Behailu Yohannes Shikur
Max-Planck-Straße 4A
D–18059 Rostock

# Exploiting the Spatio-temporal Channel Properties of Multiple Antenna Systems

Dissertation

zur Erlangung des akademischen Grades
**Doktor-Ingenieur (Dr.-Ing.)**
der Fakultät für Elektrotechnik und Informatik
der Universität Rostock

vorgelegt von
Behailu Yohannes Shikur
geb. am 12.09.1985 in Äthiopien

Tag der Abgabe: 23. Oktober 2015

urn:nbn:de:gbv:28-diss2016-0047-8

Als Dissertation genehmigt von
der Fakultät für Informatik und Elektrotechnik
der Universität Rostock

**Gutachter:**

- Prof. Dr.-Ing. habil. Tobias Weber (Universität Rostock)
- Prof. Dr.-Ing. Gerhard Bauch (Technische Universität Hamburg)
- Prof. Dr.-Ing. habil. Volker Kühn (Universität Rostock)

**Tag der öffentlichen Verteidigung:**

18. April 2016

# Erklärung

Ich erkläre, diese Arbeit selbständig angefertigt und die benutzten Unterlagen vollständig angegeben zu haben.

Rostock, den 23. Oktober 2015

\- - - - - - - - - - - - - - - - - - - - - - - -

(Behailu Yohannes Shikur)

# Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr.-Ing. Tobias Weber for his guidance and encouragement throughout my Ph.D. study. I have learned a lot and benefitted from his uncommon talent for systematic approach to quickly grasping the gist of different concepts and from his demand for a rigorous research work. His abilities have left a strong impression on me which I believe would continue to influence me for a long time. I am also indebted to all members of the Institute of Communications Engineering for providing a cordial and inspiring research environment. Most notably, I am greatly indebted to Prof. Dr.-Ing. Volker Kühn for reviewing my Ph.D. thesis and for the fruitful suggestions he has made throughout my Ph.D. study which have been worthwhile inputs to this thesis. I would also like to extend my gratitude to Prof. Dr.-Ing. Gerhard Bauch from Hamburg University of Technology for kindly devoting his valuable time to review my Ph.D. thesis. I would be remiss if I fail to extend my gratitude to Prof. Dr.-Ing. Sascha Spors and Prof. Dr.-Ing. Erika Müller for their noteworthy suggestions and comments. My sincere appreciation also goes to my colleagues: Xiang Li, Karsten Wiedmann, Hussein Al-Shatri, Nico Palleit and Enrico Ihde. It was a great pleasure to work with you. Last but not least, I would like to thank my parents, Tsigé Markos and Yohannes Shikur, for their unwavering love and support throughout my life.

# Contents

# Chapter 1

# Introduction

## 1.1 Motivation

Multiple antenna wireless communication systems, herein called multiple-input multiple-output (MIMO) systems, offer dramatic improvements in bit rate [FG98, Tel99] and error rate [Ala98, TSC98] over single antenna wireless communication systems, herein called single-input single-output (SISO) systems, by exploiting the spatio-temporal channel properties of MIMO systems. In this thesis, the spatio-temporal channel properties of MIMO systems are exploited to yield valuable insights and new approaches to localization and channel prediction in MIMO systems. Localization and channel prediction may seem to be different problems at first glance. However, both problems are related in that radio localization is achieved by exploiting the properties of the radio channel, spatial and/or temporal. Consequently, though not within the scope of this thesis, joint estimation of the location and the channel of a mobile radio user can yield an improved localization and channel estimation performance as compared to independent estimation. In [DAB+13, Slo12], it has been shown that significant reduction in signaling overhead can be obtained in cellular networks, mobile ad hoc networks and cognitive radio by exploiting the location information of the users. For example, the user's position can be used to get information about the channel statistics from a location based channel statistics database, which in turn can be used to improve channel estimation [Slo12].

Location information is integral to advanced cellular and sensor network applications. The direct implication of MIMO systems for localization is the ability to estimate the angle of departure (AOD) and the angle-of-arrival (AOA) at the transmitter and receiver, respectively using antenna array processing techniques [VT02]. The AOD and the AOA measurements can be augmented with the conventionally used time-of-arrival (TOA) or time-difference-of-arrival (TDOA) measurements to get a more precise position estimate. A rather more systematic approach, studying the spatio-temporal channel properties of MIMO systems, reveals that there is a lot more to be gained from the estimates of the AOD and the AOA than a mere extra information in the spatial dimension. It will be shown that, by systematic spatio-temporal channel modeling, MIMO localization systems can deliver satisfactory performance in multipath environments, e.g., in terrestrial and indoor scenarios, where satellite based navigation systems often

show abysmal performance. The interesting result is that, just as in the case of MIMO communication systems where the multipath fading effect yields an increase in capacity, it will be shown that in MIMO localization systems the multipath propagations can be exploited to yield an improved localization performance. This result contrasts with localization in SISO systems where the detrimental multipath propagations are to be identified and mitigated. Consequently, unlike SISO localization systems which require at least three base stations (BSs) of known positions to estimate the unknown positions of a mobile station (MS), it will be shown that in MIMO localization systems a single BS is sufficient to estimate the position of a MS.

Reliable transmitter side channel state information, which in some cases can only be obtained by channel prediction, is required to achieve the potential high performance of transmission techniques like transmitter beamforming, precoding, adaptive modulation and coding. The straightforward approach to MIMO channel prediction would be the application of the typical SISO channel prediction algorithms to predict the SISO subchannels of the MIMO channel. However, this approach leaves the spatial properties of the MIMO channel unexploited. By exploiting the spatio-temporal channel properties of MIMO systems, it will be shown that the MIMO channel can be predicted satisfactorily using outdated MIMO channel estimates. Analysis of the spatio-temporal channel properties reveals that prediction of the MIMO channel coefficients can be performed using a linear and time and frequency independent filter with a dimension at least equal to the number of propagation paths in the MIMO channel. This result has crucial importance in the design of optimal channel predictors.

Currently, antenna arrays are conveniently implemented in fixed stations, such as base stations and WiFi access points, or mobile stations that are fixed on vehicles [MSL+09]. Due to physical space constraints, it has been so far difficult to implement antenna arrays in hand-held devices. Nevertheless, it has been shown that integration of antenna arrays in hand-held devices is feasible [KBF+05]. It is expected that future wireless communication systems will employ antenna arrays in hand-held devices fabricated by microstrip technology [MSL+09]. Thus it is not too optimistic to envisage MIMO systems being ubiquitous in the near future. The ideas presented in thesis will be of paramount importance in practice as they yield elegant solutions to localization and channel prediction in MIMO systems.

## 1.2 State of the art and open problems

### 1.2.1 Localization

Radio localization, which uses radio signals to obtain information about the position of a user, is an essential radio technology which has applications in many areas. In cellular networks, localization enables location based services such as localization of emergency callers, navigational service, location based advertising, location based billing, etc. In sensor networks, which monitor and record physical or environmental conditions, localization provides information about the positions of the sensors which is essential to make a meaningful interpretation of the sensed data. The importance of localization is not limited to cellular and sensor networks; there are a number of wireless communication services which rely on the accurate location information of the users. These include context-aware services, augmented reality, guided tour of museums, etc. Of course, the issue of user's privacy is a cause for concern in location based services. It is imperative that the user shall have complete control over the acquisition, storage, retrieval and exchange of location data.

There are several radio localization techniques which yield satisfactory performance for scenarios where the transmitted signals propagate directly from the transmitter to the receiver. The two major localization techniques are the range based and the bearing based localization techniques. There also exist hybrid localization algorithms which jointly employ the range and bearing localization techniques to get an improved localization performance.

In range based localization, the distance between the BS and the MS is estimated from either the time-of-arrival (TOA) or the received signal strength measurements. The TOA can be estimated from the cross-correlation of the received signal with a locally generated replica of the transmitted signal, preferably a wide-band transmitted signal [SY03]. Moreover, super-resolution spectral estimation algorithms, e.g., the MUSIC (MUltiple SIgnal Classification) algorithm [Sch86], can also be used to obtain improved TOA estimates. The aforementioned approaches to estimating the TOA require synchronized transmitter and receiver clocks. The TOA can also be estimated without synchronized transmitter and receiver clocks from the round trip delay time taken of a signal being sent to the receiver and the acknowledgment being sent back to the transmitter. Alternatively, the distance between the BS and the MS can be estimated from the received signal strength using either an empirical path loss formula which has been determined for different environments [OOF68, Hat80] or theoretical path loss formula from radio propagation models.

If the distance between a MS and at least three BSs is known, then the position of the MS can be determined using trilateration. Given the distance between the BS and the MS, the possible positions of the MS are found on the perimeter of a circle with the BS at the center and a radius equal to the distance between the MS and the BS. Given at least three BSs, the position of the MS can be found at the intersection point of the three circles. Since the distance equation is nonlinear in the position of the MS, it is difficult to find the maximum likelihood estimate of the position of the MS. Hence, an iterative least squares algorithm based on linearizing the distance equation using a Taylor series expansion can be used to estimate the position of the MS. In [CSMC04], the distance equation is linearized by introducing a nuisance variable and the method of least squares is used to determine the position of the MS.

If the transmitter and the receiver clocks are not synchronized, then the time-difference-of-arrival (TDOA) of the signals from the different BSs to the MS can be measured. However, the clocks of the BSs need to be synchronized. Thus the distance differences between the different BSs to the MS are available. The position of the MS is estimated from the distance differences by multilateration. The distance difference between two BSs to the MS defines a hyperbola as the possible positions of the MS. Given at least four BSs, the position of the MS can be found at the intersection point of the three hyperbolas. Mathematically, the linearized least squares algorithm can be used to determine the position of the MS. A closed form solution, using the method of least squares, is given in [CH94] by reducing the distance difference equation to a linear form by introducing a nuisance variable.

In bearing based localization techniques, the bearings between the BSs and the MS are used to estimate the position of the MS by triangulation [SAK+92]. The bearing between the BS and the MS defines a straight line of possible positions of the MS. Given at least two bearings from two BSs, the position of the MS can be found at the intersection point of the straight lines. Mathematically, the nonlinear equations of the bearings can easily be reduced to a linear form in the position of the MS and the least squares method can be used to estimate the position of the MS.

In the discussion of the aforementioned localization techniques, it has been assumed that the transmitted signal propagates directly from the transmitter to the receiver. However, in indoor scenarios, e.g., shopping malls, museums, airports, and terrestrial outdoor scenarios, e.g., mountainous, urban and metropolitan areas, the transmitted signal may experience reflection, scattering and diffraction before arriving at the receiver [Mol05]. This results in multipath propagation where several copies of the transmitted signal superpose at the receiver with different delays, attenuations and phases. Furthermore, the line-of-sight (LOS) path between the transmitter and the receiver

might be obstructed and transmission is possible only via the non-line-of-sight (NLOS) multipath propagations.

If the LOS paths between the BSs and the MS are not obstructed, then satisfactory localization performance can still be obtained by identifying the LOS paths and simply discarding the multipath propagations. In such cases, the LOS path between a BS and a MS can be simply identified as the first arriving signal. However, localization of a MS becomes complicated if some or all of the LOS paths between the BSs and the MS are obstructed. Consequently, localization of a MS in NLOS multipath environments has been a hot research topic.

Signature based localization techniques are alternative localization techniques to range and bearing based localization techniques which can be used in both LOS and NLOS multipath environments. Signature based localization techniques estimate the position of a MS by finding the best match of the position dependent signatures of the received signal, e.g., received signal strength, path loss, power delay profile, etc. or a combination thereof, with the entries of position dependent signatures stored in a database [BP00, LJ05]. The database can be developed either from measurement campaigns or from radio propagation models. Signature based localization techniques yield a satisfactory performance also in NLOS multipath propagation environments as the actual channel information is already collected in the database. However, signature based localization techniques have the downside that they involve an arduous task of creating the database and maintaining the database as the environment changes.

There have been several research works on range based localization [GC09], bearing based localization [Xio98, GAMD04, AM08] and hybrid localization [CZ04, TQ08] in NLOS multipath environments. It has been assumed that the LOS propagation paths between some of the BSs and the MS are obstructed, whereas there are unobstructed LOS propagation paths between the remaining BSs and the MS. A straightforward approach to localization in NLOS multipath environments is to treat the estimated ranges and/or bearings as erroneous measurements and "average out" the influence of the NLOS multipath propagations by considering more measurements from several BSs. However, the localization accuracy of this approach would be susceptible to the number of BSs with NLOS propagation links to the MS and the deviations of the NLOS propagation paths from the LOS propagation paths. A more subtle approach would be to identify and discard the measurements from the BSs with NLOS propagation links to the MS. Localization of the MS would be possible as long as there are a sufficient number of BSs with LOS propagation links to the MS. However, since the BSs with NLOS propagation links to the MS are not known a priori, there is a possibility to misidentify a BS with NLOS propagation link as having a LOS propagation link and

a BS with LOS propagation link as having a NLOS propagation link [GC09]. This of course leads to performance degradation of the localization accuracy.

Systematic approaches to localization in NLOS multipath environments employ models to describe the influence of the NLOS multipath propagations and propose different algorithms to mitigate the influence of the NLOS multipath propagations. For range measurements, the effect of the NLOS multipath propagations on the measured distance between a BS and a MS is described by a positive distance bias. The positive distance bias has been modelled by different distributions using theoretical models and empirical models from measurements [GC09]. The variance of the measured distances for the NLOS propagation paths is assumed to be significantly higher than the variance of the measured distances for the LOS propagation paths. Likewise for bearing measurements, the effect of the NLOS multipath propagations on the measured bearing between a BS and a MS is modelled by an angular dispersion factor. Different distributions from geometrical models and empirical models from measurements have been proposed for the angular spread statistics [SJJS00, ECS+98]. Using the aforementioned models for range and bearing measurements in NLOS multipath environments, several classes of NLOS multipath mitigation techniques have been proposed [GC09, Xio98, GAMD04, AM08, CZ04, TQ08]:

- Least squares estimation techniques where the measurements are weighted in such a way that the NLOS multipath bias is suppressed.
- Constrained localization techniques which employ convex optimization techniques to localize a MS by utilizing the NLOS multipath measurements to define a constraint on the feasible region of the MS location.
- Robust estimation techniques which suppress the impact of the NLOS multipath propagations while maintaining a high estimation efficiency, i.e., low estimation variance regardless of the NLOS multipath propagation scenario.

The aforementioned approaches to localization in NLOS multipath environments rely on mitigating the influence of the NLOS multipath propagations on localizing a MS. In this thesis, an alternative approach to localization in NLOS multipath environments is considered by exploiting the spatio-temporal channel properties of MIMO systems. The alternative approach enables localization of a MS under the explicit consideration of scatterers based on the assumption that the propagation paths which are considered for localization contain a maximum of one scattering. The explicit consideration of scatterers enables exploitation of multipath propagations which yields a significant

improvement in localization performance. This approach contrasts with the NLOS multipath identification and mitigation approaches where NLOS multipath propagations are to be mitigated rather than exploited. Furthermore, unlike the NLOS multipath identification and mitigation approaches, there is no requirement for an unobstructed LOS path between some of the BSs and the MS. Owing to the explicit consideration of scatterers, satisfactory localization performance can be obtained in multipath environments even when all the propagation paths are NLOS.

The basic idea of MS localization in NLOS multipath environments under the explicit consideration of scatterers has already been considered. In [TCL01, PS09], the time-of-arrival, the angle-of-departure and the Doppler shift measurements have been used to localize a MS in NLOS multipath environments under the explicit consideration of scatterers. In [MYJ07, ST08, WPW11], the time-of-arrival, the angle-of-departure and the angle-of-arrival measurements have been used to localize a MS in NLOS multipath environments under the explicit consideration of scatterers. Furthermore, performance assessments using real world measurements have been performed [ST08, WPW11]. However, the idea of MS localization in NLOS multipath environments under the explicit consideration of scatterers has not been extensively researched. In this thesis, a comprehensive study of localization in NLOS multipath environments under the explicit consideration of scatterers is presented. Furthermore, the following open problems are addressed:

- How to exploit the spatio-temporal channel properties of MIMO systems to localize a MS in NLOS multipath environments under the explicit consideration of scatterers with incomplete measurements, e.g., owing to a MS which is unsynchronized with the BSs,
- How to exploit the spatio-temporal channel properties of MIMO systems to track a MS in NLOS multipath environments under the explicit consideration of scatterers and
- How to exploit the spatio-temporal channel properties of MIMO systems to localize MSs in network localization scenarios with NLOS multipath environments.

### 1.2.2 Channel prediction

The availability of channel state information (CSI) at the transmitter and receiver is essential in many mobile radio systems. Differential transmission techniques, which

require no channel state information at the transmitter and receiver, can be used in the absence of any channel state information under the assumption of constant channel state during at least two transmit symbol durations. However, similar to the case in SISO systems, differential transmission techniques with differential detection at the receiver incur a performance loss of approximately 3 dB as compared to coherent transmission techniques in MIMO systems [MSL+09]. Thus channel state information is necessary to achieve high performance in mobile radio systems. The CSI can be the instantaneous CSI or the statistical CSI such as the channel's correlation, average channel gain, etc. In this thesis, acquiring the instantaneous CSI is of interest. The receiver side CSI can be obtained by transmission of a priori known pilot symbols which are multiplexed with the useful data symbols [Cav91]. The receiver exploits the received a priori known pilot symbols to estimate the channel at the pilot symbol locations. The channels at other locations can then be interpolated using optimal Wiener filtering, for example [Cav91]. However, advanced transmission techniques like adaptive modulation and coding, adaptive power control, bit-loading and interference alignment require transmitter side CSI. Furthermore, MIMO transmission techniques such as network MIMO (coordinated multi-point transmission), pre-coders for spatial multiplexing, transmit antenna diversity, etc., also require transmitter side CSI. In practice, the receiver side CSI is used as transmitter side CSI under the assumption that the receiver side CSI is not outdated when used as a transmitter side CSI. In time-division-duplex (TDD) systems, the receiver side CSI is merely used as transmitter side CSI in the next transmission instant, whereas in frequency-division-duplex (FDD) systems, the receiver side CSI is fed back to the transmitter to be used as transmitter side CSI.

In time-varying frequency-selective mobile radio channels, the channel changes continually in time and frequency. Thus the receiver side CSI can be potentially outdated when used as a transmitter side CSI. Hence, in both TDD and FDD systems, readily using the receiver side CSI as transmitter side CSI can result in significant performance degradation. Furthermore, in FDD systems, feedback of the receiver side CSI reduces spectral efficiency as resources which could otherwise be used for transmission of useful data are used to transmit the estimated receiver side CSI. The reduction in spectral efficiency owing to feedback of the receiver side CSI is significantly pronounced in MIMO systems as the feedback requirements of a MIMO channel, i.e., the number of SISO subchannels of a MIMO channel, generally grow with the product of the number of the transmit antennas and the receive antennas [LHSH04]. Thus it is imperative that the receiver side CSI be used to obtain the transmitter side CSI by performing channel prediction.

Before discussing the approaches to channel prediction, it is worthwhile to discuss the

system function that shall be used to represent the mobile radio channel. The time-variant mobile radio channel can be characterized by four system functions [Bel63]. The choice of the system function has no influence on the performance of the channel prediction algorithms. However, among the four system functions, the time-variant channel transfer function, herein simply referred to as the channel transfer function, is advantageous in that it can easily be measured using a network analyzer or estimated from transmission of a priori known pilot symbols as it is the only non-sparse system function. Furthermore, the channel transfer function is usually easier to analyse than the other system functions. Hence, in this thesis, the channel transfer function is used to represent the mobile radio channel.

There are two major approaches to predicting the mobile radio channel. The first approach is based on modeling the channel transfer function using an autoregressive (AR) model, whereas the second approach is based on modeling the channel transfer function using the parameters of the radio wave propagations. In the AR model based channel prediction approach, the channel coefficient to be predicted is represented as a linear combination of past channel coefficients [EDHH98, HHDH99, DHHH00, Ekm02, SM05, Aro11, AVWS12]. If long range channel prediction is required, then an iterative multi-step prediction can be performed where each of the required channel coefficients are predicted one after the other iteratively [DH07]. It is assumed that the AR model parameters do not change or change slowly compared with the rate of channel fading within the time duration and frequency band under consideration. The AR model parameters can be estimated by the method of least squares using either true channel coefficients, if available, or empirical autocorrelation functions computed from measured channel coefficients [Mak75, Hay02]. In the case of time-varying AR model parameters, adaptive filters can be used to track the AR model parameters [HHDH99, SM05, DH07]. In the absence of estimates of the channel coefficients or the channel autocorrelation function, the WSSUS (Wide-Sense Stationary Uncorrelated Scattering) models can be used to estimate the channel autocorrelation function. A commonly used statistical model for the temporal correlations is the Jakes' model [Jak74]. For wideband channels the spectral correlations are considered which are often assumed to result from the one-sided exponential power delay profile [Gol05]. Thus the AR model parameters can be calculated from the Jakes' model and the one-sided exponential power delay profile given the maximum Doppler shift and the multipath spread of the channel.

In the parametric model based prediction approach, which is inspired by the geometry of the radio propagation paths, the channel transfer function is approximated as a superposition of a finite number of plane waves which represent the discrete propagations paths [HW98, AJJF99, VTR00, TV01, SK02, WE05, WE06b, CEV07, PW09b].

Each of the propagation paths is defined by a complex weight, delay and Doppler shift. It is assumed that the propagation path parameters do not change or change slowly compared with the rate of channel fading within the time duration and frequency band under consideration. This assumption implies that the mobile radio environment does not change or changes slowly compared with the rate of channel fading within the time duration and frequency band under consideration. The assumption holds very well when the receiver is moving with a constant velocity and the positions of the scatterers remain unchanged. In such a case the path parameters remain constant for a much longer duration than the coherence time of the channel [GS04, Pal11]. If this assumption holds, then given estimates of the path parameters the channel prediction task is a mere computation of the value of the channel transfer function, calculated using the estimated propagation path parameters, at the desired time and frequency. The path parameters are commonly estimated using subspace estimation methods such as the MUSIC (MUltiple SIgnal Classification) algorithm [Sch86], the ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) algorithm [RK89] or maximum likelihood methods such as the SAGE (Space-alternating Generalized Expectation-maximization) algorithm [FH94]. It must be noted that even though the approaches are different, the parametric model based approach is a special case of the AR model approach where all the poles of the AR filter lie on the unit circle.

A simple approach to MIMO channel prediction is to consider the SISO subchannels of a MIMO channel independently and perform SISO channel prediction for each SISO subchannel of a MIMO channel using the AR model based predictor [ZG04, LGLG04, TO05, Aro11]. In [Aro11], measured SISO subchannels of a MIMO channel are jointly exploited to estimate the AR model parameters, whereas in [ZG04, LGLG04, TO05] the AR parameters are estimated from the temporal autocorrelation function calculated using the Jakes' model. This is a reasonable approach in cases where there is no dependency between the SISO subchannels of a MIMO channel. However, in practice the SISO subchannels of a MIMO channel are not independent [WE06a]. There is a spatial correlation between the SISO subchannels of a MIMO channel which could be exploited to improve performance of MIMO channel prediction. In [WE06a], an AR model based MIMO channel prediction algorithm is proposed which exploits both the temporal and the spatial correlations in MIMO correlated fading channels. It has been shown that the proposed MIMO channel predictor achieves a superior performance, at a cost of a slight increase in complexity owing to consideration of the spatial correlations, over the SISO channel predictor.

A study of the Cramér-Rao lower bound for MIMO channel prediction compared with the SISO channel prediction for the parametric model based prediction approach shows that independently predicting the SISO subchannels of the MIMO channel results in

a significant performance loss for all considered prediction widths [SS06]. In the parametric model based MIMO channel predictor, the angle-of-departure and the angle-of-arrival of each path are considered for describing the spatial channel property. Exploitation of the spatial channel property, in addition to the temporal channel property, accounts for the performance gain in MIMO channel prediction. Furthermore, the performance gain in MIMO channel prediction can be attributed to the increase in the number of measurements compared with the SISO case with only a modest increase in the number of unknowns, i.e., the angle-of-departure and the angle-of-arrival of each path. Similar results are obtained for an extension of the study to MIMO-OFDM systems in [LSS09]. The interesting result is that a MIMO-OFDM system can operate with a lower time and frequency pilot density than a corresponding SISO implementation, even when considering the fact that more pilot symbols are required for estimating the MIMO channel.

Parametric model based MIMO channel prediction algorithms have been proposed in [GS04, ONJ+08, Pal11, ATD14a, ATD14b]. The MIMO channel prediction in [ONJ+08, Pal11, ATD14a, ATD14b] is performed by merely extrapolating the computed MIMO channel transfer function at the desired prediction width. In [GS04], a short-term MIMO channel prediction, without estimating the MIMO channel path parameters, is proposed by separating the spatial and temporal components of the MIMO channel variation using the singular value decomposition (SVD). Afterwards, short-term prediction of the temporal components of the MIMO channel variations is performed using an AR model. The method of least squares is used to estimate the AR parameters from the true MIMO channel coefficients.

Parametric model based MIMO channel prediction using the double directional channel model has been studied in detail in [Pal11]. Based on the investigation of the equivalence of the double directional channel model to the linear filter based channel model, it has been claimed that a linear and time and frequency independent prediction filter, with a filter dimension at least equal to the number of paths in the MIMO channel, can be used to predict any channel coefficient of the MIMO channel. A similar result, yet unfurnished with the detailed analysis presented in [Pal11], has been claimed in [GS04]. In this thesis, the result in [GS04, Pal11] is proved and extended. Furthermore, the following open problems in MIMO channel prediction are addressed:

- How to exploit the spatio-temporal channel properties of MIMO systems to track MIMO channels for performance gains in MIMO channel prediction and
- How to exploit the spatio-temporal channel properties of MIMO systems to exploit side information for MIMO channel prediction.

## 1.3 Structure of the thesis

In this section, the organization of the remainder of this thesis is given.

In Chapter 2, a MIMO channel model for studying the spatio-temporal channel properties of MIMO systems is presented. Furthermore, a simplified channel model for localization of a MS in MIMO systems is also presented.

In Chapter 3, localization of a MS in NLOS multipath environments under the explicit consideration of scatterers is presented. The case where the MS is unsynchronized with the BSs is considered. It will be shown how the spatio-temporal channel properties of MIMO systems can be exploited to yield satisfactory performance in NLOS multipath environments. In Chapter 4, performance gains from tracking a MS in NLOS multipath environments under the explicit consideration of the scatterers are presented. Owing to the nonlinearity of the tracking problem, a suboptimal approach to tracking a MS is presented. In Chapter 5, the problem of network localization in NLOS multipath environments is studied. A cooperative localization algorithm for localization of MSs in NLOS multipath environments under the explicit consideration of scatterers is presented. Furthermore, using techniques from robust estimation theory, a robust cooperative localization algorithm for network localization in NLOS multipath environments without the explicit consideration of scatterers is presented.

In Chapter 6, prediction of the MIMO channel is presented. Furthermore, the time and frequency dynamics of the proposed MIMO channel model are studied. In this chapter, it will be proved that a linear and time and frequency independent prediction filter, with a filter dimension at least equal to the number of propagation paths, can be used to predict any channel coefficient of the MIMO channel. In Chapter 7, performance gains from tracking the MIMO channel are presented. It will be shown that tracking the MIMO channel in TDD systems can be used to improve the receiver side channel estimation and also channel prediction. In Chapter 8, exploiting side information for improving the MIMO channel prediction performance is presented. As a demonstrative example, exploiting the feedback channel in FDD systems for improving MIMO channel prediction performance is presented.

Finally, a summary of this thesis is given in Chapter 9.

# Chapter 2

# Modeling the spatio-temporal channel properties of MIMO systems

## 2.1 Double directional channel model

In this chapter, a MIMO channel model is presented which will be used to assess the potential gains from exploitation of the spatio-temporal channel properties of MIMO systems. The simplest MIMO channel model is a statistical narrowband MIMO channel model where the entries of the MIMO channel matrix are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables. This MIMO channel model is often used to model rich scattering environments characterized by independent multipath propagations uniformly distributed in all directions [ABB+07]. However, for practical MIMO systems the channels between the different antennas are often correlated owing to the limited scattering in the environment [YO02]. Consequently, other advanced MIMO channel models have been proposed which consider the spatial dimension of MIMO channels. A different approach to MIMO channel modeling is to explicitly consider physical parameters of the electromagnetic wave propagations to describe a MIMO propagation channel. These parameters include the path complex amplitude, the delay, the Doppler shift, the angle-of-departure and the angle-of-arrival. The complex amplitude describes the effect of path attenuation and phase shift. In this thesis, the double directional channel model which characterizes MIMO channels from a propagation point of view is considered.

Fig. 2.1 shows a MIMO system with $N$ transmit antennas and $M$ receive antennas. Communication between the transmitter and the receiver antennas of the MIMO system is accomplished through the wireless propagation medium. As mentioned in Section 1.2.1, in mobile radio systems the emitted electromagnetic (EM) waves generally do not propagate directly from the transmitter to the receiver via a line-of-sight path. Rather, the emitted EM waves interact with different objects such as windows, walls, buildings, cars, trees, mountains, etc., before reaching the receiver. Fig. 2.2 shows the result of the interaction of EM waves with different objects. Reflection of an EM wave occurs when an EM wave interacts with an object which has a smooth surface of infinite extent, whereas scattering of an EM wave occurs when an EM wave interacts with an object which has a rough surface. On the other hand, diffraction of an EM wave occurs at the edges of an interacting object. Consequently, the received signal

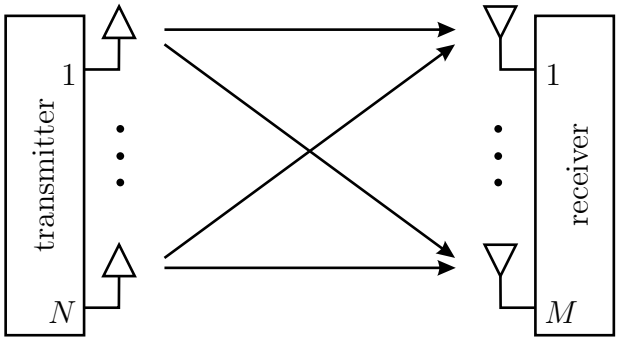



Figure 2.1. A MIMO system

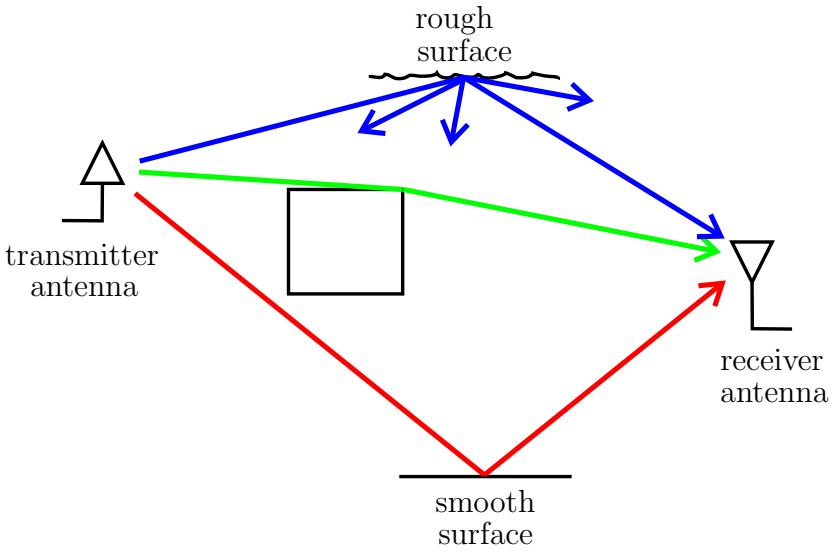



Figure 2.2. Interaction of EM waves with different objects

consists of a superposition of replicas of the transmitted signal with different complex amplitudes, delays, Doppler shifts, angles-of-departure and angles-of-arrival [Pät02]. This phenomenon is called multipath propagation. As is often done in literature all the objects with which the EM wave interacts are herein referred to as scatterers.

The multipath propagations result in the variation of the received signal power. In general, the impact of the propagation medium on the variation of the channel over distance is described using two scales: the large-scale variations and the small-scale variations [Mol05, Gol05]. The large scale variations, which are caused by path loss and shadowing, occur typically on a scale of few hundred wavelengths. Path loss is caused by the dissipation of transmit power as it radiates in space, whereas shadowing is caused by obstacles between the transmitter and the receiver. The small scale variations are caused by the interference of the multipath propagations and occur typically on a scale that is comparable with one wavelength [Mol05, Gol05]. Since the impact of

large scale variations of MIMO channels can be mitigated by power control [Mol05], the focus of this thesis is on modeling the small-scale variations of MIMO channels.

It is assumed that the base station (BS) is stationary, whereas the mobile station (MS) is moving. Without loss of generality, the downlink case where the BS is transmitting and the MS is receiving is considered in the following discussions. In general, the propagation path parameters of MIMO channels are time and frequency variant. Furthermore, the number of scatterers might change in a given time duration owing to the appearance and disappearance of scatterers. However, assuming small scale MS mobility, the geometry and the characteristics of the scatterers can be safely assumed to be constant or varying slowly for a much longer duration than the coherence time of the channel. It can also be assumed that the scatterers do not disappear and there are no new emerging scatterers. Furthermore, the scatterers can be assumed to be stationary. Consequently, the propagation path parameters are assumed to be constant within the transmission time duration and the frequency bands under consideration.

For modeling simplicity, a discrete scatterer is chosen to represent the effect of a cluster of scatterers or one large scatterer within an area. These discrete scatterers are often referred to as effective scatterers [ECS$^+$98, LR99]. Moreover, the propagation paths with gains being so small such that the received signal power from the paths is smaller than the noise power can be safely neglected. Consequently, only a finite number of discrete paths need to be considered. Assuming that the scatterers are located in the far-field of the antennas, a finite number of plane waves each of which is described by a complex amplitude, delay, Doppler shift, angle-of-departure and angle-of-arrival superpose at the receiver antennas. If the dimensions of the antenna arrays at the transmitter and the receiver are small enough, then it can be assumed that the complex amplitudes, Doppler shifts, angles-of-departure and angles-of-arrival of the plane waves do not change over the size of the array [Mol04]. Thus the MIMO channel can be characterized by the propagation medium between two reference points (RP) at the transmitter and the receiver. The transmitter and the receiver RPs can be set at the center of the antenna arrays at the transmitter and the receiver, respectively.

The aforementioned discussions result in the double directional channel model which is a generic physical MIMO channel model which characterizes the spatio-temporal channel properties of MIMO systems using the complex amplitudes, delays, Doppler shifts, angles-of-departure and angles-of-arrival of the propagation paths [Ste01, SMB01, Mol04]. Fig. 2.3 shows the double directional channel model of MIMO systems. Each propagation path between the transmitter RP and the receiver RP is characterized by a complex amplitude $\alpha_p$, delay $\tau_p$, Doppler shift $\nu_p$, azimuth angle-of-departure $\phi_{\mathrm{bs},p}$ and azimuth angle-of-arrival $\phi_{\mathrm{ms},p}, p = 1, \ldots, P$. In the case of a three-dimensional

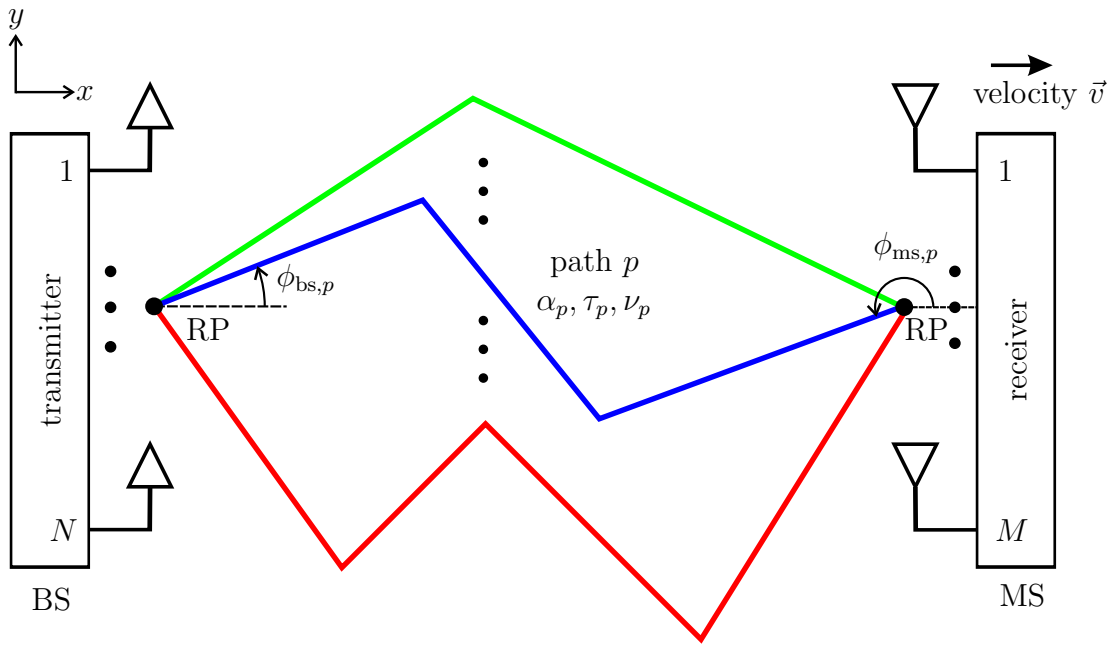



Figure 2.3. The two-dimensional double directional channel model

propagation path the solid angles $\Omega_{\mathrm{bs},p}$ and $\Omega_{\mathrm{ms},p}$ can be used to describe the angle-of-departure and the angle-of-arrival of the $p^{\mathrm{th}}$ path, respectively. The solid angle $\Omega$ is a two-dimensional angle in a three-dimensional space representing the azimuth angle $\phi$ and the elevation angle $\theta$. The double directional channel transfer function is thus defined as

$$H(f,t,\Omega_{\mathrm{bs}},\Omega_{\mathrm{ms}}) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_p} \mathrm{e}^{-\mathrm{j}2\pi\nu_p t} \delta(\Omega_{\mathrm{bs}} - \Omega_{\mathrm{bs},p}) \delta(\Omega_{\mathrm{ms}} - \Omega_{\mathrm{ms},p}), \qquad (2.1)$$

where $f_0$ is the carrier frequency, $f \in [-\frac{B}{2}, \frac{B}{2}]$ and $B$ is the bandwidth of the MIMO channel.

The channel transfer function of the channel between the $n^{\mathrm{th}}, n = 1, \ldots, N$ transmitter antenna and the $m^{\mathrm{th}}, m = 1, \ldots, M$ receiver antenna is computed by considering the constellation of the antenna arrays. Let the vectors $\mathbf{r}_{\mathrm{Tx},n} = (x_{\mathrm{Tx},n}, y_{\mathrm{Tx},n}, z_{\mathrm{Tx},n})^{\mathrm{T}}$ and $\mathbf{r}_{\mathrm{Rx},m} = (x_{\mathrm{Rx},m}, y_{\mathrm{Rx},m}, z_{\mathrm{Rx},m})^{\mathrm{T}}$ denote the position vectors of the $n^{\mathrm{th}}$ transmitter antenna and the $m^{\mathrm{th}}$ receiver antenna measured from their respective antenna array RPs, respectively. Furthermore, let the vectors $\mathbf{k}_{\mathrm{Tx},p}$ and $\mathbf{k}_{\mathrm{Rx},p}$ denote the wave number vectors of the $p^{\mathrm{th}}$ path at the transmitter and receiver, respectively. The wave number vectors of the $p^{\mathrm{th}}$ path at the transmitter and the receiver are defined as

$$\mathbf{k}_{\mathrm{Tx},p} = \frac{2\pi}{\lambda} \left( \sin(\theta_{\mathrm{bs},p}) \cos(\phi_{\mathrm{bs},p}), \quad \sin(\theta_{\mathrm{bs},p}) \sin(\phi_{\mathrm{bs},p}), \quad \cos(\theta_{\mathrm{bs},p}) \right)^{\mathrm{T}} \qquad (2.2a)$$

$$\mathbf{k}_{\mathrm{Rx},p} = \frac{2\pi}{\lambda} \left( \sin(\theta_{\mathrm{ms},p}) \cos(\phi_{\mathrm{ms},p}), \quad \sin(\theta_{\mathrm{ms},p}) \sin(\phi_{\mathrm{ms},p}), \quad \cos(\theta_{\mathrm{ms},p}) \right)^{\mathrm{T}}. \qquad (2.2b)$$

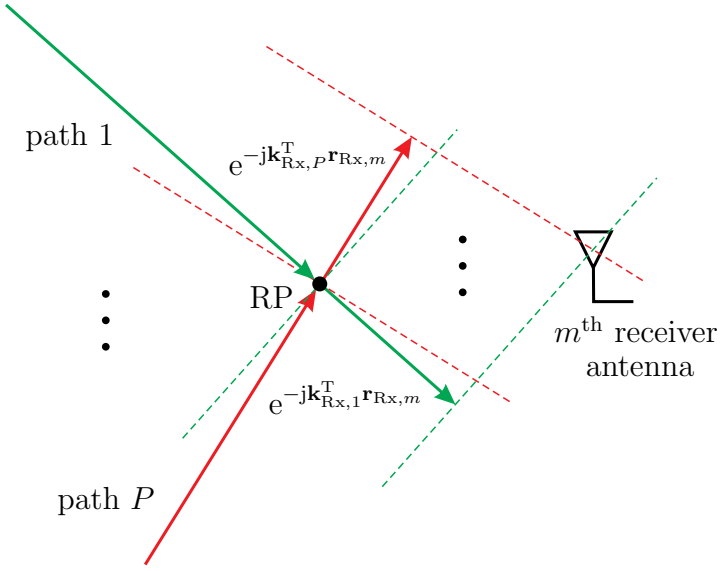



Figure 2.4. Phase shifts between the receiver antenna array RP and the $m^{\text{th}}$ receiver antenna [WABD06]

Assuming a unit antenna gain in all directions, i.e., an omnidirectional antenna, the channel transfer function of the channel between the $n^{\text{th}}$ transmitter antenna and the $m^{\text{th}}$ receiver antenna is then calculated as

$$H_{m,n}(f,t) = \sum_{p=1}^{P} \alpha_p e^{-j2\pi(f+f_0)\tau_p} e^{j2\pi\nu_p t} e^{-j\mathbf{k}_{Tx,p}^T\mathbf{r}_{Tx,n}} e^{-j\mathbf{k}_{Rx,p}^T\mathbf{r}_{Rx,m}}. \tag{2.3}$$

The phase shift $\mathbf{k}_{Tx,p}^T\mathbf{r}_{Tx,n}$ of the $p^{\text{th}}$ path at the $n^{\text{th}}$ transmitter antenna in (2.3) represents the phase shift owing to the path delay from the $n^{\text{th}}$ transmitter antenna to the transmitter antenna array RP. Likewise the phase shift $\mathbf{k}_{Rx,p}^T\mathbf{r}_{Rx,m}$ of the $p^{\text{th}}$ path at the $m^{\text{th}}$ receiver antenna in (2.3) represents the phase shift owing to the path delay from the $m^{\text{th}}$ receiver antenna to the receiver antenna array RP. Fig. 2.4 shows the phase shifts of the plane waves between the $m^{\text{th}}$ receiver antenna and the receiver antenna array RP.

The phase shift $\mathbf{k}_{Tx,p}^T\mathbf{r}_{Tx,n}$ can be simplified as

$$\begin{aligned}
\mathbf{k}_{Tx,p}^T\mathbf{r}_{Tx,n} &= \frac{2\pi}{\lambda}(x_{Tx,n}\sin(\theta_{bs,p})\cos(\phi_{bs,p}) + y_{Tx,n}\sin(\theta_{bs,p})\sin(\phi_{bs,p}) + z_{Tx,n}\cos(\theta_{bs,p})) \\
&= 2\pi(f+f_0) \\
&\quad \cdot \frac{x_{Tx,n}\sin(\theta_{bs,p})\cos(\phi_{bs,p}) + y_{Tx,n}\sin(\theta_{bs,p})\sin(\phi_{bs,p}) + z_{Tx,n}\cos(\theta_{bs,p})}{c_0} \\
&= 2\pi(f+f_0)\tau_{Tx}^{(n,p)},
\end{aligned} \tag{2.4}$$

where $c_0$ is the speed of light and $\tau_{Tx}^{(n,p)}$ is the delay of the $p^{\text{th}}$ propagation path from the transmitter antenna array RP to the $n^{\text{th}}$ transmitter antenna. Likewise the phase

shift $\mathbf{k}_{\mathrm{Rx},p}^{\mathrm{T}}\mathbf{r}_{\mathrm{Rx},m}$ can be simplified as

$$\mathbf{k}_{\mathrm{Rx},p}^{\mathrm{T}}\mathbf{r}_{\mathrm{Rx},m} = 2\pi(f+f_0)\tau_{\mathrm{Rx}}^{(m,p)}, \tag{2.5}$$

where $\tau_{\mathrm{Rx}}^{(m,p)}$ is the delay of the $p^{\mathrm{th}}$ propagation path from the receiver antenna array RP to the $m^{\mathrm{th}}$ receiver antenna. Using (2.4) and (2.5), (2.3) can be simplified as

$$H_{m,n}(f,t) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p t} \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_{\mathrm{Rx}}^{(m,p)}}. \tag{2.6}$$

In array signal processing, it is commonly assumed that the bandwidth $B$ of the signal is small as compared to the center frequency $f_0$ and the antenna array dimensions are small enough so that $\left| f \cdot \tau_{\mathrm{Tx}}^{(n,p)} \right|$ and $\left| f \cdot \tau_{\mathrm{Rx}}^{(m,p)} \right|$ are close to zero. This assumption is called the narrowband assumption for array signal processing [God97]. The narrowband assumption yields

$$\mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_{\mathrm{Tx}}^{(n,p)}} \approx \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \tag{2.7a}$$

$$\mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_{\mathrm{Rx}}^{(m,p)}} \approx \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}. \tag{2.7b}$$

The variables $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}}$ and $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}$ are referred to as steering factors in array signal processing. Using the result of the narrowband assumption in (2.7) to simplify (2.6) results in

$$H_{m,n}(f,t) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p t} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}. \tag{2.8}$$

The double directional channel model presented here is rather generic. Adaptation of the double directional channel model to different scenarios can be performed by determining the typical values of the propagation path parameters as discussed in [ECS+98, YO02]. Furthermore, large-scale variations of MIMO channels can be modeled using the double directional channel model by allowing for the path parameters to change over the range of movement of the MS [Mol04]. In this case, the number of propagation paths is time-variant, i.e., $P(t)$.

It must be noted that the double directional channel is reciprocal, i.e., exchanging the positions of the transmitter and the receiver results in the same path parameters with the angle-of-arrival now being the angle-of-departure and vice versa [Ste01, Mol04]. Furthermore, the double directional channel model models not only the multipath propagations but also the line-of-sight path. The line-of-sight path is implicitly modeled in the double directional channel model. In the line-of-sight path case, virtual scatterers can be assumed to be located on the line joining the RPs at the transmitter and the receiver and the emitted EM wave propagates through the virtual scatterers without attenuation or phase shift.

## 2.2 Single bounce scattering model

The double directional channel model presented in the previous section is comprehensive and can be used to study channel prediction in MIMO systems. However, the comprehensive feature of the double directional channel model, which is essential for channel prediction, is rather too detailed to systematically exploit the spatio-temporal channel properties of MIMO systems for localization. Thus in this section, the single bounce scattering model is introduced which contains a level of detail that is sufficient for localization in MIMO systems.

The double directional channel model considers, in addition to the line-of-sight path, multipath propagations which are possibly scattered by several scatterers before arriving at the receiver. In the single bounce scattering model, it is assumed that the first few arriving signals propagate from the transmitter to the receiver after being scattered by a maximum of one scatterer. The multiple bounce signals have low power as compared to the single bounce signals owing to the severe attenuations and scattering loss as they bounce off several scatterers. Furthermore, multiple bounce signals often have longer delays which results in a signal with low power due to path attenuation. Thus the multiple bounce signals can be safely neglected by considering only the first few arriving signals. In practice, the two step proximity detection algorithm presented in [ST08] can be applied to detect and discard multiple bounce scattering propagations.

