 93–108, Aug. 2012.

[73] A. B. I. Pereira, "A pragmatic approach of localization and tracking algorithms in wireless sensor networks," Ph.D. dissertation, Universitat Autonoma de Barcelona, 2012.

[74] A. Oracevic and S. Ozdemir, "A survey of secure target tracking algorithms for wireless sensor networks," in *Proc. World Congress on Computer Applicat. and Inform. Syst.*, Jan. 2014, pp. 1–6.

[75] R. O. Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. of the 44th IEEE Conf. on Decision and Control*, Dec. 2005, pp. 8179–8184.

[76] A. T. Kamal, J. A. Farrell, and A. K. Roy-Chowdhury, "Information consensus for distributed multi-target tracking," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Jun. 2013, pp. 2403–2410.

[77] M. A. Wahdan, M. F. Al-Mistarihi, and M. Shurman, "Static cluster and dynamic cluster head (scdch) adaptive prediction-based algorithm for target tracking in wireless sensor networks," in *Proc. of the 38th Int. Convention on Inform. and Commun. Technology, Electronics and Microelectronics*, May. 2015, pp. 596–600.

[78] W. P. Chen, J. C. Hou, and L. Sha, "Dynamic clustering for acoustic target tracking in wireless sensor networks," *IEEE Trans. Mobile Computing*, vol. 3, no. 3, pp. 258–271, Jul. 2004.

[79] J. Rosswog and K. Ghose, "Detecting and tracking spatio-temporal clusters with adaptive history filtering," in *Proc. IEEE Int. Conf.on Data Mining Workshops*, Dec. 2008, pp. 448–457.

[80] H. T. Kung and D. Vlah, "Efficient location tracking using sensor networks," in *Proc. IEEE Wireless Commun. and Networking*, vol. 3, Mar. 2003, pp. 1954–1961.

[81] C. Y. Lin, W. C. Peng, and Y. C. Tseng, "Efficient in-network moving object tracking in wireless sensor networks," *Proc. IEEE Trans. Mobile Computing*, vol. 5, no. 8, pp. 1044–1056, Aug. 2006.

[82] H. W. Tsai, C. P. Chu, and T. S. Chen, "Mobile object tracking in wireless sensor networks," *Comput. Commun.*, vol. 30, no. 8, pp. 1811–1825, Jun. 2007.

[83] W. Zhang and G. Cao, "Dctc: dynamic convoy tree-based collaboration for target tracking in sensor networks," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1689–1701, Sept. 2004.

[84] B. Jia, K. D. Pham, E. Blasch, D. Shen, Z. Wang, and G. Chen, "Cooperative space object tracking using consensus-based filters," in *Proc. of the 17th Int. Conf. on Inform. Fusion*, Jul. 2014, pp. 1–8.

[85] A. Petitti, D. D. Paola, A. Rizzo, and G. Cicirelli, "Consensus-based distributed estimation for target tracking in heterogeneous sensor networks," in *Proc. of the 50th IEEE Conf. on Decision and Control and European Control*, Dec. 2011, pp. 6648–6653.

[86] R. V. D. Merwe, A. Doucet, N. de Freitas, and E. Wan, "The unscented particle filter," in *Adv. Neural Inform. Process. Syst.*, Dec. 2000.

[87] Y. Wang, S. Sun, and L. Li, "Adaptively robust unscented Kalman filter for tracking a maneuvering vehicle," *J. of Guidance, Control, and Dynamics*, 2014.

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

[89] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York, USA: John Wiley & Sons, Inc., 2001.

[90] X. Li and V. Jilkov, "Survey of maneuvering target tracking. part I. dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct 2003.

[91] R. Moose, "An adaptive state estimation solution to the maneuvering target problem," *IEEE Trans. on Autom. Control*, vol. 20, no. 3, pp. 359–362, Jun. 1975.