Fig. 2.5 shows the single bounce scattering model where the $p^{\text{th}}$ propagation path is the path from the BS to the MS via the $p^{\text{th}}$ scatterer. The propagation path parameters of the single bounce scattering model can be described as a function of the position of the BS $\mathbf{p}_{\text{bs}} = (x_{\text{bs}}, y_{\text{bs}}, z_{\text{bs}})^{\text{T}}$, the position of the $p^{\text{th}}$ scatterer $\mathbf{p}_{\text{sc},p} = (x_{\text{sc},p}, y_{\text{sc},p}, z_{\text{sc},p})^{\text{T}}$, the position of the MS $\mathbf{p}_{\text{ms}} = (x_{\text{ms}}, y_{\text{ms}}, z_{\text{ms}})^{\text{T}}$ and the velocity of the MS $\mathbf{v}_{\text{ms}} = (v_{\text{ms},x}, v_{\text{ms},y}, v_{\text{ms},z})^{\text{T}}$ as

$$\begin{aligned} d_p &= \tau_p \cdot c \\ &= \|\mathbf{p}_{\text{bs}} - \mathbf{p}_{\text{sc},p}\| + \|\mathbf{p}_{\text{sc},p} - \mathbf{p}_{\text{ms}}\| && \text{(2.9a)} \\ \nu_p &= \frac{f_0}{c} \frac{\mathbf{v}_{\text{ms}}^{\text{T}}(\mathbf{p}_{\text{sc},p} - \mathbf{p}_{\text{ms}})}{\|\mathbf{p}_{\text{sc},p} - \mathbf{p}_{\text{ms}}\|} && \text{(2.9b)} \\ \phi_{\text{bs},p} &= \frac{\pi}{2}(1 - \text{sgn}(x_{\text{sc},p} - x_{\text{bs}})) + \tan^{-1}\left(\frac{y_{\text{sc},p} - y_{\text{bs}}}{x_{\text{sc},p} - x_{\text{bs}}}\right) && \text{(2.9c)} \\ \phi_{\text{ms},p} &= \frac{\pi}{2}(1 - \text{sgn}(x_{\text{sc},p} - x_{\text{ms}})) + \tan^{-1}\left(\frac{y_{\text{sc},p} - y_{\text{ms}}}{x_{\text{sc},p} - x_{\text{ms}}}\right) && \text{(2.9d)} \\ \theta_{\text{bs},p} &= \frac{\pi}{2} - \tan^{-1}\left(\frac{z_{\text{sc},p} - z_{\text{bs}}}{\sqrt{(x_{\text{sc},p} - x_{\text{bs}})^2 + (y_{\text{sc},p} - y_{\text{bs}})^2}}\right) && \text{(2.9e)} \end{aligned}$$

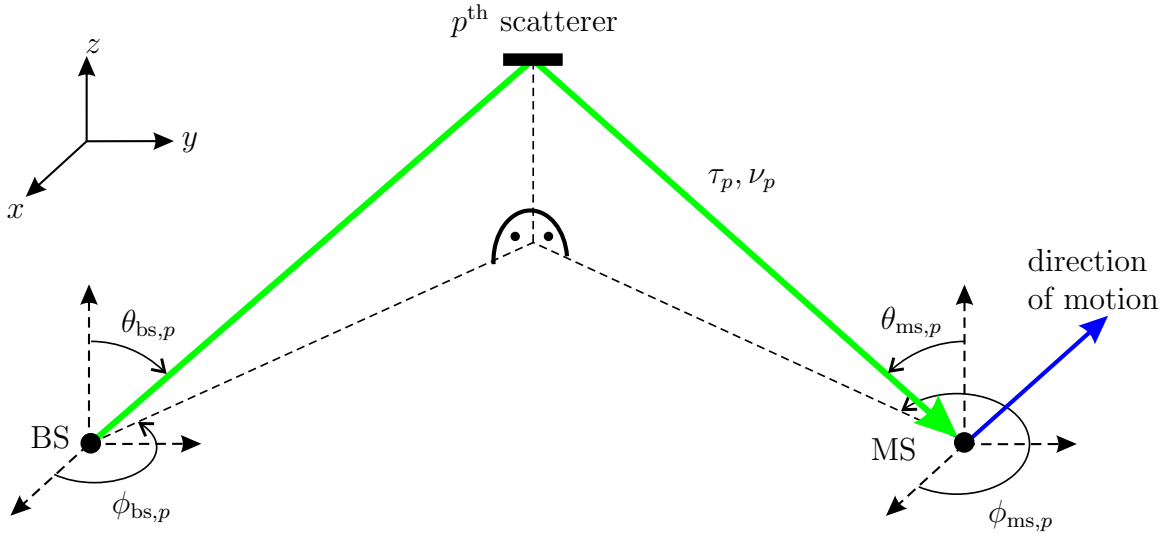



Figure 2.5. The single bounce scattering model

$$\theta_{\text{ms},p} = \frac{\pi}{2} - \tan^{-1}\left(\frac{z_{\text{sc},p} - z_{\text{ms}}}{\sqrt{(x_{\text{sc},p} - x_{\text{ms}})^2 + (y_{\text{sc},p} - y_{\text{ms}})^2}}\right), \tag{2.9f}$$

where $d_p$ is the length of the $p^{\text{th}}$ propagation path.

The propagation path parameters can be estimated using antenna array processing techniques. Owing to erroneous channel estimates, modeling imperfections, receiver noise, etc., the estimated path parameters are noisy versions of the true path parameters. Several measurement campaigns have been performed to model the estimated path parameters. For instance, the Gaussian distribution has been proposed for the angle-of-arrival statistics in rural and suburban environments [ECS+98], whereas the Laplacian distribution has been proposed for the angle-of-departure and angle-of-arrival statistics in indoor environments [YO02]. In this thesis, for the sake of clearer exposition and analysis, it is assumed that the estimated propagation path parameters, $\hat{d}_p$, $\hat{\nu}_p$, $\hat{\phi}_{\text{bs},p}$, $\hat{\phi}_{\text{ms},p}$, $\hat{\theta}_{\text{bs},p}$, and $\hat{\theta}_{\text{ms},p}$ , are i.i.d. Gaussian random variables, i.e.,

$$\hat{d}_p \sim \mathcal{N}(d_p, \sigma_{\text{d}}^2) \tag{2.10a}$$
$$\hat{\nu}_p \sim \mathcal{N}(\nu_p, \sigma_{\nu}^2) \tag{2.10b}$$
$$\hat{\phi}_{\text{bs},p} \sim \mathcal{N}(\phi_{\text{bs},p}, \sigma_{\phi_{\text{bs}}}^2) \tag{2.10c}$$
$$\hat{\phi}_{\text{ms},p} \sim \mathcal{N}(\phi_{\text{ms},p}, \sigma_{\phi_{\text{ms}}}^2) \tag{2.10d}$$
$$\hat{\theta}_{\text{bs},p} \sim \mathcal{N}(\theta_{\text{bs},p}, \sigma_{\theta_{\text{bs}}}^2) \tag{2.10e}$$
$$\hat{\theta}_{\text{ms},p} \sim \mathcal{N}(\theta_{\text{ms},p}, \sigma_{\theta_{\text{ms}}}^2). \tag{2.10f}$$

# Chapter 3

# Localization in NLOS multipath environments under the explicit consideration of scatterers

## 3.1 Localization with incomplete measurement data

In this chapter, localization in NLOS multipath environments under the explicit consideration of scatterers is considered. The explicit consideration of scatterers is enabled by the single bounce scattering model. The propagation path parameters, i.e., the complex amplitudes, the delays, the Doppler shifts, the angles-of-departure and the angles-of-arrival can be exploited using the single bounce scattering model to localize a MS in NLOS multipath environments. However, in this thesis, only the delays, the angles-of-departure and the angles-of-arrival are exploited for localization of a MS in NLOS multipath environments under the explicit consideration of scatterers. As mentioned in Section 1.2.1, information about the propagation path length can be estimated from the propagation path's complex amplitude using theoretical or empirical path loss formula. The path loss formula have to include, in addition to the free space path loss, the loss caused by scatterers as they scatter some of the transmitted signal energy into all directions. Developing a theoretical path loss formula which accounts for the scattered energy is rather a difficult task. However, empirical path loss formula which account for the scattered energy, in addition to the free space path loss, can be obtained from measurement campaigns. However, since such measurement campaigns are site specific, they have not been considered in this thesis for the sake of maintaining the generality of the ideas presented. Furthermore, performing the measurement campaigns is an arduous task and requires a concerted effort of specialists in measurement techniques. Thus the complex amplitudes of the propagation paths have not been explicitly exploited for localization in this thesis. Nevertheless, the complex amplitudes can be used as an aid in selecting single bounce propagation paths based on the assumption that the single bounce propagation paths are less attenuated than the multiple bounce propagation paths. In addition to the propagation paths complex amplitudes, the Doppler shifts of the propagation paths have also not been explicitly exploited for localization in this thesis. The Doppler shifts of the propagation paths are mainly exploited in tracking a MS as shown in [PS09, SW14a].

Localization in NLOS multipath environments under the explicit consideration of scatterers with complete measurement data has already been considered in [MYJ07, ST08, WPW11]. The case with complete measurement data requires that the delay, the angle-of-departure and the angle-of-arrival are estimated for synchronized clocks at the BSs and the MS with the orientation of the BSs and the MS assumed to be known. However, in real world localization scenarios, the aforementioned assumptions might not be fulfilled, i.e., the BSs and the MS might be unsynchronized and/or the orientation of the BSs and the MS might be unknown. If there are unsynchronized BSs and MS, then the time-differences-of-arrival of the propagation paths can be measured rather than the times-of-arrival of the propagation paths between the BSs and the MS. Moreover, if the orientations of the BSs and the MS are unknown, then only the relative angles-of-departure and angles-of-arrival of the propagation paths can be estimated. In some cases, a MS may have only a single antenna and the angles-of-arrival of the propagation paths cannot be estimated. These cases are herein referred to as cases with incomplete measurement data. Localization of a MS with incomplete measurement data is obviously trickier than localization with complete measurement data.

The case where the orientations of the BSs and the MS are unknown is actually of little practical importance as it can be easily solved by employing a compass at both the BSs and the MS to establish a reference that can be used to measure the absolute angles-of-departure and angles-of-arrival. Furthermore, for a MS with a single antenna, it has been shown that the Doppler shifts of the propagation paths can be used instead of the angles-of-arrival of the propagation paths with only a minor performance lose, i.e., the Doppler shifts of the propagation paths which can easily be measured at a MS with a single antenna contain comparable information about the position of a MS as the angles-of-arrival of the propagation paths [SW13]. In [PS09, SW14a], algorithms for estimating the position of a MS based on the delays, the angles-of-departure and the Doppler shifts of the propagation paths have been proposed. In this thesis, the most crucial case where there are unsynchronized BSs and MS is considered in detail. Since the MS's clock is unsynchronized with the BSs', the times-of-arrival of the propagation paths and hence the path lengths of the propagation paths cannot be estimated. Nevertheless, since the BSs are assumed to have synchronized clocks, it is possible to measure the time-differences-of-arrival of the propagation paths between the BSs and the MS. If there is only a single BS, then the time-differences-of-arrival of the propagation paths between the BS and the MS can be estimated. The path length differences between the propagation paths can be estimated from the time-differences-of-arrival of the propagation paths.

In this thesis, the estimated time-differences-of-arrival of the propagation paths are assumed to be obtained from the differences of biased estimates of the times-of-arrival

with a reference biased estimate of time-of-arrival. As it can be seen from (2.10), the estimated times-of-arrival are assumed to be i.i.d. Gaussian random variables. Thus the estimated time-differences-of-arrival form a multivariate Gaussian random variable with a non-diagonal covariance matrix [QKS06]. The covariance matrix is non-diagonal owing the correlation stemming from the reference biased estimate of time-of-arrival. Nevertheless, in the following, a diagonal covariance matrix is assumed for the estimated time-differences-of-arrival without loss of generality.

The single bounce scattering model presented in Section 2.2 considers the three dimensional propagation paths from the transmitter to the receiver. The necessity of consideration of a two dimensional or a three dimensional localization of a MS depends on the scenario under consideration. Owing to the relative dimensions of the localization scenarios, three dimensional localization is required in indoor localization scenarios, whereas two dimensional localization is sufficient for most outdoor localization scenarios. The localization principle and the localization algorithms presented in this thesis can be applied to both the two dimensional and the three dimensional localization scenarios without extra BSs or any other additional requirements. Thus for the sake of clearer exposition, the two dimensional localization is considered in this thesis. An extension of the presented two dimensional localization ideas to the three dimensional localization is straightforward.

## 3.2 Localization principle

In this section, the localization principle for localization of a MS in NLOS multipath environments with unsynchronized BSs and MS is presented. The localization of a MS is carried out under the explicit consideration of scatterers whose positions are also unknown. Hence the positions of the scatterers are implicitly estimated along with the position of the MS. As is often the case, the localization principle is described pictorially. Mathematical formulation of the localization principle is given in the next section. Furthermore, the true values of the propagation path parameters are considered here. Mitigation of the estimation noise of the propagation path parameters is considered in the next section.

Fig. 3.1 shows how the path length differences, the angles-of-departure and the angles-of-arrival of two propagation paths from two BSs to a MS can be exploited to obtain the possible positions of the MS in a NLOS multipath environment. Let's consider the propagation path parameters corresponding to the propagation path from the first BS to the MS via the first scatterer. From the angle-of-departure of the propagation path

from the first BS to the MS, it can be concluded that the first scatterer is located on a line extending from the first BS at an angle equal to the angle-of-departure. Since the path length of the propagation path from the first BS to the MS is unknown, the length of the line extending from the first BS, on which the first scatterer is possibly located, cannot be determined. Nevertheless, let's consider a set of possible path lengths of the propagation path from the first BS to the MS. These sets of possible path lengths of the propagation path from the first BS to the MS yield sets of line segments on which the first scatterer is possibly located. From each of these line segments on which the first scatterer is possibly located, a curve of the possible positions of the MS can be determined by considering the angle-of-arrival of the first propagation path. For each possible position of the first scatterer, the corresponding possible position of the MS is determined by drawing a line segment from the first scatterer at an angle which corresponds with the angle-of-arrival of the first propagation path and with a distance such that the assumed propagation path length is maintained. Performing this operation results in a line segment on which the MS is possibly located as shown in Fig. 3.1. For different possible path lengths of the propagation path from the first BS to the MS, parallel line segments of possible positions of the MS can be obtained as shown in Fig. 3.1.

The different possible path lengths of the propagation paths from the first BS to the MS and from the second BS to the MS are constrained by the path length differences, i.e., for a given path length of the propagation path from the first BS to the MS, there is a corresponding possible path length of the propagation path from the second BS to the MS defined by the path length difference of the propagation paths from the two BSs to the MS. Thus considering the path length difference of the propagation paths from the BSs to the MS, the line segment of possible MS positions from the first BS intersects with the corresponding line segment of possible MS positions from the second BS. Fig. 3.1 shows the intersection points of three line segments from the first and the second BSs. The intersection of all line segments of possible MS positions from the first and the second BSs results in a line of possible MS positions. Another line of possible MS positions can be obtained by considering a third BS. The intersection point of the two lines of possible MS positions yields the unknown position of the MS.

It must be noted that the LOS localization is a special case of the NLOS localization depicted in Fig. 3.1. If there are LOS propagation paths between the BSs and the MS, then the line in Fig. 3.1 denoting the possible MS positions collapses to a point denoting the position of the MS. The LOS localization scenario is actually a LOS hybrid localization where the bearing and the path length differences of the LOS propagation paths are used to estimate the position of the MS. However, the bearings from the BSs to the MS are adequate for localizing the MS.

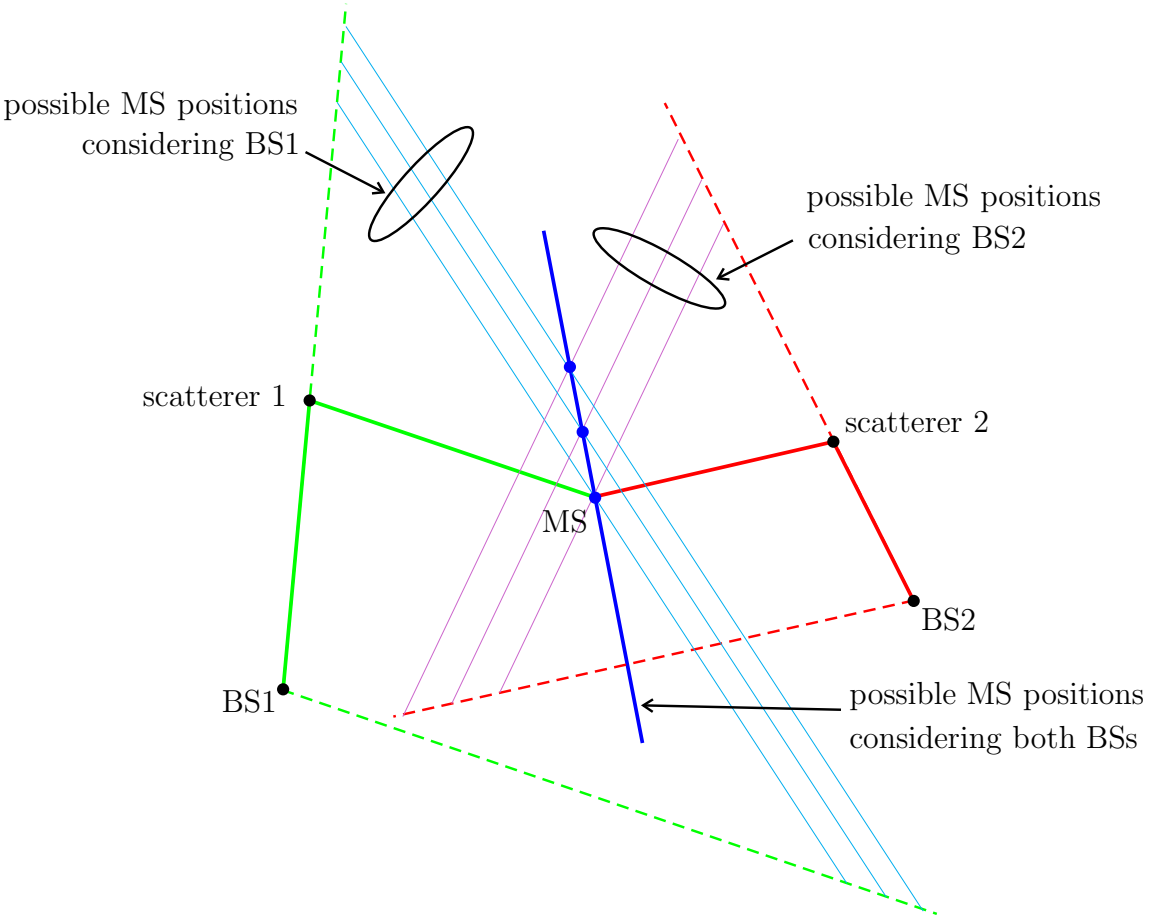



Figure 3.1. Possible positions of the MS using path parameter estimates from unsynchronized BSs and MS

In the preceding discussion, three propagation paths from three BSs are considered only for the sake of simpler exposition. Localization of the MS is also possible if three propagation paths are considered from a single BS. Consideration of multiple BSs and more than three propagation paths from a single BS or multiple BSs can be used to improve performance of the MS localization accuracy. Furthermore, an extension of the presented localization principle to a three dimensional localization scenario is straightforward. The required number of BSs or number of propagation paths remain the same in the three dimensional localization scenario.

## 3.3 Localization algorithm

In this section, the mathematical formulation of the localization principle discussed in the previous section is presented. In the previous section, it has been shown that

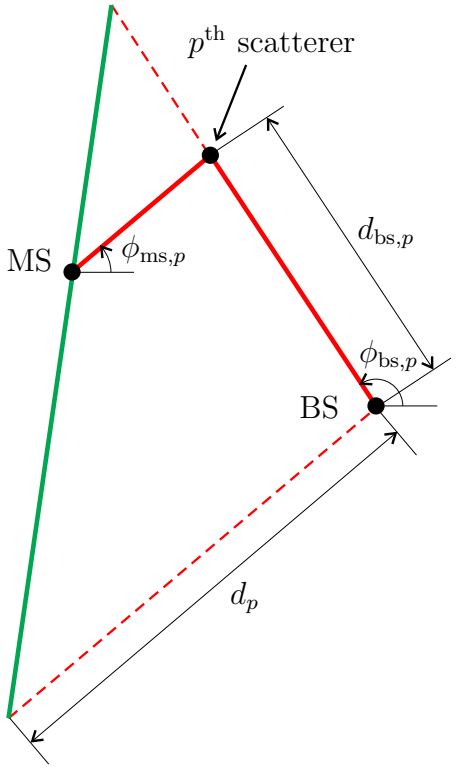



Figure 3.2. The two dimensional single bounce scattering model

there is a linear relation between the positions of the BS, the scatterer and the MS. This linear relation will be used to derive the position of the MS as a function of the position of the BS and the path parameters.

In the following, a MS which is unsynchronized with a BS is considered for simplicity. Fig. 3.2 shows the two dimensional single bounce scattering model. From Fig. 3.2 the linear relation between the position of the MS and the position of the $p^{\text{th}}$ scatterer can be determined as

$$x_{\text{ms}} = x_{\text{sc},p} - (d_p - d_{\text{bs},p})\cos(\phi_{\text{ms},p}) \tag{3.1}$$

$$y_{\text{ms}} = y_{\text{sc},p} - (d_p - d_{\text{bs},p})\sin(\phi_{\text{ms},p}), \tag{3.2}$$

where $d_{\text{bs},p}$ is the distance between the BS and the $p^{\text{th}}$ scatterer. As the BS and the MS are unsynchronized, the path length $d_p$ is not available. Rather the propagation path length difference $d_{p,q}$ between the $p^{\text{th}}$ and the $q^{\text{th}}$ paths is available, i.e.,

$$d_{p,q} = d_p - d_q. \tag{3.3}$$

Using (3.3) to find an expression for $d_p$ and substituting it in (3.1) and (3.2) results in

$$x_{\text{ms}} = x_{\text{sc},p} - (d_q + d_{p,q} - d_{\text{bs},p})\cos(\phi_{\text{ms},p}) \tag{3.4}$$

$$y_{\mathrm{ms}} = y_{\mathrm{sc},p} - (d_q + d_{p,q} - d_{\mathrm{bs},p})\sin(\phi_{\mathrm{ms},p}). \tag{3.5}$$

Furthermore, the linear relation between the position of the BS and the position of the $p^{\mathrm{th}}$ scatterer can be determined as

$$x_{\mathrm{sc},p} = x_{\mathrm{bs}} + d_{\mathrm{bs},p}\cos(\phi_{\mathrm{bs},p}), \tag{3.6}$$
$$y_{\mathrm{sc},p} = y_{\mathrm{bs}} + d_{\mathrm{bs},p}\sin(\phi_{\mathrm{bs},p}). \tag{3.7}$$

Using (3.6) and (3.7) to substitute the corresponding scatterer position in (3.4) and (3.5) results in

$$x_{\mathrm{ms}} = x_{\mathrm{bs}} + d_{\mathrm{bs},p}(\cos(\phi_{\mathrm{bs},p}) + \cos(\phi_{\mathrm{ms},p})) - (d_q + d_{p,q})\cos(\phi_{\mathrm{ms},p}), \tag{3.8}$$
$$y_{\mathrm{ms}} = y_{\mathrm{bs}} + d_{\mathrm{bs},p}(\sin(\phi_{\mathrm{bs},p}) + \sin(\phi_{\mathrm{ms},p})) - (d_q + d_{p,q})\sin(\phi_{\mathrm{ms},p}). \tag{3.9}$$

Finding the expression for $d_{\mathrm{bs},p}$ from (3.8) results in

$$d_{\mathrm{bs},p} = \frac{x_{\mathrm{ms}} - x_{\mathrm{bs}} + (d_q + d_{p,q})\cos(\phi_{\mathrm{ms},p})}{(\cos(\phi_{\mathrm{bs},p}) + \cos(\phi_{\mathrm{ms},p}))}. \tag{3.10}$$

Using (3.10) to substitute $d_{\mathrm{bs},p}$ in (3.9) results in

$$\begin{aligned}(d_q + d_{p,q})\sin(\phi_{\mathrm{bs},p} - \phi_{\mathrm{ms},p}) = &-(\sin(\phi_{\mathrm{bs},p}) + \sin(\phi_{\mathrm{ms},p}))(x_{\mathrm{ms}} - x_{\mathrm{bs}}) \\ &+ (\cos(\phi_{\mathrm{bs},p}) + \cos(\phi_{\mathrm{ms},p}))(y_{\mathrm{ms}} - y_{\mathrm{bs}}).\end{aligned} \tag{3.11}$$

Equation (3.11) shows that by considering a single propagation path, a linear relation between the position of the MS and the position of the BS with the path parameters as coefficients can be obtained. The path length $d_q$ of the $q^{\mathrm{th}}$ propagation path is a nuisance variable. Hence by considering the measurements from the $P$ single scattering propagation paths, a system of $P$ linear equations can be set up as follows

$$\mathbf{A}_0 \cdot \mathbf{r} = \mathbf{b}_0, \tag{3.12}$$

where

$$\mathbf{A}_0 = \begin{pmatrix} -\sin(\phi_{\mathrm{bs},1}) - \sin(\phi_{\mathrm{ms},1}) & \cos(\phi_{\mathrm{bs},1}) + \cos(\phi_{\mathrm{ms},1}) & -\sin(\phi_{\mathrm{bs},1} - \phi_{\mathrm{ms},1}) \\ \vdots & \vdots & \vdots \\ -\sin(\phi_{\mathrm{bs},P}) - \sin(\phi_{\mathrm{ms},P}) & \cos(\phi_{\mathrm{bs},P}) + \cos(\phi_{\mathrm{ms},P}) & -\sin(\phi_{\mathrm{bs},P} - \phi_{\mathrm{ms},P}) \end{pmatrix}, \tag{3.13}$$

$$\begin{aligned}\mathbf{b}_0 = &\begin{pmatrix} -\sin(\phi_{\mathrm{bs},1}) - \sin(\phi_{\mathrm{ms},1}) & \cos(\phi_{\mathrm{bs},1}) + \cos(\phi_{\mathrm{ms},1}) \\ \vdots & \vdots \\ -\sin(\phi_{\mathrm{bs},P}) - \sin(\phi_{\mathrm{ms},P}) & \cos(\phi_{\mathrm{bs},P}) + \cos(\phi_{\mathrm{ms},P}) \end{pmatrix} \begin{pmatrix} x_{\mathrm{bs}} \\ y_{\mathrm{bs}} \end{pmatrix} \\ &+ \begin{pmatrix} d_{1,q}\sin(\phi_{\mathrm{bs},1} - \phi_{\mathrm{ms},1}) \\ \vdots \\ d_{P,q}\sin(\phi_{\mathrm{bs},P} - \phi_{\mathrm{ms},P}) \end{pmatrix},\end{aligned} \tag{3.14}$$

$$\mathbf{r} = (x_{\mathrm{ms}}, y_{\mathrm{ms}}, d_q)^{\mathrm{T}}. \tag{3.15}$$

If $\mathbf{A}_0$ has full column rank, then the position of the MS can be estimated from (3.12) using the pseudo-inverse of the matrix $\mathbf{A}_0$ as

$$\hat{\mathbf{r}} = \left(\mathbf{A}_0^{\mathrm{T}} \cdot \mathbf{A}_0\right)^{-1} \cdot \mathbf{A}_0^{\mathrm{T}} \cdot \mathbf{b}_0. \tag{3.16}$$

It must be noted that if there is a LOS propagation path, then

$$|\phi_{\mathrm{bs,LOS}} - \phi_{\mathrm{ms,LOS}}| = \pi. \tag{3.17}$$

For a LOS propagation path (3.10) is not defined and hence (3.11), which is derived using (3.10), does not hold. A LOS propagation path can easily be identified using the relation of (3.17). For a LOS propagation path, a linear relation in the position of the MS can be easily obtained using the path length difference $d_{\mathrm{LOS},q}$ between the LOS path and the $q^{\mathrm{th}}$ propagation path as

$$x_{\mathrm{ms}} = x_{\mathrm{bs}} + (d_q + d_{\mathrm{LOS},q})\cos\left(\phi_{\mathrm{bs,LOS}}\right), \tag{3.18}$$

$$y_{\mathrm{ms}} = y_{\mathrm{bs}} + (d_q + d_{\mathrm{LOS},q})\sin\left(\phi_{\mathrm{bs,LOS}}\right), \tag{3.19}$$

Equations (3.18) and (3.19) can be combined with other NLOS propagation paths to obtain the position of the MS. However, as mentioned in the previous section, for the LOS propagation case the angle-of-departure and the angle-of-arrival of the LOS propagation paths from two BSs are sufficient to localize a MS.

The matrix $\mathbf{A}_0$ and the vector $\mathbf{b}_0$ are constructed using the true values of the path length differences, the angles-of-departure and the angles-of-arrival of the propagation paths. However, in practice, the estimated path length differences, angles-of-departure and angles-of-arrival are noisy versions of the true path length differences, angles-of-departure and angles-of-arrival, respectively. Thus the position of the MS is estimated based on the noisy versions of the matrix $\mathbf{A}_0$ and vector $\mathbf{b}_0$ which are denoted by $\mathbf{A}$ and $\mathbf{b}$, respectively. The system of $P$ linear equations in (3.12) is rewritten for the noisy measurements as

$$\mathbf{A} \cdot \mathbf{r} \approx \mathbf{b}. \tag{3.20}$$

An estimator which minimizes the Euclidean distance $\|\mathbf{b} - \mathbf{A} \cdot \mathbf{r}\|_2$ is required, i.e.,

$$\hat{\mathbf{r}} = \arg\min_{\mathbf{r}} \quad \|\mathbf{b} - \mathbf{A} \cdot \mathbf{r}\|_2 . \tag{3.21}$$

Given that $\mathbf{A}$ has full column rank, the least squares solution of (3.21) can be determined as

$$\hat{\mathbf{r}}_{\mathrm{LS}} = \left(\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^{\mathrm{T}} \cdot \mathbf{b}. \tag{3.22}$$

The downside of the least squares estimator in (3.22) is that it assumes that the matrix $\mathbf{A}$ is error free. The explicit formulation of the least squares problem of (3.21) is given as

$$\begin{aligned} \{\hat{\mathbf{r}}, \Delta\hat{\mathbf{b}}\} = \underset{\mathbf{r},\Delta\mathbf{b}}{\arg\min} \quad & \|\Delta\mathbf{b}\|_2 \\ \text{subject to} \quad & \mathbf{A}\cdot\mathbf{r} = \mathbf{b} + \Delta\mathbf{b}. \end{aligned} \tag{3.23}$$

Thus a correction, as little as possible in the Euclidean norm sense, is made only to the vector $\mathbf{b}$. However, since both $\mathbf{A}$ and $\mathbf{b}$ are perturbed, an appropriate estimator is the one which makes corrections to both $\mathbf{A}$ and $\mathbf{b}$. This can be achieved by using the total least squares estimator [MVH07, GVL13]. In the total least squares estimation, the MS position estimation problem is given as

$$\begin{aligned} \{\hat{\mathbf{r}}, \Delta\hat{\mathbf{A}}, \Delta\hat{\mathbf{b}}\} = \underset{\mathbf{r},\Delta\mathbf{A},\Delta\mathbf{b}}{\arg\min} \quad & \|(\Delta\mathbf{A} \,|\, \Delta\mathbf{b})\|_{\mathrm{F}} \\ \text{subject to} \quad & (\mathbf{A} + \Delta\mathbf{A})\cdot\mathbf{r} = \mathbf{b} + \Delta\mathbf{b}. \end{aligned} \tag{3.24}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix. The total least squares estimator makes as little as possible corrections in the Frobenius norm sense to both $\mathbf{A}$ and $\mathbf{b}$.

As shown in Section 3.2, in order to estimate the position of the MS, $P \geq 3$ propagation paths are required. In order to obtain reliable estimates of the MS position, it is assumed that the system of linear equations in (3.20) is overdetermined. Consequently, for the $P \times 3$ matrix $\mathbf{A}$, it is assumed, in the following, that $P \geq 4$.

The constraint of the total least squares problem in (3.24) can be written in the homogenous form as

$$\begin{aligned} (\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b}) \cdot \begin{pmatrix} \hat{\mathbf{r}} \\ -1 \end{pmatrix} &= \mathbf{0} \\ ((\mathbf{A} \,|\, \mathbf{b}) + (\Delta\mathbf{A} \,|\, \Delta\mathbf{b})) \cdot \begin{pmatrix} \hat{\mathbf{r}} \\ -1 \end{pmatrix} &= \mathbf{0}. \end{aligned} \tag{3.25}$$

The above equation has a solution if the augmented vector $(\hat{\mathbf{r}}^{\mathrm{T}}, -1)^{\mathrm{T}}$ lies in the nullspace of $((\mathbf{A} \,|\, \mathbf{b}) + (\Delta\mathbf{A} \,|\, \Delta\mathbf{b}))$. Furthermore, the solution is nontrivial if the perturbation $(\Delta\mathbf{A} \,|\, \Delta\mathbf{b})$ is such that $((\mathbf{A} \,|\, \mathbf{b}) + (\Delta\mathbf{A} \,|\, \Delta\mathbf{b}))$ is rank deficient. Thus the total least squares solution finds the matrix $(\Delta\mathbf{A} \,|\, \Delta\mathbf{b})$ with the smallest norm that makes the matrix $((\mathbf{A} \,|\, \mathbf{b}) + (\Delta\mathbf{A} \,|\, \Delta\mathbf{b}))$ rank deficient [MS00].

The solution to the total least squares problem can be obtained by using the singular value decomposition (SVD) [GVL13]. The SVD of $(\mathbf{A} \,|\, \mathbf{b})$ is a decomposition

$$(\mathbf{A} \,|\, \mathbf{b}) = \mathbf{U}\cdot\mathbf{\Sigma}\cdot\mathbf{V}^{\mathrm{T}}, \tag{3.26}$$

where $\mathbf{U}$ and $\mathbf{V}$ are matrices with orthonormal columns and $\mathbf{\Sigma} = \mathrm{diag}\,(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ with $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4$. The SVD of $(\mathbf{A} \,|\, \mathbf{b})$ can be partitioned as

$$(\mathbf{A} \,|\, \mathbf{b}) = \begin{pmatrix} \mathbf{U}_\mathrm{A} & \mathbf{u}_\mathrm{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{\Sigma}_\mathrm{A} & \mathbf{0} \\ \mathbf{0} & \sigma_\mathrm{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{V}_\mathrm{AA} & \mathbf{v}_\mathrm{Ab} \\ \mathbf{v}_\mathrm{bA} & v_\mathrm{bb} \end{pmatrix}^\mathrm{T}, \tag{3.27}$$

where the subscripts A and b refer to partitions corresponding to $\mathbf{A}$ and $\mathbf{b}$, respectively. The reduced-rank matrix $(\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b})$ closest to $(\mathbf{A} \,|\, \mathbf{b})$ is [MS00]

$$(\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b}) = \begin{pmatrix} \mathbf{U}_\mathrm{A} & \mathbf{u}_\mathrm{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{\Sigma}_\mathrm{A} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{V}_\mathrm{AA} & \mathbf{v}_\mathrm{Ab} \\ \mathbf{v}_\mathrm{bA} & v_\mathrm{bb} \end{pmatrix}^\mathrm{T} \tag{3.28}$$

and hence the perturbation is of the form

$$(\Delta\mathbf{A} \,|\, \Delta\mathbf{b}) = -\begin{pmatrix} \mathbf{U}_\mathrm{A} & \mathbf{u}_\mathrm{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_\mathrm{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{V}_\mathrm{AA} & \mathbf{v}_\mathrm{Ab} \\ \mathbf{v}_\mathrm{bA} & v_\mathrm{bb} \end{pmatrix}^\mathrm{T}. \tag{3.29}$$

The perturbation $(\Delta\mathbf{A} \,|\, \Delta\mathbf{b})$ can be simplified to

$$\begin{aligned} (\Delta\mathbf{A} \,|\, \Delta\mathbf{b}) &= -\sigma_\mathrm{b} \cdot \mathbf{u}_\mathrm{b} \cdot \begin{pmatrix} \mathbf{v}_\mathrm{Ab} \\ v_\mathrm{bb} \end{pmatrix}^\mathrm{T} \\ &= -\begin{pmatrix} \mathbf{A} & \mathbf{b} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}_\mathrm{Ab} \\ v_\mathrm{bb} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}_\mathrm{Ab} \\ v_\mathrm{bb} \end{pmatrix}^\mathrm{T}, \end{aligned} \tag{3.30}$$

where the property $\mathbf{V}^{-1} = \mathbf{V}^\mathrm{T}$ has been used in the above simplification. Thus the span $(\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b})$ does not contain the vector $(\mathbf{v}_\mathrm{Ab}^\mathrm{T}, v_\mathrm{bb})^\mathrm{T}$, i.e.,

$$(\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b}) \cdot \begin{pmatrix} \mathbf{v}_\mathrm{Ab} \\ v_\mathrm{bb} \end{pmatrix} = \mathbf{0}. \tag{3.31}$$

Thus if $v_\mathrm{bb}$ is non-zero, then

$$(\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b}) \cdot \begin{pmatrix} -\frac{\mathbf{v}_\mathrm{Ab}}{v_\mathrm{bb}} \\ -1 \end{pmatrix} = (\mathbf{A} + \Delta\mathbf{A} \,|\, \mathbf{b} + \Delta\mathbf{b}) \cdot \begin{pmatrix} \hat{\mathbf{r}} \\ -1 \end{pmatrix} = \mathbf{0}. \tag{3.32}$$

Thus the total least squares solution is given by

$$\hat{\mathbf{r}}_\mathrm{TLS} = -\frac{\mathbf{v}_\mathrm{Ab}}{v_\mathrm{bb}}. \tag{3.33}$$

It has been proved in [GVL80, GVL13] that if $\tilde{\sigma}_3 > \sigma_4$, where $\tilde{\sigma}_3$ is the smallest singular value of $\mathbf{A}$, then $v_\mathrm{bb}$ is non-zero and (3.33) yields a unique solution.

It has been proved in [GVL80] that the condition of the total least squares problem is always worse than the condition of the corresponding least squares problem. Thus significant improvements in performance by the total least squares algorithm over the least squares algorithm can be obtained for well-conditioned problems, i.e., for cases

where $\mathbf{A}$ has full rank with a significant gap in the singular values of $\mathbf{A}$. Furthermore, the total least squares problem is unstable whenever the smallest singular value $\tilde{\sigma}_3$ of $\mathbf{A}$ is close to the smallest singular value $\sigma_4$ of $(\mathbf{A} \,|\, \mathbf{b})$, i.e., $(\tilde{\sigma}_3 - \sigma_4)^{-1}$ measures the sensitivity of the total least squares problem [GVL80]. This happens especially for cases where there are similar propagation paths. If the propagation paths are similar, then it is highly likely that $\mathbf{A}$ is nearly rank deficient with a small gap in the singular values. In such cases, just as in the case of the least squares problem, the total least squares solution can be stabilized by adding a quadratic penalty to the total least squares objective function or by imposing a quadratic constraint bounding the size of the total least squares solution [SVHG04, BBT06]. In this thesis, a rather simple approach is considered where the least squares solution is used instead of the total least squares solution for cases where $\tilde{\sigma}_3 - \sigma_4 < 0.1$. The threshold value of 0.1 is chosen from empirical studies.

Thus using (3.33) it is possible to estimate the position of a MS in NLOS multipath environments using the path length differences, the angles-of-departure and the angles-of-arrival of the propagation paths. An extension of the algorithm of (3.33) to the case where there are multiple BSs and/or to a three dimensional localization scenario is straightforward.

It must be noted that the localization algorithms proposed in [MYJ07, ST08, WPW11, SW14c] have considered the least squares solution even though the coefficient matrix and the observation vector are contaminated by noise. Hence significant performance improvement would be obtained if the total least squares algorithm is employed as presented in this thesis.

## 3.4 Performance analysis

### 3.4.1 Cramér-Rao lower bound

The performance of the estimation accuracy is one of the essential features which determines the practicability of any estimator. Thus having a performance benchmark to which the performance of the accuracy of estimators can be compared against is of paramount importance in assessing the efficiency of estimators. The Cramér-Rao lower bound (CRLB) is one such performance benchmark which determines the lower bound of the mean square error of any unbiased estimator of a non-random parameter.

Let the $2(P+1) \times 1$ vector $\mathbf{p}_{\mathrm{sm}} = (\mathbf{p}_{\mathrm{sc},1}^{\mathrm{T}}, \ldots, \mathbf{p}_{\mathrm{sc},P}^{\mathrm{T}}, \mathbf{p}_{\mathrm{ms}}^{\mathrm{T}})^{\mathrm{T}}$ denote the unknown non-random two dimensional positions of the MS and the scatterers. Furthermore, let the $3P \times 1$ vector-valued function $\mathbf{g}(\mathbf{p}_{\mathrm{sm}}) = (d_{1,q}, \ldots, d_{P,q}, \phi_{\mathrm{bs},1}, \ldots, \phi_{\mathrm{bs},P}, \phi_{\mathrm{ms},1}, \ldots, \phi_{\mathrm{ms},P})^{\mathrm{T}}$ denote the path length differences, the angles-of-departure and the angles-of-arrival of the $P$ propagation paths and let the $3P \times 1$ vector $\mathbf{z}$ denote the measured path parameters. As mentioned in Section 2.2, the measured path parameters are assumed to be i.i.d. Gaussian random variables. Thus the relation between the measured path parameters $\mathbf{z}$ and the unknown non-random vector parameter $\mathbf{p}_{\mathrm{sm}}$ is given as

$$\mathbf{z} = \mathbf{g}(\mathbf{p}_{\mathrm{sm}}) + \mathbf{w}, \tag{3.34}$$

where $\mathbf{w}$ is the measurement noise which is assumed to be multivariate zero-mean Gaussian distributed with covariance matrix $\mathbf{R}_{\mathrm{ww}}$, i.e.,

$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\mathrm{ww}}). \tag{3.35}$$

Thus the conditional probability density function (pdf) of $\mathbf{z}$ given $\mathbf{p}_{\mathrm{sm}}$ is given as

$$\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right) = \frac{1}{\sqrt{(2\pi)^{3P}\det(\mathbf{R}_{\mathrm{ww}})}} \exp\left(-\frac{1}{2}\left(\mathbf{z}-\mathbf{g}(\mathbf{p}_{\mathrm{sm}})\right)^{\mathrm{T}} \mathbf{R}_{\mathrm{ww}}^{-1}\left(\mathbf{z}-\mathbf{g}(\mathbf{p}_{\mathrm{sm}})\right)\right). \tag{3.36}$$

The CRLB relates the likelihood function $\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)$ to the minimum mean square error of an unbiased estimator $\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z})$ of $\mathbf{p}_{\mathrm{sm}}$, under some regularity conditions on the pdf $\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)$ [Kay93]. Mathematically, for an unbiased estimator $\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z})$, i.e., $\mathbb{E}\{\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z})\} = \mathbf{p}_{\mathrm{sm}}$, the covariance matrix is bounded from below as

$$\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\{(\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z}) - \mathbf{p}_{\mathrm{sm}})(\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z}) - \mathbf{p}_{\mathrm{sm}})^{\mathrm{T}}\} - \mathbf{J}^{-1}(\mathbf{p}_{\mathrm{sm}}) \succeq 0, \tag{3.37}$$

where $\mathbf{J}(\mathbf{p}_{\mathrm{sm}})$ is the Fisher information matrix defined as

$$\begin{aligned}\mathbf{J}(\mathbf{p}_{\mathrm{sm}}) &= -\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\left\{\left(\frac{\partial^2 \ln\left(\mathrm{p}(\mathbf{z}|\mathbf{p}_{\mathrm{sm}})\right)}{\partial \mathbf{p}_{\mathrm{sm}} \partial \mathbf{p}_{\mathrm{sm}}^{\mathrm{T}}}\right)\right\} \\ &= \mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)\right)}{\partial \mathbf{p}_{\mathrm{sm}}}\right)\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)\right)}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}}\right\}.\end{aligned} \tag{3.38}$$

The symbol $\succeq$ in (3.37) denotes that the difference of the two matrices is positive semi-definite and $\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)$. Substituting (3.36) in (3.38) and computing the gradients [PP12] results in

$$\begin{aligned}&\mathbf{J}(\mathbf{p}_{\mathrm{sm}}) = \\ &\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\left\{\left(\left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left(\mathbf{z}-\mathbf{g}(\mathbf{p}_{\mathrm{sm}})\right)\right)\left(\left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left(\mathbf{z}-\mathbf{g}(\mathbf{p}_{\mathrm{sm}})\right)\right)^{\mathrm{T}}\right\}.\end{aligned} \tag{3.39}$$

Re-arranging the above equation results in

$$\mathbf{J}(\mathbf{p}_{\mathrm{sm}}) = \left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}} \left\{(\mathbf{z} - \mathbf{g}(\mathbf{p}_{\mathrm{sm}}))(\mathbf{z} - \mathbf{g}(\mathbf{p}_{\mathrm{sm}}))^{\mathrm{T}}\right\} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}}. \tag{3.40}$$

The expectation in the above equation is nothing but the covariance matrix $\mathbf{R}_{\mathrm{ww}}$ of the measurement noise. Substituting $\mathbf{R}_{\mathrm{ww}}$ in place of the expectation results in

$$\mathbf{J}(\mathbf{p}_{\mathrm{sm}}) = \left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right). \tag{3.41}$$

Thus the CRLB for the conditional pdf $\mathrm{p}\left(\mathbf{z}|\mathbf{p}_{\mathrm{sm}}\right)$ given in (3.36) is given as

$$\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{sm}}}\{(\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z}) - \mathbf{p}_{\mathrm{sm}})(\hat{\mathbf{p}}_{\mathrm{sm}}(\mathbf{z}) - \mathbf{p}_{\mathrm{sm}})^{\mathrm{T}}\} \succeq \left(\left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left(\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}\right)\right)^{-1}. \tag{3.42}$$

The Jacobian matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ can be calculated using the expressions given in Appendix A.

The Fisher information matrix $\mathbf{J}(\mathbf{p}_{\mathrm{sm}})$ can be interpreted as a matrix quantifying the amount of information [BSKL01] contained in the measured path parameters $\mathbf{z}$ about the unknown vector parameter $\mathbf{p}_{\mathrm{sm}}$. Hence (3.37) is intuitive in that the lower bound of the covariance matrix of an unbiased estimator is inversely related with the amount of information available about the unknown parameter. Consequently, an estimator is said to be efficient if the covariance matrix of the estimator is equal to the CRLB.

From the discussion of the localization principle in Section 3.2, it is known that at least three NLOS propagation paths are required to localize a MS, and implicitly the scatterers. This result can also be confirmed from the CRLB. The vector parameter $\mathbf{p}_{\mathrm{sm}}$ is identifiable if the Fisher information matrix $\mathbf{J}(\mathbf{p}_{\mathrm{sm}})$ derived from the likelihood function in (3.36) is nonsingular [Rot71]. As it can be inferred from (3.41), $\mathbf{J}(\mathbf{p}_{\mathrm{sm}})$ is nonsingular if $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ has full column rank. The $3P \times 2(P+1)$ matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ has a zero row vector owing to the zero path length difference of the $q^{\mathrm{th}}$ path as it can be inferred from (3.3). Considering the zero row vector in the matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$, the matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ can have a full column rank of $2(P+1)$ only for $P \geq 3$. Consequently, the position of the MS is identifiable for $P \geq 3$ if the matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ has full column rank. For random positions of MS and scatterers, the matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ can be considered to have full column rank for $P \geq 3$. Simulation results have shown that the matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ can indeed be considered to have full column rank for random positions of MS and scatterers for $P \geq 3$.

The vector parameter $\mathbf{p}_{\mathrm{sm}} = (\mathbf{p}_{\mathrm{sc},1}^{\mathrm{T}}, \ldots, \mathbf{p}_{\mathrm{sc},P}^{\mathrm{T}}, \mathbf{p}_{\mathrm{ms}}^{\mathrm{T}})^{\mathrm{T}}$ contains the unknown position of the MS and the scatterers. Since only the position of the MS is of interest, the CRLB of

the estimate of the position of the MS is required. This can be obtained by partitioning the Fisher information matrix $\mathbf{J}(\mathbf{p}_{\mathrm{sm}})$ into block matrices

$$\mathbf{J}(\mathbf{p}_{\mathrm{sm}}) = \left( \begin{array}{c|c} \mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}}) & \mathbf{J}_{12}(\mathbf{p}_{\mathrm{sm}}) \\ \hline \mathbf{J}_{21}(\mathbf{p}_{\mathrm{sm}}) & \mathbf{J}_{22}(\mathbf{p}_{\mathrm{sm}}) \end{array} \right), \tag{3.43}$$

where $\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}})$ is equal to the Fisher information matrix with the position of the MS as the unknown non-random parameter, i.e., $\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}}) = \mathbf{J}(\mathbf{p}_{\mathrm{ms}})$. The inverse of the submatrix $\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}})$ can be obtained from the inverse of the Schur complement of the submatrix $\mathbf{J}_{22}(\mathbf{p}_{\mathrm{sm}})$ [PP12]. Since $\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}}) = \mathbf{J}(\mathbf{p}_{\mathrm{ms}})$, the inverse of the submatrix $\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}})$ yields the CRLB of the covariance matrix of an unbiased MS position estimator $\hat{\mathbf{p}}_{\mathrm{ms}}(\mathbf{z})$, i.e.,

$$\begin{aligned} &\mathbb{E}_{\mathbf{z}|\mathbf{p}_{\mathrm{ms}}}\{(\hat{\mathbf{p}}_{\mathrm{ms}}(\mathbf{z}) - \mathbf{p}_{\mathrm{ms}})(\hat{\mathbf{p}}_{\mathrm{ms}}(\mathbf{z}) - \mathbf{p}_{\mathrm{ms}})^{\mathrm{T}}\} \succeq \\ &\quad \left(\mathbf{J}_{11}(\mathbf{p}_{\mathrm{sm}}) - \mathbf{J}_{12}(\mathbf{p}_{\mathrm{sm}}) \cdot \mathbf{J}_{22}^{-1}(\mathbf{p}_{\mathrm{sm}}) \cdot \mathbf{J}_{21}(\mathbf{p}_{\mathrm{sm}})\right)^{-1}. \end{aligned} \tag{3.44}$$

As it can be seen from (3.42), the CRLB of an unbiased MS and scatterers position estimator is related to the error in the path parameter estimation appropriately weighted using the Jacobian matrix $\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}}$ which relates the errors in the path parameter estimates to the errors in the estimates of the position of the MS and the scatterers. Thus the estimation error of the position of the MS and scatterers depends not only the path parameter estimation error but also on the relative geometry of the BSs, the MS and the scatterers. This can be explained quite intuitively by considering the impact of the same path parameter estimation error on two localization scenarios with different geometries between a BS, a MS and two scatterers as shown in Fig. 3.3. For the sake of simpler exposition, only the position of the second scatterer is changed to obtain a different localization geometry. Furthermore, it is assumed that the clocks of the BS and MS are synchronized and hence estimates of the propagation path lengths are available. The two line segments of possible positions of the MS are shown pursuant to the propagation paths via the two scatterers for both localization scenarios. For the sake of simplicity, only errors in the estimation of the path lengths shall be considered. Consequently, two error bounds on the line segments of the possible positions of the MS arising from error bounds on the two propagation path lengths are shown by dotted line segments for both localization scenarios. The error bounds are the same for the two propagation paths for both localization scenarios. The areas of possible position estimates of the MS for the error bounds on the estimates of the two propagation path lengths are shown by the shaded areas in Fig. 3.3. It is clear that for the localization scenario shown in Fig. 3.3a, the shaded area is considerably smaller than for the localization scenario shown in Fig. 3.3b. Thus for the same error bound, different accuracies on the estimated positions of the MS are obtained owing to the difference of the relative geometry between the BSs, MS and scatterers. This property has been studied in

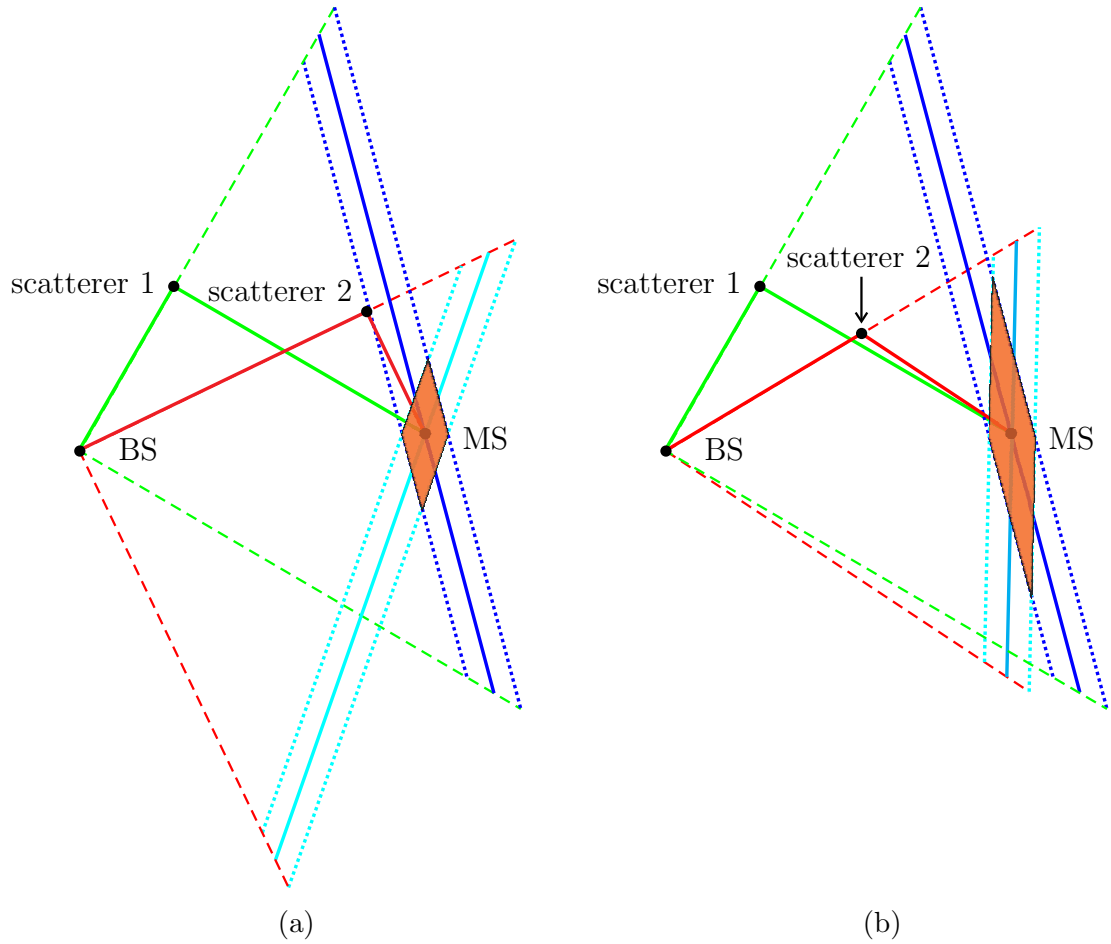



Figure 3.3. Influence of the relative geometry of a BS, a MS and two scatterers on the accuracy of the MS position estimate given the same error on path parameters estimates: (a) localization scenario with good relative geometry, (b) localization scenario with bad relative geometry

satellite navigation systems and it is referred to as dilution of precision [KH06]. Thus the localization geometry shown in Fig. 3.3a has a smaller dilution of precision than the localization geometry shown in Fig. 3.3b.