[92] G. Fuquan, D. Chuanhong, and L. Jianfeng, "Initial alignment of strap down inertial navigation system using Kalman filter," in *Proc. Int. Conf. on Computer Applicat. and Syst. Modeling*, vol. 3, Oct. 2010, pp. 629–633.

[93] J. L. Anderson, "Ensemble Kalman filters for large geophysical applications," *IEEE Control Syst.*, vol. 29, no. 3, pp. 66–82, Jun. 2009.

[94] M. Mathe, S. P. Nandyala, and T. K. Kumar, "Speech enhancement using Kalman filter for white, random and color noise," in *Proc. Int. Conf. on Devices, Circuits and Syst.*, Mar. 2012, pp. 195–198.

[95] Z. Xu and M. F. Rahman, "Comparison of a sliding observer and a Kalman filter for direct-torque-controlled ipm synchronous motor drives," *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4179–4188, Nov. 2012.

[96] N. Mahler, "Modeling the s & p 500 index using the Kalman filter and the laglasso," in *Proc. IEEE Int. Workshop on Machine Learning for Signal Process.*, Sept. 2009, pp. 1–6.

[97] P. Lall, R. Lowe, and K. Goebel, "Extended Kalman filter models and resistance spectroscopy for prognostication and health monitoring of leadfree electronics under vibration," in *Proc. IEEE Conf. on Prognostics and Health Management*, Jun. 2011, pp. 1–12.

[98] F. Kong, Y. Chen, J. Xie, G. Zhang, and Z. Zhou, "Mobile robot localization based on extended Kalman filter," in *Proc. of the 6th World Congress on Intell. Control and Automation*, vol. 2, 2006, pp. 9242–9246.

[99] T. Gehrig, K. Nickel, H. K. Ekenel, U. Klee, and J. McDonough, "Kalman filters for audio-video source localization," in *IEEE Workshop on Applicat. of Signal Process. to Audio and Acoustics*, Oct. 2005, pp. 118–121.

[100] Y. Du and F. Yuan, "Real-time vehicle tracking by Kalman filtering and gabor decomposition," in *Proc. of the 1st Int. Conf. on Information Science and Engineering*, Dec. 2009, pp. 1386–1390.

[101] F. A. Faruqi and R. C. Davis, "Kalman filter design for target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-16, no. 4, pp. 500–508, July. 1980.

[102] F. Daum, "Nonlinear filters: beyond the Kalman filter," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 8, pp. 57–69, Aug. 2005.

[103] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[104] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[105] M. B. Luca, S. Azou, G. Burel, and A. Serbanescu, "On exact Kalman filtering of polynomial systems," *IEEE Trans. Circuits Syst. I: Regular Papers*, vol. 53, no. 6, pp. 1329–1340, Jun. 2006.

[106] M. Baum, B. Noack, F. Beutler, D. Itte, and U. D. Hanebeck, "Optimal Gaussian filtering for polynomial systems applied to association-free multi-target tracking," in *Proc. of the 14th Int. Conf. on Inform. Fusion*, July. 2011, pp. 1–8.

[107] R. Van Der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, 2004.

[108] O. Cappe, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. of the IEEE*, vol. 95, no. 5, pp. 899–924, May. 2007.

[109] G. Oppenheim, A. Philippe, and J. de Rigal, "The particle filters and their applications," *Chemometrics and Intell. Laboratory Syst.*, vol. 91, no. 1, pp. 87 – 93, 2008.

[110] W. Li, Z. Wang, Y. Yuan, and L. Guo, "Particle filtering with applications in networked systems: a survey," *Complex & Intell. Syst.*, vol. 2, no. 4, pp. 293–315, Dec. 2016.

[111] D. B. Rubin, "A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when the fraction of missing information is modest: The SIR algorithm." *J. of American Stat. Assoc.*, vol. 82, no. 398, pp. 543–546, 1987.

[112] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-

Gaussian Bayesian state estimation," *IEE Proc. F - Radar and Signal Process.*, vol. 140, no. 2, pp. 107–113, April 1993.

[113] R. Chen and J. S. Liu, "Predictive updating methods with application to Bayesian classification," *J. of Royal Stat. Soc. Series B (Methodological)*, vol. 58, no. 2, pp. 397–415, 1996.

[114] A. Kong, J. S. Liu, and W. H. Wong, "Sequential imputations and Bayesian missing data problems," *J. of the American Stat. Assoc.*, vol. 89, no. 425, pp. 278–288, 1994.

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

[116] 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.

[117] J. D. Hol, T. B. Schon, and F. Gustafsson, "On resampling algorithms for particle filters," in *Proc. IEEE Nonlinear Statistical Signal Process. Workshop*, Sept. 2006, pp. 79–82.

[118] J. S. Liu, R. Chen, and T. Logvinenko, "A theoretical framework for sequential importance sampling with resampling," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. New York: Springer, 2001, pp. 225–246.

[119] L. Mihaylova, D. Bull, D. Angelova, and C. Canagarajah, "Mobility tracking in cellular networks with sequential Monte Carlo filters," in *Proc. of the 8th Int. Conf. on Inform. Fusion*, 2005.

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

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

[122] O. Ledoit and M. N. Wolf, "Honey, I shrunk the sample covariance matrix," *J. Portfolio Manage.*, vol. 30, pp. 110–119, 2003.

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

[124] J. Schäfer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Stat. Applicat. in Genetics and Molecular Biology*, vol. 4, Nov. 2005.

[125] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 2, pp. 115–127, Mar. 1986.

[126] C. Stein, "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution," in *Proc. of the 3rd Berkeley Symp. on Mathematical Stat. and Probability, Volume 1: Contributions to the Theory of Stat.* University of California Press, 1956, pp. 197–206.

[127] W. F. Sharpe, "A simplified model for portfolio analysis," *Manage. Sci.*, vol. 9, no. 2, pp. 277–293, 1963.

[128] P. A. Frost and J. E. Savarino, "An empirical Bayes approach to efficient portfolio selection," *J. of Financial and Quantitative Anal.*, vol. 21, no. 03, pp. 293–305, 1986.

[129] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portofolio selection," Department of Economics and Business, Universitat Pompeu Fabra, Economics Working Papers 586, Nov. 2001.

[130] O. Ledoit and M. N. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. of Multivariate Anal.*, vol. 88, no. 2, pp. 365 – 411, 2004.

[131] T. J. Fisher and X. Sun, "Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix," *Computational Stat. and Data Anal.*, vol. 55, no. 5, pp. 1909 – 1918, 2011.

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

[133] A. Tulsyan, B. Huang, R. Gopaluni, and J. 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.

[134] M. Lei, B. J. V. 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.

[135] 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.

[136] B. Ristic, S. Arulampalam, and N. C. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Application*. Artech House, 2004.

[137] 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.

[138] Linkoping University, *Sensor Fusion App*, online at https://play.google.com/store/apps.

[139] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part ii," *IEEE Robot. & Autom. Mag.*, vol. 13, no. 3, pp. 108–117, Sept. 2006.

[140] A. Gil, O. Reinoso, O. M. Mozos, C. Stachniss, and W. Burgard, "Improving data association in vision-based SLAM," in *Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Syst.*, Oct. 2006, pp. 2076–2081.

[141] S. A. Berrani and C. Garcia, "Robust detection of outliers for projection-based face recognition methods," *Multimedia Tools and Applications*, vol. 38, no. 2, pp. 271–291, Jun. 2008.

[142] J. A. Ting, A. D'Souza, and S. Schaal, "Automatic outlier detection: A Bayesian approach," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007.

[143] R. K. Pearson, "Outliers in process modeling and identification," *IEEE Trans. on Control Syst. Technol.*, vol. 10, no. 1, pp. 55–63, Jan. 2002.

[144] H. J. Motulsky and R. E. Brown, "Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate," *BMC Bioinformatics*, vol. 7, no. 1, p. 123, Mar. 2006.