The CRLB determines the lower bound of the mean square error of any unbiased estimator of a non-random parameter. As it can be seen from (3.11), the measured propagation path parameters are processed nonlinearly to obtain a linear relation between the position of the MS and the position of the BS. Furthermore, the propagation path parameters are nonlinearly related with the position of the MS. Thus a bias should be expected in the proposed localization algorithm. Consequently, the derived CRLB is not tight.

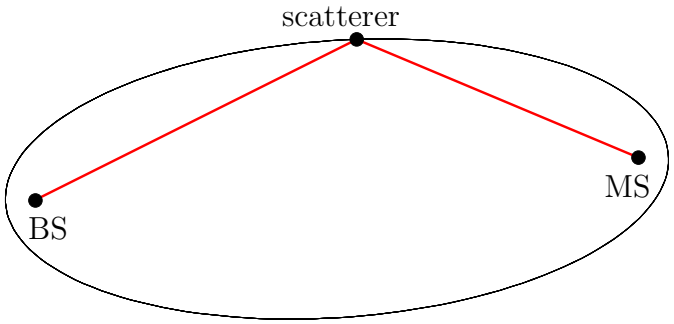



Figure 3.4. One possible scatterer position and NLOS propagation path according to the single bounce elliptical model

### 3.4.2 Simulation results

In this section, the performance of the proposed localization algorithm is analysed considering a picocell localization scenario using Monte Carlo simulations with $10^5$ independent trials. It is assumed that the radius of the cell is 100 m and it is served by a single BS. Furthermore, the probability distribution of the position of the MS is assumed to be uniform within the cell. The positions of the scatterers are generated randomly using the single bounce elliptical model [LR99]. In the single bounce elliptical model, each propagation path with delay $\tau_p$ is assumed to be a result of a single reflection. Consequently, the $p^{\text{th}}$ scatterer lies on the perimeter of an ellipse with the BS and the MS located at the focal points of the ellipse. Fig. 3.4 shows one possible scatterer position on the perimeter of an ellipse and the corresponding NLOS propagation path. If a limit on the maximum propagation path delay $\tau_{\max}$ is set, then all the propagation paths have a delay between the LOS path delay $\tau_{\text{LOS}}$ and the maximum path delay $\tau_{\max}$. Furthermore, scatterers corresponding to the propagation path delays $\tau_{\text{LOS}} \leq \tau_p < \tau_{\max}$ are located within the ellipse defined by the maximum path delay $\tau_{\max}$. In the simulations, the positions of scatterers are assumed to be uniformly distributed within an ellipse defined by $\tau_{\max}$. The simulations were performed using a normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}} = 2.5$. It must be noted that even though the LOS path exists between the BS and the MS, the probability of generating a scatterer located at the LOS path is zero.

The performance metric is the root mean square error $\text{rmse}(\hat{\mathbf{p}}_{\text{ms}})$ of the estimated position $\hat{\mathbf{p}}_{\text{ms}}$ of the MS which is calculated as

$$\text{rmse}(\hat{\mathbf{p}}_{\text{ms}}) = \sqrt{\mathbb{E}_{\mathbf{z},\mathbf{p}_{\text{ms}}}\left\{(\hat{\mathbf{p}}_{\text{ms}} - \mathbf{p}_{\text{ms}})^{\text{T}}(\hat{\mathbf{p}}_{\text{ms}} - \mathbf{p}_{\text{ms}})\right\}}, \tag{3.45}$$

where $\mathbf{p}_{\text{ms}}$ is the true position of the MS and $\mathbb{E}_{\mathbf{z},\mathbf{p}_{\text{ms}}}\{\cdot\}$ denotes the expectation with respect to the pdf $\text{p}(\mathbf{z}, \mathbf{p}_{\text{ms}})$. In the simulations, the expectations in (3.45) are computed

using averages. Since the localization scenarios are randomly generated, the averaged CRLB of the randomly generated localization scenarios is considered as a performance bound of the proposed localization algorithm.

In the following, the performances of the least squares (LS) and the total least squares (TLS) localization algorithms will be studied. The LS localization algorithm is presented to assess the performance improvement brought about by the TLS localization algorithm.

Fig. 3.5 shows the performances of the LS and the TLS localization algorithms versus the standard deviation $\sigma_\phi = \sigma_{\phi_{\mathrm{bs}},p} = \sigma_{\phi_{\mathrm{ms}},p}$ of the angles-of-departure and angles-of-arrival measurement noise for different numbers of propagation paths $P$. In the simulations, the standard deviation $\sigma_{\mathrm{dd}}$ of the path length differences measurement noise was set to 5 m. It can be seen from Fig. 3.5 that the performances of the LS and the TLS localization algorithms are bounded by the CRLB. Furthermore, there are significant performance improvements by both the LS and the TLS localization algorithms as the number of propagation paths $P$ is increased owing to the extra information brought about the position of the MS from the extra propagation paths. This shows that the NLOS multipath propagations are exploited for localization of the MS.

It can be seen from Fig. 3.5 that the TLS localization algorithm results in a significant performance improvement over the LS localization algorithm as the number of propagation paths $P$ is increased. As mentioned in Section 3.3, the condition of the TLS problem is always worse than the condition of the corresponding LS problem. Consequently, a significant improvement in performance by the TLS algorithm over the LS algorithm is achieved for well-conditioned problems. For the randomly generated localization scenario, the probability of obtaining a well-conditioned problem increases as the number of propagation paths increases. This explains for the improvement in performance by the TLS localization algorithm over the LS localization algorithm as the number of paths is increased.

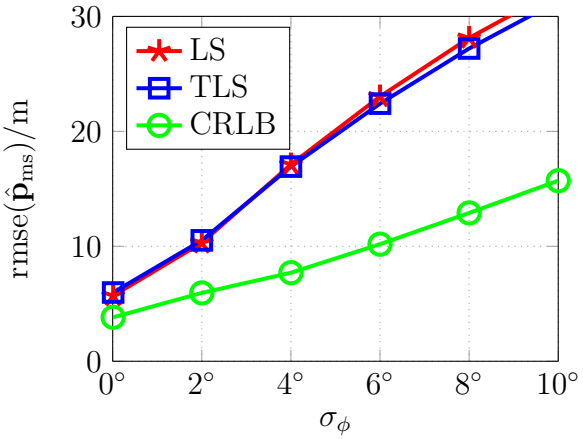
It can also be seen from Fig. 3.5 that the performance improvement brought about by the TLS localization algorithm over the LS localization algorithm is more pronounced as the standard deviation $\sigma_\phi$ of the angles-of-departure and angles-of-arrival measurement noise is increased. This result happens owing to the fact that the performance degradation of the LS localization algorithm for the assumption of error-free matrix $\mathbf{A}$ increases as the perturbation in $\mathbf{A}$ increases.

Fig. 3.6 shows the performances of the LS and the TLS localization algorithms versus the standard deviation $\sigma_{\mathrm{dd}}$ of the path length differences measurement noise for

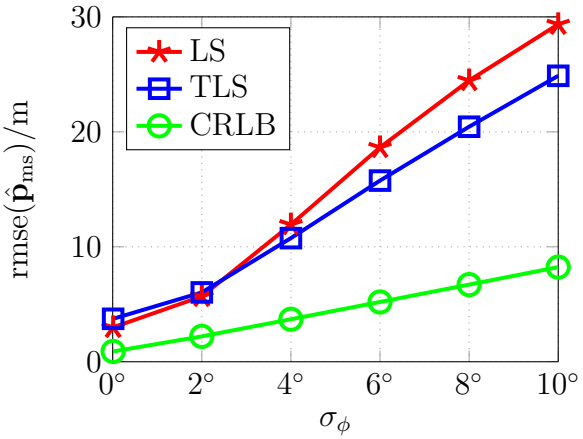
different numbers of propagation paths $P$. In the simulations, the standard deviations $\sigma_{\phi_{\text{bs}}}$ and $\sigma_{\phi_{\text{ms}}}$ of the angles-of-departure and angles-of-arrival measurement noise were set to 5°. It can be seen from Fig. 3.6 that the performances of the LS and the TLS localization algorithms are bounded by the CRLB. Furthermore, confirming the exploitation of the NLOS propagation paths for localization, there are significant performance improvements by both the LS and the TLS localization algorithms as the number of propagation paths $P$ is increased. Moreover, similar to the simulation results shown in Fig. 3.5, the TLS localization algorithm results in a significant performance improvement over the LS localization algorithm as the number of propagation paths $P$ is increased.

In contrast to the simulation results shown in Fig. 3.5, the performance improvement brought about by the TLS localization algorithm over the LS localization algorithm diminishes as the standard deviation $\sigma_{\text{dd}}$ of the path length differences measurement noise is increased. This result happens owing to the differences in perturbations of $\mathbf{A}$ and $\mathbf{b}$. As it can be inferred from the non-perturbed versions of $\mathbf{A}$ and $\mathbf{b}$ in (3.13) and (3.14), respectively the perturbation in $\mathbf{A}$ comes solely from the noises in the measured angles-of-departure and angles-of-arrival of the propagation paths, whereas the perturbation in $\mathbf{b}$ comes from the noises in the measured path length differences, angles-of-departure and angles-of-arrival of the propagation paths. Since the simulation result in Fig. 3.6 is shown for a fixed standard deviation $\sigma_{\phi}$ of the angles-of-departure and angles-of-arrival measurement noise and varying standard deviation $\sigma_{\text{dd}}$ of the path length differences measurement noise, the perturbation power in $\mathbf{A}$ is fixed and varying in $\mathbf{b}$. As $\sigma_{\text{dd}}$ is increased, the perturbation in $\mathbf{b}$ becomes more dominant than the perturbation in $\mathbf{A}$. Consequently, the performance improvement from considering the perturbation in $\mathbf{A}$ by using the total least squares estimator over the least squares estimator diminishes as $\sigma_{\text{dd}}$ is increased.

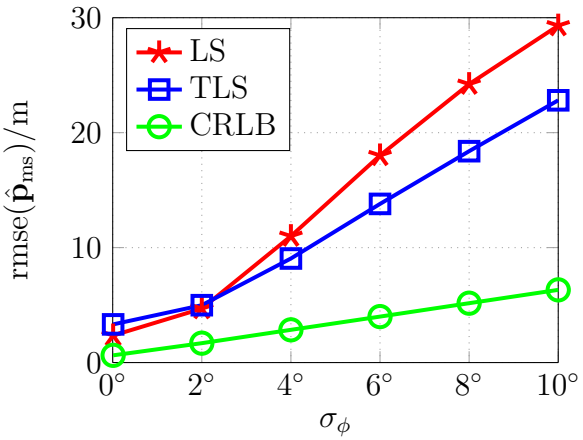
The differences of perturbations in $\mathbf{A}$ and $\mathbf{b}$ also explains for the rather similar performances by both the LS and the TLS localization algorithms in Fig. 3.5 when $\sigma_{\phi} \leq 2°$. Since the perturbations in $\mathbf{A}$ are low at low values of $\sigma_{\phi}$, the LS and the TLS problems are essentially similar. The small performance differences in Fig. 3.5 for such cases are caused by the rather simple approach taken for mitigating cases where the total least squares problem is unstable.




(a) Number of paths $P = 6$

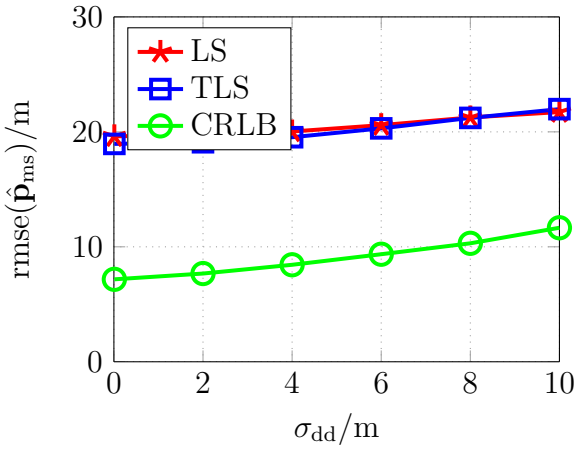



(b) Number of paths $P = 12$

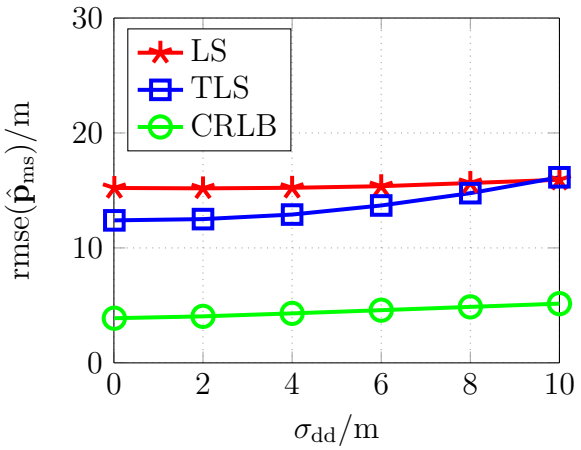



(c) Number of paths $P = 18$

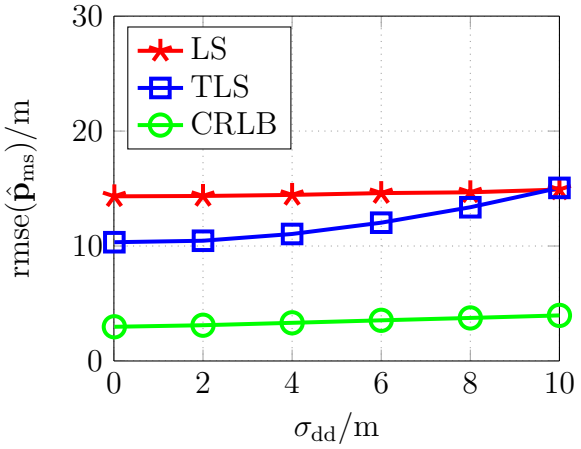
Figure 3.5. Influence of the number of paths and the standard deviation of the angles-of-departure and angles-of-arrival measurement noise on the proposed localization algorithms, $\sigma_{\mathrm{dd}} = 5$ m




(a) Number of paths $P = 6$




(b) Number of paths $P = 12$




(c) Number of paths $P = 18$

Figure 3.6. Influence of the number of paths and the standard deviation of the path length differences measurement noise on the proposed localization algorithms, $\sigma_\phi = 5°$ (figure drawn not to scale)

# Chapter 4

# Performance gains from tracking a mobile station

## 4.1 The mobile station tracking problem

In the previous chapter, localization of a MS in NLOS multipath environments under the explicit consideration of scatterers was discussed. The position of the MS was estimated by exploiting the position dependent parameters of the radio propagation paths using the single bounce scattering model. Owing to the mobility of MSs, the position of a MS can change with time. Consequently, tracking a MS is of interest in many applications. A straightforward approach to tracking a MS would be to make successive independent position estimates for each position constituting the trajectory of a MS using the localization algorithm discussed in the previous chapter. However, the straightforward successive independent MS position estimation approach does not exploit the possible statistical dependency between the consecutive positions constituting the trajectory of a MS. Owing to the physical mobility constraints of MSs, there is a strong statistical dependency between the consecutive positions of a MS. By proper modeling of the mobility of a MS, significant performance gains over the successive independent MS position estimation approach can be obtained by exploiting the statistical dependency between successive positions of the MS [SFW12, SW12, SW14a]. It has been shown in [WM05] that even post processing of the MS positions estimates obtained from successive independent position estimation by using a proper mobility model of a MS can result in improved MS positions estimates. In this chapter, the performance improvements brought about by considering the mobility of a MS for tracking a MS in NLOS multipath environments under the explicit consideration of scatterers are presented.

Before discussing the approach to tracking a MS in NLOS multipath environments under the explicit consideration of scatterers, it is essential to discuss the MS tracking problem first. In practice, tracking a MS in discrete-time domain is of interest. Thus the trajectory of a MS is tracked using noisy estimates of the propagation path parameters available at discrete times $\ell$. The evolution of the positions of a MS with time can well be described using a discrete time dynamical model. Owing to its convenience for handling multivariate data and nonlinear functions, the state-space approach to modeling dynamical systems is considered for the MS tracking problem. Furthermore,

the state-space approach gives an intuitive system description which is close to the physical reality [GKN+74]. The state-space model describes a dynamical system using a state transition model and a measurement model. The state transition model describes the change of the state of the system with time, whereas the measurement model describes the relation between the measurements and the state of the system.

The state vector $\boldsymbol{\theta}(\ell)$ represents the state of the system at discrete time $\ell$. Consequently, for the considered tracking problem, i.e., tracking a MS under the explicit consideration of scatterers, the state vector shall contain the positions of the MS and scatterers. Furthermore, as the position of the MS is changing with time, the derivatives of the position of the MS with time, e.g., velocity and acceleration, shall also be included in the state vector. As mentioned in Section 2.1, stationary scatterers are considered and hence the derivatives of the positions of the scatterers with time are not included in the state vector.

For the MS tracking problem, since stationary scatterers are considered, the state transition model is essentially defined by the mobility model of the MS. Owing to the physical mobility constraints of MSs, a MS cannot move in a totally random way. Furthermore, a MS to be tracked seldom moves in a totally deterministic way. In practice, the MS tracking problem involves a MS trajectory with some statistical dependency between the successive positions and some randomness accounting for manoeuvering operations by the MS. Thus the positions of the MS can be considered as realizations of a random process with some statistical dependency between the consecutive positions of the MS. The dependency between the successive positions of the MS precludes the MS from moving in a totally random way, whereas the randomness precludes the MS from moving in a totally deterministic way. A general mobility model would consider the statistical dependency between all the positions of the MS constituting the trajectory of the MS, i.e., the joint pdf of the positions of the MS. However, modeling and analysing such a mobility model is a formidable task and computationally demanding. Hence in this thesis a rather simple approach is considered where the statistical dependency between the positions of the MS is described using a first-order Markov model. Thus ignoring the randomness in the positions of the MS, the future positions of the MS would only depend on the current position of the MS and the current non-zero derivatives of the position of the MS with time. The random change in the position of the MS is modeled using Gaussian noise. Hence the mobility of the MS is described using a first-order Gauss-Markov process [Kay93]. Using the first-order Gauss-Markov mobility model, the state transition equation can be defined as

$$\boldsymbol{\theta}(\ell) = \mathbf{B} \cdot \boldsymbol{\theta}(\ell - 1) + \mathbf{u}(\ell), \tag{4.1}$$

where the matrix $\mathbf{B}$ is the state transition matrix which is determined by the dependency between the consecutive state vectors and $\mathbf{u}(\ell)$ is the state transition noise. The state vector $\boldsymbol{\theta}(\ell-1)$ summarizes the effect of all past inputs to the system. Furthermore, together with the state transition noise at time $\ell$ and beyond, the state vector $\boldsymbol{\theta}(\ell-1)$ determines the future outputs of the system [Kay93].

There are different classes of mobility models derived from basic equations of motion, viz. constant velocity and constant acceleration models [BSKL01]. In order to account for adjustments to disturbances, the constant velocity and constant acceleration models model the second- and third-order derivatives of the position of a MS by a zero mean random process, respectively. In this thesis, a simplified mobility model of a MS is considered. The MS is assumed to move with a constant velocity $\mathbf{v}_{\text{ms}}(\ell)$ during the tracking interval $\Delta t$. In order to account for some manoeuvering operations made by the MS, random changes in the velocity of the MS are considered in every tracking interval $\Delta t$. The random changes in the velocity of the MS are modeled by a white Gaussian noise, i.e., white noise acceleration. Thus the position $\mathbf{p}_{\text{ms}}(\ell)$ and the velocity $\mathbf{v}_{\text{ms}}(\ell)$ of the MS evolve as follows

$$\mathbf{v}_{\text{ms}}(\ell) = \mathbf{v}_{\text{ms}}(\ell-1) + \mathbf{u}_{\text{ms}}(\ell) \tag{4.2}$$

$$\mathbf{p}_{\text{ms}}(\ell) = \mathbf{p}_{\text{ms}}(\ell-1) + \Delta t \cdot \mathbf{v}_{\text{ms}}(\ell), \tag{4.3}$$

where the $2 \times 1$ vector $\mathbf{u}_{\text{ms}}(\ell)$ is a multivariate white Gaussian noise, i.e.,

$$\mathbf{u}_{\text{ms}}(\ell) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\text{uu}}). \tag{4.4}$$

Thus the state vector includes the positions of the MS and scatterers and the velocity of the MS, i.e., the $2(P+2) \times 1$ vector $\boldsymbol{\theta}(\ell) = (\mathbf{p}_{\text{ms}}^{\text{T}}(\ell), \mathbf{v}_{\text{ms}}^{\text{T}}(\ell), \mathbf{p}_{\text{sc},1}^{\text{T}}, \ldots, \mathbf{p}_{\text{sc},P}^{\text{T}})^{\text{T}}$ defines the state vector. Based on the above mobility model of the MS, the first order Gauss-Markov mobility model of (4.1) can be re-defined as

$$\boldsymbol{\theta}(\ell) = \mathbf{B} \cdot \boldsymbol{\theta}(\ell-1) + \boldsymbol{\Gamma} \cdot \mathbf{u}_{\text{ms}}(\ell), \tag{4.5}$$

where the $2(P+2) \times 2(P+2)$ state transition matrix $\mathbf{B}$ is defined as

$$\mathbf{B} = \begin{pmatrix} \mathbf{I}_2 & \Delta t \cdot \mathbf{I}_2 & \mathbf{0}_{2\times 2P} \\ \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_{2\times 2P} \\ \mathbf{0}_{2P\times 2} & \mathbf{0}_{2P\times 2} & \mathbf{I}_{2P} \end{pmatrix} \tag{4.6}$$

and the matrix $\boldsymbol{\Gamma}$ is defined, considering the stationary scatterers, as

$$\boldsymbol{\Gamma} = \begin{pmatrix} \Delta t & 0 \\ 0 & \Delta t \\ 1 & 0 \\ 0 & 1 \\ \mathbf{0}_{2P\times 1} & \mathbf{0}_{2P\times 1} \end{pmatrix}. \tag{4.7}$$

For the sake of simplicity, the initial state vector $\boldsymbol{\theta}(0)$ is modeled as multivariate Gaussian distributed random vector, i.e.,

$$\boldsymbol{\theta}(0) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\theta\theta}(0)). \tag{4.8}$$

Thus given the initial state vector $\boldsymbol{\theta}(0)$, the state transition equation (4.5) models the movement of a MS by making random changes to the velocity of the MS which is assumed to be constant within the tracking interval $\Delta t$.

The measurement vector $\mathbf{z}(\ell)$ contains noisy estimates of the propagation path parameters. In the previous chapter, localization of a MS for unsynchronized BS and MS was considered. In this chapter, tracking a MS for synchronized BS and MS is considered. Hence the path lengths of the propagation paths are considered for tracking the MS rather than the path length differences of the propagation paths. Needless to say, the angles-of-departure and angles-of-arrivals of the propagation paths are also considered for tracking the MS. As given in (2.10), the measured path parameters are assumed to be i.i.d. Gaussian random variables. Thus the relation between the measured path parameters $\mathbf{z}(\ell)$ and the state vector $\boldsymbol{\theta}(\ell)$ is defined as

$$\mathbf{z}(\ell) = \mathbf{g}(\boldsymbol{\theta}(\ell)) + \mathbf{w}(\ell), \tag{4.9}$$

where $\mathbf{g}(\boldsymbol{\theta}(\ell)) = (d_1(\ell), \ldots, d_P(\ell), \phi_{\mathrm{bs},1}(\ell), \ldots, \phi_{\mathrm{bs},P}(\ell),\ \phi_{\mathrm{ms},1}(\ell), \ldots, \phi_{\mathrm{ms},P}(\ell))^{\mathrm{T}}$ is the measurement vector-valued function and $\mathbf{w}(\ell)$ is the measurement noise, which is assumed to be multivariate zero-mean white Gaussian distributed, i.e.,

$$\mathbf{w}(\ell) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\mathrm{ww}}). \tag{4.10}$$

It must be noted that the initial state vector $\boldsymbol{\theta}(0)$ and the noises $\mathbf{u}_{\mathrm{ms}}(\ell)$ and $\mathbf{w}(\ell)$ are assumed to be independent of each other.

Since both the state vector $\boldsymbol{\theta}(\ell)$ and the measurement vector $\mathbf{z}(\ell)$ are realizations of random processes, the Bayesian approach is appropriate for tracking the MS. The state transition equation in (4.5) characterizes the probability density function (pdf) $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1)\right)$, whereas the measurement equation in (4.9) characterizes the conditional pdf $\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right)$ of the measured propagation path parameters $\mathbf{z}(\ell)$ given the state vector $\boldsymbol{\theta}(\ell)$.

The objective of an estimator which tracks a MS is to estimate the optimal state vector at discrete time $\ell$ based on measurements $\mathbf{Z}(\ell) = (\mathbf{z}(1) \cdots \mathbf{z}(\ell)))$ up to time instant $\ell$, i.e., the optimal state vector is estimated from the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$. Using the Bayesian approach, the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ can be recursively computed using

the prediction and update steps given the initial prior $\mathrm{p}\left(\boldsymbol{\theta}(0), \mathbf{z}(0)\right) = \mathrm{p}\left(\boldsymbol{\theta}(0)\right)$, the transition pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1)\right)$ and the likelihood $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{z}(\ell)\right)$ [Che03, AMGC02]. In the prediction step, the pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right)$ is predicted from the pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell-1)|\mathbf{Z}(\ell-1)\right)$ using the Chapman-Kolmogorov equation

$$\begin{aligned}
\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right) &= \int \mathrm{p}\left(\boldsymbol{\theta}(\ell), \boldsymbol{\theta}(\ell-1)|\mathbf{Z}(\ell-1)\right)\, \mathrm{d}\boldsymbol{\theta}(\ell-1) \\
&= \int \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1), \mathbf{Z}(\ell-1)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell-1)|\mathbf{Z}(\ell-1)\right)\, \mathrm{d}\boldsymbol{\theta}(\ell-1) \\
&= \int \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell-1)|\mathbf{Z}(\ell-1)\right)\, \mathrm{d}\boldsymbol{\theta}(\ell-1). \qquad (4.11)
\end{aligned}$$

The simplification $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1), \mathbf{Z}(\ell-1)\right) = \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\boldsymbol{\theta}(\ell-1)\right)$ above is made using the state transition equation in (4.5). In the update step, the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ is computed from the predicted pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right)$ using Bayes' rule

$$\begin{aligned}
\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right) &= \frac{\mathrm{p}\left(\boldsymbol{\theta}(\ell), \mathbf{Z}(\ell)\right)}{\mathrm{p}\left(\mathbf{Z}(\ell)\right)} \\
&= \frac{\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell), \mathbf{Z}(\ell-1)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell), \mathbf{Z}(\ell-1)\right)}{\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right) \cdot \mathrm{p}\left(\mathbf{Z}(\ell-1)\right)} \\
&= \frac{\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right) \cdot \mathrm{p}\left(\mathbf{Z}(\ell-1)\right)}{\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right) \mathrm{p}\left(\mathbf{Z}(\ell-1)\right)} \\
&= \frac{\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right)}{\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right)}, \qquad (4.12)
\end{aligned}$$

where $\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right)$ is a normalizing constant calculated as

$$\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right) = \int \mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell-1)\right)\, \mathrm{d}\boldsymbol{\theta}(\ell). \qquad (4.13)$$

The simplification $\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell), \mathbf{Z}(\ell-1)\right) = \mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right)$ above is made using the measurement equation in (4.9). The normalizing constant $\mathrm{p}\left(\mathbf{z}(\ell)|\mathbf{Z}(\ell-1)\right)$ is derived following similar steps as in (4.11).

The posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ can be determined by recursively solving (4.11) and (4.12). However, owing to the noisy measurements that are nonlinearly related to the state vector, the analytic solution of the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ is intractable. Several algorithms, including the recently developed particle filters, have been proposed for solving such problems [AMGC02, Che03]. Particle filters are more appropriate for nonlinear and non-Gaussian Bayesian tracking problems. In this thesis, the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ is approximately computed by linearizing the nonlinear measurement equation using a first order Taylor series and subsequently applying the Kalman filter.

## 4.2 Mobile station tracking algorithm

The nonlinear measurement equation given in (4.9) can be linearized using a first order Taylor series. The linear approximation of the vector function $\mathbf{g}(\boldsymbol{\theta}(\ell))$ at a reference point $\boldsymbol{\theta}_0(\ell)$, which shall be close to the true value of the vector $\boldsymbol{\theta}(\ell)$ to yield a reliable linear approximation, is given as

$$\mathbf{g}(\boldsymbol{\theta}(\ell)) \approx \mathbf{g}(\boldsymbol{\theta}_0(\ell)) + \mathbf{G}_0(\ell) \cdot (\boldsymbol{\theta}(\ell) - \boldsymbol{\theta}_0(\ell)), \tag{4.14}$$

where the matrix $\mathbf{G}_0(\ell)$ is defined as

$$\mathbf{G}_0(\ell) = \left.\frac{\partial \mathbf{g}(\boldsymbol{\theta}(\ell))}{\partial \boldsymbol{\theta}(\ell)}\right|_{\boldsymbol{\theta}_0(\ell)}. \tag{4.15}$$

The matrix $\mathbf{G}_0(\ell)$ can be calculated using the expressions given in Appendix A. Substituting (4.14) in (4.9) results in the linearized measurement equation

$$\mathbf{z}(\ell) = \mathbf{G}_0(\ell) \cdot \boldsymbol{\theta}(\ell) + \mathbf{g}(\boldsymbol{\theta}_0(\ell)) - \mathbf{G}_0(\ell) \cdot \boldsymbol{\theta}_0(\ell) + \mathbf{w}(\ell). \tag{4.16}$$

It must be noted that the linearization of the measurement equation is to be performed at every tracking interval.

For a tracking problem with a linear Gaussian state transition and measurement equations, the posterior pdf of the current state vector given the current and past measurement vectors is Gaussian. Thus the posterior pdf is parameterized by its mean and covariance matrix. Furthermore, the mean and the covariance matrix of the posterior pdf constitute estimates of the state vector and the covariance matrix of the estimation error, respectively [BSKL01]. The Kalman filter calculates estimates of the state vector and the covariance matrix of the estimation error using the prediction and update steps [Kal60]. In the prediction step, one time unit prediction of the estimates of the state vector and the covariance matrix of the estimation error is performed using the state transition model, whereas in the update step, an update of the predicted estimates of the state vector and the covariance matrix of the estimation error is made using the current measurement vector. However, for the MS tracking problem under consideration, the noisy measurements are nonlinearly related to the state vector. Thus the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ is not Gaussian. Nevertheless, by linearization of the measurement equation using a first order Taylor series, the posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ is approximated by a Gaussian distribution. Estimates of the mean and the covariance matrix of the approximately Gaussian posterior pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell)|\mathbf{Z}(\ell)\right)$ are recursively computed using the prediction and update steps of the Kalman filter. The Kalman filter when applied to filtering problems with linearized state transition and/or measurement equations is commonly referred to as the extended Kalman filter [Kay93, AMGC02, Che03].

Needless to say, it is imperative that the linearization of the measurement equation be performed sufficiently close to the true value of the state vector so that the error in performing the linearization can be ignored. Consequently, in the extended Kalman filter, the predicted state vector at time $\ell$ based on measurements up to $\ell - 1$ denoted by $\hat{\boldsymbol{\theta}}(\ell|\ell-1)$ is used to linearize the measurement function, i.e.,

$$\mathbf{G}(\ell) = \left.\frac{\partial \mathbf{g}(\boldsymbol{\theta}(\ell))}{\partial \boldsymbol{\theta}(\ell)}\right|_{\boldsymbol{\theta}(\ell)=\hat{\boldsymbol{\theta}}(\ell|\ell-1)}. \tag{4.17}$$

The extended Kalman filter starts with an initial estimate $\hat{\boldsymbol{\theta}}(0)$ of the state vector $\boldsymbol{\theta}(0)$ and an initial covariance matrix $\mathbf{M}(0|0)$ of the estimation error and recursively computes estimates $\hat{\boldsymbol{\theta}}(\ell|\ell)$ of the state vector $\boldsymbol{\theta}(\ell)$ and the covariance matrix $\mathbf{M}(\ell|\ell)$ of the estimation error using the following equations

- prediction step:

$$\hat{\boldsymbol{\theta}}(\ell|\ell-1) = \mathbf{B} \cdot \hat{\boldsymbol{\theta}}(\ell-1|\ell-1) \tag{4.18}$$

$$\mathbf{M}(\ell|\ell-1) = \mathbf{B} \cdot \mathbf{M}(\ell-1|\ell-1) \cdot \mathbf{B}^{\mathrm{T}} + \boldsymbol{\Gamma} \cdot \mathbf{R}_{\mathrm{uu}} \cdot \boldsymbol{\Gamma}^{\mathrm{T}} \tag{4.19}$$

- update step:

$$\mathbf{K}(\ell) = \frac{\mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{\mathrm{T}}(\ell)}{\mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{\mathrm{T}}(\ell) + \mathbf{R}_{\mathrm{ww}}} \tag{4.20}$$

$$\hat{\boldsymbol{\theta}}(\ell|\ell) = \hat{\boldsymbol{\theta}}(\ell|\ell-1) + \mathbf{K}(\ell)\left(\mathbf{z}(\ell) - \mathbf{g}(\hat{\boldsymbol{\theta}}(\ell|\ell-1))\right) \tag{4.21}$$

$$\mathbf{M}(\ell|\ell) = \mathbf{M}(\ell|\ell-1) - \mathbf{K}(\ell) \cdot \mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1), \tag{4.22}$$

where the matrix $\mathbf{K}(\ell)$ is the Kalman gain. The Kalman gain quantifies the relative accuracy of the predicted state vector and the measurement vector. Thus a "small" Kalman gain implies that the predicted state vector is more reliable than the measurement vector and hence more weight is given to the predicted state vector than the measurement vector. On the other hand, a "large" Kalman gain implies that the measurement vector is more reliable than the predicted state vector and hence more weight is given to the measurement vector than the predicted state vector [BSKL01].

It must be noted that the prediction step which is based on the state transition equation in (4.5) leaves the estimated positions of the stationary scatterers unchanged. However, in the update step the estimated positions of the stationary scatterers is updated, i.e.,

$$\hat{\boldsymbol{\theta}}(\ell|\ell) = \hat{\boldsymbol{\theta}}(\ell|\ell-1) + \mathbf{K}(\ell)\left(\mathbf{z}(\ell) - \mathbf{g}(\hat{\boldsymbol{\theta}}(\ell|\ell-1))\right) \tag{4.23}$$

updates the estimated positions of the stationary scatterers based on measurement data. Thus the proposed MS tracking algorithm corrects the initially estimated positions of scatterers based on the information obtained from measurement data.

The initial estimate of the position of the MS is obtained by using the total least squares (TLS) localization algorithm discussed in Section 3.3 for synchronized BSs and MS. The initial estimate of the positions of the scatterers is obtained from the estimated position of the MS by exploiting the linear relation between the positions of the MS and scatterers. Furthermore, the initial estimate of the velocity of the MS is obtained from a multivariate Gaussian distribution. The initial estimate $\hat{\boldsymbol{\theta}}(0)$ of the state vector is constructed from the initial estimate of the position of the MS, the velocity of the MS and the positions of the scatterers. In some cases, the initial estimate of the state vector may not be in the neighborhood of the true value of the initial state vector. Hence the linear approximation of the measurement function would be rather coarse and the extended Kalman filter may diverge. In such cases, estimates of the trajectory of the MS obtained from the successive independent position estimates by the TLS localization algorithm are used as MS trajectory estimates of the extended Kalman filter.

It must be noted that unlike the initial estimate of the position of the MS which is determined by exploiting all the measurements, the initial estimate of the position of each scatterer is determined by exploiting the measurement which corresponds to each scatterer. Thus the initial estimates of the positions of the scatterers are not as reliable as the initial estimate of the position of the MS. Consequently, the entries of the initial covariance matrix $\mathbf{M}(0|0)$ of the state vector for the positions of the scatterers should be much larger as compared to the entries for the positions of the MS. This would mitigate the propagation of the estimation errors arising from the possibly erroneous initial estimate of the positions of the scatterers to the estimated positions of the MS during tracking.

## 4.3 Performance analysis

### 4.3.1 Posterior Cramér-Rao lower bound

Performance bounds give useful information about attainability of a given performance requirement. Furthermore, performance bounds give useful information about the efficiency of an estimator in terms of the amount of information extracted from the

measurements about the parameter in question. In Section 3.4.1, the lower bound of the mean square error of an unbiased estimator of a non-random parameter has been presented using the CRLB. In the MS tracking problem considered here, the unknown parameter is random as it tracks the state of a nonlinear randomly driven dynamical system. The lower bound of the mean-square error of a biased estimator of a random scalar or vector parameter is commonly referred to as the Van Trees bound [VT68] or the posterior CRLB (PCRLB) [TMN98]. The bias in the estimator comes from the random nature of the parameter to be estimated which is characterized by the a priori pdf of the parameter. Thus unlike the CRLB which considers the conditional pdf, the PCRLB considers the conditional and the a priori pdfs when computing the mean square error of a biased estimator. Hence the posterior CRLB is more general than the conventional CRLB.

Mathematically, for a biased estimator $\hat{\boldsymbol{\theta}}(\mathbf{z}(\ell))$ which estimates the random parameter $\boldsymbol{\theta}(\ell)$ based on the measurement vector $\mathbf{z}(\ell)$, the covariance matrix of the estimator is bounded from below as

$$\mathbb{E}_{\mathbf{z}(\ell),\boldsymbol{\theta}(\ell)}\{(\hat{\boldsymbol{\theta}}(\mathbf{z}(\ell)) - \boldsymbol{\theta}(\ell))(\hat{\boldsymbol{\theta}}(\mathbf{z}(\ell)) - \boldsymbol{\theta}(\ell))^{\mathrm{T}}\} - \mathbf{J}^{-1}(\boldsymbol{\theta}(\ell)) \succeq 0, \tag{4.24}$$

where $\mathbf{J}(\boldsymbol{\theta}(\ell))$ is the Fisher information matrix defined as

$$\begin{aligned}\mathbf{J}(\boldsymbol{\theta}(\ell)) &= \mathbb{E}_{\mathbf{z}(\ell),\boldsymbol{\theta}(\ell)}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\boldsymbol{\theta}(\ell),\mathbf{z}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)\left(\frac{\partial \ln\left(\mathrm{p}\left(\boldsymbol{\theta}(\ell),\mathbf{z}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)^{\mathrm{T}}\right\} \\ &= -\mathbb{E}_{\mathbf{z}(\ell),\boldsymbol{\theta}(\ell)}\left\{\left(\frac{\partial^2 \ln\left(\mathrm{p}\left(\boldsymbol{\theta}(\ell),\mathbf{z}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)\partial \boldsymbol{\theta}^{\mathrm{T}}(\ell)}\right)\right\}. \end{aligned} \tag{4.25}$$

The symbol $\succeq$ in (4.24) denotes that the difference of the two matrices is positive semi-definite and $\mathbb{E}_{\mathbf{z}(\ell),\boldsymbol{\theta}(\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\boldsymbol{\theta}(\ell),\mathbf{z}(\ell)\right)$. Similar to the derivation of the CRLB, the PCRLB is derived under certain regularity conditions described in Appendix B. Using the relation $\mathrm{p}\left(\boldsymbol{\theta}(\ell),\mathbf{z}(\ell)\right) = \mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right)\mathrm{p}\left(\boldsymbol{\theta}(\ell)\right)$, the Fisher information matrix $\mathbf{J}(\boldsymbol{\theta}(\ell))$ can be decomposed into the sum of the measurement $\mathbf{J}_{\mathrm{m}}(\boldsymbol{\theta}(\ell))$ and the a priori $\mathbf{J}_{\mathrm{a}}(\boldsymbol{\theta}(\ell))$ information matrices as follows

$$\mathbf{J}(\boldsymbol{\theta}(\ell)) = \mathbf{J}_{\mathrm{m}}(\boldsymbol{\theta}(\ell)) + \mathbf{J}_{\mathrm{a}}(\boldsymbol{\theta}(\ell)), \tag{4.26}$$

where

$$\mathbf{J}_{\mathrm{m}}(\boldsymbol{\theta}(\ell)) = \mathbb{E}_{\mathbf{z}(\ell),\boldsymbol{\theta}(\ell)}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z}(\ell)|\boldsymbol{\theta}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)^{\mathrm{T}}\right\} \tag{4.27}$$

$$\mathbf{J}_{\mathrm{a}}(\boldsymbol{\theta}(\ell)) = \mathbb{E}_{\boldsymbol{\theta}(\ell)}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\boldsymbol{\theta}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)\left(\frac{\partial \ln\left(\mathrm{p}\left(\boldsymbol{\theta}(\ell)\right)\right)}{\partial \boldsymbol{\theta}(\ell)}\right)^{\mathrm{T}}\right\}, \tag{4.28}$$

and $\mathbb{E}_{\boldsymbol{\theta}(\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\mathbf{z}(\ell)\right)$.

For the MS tracking problem which is considered as tracking the state of a dynamical system, the PCRLB is calculated by considering the evolution of the state vectors, the measurements and the a priori distribution of the initial state vector. To this end, let $\boldsymbol{\Theta}(\ell) = (\boldsymbol{\theta}(0) \cdots \boldsymbol{\theta}(\ell))$ and $\mathbf{Z}(\ell) = (\mathbf{z}(0) \cdots \mathbf{z}(\ell))$ denote the sequences of the state and measurement vectors, respectively, up to time instance $\ell$. The initial state vector $\boldsymbol{\theta}(0)$ has a priori pdf $\mathrm{p}\left(\boldsymbol{\theta}(0)\right)$. Furthermore, it is assumed that the a priori pdf $\mathrm{p}\left(\mathbf{z}(0), \boldsymbol{\theta}(0)\right)$ equals $\mathrm{p}\left(\boldsymbol{\theta}(0)\right)$. The expression for the joint probability density function of the pair $(\boldsymbol{\Theta}(\ell), \mathbf{Z}(\ell))$ can be simplified using (4.5) and (4.9) as

$$\mathrm{p}\left(\boldsymbol{\Theta}(\ell), \mathbf{Z}(\ell)\right) = \prod_{i=1}^{\ell} \mathrm{p}\left(\mathbf{z}(i)|\boldsymbol{\theta}(i)\right) \prod_{j=1}^{\ell} \mathrm{p}\left(\boldsymbol{\theta}(j)|\boldsymbol{\theta}(j-1)\right) \cdot \mathrm{p}\left(\boldsymbol{\theta}(0)\right). \tag{4.29}$$

Thus the PCRLB of the parameter $\boldsymbol{\Theta}(\ell)$ can be computed from the Fisher information matrix of the pdf of (4.29). Furthermore, using the matrix inversion lemma, it is possible to recursively compute the PCRLB for each vector parameter $\boldsymbol{\theta}(\ell)$ at time instant $\ell$ instead of finding the PCRLB for the whole sequence of observations $\boldsymbol{\Theta}(\ell)$ in a single step [TMN98]. The recursive approach is easier to compute and helps to analyze the convergence properties of the PCRLB.

It has been assumed that the state transition and the measurement noises are multivariate Gaussian random variables. For Gaussian state transition and measurement noises, computation of the PCRLB involves the inversion of the covariance matrix of the state transition noise $\boldsymbol{\Gamma} \cdot \mathbf{u}_{\mathrm{ms}}(\ell)$ and the measurement noise $\mathbf{w}(\ell)$. From (4.4) and (4.5), the covariance matrix of the state transition noise $\boldsymbol{\Gamma} \cdot \mathbf{u}_{\mathrm{ms}}(\ell)$ can be calculated as

$$\mathbb{E}\left\{\boldsymbol{\Gamma} \cdot \mathbf{u}_{\mathrm{ms}}(\ell) \cdot \left(\boldsymbol{\Gamma} \cdot \mathbf{u}_{\mathrm{ms}}(\ell)\right)^{\mathrm{T}}\right\} = \boldsymbol{\Gamma} \cdot \mathbf{R}_{\mathrm{uu}} \cdot \boldsymbol{\Gamma}^{\mathrm{T}}. \tag{4.30}$$

For $\boldsymbol{\Gamma}$ defined in (4.7), i.e.,

$$\boldsymbol{\Gamma} = \begin{pmatrix} \Delta t & 0 \\ 0 & \Delta t \\ 1 & 0 \\ 0 & 1 \\ \mathbf{0}_{2P\times 1} & \mathbf{0}_{2P\times 1} \end{pmatrix},$$

the state transition noise covariance matrix is singular. Thus the computation of the Fisher information matrix of the pdf given in (4.29) fails. This calls for a modification of the computation of the PCRLB. In [TMN98], a systematic approach to computing the PCRLB for such cases has been proposed. In the following, the proposed approach is used to derive the PCRLB for the MS tracking problem. The proposed approach works

by separating the vector parameter in block form as $\boldsymbol{\theta}(\ell) = (\boldsymbol{\theta}_1^{\mathrm{T}}(\ell), \boldsymbol{\theta}_2^{\mathrm{T}}(\ell))^{\mathrm{T}}$ where $\boldsymbol{\theta}_1(\ell)$ denotes the part of the vector parameter which is driven by the system driving noise and $\boldsymbol{\theta}_2(\ell)$ denotes the part of the vector parameter which does not evolve with the system dynamics and/or evolves as a result of its relation with $\boldsymbol{\theta}_1(\ell)$. For the considered MS tracking problem, the scatterers are assumed stationary and the movement of the MS is modelled by making random changes to the velocity of the MS which is assumed to be constant within the tracking interval. Consequently, $\boldsymbol{\theta}_1(\ell) = \mathbf{v}_{\mathrm{ms}}(\ell)$ and $\boldsymbol{\theta}_2(\ell) = (\mathbf{p}_{\mathrm{ms}}^{\mathrm{T}}(\ell), \mathbf{p}_{\mathrm{sc},1}^{\mathrm{T}}, \ldots, \mathbf{p}_{\mathrm{sc},P}^{\mathrm{T}})^{\mathrm{T}}$. Hence the state transition equation given in (4.5) is modified as

$$\boldsymbol{\theta}_1(\ell) = \boldsymbol{\theta}_1(\ell-1) + \mathbf{u}_{\mathrm{ms}}(\ell) \tag{4.31}$$

$$\boldsymbol{\theta}_2(\ell) = \boldsymbol{\theta}_2(\ell-1) + \tilde{\mathbf{B}} \cdot \boldsymbol{\theta}_1(\ell), \tag{4.32}$$

where

$$\tilde{\mathbf{B}} = \begin{pmatrix} \Delta t & 0 \\ 0 & \Delta t \\ \mathbf{0}_{2P\times 1} & \mathbf{0}_{2P\times 1} \end{pmatrix}. \tag{4.33}$$

The measurement vector $\mathbf{z}(\ell)$ consists of the path lengths, the angles-of-departure and the angles-of-arrival of the propagation paths which are not functions of the velocity of the MS. Since $\boldsymbol{\theta}_1(\ell)$ represents only the velocity of the MS, the measurement equation given in (4.9) is modified as

$$\mathbf{z}(\ell) = \mathbf{g}(\boldsymbol{\theta}_2(\ell)) + \mathbf{w}(\ell). \tag{4.34}$$

The calculation of the PCRLB for the modified MS tracking problem is carried out by subsequent recursive calculation of the Fisher information matrices (FIMs) of the pdfs

$$\mathrm{p}(\ell) = \mathrm{p}(\boldsymbol{\Theta}_1(\ell-1), \boldsymbol{\theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \mathbf{Z}(\ell)) \tag{4.35}$$

$$\tilde{\mathrm{p}}(\ell+1) = \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1), \mathbf{Z}(\ell+1)) \tag{4.36}$$

$$\mathrm{p}(\ell+1) = \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_1(\ell+1), \boldsymbol{\theta}_2(\ell+1), \mathbf{Z}(\ell+1)). \tag{4.37}$$

The FIM of the pdf $\mathrm{p}(\ell)$ can be used to calculate the FIM of the pdf $\tilde{\mathrm{p}}(\ell+1)$ by relating the two pdfs as follows:

$$\begin{aligned} \tilde{\mathrm{p}}(\ell+1) &= \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1), \mathbf{Z}(\ell+1)) \\ &= \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \mathbf{Z}(\ell)) \cdot \mathrm{p}(\boldsymbol{\theta}_1(\ell+1), \mathbf{z}(\ell+1) | \boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \mathbf{Z}(\ell)) \\ &= \mathrm{p}(\ell) \cdot \mathrm{p}(\boldsymbol{\theta}_1(\ell+1) | \boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \mathbf{Z}(\ell)) \\ &\quad \cdot \mathrm{p}(\mathbf{z}(\ell+1) | \boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1), \mathbf{Z}(\ell)). \end{aligned} \tag{4.38}$$

Using (4.31), (4.32) and (4.34) one obtains

$$\tilde{\mathrm{p}}(\ell+1) = \mathrm{p}(\ell) \cdot \mathrm{p}(\boldsymbol{\theta}_1(\ell+1) | \boldsymbol{\theta}_1(\ell)) \cdot \mathrm{p}(\mathbf{z}(\ell+1) | \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)). \tag{4.39}$$

Furthermore, the FIM of the pdf $\tilde{\mathrm{p}}(\ell+1)$ can be used to calculate the FIM of the pdf $\mathrm{p}(\ell+1)$ by using the rule for change of coordinates of estimated parameters. The coordinates of the pdfs $\tilde{\mathrm{p}}(\ell+1) = \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1), \mathbf{Z}(\ell+1))$ and $\mathrm{p}(\ell+1) = \mathrm{p}(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_1(\ell+1), \boldsymbol{\theta}_2(\ell+1), \mathbf{Z}(\ell+1))$ are related, using (4.32), as

$$\begin{pmatrix} \mathrm{vec}(\boldsymbol{\Theta}_1(\ell-1)) \\ \boldsymbol{\theta}_1(\ell) \\ \boldsymbol{\theta}_1(\ell+1) \\ \boldsymbol{\theta}_2(\ell+1) \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{I}_{2\ell} & \mathbf{0}_{2\ell\times 2} & \mathbf{0}_{2\ell\times 2(P+1)} & \mathbf{0}_{2\ell\times 2} \\ \mathbf{0}_{2\times 2\ell} & \mathbf{I}_2 & \mathbf{0}_{2\times 2(P+1)} & \mathbf{0}_2 \\ \mathbf{0}_{2\times 2\ell} & \mathbf{0}_2 & \mathbf{0}_{2\times 2(P+1)} & \mathbf{I}_2 \\ \mathbf{0}_{2(P+1)\times 2\ell} & \tilde{\mathbf{B}} & \mathbf{I}_{2(P+1)} & \mathbf{0}_{2(P+1)\times 2} \end{pmatrix}}_{=\tilde{\boldsymbol{\Upsilon}}} \begin{pmatrix} \mathrm{vec}(\boldsymbol{\Theta}_1(\ell-1)) \\ \boldsymbol{\theta}_1(\ell) \\ \boldsymbol{\theta}_2(\ell) \\ \boldsymbol{\theta}_1(\ell+1) \end{pmatrix}, \tag{4.40}$$

where $\mathrm{vec}(\cdot)$ is a matrix vectorization operator which transforms a matrix into a vector by stacking the columns of the matrix one underneath the other. Obviously, the matrix $\tilde{\boldsymbol{\Upsilon}}$ is nonsingular. For a nonsingular matrix $\tilde{\boldsymbol{\Upsilon}}$ the FIMs of the pdfs $\tilde{\mathrm{p}}(\ell+1)$ and $\mathrm{p}(\ell+1)$ are related as [LC98]

$$\mathbf{J}\left(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_1(\ell+1), \boldsymbol{\theta}_2(\ell+1)\right) = (\tilde{\boldsymbol{\Upsilon}}^{-1})^{\mathrm{T}} \cdot \mathbf{J}\left(\boldsymbol{\Theta}_1(\ell), \boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)\right) \cdot \tilde{\boldsymbol{\Upsilon}}^{-1}. \tag{4.41}$$

Thus it remains to make use of the discussion above to recursively calculate the Fisher information matrix (FIM) of the state vector $\boldsymbol{\theta}(\ell)$. To this end, let $\mathbf{J}(\ell)$ be the FIM of the state vector $\left(\boldsymbol{\theta}_1^{\mathrm{T}}(\ell), \boldsymbol{\theta}_2^{\mathrm{T}}(\ell)\right)^{\mathrm{T}}$. The decomposition of the matrix $\mathbf{J}(\ell)$ into blocks corresponding to the vectors $\boldsymbol{\theta}_1(\ell)$ and $\boldsymbol{\theta}_2(\ell)$ reads

$$\mathbf{J}(\ell) = \begin{pmatrix} \mathbf{J}_{11}(\ell) & \mathbf{J}_{12}(\ell) \\ \mathbf{J}_{21}(\ell) & \mathbf{J}_{22}(\ell) \end{pmatrix}. \tag{4.42}$$

Furthermore, let $\mathbf{Q}(\ell)$ be the FIM of the vector $\left(\boldsymbol{\theta}_1^{\mathrm{T}}(\ell-1), \boldsymbol{\theta}_1^{\mathrm{T}}(\ell), \boldsymbol{\theta}_2^{\mathrm{T}}(\ell)\right)^{\mathrm{T}}$. The decomposition of the matrix $\mathbf{Q}(\ell)$ into blocks corresponding to the vectors $\boldsymbol{\theta}_1(\ell-1)$, $\boldsymbol{\theta}_1(\ell)$ and $\boldsymbol{\theta}_2(\ell)$ reads

$$\mathbf{Q}(\ell) = \begin{pmatrix} \mathbf{Q}_{11}(\ell) & \mathbf{Q}_{12}(\ell) & \mathbf{Q}_{13}(\ell) \\ \mathbf{Q}_{21}(\ell) & \mathbf{Q}_{22}(\ell) & \mathbf{Q}_{23}(\ell) \\ \mathbf{Q}_{31}(\ell) & \mathbf{Q}_{32}(\ell) & \mathbf{Q}_{33}(\ell) \end{pmatrix}. \tag{4.43}$$

Based on the discussion made in the previous paragraph, the FIM $\mathbf{J}(\ell)$ can be computed recursively using the FIM $\mathbf{Q}(\ell)$ as [TMN98]:

$$\mathbf{J}(\ell) = \begin{pmatrix} \mathbf{Q}_{22}(\ell) & \mathbf{Q}_{23}(\ell) \\ \mathbf{Q}_{32}(\ell) & \mathbf{Q}_{33}(\ell) \end{pmatrix} - \begin{pmatrix} \mathbf{Q}_{21}(\ell) \\ \mathbf{Q}_{31}(\ell) \end{pmatrix} \mathbf{Q}_{11}^{-1}(\ell) \begin{pmatrix} \mathbf{Q}_{12}(\ell) & \mathbf{Q}_{13}(\ell) \end{pmatrix} \tag{4.44}$$

$$\mathbf{Q}(\ell+1) = (\boldsymbol{\Upsilon}^{-1})^{\mathrm{T}} \cdot \begin{pmatrix} \mathbf{J}_{11}(\ell) + \mathbf{H}_{11}(\ell) & \mathbf{J}_{12}(\ell) + \mathbf{H}_{12}(\ell) & \mathbf{H}_{13}(\ell) \\ (\mathbf{J}_{12}(\ell) + \mathbf{H}_{12}(\ell))^{\mathrm{T}} & \mathbf{J}_{22}(\ell) + \mathbf{H}_{22}(\ell) & \mathbf{H}_{23}(\ell) \\ \mathbf{H}_{13}^{\mathrm{T}}(\ell) & \mathbf{H}_{23}^{\mathrm{T}}(\ell) & \mathbf{H}_{33}(\ell) \end{pmatrix} \cdot \boldsymbol{\Upsilon}^{-1}, \tag{4.45}$$

where

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0}_{2\times 2(P+1)} & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_{2\times 2(P+1)} & \mathbf{I}_2 \\ \tilde{\mathbf{B}} & \mathbf{I}_{2(P+1)} & \mathbf{0}_{2(P+1)\times 2} \end{pmatrix} \tag{4.46}$$

$$\mathbf{H}_{11}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_1(\ell) \partial \boldsymbol{\theta}_1^{\mathrm{T}}(\ell)} \right) \right\} \tag{4.47}$$

$$\mathbf{H}_{12}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_1(\ell) \partial \boldsymbol{\theta}_2^{\mathrm{T}}(\ell)} \right) \right\} \tag{4.48}$$

$$\mathbf{H}_{13}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_1(\ell) \partial \boldsymbol{\theta}_1^{\mathrm{T}}(\ell+1)} \right) \right\} \tag{4.49}$$

$$\mathbf{H}_{22}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_2(\ell) \partial \boldsymbol{\theta}_2^{\mathrm{T}}(\ell)} \right) \right\} \tag{4.50}$$

$$\mathbf{H}_{23}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_2(\ell) \partial \boldsymbol{\theta}_1^{\mathrm{T}}(\ell+1)} \right) \right\} \tag{4.51}$$

$$\mathbf{H}_{33}(\ell) = \mathbb{E}_{\mathbf{z}(\ell+1),\boldsymbol{\theta}(\ell+1)} \left\{ - \left( \frac{\partial^2 \ln(\bar{\mathrm{p}})}{\partial \boldsymbol{\theta}_1(\ell+1) \partial \boldsymbol{\theta}_1^{\mathrm{T}}(\ell+1)} \right) \right\} \tag{4.52}$$

$$\bar{\mathrm{p}} = \mathrm{p}(\boldsymbol{\theta}_1(\ell+1)|\boldsymbol{\theta}_1(\ell)) \cdot \mathrm{p}(\mathbf{z}(\ell+1)|\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)). \tag{4.53}$$

The partial derivatives of $\ln(\mathrm{p}(\boldsymbol{\theta}_1(\ell+1)|\boldsymbol{\theta}_1(\ell)) \cdot \mathrm{p}(\mathbf{z}(\ell+1)|\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)))$ can be evaluated using the modified MS tracking problem given in (4.31), (4.32) and (4.34). From (4.4) and (4.31) the expression for the pdf $\mathrm{p}(\boldsymbol{\theta}_1(\ell+1)|\boldsymbol{\theta}_1(\ell))$ is given as

$$\ln\left(\mathrm{p}\left(\boldsymbol{\theta}_1(\ell+1)|\boldsymbol{\theta}_1(\ell)\right)\right) = c_1 - \frac{1}{2}\left(\boldsymbol{\theta}_1(\ell+1) - \boldsymbol{\theta}_1(\ell)\right)^{\mathrm{T}} \mathbf{R}_{\mathrm{uu}}^{-1} \left(\boldsymbol{\theta}_1(\ell+1) - \boldsymbol{\theta}_1(\ell)\right), \tag{4.54}$$

where $c_1$ is a constant. On the other hand, from (4.10), (4.32) and (4.34)

$$\begin{aligned} \ln\left(\mathrm{p}\left(\mathbf{z}(\ell+1)|\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)\right)\right) = c_2 - \frac{1}{2}&\left(\mathbf{z}(\ell+1) - \mathbf{g}(\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1))\right)^{\mathrm{T}} \\ &\cdot \mathbf{R}_{\mathrm{ww}}^{-1} \left(\mathbf{z}(\ell+1) - \mathbf{g}(\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1))\right), \end{aligned} \tag{4.55}$$

where $c_2$ is a constant. Hence using (4.54) and (4.55) the partial derivatives of $\ln(\mathrm{p}(\boldsymbol{\theta}_1(\ell+1)|\boldsymbol{\theta}_1(\ell)) \cdot \mathrm{p}(\mathbf{z}(\ell+1)|\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1)))$ can be calculated as

$$\mathbf{H}_{11}(\ell) = \mathbf{R}_{\mathrm{uu}}^{-1} \tag{4.56}$$

$$\mathbf{H}_{12}(\ell) = \mathbf{0}_{2\times 2(P+1)} \tag{4.57}$$

$$\mathbf{H}_{13}(\ell) = -\mathbf{R}_{\mathrm{uu}}^{-1} \tag{4.58}$$

$$\mathbf{H}_{22}(\ell) = \left( \frac{\partial \mathbf{g}(\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1))}{\partial \boldsymbol{\theta}_2(\ell)} \right)^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{ww}}^{-1} \cdot \left( \frac{\partial \mathbf{g}(\boldsymbol{\theta}_2(\ell), \boldsymbol{\theta}_1(\ell+1))}{\partial \boldsymbol{\theta}_2(\ell)} \right) \tag{4.59}$$

$$\mathbf{H}_{23}(\ell) = \mathbf{0}_{2(P+1)\times 2} \tag{4.60}$$

$$\mathbf{H}_{33}(\ell) = \mathbf{R}_{\mathrm{uu}}^{-1}. \tag{4.61}$$

The above Hessian matrices were derived in a similar manner as in (3.42). The Jacobian matrix $\frac{\partial \mathbf{g}(\boldsymbol{\theta}_2(\ell))}{\partial \boldsymbol{\theta}_2(\ell)}$ can be calculated using the expressions given in Appendix A.

The initial recursion Fisher information matrix (FIM) $\mathbf{J}(0)$ is defined by using the Jacobian of the a priori pdf $\mathrm{p}(\boldsymbol{\theta}_1(0), \boldsymbol{\theta}_2(0), \mathbf{z}(0)) = \mathrm{p}(\boldsymbol{\theta}_1(0), \boldsymbol{\theta}_2(0))$. From (4.8), the

initial state vector is a multivariate Gaussian distributed random variable and hence

$$\mathbf{J}(0) = \mathbf{R}_{\theta\theta}^{-1}(0). \tag{4.62}$$

The Fisher information matrix (FIM) $\mathbf{J}(\ell)$ represents the FIM of the positions of the MS and scatterers and the velocity of the MS. Since only the positions of the MS are of interest, the PCRLB of estimates of the positions of the MS is required. This can be obtained by partitioning the FIM $\mathbf{J}(\ell)$ into block submatrices and using the matrix inversion lemma as shown in Section 3.4.1.

### 4.3.2 Simulation results

In this section, the performance of the proposed MS tracking algorithm is analysed considering a picocell NLOS multipath localization scenario using Monte Carlo simulations with $10^5$ independent trials. The radius of the cell is assumed to be 100 m and it is served by a single BS. The probability distribution of the initial position of the MS is assumed to be uniform within the cell. In the simulations, the MS to be tracked is assumed to be a pedestrian. Consequently, the initial velocity of the MS is generated randomly from the distribution $\mathbf{v}_{\text{ms}}(0) \sim \mathcal{N}\left(\mathbf{0}, \text{diag}\left(1\ \text{m}^2/\text{sec}^2, 1\ \text{m}^2/\text{sec}^2\right)\right)$. The positions of the scatterers are generated randomly using the single bounce elliptical model discussed in Section 3.4.2 for a normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}} = 2.5$. The MS trajectory consists of $L = 10$ positions with a tracking interval of $\Delta t = 0.5$ sec. The MS trajectory has been generated using the first order Gauss-Markov mobility model defined in (4.5). The state transition noise $\mathbf{u}_{\text{ms}}(\ell)$ in (4.5) has a covariance matrix $\mathbf{R}_{\text{uu}} = \text{diag}\left(0.1^2\ \text{m}^2/\text{sec}^2, 0.1^2\ \text{m}^2/\text{sec}^2\right)$.

As mentioned in Section 4.2, the initial estimate $\hat{\boldsymbol{\theta}}(0)$ of the state vector $\boldsymbol{\theta}(0)$ is obtained using initial MS and scatterer position estimates from the total least squares (TLS) localization algorithm. The initial estimates of the velocity of the MS are obtained using a zero mean Gaussian distribution, i.e., $\hat{\mathbf{v}}_{\text{ms}}(0) \sim \mathcal{N}\left(\mathbf{0}, \text{diag}\left(1\ \text{m}^2/\text{sec}^2, 1\ \text{m}^2/\text{sec}^2\right)\right)$. Furthermore, the initial covariance matrix $\mathbf{M}(0|0)$ of the state vector was set to $\text{diag}(\underbrace{10, \ldots, 10}_{4}, \underbrace{10^3, \ldots, 10^3}_{2P})$. As discussed in Section 4.2, the magnitude of the diagonal elements of the covariance matrix $\mathbf{M}(0|0)$ was set in such a way that its inverse reflects the confidence on the initial estimates of the parameter in question, i.e., the position of the MS, the velocity of the MS or the positions of the scatterers. For cases where the extended Kalman filter diverges, MS trajectory estimates obtained using the TLS localization algorithm are used as MS trajectory estimates of the extended Kalman filter.

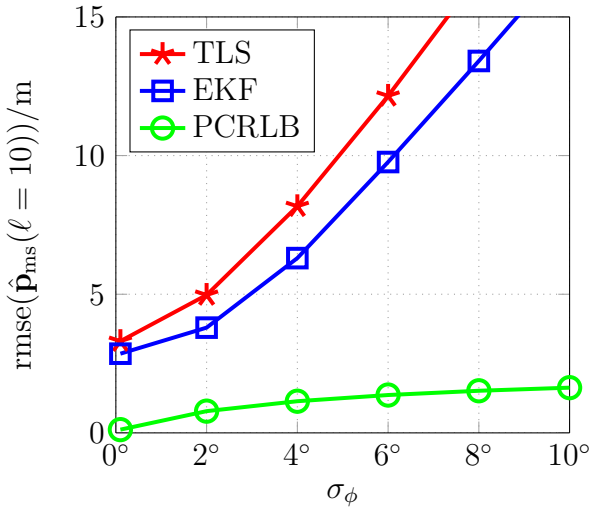
The extended Kalman filter is said to have diverged if $\|\hat{\mathbf{p}}_{\mathrm{ms}}(\ell) - \hat{\mathbf{p}}_{\mathrm{ms}}(\ell-1)\|_2 \geq 10$ m, where $\hat{\mathbf{p}}_{\mathrm{ms}}(\ell)$ is the estimated position of the MS at the $\ell^{\mathrm{th}}$ time instant. It has been observed that depending on the magnitude of the standard deviations of the measurement noises, the extended Kalman filter converges for 50–75 % of the simulation runs. It will be shown that the extended Kalman filter nevertheless yields significant performance gains over the TLS localization algorithm.

The performance metric is the root mean square error $\mathrm{rmse}(\hat{\mathbf{p}}_{\mathrm{ms}}(\ell))$ of the estimated position $\hat{\mathbf{p}}_{\mathrm{ms}}(\ell)$ of the MS which is calculated as

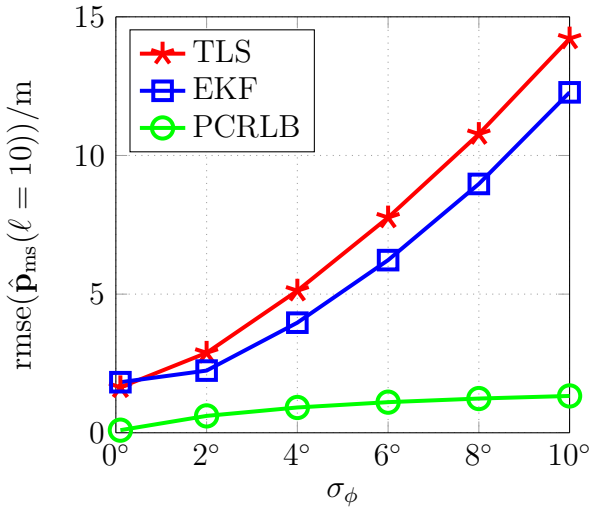
$$\mathrm{rmse}(\hat{\mathbf{p}}_{\mathrm{ms}}(\ell)) = \sqrt{\mathbb{E}_{\mathbf{z}(\ell),\mathbf{p}_{\mathrm{ms}}(\ell)}\left\{(\hat{\mathbf{p}}_{\mathrm{ms}}(\ell) - \mathbf{p}_{\mathrm{ms}}(\ell))^{\mathrm{T}}(\hat{\mathbf{p}}_{\mathrm{ms}}(\ell) - \mathbf{p}_{\mathrm{ms}}(\ell))\right\}}, \tag{4.63}$$

where $\mathbf{p}_{\mathrm{ms}}(\ell)$ is the true position of the MS at the $\ell^{\mathrm{th}}$ time instant and $\mathbb{E}_{\mathbf{z}(\ell),\mathbf{p}_{\mathrm{ms}}(\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\mathbf{z}(\ell), \mathbf{p}_{\mathrm{ms}}(\ell)\right)$. In the simulations, the expectations of (4.63) are computed using averages. To assess the performance gains obtained from tracking the MS, the performance of successive independent position estimates of the trajectory of the MS using the total least squares (TLS) localization algorithm is also considered. In the following, the root mean square errors $\mathrm{rmse}(\hat{\mathbf{p}}_{\mathrm{ms}}(\ell = 10))$ of the last position on the trajectory of the MS of the extended Kalman filter tracking, which is close to the steady state root mean square error, and the TLS localization algorithms are compared for different numbers of paths and noise powers. In all cases, the PCRLB is used as a performance benchmark. The initial recursion of the PCRLB is computed using the covariance matrix $\mathbf{R}_{\theta\theta}(0)$ of the initial state vector which is assumed to be a diagonal matrix with a variance of 1 $\mathrm{m}^2/\mathrm{sec}^2$ for the velocity and a variance of $50^2$ $\mathrm{m}^2$ for the positions of the MS and scatterers.

Fig. 4.1 shows the performances of the extended Kalman filter (EKF) tracking and the TLS localization algorithms versus the standard deviation $\sigma_\phi = \sigma_{\phi_{\mathrm{bs}},p} = \sigma_{\phi_{\mathrm{ms}},p}$ of the angles-of-departure and angles-of-arrival measurement noise for different numbers of propagation paths $P$. In the simulations, the standard deviation $\sigma_{\mathrm{d}}$ of the path lengths measurement noise was set to 5 m. It can be seen from Fig. 4.1 that the performances of the EKF tracking and the TLS localization algorithms are bounded by the PCRLB. The performances of the EKF tracking and the TLS localization algorithms approach the PCRLB when $\sigma_\phi$ is close to zero. Furthermore, it can be concluded from the superior performance of the EKF tracking algorithm over the TLS localization algorithm that there is an appreciable performance gain obtained from tracking the MS. Moreover, there are significant performance improvements by both the EKF tracking algorithm and the TLS localization algorithm as the number of propagations paths $P$ is increased owing to the extra information about the position of the MS from the extra propagation paths. This shows that the NLOS multipath propagations are effectively exploited as an aid in tracking the MS.



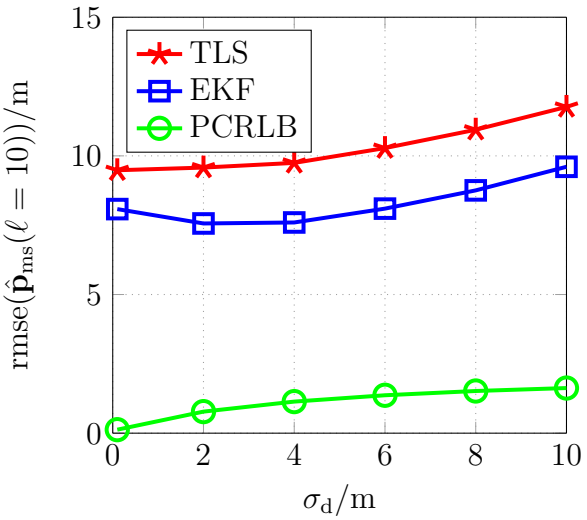

(a) Number of paths $P = 4$



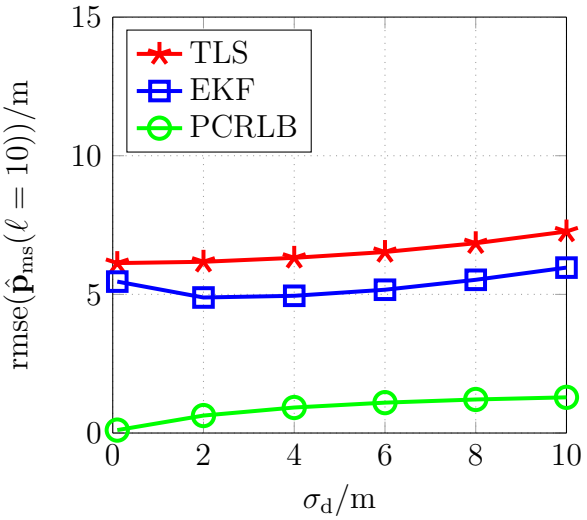

(b) Number of paths $P = 6$

Figure 4.1. Influence of the number of paths and the standard deviation of the angles-of-departure and angles-of-arrival measurement noise on the proposed MS tracking algorithm, $\sigma_{\mathrm{d}} = 5$ m

Fig. 4.2 shows the performances of the EKF tracking and the TLS localization algorithms versus the standard deviation $\sigma_{\mathrm{d}}$ of the path lengths measurement noise for different numbers of propagation paths $P$. In the simulations, the standard deviations $\sigma_{\phi_{\mathrm{bs}}}$ and $\sigma_{\phi_{\mathrm{ms}}}$ of the angles-of-departure and angles-of-arrival measurement noises were set to 5°. It can be seen from Fig. 4.2 that the performances of the EKF tracking and the TLS localization algorithms are bounded by the PCRLB. Furthermore, similar to the simulation results shown in Fig. 4.1, the EKF tracking algorithm results in an appreciable performance improvement over the TLS localization algorithm. Moreover,




(a) Number of paths $P = 4$




(b) Number of paths $P = 6$

Figure 4.2. Influence of the number of paths and the standard deviation of the path lengths measurement noise on the proposed MS tracking algorithm, $\sigma_\phi = 5°$ (figure drawn not to scale)

confirming the exploitation of the NLOS propagation paths for localization and tracking, there are significant performance improvements by both the EKF tracking and the TLS localization algorithms as the number of propagations paths $P$ is increased.

It can be seen from Fig. 4.1 and Fig. 4.2 that generally there is only a minor improvement in performance brought about by the EKF tracking algorithm over the TLS localization algorithm when $\sigma_{\phi_{\mathrm{ms}}}$ and $\sigma_{\mathrm{d}}$ are near zero, respectively. This occurs owing to the fact that for such cases, the standard deviations $\sigma_{\phi_{\mathrm{ms}}}$ and $\sigma_{\mathrm{d}}$ of the

measurement noises are comparable with the standard deviation of the state transition noise. As mentioned earlier, the state transition noise has a covariance matrix $\mathbf{R}_{\text{uu}} = \text{diag}\,(0.1^2\ \text{m}^2/\text{sec}^2, 0.1^2\ \text{m}^2/\text{sec}^2)$. This means that the predicted state vector and the measurement vector have comparable reliability and the Kalman filter does not significantly outweigh the predicted state vector over the measurement vector. This reduces the performance gain that is obtained by exploiting the dependency between the successive positions of the MS.

Owing to the reliance of the extended Kalman filter on linearization of the nonlinear measurement function, estimates of the initial state vector play a crucial role in the convergence and accuracy of the extended Kalman filter. The extended Kalman filter obtains estimates of the initial state vector using estimates of the TLS localization algorithm. The accuracy of estimates of the initial position of the MS can be increased by considering more NLOS propagation paths. However, the increase in the accuracy of the estimates of the initial position of the MS from consideration of more NLOS propagation paths is accompanied by an increase in the number of scatterers whose positions are also to be tracked by the extended Kalman filter. For each considered scatter, three measured propagation path parameters are obtained while the dimension of the state vector is increased by two for the position of the considered scatter. Thus the reliability of estimates of the initial state vector does not increase dramatically when considering more NLOS propagation paths. Nevertheless, significant gains can be obtained when considering more NLOS propagation paths as shown in Figs. 4.1 and 4.2.

It must be noted that the simulations have been performed for synchronized BS and MS. It has been observed that, unlike MS localization for unsynchronized BS and MS, MS localization for synchronized BS and MS does not result in an unstable TLS problem. Thus the simulations were carried out without having to mitigate such cases. Furthermore, there was no need to consider a high number of propagation paths to obtain a well posed TLS problem. This is due to the fact that estimates of the propagation path lengths obtained from synchronized BS and MS yield more information about the position of a MS than estimates of the propagation path length differences obtained from unsynchronized BS and MS. Consequently, for synchronized BS and MS at least two NLOS propagation paths are required for MS localization unlike for unsynchronized BS and MS where at least three NLOS propagation path are required for MS localization.

# Chapter 5

# Network localization in NLOS multipath environments

## 5.1 Cooperative localization

In the previous chapters, the problem of localizing and tracking a MS in NLOS multipath environments was considered. It has been shown that while a single BS is adequate for localizing and tracking a MS, additional BSs could be considered to increase the accuracy of the estimated positions of a MS. In this chapter, a network localization scenario is considered where there are at least two MSs and two BSs. The unknown positions of the MSs are to be estimated using the known positions of the BSs. Network localization plays a key role in sensor networks where the knowledge of the locations of the sensors is essential to make a meaningful interpretation of the sensed data.

In a network localization scenario, the positions of the MSs can be estimated using either a centralized or a distributed localization. In centralized localization, a central processing unit collects all measurements and estimates the positions of the MSs. Centralized localization algorithms have the downside of not being scalable and hence don't lend themselves to applications in large scale networks [WLW09]. In distributed localization, each MS estimates its own position based on its own acquired measurements. In some cases, the MSs may iteratively update their estimated positions by exchanging the processed measurement data with neighbouring MSs. In practice, distributed localization algorithms have inferior performance over centralized localization algorithms owing to the exchange of processed measurement data, rather than the raw measurement data, which may suffer from information loss during processing by each MS. Furthermore, distributed localization algorithms may require a substantial number of iterations to have comparable performance to centralized localization algorithms. In [SCC+14], it has been shown that after infinite iterations a distributed localization algorithm has the same performance bound as a centralized localization algorithm. In the following, a centralized localization algorithm is considered for network localization in NLOS multipath environments.

In a network localization, the MSs can cooperate with each other to determine their positions [WLW09] and/or to obtain a significant improvement in localization accuracy [MBW+04]. The MSs cooperate by directly communicating with each other to

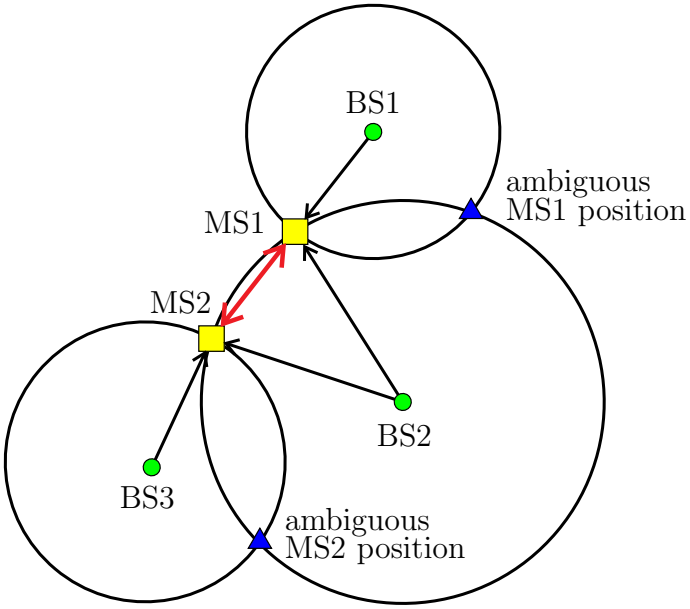



Figure 5.1. Cooperative localization can be used to resolve ambiguity in the positions of the MSs by considering the distance between the MSs [WLW09]

obtain estimates of the propagation path parameters between the MSs. The measurements obtained from consideration of the MS-MS links can be used to complement the measurements from the BS-MS links so that it is possible to localize MSs where the measurements from the BS-MS links are not sufficient for localization, see Fig. 5.1. Thus cooperation between the MSs can be used to increase localization coverage. Furthermore, the measurements obtained from cooperation between the MSs can be used to improve localization accuracy by exploiting the information about the relative positions of the MSs.

In [WLW09] and [MBW+04], cooperation between the MSs has been exploited to increase localization coverage and accuracy for network localization in LOS environments, respectively. In this chapter, network localization in NLOS multipath environments is of interest. The idea of using cooperation between the MSs to increase localization coverage and accuracy can also be applied in NLOS multipath environments. The increased localization accuracy obtained as a result of the cooperation between the MSs is especially important in NLOS multipath environments. Cooperative localization algorithms for network localization in LOS environments can be applied for network localization in NLOS environments. However, the NLOS multipath propagations are detrimental to the localization performance of localization algorithms developed for LOS environments [GC09, SW14b]. Consequently, new approaches to network localization in NLOS multipath environments are required. In this chapter, two new approaches to cooperative localization in NLOS multipath environments are proposed.

## 5.2 Cooperative localization in NLOS multipath environments under the explicit consideration of scatterers

The potential performance gain from cooperation between MSs can be demonstrated for network localization in NLOS multipath environments under the explicit consideration of scatterers. Towards this end, consider a network consisting of $N_{\mathrm{bs}}$ BSs and $N_{\mathrm{ms}}$ MSs. For ease of exposition, let the number of NLOS propagation paths between each BS-MS and MS-MS links be $P$. Furthermore, let $C$ denote the number of measured propagation path parameters for each BS-MS and MS-MS links, i.e., the times-of-arrival, the angles-of-departure, the angles-of-arrival or a combination thereof. It is assumed that each MS cooperates with all other MSs, i.e., there is full connectivity between the MSs. Furthermore, it is assumed that there is full connectivity between the BSs and the MSs. For such a network localization scenario, the number of unknown parameters, i.e., the positions of the MSs and the scatterers, is

$$2N_{\mathrm{ms}} + 2 \cdot P \cdot N_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + 2 \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2},$$

whereas the number of measured propagation path parameters is

$$C \cdot P \cdot N_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + C \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}.$$

The terms $2 \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}$ and $C \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}$ denote the extra unknown parameters and the extra measured propagation path parameters incurred due to the consideration of the MS-MS connections, respectively. If $C \geq 3$, then the number of measured path parameters, i.e., the number of equations, is greater than the number of unknown parameters and hence localization of the MSs is possible. Furthermore, if $C \geq 3$, then consideration of the MS-MS links results in an increase in the number of unknown parameters by $2 \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}$, while the increase in the number of measured propagation path parameters is $C \cdot P\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}$. Thus consideration of the MS-MS links results in a larger increase of the number of measured propagation path parameters than the number of unknown parameters. Consequently, the MS-MS links can be considered to obtain significant performance gain in the accuracy of the estimated positions of the MSs.

In Section 3.3, an algorithm for MS localization in NLOS multipath environments under the explicit consideration of scatterers has been proposed. The proposed MS localization algorithm exploits the linear relations between the positions of the MS, the BS and the scatterers to determine the position of the MS. The proposed localization algorithm can be easily extended to network localization in NLOS multipath environments.

This results in a geometric cooperative localization algorithm for network localization in NLOS multipath environments which determines the positions of the MSs under the explicit consideration of scatterers.

For cooperative localization in NLOS multipath environments under the explicit consideration of scatterers, the increase in the number of unknown parameters and the number of measured propagation path parameters owing to consideration of the MS-MS links is quadratic in the number of MSs. Thus even though there is a significant performance gain from consideration of the MS-MS links, the comparable increase in the number of unknown parameters and the number of measured path parameters limits the performance gain from consideration of the MS-MS links. Furthermore, the quadratic increase in the number of unknown parameters could result in a significant computational burden on network localization algorithms which would have to explicitly estimate the positions of the scatterers, for example a maximum-likelihood estimation algorithm. Nevertheless, it must be kept in mind that a satisfactory network localization performance can be guaranteed for all NLOS multipath network localization scenarios only under the explicit consideration of scatterers.

## 5.3 Cooperative localization in NLOS multipath environments without the explicit consideration of scatterers

Had it not been for the explicit consideration of scatterers, the increase in the number of measurements which is quadratic in the number of the MSs would have not been accompanied by an increase in the number of unknown parameters. Hence if scatterers are not explicitly considered by a network localization algorithm, the number of unknown parameters, i.e., the positions of the MSs, is $2N_{\mathrm{ms}}$, whereas the number of measured propagation path parameters is

$$CN_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + C\frac{N_{\mathrm{ms}}(N_{\mathrm{ms}}-1)}{2}.$$

Thus a network localization algorithm that does not explicitly consider scatterers would strongly benefit from the number of measurements which grow quadratically with the number of the MSs without being accompanied by an increase in the number of unknowns. However, without the explicit consideration of scatterers, the NLOS multipath propagations become detrimental to the performance of the network localization algorithm. To mitigate this, the straightforward approach would be to use LOS network

localization algorithms under the assumption that the NLOS multipath propagations cause erroneous path parameter measurements. The extra measurements obtained from cooperation can help to "average out" the influence of the NLOS propagations. However, this approach is susceptible to the proportion of the LOS and NLOS propagation paths; hence it does not yield a satisfactory localization performance [SW14b]. A more subtle approach to exploiting the extra measurements can potentially yield a satisfactory localization performance. In the following, a systematic approach to network localization in NLOS multipath environments is proposed without the explicit consideration of scatterers.

Since the scatterers are not explicitly considered, it is necessary to consider models which describe the influence of the NLOS propagations. Different models have been considered to describe the effect of the NLOS propagations on the measured distances [GC09] and directions of arrivals [ECS+98, SJJS00]. In the following, the measured distance between the BS-MS and the MS-MS communication links is considered for simplicity without loss of any generality of the cooperative localization algorithm to be proposed, hence $C = 1$ holds. Consideration of only the times-of-arrival of the propagation paths for the network localization enables the cooperative localization algorithm to be proposed to be applied also for sensor networks where the sensor nodes in general do not possess antenna arrays which are required to estimate the directions of arrival.

Let the position of the $n_{\mathrm{bs}}^{\mathrm{th}}$ BS, $n_{\mathrm{bs}} = 1, \ldots, N_{\mathrm{bs}}$ be denoted by $\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}})$, whereas the position of the $n_{\mathrm{ms}}^{\mathrm{th}}$ MS, $n_{\mathrm{ms}} = 1, \ldots, N_{\mathrm{ms}}$ be denoted by $\mathbf{p}_{\mathrm{ms}}(n_{\mathrm{ms}})$. For the network localization scenario with $N_{\mathrm{bs}}$ BSs and $N_{\mathrm{ms}}$ MSs, there are

$$N_{\mathrm{d}} = N_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + \frac{N_{\mathrm{ms}}(N_{\mathrm{ms}} - 1)}{2} \tag{5.1}$$

measured distances considering the BS-MS and MS-MS links. Let $\tilde{d}(n_{\mathrm{d}})$, $n_{\mathrm{d}} = 1, \ldots, N_{\mathrm{d}}$ denote the measured distances of the BS-MS and the MS-MS links. The measured distance $\tilde{d}(n_{\mathrm{d}})$, $\forall n_{\mathrm{d}} = (n_{\mathrm{bs}} - 1)N_{\mathrm{ms}} + n_{\mathrm{ms}} \in \{1, \ldots, N_{\mathrm{bs}} \cdot N_{\mathrm{ms}}\}$ of the BS-MS links is given by

$$\tilde{d}(n_{\mathrm{d}}) = \|\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}}) - \mathbf{p}_{\mathrm{ms}}(n_{\mathrm{ms}})\|_2 + w(n_{\mathrm{d}}), \tag{5.2}$$

whereas the measured distance $\tilde{d}(n_{\mathrm{d}})$, $\forall n_{\mathrm{d}} = N_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + (n_{\mathrm{ms}} - 1)N_{\mathrm{ms}} - \frac{1}{2}n_{\mathrm{ms}}(n_{\mathrm{ms}} + 1) + \acute{n}_{\mathrm{ms}} \in \{N_{\mathrm{bs}} \cdot N_{\mathrm{ms}} + 1, \ldots, N_{\mathrm{d}}\}$ of the MS-MS links, i.e., the $n_{\mathrm{ms}}^{\mathrm{th}}$ and the $\acute{n}_{\mathrm{ms}}^{\mathrm{th}}$ MSs, $\acute{n}_{\mathrm{ms}} > n_{\mathrm{ms}}$, $\acute{n}_{\mathrm{ms}} \in \{1 \ldots N_{\mathrm{ms}}\}$, is given by

$$\tilde{d}(n_{\mathrm{d}}) = \|\mathbf{p}_{\mathrm{ms}}(n_{\mathrm{ms}}) - \mathbf{p}_{\mathrm{ms}}(\acute{n}_{\mathrm{ms}})\|_2 + w(n_{\mathrm{d}}). \tag{5.3}$$

$w(n_{\mathrm{d}})$ is the measurement noise.

In the previous chapters, the distance measurement noises were assumed to be independent and identically distributed zero mean Gaussian distributed random variables. There is no need to modify this assumption if the BS-MS and MS-MS links have unobstructed LOS propagation paths. However, if the BS-MS and MS-MS links are multipath links, then the measurement noises shall be modified to describe the effect of the NLOS propagations on the measured distances. Consequently, different distributions have been assumed for the NLOS propagation noises based on theoretical models and empirical models based on measurements [GC09]. In the following, the measurement noises for the NLOS propagation paths are assumed to be independent and identically distributed Gaussian noises for simplicity. The impact of the NLOS propagation path lengths which are longer than the LOS propagation path lengths is described by a positive distance bias $b$ whose value depends on the localization scenario under consideration. Furthermore, owing to the various possible NLOS propagation paths, the variance of the measured distances for the NLOS propagation paths are assumed to be significantly higher than the variance of the measured distances for the LOS propagation paths. Thus the measurement noise for the NLOS propagation paths are identical and independently distributed Gaussian noises with a non-zero mean described by the positive distance bias $b$ and a variance which is much larger than the measurement noises for the LOS propagation paths. Consequently, the measurement noise $w(n_\mathrm{d})$ can be described by a Gaussian mixture noise:

$$w(n_\mathrm{d}) \sim (1-\varepsilon)\mathcal{N}(0, \sigma^2_\mathrm{LOS}) + \varepsilon\mathcal{N}(b(n_\mathrm{d}), \sigma^2_\mathrm{NLOS}), \tag{5.4}$$

where $b(n_\mathrm{d})$ is the positive distance bias and the parameter $\varepsilon$ determines the fraction of the BS-MS and MS-MS links which are NLOS. The parameters $\sigma_\mathrm{LOS}$ and $\sigma_\mathrm{NLOS}$ denote the standard deviations of the LOS and the NLOS measurement noises, respectively.

The NLOS propagation paths, unlike the LOS propagation paths, result in distance measurements which tend to be much larger than the true distances between the BS-MS and the MS-MS links. Thus if the majority of the propagation paths are NLOS, then the estimated positions of the MSs will be biased by the measured distances from the NLOS propagations; hence satisfactory localization performance cannot be guaranteed. Consequently, as is often the case, in the following, the case where the majority of the links are LOS propagation links, i.e., $\varepsilon < 0.5$, is considered. In this case, the distance measurements from the NLOS propagation paths result in outliers which do not agree with the majority of the distance measurements. The straightforward approach to mitigate the impact of the NLOS propagation paths would be to identify and eliminate the distance measurements from the NLOS propagations and perform cooperative localization using the measurements from the LOS propagations. However, this approach has the downside that it results in a loss of information which otherwise

could be exploited for localization. Furthermore, there is always the possibility of misidentification of distance measurements from LOS propagations as being from NLOS propagations and vice versa. Thus more subtle approaches are required for mitigating the NLOS propagations.

The maximum likelihood estimator which estimates the positions of the MSs $\mathbf{p}_{\mathrm{mss}} = \left(\mathbf{p}_{\mathrm{ms}}^{\mathrm{T}}(1), \ldots, \mathbf{p}_{\mathrm{ms}}^{\mathrm{T}}(N_{\mathrm{ms}})\right)^{\mathrm{T}}$ based on the distance measurements $\tilde{\mathbf{d}} = (\tilde{d}(1), \ldots, \tilde{d}(N_{\mathrm{d}}))$ is given as

$$\hat{\mathbf{p}}_{\mathrm{mss}} = \underset{\mathbf{p}_{\mathrm{mss}}}{\arg\max} \;\; \mathrm{p}\left(\tilde{\mathbf{d}}|\mathbf{p}_{\mathrm{mss}}\right), \tag{5.5}$$

where $\mathrm{p}\left(\tilde{\mathbf{d}}|\mathbf{p}_{\mathrm{mss}}\right)$ is the conditional pdf of the measured distances given the positions of the MSs. The maximum likelihood estimator for the cooperative localization problem given in (5.5) is intractable as it is a highly nonlinear and non-convex optimization problem. Consequently, several suboptimal cooperative localization algorithms have been proposed [SG04, LTZN05, AM08, YES09, ER10, VB12]. In the following, a cooperative localization algorithm is proposed which yields nearly optimal estimates of the positions of the MSs. The proposed cooperative localization algorithm uses techniques from robust estimation theory where the measurements are systematically weighted so that the impact of outliers resulting from the NLOS propagation paths is minimized.

## 5.4 Robust cooperative localization

The goal of robust estimators is to yield nearly optimal estimates of an unknown parameter not only when the measurements follow a given distribution exactly, but also when the measurements follow a given distribution only approximately [MMY06]. Consequently, for the network localization problem given in (5.5), nearly optimal estimates of the positions of the MSs could be obtained by using techniques from robust estimation theory without complete information about the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\mathbf{p}_{\mathrm{mss}}\right)$, i.e., $\varepsilon$, $b(n_{\mathrm{d}})$, $\sigma_{\mathrm{LOS}}$ and $\sigma_{\mathrm{NLOS}}$. It is only required that the majority of the BS-MS and the MS-MS links are LOS. The maximum likelihood estimator, on the other hand, requires full information about the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\mathbf{p}_{\mathrm{mss}}\right)$.

For the network localization problem under consideration, the measurement noise is modeled by a Gaussian mixture given in (5.4) as

$$w(n_{\mathrm{d}}) \sim (1-\varepsilon)\mathcal{N}(0, \sigma_{\mathrm{LOS}}^2) + \varepsilon\mathcal{N}(b(n_{\mathrm{d}}), \sigma_{\mathrm{NLOS}}^2),$$

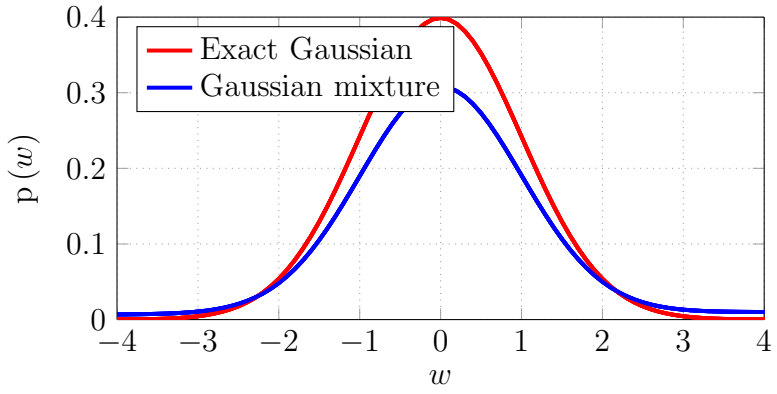



Figure 5.2. Plot of exact Gaussian and Gaussian mixture pdfs

which can be considered to be approximately Gaussian distributed. Fig. 5.2 shows plot of the pdf of an exact Gaussian and a Gaussian mixture for $\varepsilon = 0.25$. The parameters of the Gaussian mixture pdf are $b = 5$, $\sigma_{\text{LOS}} = 1$ and $\sigma_{\text{NLOS}} = 10$, whereas the exact Gaussian pdf has a standard deviation $\sigma = 1$. It can be seen that the Gaussian mixture pdf has heavy tails. A robust estimator can be used to deliver nearly optimal estimates when the measurement noise is not only exactly Gaussian, but also Gaussian mixture. Since any robust estimator is developed based on minimizing the impact of outliers, the distribution of the outliers does not affect the performance of a robust estimator. Thus the distribution of the measurement noises for the NLOS propagations, which have been assumed to be Gaussian for simplicity, need not necessarily be Gaussian. This has great practical importance as the cooperative localization algorithm to be proposed is not constrained by the assumption of the Gaussian measurement noise for the NLOS propagations.