[145] J. Morris, "The Kalman filter: A robust estimator for some classes of linear quadratic problems," *IEEE Trans. Inf. Theory*, vol. 22, no. 5, pp. 526–534, Sept. 1976.

[146] S. Sarkka and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational Bayesian approximations," *IEEE Trans. Autom. Control*, vol. 54, no. 3, pp. 596–600, Mar. 2009.

[147] S. Roujol, B. D. de Senneville, S. Hey, C. Moonen, and M. Ries, "Robust adaptive extended Kalman filtering for real time mr-thermometry guided hifu interventions," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 533–542, Mar. 2012.

[148] Z. M. Durovic and B. D. Kovacevic, "Robust estimation with unknown noise statistics," *IEEE Trans. Autom. Control*, vol. 44, no. 6, pp. 1292–1296, Jun. 1999.

[149] S. C. Chan, Z. G. Zhang, and K. W. Tse, "A new robust Kalman filter algorithm under outliers and system uncertainties," in *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 5, May. 2005, pp. 4317–4320.

[150] J. A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," in *Proc. IEEE Int. Conf. Intell. Robots and Syst.*, Oct. 2007.

[151] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. of Royal Stat. Soc. Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[152] G. Kitagawa, "Non-Gaussian state-space modeling of nonstationary time series," *J. of the American Statistical Assoc.*, vol. 82, no. 400, pp. 1032–1041, 1987.

[153] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *J. of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[154] N. J. Gordon and A. F. M. Smith, "Approximate non-gaussian Bayesian estimation and modal consistency," *J. of the Royal Stat. Soc. Series B (Methodological)*, vol. 55, no. 4, pp. 913–918, 1993.

[155] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," in *Proc. IEEE Int. Conf. on Robotics and Automation*, May. 2011, pp. 1551–1558.

[156] M. J. Beal, "Variational algorithms for approximate Bayesian inference," Ph.D. dissertation, University College London, 2003.

[157] R. Piché, S. Särkkä, and J. Hartikainen, "Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate student-t distribution," in *Proc. IEEE Int. Workshop on Machine Learning for Signal Process.*, Sept. 2012, pp. 1–6.

[158] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "Approximate inference in state-space models with heavy-tailed noise," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5024–5037, Oct. 2012.

[159] M. Roth, E. Özkan, and F. Gustafsson, "A student's t filter for heavy tailed process and measurement noise," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process.*, May. 2013, pp. 5770–5774.

[160] H. Nurminen, T. Ardeshiri, R. Piché, and F. Gustafsson, "Robust inference for state-space models with skewed measurement noise," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1898–1902, Nov. 2015.

[161] B. Liu, "Robust particle filter by dynamic averaging of multiple noise models," in *Proc. 42nd IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Mar. 2017, pp. 4034–4038.

[162] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: a tutorial," *Stat. Sci.*, vol. 14, no. 4, pp. 382–417, Nov. 1999.

[163] T. Klingenbrunn and P. Mogensen, "Modelling cross-correlated shadowing in network simulations," in *Proc. IEEE VTC*, vol. 3, Sept. 1999, pp. 1407–1411.

[164] S. Kotz and S. Nadarajah, *Multivariate t Distributions and Their Applications*. Cambridge University Press, 2004.

[165] T. Himeno and T. Yamada, "Estimations for some functions of covariance matrix in high dimension under non-normality and its applications," *J. of Multivariate Anal.*, vol. 130, pp. 27 – 44, 2014.

[166] B. Liu, "Instantaneous frequency tracking under model uncertainty via dynamic model averaging and particle filtering," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1810–1819, June. 2011.

[167] M. S. Srivastava, H. Yanagihara, and T. Kubokawa, "Tests for covariance matrices in high dimension with less sample size," *J. of Multivariate Anal.*, vol. 130, pp. 289 – 309, 2014.

[168] Y. Ikeda, T. Kubokawa, and M. S. Srivastava, "Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions," *Computational Stat. and Data Anal.*, vol. 95, pp. 95 – 108, 2016.