The maximum likelihood estimator of the network localization problem of (5.5) solves a nonlinear and non-convex optimization problem. In the following, an approximate solution to the estimation problem of (5.5) is proposed. Techniques from robust estimation theory are proposed to solve the non-convexity of the optimization problem. The nonlinearity of the estimation problem can be handled by approximating the nonlinear measurement function as a linear function using a first order Taylor series. Towards this end, let $\mathbf{g}(\mathbf{p}_{\text{mss}}) = (d(1), \ldots, d(N_{\text{d}}))^{\text{T}}$ denote the vector-valued function which yields the true distances between the BS-MS and the MS-MS links. Thus the network localization measurement equation is given as

$$\tilde{\mathbf{d}} = \mathbf{g}(\mathbf{p}_{\text{mss}}) + \mathbf{w}, \tag{5.6}$$

where $\mathbf{w} = (w(1), \ldots, w(N_{\text{d}}))^{\text{T}}$ is a multivariate Gaussian mixture measurement noise. The linear approximation of the vector-valued function $\mathbf{g}(\mathbf{p}_{\text{mss}})$ at a reference point

$\mathbf{p}_{\text{mss}_0}$, which shall be close to the true value of the vector $\mathbf{p}_{\text{mss}}$ to yield a reliable linear approximation, is given as

$$\mathbf{g}(\mathbf{p}_{\text{mss}}) \approx \mathbf{g}(\mathbf{p}_{\text{mss}_0}) + \mathbf{G}_0 \cdot (\mathbf{p}_{\text{mss}} - \mathbf{p}_{\text{mss}_0}), \tag{5.7}$$

where the matrix $\mathbf{G}_0$ is defined as

$$\mathbf{G}_0 = \left.\frac{\partial \mathbf{g}(\mathbf{p}_{\text{mss}})}{\partial \mathbf{p}_{\text{mss}}}\right|_{\mathbf{p}_{\text{mss}_0}}. \tag{5.8}$$

The matrix $\mathbf{G}_0$ can be calculated using the expressions given in Appendix A. Substituting (5.7) in (5.6) results in

$$\tilde{\mathbf{d}} = \mathbf{g}(\mathbf{p}_{\text{mss}_0}) + \mathbf{G}_0 \cdot (\mathbf{p}_{\text{mss}} - \mathbf{p}_{\text{mss}_0}) + \mathbf{w}. \tag{5.9}$$

Re-arranging the above equation results in

$$\grave{\mathbf{d}}_0 = \mathbf{G}_0 \cdot \mathbf{p}_{\text{mss}} + \mathbf{w}, \tag{5.10}$$

where

$$\grave{\mathbf{d}}_0 = \tilde{\mathbf{d}} - \mathbf{g}(\mathbf{p}_{\text{mss}_0}) + \mathbf{G}_0 \cdot \mathbf{p}_{\text{mss}_0}. \tag{5.11}$$

The conditional pdf $\mathrm{p}\left(\grave{\mathbf{d}}_0 | \mathbf{p}_{\text{mss}}\right)$ for the linearized network localization scenario in (5.10) is given as

$$\mathrm{p}\left(\grave{\mathbf{d}}_0 | \mathbf{p}_{\text{mss}}\right) = \prod_{n_{\text{d}}=1}^{N_{\text{d}}} \mathrm{p}\left(\grave{d}_0(n_{\text{d}}) - \mathbf{g}_0^{\text{T}}(n_{\text{d}}) \cdot \mathbf{p}_{\text{mss}}\right), \tag{5.12}$$

where $\mathbf{g}_0^{\text{T}}(n_{\text{d}})$ and $\grave{d}_0(n_{\text{d}})$ are the $n_{\text{d}}^{\text{th}}$ row of the matrix $\mathbf{G}_0$ and the $n_{\text{d}}^{\text{th}}$ element of the observation vector $\grave{\mathbf{d}}_0$, respectively. The maximum likelihood estimator for the linearized network localization scenario is given as

$$\hat{\mathbf{p}}_{\text{mss}} = \underset{\mathbf{p}_{\text{mss}}}{\arg\max} \prod_{n_{\text{d}}=1}^{N_{\text{d}}} \mathrm{p}\left(\grave{d}_0(n_{\text{d}}) - \mathbf{g}_0^{\text{T}}(n_{\text{d}}) \cdot \mathbf{p}_{\text{mss}}\right). \tag{5.13}$$

Since the logarithm function is a monotonically increasing function, (5.13) can be rewritten as

$$\hat{\mathbf{p}}_{\text{mss}} = \underset{\mathbf{p}_{\text{mss}}}{\arg\min} \sum_{n_{\text{d}}=1}^{N_{\text{d}}} \rho\left(\grave{d}_0(n_{\text{d}}) - \mathbf{g}_0^{\text{T}}(n_{\text{d}}) \cdot \mathbf{p}_{\text{mss}}\right), \tag{5.14}$$

where $\rho(\theta)$ is the loss function defined as

$$\rho(\theta) = -\ln\left(\mathrm{p}(\theta)\right). \tag{5.15}$$

Differentiating (5.14) with respect to $\mathbf{p}_{\text{mss}}$ and equating the resulting expression to zero yields

$$\sum_{n_{\text{d}}=1}^{N_{\text{d}}} \eta\left(\grave{d}_0(n_{\text{d}}) - \mathbf{g}_0^{\text{T}}(n_{\text{d}}) \cdot \hat{\mathbf{p}}_{\text{mss}}\right) \cdot \mathbf{g}_0(n_{\text{d}}) = \mathbf{0}, \tag{5.16}$$

where

$$\eta(\theta) = \frac{\partial \rho(\theta)}{\partial \theta}. \tag{5.17}$$

In order to solve (5.16), exact knowledge of the conditional pdf $\text{p}\left(\tilde{\mathbf{d}}|\mathbf{p}_{\text{mss}}\right)$, i.e., the values of $\varepsilon$, $b$, $\sigma_{\text{LOS}}$ and $\sigma_{\text{NLOS}}$, is required. As mentioned earlier, the goal of robust estimators is to yield nearly optimal estimates of an unknown parameter when the measurements follow a given distribution not only exactly, but also only approximately [MMY06]. Thus robust estimators consider functions $\rho(\theta)$ which do not necessarily yield the maximum likelihood estimates for any distribution but which ensure that nearly optimal estimates are obtained when the pdf is exactly Gaussian and approximately Gaussian. It has been shown that a continuous and increasing $\eta(\theta)$ yields a unique solution to (5.16) and hence (5.14) [MMY06]. One such family of $\rho(\theta)$ and $\eta(\theta)$ functions is the Huber functions [Hub64]. The Huber function $\rho_{\text{H},\kappa}(\theta)$ is defined as

$$\rho_{\text{H},\kappa}(\theta) = \begin{cases} \frac{1}{2}\theta^2 & \text{if } |\theta| \leq \kappa \\ \kappa|\theta| - \frac{1}{2}\kappa^2 & \text{if } |\theta| > \kappa, \end{cases} \tag{5.18}$$

which has the derivative

$$\eta_{\text{H},\kappa}(\theta) = \begin{cases} \theta & \text{if } |\theta| \leq \kappa \\ \text{sgn}(\theta) \cdot \kappa & \text{if } |\theta| > \kappa. \end{cases} \tag{5.19}$$

It must be noted that the Huber functions $\rho_{\text{H},\kappa}(\theta)$ and $\eta_{\text{H},\kappa}(\theta)$ are defined for a dispersion parameter $\sigma$ of one. The dispersion parameter is a robust alternative to measuring the deviations in a given data which is conventionally measured using the standard deviation. An example of a robust dispersion parameter estimate is the median absolute deviation (MAD) which is defined for a data set $\mathbf{x}$ as

$$\text{MAD}(\mathbf{x}) = \text{median}\left(|\mathbf{x} - \text{median}(\mathbf{x})|\right). \tag{5.20}$$

If the dispersion parameter $\sigma$ is different from one, then the Huber functions $\rho_{\text{H},\kappa}(\theta)$ and $\eta_{\text{H},\kappa}(\theta)$ should be scaled by the dispersion parameter $\sigma$.

Fig. 5.3 shows plot of the Huber functions $\rho_{\text{H},\kappa}(\theta)$ and $\eta_{\text{H},\kappa}(\theta)$ for $\kappa = 1.4$ and $\kappa = \infty$. It must be noted that the Huber loss function $\rho_{\text{H},\kappa}(\theta)$ for $\kappa = \infty$ corresponds to the square loss function which yields optimal estimates if the pdf p is exactly Gaussian. However, the Huber loss function $\rho_{\text{H},\kappa}(\theta)$ for $\kappa = \infty$ does not yield optimal estimates

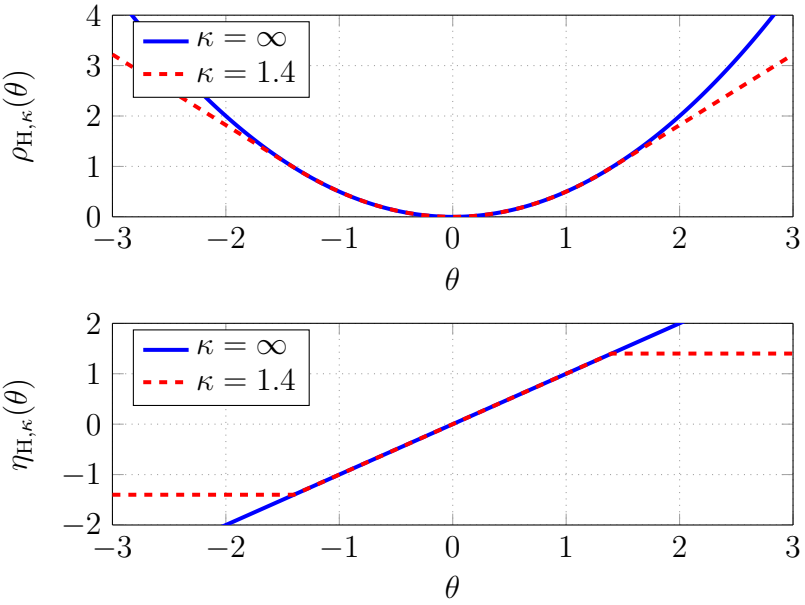



Figure 5.3. The Huber functions $\rho_{\mathrm{H},\kappa}(\theta)$ and $\eta_{\mathrm{H},\kappa}(\theta)$

when the pdf p is approximately Gaussian. On the other hand, the Huber function $\rho_{\mathrm{H},\kappa}(\theta)$ for $\kappa = 1.4$ yields nearly optimal estimates when the pdf p is exactly Gaussian and approximately Gaussian. This can be inferred by considering the equation given in (5.16), i.e.,

$$\sum_{n_{\mathrm{d}}=1}^{N_{\mathrm{d}}} \eta\left(\grave{d}_0(n_{\mathrm{d}}) - \mathbf{g}_0^{\mathrm{T}}(n_{\mathrm{d}}) \cdot \hat{\mathbf{p}}_{\mathrm{mss}}\right) \cdot \mathbf{g}_0(n_{\mathrm{d}}) = \mathbf{0},$$

and the plot of the Huber function $\eta_{\mathrm{H},\kappa}(\theta)$ in Fig. 5.3. For $\kappa = \infty$, the Huber function $\eta_{\mathrm{H},\kappa}(\theta)$ weights all the residuals $\grave{d}_0(n_{\mathrm{d}}) - \mathbf{g}_0^{\mathrm{T}}(n_{\mathrm{d}}) \cdot \hat{\mathbf{p}}_{\mathrm{mss}}$ in (5.16) in proportion to their values, whereas for $\kappa = 1.4$, the Huber function $\eta_{\mathrm{H},\kappa}(\theta)$ weights the residuals for $|\theta| \leq \kappa$ in proportion to their values and the rest of the residuals are weighted equally. Thus, for the case where $\kappa = \infty$, the outliers resulting from the NLOS propagation path measurements are weighted in proportion to their values which results in estimates of the positions of the MSs dominated by outliers. On the other hand, for the case where $\kappa = 1.4$, the outliers resulting from the NLOS propagation path measurements are not weighted in proportion to their residuals and hence their impact on the estimates of the positions of the MSs is rather limited.

Nevertheless, the robustness of the Huber function $\rho_{\mathrm{H},\kappa}(\theta)$ for $\kappa = 1.4$ is accompanied by an increase in the variance of the estimates when the pdf p is exactly Gaussian. In robust estimation theory, the efficiency of a robust estimator is defined as the ratio of the variance of the optimal estimator and the robust estimator when the pdf p is exactly Gaussian. Thus, the Huber function $\rho_{\mathrm{H},\kappa}(\theta)$ for $\kappa = 1.4$ yields an estimator

with reduced efficiency when the pdf p is exactly Gaussian [MMY06]. This happens owing to the saturation of the Huber function $\rho_{\mathrm{H},\kappa}(\theta)$ for $|\theta| > \kappa$ which results in loss of information from some residuals which ultimately leads to a loss in efficiency. Needless to say, the value of the parameter $\kappa$ determines the efficiency and robustness of an estimator. By proper selection of the parameter $\kappa$, it is possible to have an estimator which has a comparable variance with the square loss function when the pdf p is exactly Gaussian and yields robust estimates when the pdf p is approximately Gaussian. For example, for an estimator with $\kappa = 1.4$, the variance is only 4.7 % larger than for $\kappa = \infty$ when the pdf p is exactly Gaussian. However, an estimator with $\kappa = 1.4$ is robust unlike the one of with $\kappa = \infty$. Consequently, there is a trade-off between robustness and efficiency.

In order to achieve more robustness towards outliers, the function $\eta$ should be designed in such a way that it tends to zero at infinity so that the outliers are given small or no weight in the estimation. Consequently, the bisquare family of functions is a popular choice for $\rho$ and $\eta$ as they result in a bounded $\rho$ function and hence an $\eta$ function which is zero at infinity. Thus the bisquare family of functions is more robust towards outliers than the Huber functions. A bisquare $\rho_{\mathrm{bis},\kappa}(\theta)$ function is defined as

$$\rho_{\mathrm{bis},\kappa}(\theta) = \begin{cases} 1 - \left(1 - (\frac{\theta}{\kappa})^2\right)^3 & \text{if } |\theta| \leq \kappa \\ 1 & \text{if } |\theta| > \kappa, \end{cases} \tag{5.21}$$

which has a derivative $\frac{\partial \rho_{\mathrm{bis},\kappa}(\theta)}{\partial \theta} = \frac{6\eta_{\mathrm{bis},\kappa}(\theta)}{\kappa^2}$ where

$$\eta_{\mathrm{bis},\kappa}(\theta) = \begin{cases} \theta \cdot \left(1 - (\frac{\theta}{\kappa})^2\right)^2 & \text{if } |\theta| \leq \kappa \\ 0 & \text{if } |\theta| > \kappa. \end{cases} \tag{5.22}$$

As in the case of the Huber functions $\rho_{\mathrm{H},\kappa}(\theta)$ and $\eta_{\mathrm{H},\kappa}(\theta)$, the bisquare functions $\rho_{\mathrm{bis},\kappa}(\theta)$ and $\eta_{\mathrm{bis},\kappa}(\theta)$ are defined for a dispersion parameter $\sigma$ of one. Fig. 5.4 shows plot of the bisquare functions $\rho_{\mathrm{bis},\kappa}(\theta)$ and $\eta_{\mathrm{bis},\kappa}(\theta)$, respectively. As mentioned earlier, a continuous and increasing $\eta(\theta)$ yields a unique solution to (5.16) and hence (5.14). Since the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ is not an increasing function, it is possible to have multiple solutions of (5.16). Moreover, some solutions may not satisfy the absolute minimum criterion in (5.16). In addition, the bisquare functions do not yield the maximum likelihood estimates for any distribution. Nevertheless, in the following, the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ is considered to solve (5.16) as it is more robust towards outliers than the Huber function $\eta_{\mathrm{H},\kappa}(\theta)$.

It can be inferred from (5.22) that the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ can be written as

$$\eta_{\mathrm{bis},\kappa}(\theta) = c_\kappa(\theta) \cdot \theta, \tag{5.23}$$

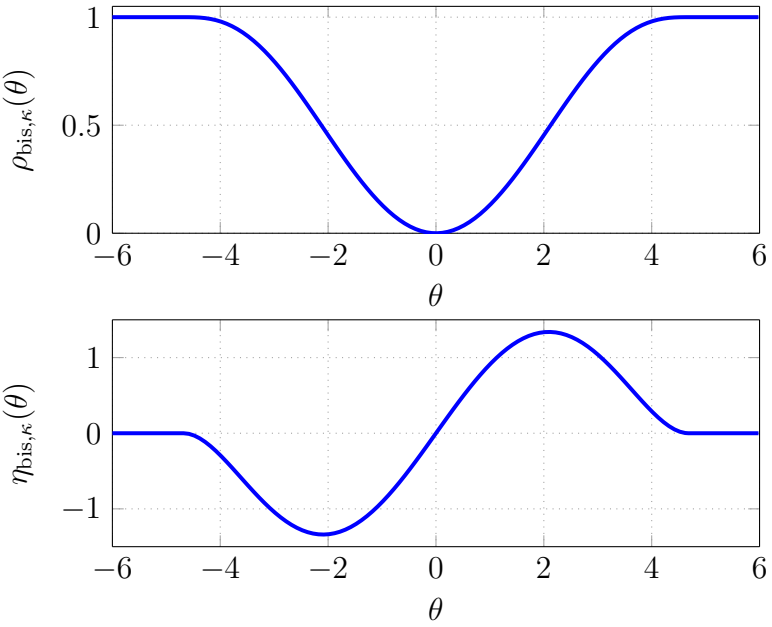



Figure 5.4. The bisquare functions $\rho_{\text{bis},\kappa}(\theta)$ and $\eta_{\text{bis},\kappa}(\theta)$, $\kappa = 4.68$

where $c_\kappa(\theta)$ is a weighting function. For the bisquare function $\eta_{\text{bis},\kappa}(\theta)$, $\eta_{\text{bis},\kappa}(0) = 0$ and $\frac{\partial \eta_{\text{bis},\kappa}(\theta)}{\partial \theta}\Big|_{\theta=0}$ exists, so that $\eta_{\text{bis},\kappa}(\theta)$ can be considered to be linear at the origin. Thus the weight function can be derived from (5.23) as

$$c_\kappa(\theta) = \begin{cases} \frac{\eta_{\text{bis},\kappa}(\theta)}{\theta} & \text{if } \theta \neq 0 \\ \frac{\partial \eta_{\text{bis},\kappa}(\theta)}{\partial \theta}\Big|_{\theta=0} & \text{if } \theta = 0. \end{cases} \tag{5.24}$$

Using the formula for the bisquare function $\eta_{\text{bis},\kappa}(\theta)$ in (5.22), the bisquare weight function is given as

$$c_\kappa(\theta) = \begin{cases} \left(1 - (\frac{\theta}{\kappa})^2\right)^2 & \text{if } |\theta| \leq \kappa \\ 0 & \text{if } |\theta| > \kappa. \end{cases} \tag{5.25}$$

Fig. 5.5 shows plot of the bisquare weight function. It can be seen from Fig. 5.5 that the residuals of the outliers are given small or zero weight.

Using the bisquare weight function, (5.16) can be re-written as

$$\sum_{n_\text{d}=1}^{N_\text{d}} c_\kappa(n_\text{d}) \cdot \left(\grave{d}_0(n_\text{d}) - \mathbf{g}_0^\text{T}(n_\text{d}) \cdot \hat{\mathbf{p}}_\text{mss}\right) \cdot \mathbf{g}_0(n_\text{d}) = \mathbf{0}, \tag{5.26}$$

where

$$c_\kappa(n_\text{d}) = c_\kappa\left(\frac{\grave{d}_0(n_\text{d}) - \mathbf{g}_0^\text{T}(n_\text{d}) \cdot \hat{\mathbf{p}}_\text{mss}}{\sigma}\right). \tag{5.27}$$

The weights $c_\kappa(n_\text{d})$ are scaled by the dispersion parameter $\sigma$ as the bisquare function $\eta_{\text{bis},\kappa}(\theta)$ is defined for a dispersion parameter $\sigma$ of one. The system of equations in (5.26)

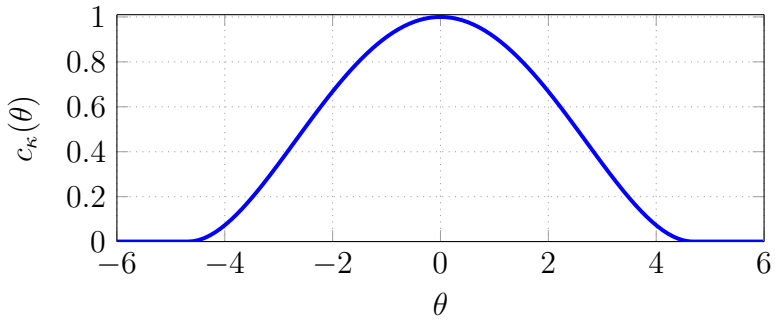



Figure 5.5. The bisquare weight function $c_\kappa(\theta)$, $\kappa = 4.68$

is nonlinear in the estimates of the positions of the MSs $\hat{\mathbf{p}}_{\text{mss}}$ as the weights $c_\kappa(n_\text{d})$ are calculated using estimates of the positions of the MSs $\hat{\mathbf{p}}_{\text{mss}}$. The system of nonlinear equations can be solved using a method called the iteratively re-weighted least squares (IRLS) algorithm [MMY06]. The IRLS algorithm starts with a initial estimate $\hat{\mathbf{p}}_{\text{mss}_0}$ from which the weights $c_\kappa(n_\text{d})$ are calculated. Then (5.26) solved iteratively until a desired threshold is reached. Towards this end, (5.26) can be re-formulated as follows

$$(\mathbf{C}_\kappa \odot \mathbf{G}_0)^\text{T} \cdot \grave{\mathbf{d}}_0 = (\mathbf{C}_\kappa \odot \mathbf{G}_0)^\text{T} \cdot \mathbf{G}_0 \cdot \hat{\mathbf{p}}_{\text{mss}}, \tag{5.28}$$

where $\mathbf{C}_\kappa$ is a $N_\text{d} \times 2N_\text{ms}$ matrix defined as

$$\mathbf{C}_\kappa = \begin{pmatrix} c_\kappa(1) & c_\kappa(1) & \cdots & c_\kappa(1) \\ c_\kappa(2) & c_\kappa(2) & \cdots & c_\kappa(2) \\ \vdots & \vdots & & \vdots \\ c_\kappa(N_\text{d}) & c_\kappa(N_\text{d}) & \cdots & c_\kappa(N_\text{d}) \end{pmatrix}. \tag{5.29}$$

The symbol $\odot$ in (5.28) denotes the Hadamard (element-wise) matrix product. Thus $\hat{\mathbf{p}}_{\text{mss}}$ can be estimated as

$$\hat{\mathbf{p}}_{\text{mss}} = \left((\mathbf{C}_\kappa \odot \mathbf{G}_0)^\text{T} \cdot \mathbf{G}_0\right)^+ \cdot (\mathbf{C}_\kappa \odot \mathbf{G}_0)^\text{T} \cdot \grave{\mathbf{d}}_0, \tag{5.30}$$

where the superscript + stands for the Moore-Penrose pseudoinverse. Thus the IRLS algorithm starts with a initial estimate $\hat{\mathbf{p}}_{\text{mss}_0}$ and solves (5.30) iteratively until a desired threshold is reached.

The nonlinear network localization problem has been handled by approximating it as a linear estimation problem using a first order Taylor series. In order to obtain reliable estimates of the positions of the MSs, the linearization of the nonlinear network localization problem and estimation of the positions of the MSs using the IRLS algorithm shall be performed iteratively. The initial iteration point is crucial to reliable linearization of the nonlinear network localization problem and to convergence of the

IRLS algorithm. In the following, a means of obtaining robust initial estimates $\hat{\mathbf{p}}_{\mathrm{mss}_0}$ of the positions of the MSs is proposed by considering the BS-MS links. Towards this end, the measurement equation for the BS-MS links given in (5.2) can be linearized as follows [CSMC04]

$$\begin{aligned}
\tilde{d}(n_{\mathrm{d}}) &= \|\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}}) - \mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\| + w(n_{\mathrm{d}}) \\
\Big(\tilde{d}(n_{\mathrm{d}}) - w(n_{\mathrm{d}})\Big)^2 &= \|\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}}) - \mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\|^2 \\
\Big(\tilde{d}(n_{\mathrm{d}}) - w(n_{\mathrm{d}})\Big)^2 &= \|\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}})\|^2 - 2\mathbf{p}_{\mathrm{bs}}^{\mathrm{T}}(n_{\mathrm{bs}})\mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}}) + \|\mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\|^2 \\
\check{d}(n_{\mathrm{d}}) &= \begin{pmatrix} -2\mathbf{p}_{\mathrm{bs}}^{\mathrm{T}}(n_{\mathrm{bs}}) & 1 \end{pmatrix} \begin{pmatrix} \mathbf{p}_{\mathrm{mss}}^{\mathrm{T}}(n_{\mathrm{ms}}) & \|\mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\|^2 \end{pmatrix}^{\mathrm{T}} + \check{w}(n_{\mathrm{d}}), \quad (5.31)
\end{aligned}$$

where

$$\check{d}(n_{\mathrm{d}}) = \tilde{d}^2(n_{\mathrm{d}}) - \|\mathbf{p}_{\mathrm{bs}}(n_{\mathrm{bs}})\|^2 \quad (5.32)$$

$$\check{w}(n_{\mathrm{d}}) = 2\tilde{d}(n_{\mathrm{d}})w(n_{\mathrm{d}}) - w^2(n_{\mathrm{d}}). \quad (5.33)$$

The squared norm $\|\mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\|^2$ in (5.31) is a nuisance variable. The least squares estimator can be used to obtain initial estimates $\hat{\mathbf{p}}_{\mathrm{mss}_0}$ of the positions of the MSs from the linearized BS-MS measurement equation in (5.31). However, the square loss function of the least squares estimator is not robust to outliers. Consequently, the absolute loss function is considered to obtain robust initial estimates $\hat{\mathbf{p}}_{\mathrm{mss}_0}$ of the positions of the MSs, i.e.,

$$\hat{\mathbf{p}}_{\mathrm{mss}_0} = \underset{\mathbf{p}_{\mathrm{mss}},\ \|\mathbf{p}_{\mathrm{mss}}\|^2}{\arg\min} \sum_{n_{\mathrm{d}}=1}^{N_{\mathrm{bs}} \cdot N_{\mathrm{ms}}} \left| \check{d}(n_{\mathrm{d}}) - \begin{pmatrix} -2\mathbf{p}_{\mathrm{bs}}^{\mathrm{T}}(n_{\mathrm{bs}}) & 1 \end{pmatrix} \begin{pmatrix} \mathbf{p}_{\mathrm{mss}}^{\mathrm{T}}(n_{\mathrm{ms}}) & \|\mathbf{p}_{\mathrm{mss}}(n_{\mathrm{ms}})\|^2 \end{pmatrix}^{\mathrm{T}} \right|. \quad (5.34)$$

The absolute loss function is robust to outliers as it scales the loss only linearly, instead of quadratically as in the square loss function, by the value of the residuals. Minimization of the absolute loss function in (5.34) can be evaluated using linear programming. The MATLAB function *fminsearch*, which uses the simplex method to iteratively solve linear programming problems, can be used to solve (5.34).

Algorithm 1 shows the pseudocode of the robust cooperative localization algorithm. The robustness of the proposed estimator is due to the initial robust estimates and the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ which systematically weights the distance measurements from the LOS and the NLOS propagations so that the impact of outliers resulting from the NLOS propagation paths is minimized. As it can be seen from algorithm 1, the proposed robust cooperative localization algorithm is semi-parametric in that it only requires estimates of the dispersion parameter $\hat{\sigma}$ to yield robust estimates of the positions of the MSs. On the other hand, the maximum likelihood estimator requires an estimate of the fraction $\varepsilon$ of the propagation paths which are NLOS, the NLOS

**Algorithm 1** Pseudocode for the robust cooperative localization algorithm

use the MATLAB function *fminsearch* to solve (5.34)
$i = 0$
**repeat**
  $\mathbf{G}_i = \frac{\partial \mathbf{g}(\mathbf{p}_{\text{mss}})}{\partial \mathbf{p}_{\text{mss}}}\Big|_{\hat{\mathbf{p}}_{\text{mss}_i}}$
  use the IRWLS algorithm to solve (5.30)
  $i = i+1$
**until** $\frac{1}{\sqrt{N_{\text{ms}}}}\|\hat{\mathbf{p}}_{\text{mss}_i} - \hat{\mathbf{p}}_{\text{mss}_{i-1}}\| < \Delta$ or $N_{\max}$
$\hat{\mathbf{p}}_{\text{mss}} = \hat{\mathbf{p}}_{\text{mss}_i}$

biases $b(n_{\text{d}})$, the standard deviation of the LOS paths $\sigma_{\text{LOS}}$ and the NLOS paths $\sigma_{\text{NLOS}}$ to yield optimal estimates of the positions of the MSs.

It must be noted that the performance of the robust cooperative localization algorithm depends on the accuracy of the initial estimates $\hat{\mathbf{p}}_{\text{mss}_0}$ of the positions of the MSs. Bad initial estimates and unfavorable localization scenarios can result in ill-posed inverse problems in (5.30). Such cases can be alleviated using the Thikonov regularization [TA77], for example. Empirical studies have shown that given a sufficiently large network localization scenario, (5.30) is well-posed.

## 5.5 Performance analysis

### 5.5.1 Cramér-Rao lower bound

The Cramér-Rao lower bound (CRLB) is considered here to determine the lower bound of the covariance matrix of an unbiased estimate of the positions of the MSs. The CRLB of the geometric cooperative localization algorithm can be easily obtained from a straightforward extension of the CRLB of the MS localization algorithm for NLOS multipath environments under the explicit consideration of scatterers presented in Section 3.4.1. On the contrary, the CRLB of the robust cooperative localization algorithm is rather complicated and needs special treatment. In the following, the CRLB of the robust cooperative localization algorithm is presented.

For network localization in multipath environments with the effect of the NLOS propagations on the measured NLOS propagation path parameters modeled by a positive distance bias, the unknown parameters are the positions of the MSs and the NLOS biases. It is assumed that the parameter $\varepsilon$ which determines the proportion of the

LOS and the NLOS propagation paths is known. Let the $2N_{\mathrm{ms}} + N_{\mathrm{d}} \times 1$ vector $\boldsymbol{\rho} = \left(\mathbf{p}_{\mathrm{mss}}^{\mathrm{T}}, b(1), \ldots, b(N_{\mathrm{d}})\right)^{\mathrm{T}}$ denote the unknown parameters. Furthermore, let the pdf $\mathrm{p}\left(\hat{\mathbf{d}}|\boldsymbol{\rho}\right)$ denote the conditional pdf of the measurements $\tilde{\mathbf{d}} = \left(\tilde{d}(1), \ldots, \tilde{d}(N_{\mathrm{d}})\right)^{\mathrm{T}}$ given the unknown parameters $\boldsymbol{\rho}$. The parameter $\varepsilon$ determines the proportion of the LOS and the NLOS distance measurements of the measurement vector $\tilde{\mathbf{d}}$. Consequently, the parameter $\varepsilon$ determines the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\boldsymbol{\rho}\right)$.

Given the measurement vector $\tilde{\mathbf{d}}$, the composition of the LOS and the NLOS distance measurements of the measurement vector $\tilde{\mathbf{d}}$ is unknown, i.e., it is not known which distance measurements are from LOS and NLOS propagation paths. For a measurement vector $\tilde{\mathbf{d}}$ of length $N_{\mathrm{d}}$, there are $2^{N_{\mathrm{d}}}$ possible combinations of LOS and NLOS distance measurements. Let $\mathrm{p}(\gamma)$, $\gamma = 1, \ldots, 2^{N_{\mathrm{d}}}$ denote the probability mass function (pmf) of the possible combinations of LOS and NLOS distance measurements. Thus the pmf $\mathrm{p}(\gamma)$ can be written as

$$\mathrm{p}(\gamma) = (1-\varepsilon)^{n_{\mathrm{LOS}}} \varepsilon^{N_{\mathrm{d}} - n_{\mathrm{LOS}}}, \tag{5.35}$$

where $n_{\mathrm{LOS}} = 0, \ldots, N_{\mathrm{d}}$ is the number of LOS propagation path measurements. Thus considering the different possible distance measurements, the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\boldsymbol{\rho}\right)$ can be written using the total probability theorem as

$$\mathrm{p}\left(\tilde{\mathbf{d}}|\boldsymbol{\rho}\right) = \sum_{\gamma=1}^{2^{N_{\mathrm{d}}}} \mathrm{p}(\gamma)\, \mathrm{p}\left(\tilde{\mathbf{d}}|\gamma, \boldsymbol{\rho}\right), \tag{5.36}$$

where the pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\gamma, \boldsymbol{\rho}\right)$ is a product of Gaussian pdfs of the LOS and the NLOS distance measurements whose composition is determined by $\gamma$.

The CRLB of the robust cooperative localization algorithm can be computed using the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\boldsymbol{\rho}\right)$ defined in (5.36). However, the computation involves cumbersome expressions whose computations prove to be rather an arduous task. Consequently, in the following, the CRLB shall be investigated under the assumption that the decomposition of the distance measurements are known a priori, i.e., it is known which of the distance measurements resulted from a LOS or a NLOS propagation path. For such cases, the CRLB can be computed in a similar manner as the derivation of the CRLB discussed in Section 3.4.1. It must be noted that this CRLB is not tight as the robust cooperative localization algorithm discussed in the previous section does not have a priori information about the composition of the distance measurements.

It has been shown in [QKS06] that the CRLB for a MS localization in multipath environments, when the BS-MS links with NLOS propagation paths are known a priori, depends only on the propagation paths which are LOS. Thus the distance measurements

from the NLOS propagation paths are not to be exploited for MS localization. This result, somewhat surprising, is justified in that no prior information is assumed about the NLOS biases which could be exploited for localization of the MS. The fact that the NLOS biases are positive does not give any information about the locations of the MSs. Thus an estimator can only exploit the distance measurements from the LOS propagation paths, by identifying the LOS distance measurements, for localizing the MS. This result also holds for the network localization case. Thus the CRLB depends on the BS-MS and the MS-MS links which are LOS which in turn is determined by the parameter $\varepsilon$. Thus the CRLB depends implicitly on the parameter $\varepsilon$. It must be noted that if the CRLB were derived from the conditional pdf $\mathrm{p}\left(\tilde{\mathbf{d}}|\boldsymbol{\rho}\right)$ defined in (5.36), then the CRLB would depend explicitly on the parameter $\varepsilon$.

### 5.5.2 Simulation results

In this section, the performances of the geometric and the robust cooperative localization algorithms are analyzed considering a picocell network localization scenario using Monte Carlo simulations with $10^4$ independent trials. The picocell network localization scenario consists of 14 cells. Each cell has a radius of 100 m and is served by a single BS. Furthermore, in each cell there are 4 MSs. Thus there are 14 BSs and 56 MSs. The BSs are located at the center of the cell. For each run of the Monte Carlo simulations, the positions of the MSs are generated randomly from a distribution in which each MS is assumed to be uniformly distributed within the cell. Fig. 5.6 shows an exemplary realization of the randomly generated picocell network localization scenario. The BSs are marked with boxes, whereas the MSs are marked with circles. It is assumed that full connectivity exists between the BSs and MSs. It has been discussed that dramatic improvement in performance could be obtained from cooperation between all the MSs, i.e., full connectivity is assumed between the MSs. In practice, the MSs have limited power and it may not be possible to have full connectivity between the MSs. Consequently, it is assumed that two MSs cooperate only if the distance between the MSs is less than the radius of the cell.

The performance metric is the root mean square error (RMSE) $\mathrm{rmse}(\hat{\mathbf{p}}_{\mathrm{mss}})$ of the estimated positions $\hat{\mathbf{p}}_{\mathrm{mss}}$ of the MSs which is calculated as

$$\mathrm{rmse}(\hat{\mathbf{p}}_{\mathrm{mss}}) = \sqrt{\mathbb{E}_{\tilde{\mathbf{d}},\mathbf{p}_{\mathrm{mss}}}\{(\hat{\mathbf{p}}_{\mathrm{mss}} - \mathbf{p}_{\mathrm{mss}})^{\mathrm{T}}(\hat{\mathbf{p}}_{\mathrm{mss}} - \mathbf{p}_{\mathrm{mss}})\}}, \tag{5.37}$$

where $\mathbf{p}_{\mathrm{mss}}$ denotes the true positions of the MSs and $\mathbb{E}_{\tilde{\mathbf{d}},\mathbf{p}_{\mathrm{mss}}}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\tilde{\mathbf{d}}, \mathbf{p}_{\mathrm{mss}}\right)$. In the simulations, the expectations in (5.37) are computed using averages.

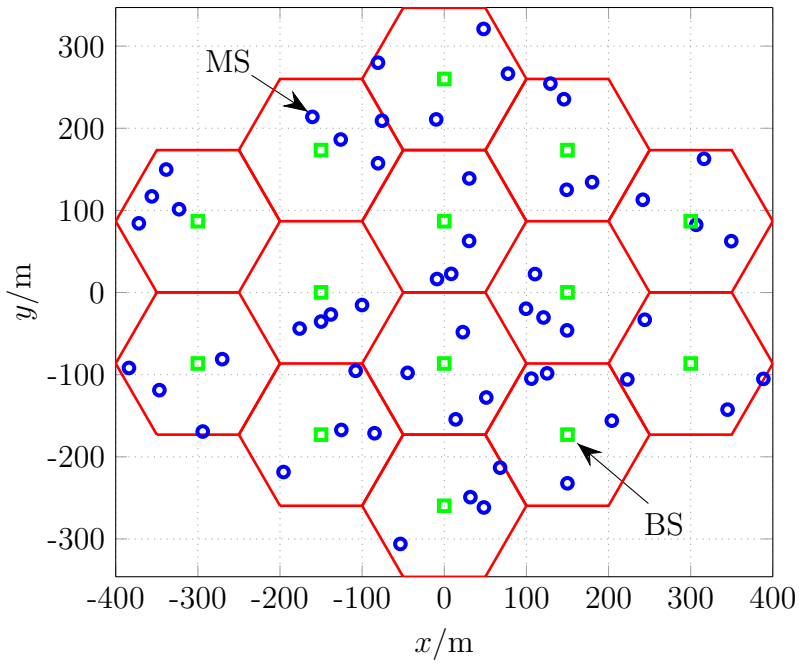



Figure 5.6. A sample realization of the random generation of the network localization scenario

The geometric cooperative localization algorithm has been proposed for network localization in NLOS multipath environments under the explicit consideration of scatterers. The explicit consideration of scatterers is effected via the geometric description of the NLOS propagations. On the other hand, the robust cooperative localization algorithm has been proposed without the explicit consideration of scatterers. Since the scatterers are not explicitly considered, the effect of the NLOS propagations on the measured NLOS propagation path parameters has been described using a probabilistic model. In order to obtain significant insights about the performance of the robust cooperative localization algorithm, its performance is analysed first using the probabilistic description of the effect of the NLOS propagations on the measured NLOS propagation path parameters. Later, the performances of the geometric and the robust cooperative localization algorithms are analysed using the geometric description of the NLOS propagations.

As discussed in Section 5.3, the probabilistic description of the effect of the NLOS propagations on the measured NLOS propagation path parameters is modeled using a Gaussian mixture measurement noise given in (5.4). In the simulations, the LOS path noise standard deviation $\sigma_{\mathrm{LOS}}$ is assumed to be 20 m. Furthermore, it is assumed that the NLOS propagation path bias $b(n_{\mathrm{d}})$ is 40 m. This assumption is made only for the sake of simplicity. In reality, the NLOS path biases depend on the BS-MS and the MS-MS link distances, among other factors. The performance of the robust cooperative

localization algorithm is compared with the one of a non-robust cooperative localization algorithm proposed in [MBW+04]. The non-robust cooperative localization algorithm is based on linearization of the measurement equation in (5.6) and iteratively solving the resulting linearized least squares problem. To assess the performance gain obtained from cooperation between the MSs, the performance of a robust non-cooperative localization algorithm given in (5.34) is also considered. The performance of a non-robust non-cooperative localization algorithm obtained from the least squares estimates of the linearized BS-MS links given in (5.31) is also considered. The non-robust cooperative localization algorithm from [MBW+04] is initialized using estimates of the MSs from the non-robust non-cooperative localization algorithm. Since the network localization scenarios are randomly generated, the averaged CRLB of the randomly generated network localization scenarios is considered as a performance benchmark of the considered cooperative localization algorithms.

In all the simulations, the robust cooperative localization algorithm is tuned for an efficiency of 95 %, thus $\kappa = 4.68$ in (5.22). For a fair comparison between the robust and the non-robust cooperative localization algorithms, the LOS path noise standard deviation $\sigma_{\mathrm{LOS}}$ is used as the dispersion parameter for both cooperative localization algorithms. Furthermore, the maximum number of iterations $N_{\max}$ in both the robust and non-robust cooperative localization algorithms is set to 10. For cases where the robust and the non-robust cooperative localization algorithms diverge, the estimates of the positions of the MSs are obtained from the robust and the non-robust non-cooperative localization algorithms, respectively. The robust and the non-robust cooperative localization algorithms are said to have diverged if $\frac{1}{\sqrt{N_{\mathrm{ms}}}}\|\hat{\mathbf{p}}_{\mathrm{mss}_i} - \hat{\mathbf{p}}_{\mathrm{mss}_{i-1}}\| > 20$ m.

Fig. 5.7 shows the RMSE performances of the aforementioned localization algorithms for different fractions $\varepsilon$ of the propagation paths which are NLOS. The NLOS path noise standard deviation $\sigma_{\mathrm{NLOS}}$ is assumed to be 80 m. It can be seen that the cooperation between the MSs results in a significant improvement in performance for both the robust and the non-robust localization algorithms. Furthermore, the robust cooperative and the robust non-cooperative localization algorithms result in a significant performance improvement over the non-robust cooperative and the non-robust non-cooperative localization algorithms, respectively. Moreover, the performances of the localization algorithms are bounded by the CRLB. In general, the robust cooperative localization algorithm results in a significant performance improvement over the non-robust cooperative localization algorithm. It can be seen that the improvement in performance begins to decrease as the fraction $\varepsilon$ of the propagation paths which are NLOS approaches 0.5. This is to be expected as the number of the LOS propagations decreases with increasing $\varepsilon$. A closer look at Fig. 5.7 reveals that when the fraction $\varepsilon$ of the propagation paths which are NLOS is close to zero, the non-robust cooperative

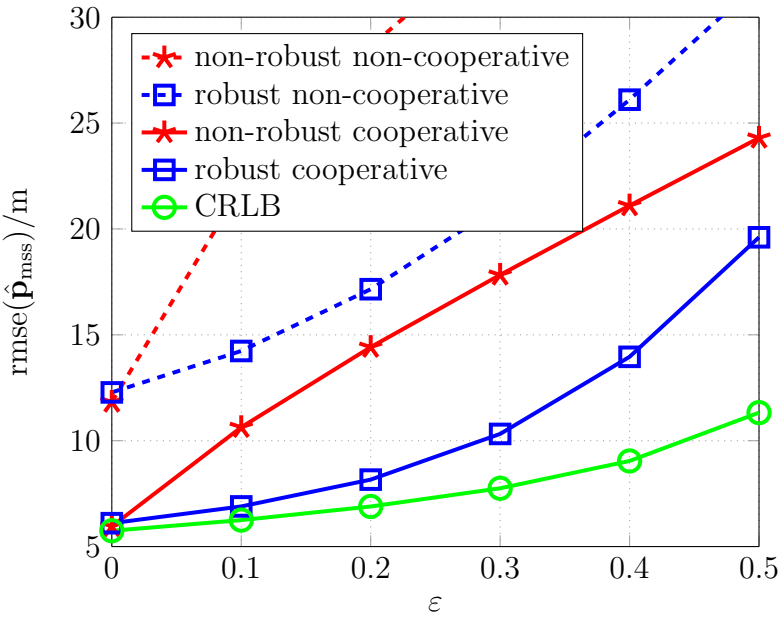



Figure 5.7. RMSE performances of different localization algorithms for different fractions $\varepsilon$ of the propagation paths which are NLOS, $\sigma_{\text{NLOS}} = 80$ m

localization algorithm results in a slight improvement in performance over the robust cooperative localization algorithm. This happens owing to the fact that the robust cooperative localization algorithm is derived to be only nearly optimal in the presence or absence of the NLOS propagations. Thus in the absence of the NLOS propagations, the non-robust cooperative localization algorithm, which assumes only LOS propagations, performs better than the robust cooperative localization algorithm. The robustness of the robust cooperative localization algorithm is paid for by the loss of efficiency when there are only LOS propagation paths.

It must be noted that the robustness of the robust cooperative localization algorithm is guaranteed only for cases where the fraction $\varepsilon$ of the propagation paths which are NLOS is less than 0.5. For $\varepsilon \geq 0.5$, the robust estimator cannot distinguish the noise distribution for the measured LOS propagation path parameters from the measured propagation path parameters. Thus estimates of the dispersion parameter $\hat{\sigma}$ will be biased towards the standard deviation $\sigma_{\text{NLOS}}$ of the NLOS propagation path parameter measurement noise. Consequently, estimates of the positions of the MSs will be severely influenced by the measured NLOS propagation path parameters. However, this is not reflected in the simulation results shown in Fig. 5.7 as the LOS path noise standard deviation $\sigma_{\text{LOS}}$ is used as an estimate of the dispersion parameter $\hat{\sigma}$. Thus the robust cooperative localization algorithm is still able to mitigate the measured path parameters from the NLOS propagation paths. Furthermore, the NLOS propagation path bias

$b(n_\mathrm{d})$ is assumed to be 40 m regardless of the BS-MS and the MS-MS link distances. Thus, for cases where the BS-MS and the MS-MS link distances are large as compared to the NLOS propagation path bias $b(n_\mathrm{d})$, the impact of the NLOS propagation paths on the estimated positions of the MSs is rather small.

Fig. 5.8 shows the RMSE performances of the aforementioned localization algorithms for different NLOS propagation path noise standard deviations $\sigma_\mathrm{NLOS}$ when the fraction $\varepsilon$ of the propagation paths which are NLOS is 0.1. As in the simulation results shown in Fig. 5.7, cooperation between the MSs results in a significant improvement in performance for both the robust and the non-robust localization algorithms. Furthermore, the robust localization algorithms result in a significant performance improvement over their non-robust counterparts. In addition, the performances of the localization algorithms are bounded by the CRLB. The CRLB curve is flat as it is independent of $\sigma_\mathrm{NLOS}$. It can be seen that as the NLOS propagation path noise standard deviation $\sigma_\mathrm{NLOS}$ increases, there is a significant decline in performance by the non-robust localization algorithms, whereas the robust localization algorithms show little or no change in performance. This shows that the robust localization algorithms mitigate the outliers resulting from the NLOS propagation paths regardless of their strength. A close examination of Fig. 5.8 reveals that up to $\sigma_\mathrm{NLOS} = 60$ m the robust cooperative localization algorithm shows improvement in performance with increasing $\sigma_\mathrm{NLOS}$. This result which may appear rather strange occurs owing to the fact that the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ is tuned to have an efficiency of 95 %, i.e., $\kappa = 4.68$. Owing to the trade-off between robustness and efficiency, the bisquare function $\eta_{\mathrm{bis},\kappa}(\theta)$ with $\kappa = 4.68$ is not very sensitive to outliers. Thus a considerable number of distance measurements from the NLOS propagation paths are not zero-weighted when $\sigma_\mathrm{NLOS}$ is comparable to $\sigma_\mathrm{LOS}$. Such distance measurements negatively impact the performance of the robust cooperative localization algorithm. However, as $\sigma_\mathrm{NLOS}$ increases, most of the outliers from the NLOS propagation paths are zero-weighted and eliminated completely. Consequently, the performance of the robust cooperative localization algorithm stabilizes.

In the previous simulations, the probabilistic description of the effect of the NLOS propagations on the measured NLOS propagation path parameters has been considered. As the robust cooperative localization algorithm has been derived based on this probabilistic description, it was seen that it results in a satisfactory localization performance. Nevertheless, a more accurate description of the NLOS propagations is obtained from the geometric description of the NLOS propagations. Thus in the following, the performance of the robust cooperative localization algorithm is analysed using the geometric description of the NLOS propagations. In doing so, the versatility of the robust cooperative localization algorithm can be assessed. As the geometric

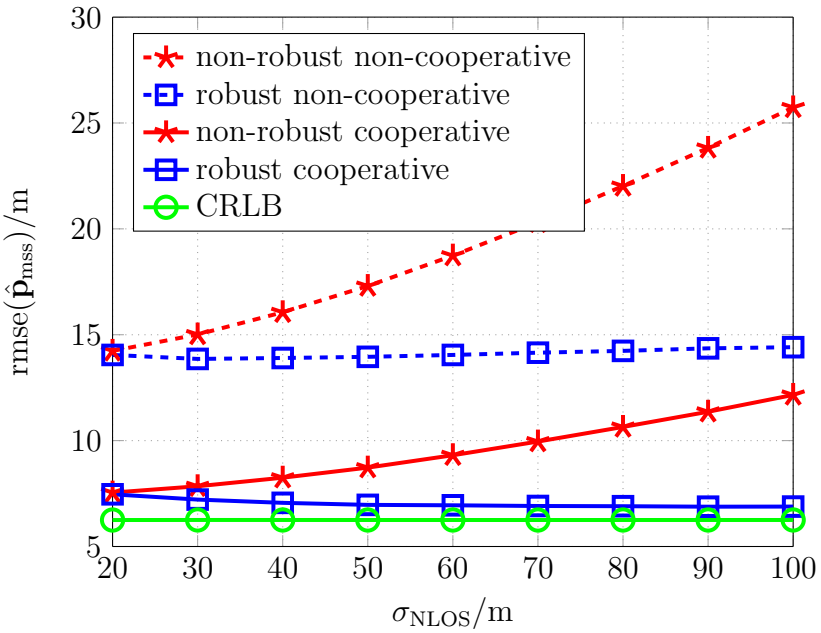



Figure 5.8. RMSE performances of different localization algorithms for different standard deviations of the NLOS propagation noise $\sigma_{\text{NLOS}}$, $\varepsilon = 0.1$ (figure drawn not to scale)

cooperative localization algorithm is based on the geometric description of the NLOS propagations, its performance is also analysed.

The geometric description of the NLOS propagations is obtained using the single bounce elliptical model discussed in Section 3.4.2. Thus, in the simulations, the positions of the scatterers are generated randomly using the single bounce elliptical model where the scatterers are assumed to be uniformly distributed within an ellipse defined by the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ of the NLOS propagation paths. For the sake of simplicity, it is assumed that each of the BS-MS and the MS-MS links has only one propagation path. As the robust cooperative localization algorithm is derived under the assumption that the majority of the propagation paths are LOS, the parameter $\varepsilon$ is used to determine the fraction of the propagation paths which are NLOS.

It must be noted that it is difficult to fairly compare the performances of the robust and the geometric cooperative localization algorithms. The geometric cooperative localization algorithm uses estimates of the times-of-arrival, angles-of-departure and angles-of-arrival of the propagation paths, whereas the robust cooperative localization algorithm uses only the times-of-arrival of the propagation paths. Thus, in the following, the relative behaviour of the robust and the geometric cooperative localization algorithms is studied. Except for the generation of the NLOS propagations using the single bounce elliptical model, the simulation setup is similar to the one considered earlier.

Fig. 5.9 shows the RMSE performances of the robust and the geometric cooperative localization algorithms for different normalized maximum path delays $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ of the NLOS propagation paths and fractions $\varepsilon$ of the propagation paths which are NLOS. The standard deviation $\sigma_d$ of the measured times-of-arrival is 20 m, whereas the standard deviation $\sigma_\phi$ of the measured angles-of-departure and angles-of-arrival is 5°. It can be seen that the geometric cooperative localization algorithm yields a satisfactory performance regardless of the fractions $\varepsilon$ of the propagation paths which are NLOS and the value of the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ of the NLOS propagation paths. Even though the geometric cooperative localization algorithm benefits from the extra information from the spatial dimension, i.e., the angles-of-departure and angles-of-arrivals of the propagation paths, its satisfactory performance stems from the explicit consideration of scatterers rather than a mere direct usage of the extra information obtained from the spatial dimension.

It can be seen from Fig. 5.9 that the performance of the robust cooperative localization algorithm strongly depends on the fraction $\varepsilon$ of the propagation paths which are NLOS. In general, for $\varepsilon = 0.1$ and $\varepsilon = 0.3$, the robust cooperative localization algorithm yields a satisfactory performance. However, for $\varepsilon = 0.5$, the robust cooperative localization algorithm deteriorates severely at higher values of $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$. This occurs owing to the poor initial estimates obtained from the BS-MS links using the absolute loss function given in (5.34). For $\varepsilon = 0.5$, the measured propagation path parameters from the LOS and the NLOS propagation paths are equal in proportion. Thus the absolute loss function, which weights the residuals linearly, is severely influenced by the measured NLOS propagation path parameters. Consequently, for $\varepsilon = 0.5$, the poor initial estimates obtained from the absolute loss function hinder the robust cooperative localization algorithm from converging.

As discussed previously, the robustness of the robust cooperative localization algorithm is guaranteed only for cases where the fraction $\varepsilon$ of the propagation paths which are NLOS is less than 0.5. Nevertheless, it can be seen from Fig. 5.7 and Fig. 5.9 that the robust cooperative localization algorithm is relatively robust when $\varepsilon = 0.5$ as the LOS path noise standard deviation $\sigma_{\text{LOS}}$ is used as an estimate of the dispersion parameter $\hat{\sigma}$. Thus the robust cooperative localization algorithm is still able to mitigate the measured NLOS propagation path parameters. However, the performance of the robust cooperative localization algorithm suffers relatively more for the case where the geometric description of the NLOS propagations is considered. This occurs owing to the fact that in the geometric description of the NLOS propagations, the extra propagation path lengths due to the NLOS propagations correspond with the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ of the NLOS propagation paths and the distances between the BS-MS and the MS-MS links. Hence the impact of the measured NLOS propagation path

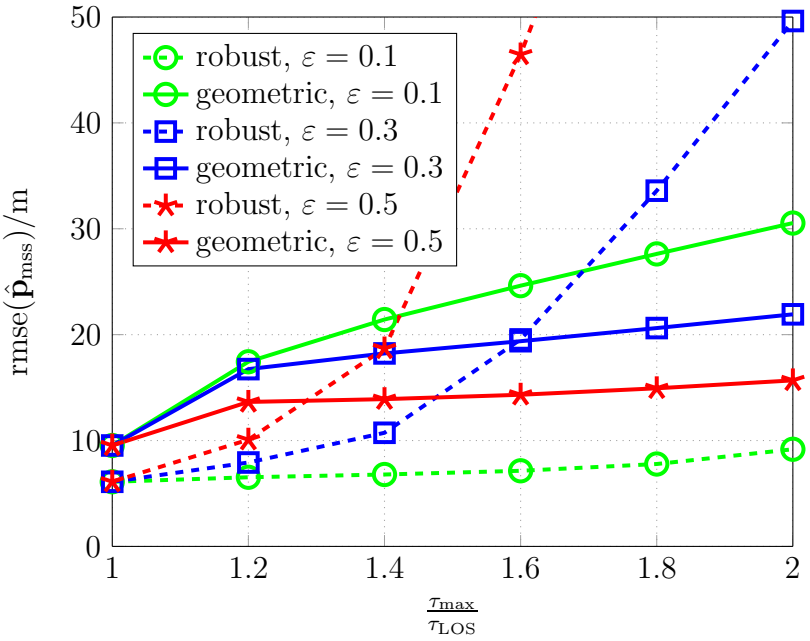



Figure 5.9. RMSE performances of the robust and the geometric cooperative localization algorithms versus the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ for different fractions $\varepsilon$ of the propagation paths which are NLOS, $\sigma_d = 20$ m and $\sigma_\phi = 5°$

parameters on the estimated positions of the MSs is severe. On the other hand, in the probabilistic description of the effect of the NLOS propagations on the measured NLOS propagation path parameters, a NLOS propagation path bias $b(n_{\mathrm{d}})$ of 40 m was assumed regardless of the BS-MS and the MS-MS link distances. This limits the impact of the NLOS propagations on the estimated positions of the MSs.

It can also be seen from Fig. 5.9 that the performance of the geometric cooperative localization algorithm somewhat deteriorates as the value of the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ of the NLOS propagation paths is increased. As a result of the explicit consideration of scatterers, the geometric cooperative localization algorithm is not influenced by the value of $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ except for the impact of the change in the geometry of the localization scenario with a change in $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$. However, the impact of the measurement noises on the geometric cooperative localization algorithm varies with the change in $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$. If the standard deviation $\sigma_d$ of the times-of-arrival measurement noises is fixed, then an increase in $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ results in NLOS propagation paths with increased path lengths while the value of $\sigma_d$ is unchanged. Hence, for a given $\sigma_d$, the geometric cooperative localization algorithm would show a performance gain with an increase in $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ owing to the decrease in the relative value of the NLOS propagation path lengths and $\sigma_d$. On the other hand, if the standard deviation $\sigma_\phi$ of the angles-of-departure and angles-of-arrival measurement noises is fixed, then an increase in $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}}$ results in a localization

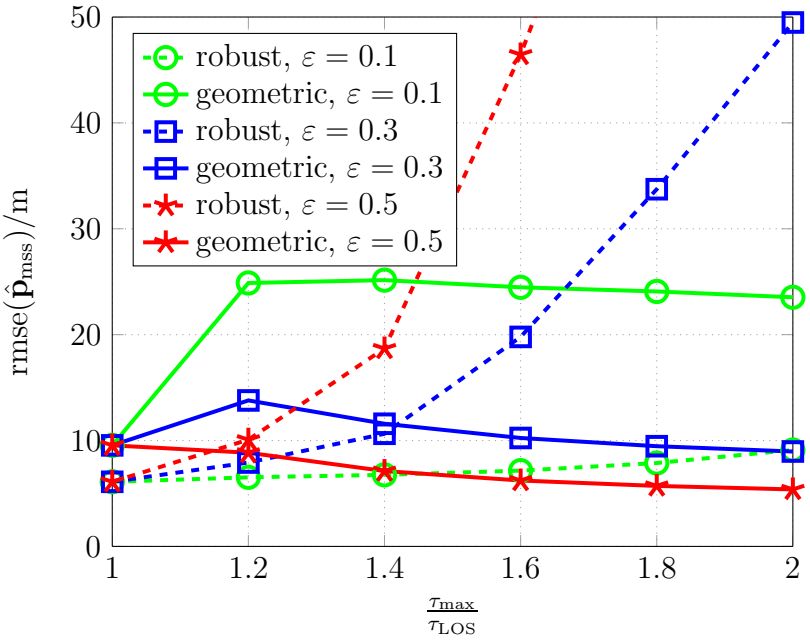



Figure 5.10. RMSE performances of the robust and the geometric cooperative localization algorithms versus the normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ for different fractions $\varepsilon$ of the propagation paths which are NLOS, $\sigma_d = 20$ m and $\sigma_\phi = 1°$

scenario with relatively large distances between a BS, a scatterer and a MS which is more sensitive to the errors in the angles-of-departure and angles-of-arrival estimates. Hence, for a given $\sigma_\phi$, the geometric cooperative localization algorithm would show a performance loss with an increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ owing to the increase of the impact of $\sigma_\phi$. Setting aside the impact of the change in the geometry of the localization scenario with a change in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$, the overall impact of the increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ is determined by the relative values of $\sigma_d$ and $\sigma_\phi$. Thus, if the times-of-arrival measurement noise is dominant, then the increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ results in a performance gain, whereas if the angles-of-departure and angles-of-arrival measurement noise is dominant, then the increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ results in a performance loss. This effect can be seen in Fig. 5.9 where $\sigma_d = 20$ m and $\sigma_\phi = 5°$ and Fig. 5.10 where $\sigma_d = 20$ m and $\sigma_\phi = 1°$. It can be considered that the angles-of-departure and angles-of-arrival measurement noise is dominant in Fig. 5.9, whereas the times-of-arrival measurement noise is dominant in Fig. 5.10. Consequently, in Fig. 5.9, the increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ results in a performance loss for all considered fractions $\varepsilon$ of the propagation paths which are NLOS. On the other hand, in Fig. 5.10, the increase in $\frac{\tau_{\max}}{\tau_{\text{LOS}}}$ results in performance gain for all considered fractions $\varepsilon$ of the propagation paths which are NLOS.

In both Fig. 5.9 and Fig. 5.10, it can be seen that the increase in the fractions $\varepsilon$ of the propagation paths which are NLOS results in a performance gain for the geometric

cooperative localization algorithm. However, this does not hold true for all cases. The result occurs owing to the interplay of the impacts of $\sigma_d$ and $\sigma_\phi$ on the LOS and the NLOS propagation paths and the changing localization scenario with the fraction $\varepsilon$ of the propagation paths which are NLOS.

It can be concluded that, if the majority of the propagation paths are LOS, then the proposed robust cooperative localization algorithm can be used to get a satisfactory localization performance. However, a satisfactory network localization performance can be guaranteed for all NLOS multipath network localization scenarios only under the explicit consideration of scatterers.

In the previous chapters, it has been assumed that the measurement noises are exactly Gaussian distributed. This assumption has been made for simplicity only and in reality the measurement noises are not exactly Gaussian distributed. Measurement campaigns in indoor [BKR97] and outdoor [Mid99] mobile radio communication channels have shown the presence of outliers which result in heavy-tailed measurement noises [ZKCM12]. Hence the robust localization technique discussed in this chapter can be easily adapted to the localization techniques discussed in the previous chapters to guarantee a satisfactory performance in practice.

# Chapter 6

# MIMO channel prediction

## 6.1 Channel prediction using the double directional channel model

In the previous chapters, the spatio-temporal channel properties of MIMO systems have been exploited for localizing and tracking a MS in NLOS multipath environments under the explicit consideration of scatterers. It has been shown that exploiting the spatio-temporal channel properties of MIMO systems offers new insights which result in significant performance gains in localizing and tracking a MS in NLOS multipath environments. In this chapter, the spatio-temporal channel properties of MIMO systems are investigated to obtain new insights in MIMO channel prediction. In the following chapters, the new insights will be exploited for developing MIMO channel prediction algorithms.

In this section, MIMO channel prediction based on the double directional channel model presented in Section 2.1 is considered. For the double directional channel model, the channel transfer function of the channel between the $n^{\text{th}}$ transmitter antenna and the $m^{\text{th}}$ receiver antenna is given in (2.8) as

$$H_{m,n}(f,t) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p t} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}.$$

In this thesis, multicarrier transmission scheme is considered. The time index is denoted by $\ell = 1, \ldots, L$, whereas the subcarrier index is denoted by $k = 1, \ldots, K$. Moreover, the transmit symbol has a duration $T$ and the subcarrier spacing is $F = \frac{1}{T}$. Thus the channel coefficient of the channel between the $n^{\text{th}}$ transmitter antenna and the $m^{\text{th}}$ receiver antenna for the $k^{\text{th}}$ subcarrier at the $\ell^{\text{th}}$ time instant is given as

$$H_{m,n}(k,\ell) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}. \tag{6.1}$$

Consequently, the channel coefficient of the channel between the $n^{\text{th}}$ transmitter antenna and the $m^{\text{th}}$ receiver antenna for the $(k+\Delta k)^{\text{th}}$ subcarrier at the $(\ell+\Delta\ell)^{\text{th}}$ time instant, i.e., for a prediction depth of $\Delta\ell$ in time and $\Delta k$ in frequency, is given as

$$H_{m,n}(k+\Delta k, \ell+\Delta\ell) =$$

$$\sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell}. \quad (6.2)$$

Thus given estimates of the propagation path parameters, the channel prediction task is a mere computation of the value of the channel coefficient at the prediction depths $\Delta\ell$ in time and $\Delta k$ in frequency.

It must be noted that in the derivation of the channel transfer function the narrowband assumption for array signal processing has been considered, i.e., the bandwidth $B$ of the signal is small as compared to the center frequency $f_0$ and the antenna array dimensions are small enough so that $\left|F \cdot (k+\Delta k) \cdot \tau_{\mathrm{Tx}}^{(n,p)}\right|$ and $\left|F \cdot (k+\Delta k) \cdot \tau_{\mathrm{Rx}}^{(m,p)}\right|$ are close to zero. Consequently, under the narrowband assumption for array signal processing the relation

$$\mathrm{e}^{-\mathrm{j}2\pi(F(k+\Delta k)+f_0)\tau_{\mathrm{Tx}}^{(n,p)}} \approx \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \quad (6.3a)$$

$$\mathrm{e}^{-\mathrm{j}2\pi(F(k+\Delta k)+f_0)\tau_{\mathrm{Rx}}^{(m,p)}} \approx \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \quad (6.3b)$$

holds. Thus it was not necessary to modify the steering factors $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}}$ and $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}$ in (6.2) in order to account for the prediction in frequency direction.

## 6.2 Channel dynamics of the double directional channel model

Given reliable estimates of the propagation path parameters, (6.2) can be used to predict the channel coefficients at a prediction depth of $\Delta\ell$ in time and $\Delta k$ in frequency. However, obtaining reliable estimates of the propagation path parameters requires usage of array signal processing techniques which tend to be computationally intensive. Consequently, far less computationally intensive channel prediction approaches have been considered which model the mobile radio channel as a stochastic process and exploit the statistical dependency between the channel coefficients to perform channel prediction. For the sake of simplicity, the channel coefficients are modeled using an autoregressive model with the channel correlation describing the statistical dependency between the channel coefficients. Consequently, linear predictors have been employed to exploit the linear correlation between the channel coefficients. While many of the channel prediction algorithms have considered the linear correlation between the channel coefficients, some works have considered the possible nonlinear statistical dependencies between the channel coefficients for channel prediction [EK99].

In this section, the spatio-temporal channel properties of MIMO systems are analyzed to exploit the dependency between the channel coefficients for channel prediction without making any statistical assumption about the channel coefficients. The dependency between the channel coefficients is studied by considering the channel dynamics of the double directional channel model in time and frequency. Consequently, the double directional channel model based predictor in (6.2) is investigated to obtain significant insights about the dependency between the channel coefficients. These insights shall be used to develop optimal channel prediction algorithms.

The channel coefficient in (6.1), i.e.,

$$H_{m,n}(k,\ell) = \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}$$

is a weighted sum, with weights $\alpha_p$, of products of complex exponentials. The prediction of the channel coefficient at a prediction depth of $\Delta\ell$ in time and $\Delta k$ in frequency is given in (6.2), i.e.,

$$\begin{aligned} &H_{m,n}(k+\Delta k, \ell+\Delta\ell) = \\ &\quad \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell}. \end{aligned}$$

It can be seen that the channel coefficient $H_{m,n}(k+\Delta k, \ell+\Delta\ell)$ can be obtained from the channel coefficient $H_{m,n}(k,\ell)$ by rotating the phase of each of the products of the complex exponentials of $H_{m,n}(k,\ell)$ by an amount $\mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p}\mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell}$ which depends on the delays $\tau_p$ and the Doppler shifts $\nu_p$ of the propagation paths and the prediction depths $\Delta\ell$ and $\Delta k$. Thus the channel prediction process is a mere phase rotation of each of the products of the complex exponentials. Since each channel coefficient is a linear combination of products of complex exponentials with weights $\alpha_p$, it is possible to perform a linear combination of a sufficient number of channel coefficients such that the channel coefficient to be predicted with the required phase rotation can be obtained. Thus the channel coefficient at a prediction depth of $\Delta\ell$ and $\Delta k$ can be obtained by finding weights such that the linear combination of the products of the complex exponentials corresponding to each propagation path for each channel coefficient, i.e.,

$$\alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}},$$

yields the products of the complex exponentials corresponding to each propagation path for the channel coefficient to be predicted, i.e.,

$$\alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell}.$$

The question is then what is the minimum number of channel coefficients that are required to predict any channel coefficient. It is obvious that the minimum number of required channel coefficients depends on the number of propagation paths, i.e., the number of the products of complex exponentials constituting the channel coefficients. In fact, in the following, it will be proved that the minimum number of required channel coefficients is equal to the number of propagation paths. This result has been claimed in [Pal11] based on the investigation of the equivalence of the double directional channel model to the linear filter based channel model. Furthermore, similar arguments as in the preceding discussions have been used for performing temporal channel prediction of the narrowband version of the double directional channel model using an autoregressive analysis [GS04]. In the following, the proof of the preceding intuitive explanation, i.e., any channel coefficient can be predicted using a linear combination of other known channel coefficients, and the results in [GS04, Pal11] is given.

Fig. 6.1 shows an example channel prediction scenario. The unknown channel coefficient is to be predicted using $\acute{K}$ known channel coefficients considered in $\acute{L}$ time instants. The unknown channel coefficient is located at a distance of one time unit and minus one subcarrier unit from the latest channel coefficient. Thus the prediction depth is of $\Delta\ell = 1$ in time and $\Delta k = -1$ in frequency from the latest channel coefficient. It shall be proved that a filter which only depends on the prediction depths $\Delta\ell$ in time and $\Delta k$ frequency but which is independent of the actual time and frequency indices can be used to predict the unknown channel coefficient. As the filter coefficients are time and frequency independent, channel coefficients at the same prediction depths in time and frequency are predicted using the same filter coefficients. Let the known channel coefficients be denoted by a matrix

$$\mathbf{H}_{m,n}(k,\ell) = \begin{pmatrix} H_{m,n}(k,\ell) & \cdots & H_{m,n}(k,\ell-\acute{L}+1) \\ \vdots & & \vdots \\ H_{m,n}(k-\acute{K}+1,\ell) & \cdots & H_{m,n}(k-\acute{K}+1,\ell-\acute{L}+1) \end{pmatrix}. \tag{6.4}$$

Each of these channel coefficients shall be weighted by a filter coefficient matrix

$$\mathbf{\Phi}_{\Delta k,\Delta\ell} = \begin{pmatrix} \varphi_{\Delta k,\Delta\ell}^{(0,0)} & \cdots & \varphi_{\Delta k,\Delta\ell}^{(0,\acute{L}-1)} \\ \vdots & & \vdots \\ \varphi_{\Delta k,\Delta\ell}^{(\acute{K}-1,0)} & \cdots & \varphi_{\Delta k,\Delta\ell}^{(\acute{K}-1,\acute{L}-1)} \end{pmatrix} \tag{6.5}$$

to predict the unknown channel coefficient $H_{m,n}(k+\Delta k,\ell+\Delta\ell)$, i.e.,

$$H_{m,n}(k+\Delta k,\ell+\Delta\ell) = \boldsymbol{\varphi}_{\Delta k,\Delta\ell}^{\mathrm{T}} \cdot \mathrm{vec}(\mathbf{H}_{m,n}(k,\ell)), \tag{6.6}$$

where $\boldsymbol{\varphi}_{\Delta k,\Delta\ell} = \mathrm{vec}(\mathbf{\Phi}_{\Delta k,\Delta\ell})$ and $\mathrm{vec}(\cdot)$ is a matrix vectorization operator which transforms a matrix into a vector by stacking the columns of the matrix one underneath the

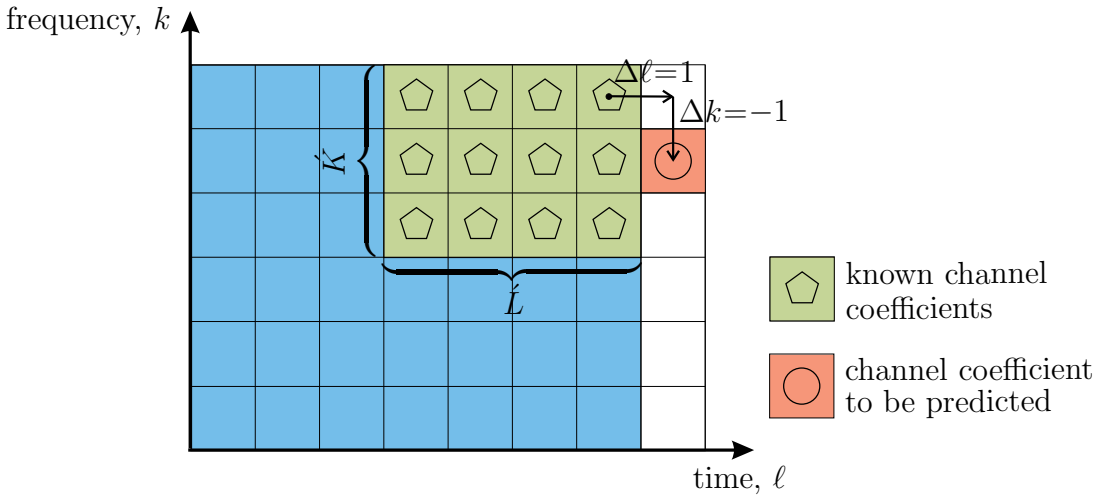



Figure 6.1. Prediction using $\acute{L} \cdot \acute{K}$ known channel coefficients

other. In derivation of the channel prediction based on the double directional channel model, it was not necessary to modify the steering factors $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}}$ and $\mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}$ owing to the narrowband assumption for array signal processing. Consequently, the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ is independent of the antenna pair under consideration, i.e., the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ is valid for each of the $M \cdot N$ transmitter and receiver antenna pairs.

The linear channel prediction in (6.6) can be re-written as

$$H_{m,n}(k+\Delta k, \ell+\Delta\ell) = \sum_{\acute{k}=0}^{\acute{K}-1} \sum_{\acute{\ell}=0}^{\acute{L}-1} \varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})} \cdot H_{m,n}(k-\acute{k}, \ell-\acute{\ell}). \tag{6.7}$$

Substituting for the channel coefficients using (6.1) results in

$$\begin{aligned} &\sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell} \\ &= \sum_{\acute{\ell}=0}^{\acute{L}-1} \sum_{\acute{k}=0}^{\acute{K}-1} \varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})} \sum_{p=1}^{P} \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \mathrm{e}^{\mathrm{j}2\pi F\acute{k}\tau_p} \mathrm{e}^{-\mathrm{j}2\pi\nu_p T\acute{\ell}}. \end{aligned} \tag{6.8}$$

Let the variable

$$\xi_{m,n,p}^{(k,\ell)} = \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}} \tag{6.9}$$

denote the common terms on the left and the right hand side of (6.8). Then (6.8) can be re-written as

$$\sum_{p=1}^{P} \xi_{m,n,p}^{(k,\ell)} \cdot \mathrm{e}^{-\mathrm{j}2\pi F\Delta k\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\Delta\ell} = \sum_{p=1}^{P} \xi_{m,n,p}^{(k,\ell)} \sum_{\acute{\ell}=0}^{\acute{L}-1} \sum_{\acute{k}=0}^{\acute{K}-1} \varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})} \mathrm{e}^{\mathrm{j}2\pi F\acute{k}\tau_p} \mathrm{e}^{-\mathrm{j}2\pi\nu_p T\acute{\ell}}. \tag{6.10}$$

Furthermore, defining

$$\psi_{\tau_p} = \mathrm{e}^{\mathrm{j}2\pi F\tau_p} \tag{6.11}$$

$$\psi_{\nu_p} = \mathrm{e}^{-\mathrm{j}2\pi T\nu_p} \tag{6.12}$$

and using these expressions to simplify (6.10) results in

$$\sum_{p=1}^{P} \xi_{m,n,p}^{(k,\ell)} \cdot \psi_{\tau_p}^{-\Delta k} \cdot \psi_{\nu_p}^{-\Delta \ell} = \sum_{p=1}^{P} \xi_{m,n,p}^{(k,\ell)} \sum_{\acute{\ell}=0}^{\acute{L}-1} \sum_{\acute{k}=0}^{\acute{K}-1} \varphi_{\Delta k,\Delta \ell}^{(\acute{k},\acute{\ell})} \cdot \psi_{\tau_p}^{\acute{k}} \cdot \psi_{\nu_p}^{\acute{\ell}}. \tag{6.13}$$

Equation (6.13) can be re-arranged to obtain a simplified expression as

$$\left(\boldsymbol{\xi}_{m,n}^{(k,\ell)}\right)^{\mathrm{T}} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell} = \left(\boldsymbol{\xi}_{m,n}^{(k,\ell)}\right)^{\mathrm{T}} \cdot \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta \ell}, \tag{6.14}$$

where

$$\boldsymbol{\Psi} = \begin{pmatrix} 1 & \psi_{\tau_1} & \cdots & \psi_{\tau_1}^{\acute{K}-1} & \psi_{\nu_1} & \psi_{\nu_1}\psi_{\tau_1} & \cdots & \psi_{\nu_1}\psi_{\tau_1}^{\acute{K}-1} & \cdots & \psi_{\nu_1}^{\acute{L}-1}\psi_{\tau_1}^{\acute{K}-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1 & \psi_{\tau_P} & \cdots & \psi_{\tau_P}^{\acute{K}-1} & \psi_{\nu_P} & \psi_{\nu_P}\psi_{\tau_P} & \cdots & \psi_{\nu_P}\psi_{\tau_P}^{\acute{K}-1} & \cdots & \psi_{\nu_P}^{\acute{L}-1}\psi_{\tau_P}^{\acute{K}-1} \end{pmatrix} \tag{6.15}$$

$$\boldsymbol{\xi}_{m,n}^{(k,\ell)} = \left(\xi_{m,n,1}^{(k,\ell)}, \ldots, \xi_{m,n,P}^{(k,\ell)}\right)^{\mathrm{T}} \tag{6.16}$$

$$\boldsymbol{\psi}_{\Delta k,\Delta \ell} = \left(\psi_{\tau_1}^{-\Delta k} \cdot \psi_{\nu_1}^{-\Delta \ell}, \ldots, \psi_{\tau_P}^{-\Delta k} \cdot \psi_{\nu_P}^{-\Delta \ell}\right)^{\mathrm{T}}. \tag{6.17}$$

As the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ is time and frequency independent, the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ holds for all $K$ channel coefficients considered in $L$ time instants. Furthermore, owing to the narrowband assumption for array signal processing, the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ is independent of the transmitter and receiver antenna pair under consideration. Consequently, a system of $M \cdot N \cdot K \cdot L$ linear equations can be obtained from (6.14). The resulting system of linear equations can be written as

$$\boldsymbol{\Xi} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell} = \boldsymbol{\Xi} \cdot \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta \ell}, \tag{6.18}$$

where $\boldsymbol{\Xi}$ is a $M \cdot N \cdot K \cdot L \times P$ matrix with elements obtained from (6.9), i.e.,

$$\xi_{m,n,p}^{(k,\ell)} = \alpha_p \mathrm{e}^{-\mathrm{j}2\pi(Fk+f_0)\tau_p} \mathrm{e}^{\mathrm{j}2\pi\nu_p T\ell} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Tx}}^{(n,p)}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{\mathrm{Rx}}^{(m,p)}}.$$

The matrix $\boldsymbol{\Xi}$ can be represented as

$$\boldsymbol{\Xi} = \begin{pmatrix} \boldsymbol{\Xi}_{1,1} \\ \boldsymbol{\Xi}_{1,2} \\ \vdots \\ \boldsymbol{\Xi}_{M,N} \end{pmatrix} \odot \begin{pmatrix} \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \\ \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \end{pmatrix}, \tag{6.19}$$

where

$$\boldsymbol{\Xi}_{m,n} = \begin{pmatrix} \alpha_1 \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{n,1}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{m,1}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_1} & \cdots & \alpha_P \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{n,P}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{m,P}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_P} \\ \vdots & & \vdots \\ \alpha_1 \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{n,1}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{m,1}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_1} & \cdots & \alpha_P \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{n,P}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_{m,P}} \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau_P} \end{pmatrix} \tag{6.20}$$

$$\boldsymbol{\Xi}_{\nu,\tau} = \underbrace{\begin{pmatrix} \mathrm{e}^{\mathrm{j}2\pi T \nu_1} & \cdots & \mathrm{e}^{\mathrm{j}2\pi T L \nu_1} \\ \vdots & & \vdots \\ \mathrm{e}^{\mathrm{j}2\pi T \nu_P} & \cdots & \mathrm{e}^{\mathrm{j}2\pi T L \nu_P} \end{pmatrix}}_{\boldsymbol{\Xi}_\nu} * \underbrace{\begin{pmatrix} \mathrm{e}^{-\mathrm{j}2\pi F \tau_1} & \cdots & \mathrm{e}^{-\mathrm{j}2\pi F K \tau_1} \\ \vdots & & \vdots \\ \mathrm{e}^{-\mathrm{j}2\pi F \tau_P} & \cdots & \mathrm{e}^{-\mathrm{j}2\pi F K \tau_P} \end{pmatrix}}_{\boldsymbol{\Xi}_\tau}. \tag{6.21}$$

The matrix $\boldsymbol{\Xi}_{m,n}$ is of dimension $K \cdot L \times P$ and the symbols $\odot$ and $*$ in (6.19) denote the Hadamard (element-wise) and Khatri-Rao (row-wise Kronecker) matrix products, respectively.

If there is a filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ which satisfies

$$\boldsymbol{\psi}_{\Delta k,\Delta \ell} = \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta \ell}, \tag{6.22}$$

then (6.18), i.e.,

$$\boldsymbol{\Xi} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell} = \boldsymbol{\Xi} \cdot \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta \ell}$$

is also satisfied. Thus (6.22) is a sufficient condition for (6.18). If the matrix $\boldsymbol{\Xi}$ has full column rank, then (6.22) is a necessary and sufficient condition for (6.18). This is due to the fact that if the matrix $\boldsymbol{\Xi}$ has full column rank, then the pseudo inverse of matrix $\boldsymbol{\Xi}$, i.e.,

$$(\boldsymbol{\Xi}^{*\mathrm{T}} \cdot \boldsymbol{\Xi})^{-1} \cdot \boldsymbol{\Xi}^{*\mathrm{T}}$$

can be multiplied on the left of both sides of (6.18) to obtain (6.22). Thus (6.22) is a necessary and sufficient condition if (6.7) shall hold for all $k$ and $\ell$ and all $M \cdot N$ transmitter and receiver antenna pairs, i.e., if the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ shall be independent of $k$ and $\ell$ and the transmitter and receiver antenna pair. In the following, it will be shown that the matrix $\boldsymbol{\Xi}$ has full column rank and hence (6.22) is a necessary and sufficient condition for (6.7) to hold.

To show that the matrix $\boldsymbol{\Xi}$ in (6.19), i.e.,

$$\boldsymbol{\Xi} = \begin{pmatrix} \boldsymbol{\Xi}_{1,1} \\ \boldsymbol{\Xi}_{1,2} \\ \vdots \\ \boldsymbol{\Xi}_{M,N} \end{pmatrix} \odot \begin{pmatrix} \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \\ \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}} \end{pmatrix}$$

has a full column rank, it is sufficient to show that the matrix $\boldsymbol{\Xi}_{m,n} \odot \boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}}$ has a full column rank. As it can be seen from (6.21), the matrix $\boldsymbol{\Xi}_{\nu,\tau}$ is the Khatri-Rao product

of the matrices $\boldsymbol{\Xi}_\nu$ and $\boldsymbol{\Xi}_\tau$. Furthermore, it can be seen that the matrices $\boldsymbol{\Xi}_\nu$ and $\boldsymbol{\Xi}_\tau$ are Vandermonde matrices. If the generators of the Vandermonde matrices, i.e., $\mathrm{e}^{\mathrm{j}2\pi T\ell\nu_p}$ for $\boldsymbol{\Xi}_\nu$ and $\mathrm{e}^{-\mathrm{j}2\pi Fk\tau_p}$ for $\boldsymbol{\Xi}_\tau$, are distinct, then the Vandermonde matrices $\boldsymbol{\Xi}_\nu$ and $\boldsymbol{\Xi}_\tau$ have full rank [GVL13]. Moreover, the Khatri-Rao product of the Vandermonde matrices $\boldsymbol{\Xi}_\nu \in \mathbb{C}^{P\times L}$ and $\boldsymbol{\Xi}_\tau \in \mathbb{C}^{P\times K}$ whose $2P$ complex exponential parameters $\nu_p$ and $\tau_p$ are drawn from a continuous distribution has almost surely full rank, i.e., for the Doppler shifts $\nu_p$ and delays $\tau_p$ drawn from a continuous distribution, the probability of the matrix $\boldsymbol{\Xi}_{\nu,\tau}$ being a full rank matrix is one [JStB01]. Consequently, for $K \cdot L > P$ the matrix $\boldsymbol{\Xi}_{\nu,\tau}$ has almost surely a full row rank of $P$ and hence the matrix $\boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}}$ has almost surely a full column rank of $P$. On the other hand, as it can be seen from (6.20) the matrix $\boldsymbol{\Xi}_{m,n}$ whose columns consist of identical elements has a rank of one. The Hadamard (element-wise) product of the matrices $\boldsymbol{\Xi}_{m,n}$ and $\boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}}$ results in a matrix which has almost surely full column rank as the Hadamard product only scales the columns of the matrix $\boldsymbol{\Xi}_{\nu,\tau}^{\mathrm{T}}$.

In the preceding discussions, it has been proved that the matrix $\boldsymbol{\Xi}$ has almost surely a full column rank. Thus it is left to show that the necessary and sufficient condition given in (6.22), i.e.,

$$\boldsymbol{\psi}_{\Delta k,\Delta\ell} = \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta\ell}$$

holds. According to the Rouché-Capelli theorem, (6.22), i.e., the above equation, has at least one solution if the rank of the augmented matrix $(\boldsymbol{\Psi}|\boldsymbol{\psi}_{\Delta k,\Delta\ell})$ is equal to the rank of the coefficient matrix $\boldsymbol{\Psi}$ [Ser93]. The matrix $\boldsymbol{\Psi}$ is the Khatri-Rao (row-wise Kronecker) product of two Vandermonde matrices, i.e.,

$$\begin{aligned}
\boldsymbol{\Psi} &= \begin{pmatrix} 1 & \psi_{\tau_1} & \cdots & \psi_{\tau_1}^{\acute{K}-1} & \psi_{\nu_1} & \psi_{\nu_1}\psi_{\tau_1} & \psi_{\nu_1}\psi_{\tau_1}^2 & \cdots & \psi_{\nu_1}^{\acute{L}-1}\psi_{\tau_1}^{\acute{K}-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & \psi_{\tau_P} & \cdots & \psi_{\tau_P}^{\acute{K}-1} & \psi_{\nu_P} & \psi_{\nu_P}\psi_{\tau_P} & \psi_{\nu_P}\psi_{\tau_P}^2 & \cdots & \psi_{\nu_P}^{\acute{L}-1}\psi_{\tau_P}^{\acute{K}-1} \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} 1 & \psi_{\nu_1} & \psi_{\nu_1}^2 & \cdots & \psi_{\nu_1}^{\acute{L}-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \psi_{\nu_P} & \psi_{\nu_P}^2 & \cdots & \psi_{\nu_P}^{\acute{L}-1} \end{pmatrix}}_{\boldsymbol{\Psi}_\nu} * \underbrace{\begin{pmatrix} 1 & \psi_{\tau_1} & \psi_{\tau_1}^2 & \cdots & \psi_{\tau_1}^{\acute{K}-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \psi_{\tau_P} & \psi_{\tau_P}^2 & \cdots & \psi_{\tau_P}^{\acute{K}-1} \end{pmatrix}}_{\boldsymbol{\Psi}_\tau}.
\end{aligned} \tag{6.23}$$

The Khatri-Rao product of the Vandermonde matrices $\boldsymbol{\Psi}_\nu \in \mathbb{C}^{P\times\acute{L}}$ and $\boldsymbol{\Psi}_\tau \in \mathbb{C}^{P\times\acute{K}}$ whose $2P$ complex exponential parameters $\nu_p$ and $\tau_p$ are drawn from a continuous distribution has almost surely full rank [JStB01]. Thus the matrix $\boldsymbol{\Psi}$ has almost surely full rank. Furthermore, for $\acute{L} \cdot \acute{K} \geq P$ the augmented matrix $(\boldsymbol{\Psi}|\boldsymbol{\psi}_{\Delta k,\Delta\ell})$ has a rank equal to the rank of the coefficient matrix $\boldsymbol{\Psi}$. Thus there exists almost surely a unique filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ which satisfies (6.22) if $\acute{L} \cdot \acute{K} = P$. For $\acute{L} \cdot \acute{K} > P$ there are almost surely infinitely many filter coefficient vectors $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ which satisfy (6.22).

For the case where the number of unknowns in (6.22) is equal to the number of equations, i.e., $\acute{L} \cdot \acute{K} = P$, there is a unique filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$. The optimal estimate $\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell}$ of the unique filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$, in the least squares sense, can be determined from (6.22) as

$$\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell} = \left(\boldsymbol{\Psi}^{*\mathrm{T}} \cdot \boldsymbol{\Psi}\right)^{-1} \cdot \boldsymbol{\Psi}^{*\mathrm{T}} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell}. \tag{6.24}$$

On the other hand, for the case where the number of unknowns in (6.22) is greater than the number of equations, i.e., $\acute{L} \cdot \acute{K} > P$, there are an infinite number of filter coefficient vectors $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ which satisfy the underdetermined system of linear equations of (6.22). Out of the infinitely many possible filter coefficient vectors, the filter coefficient vector with the minimum norm is of interest. The choice of the norm has crucial influence on the complexity of finding the filter coefficient vector and the qualitative behaviour of the filter in the presence of noise. For example, the filter coefficient vector with the minimum $l0$-norm, i.e., the filter coefficient vector with the smallest possible number of non-zero elements, has no closed-form solution. Furthermore, the filter coefficient vector with the minimum $l0$-norm is disadvantageous in that not all channel coefficients are exploited in predicting the required channel coefficient. This results in severe performance degradation in the presence of noise. In the following, owing to the ease of finding a closed-form solution and its stability in the presence of noise, the minimum Euclidean norm is considered. The filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ with the minimum Euclidean norm can be determined by solving

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell} = \underset{\boldsymbol{\varphi}_{\Delta k,\Delta \ell}}{\arg\min} \quad & \left\|\boldsymbol{\varphi}_{\Delta k,\Delta \ell}\right\|_2 \\ \text{subject to} \quad & \boldsymbol{\psi}_{\Delta k,\Delta \ell} = \boldsymbol{\Psi} \cdot \boldsymbol{\varphi}_{\Delta k,\Delta \ell}. \end{aligned} \tag{6.25}$$

The minimum Euclidean norm solution to the system of linear equations of (6.25) can be determined using the Moore-Penrose pseudoinverse as

$$\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell} = \boldsymbol{\Psi}^{+} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell}, \tag{6.26}$$

where the superscript + denotes the Moore-Penrose pseudoinverse. For $\acute{L} \cdot \acute{K} > P$, the rows of $\boldsymbol{\Psi}$ are linearly independent and hence the Moore-Penrose pseudoinverse $\boldsymbol{\Psi}^{+}$ is the right inverse of $\boldsymbol{\Psi}$. Consequently, the estimate $\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell}$ of the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ which minimizes the Euclidean norm is given as

$$\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell} = \boldsymbol{\Psi}^{*\mathrm{T}} \cdot \left(\boldsymbol{\Psi} \cdot \boldsymbol{\Psi}^{*\mathrm{T}}\right)^{-1} \cdot \boldsymbol{\psi}_{\Delta k,\Delta \ell}. \tag{6.27}$$

It must be noted that even though the matrix $\boldsymbol{\Psi}$ has almost surely full rank, finding estimates of the filter coefficient vector $\hat{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell}$ from (6.22) can in some cases result in an ill-posed problem. Consequently, consideration of regularization methods might be

necessary. A well-posed problem in (6.22) can be obtained by considering cases where $\acute{L} \cdot \acute{K} > P$.

In the discussions so far, it has been assumed that the unknown channel is to be predicted based on $\acute{K}$ known channel coefficients considered in $\acute{L}$ time instants. The consideration of consecutive channel coefficients in time and frequency has resulted in the matrix $\mathbf{\Psi}$ which is the Khatri-Rao product two Vandermonde matrices. Owing to its special structure, the matrix $\mathbf{\Psi}$ has almost surely full rank. Consequently, it has been shown that there exists a filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ with dimension of at least $P$. As discussed earlier, the linear relationship between the channel coefficients arises due to their description as a weighted sum of products of complex exponentials. Thus the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ shall exist regardless of the sequence of the known channel coefficients. Towards this end, consider the channel prediction scenario in Fig. 6.2 with arbitrarily chosen known channel coefficients. The unknown channel coefficient is to be predicted using $R$ arbitrarily chosen known channel coefficients. Without loss of any generality, the unknown channel coefficient is chosen to be located at a distance of one time unit and minus one subcarrier unit from the latest channel coefficient, i.e., $\Delta \ell = 1$ and $\Delta k = -1$. In the following, it will be proved that there exists a filter coefficient vector which can predict the unknown channel coefficient using $R \geq P$ arbitrarily chosen known channel coefficients.

Following a similar argument as in the preceding discussion, a system of linear equations

$$\grave{\boldsymbol{\psi}}_{\Delta k,\Delta \ell} = \grave{\mathbf{\Psi}} \cdot \grave{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell}, \tag{6.28}$$

can be obtained for the case of arbitrarily chosen known channel coefficients. It can be shown that the matrix $\grave{\mathbf{\Psi}}$ has almost surely full rank. Towards this end, $\acute{K}$ consecutive channel coefficients in $\acute{L}$ time instants such that all of the $R$ arbitrarily chosen known channel coefficients are contained in the $\acute{L} \cdot \acute{K}$ channel coefficients shall be considered. Then the matrix $\grave{\mathbf{\Psi}}$ is nothing but a submatrix of

$$\mathbf{\Psi} = \begin{pmatrix} 1 & \psi_{\tau_1} & \cdots & \psi_{\tau_1}^{\acute{K}-1} & \psi_{\nu_1} & \psi_{\nu_1}\psi_{\tau_1} & \cdots & \psi_{\nu_1}\psi_{\tau_1}^{\acute{K}-1} & \cdots & \psi_{\nu_1}^{\acute{L}-1}\psi_{\tau_1}^{\acute{K}-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1 & \psi_{\tau_P} & \cdots & \psi_{\tau_P}^{\acute{K}-1} & \psi_{\nu_P} & \psi_{\nu_P}\psi_{\tau_P} & \cdots & \psi_{\nu_P}\psi_{\tau_P}^{\acute{K}-1} & \cdots & \psi_{\nu_P}^{\acute{L}-1}\psi_{\tau_P}^{\acute{K}-1} \end{pmatrix}, \tag{6.29}$$

with $R$ columns drawn from $\mathbf{\Psi}$ corresponding to the positions of the $R$ channel coefficients in the $\acute{L} \cdot \acute{K}$ channel coefficients. Thus the matrix $\grave{\mathbf{\Psi}}$ has almost surely full rank. Consequently, for $R \geq P$ there is almost surely at least one filter coefficient vector $\grave{\boldsymbol{\varphi}}_{\Delta k,\Delta \ell}$ which can be used to predict the unknown channel coefficient.

Thus it can be concluded that any unknown channel coefficient can be predicted as a linear combination of at least $P$ other known channel coefficients. The matrix $\mathbf{\Psi}$ and

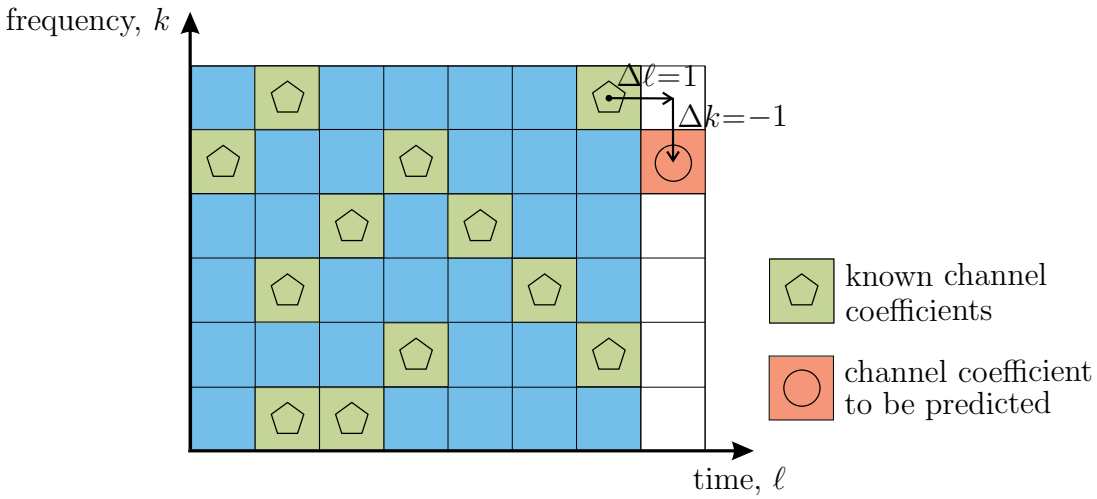



Figure 6.2. Prediction using arbitrarily chosen known channel coefficients

the vector $\boldsymbol{\psi}_{\Delta k,\Delta\ell}$ are determined using the delays $\tau_p$ and the Doppler shifts $\nu_p$ of the propagation paths. Thus the prediction filter vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ depends on the delays $\tau_p$ and the Doppler shifts $\nu_p$ of the propagation paths and the prediction depths $\Delta\ell$ and $\Delta k$ but is independent of the actual time and frequency.

It may seem counter-intuitive that even though the double directional channel model has $P$ degrees of freedom in time and in frequency, it is possible to predict a channel coefficient in time, in frequency and in time and frequency using only $P$ known channel coefficients. This result stems from the fact that each channel coefficient is a weighted sum of products of complex exponentials. The products of complex exponentials describe the phase rotations in time and frequency. As small scale MS mobility is assumed, the propagation path parameters are assumed to be constant. Thus each channel coefficient is a weighted sum of products of complex exponentials with the same weights for each channel coefficient. Consequently, for predicting a channel coefficient it is only required to obtain the $P$ appropriate overall phase rotations in time and frequency. For a given channel coefficient to be predicted, this can be obtained from a proper linear combination of $P$ known channel coefficients.

The consequence of the preceding discussions is that MIMO channel prediction can be performed by exploiting the linear relation between the channel coefficients without having to estimate the propagation path parameters. Owing to the linear relation between the channel coefficients, given a sufficient number of estimated channel coefficients, a system of linear equations can be set up from the estimated channel coefficients. Then the filter coefficient vector can be estimated from the set up system of linear equations by solving for the filter coefficient vector. Linear estimation techniques can be used to

mitigate the impact of noises in the estimated channel coefficients in the estimation of the filter coefficient vector.

In the discussions so far, it has been assumed that the propagation path parameters are time and frequency independent. In practice, a variation of the propagation path parameters can occur with time and frequency. It is possible to account for the change in the path parameters by introducing noise, i.e.,

$$H_{m,n}(k+\Delta k, \ell+\Delta\ell) = \sum_{\acute{k}=0}^{\acute{K}-1} \sum_{\acute{\ell}=0}^{\acute{L}-1} \varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})} \cdot H_{m,n}(k-\acute{k}, \ell-\acute{\ell}) + u_{m,n}(k+\Delta k, \ell+\Delta\ell), \tag{6.30}$$

where $u_{m,n}(k+\Delta k, \ell+\Delta\ell)$ is a zero mean complex white Gaussian noise with variance $\sigma_u^2$. It must be noted that for time and frequency independent path parameters, the poles of the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ are on the unit circle, whereas for the case where the change in the path parameters is accounted for as in (6.30), the poles of the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k,\Delta\ell}$ are inside the unit circle [Mak75].

# Chapter 7

# Performance gains from tracking a MIMO channel

## 7.1 Tracking a MIMO channel in time-division-duplex systems

In the previous chapter, it has been shown, based on the double directional channel model, that there is a linear relation between the channel coefficients in time and frequency. Consequently, linear channel estimation algorithms can be used to obtain optimal estimates of the channel coefficients. In the following, training signal based channel estimation techniques are considered to obtain estimates of the channel coefficients. In training signal based channel estimation, the transmitter transmits a priori known pilot symbols multiplexed with useful data symbols [Cav91]. The receiver exploits the received a priori known pilot symbols to estimate the channel at the pilot symbol locations. The channels at other locations can then be estimated using interpolation techniques. Furthermore, channel prediction at a desired point in time and frequency can be performed by exploiting the linear relation between the channel coefficients. In performing channel prediction, it is possible to exploit only the current received pilot symbols. This approach would make sense if the radio propagation scenario changes so swiftly that the channel coefficients at different pilot symbol transmission times are independent. However, in this thesis small scale MS mobility is assumed. Consequently, the radio propagation scenario remains unchanged or changes slightly for a duration that is larger than multiples of the transmission duration of the pilot symbols. Thus the present and the past received pilot symbols could be exploited in predicting the required channel coefficients. Consequently, significant performance gains in channel estimation can be obtained from tracking the mobile radio channel.

In the following, tracking a MIMO channel in time-division-duplex (TDD) systems is considered. In TDD systems, the uplink data and the downlink data are transmitted at different times. The uplink and the downlink channel are identical if the duplexing time is much smaller than the coherence time of the channel [Mol05]. Thus, in general, the uplink and the downlink channels are not identical. In the following, the uplink and the downlink MIMO channels are to be tracked using the received pilot symbols at the BS and the MS, respectively. As it has been discussed in Chapter 4, owing to its convenience and intuitive description of dynamical systems, the state-space approach

is preferable for modeling dynamical systems. Thus the MIMO channel tracking problem in TDD systems is described using a state-space model. The state-space model describes a dynamical system using a state transition model and a measurement model. The state transition model describes the change of the state of the system with time, whereas the measurement model describes the relation between the measurements and the state of the system.

Since training signal based channel estimation is considered, the channel coefficients that are considered are the channel coefficients of the pilot symbol subcarriers. Consequently, the subcarrier index $k = 1, \ldots, K$ is re-used to index the pilot symbol subcarriers instead of all the subcarriers. Furthermore, the time index $\ell = 1, \ldots, L$ is re-used to index the sequence of the pilot symbols in time. The state vector $\boldsymbol{\theta}(\ell)$ represents the state of the system at discrete time $\ell$. As it has been shown in the previous chapter, each of the channel coefficients can be represented as a linear combination of other channel coefficients which are at least equal, in number, to the number of the propagation paths. Thus the state vector representing each channel coefficient shall consider other channel coefficients which are at least equal, in number, to the number of the propagation paths. Nevertheless, significant performance gains could be obtained if the relation between all channel coefficients of the pilot symbol subcarriers is exploited in tracking the evolution of the channel coefficients. To this end, all the channel coefficients at a discrete time instant $\ell$ between the $m^{\text{th}}$ receive antenna and the $n^{\text{th}}$ transmit antenna shall be considered, i.e.,

$$\mathbf{h}_{m,n}(\ell) = (H_{m,n}(K, \ell), H_{m,n}(K-1, \ell), \ldots, H_{m,n}(1, \ell))^{\mathrm{T}}. \tag{7.1}$$

Stacking such vectors from all $M \cdot N$ transmit and receive antennas results in a vector

$$\mathbf{h}(\ell) = (\mathbf{h}_{1,1}^{\mathrm{T}}(\ell), \mathbf{h}_{1,2}^{\mathrm{T}}(\ell), \ldots, \mathbf{h}_{M,N}^{\mathrm{T}}(\ell))^{\mathrm{T}}. \tag{7.2}$$

Considering the vector $\mathbf{h}(\ell)$ for $\acute{L}$ time instants means that the evolution of each channel coefficient, for each of the $M \cdot N$ transmit and receive antennas, is represented by $\acute{K} \cdot \acute{L}$ other channel coefficients where $\acute{K} = K$, i.e., all pilot symbol subcarriers. Each channel coefficient can be represented as a linear combination of $\acute{K} \cdot \acute{L}$ other channel coefficients as long as $\acute{K} \cdot \acute{L} \geq P$, where $P$ is the number of propagation paths. Thus considering the vector $\mathbf{h}(\ell)$ for $\acute{L}$ time instants results in the state vector

$$\boldsymbol{\theta}(\ell) = (\mathbf{h}^{\mathrm{T}}(\ell), \mathbf{h}^{\mathrm{T}}(\ell-1), \ldots, \mathbf{h}^{\mathrm{T}}(\ell-\acute{L}+1))^{\mathrm{T}}. \tag{7.3}$$

The state transition equation describes the evolution of each of the channel coefficients with time. If the path parameters are constant within the time duration and frequency

band under consideration, then each of the channel coefficients can be described, without error, as a linear combination of a sufficient number of other channel coefficients. Otherwise, it is possible to account for the change in the path parameters by introducing noise as given in (6.30), i.e.,

$$H_{m,n}(k+\Delta k,\ell+\Delta\ell)=\sum_{\acute{k}=0}^{\acute{K}-1}\sum_{\acute{\ell}=0}^{\acute{L}-1}\varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})}\cdot H_{m,n}(k-\acute{k},\ell-\acute{\ell})+u_{m,n}(k+\Delta k,\ell+\Delta\ell),$$

where $u_{m,n}(k+\Delta k,\ell+\Delta\ell)$ is a zero mean complex white Gaussian noise with variance $\sigma_u^2$. Consequently, the state transition equation is linear. Thus using the above equation the state transition equation of the state space model can be defined as

$$\boldsymbol{\theta}(\ell)=\mathbf{B}\cdot\boldsymbol{\theta}(\ell-1)+\boldsymbol{\Gamma}\cdot\mathbf{u}(\ell), \tag{7.4}$$

where the $M\cdot N\cdot\acute{K}\times 1$ vector $\mathbf{u}(\ell)$ is assumed to be zero mean complex multivariate white Gaussian distributed with covariance matrix $\mathbf{R}_{\mathrm{uu}}$. The $M\cdot N\cdot\acute{K}\cdot\acute{L}\times M\cdot N\cdot\acute{K}\cdot\acute{L}$ state transition matrix $\mathbf{B}$ and the $M\cdot N\cdot\acute{K}\cdot\acute{L}\times M\cdot N\cdot\acute{K}$ matrix $\boldsymbol{\Gamma}$ are defined as

$$\mathbf{B}=\left(\begin{array}{cccccccc|ccc}\mathbf{D}^{(1)}&\cdots&\mathbf{0}_{\acute{K}}&\mathbf{D}^{(2)}&\cdots&\mathbf{0}_{\acute{K}}&\cdots&\mathbf{D}^{(\acute{L})}&\cdots&\mathbf{0}_{\acute{K}}\\ \vdots&\ddots&\vdots&\vdots&\ddots&\vdots&&\vdots&\ddots&\vdots\\ \mathbf{0}_{\acute{K}}&\cdots&\mathbf{D}^{(1)}&\mathbf{0}_{\acute{K}}&\cdots&\mathbf{D}^{(2)}&\cdots&\mathbf{0}_{\acute{K}}&\cdots&\mathbf{D}^{(\acute{L})}\\ \hline \multicolumn{8}{c|}{\mathbf{I}_{M\cdot N\cdot\acute{K}\cdot(\acute{L}-1)}}&\multicolumn{3}{c}{\mathbf{0}_{M\cdot N\cdot\acute{K}\cdot(\acute{L}-1)\times M\cdot N\cdot\acute{K}}}\end{array}\right) \tag{7.5}$$

$$\boldsymbol{\Gamma}=\left(\begin{array}{cc}\mathbf{I}_{M\cdot N\cdot\acute{K}}&\mathbf{0}_{M\cdot N\cdot\acute{K}\times M\cdot N\cdot\acute{K}\cdot(\acute{L}-1)}\end{array}\right)^{\mathrm{T}}. \tag{7.6}$$

The matrices $\mathbf{I}_{M\cdot N\cdot\acute{K}}$ and $\mathbf{0}_{M\cdot N\cdot\acute{K}}$ denote the identity and the zero matrices of dimension $M\cdot N\cdot\acute{K}$, respectively. The $\acute{K}\times\acute{K}\cdot\acute{L}$ filter coefficient matrix

$$\mathbf{D}=\left(\begin{array}{ccc}\mathbf{D}^{(1)}&\cdots&\mathbf{D}^{(\acute{L})}\end{array}\right) \tag{7.7}$$

weights the channel coefficient vector $\left(\mathbf{h}_{m,n}^{\mathrm{T}}(\ell-1),\ldots,\mathbf{h}_{m,n}^{\mathrm{T}}(\ell-\acute{L})\right)^{\mathrm{T}}$ constituting the state vector $\boldsymbol{\theta}(\ell-1)$ to yield the channel coefficient vector $\mathbf{h}_{m,n}(\ell)$. The filter coefficient matrix $\mathbf{D}$ can be calculated, given the delays $\tau_p$ and Doppler shifts $\nu_p$, using (6.27) as

$$\mathbf{D}=\left(\begin{array}{ccc}\boldsymbol{\varphi}_{\Delta k=0,\Delta\ell=1}&\cdots&\boldsymbol{\varphi}_{\Delta k=1-K,\Delta\ell=1}\end{array}\right)^{\mathrm{T}}. \tag{7.8}$$

The $\acute{K}\cdot\acute{L}\times 1$ filter coefficient vector $\boldsymbol{\varphi}_{\Delta k=k-K,\Delta\ell=1}$ weights the channel coefficient vector $\left(\mathbf{h}_{m,n}^{\mathrm{T}}(\ell-1),\ldots,\mathbf{h}_{m,n}^{\mathrm{T}}(\ell-\acute{L})\right)^{\mathrm{T}}$ to yield the channel coefficient $H_{m,n}(k,\ell)$. The prediction depths $\Delta\ell=1$ in time and $\Delta k=k-K$ in frequency are chosen as the channel coefficient $H_{m,n}(k,\ell)$ is located at a prediction depths of $\Delta\ell=1$ in time and $\Delta k=k-K$ in frequency from the reference channel coefficient $H_{m,n}(K,\ell-1)$, see Fig. 6.1.

The measurement equation of the state-space model can be developed by considering the signal model. It is assumed that training signal based channel estimation techniques are employed to obtain measurements of the channel coefficients. Let the matrix $\mathbf{S}(\ell)$ denote a diagonal matrix of size $M \cdot N \cdot \acute{K}$ with elements on the diagonal being the transmitted pilot symbols. The $M \cdot N \cdot \acute{K} \times 1$ received pilot symbol vector $\mathbf{e}(\ell)$ is then given as

$$\mathbf{e}(\ell) = \underbrace{\begin{pmatrix} \mathbf{S}(\ell) & \mathbf{0}_{M\cdot N\cdot \acute{K} \times M\cdot N\cdot \acute{K}\cdot(\acute{L}-1)} \end{pmatrix}}_{\mathbf{G}(\ell)} \cdot \boldsymbol{\theta}(\ell) + \mathbf{n}(\ell), \tag{7.9}$$

where $\mathbf{n}(\ell)$ is the measurement noise which is a multivariate zero mean complex white Gaussian noise with covariance matrix $\mathbf{R}_{\mathrm{nn}}$. Thus (7.9) defines the measurement equation of the state space model with the vector $\mathbf{e}(\ell)$ being the measurement vector.

It can be seen that both the state transition and the measurement equations are linear and Gaussian. Consequently, the Kalman filter can be used to obtain optimal estimates, in the minimum mean square error sense, of the state vector. The Kalman filter tracks the estimate of the state vector $\hat{\boldsymbol{\theta}}(\ell|\ell)$ and the correlation matrix of the estimation error $\mathbf{M}(\ell|\ell)$ based on the measurements $\mathbf{e}(1), \ldots, \mathbf{e}(\ell)$. The Kalman filter starts with an initial estimate $\hat{\boldsymbol{\theta}}(0|0) = \mathbf{0}_{M\cdot N\cdot \acute{K}\cdot \acute{L}\times 1}$ of the initial state vector $\boldsymbol{\theta}(0|0)$ and an initial covariance matrix $\mathbf{M}(0|0) = \mathbf{I}_{M\cdot N\cdot \acute{K}\cdot \acute{L}}$ of the estimation error and recursively computes estimates $\hat{\boldsymbol{\theta}}(\ell|\ell)$ of the state vector $\boldsymbol{\theta}(\ell)$ and the covariance matrix $\mathbf{M}(\ell|\ell)$ of the estimation error using the following equations

- prediction step:

$$\mathbf{M}(\ell|\ell-1) = \mathbf{B} \cdot \mathbf{M}(\ell-1|\ell-1) \cdot \mathbf{B}^{*\mathrm{T}} + \boldsymbol{\Gamma} \cdot \mathbf{R}_{\mathrm{uu}} \cdot \boldsymbol{\Gamma}^{*\mathrm{T}} \tag{7.10}$$

$$\hat{\boldsymbol{\theta}}(\ell|\ell-1) = \mathbf{B} \cdot \hat{\boldsymbol{\theta}}(\ell-1|\ell-1) \tag{7.11}$$

- update step:

$$\mathbf{K}(\ell) = \frac{\mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell)}{\mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell) + \mathbf{R}_{\mathrm{nn}}} \tag{7.12}$$

$$\hat{\boldsymbol{\theta}}(\ell|\ell) = \hat{\boldsymbol{\theta}}(\ell|\ell-1) + \mathbf{K}(\ell) \cdot \Big(\mathbf{e}(\ell) - \mathbf{G}(\ell) \cdot \hat{\boldsymbol{\theta}}(\ell|\ell-1)\Big) \tag{7.13}$$

$$\mathbf{M}(\ell|\ell) = \mathbf{M}(\ell|\ell-1) - \mathbf{K}(\ell) \cdot \mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1). \tag{7.14}$$

The matrix $\mathbf{K}(\ell)$ is the Kalman gain. Estimates of the channel coefficients at time instant $\ell$ are obtained from $\hat{\boldsymbol{\theta}}(\ell|\ell)$, whereas the predicted channel coefficients at time instant $\ell$ are obtained from $\hat{\boldsymbol{\theta}}(\ell|\ell-1)$.

## 7.2 Channel prediction using an adaptive Kalman filter

The Kalman filter can be used to obtain optimal estimates of the channel coefficients. However, in order to perform optimal estimation of the channel coefficients using the Kalman filter, estimates of the delays and the Doppler shifts of each propagation path are required to calculate the state transition matrix $\mathbf{B}$. For the estimation problem under consideration, it is assumed that only estimates of the received pilot symbols are available. In such cases, an estimate of the Kalman filter parameter, i.e., the state transition matrix, is obtained by considering stochastic models of the mobile radio channel. Many stochastic channel models are based on the wide-sense stationary uncorrelated scattering (WSSUS) model which is valid for many mobile radio channels of interest [Bel63]. The WSSUS model assumes that the second-order moment, i.e., the autocorrelation function, of the channel does not change with time and the scattering at two different paths are statistically uncorrelated. Considering zero mean normalized channel coefficients, i.e., $\mathbb{E}\{H_{m,n}(k,\ell)\} = 0$ and $\mathbb{E}\{|H_{m,n}(k,\ell)|^2\} = 1$, the autocorrelation function for the WSSUS model defined as

$$r_{\mathrm{f,t}}(\breve{k},\breve{\ell}) = \mathbb{E}\{H_{m,n}(k,\ell) \cdot H^*_{m,n}(k-\breve{k},\ell-\breve{\ell})\}. \tag{7.15}$$

The WSSUS model is thus characterized by the autocorrelation function of the channel.

In the stochastic approach to mobile radio channel modeling, the current channel coefficients are often assumed to be linear functions of past channel coefficients. Consequently, the time and frequency dynamics of the channel are modeled using an autoregressive model. For autoregressive processes with stationary autocorrelation function, the parameters of the autoregressive model can be easily determined from the so-called Yule-Walker equations. A brief overview of autoregressive modeling and determination of the model parameters using the Yule-Walker equations is given in Appendix C. The parameters of the autoregressive model, i.e., the filter coefficients and the standard deviation of the process noise, can be easily used to determine the parameters of the Kalman filter, i.e., state transition matrix and the covariance matrix of the state transition noise. Thus using stochastic channel models it is possible to determine the parameters of the Kalman filter using the autocorrelation function of the channel derived under the WSSUS assumption.

It is often assumed that the autocorrelation function of (7.15) is separable in to time-domain and frequency-domain correlations, i.e.,

$$r_{\mathrm{f,t}}(\breve{k},\breve{\ell}) = r_{\mathrm{f}}(\breve{k}) \cdot r_{\mathrm{t}}(\breve{\ell}), \tag{7.16}$$

where $r_{\mathrm{f}}(\breve{k}) = \mathbb{E}\{H_{m,n}(k,\ell) \cdot H^*_{m,n}(k-\breve{k},\ell)\}$ and $r_{\mathrm{t}}(\breve{\ell}) = \mathbb{E}\{H_{m,n}(k,\ell) \cdot H^*_{m,n}(k,\ell-\breve{\ell})\}$. A commonly used model for the temporal correlations is the Jakes' model where it is assumed that the propagation paths impinge uniformly from all directions [Jak74]. Each of these paths has an associated Doppler shift dependent on the direction of arrival at the receiver. For the Jakes' model the temporal correlation function becomes

$$r_{\mathrm{t}}(\breve{\ell}) = J_0\left(2\pi\nu_{\max}\breve{\ell}T\right), \tag{7.17}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind and $\nu_{\max}$ is the maximum Doppler frequency. The symbol $T$ stands for the transmit symbol duration. On the other hand, the spectral correlations are often derived from the truncated one-sided exponential power delay profile (PDP) [Pro95, WPY07]

$$\begin{aligned} A_h(\tau) &= \mathbb{E}\left\{h_{m,n}(\tau,t)h^*_{m,n}(\tau,t)\right\} \\ &= \begin{cases} \frac{c}{b}\mathrm{e}^{-\frac{\tau}{b}} & 0 \le \tau \le \tau_{\max} \\ 0 & \text{else}\ , \end{cases} \end{aligned} \tag{7.18}$$

where $h_{m,n}(\tau,t)$ is the time varying impulse response of the channel. The parameters $b$ and $\tau_{\max}$ depend on the radio channel scenario and $c = 1/(1-\mathrm{e}^{-\frac{\tau_{\max}}{b}})$ is a normalization constant which normalizes $A_h(\tau)$ to unity. The Fourier transform of the power delay profile yields the spectral correlation function

$$r_{\mathrm{f}}(\breve{k}) = c\frac{b(1-\mathrm{e}^{-\tau_{\max}(b+\mathrm{j}2\pi\breve{k}F)})}{b+\mathrm{j}2\pi\breve{k}F}. \tag{7.19}$$

Thus using the WSSUS model, information about the maximum Doppler shift and the multipath spread of the channel can be used to obtain estimates of the state transition matrix and the covariance matrix of the state transition noise. Estimating the maximum Doppler shift and multipath spread of a channel is considerably easier than estimating the Doppler shift and the delay of each propagation path of a mobile radio channel. Consequently, the Jakes' model and the truncated one-sided exponential PDP are commonly used in channel tracking and prediction [CZ04, MCCK07]. However, the stochastic process for many mobile radio channels of interest is not ergodic. Thus the stochastic channel models describe only the statistics of the whole ensemble of the channel realizations but not the individual channel realizations. For example, several channel measurements have shown that there are only a handful of propagation paths with significant powers [HRS+99, PW09a]. This is unlike the assumption in the Jakes' model where infinitely many propagation paths impinge the receiver uniformly from all directions. Thus the stochastic channel models do not reflect the statistics of a single channel realization and the estimated channel autocorrelation function may not be accurate enough to be used in practice with high reliability. In order to alleviate

this problem, in this thesis, the Jakes' model and the truncated one-sided exponential PDP are used only for initially estimating the parameters of the Kalman filter. The estimated channel coefficients, which are obtained from the Kalman filter, are then used to periodically update the parameters of the Kalman filter. This results in an adaptive Kalman filter which matches the actual mobile radio channel. Fig. 7.1 shows a block diagram of the adaptive Kalman filter for MIMO channel tracking in TDD systems.

Let $\hat{H}_{m,n}(k,\ell)$ denote the estimated channel coefficient of the $k^{\text{th}}$ subcarrier at the $\ell^{\text{th}}$ time instant of the $m^{\text{th}}$ receive antenna and the $n^{\text{th}}$ transmit antenna. The filter coefficients $\varphi_{\Delta k,\Delta\ell}^{(\acute{k},\acute{\ell})}$ which constitute the coefficient matrices $\mathbf{D}^{(\acute{\ell})}$ can be estimated from the estimated channel coefficients $\hat{H}_{m,n}(k,\ell)$ by exploiting the linear relation between the channel coefficients. Towards this end, a system of linear equations can be set up from the estimated channel coefficients using (6.7), i.e.,

$$\hat{H}_{m,n}(k,\ell) \approx \sum_{\acute{k}=0}^{\acute{K}-1} \sum_{\acute{\ell}=1}^{\acute{L}} \varphi_{\Delta k=k-K,\Delta\ell=1}^{(\acute{k},\acute{\ell})} \cdot \hat{H}_{m,n}(K-\acute{k},\ell-\acute{\ell}). \tag{7.20}$$

As the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k=k-K,\Delta\ell=1}$ is of dimension $\acute{K}\cdot\acute{L} \geq P$, then a system of at least $\acute{K}\cdot\acute{L} \geq P$ linear equations is required to solve for the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k=k-K,\Delta\ell=1}$. The required system of linear equations can be obtained by considering different sets of estimated channel coefficients which overlap in time. Fig. 7.2 shows different sets of estimated channel coefficients which overlap in time. Thus using $\acute{K}\cdot\acute{L} \geq P$ overlapping sets of estimated channel coefficients, a system of $\acute{K}\cdot\acute{L} \geq P$ linear equations can be generated which can be solved for the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k=k-K,\Delta\ell=1}$.

A system of linear equations can also be set up which can be used to directly solve for the coefficient matrices $\mathbf{D}^{(\acute{\ell})}$. Equation (7.21) shows a system of linear equations set up using the estimated channel coefficients from all $M\cdot N$ transmit and receive antennas.

$$\underbrace{\begin{pmatrix} \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}+1) \\ \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}+2) \\ \vdots \\ \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\ell) \\ \vdots \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}+1) \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}+2) \\ \vdots \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\ell) \end{pmatrix}}_{\mathbf{Y}(\ell)} \approx \underbrace{\begin{pmatrix} \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}) & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}-1) & \cdots & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(1) \\ \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}+1) & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\acute{L}) & \cdots & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(2) \\ \vdots & \vdots & & \vdots \\ \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\ell-1) & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\ell-2) & \cdots & \hat{\mathbf{h}}_{1,1}^{*\mathrm{T}}(\ell-\acute{L}) \\ \vdots & \vdots & & \vdots \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}) & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}-1) & \cdots & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(1) \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}+1) & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\acute{L}) & \cdots & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(2) \\ \vdots & \vdots & & \vdots \\ \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\ell-1) & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\ell-2) & \cdots & \hat{\mathbf{h}}_{M,N}^{*\mathrm{T}}(\ell-\acute{L}) \end{pmatrix}}_{\mathbf{X}(\ell)} \cdot \underbrace{\begin{pmatrix} (\mathbf{D}^{(1)})^{*\mathrm{T}} \\ \vdots \\ (\mathbf{D}^{(\acute{L})})^{*\mathrm{T}} \end{pmatrix}}_{\mathbf{D}^{*\mathrm{T}}}. \tag{7.21}$$

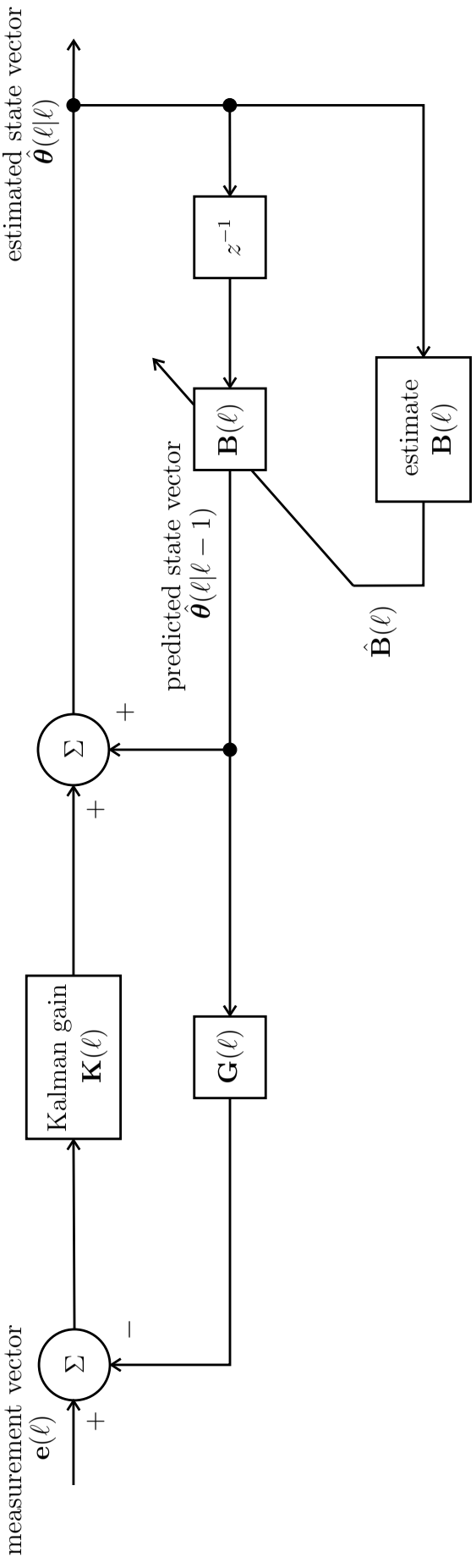



Figure 7.1. A block diagram of the adaptive Kalman filter for MIMO channel tracking in TDD systems

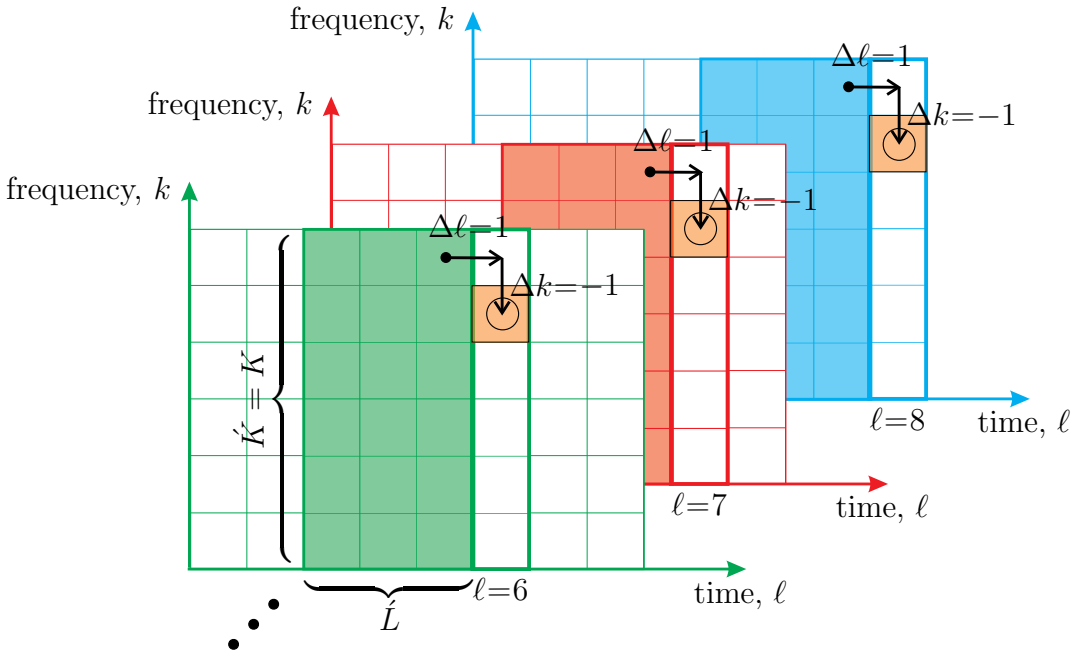



Figure 7.2. Three linear equations as a function of $\varphi_{\Delta k=-1,\Delta\ell=1}$ can be obtained from three overlapping sets of estimated channel coefficients from time instants $\ell = 6$, $\ell = 7$ and $\ell = 8$

The filter coefficient matrix $\mathbf{D}^{*\mathrm{T}}$ can be determined from the overdetermined system of linear equations of (7.21) using the method of least squares when $M \cdot N \cdot (\ell - \acute{L}) > \acute{K} \cdot \acute{L}$. However, both the matrices $\mathbf{X}(\ell)$ and $\mathbf{Y}(\ell)$ have errors owing to the estimation error in the estimated channel coefficient vector $\hat{\mathbf{h}}_{m,n}(\ell)$. Consequently, the total least squares estimator, which accounts for errors on both sides of (7.21), is a more appropriate estimator of the filter coefficient matrix $\mathbf{D}^{*\mathrm{T}}$ than the least squares estimator. In the total least squares estimation, the estimation problem is given as

$$\begin{aligned} \{\hat{\mathbf{D}}, \Delta\hat{\mathbf{X}}(\ell), \Delta\hat{\mathbf{Y}}(\ell)\} = \underset{\mathbf{D},\Delta\mathbf{X}(\ell),\Delta\mathbf{Y}(\ell)}{\arg\min} \quad & \|(\Delta\mathbf{X}(\ell) \,|\, \Delta\mathbf{Y}(\ell))\|_{\mathrm{F}} \\ \text{subject to} \quad & (\mathbf{X}(\ell) + \Delta\mathbf{X}(\ell)) \cdot \mathbf{D}^{*\mathrm{T}} = \mathbf{Y}(\ell) + \Delta\mathbf{Y}(\ell). \end{aligned} \tag{7.22}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix. Thus the total least squares estimator makes as little as possible corrections $\Delta\mathbf{X}(\ell)$ and $\Delta\mathbf{Y}(\ell)$ in the Frobenius norm sense to both matrices $\mathbf{X}(\ell)$ and $\mathbf{Y}(\ell)$, respectively. The solution to the total least squares problem can be determined using the singular value decomposition (SVD). The SVD of the augmented matrix $(\mathbf{X}(\ell) \,|\, \mathbf{Y}(\ell))$, where the matrix $\mathbf{X}(\ell)$ is of dimension $M \cdot N \cdot (\ell - \acute{L}) \times \acute{K} \cdot \acute{L}$ and the matrix $\mathbf{Y}(\ell)$ is of dimension $M \cdot N \cdot (\ell - \acute{L}) \times \acute{K}$, is a decomposition

$$(\mathbf{X}(\ell) \,|\, \mathbf{Y}(\ell)) = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^{*\mathrm{T}}, \tag{7.23}$$

where the $M \cdot N \cdot (\ell - \acute{L}) \times (\acute{K} \cdot \acute{L} + \acute{K})$ matrix $\mathbf{U}$ and the $(\acute{K} \cdot \acute{L} + \acute{K}) \times (\acute{K} \cdot \acute{L} + \acute{K})$ matrix $\mathbf{V}$ have orthonormal columns and $\mathbf{\Sigma} = \mathrm{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_{\acute{K}\cdot\acute{L}+\acute{K}}\right)$ with $\sigma_1 \geq$

$\cdots \geq \sigma_{\acute{K}\cdot\acute{L}+\acute{K}} \geq 0$. As described in Section 3.3, the total least squares solution finds the matrix $(\Delta\mathbf{X}(\ell)|\Delta\mathbf{Y}(\ell))$ with the minimum norm that makes the matrix $(\mathbf{X}(\ell)+\Delta\mathbf{X}(\ell)|\mathbf{Y}(\ell)+\Delta\mathbf{Y}(\ell))$, owing to the constraint in (7.22), rank deficient.

Each channel coefficient can be represented as a linear combination of $P$ other channel coefficients. Thus had it not been for the estimation error in the estimated channel coefficients vector $\hat{\mathbf{h}}_{m,n}(\ell)$, the rank of the matrix $\mathbf{X}(\ell)$ would have been $P$. Thus the SVD of the augmented matrix $(\mathbf{X}(\ell)\,|\,\mathbf{Y}(\ell))$ has $\acute{K}\cdot\acute{L}+\acute{K}$ singular values out of which $P$ are a superposition of the singular values corresponding to the propagation paths and the estimation error in $\hat{\mathbf{h}}_{m,n}(\ell)$, whereas the remaining $\acute{K}\cdot\acute{L}+\acute{K}-P$ singular values stem from the estimation error in $\hat{\mathbf{h}}_{m,n}(\ell)$, i.e., $\sigma_1 \geq \cdots \geq \sigma_P > \sigma_{P+1} \geq \cdots \geq \sigma_{\acute{K}\cdot\acute{L}+\acute{K}}$. Consequently, the largest $P$ singular values are of interest in the total least squares solution. One approach to such kinds of problems is the truncated total least squares [FGHO97] where only the largest $P$ singular values corresponding to the $P$ propagation paths are considered, whereas the remaining singular values are treated as zeros. Towards this end, the matrix $\mathbf{V}$ is partitioned, as shown in Fig. 7.3, as

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix} \tag{7.24}$$

where $\mathbf{V}_{11} \in \mathbb{C}^{\acute{K}\cdot\acute{L}\times P}$, $\mathbf{V}_{12} \in \mathbb{C}^{\acute{K}\cdot\acute{L}\times(\acute{K}\cdot\acute{L}+\acute{K}-P)}$, $\mathbf{V}_{21} \in \mathbb{C}^{\acute{K}\times P}$ and $\mathbf{V}_{22} \in \mathbb{C}^{\acute{K}\times(\acute{K}\cdot\acute{L}+\acute{K}-P)}$. If $\mathbf{V}_{22}$ has full rank, then the total least squares solution $\hat{\mathbf{D}}^{*\mathrm{T}}$, which has the minimum Frobenius norm $\|\hat{\mathbf{D}}^{*\mathrm{T}}\|_{\mathrm{F}}$, is given as

$$\hat{\mathbf{D}}^{*\mathrm{T}} = -\mathbf{V}_{12}\cdot\mathbf{V}_{22}^{+}, \tag{7.25}$$

where the superscript + denotes the Moore-Penrose pseudoinverse [VHV91]. If the $P^{\mathrm{th}}$ singular value of the matrix $\mathbf{X}(\ell)$ is greater than the $(P+1)^{\mathrm{th}}$ singular value of the augmented matrix $(\mathbf{X}(\ell)\,|\,\mathbf{Y}(\ell))$, i.e., $\sigma_P(\mathbf{X}(\ell)) > \sigma_{P+1}((\mathbf{X}(\ell)\,|\,\mathbf{Y}(\ell)))$, then $\mathbf{V}_{22}$ has full rank and the total least squares solution exists.

From the preceding discussions, it can be concluded that the adaptive Kalman filter is guaranteed to converge if the singular values corresponding to the propagation paths are not negligible compared with the singular values of the estimation error in the estimated channel coefficients vector $\hat{\mathbf{h}}_{m,n}(\ell)$.

In the discussions so far, the total least squares solution has been derived under the assumption of having an overdetermined system of linear equations, i.e., $M\cdot N\cdot(\ell-\acute{L}) > \acute{K}\cdot\acute{L}$. However, the total least squares can also be applied to underdetermined system of linear equations, i.e., $M\cdot N\cdot(\ell-\acute{L}) < \acute{K}\cdot\acute{L}$, as long as $M\cdot N\cdot(\ell-\acute{L}) > P$. In this case, the total least squares yields a solution, out of possible infinite number

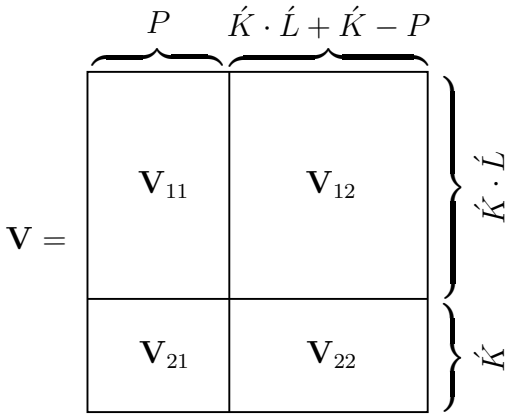



Figure 7.3. Pictorial representation of partitioning of the matrix $\mathbf{V}$

of solutions, with minimum norm [MS00]. Thus it is possible to start update of the adaptive Kalman filter parameters at the discrete time instant $\ell = \acute{L} + 1$. This is beneficial in the case where $\acute{K} \cdot \acute{L}$ is large and the adaptive Kalman filter has to wait a considerable time, i.e., $M \cdot N \cdot (\ell - \acute{L}) > \acute{K} \cdot \acute{L}$, before the onset of the periodic update of the adaptive Kalman filter parameters. Nevertheless, the onset of update of the adaptive Kalman filter parameters should be made when the there is a sufficient information that could be obtained from the estimated channel coefficients.

The reliability of the estimates $\hat{\mathbf{D}}$ of the filter coefficient matrix $\mathbf{D}$ improves periodically as more estimated channel coefficients are used to estimate the filter coefficient matrix $\mathbf{D}$. Thus the adaptive Kalman filter has to consider the error in the estimates of the filter coefficient matrix $\mathbf{D}$ and hence the state transition matrix $\mathbf{B}$. Assuming that the estimates $\hat{\mathbf{D}}$ of the filter coefficient matrix $\mathbf{D}$ as the superposition of the true filter coefficient matrix $\mathbf{D}$ and an estimation error modelled as Gaussian noise, the state transition noise can also be used to model the error in the estimates of the filter coefficient matrix $\mathbf{D}$. The covariance matrix of the state transition noise can be estimated using the Yule-Walker equations as shown in Appendix C. Initial estimates of the autocorrelation function of the channel can be obtained from the Jakes' model and the truncated one-sided exponential PDP, while at the later stages estimates of the autocorrelation function of the channel can be obtained from the estimated channel coefficients using time averaging. Thus at the early recursions of the adaptive Kalman filter the covariance matrix of the state transition noise is "large" to reflect the less reliability of the early estimates of the state transition matrix. This enables the adaptive Kalman filter not to over-rely on the early estimates of the state transition matrix. However, as the tracking progresses, the estimates of the covariance matrix of the state transition noise get "smaller" to reflect the increasing confidence on the estimates of the state transition matrix.

## 7.3 Performance analysis

In this section, the performance of the proposed MIMO channel prediction algorithm for TDD systems is analyzed considering a microcellular mobile radio system using Monte Carlo simulations with $10^4$ independent trials. It is assumed that the BS is located at the origin, whereas the MS is located at (1000, 0) m. Furthermore, it is assumed that the BS and the MS employ uniform linear antenna arrays. The minimum distance between the BS antennas is 0.03 m, whereas the minimum distance between the MS antennas is 0.02 m. It is assumed that the center frequency is $f_0 = 2.4$ GHz and hence the wavelength is $\lambda = 0.125$ m. Thus the minimum distance between the BS antennas and the MS antennas is less than half the wavelength. The MS is assumed to move with a speed of 30 km/hr during the tracking time and the direction of motion does not change. The positions of the scatterers are generated randomly using the single bounce elliptical model discussed in Section 3.4.2 for a normalized maximum path delay $\frac{\tau_{\max}}{\tau_{\mathrm{LOS}}} = 2.5$. In the single bounce elliptical model, the scatterers are assumed to lie in a horizontal plane [LR96]. Thus the elevation angles-of-departure and angles-of-arrivals are 90° and only the azimuth angles-of-departure and angles-of-arrivals are considered to characterize the directions of propagations. The propagation path parameters, i.e., the delays, the Doppler shifts, the angles-of-departure and the angles-of-arrival are calculated from the geometry of the MS, BS and scatterers positions using (2.9).

It is assumed that both the direct path and the multipath components experience $d^\beta$ path loss, where $\beta$ is the path loss exponent. The received power $P_{\mathrm{Rx,p}}$ for the $p^{\mathrm{th}}$ multipath component relative to the power $P_{\mathrm{Rx,0}}$ of the direct path component is given as

$$P_{\mathrm{Rx,p}} - P_{\mathrm{Rx,0}} = -10\beta \log_{10}\left(\frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{LOS}}}\right) - L_{\mathrm{r}}, \tag{7.26}$$

where $L_{\mathrm{r}}$ is the loss owing to the reflection from a scatterer [LR96]. If the complex amplitudes of the multipath components are normalized such that $|\alpha_0| = 1$, then the complex amplitudes of the $p^{\mathrm{th}}$ multipath component can be determined as

$$\alpha_p = 10^{\frac{P_{\mathrm{Rx,p}} - P_{\mathrm{Rx,0}}}{20}} \mathrm{e}^{\mathrm{j}\vartheta}, \tag{7.27}$$

where $\vartheta$ is the random phase of the $p^{\mathrm{th}}$ path which is assumed to be drawn from a uniform distribution in $[0,\ 2\pi)$. In the simulations, a path loss exponent of $\beta = 4$ and a reflection loss of $L_{\mathrm{r}} = 6$ dB are assumed.

It is assumed that the subcarrier spacing is $F = 2$ kHz and the transmit symbol duration is $T = \frac{1}{F} = 500$ $\mu$sec. The channel coefficients $H_{m,n}(k, \ell)$ are calculated using (6.1) based on the randomly generated parameters of the propagation paths. The generated

channel coefficients are normalized to one, i.e., $\mathbb{E}\{|H_{m,n}(k,\ell)|^2\} = 1$. $K = 32$ pilot subcarriers are considered in 50 time instants. The time and frequency spacing of the pilot symbols is $15T$ and $30F$, respectively. Thus the pilot symbols sample the channel transfer function at the Nyquist rate in time and frequency, i.e., $15T \cdot \nu_{\max} \approx \frac{1}{2}$ and $30F \cdot \tau_{\max} \approx \frac{1}{2}$. In the simulations, it is assumed that the propagation path parameters are unchanged during the tracking time. Thus the state transition noise $\mathbf{u}(\ell)$ is zero. However, for the adaptive Kalman filter a non-zero state transition noise, computed using the Yule-Walker equations, is considered in order to account for the imperfect estimates of the state transition matrix.

The performance metric is the normalized mean square error (NMSE) for the $\ell = 50^{\text{th}}$ time instant between the true $H_{m,n}(k, \ell = 50)$ and the predicted $\check{H}_{m,n}(k, \ell = 50)$ channel coefficients:

$$\text{nmse}(\check{H}(\ell = 50)) = \frac{\mathbb{E}_{\mathbf{e}(\ell),H_{m,n}(k,\ell)}\left\{\left|\check{H}_{m,n}(k,\ell = 50) - H_{m,n}(k,\ell = 50)\right|^2\right\}}{\mathbb{E}\left\{\left|H_{m,n}(k,\ell = 50)\right|^2\right\}}. \tag{7.28}$$

The operator $\mathbb{E}_{\mathbf{e}(\ell),H_{m,n}(k,\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\text{p}\left(\mathbf{e}(\ell), H_{m,n}(k,\ell)\right)$. In the simulations, the expectations are computed using time averages. The prediction performance of the adaptive Kalman filter is analyzed for different pseudo signal-to-noise-ratios (PSNRs)

$$\text{PSNR} = \mathbb{E}\left\{|s(k,\ell)|^2\right\}/\sigma_n^2, \tag{7.29}$$

where $s(k,\ell)$ is the transmitted pilot symbol and $\sigma_n$ is the standard deviation of the measurement noise.

A reference Kalman filter where the Kalman filter parameters are calculated from the true path parameters for each run of the Monte Carlo simulations is used as a performance benchmark. The prediction performance of the Kalman filter based solely on the Jakes' model and the truncated one-sided exponential PDP is rather poor regardless of the considered dimensions of the MIMO channels or the numbers of propagation paths. Consequently, the performance of the Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP is not considered. Furthermore, for the same reason, the performance of independent prediction of the SISO subchannels of the MIMO channel is also not considered.

Fig. 7.4 shows the prediction performances of the reference and the adaptive Kalman filters. $2\times2$ and $4\times4$ MIMO channels are considered for $P = 5$ and $P = 10$ propagation paths. In the simulations, the linear relation between all the pilot symbol subcarriers for two time instants is exploited, i.e., $\acute{K} = K$ and $\acute{L} = 2$. Thus the filter dimension

$\acute{K} \cdot \acute{L} = 64$ is well beyond the minimum required filter dimension, i.e., $P = 5$ and $P = 10$. The periodical update of the parameters of the adaptive Kalman filter is initiated at the time instant $\ell = 2\acute{L} + 1$. It can be seen from Fig. 7.4 that the performance of the adaptive Kalman filter is bounded by the reference Kalman filter. It can also be seen that the performance of the reference Kalman filter depends only the number of propagations paths but not on the dimension of the MIMO channel. This occurs owing to the narrowband assumption for array signal processing which results in prediction filter coefficients which are independent of the transmitter and receiver antennas under consideration as discussed in Section 6.1. It can also be seen from Fig. 7.4 that the performance of the adaptive Kalman filter improves as the dimension of the MIMO channel is increased. This arises owing to the exploitation of the spatial property of the MIMO channel in predicting the channel coefficients.

Fig. 7.4 also shows the impact of the number of the propagation paths on the performances of the reference and the adaptive Kalman filters. It can be seen that, for a given MIMO channel, as the number of the propagation paths increases, the performance of both the reference and the adaptive Kalman filter deteriorates. This result stems from the fact that for the same filter dimensions, i.e., $\acute{K} \cdot \acute{L} = 64$, more measurements than the minimum required number of measurements are exploited for the case where the number of propagation paths is smaller, i.e., $P = 5$, than for the case where the number of propagation paths is higher, i.e., $P = 10$. This results in an improved performance from noise suppression.

Fig. 7.5 and Fig. 7.6 show the convergence behaviours of the reference and the adaptive Kalman filters for different numbers of propagation paths and dimensions of the MIMO channels when PSNR = 25 dB. It can be seen that the adaptive Kalman filter converges quickly after $\ell > 2\acute{L}$ iterations. In general, the adaptive Kalman filter starts to converge quickly after the onset of update of the adaptive Kalman filter parameters. Furthermore, it has been observed from the simulation results that the rate of convergence increases considerably as the PSNR increases.

The Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP would show a significant improvement in performance if the transmitted pilot symbols were to sample the channel transfer function at several multiples of the Nyquist rate in time and frequency. Fig. 7.7 shows the NMSE prediction performances of the Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP and the adaptive Kalman filter when the pilot symbols sample the channel transfer function at eight times the Nyquist rate in time and frequency. In the previous simulation results, the performance of the Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP was rather too poor to be considered. However,

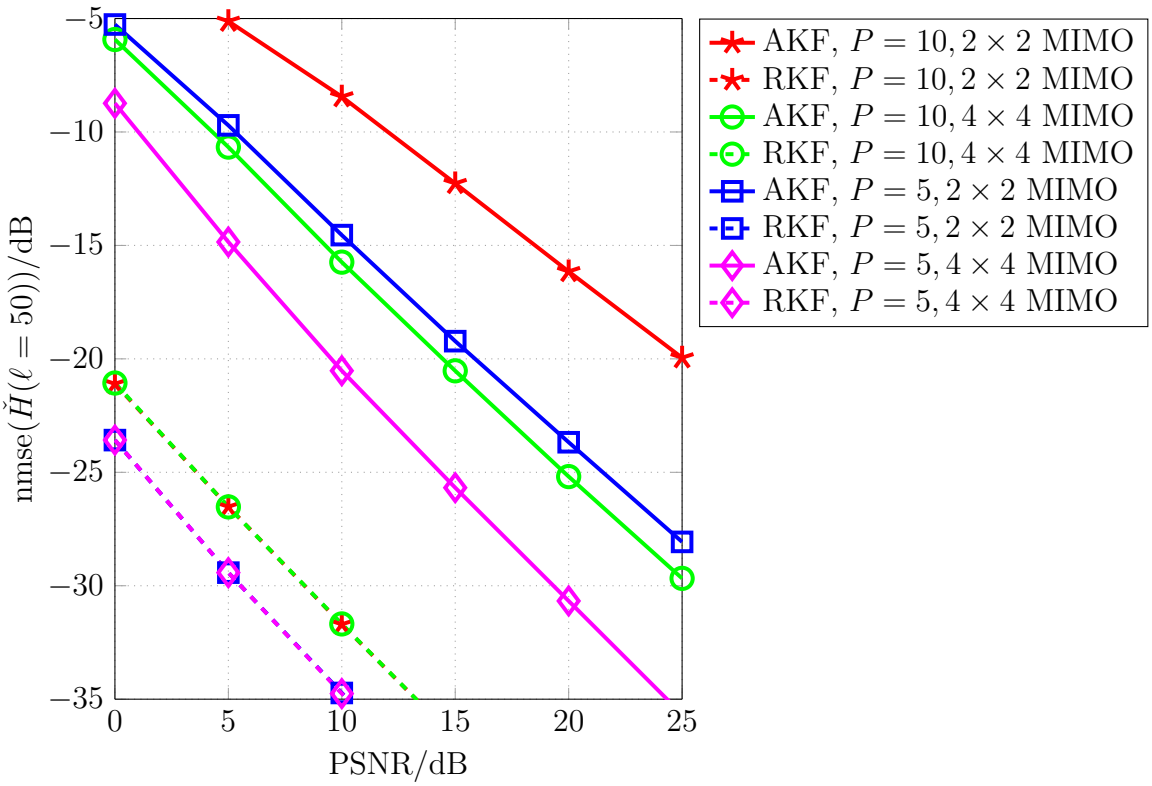
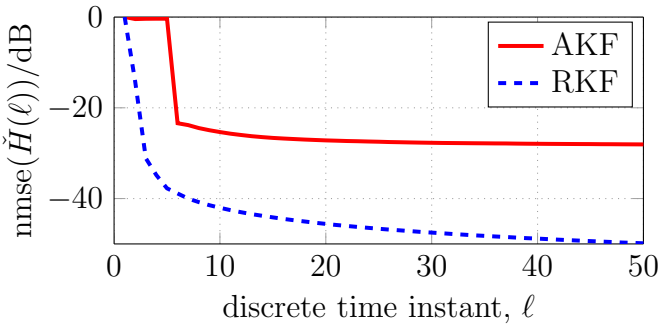


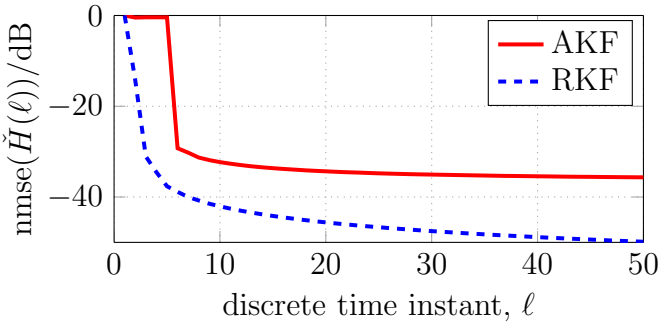

Figure 7.4. The NMSE prediction performances of the reference Kalman filter (RKF) and the adaptive Kalman filter (AKF)

for the over-sampled channel case shown in Fig. 7.7, it can be seen that the Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP yields an appreciable prediction performance. Consequently, in practice, channel prediction and estimation algorithms based on stochastic models commonly consider pilot symbols which sample the channel transfer function at several multiples of the Nyquist rate so that the channel coefficients have strong correlations, thereby making the prediction problem relatively easy [CZ04, SM05, MCCK07].

A note in passing: It has been assumed that training signal based channel estimation techniques are employed. The received pilot symbols have been exploited to estimate the channel at the pilot symbol locations. The channels at other locations are to be estimated using interpolation techniques. Towards this end, the proposed adaptive Kalman filter can also be used as an interpolation filter in the frequency domain. In the adaptive Kalman filter implementation, the measurements for the channel coefficients to be interpolated can be assumed to be unavailable. Thus the interpolation is to be performed by the cumulative effect of extracting the information from the measurement data about the channel coefficients to be interpolated and recursively incorporating this information with the next measurement data to predict the channel coefficients to be



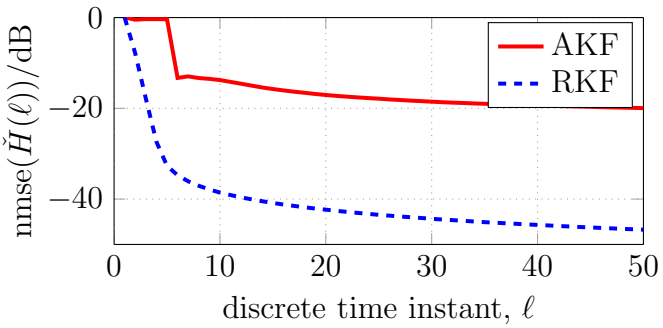

(a) The convergence behaviours for a $2 \times 2$ MIMO channel



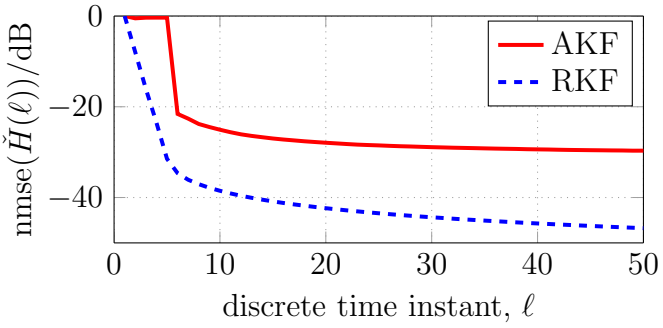

(b) The convergence behaviours for a $4 \times 4$ MIMO channel

Figure 7.5. The convergence behaviours of the reference Kalman filter (RKF) and the adaptive Kalman filter (AKF), $P = 5$ and PSNR $= 25$ dB

interpolated. It has been shown in [SW15a] that satisfactory interpolation performance can be obtained by the adaptive Kalman filter. Needless to say, the interpolated channel coefficients have a higher NMSE than the channel coefficients whose measurement data is available.




(a) The convergence behaviours for a $2 \times 2$ MIMO channel




(b) The convergence behaviours for a $4 \times 4$ MIMO channel

Figure 7.6. The convergence behaviours of the reference Kalman filter (RKF) and the adaptive Kalman filter (AKF), $P = 10$ and PSNR $= 25$ dB

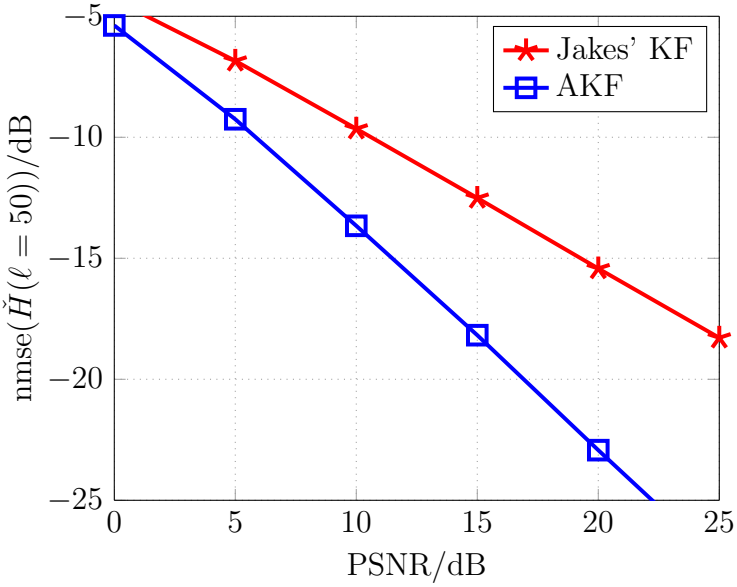



Figure 7.7. The NMSE prediction performances of the Kalman filter based on the Jakes' model and the truncated one-sided exponential PDP (Jakes' KF) and the adaptive Kalman filter (AKF) for an eight times over-sampled $2 \times 2$ MIMO channel, $P = 5$

# Chapter 8

# Exploiting side information in MIMO channel prediction

## 8.1 Exploiting feedback in frequency-division-duplex systems

In the previous chapter, it has been shown that significant improvements in channel prediction performance can be obtained by tracking a MIMO channel. As an example, tracking the uplink (UL) and the downlink (DL) channels in TDD systems has been considered. In this chapter, the possible performance improvements brought about by exploiting side information in tracking a MIMO channel are considered. The double directional channel model has been derived by considering the physical parameters of the radio propagation paths. Hence information related to the propagation paths can be exploited to improve channel prediction performance. For example, nowadays it is not uncommon for a MS to be equipped with sensors which can measure the speed, acceleration and direction of motion of the MS. This information can be used to obtain significant performance improvements in tracking and predicting a MIMO channel. The adaptive Kalman filter has been derived under the assumption of a small scale MS mobility which is assumed to result in a negligible change in the radio propagation scenario. Hence the information about the change in the direction of motion of a MS can be used to accommodate the adaptive Kalman filter to the change in the radio propagation scenario.

A rather readily available side information which could be exploited for tracking a MIMO channel is the feedback information. Side information about the channel obtained from the feedback channel is particularly useful for frequency-division-duplex (FDD) systems. In FDD systems, the UL data and the DL data are transmitted in different frequency bands. The UL and the DL channels are identical only if the frequency duplexing distance is much smaller than the coherence bandwidth of the channel [Mol05]. However, in practice, there is a relatively large frequency duplexing distance between the UL and the DL bands to enable easy separation between transmission and reception using affordable filters. Thus, in general, the UL and the DL channels are not identical. Thus using training signal based channel estimation techniques, the BS and the MS would only have an estimate of the UL and the DL channel, respectively, i.e., only receiver side channel information is available. However, at the BS and the MS

the required transmitter side channel state information (CSI) is that of the DL and the UL channel, respectively. Towards this end, the BS and the MS transmit their receiver side CSI via the feedback channel to be used as a transmitter side CSI. However, this approach yields outdated transmitter side CSI.

Thus the outdated transmitter side CSI obtained via the feedback channel shall be exploited to obtain an up-to-date transmitter side CSI. In the following, two approaches to exploiting the outdated transmitter side CSI are considered. In the first approach, only the outdated transmitter side CSI is exploited to obtain an up-to-date transmitter side CSI. In the second approach, the outdated transmitter side CSI is jointly exploited with the up-to-date receiver side CSI to obtain an up-to-date transmitter side CSI. The second approach constitutes an UL-DL channel transformation at the BS and a DL-UL channel transformation at the MS. In the following, the UL-DL transformation of the channel transfer function (CTF) at the BS is considered. The DL-UL transformation of the channel transfer function at the MS is a similar problem. Furthermore, the first approach is also considered to obtain an up-to-date transmitter side DL CSI.

Fig. 8.1 shows the UL and DL channel estimation scenario at the BS. It is assumed that the number of UL and DL subcarriers are equal, i.e., $K_{\mathrm{UL}} = K_{\mathrm{DL}}$. A gap band between the UL and DL channels is not considered. Thus the UL and the DL channel coefficients are indexed with $1 \leq k_{\mathrm{UL}} \leq K_{\mathrm{UL}}$ and $K_{\mathrm{UL}} + 1 \leq k_{\mathrm{DL}} \leq K$, respectively. However, as mentioned earlier, in practice, there is a relatively large gap band between the UL and the DL bands. For the first approach of obtaining up-to-date transmitter side CSI using only the outdated transmitter side CSI, the presence of a gap band has no influence on estimating the up-to-date transmitter side CSI. However, for the second approach, i.e., UL-DL channel transformation, the presence of a gap band makes the UL-DL channel transformation difficult as the prediction has to span a large bandwidth. In the discussions so far, it has been assumed that the propagation path parameters remain unchanged during the frequency band under consideration. Thus the impact of the gap band on the UL-DL channel transformation is not considered. In [SW14d], UL-DL channel transformation algorithms have been proposed based on the geometric channel model under consideration of the gap band between the UL and DL channels.

It is assumed that the BS uses training signal based channel estimation to obtain up-to-date estimates of the UL channel. Furthermore, it is assumed that the BS has outdated estimates of the DL channel obtained from the MS via the feedback channel. In practice, the feedback channel is limited and hence a quantized CSI information is transmitted rather than the estimated DL channel [LHSH04]. Nevertheless, it is assumed here that the estimated DL channel at the MS is available at the BS after a delay of one time unit without loss of information. Thus at the BS the estimate of the

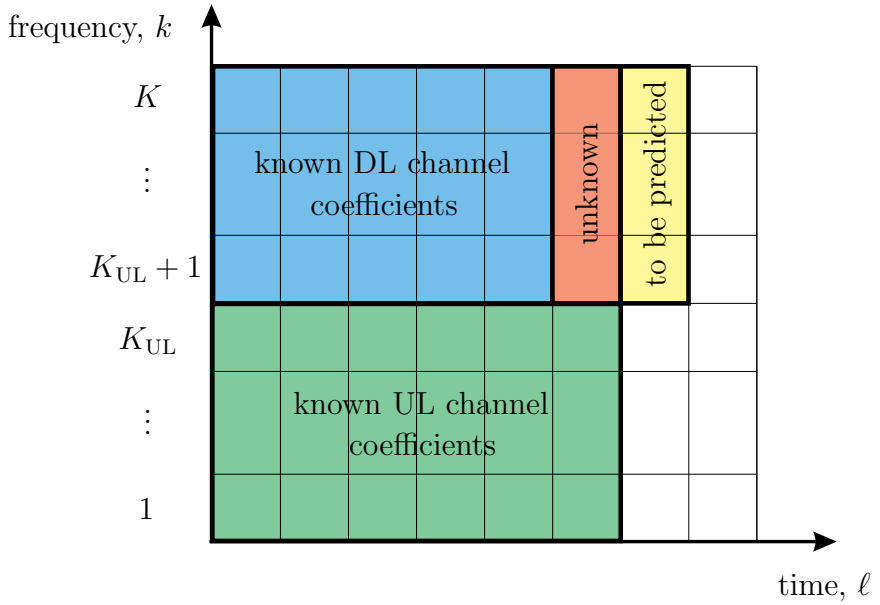



Figure 8.1. The UL and DL channel estimation scenario at the BS

current DL channel is unknown. Furthermore, the estimate of the DL channel at the next time instant is required so that the BS has an up-to-date transmitter side CSI. In the first approach, the BS shall exploit the delayed estimates of the DL channel to predict the DL channel at the next time instant, i.e., a prediction of two time units. On the other hand, in the second approach, the BS shall exploit the current estimates of the UL channel and the delayed estimates of the DL channel to predict the DL channel at the next time instant, i.e., a prediction of one time unit. The first approach is rather similar to the MIMO channel tracking discussed in the previous chapter, whereas the second approach requires further treatment. In the following, the second approach is discussed in detail with appropriate digression to the first approach when necessary.

Owing to its convenience and intuitive description of dynamical systems, the MIMO DL channel tracking problem in FDD systems is described using a state-space model. Since training signal based channel estimation is considered, the channel coefficients that are considered are the channel coefficients of the pilot symbol subcarriers. Consequently, for the sake of simplicity, the subcarrier index $k = 1, \ldots, K$ is re-used to index the pilot symbol subcarriers instead of all the subcarriers. Furthermore, the time index $\ell = 1, \ldots, L$ is re-used to index the sequence of the pilot symbols in time. Since the DL channel coefficients are to be tracked, the state vector shall contain the DL channel coefficients. It is beneficial to exploit the linear relation between all the DL channel coefficients of the pilot symbol subcarriers. To this end, all the DL channel coefficients at a discrete time instant $\ell$ between the $m^{\mathrm{th}}$ receive antenna and the $n^{\mathrm{th}}$ transmit

antenna shall be considered, i.e.,

$$\mathbf{h}_{m,n}^{(\mathrm{DL})}(\ell) = (H_{m,n}(K,\ell), H_{m,n}(K-1,\ell), \ldots, H_{m,n}(K_{\mathrm{UL}}+1,\ell))^{\mathrm{T}}. \tag{8.1}$$

Stacking such vectors from all $M \cdot N$ transmit and receive antennas results in a vector

$$\mathbf{h}_{\mathrm{DL}}(\ell) = ((\mathbf{h}_{1,1}^{(\mathrm{DL})}(\ell))^{\mathrm{T}}, (\mathbf{h}_{1,2}^{(\mathrm{DL})}(\ell))^{\mathrm{T}}, \ldots, (\mathbf{h}_{M,N}^{(\mathrm{DL})}(\ell))^{\mathrm{T}})^{\mathrm{T}}. \tag{8.2}$$

Considering the vector $\mathbf{h}_{\mathrm{DL}}(\ell)$ for $\acute{L}$ time instants means that the evolution of each DL channel coefficient, for each of the $M \cdot N$ transmit and receive antennas, is represented by $\acute{K} \cdot \acute{L}$ other DL channel coefficients where $\acute{K} = K_{\mathrm{DL}}$, i.e., the number of DL pilot symbol subcarriers. Each DL channel coefficient can be represented as a linear combination of $\acute{K} \cdot \acute{L}$ other DL channel coefficients as long as $\acute{K} \cdot \acute{L} \geq P$, where $P$ is the number of propagation paths. Thus considering the vector $\mathbf{h}_{\mathrm{DL}}(\ell)$ for $\acute{L}$ time instants results in the DL state vector

$$\boldsymbol{\theta}_{\mathrm{DL}}(\ell) = (\mathbf{h}_{\mathrm{DL}}^{\mathrm{T}}(\ell), \mathbf{h}_{\mathrm{DL}}^{\mathrm{T}}(\ell-1), \ldots, \mathbf{h}_{\mathrm{DL}}^{\mathrm{T}}(\ell-\acute{L}+1))^{\mathrm{T}}. \tag{8.3}$$

The state transition equation describes the evolution of each of the DL channel coefficients with time. It has been shown that each channel coefficient can be represented as a linear combination of a sufficient number of other channel coefficients. Consequently, the state transition equation is linear. The state transition equation can be written as

$$\boldsymbol{\theta}_{\mathrm{DL}}(\ell) = \mathbf{B} \cdot \boldsymbol{\theta}_{\mathrm{DL}}(\ell-1) + \boldsymbol{\Gamma} \cdot \mathbf{u}_{\mathrm{DL}}(\ell), \tag{8.4}$$

where the $M \cdot N \cdot \acute{K} \times 1$ vector $\mathbf{u}_{\mathrm{DL}}(\ell)$ is assumed to be zero mean complex multivariate white Gaussian distributed with covariance matrix $\mathbf{R}_{\mathrm{u_{DL}u_{DL}}}$. The $M \cdot N \cdot \acute{K} \cdot \acute{L} \times M \cdot N \cdot \acute{K} \cdot \acute{L}$ state transition matrix $\mathbf{B}$ and the $M \cdot N \cdot \acute{K} \cdot \acute{L} \times M \cdot N \cdot \acute{K}$ matrix $\boldsymbol{\Gamma}$ can be derived in a similar manner as in (7.5) and (7.6), respectively.

The measurement equation of the state-space model can be developed by considering the signal model. The UL-DL channel transformation exploits the current UL channel measurements and the one time unit delayed DL channel measurements, which are fedback from the MS to the BS, to predict the DL channel coefficients. Thus the measurement vector is defined as $\mathbf{e}(\ell) = (\mathbf{e}_{\mathrm{UL}}^{\mathrm{T}}(\ell), \mathbf{e}_{\mathrm{DL}}^{\mathrm{T}}(\ell-1))^{\mathrm{T}}$, where the vectors $\mathbf{e}_{\mathrm{UL}}(\ell)$ and $\mathbf{e}_{\mathrm{DL}}(\ell-1)$ are the current UL and the delayed DL measurements, respectively. Let the matrix $\mathbf{S}_{\mathrm{UL}}(\ell)$ and $\mathbf{S}_{\mathrm{DL}}(\ell)$ denote diagonal matrices of size $M \cdot N \cdot \acute{K}$ with elements on the diagonal being the transmitted UL and DL pilot symbols, respectively. The $2M \cdot N \cdot \acute{K} \times 1$ received pilot symbol vector $\mathbf{e}(\ell)$ is then given as

$$\mathbf{e}(\ell) = \mathbf{G}(\ell) \cdot \boldsymbol{\theta}_{\mathrm{DL}}(\ell) + \mathbf{n}(\ell), \tag{8.5}$$

where

$$\mathbf{G}(\ell) = \left( \begin{array}{c|c|c} \multicolumn{3}{c}{\mathbf{S}_{\mathrm{UL}}(\ell) \cdot \mathbf{T}} \\ \hline \mathbf{0}_{M \cdot N \cdot \acute{K}} & \mathbf{S}_{\mathrm{DL}}(\ell - 1) & \mathbf{0}_{M \cdot N \cdot \acute{K} \times M \cdot N \cdot \acute{K} \cdot (\acute{L}-2)} \end{array} \right) \tag{8.6}$$

and $\mathbf{n}(\ell)$ is the measurement noise which is a multivariate zero mean complex white Gaussian noise with covariance matrix $\mathbf{R}_{\mathrm{nn}}$. The $M \cdot N \cdot \acute{K} \times M \cdot N \cdot \acute{K} \cdot \acute{L}$ matrix $\mathbf{T}$ is the DL to UL channel transformation matrix which transforms the state vector $\boldsymbol{\theta}_{\mathrm{DL}}(\ell)$ containing the DL channel coefficients into the UL channel coefficient vector $\mathbf{h}_{\mathrm{UL}}(\ell)$ so that the current UL measurements can be exploited. The matrix $\mathbf{T}$ is given as

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(1)} & \cdots & \mathbf{0}_{\acute{K}} & \mathbf{T}^{(2)} & \cdots & \mathbf{0}_{\acute{K}} & \cdots & \mathbf{T}^{(\acute{L})} & \cdots & \mathbf{0}_{\acute{K}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ \mathbf{0}_{\acute{K}} & \cdots & \mathbf{T}^{(1)} & \mathbf{0}_{\acute{K}} & \cdots & \mathbf{T}^{(2)} & \cdots & \mathbf{0}_{\acute{K}} & \cdots & \mathbf{T}^{(\acute{L})} \end{pmatrix}. \tag{8.7}$$

The $\acute{K} \times \acute{K} \cdot \acute{L}$ coefficient matrix

$$\tilde{\mathbf{T}} = \begin{pmatrix} \mathbf{T}^{(1)} & \cdots & \mathbf{T}^{(\acute{L})} \end{pmatrix} \tag{8.8}$$

weights the DL channel coefficient vector $\left( (\mathbf{h}_{m,n}^{(\mathrm{DL})}(\ell))^{\mathrm{T}}, \ldots, (\mathbf{h}_{m,n}^{(\mathrm{DL})}(\ell - \acute{L} + 1))^{\mathrm{T}} \right)^{\mathrm{T}}$ constituting the DL state vector $\boldsymbol{\theta}_{\mathrm{DL}}(\ell)$ to yield the UL channel coefficient vector $\mathbf{h}_{m,n}^{(\mathrm{UL})}(\ell)$. The coefficient matrix $\tilde{\mathbf{T}}$ can be calculated, given the delays $\tau_p$ and Doppler shifts $\nu_p$, using (6.27) as

$$\tilde{\mathbf{T}} = \begin{pmatrix} \boldsymbol{\varphi}_{\Delta k = K_{\mathrm{UL}} - K, \Delta \ell = 0} & \cdots & \boldsymbol{\varphi}_{\Delta k = 1 - K, \Delta \ell = 0} \end{pmatrix}^{\mathrm{T}}. \tag{8.9}$$

The $\acute{K} \cdot \acute{L} \times 1$ filter coefficient vector $\boldsymbol{\varphi}_{\Delta k = k_{\mathrm{UL}} - K, \Delta \ell = 0}$ weights the DL channel coefficient vector $\left( (\mathbf{h}_{m,n}^{(\mathrm{DL})}(\ell))^{\mathrm{T}}, \ldots, (\mathbf{h}_{m,n}^{(\mathrm{DL})}(\ell - \acute{L} + 1))^{\mathrm{T}} \right)^{\mathrm{T}}$ to yield the UL channel coefficient $H_{m,n}(k_{\mathrm{UL}}, \ell)$. The prediction depths $\Delta \ell = 0$ in time and $\Delta k = k_{\mathrm{UL}} - K$ in frequency are chosen as the UL channel coefficient $H_{m,n}(k_{\mathrm{UL}}, \ell)$ is located at a prediction depths of $\Delta \ell = 0$ in time and $\Delta k = k_{\mathrm{UL}} - K$ in frequency from the reference DL channel coefficient $H_{m,n}(K, \ell)$.

As both the state transition and the measurement equations are linear and Gaussian, the Kalman filter can be used to obtain optimal estimates, in the minimum mean square error (MMSE) sense, of the state vector. The Kalman filter starts with an initial estimate $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(0|0) = \mathbf{0}_{M \cdot N \cdot \acute{K} \cdot \acute{L} \times 1}$ of the initial state vector $\boldsymbol{\theta}_{\mathrm{DL}}(0|0)$ and an initial covariance matrix $\mathbf{M}(0|0) = \mathbf{I}_{M \cdot N \cdot \acute{K} \cdot \acute{L}}$ of the estimation error and recursively computes estimates $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell)$ of the state vector $\boldsymbol{\theta}_{\mathrm{DL}}(\ell)$ and the covariance matrix $\mathbf{M}(\ell|\ell)$ of the estimation error using the following equations

- prediction step:

$$\mathbf{M}(\ell|\ell - 1) = \mathbf{B} \cdot \mathbf{M}(\ell - 1|\ell - 1) \cdot \mathbf{B}^{*\mathrm{T}} + \boldsymbol{\Gamma} \cdot \mathbf{R}_{\mathrm{u_{DL}u_{DL}}} \cdot \boldsymbol{\Gamma}^{*\mathrm{T}} \tag{8.10}$$

$$\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell-1) = \mathbf{B} \cdot \hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell-1|\ell-1) \tag{8.11}$$

- update step:

$$\mathbf{K}(\ell) = \frac{\mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell)}{\mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell) + \mathbf{R}_{\mathrm{nn}}} \tag{8.12}$$

$$\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell) = \hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell-1) + \mathbf{K}(\ell) \cdot \Big(\mathbf{e}(\ell) - \mathbf{G}(\ell) \cdot \hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell-1)\Big) \tag{8.13}$$

$$\mathbf{M}(\ell|\ell) = \mathbf{M}(\ell|\ell-1) - \mathbf{K}(\ell) \cdot \mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1). \tag{8.14}$$

Estimates of the DL channel coefficients at time instant $\ell$ are obtained from $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell)$, whereas the predicted DL channel coefficients at time instant $\ell$ are obtained from $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell|\ell-1)$.

In the previous discussion, the UL-DL channel transformation using the Kalman filter has been presented. The Kalman filter for obtaining the DL channel estimates using only the delayed DL channel measurements can be easily adapted from the UL-DL channel transformation Kalman filter. While the state transition equation is identical, the measurement equation is different in that only the delayed DL channel measurements are considered. Since delayed DL channel measurements are used to track the DL channel, the Kalman filter tracks the DL channel with one time unit delay. Hence in order to obtain an up-to-date estimate of the DL channel, two time units prediction has to be performed. In the prediction step of the Kalman filter, two time units prediction can be performed by multiplying the estimate $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell-1|\ell-1)$ of the state vector $\boldsymbol{\theta}_{\mathrm{DL}}(\ell-1)$ at time instant $\ell-1$ with the square of the state transition matrix $\mathbf{B}$, i.e.,

$$\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell+1|\ell-1) = \mathbf{B}^2 \cdot \hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell-1|\ell-1). \tag{8.15}$$

The two step predicted DL channel coefficients at time instant $\ell-1$ are obtained from $\hat{\boldsymbol{\theta}}_{\mathrm{DL}}(\ell+1|\ell-1)$.

## 8.2 Downlink channel prediction using an adaptive Kalman filter

As discussed in the previous chapter, in order to use the Kalman filter, which yields optimal estimates of the DL channel coefficients, estimates of the delays and the Doppler shifts of each propagation path are required. In this chapter, performing up-to-date estimates of the DL channel at the BS is of interest. In the two-step DL channel prediction approach, the DL channel is to be tracked and predicted using the delayed DL

channel measurements, whereas in the UL-DL channel transformation, the DL channel is to be tracked and predicted using the current UL and the delayed DL channel measurements. Consequently, the adaptive Kalman filter presented in the previous chapter shall be considered to obtain up-to-date estimates of the DL channel at the BS for both approaches. In the adaptive Kalman filter, early estimates of the DL channel coefficients are obtained using estimates of the Kalman filter parameters which are determined using stochastic channel models. The state transition matrix and the covariance matrix of the state transition noise are determined from the Yule-Walker equations based on the Jakes' model and the truncated one-sided exponential PDP. An initial estimate of the DL to UL channel transformation matrix $\mathbf{T}$ can also be obtained using the Jakes' model and the truncated one-sided exponential PDP.

The DL to UL channel transformation matrix $\mathbf{T}$ transforms the state vector $\boldsymbol{\theta}_{\mathrm{DL}}(\ell)$ into the UL channel coefficient vector $\mathbf{h}_{\mathrm{UL}}(\ell)$ so that the current UL measurements can be exploited. Initial estimates of the DL to UL channel transformation matrix $\mathbf{T}$ can be determined using the correlation matrix of the channel obtained from the Jakes' model and the truncated one-sided exponential PDP. The DL to UL channel transformation matrix $\hat{\mathbf{T}}$ which yields the MMSE estimate $\hat{\mathbf{T}} \cdot \boldsymbol{\theta}_{\mathrm{DL}}(\ell)$ of the UL channel coefficient vector $\mathbf{h}_{\mathrm{UL}}(\ell)$ can be determined by solving

$$\hat{\mathbf{T}} = \underset{\mathbf{T}}{\arg\min} \quad \mathbb{E}_{\mathbf{h}_{\mathrm{UL}}(\ell),\boldsymbol{\theta}_{\mathrm{DL}}(\ell)} \left\{ \left\| \mathbf{h}_{\mathrm{UL}}(\ell) - \mathbf{T} \cdot \boldsymbol{\theta}_{\mathrm{DL}}(\ell) \right\|_2 \right\}, \tag{8.16}$$

where the operator $\mathbb{E}_{\mathbf{h}_{\mathrm{UL}}(\ell),\boldsymbol{\theta}_{\mathrm{DL}}(\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\mathbf{h}_{\mathrm{UL}}(\ell), \boldsymbol{\theta}_{\mathrm{DL}}(\ell)\right)$. The above problem can be solved using the Wiener filter which finds a linear MMSE estimator based on the correlation between $\mathbf{h}_{\mathrm{UL}}(\ell)$ and $\boldsymbol{\theta}_{\mathrm{DL}}(\ell)$. The estimate $\hat{\mathbf{T}}$ of the DL to UL channel transformation matrix which yields the MMSE estimate of the UL channel coefficient vector $\mathbf{h}_{\mathrm{UL}}(\ell)$ is given as [Kay93]

$$\begin{aligned} \hat{\mathbf{T}} &= \mathbb{E}_{\mathbf{h}_{\mathrm{UL}}(\ell),\boldsymbol{\theta}_{\mathrm{DL}}(\ell)} \left\{ \mathbf{h}_{\mathrm{UL}}(\ell) \cdot \boldsymbol{\theta}_{\mathrm{DL}}^{*\mathrm{T}}(\ell) \right\} \cdot \left( \mathbb{E}_{\boldsymbol{\theta}_{\mathrm{DL}}(\ell)} \left\{ \boldsymbol{\theta}_{\mathrm{DL}}(\ell) \cdot \boldsymbol{\theta}_{\mathrm{DL}}^{*\mathrm{T}}(\ell) \right\} \right)^{-1} \\ &= \mathbf{R}_{\mathrm{UL,DL}} \cdot \mathbf{R}_{\mathrm{DL,DL}}^{-1}, \end{aligned} \tag{8.17}$$

where the operator $\mathbb{E}_{\boldsymbol{\theta}_{\mathrm{DL}}(\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\boldsymbol{\theta}_{\mathrm{DL}}(\ell)\right)$. Estimates of the covariance matrices $\mathbf{R}_{\mathrm{UL,DL}}(\ell)$ and $\mathbf{R}_{\mathrm{DL,DL}}(\ell)$ of the channel can be calculated using the Jakes' model and the truncated one-sided exponential PDP.

As described in the previous chapter, the stochastic channel models, i.e., the Jakes' model and the truncated one-sided exponential PDP, describe only the statistics of the whole ensemble of channel realizations but not the individual channel realizations. Thus the stochastic channel models do not reflect the statistics of a single channel realization and the estimated autocorrelation function of the channel may not be accurate enough to be used in practice with high reliability. Consequently, the estimated channel

coefficients are then used to periodically update the estimates of the state transition matrix, the covariance matrix of the state transition noise and the DL to UL channel transformation matrix.

The initial estimate of the DL to UL channel transformation matrix $\mathbf{T}$ is periodically updated using the estimated UL channel coefficients $\hat{H}_{m,n}(k_{\mathrm{UL}}, \ell)$ and the DL channel coefficients $\hat{H}_{m,n}(k_{\mathrm{DL}}, \ell)$ by exploiting the linear relation between the UL and the DL channel coefficients. Towards this end, a system of linear equations can be set up from the estimated channel coefficients using (6.7), i.e.,

$$\hat{H}_{m,n}(k_{\mathrm{UL}}, \ell) \approx \sum_{\acute{k}=0}^{\acute{K}-1} \sum_{\acute{\ell}=0}^{\acute{L}-1} \varphi^{(\acute{k},\acute{\ell})}_{\Delta k = k_{\mathrm{UL}} - K, \Delta \ell = 0} \cdot \hat{H}_{m,n}(K - \acute{k}, \ell - \acute{\ell}). \tag{8.18}$$

In order to solve for the filter coefficient vector $\boldsymbol{\varphi}_{\Delta k = k_{\mathrm{UL}} - K, \Delta \ell = 0}$, a system of linear equations should be set up by considering different sets of estimated DL and UL channel coefficients which overlap in time. The filter coefficient vector $\boldsymbol{\varphi}_{\Delta k = k_{\mathrm{UL}} - K, \Delta \ell = 0}$ can be determined from the set up system of linear equations using the method of total least squares. Following similar derivations as shown in the previous chapter, it is also possible to estimate the DL to UL channel transformation matrix $\mathbf{T}$ at once using the total least squares estimator.

The DL to UL channel transformation matrix $\mathbf{T}$ relates the DL channel coefficients and the UL channel coefficients so that the current UL channel measurements can be exploited to estimate the DL channel coefficients. The imperfect estimates of the DL to UL channel transformation matrix $\mathbf{T}$ can result in distortion of the information contained in the UL channel measurements. Thus depending on the accuracy of estimates of the DL to UL channel transformation matrix $\mathbf{T}$, it is possible that erroneous information about the DL channel coefficients can be extracted from the current UL channel measurements. Thus the adaptive Kalman filter shall consider the reliability of estimates of the DL to UL channel transformation matrix $\mathbf{T}$ so that the information obtained from the current UL channel measurements shall be weighted accordingly. Towards this end, the estimate $\hat{\mathbf{T}}$ of the DL to UL channel transformation matrix $\mathbf{T}$ is assumed to be a superposition of the true DL to UL channel transformation matrix $\mathbf{T}$ and an estimation error modelled as Gaussian noise. Hence the measurement noise can also be used to model the error in the estimates of the DL to UL channel transformation matrix $\mathbf{T}$. The covariance matrix of the estimation error of the DL to UL channel transformation matrix $\mathbf{T}$ can be estimated using the Yule-Walker equations as shown in Appendix C by considering the evolution of the channel coefficients in the frequency domain. Initial estimates of the autocorrelation function of the channel can be obtained from the Jakes' model and the truncated one-sided exponential PDP, while

at the later stages estimates of the autocorrelation function of the channel is obtained from the estimated channel coefficients using time averaging. Thus the update stage of the Kalman filter is modified as follows to include the covariance matrix $\mathbf{R}_{\mathrm{TT}}$ of the estimation error of the DL to UL channel transformation matrix $\mathbf{T}$

$$\mathbf{K}(\ell) = \frac{\mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell)}{\mathbf{G}(\ell) \cdot \mathbf{M}(\ell|\ell-1) \cdot \mathbf{G}^{*\mathrm{T}}(\ell) + \tilde{\mathbf{R}}_{\mathrm{nn}}}, \tag{8.19}$$

where

$$\tilde{\mathbf{R}}_{\mathrm{nn}} = \mathbf{R}_{\mathrm{nn}} + \begin{pmatrix} \mathbf{R}_{\mathrm{TT}} & \mathbf{0}_{M \cdot N \cdot \acute{K}} \\ \mathbf{0}_{M \cdot N \cdot \acute{K}} & \mathbf{0}_{M \cdot N \cdot \acute{K}} \end{pmatrix}. \tag{8.20}$$

Thus at the early recursions of the adaptive Kalman filter the delayed DL measurements are given more weight than the current UL channel measurements in estimating the DL channel measurements. This enables the adaptive Kalman filter to mitigate the information loss and possible inclusion of misleading information from the processed current UL channel measurements. However, as the tracking progresses, the current UL and the delayed DL channel measurements are properly weighted to reflect the reliability of extracting information from the current UL channel measurements using the improved estimate $\hat{\mathbf{T}}$ of the DL to UL channel transformation matrix.

Estimates of the state transition matrix and the covariance matrix of the state transition noise are also periodically updated using the estimated channel coefficients. In the two-step DL channel prediction, estimates of the state transition matrix and the covariance matrix of the state transition noise are updated using the estimated DL channel coefficients. For the UL-DL channel transformation approach, periodical update of estimates of the state transition matrix and the covariance matrix of the state transition noise can be estimated from the estimated DL channel coefficients. However, in the UL-DL channel transformation, the current UL channel measurements, processed by the DL to UL channel transformation matrix, and the delayed DL channel measurements are used to obtain estimates of the DL channel coefficients. As mentioned earlier, processing the current UL channel measurements using the initial imperfect estimate of the DL to UL channel transformation matrix $\mathbf{T}$ can result in distortion of the information contained by the UL channel measurements. Consequently, using the early estimated DL channel coefficients to estimate the state transition matrix and the covariance matrix of the state transition noise can result in unreliable estimates. Thus estimates of the UL channel coefficients which are estimated using the current UL channel measurements at the BS are used to obtain estimates of the state transition matrix and the covariance matrix of the state transition noise. For both approaches, the state transition matrix and the covariance matrix of the state transition noise are estimated from the estimated channel coefficients using the same technique discussed in the previous chapter.

The UL-DL channel transformation adaptive Kalman filter initially uses the state transition matrix, the covariance matrix of the state transition noise and the DL to UL channel transformation matrix calculated using the correlation matrix of the channel obtained from the Jakes' model and the truncated one-sided exponential PDP. The estimated channel coefficients, which are obtained from the adaptive Kalman filters, are then used to periodically update the parameters of the Kalman filter. This results in an UL-DL channel transformation adaptive Kalman filter which matches the actual mobile radio channel. Fig. 8.2 shows a block diagram of the UL-DL channel transformation adaptive Kalman filter for FDD systems.

## 8.3 Performance analysis

In this section, the performances of the proposed MIMO DL channel prediction algorithms for FDD systems is analyzed considering a microcellular mobile radio system using Monte Carlo simulations with $10^4$ independent trials. Except for the number of UL and DL pilot symbols, the simulation setup is identical to the one considered in the previous chapter, i.e., Section 7.3. The generated channel coefficients are normalized to one, i.e., $\mathbb{E}\{|H_{m,n}(k,\ell)|^2\} = 1$. $K_{\mathrm{UL}} = 32$ UL and $K_{\mathrm{DL}} = 32$ DL pilot subcarriers are considered in 50 time instants. The performance metric is the normalized mean square error (NMSE) for the $\ell = 50^{\mathrm{th}}$ time instant between the true $H_{m,n}(k_{\mathrm{DL}}, \ell = 50)$ and the predicted $\check{H}_{m,n}(k_{\mathrm{DL}}, \ell = 50)$ DL channel coefficients:

$$\mathrm{nmse}(\check{H}_{\mathrm{DL}}(\ell = 50)) = \frac{\mathbb{E}_{\mathbf{e}(\ell),H_{m,n}(k_{\mathrm{DL}},\ell)}\left\{\left|\check{H}_{m,n}(k_{\mathrm{DL}}, \ell = 50) - H_{m,n}(k_{\mathrm{DL}}, \ell = 50)\right|^2\right\}}{\mathbb{E}\left\{\left|H_{m,n}(k_{\mathrm{DL}}, \ell = 50)\right|^2\right\}}. \tag{8.21}$$

The operator $\mathbb{E}_{\mathbf{e}(\ell),H_{m,n}(k_{\mathrm{DL}},\ell)}\{\cdot\}$ denotes the expectation with respect to the pdf $\mathrm{p}\left(\mathbf{e}(\ell), H_{m,n}(k_{\mathrm{DL}}, \ell)\right)$. In the simulations, the expectations are computed using time averages. The prediction performance of the adaptive Kalman filter is analyzed for different pseudo signal-to-noise-ratios (PSNRs)

$$\mathrm{PSNR} = \mathbb{E}\left\{|s(k_{\mathrm{DL}}, \ell)|^2\right\}/\sigma_n^2, \tag{8.22}$$

where $s(k_{\mathrm{DL}}, \ell)$ is the transmitted DL pilot symbol and $\sigma_n$ is the standard deviation of the measurement noise. A reference two-step DL channel prediction Kalman filter and a reference UL-DL channel transformation Kalman filter where the Kalman filter parameters are calculated from the true propagation path parameters for each run of the Monte Carlo simulations are used as performance benchmarks for the two-step DL channel prediction and the UL-DL channel transformation adaptive Kalman filters, respectively.

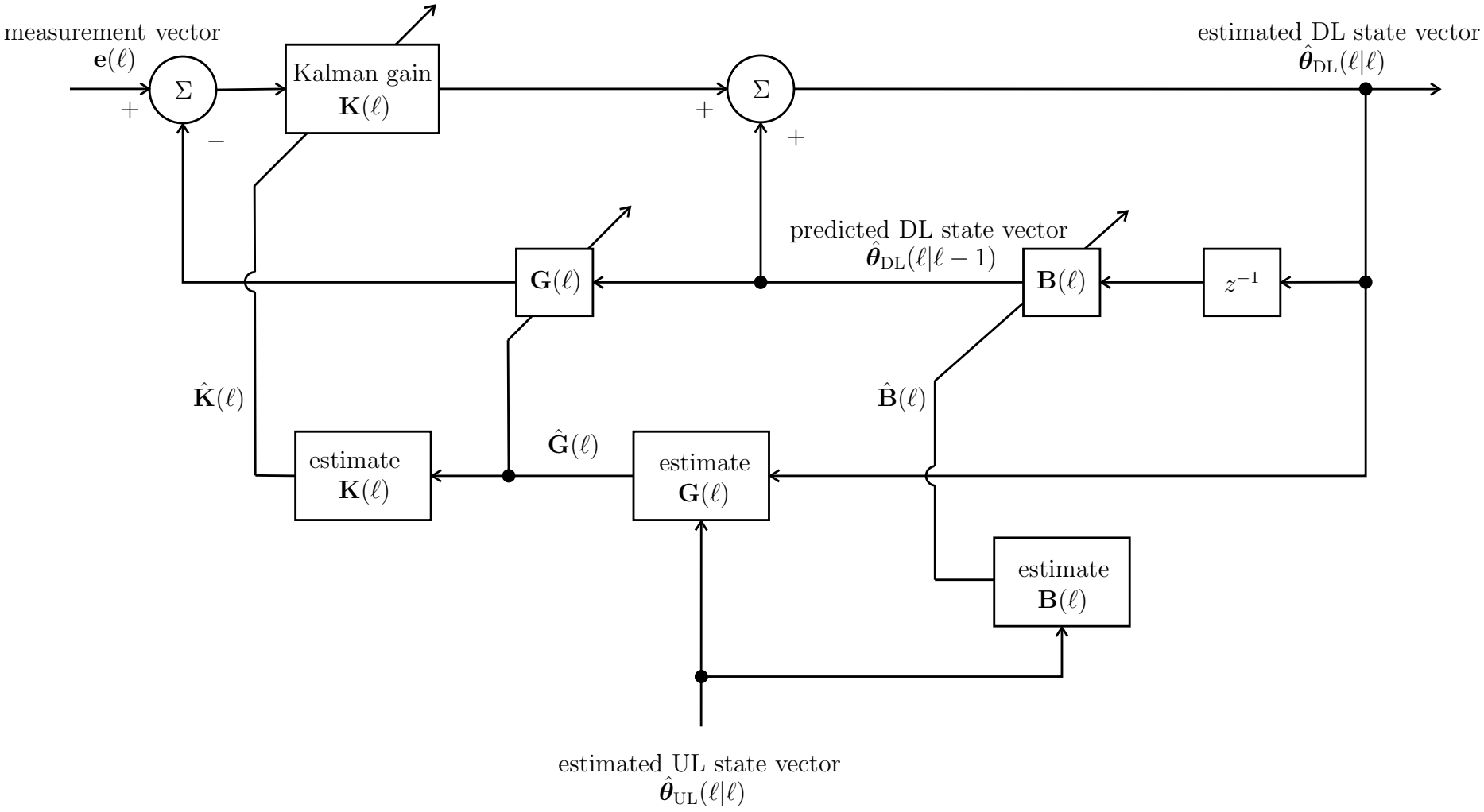



Figure 8.2. A block diagram of the UL-DL channel transformation adaptive Kalman filter for FDD systems

Fig. 8.3 shows the prediction performances of the two-step DL channel prediction and the UL-DL channel transformation adaptive Kalman filters for different numbers of propagation paths and dimensions of the MIMO channels. $2 \times 2$ and $4 \times 4$ MIMO channels are considered for $P = 5$ and $P = 10$ propagation paths. In the simulations, the linear relation between all the DL pilot symbol subcarriers for two time instants is exploited, i.e., $\acute{K} = K_{\mathrm{DL}}$ and $\acute{L} = 2$. Thus the filter dimension $\acute{K} \cdot \acute{L} = 64$ is well beyond the minimum required filter dimensions, i.e., $P = 5$ and $P = 10$. Consideration of higher filter dimensions results in an improved performance owing to the gain from noise suppression from exploitation of more measurements for each channel coefficient [SW15b]. It can be seen from Fig. 8.3 that the performances of the two-step DL channel prediction and the UL-DL channel transformation adaptive Kalman filters are bounded by their respective reference Kalman filters. It can also be seen that the performances of the reference Kalman filters depend only the number of propagations paths but not on the dimension of the MIMO channel. This occurs owing to the narrowband assumption for array signal processing which results in prediction filter coefficients which are independent of the transmitter and receiver antennas under consideration as discussed in Section 6.1. On the other hand, the performance of the adaptive Kalman filter improves as the dimension of the MIMO channel is increased. This occurs owing to the exploitation of the spatial property of the MIMO channel in predicting the channel coefficients.

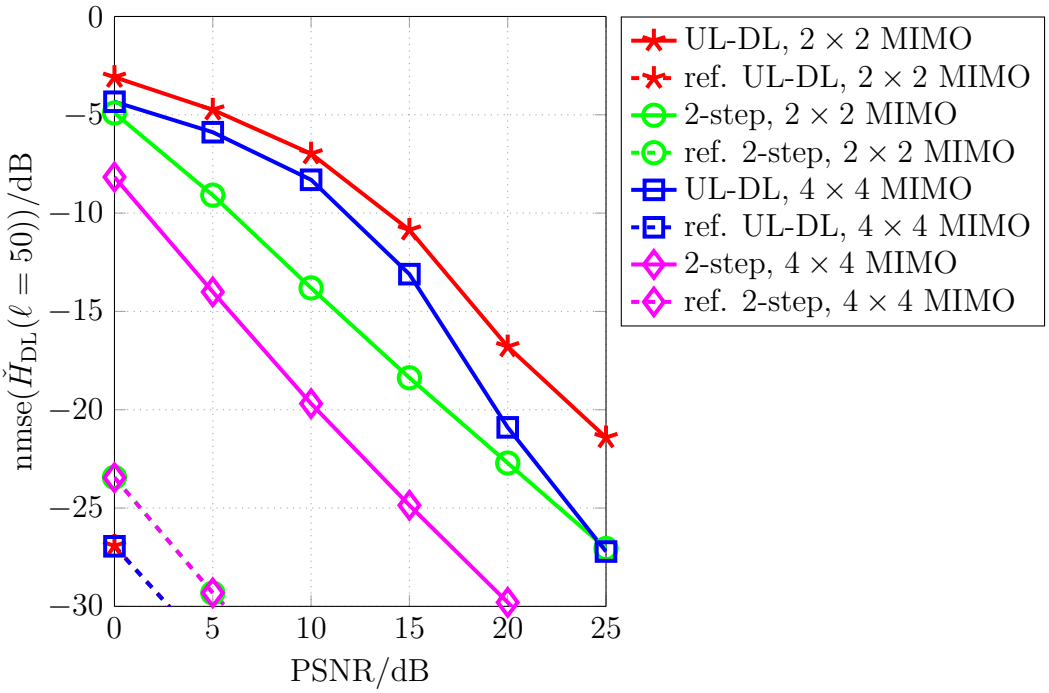
Fig. 8.3 also shows the impact of the number of the propagation paths on the performances of the reference and the adaptive Kalman filters. It can be seen that, for a given MIMO channel, as the number of the propagation paths increases, the performances of both the reference and the adaptive Kalman filters deteriorate. This result stems from the fact that for the same filter dimensions, i.e., $\acute{K} \cdot \acute{L} = 64$, more measurements than the minimum required number of measurements are exploited for the case where the number of propagation paths is smaller, i.e., $P = 5$, than for the case where the number of propagation paths is higher, i.e., $P = 10$. This results in an improved performance from noise suppression.

Comparing the two DL channel prediction approaches, it can be seen from Fig. 8.3 that the reference Kalman filter for the UL-DL channel transformation has a lower NMSE than the reference Kalman filter for the two-step DL channel prediction. This occurs owing to the fact that the UL-DL channel transformation approach exploits both the current UL and the delayed DL channel measurements, whereas the two step DL channel prediction approach exploits only the delayed DL channel measurements. Hence the UL-DL channel transformation benefits from consideration of the extra information from the current UL channel measurements. However, the two-step DL channel

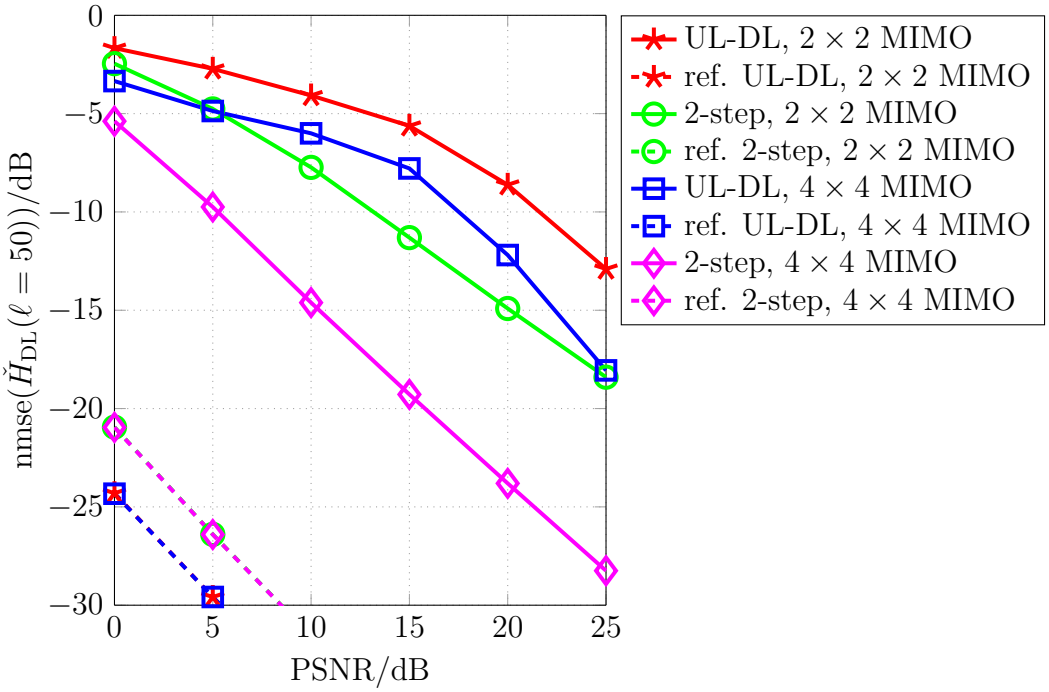
prediction adaptive Kalman filter yields a superior performance over the UL-DL channel transformation adaptive Kalman filter. This result stems from the fact that the imperfect estimates of the DL to UL channel transformation matrix $\mathbf{T}$ result in information distortion which impairs the performance of the UL-DL channel transformation adaptive Kalman filter.

Fig. 8.4 shows the convergence behaviours of the reference and the adaptive Kalman filters for different numbers of propagation paths and dimensions of the MIMO channels when PSNR = 25 dB. It can be seen that the reference two-step DL channel prediction and the UL-DL channel transformation Kalman filters converge quickly. Furthermore, the two-step DL channel prediction adaptive Kalman filter also converges quickly after $\ell > 2\acute{L}$ iterations. On the other hand, the UL-DL channel transformation adaptive Kalman filter converges gradually. This occurs owing to the fact that for the UL-DL channel transformation adaptive Kalman filter the imperfect estimates of the DL to UL channel transformation matrix $\mathbf{T}$ precludes the UL-DL channel transformation adaptive Kalman filter from exploiting all the available information. Hence it takes relatively long time before sufficient information is available to have reliable estimates of the adaptive Kalman filter parameters. In general, the two-step DL channel prediction adaptive Kalman filter starts to converge quickly after the onset of update of the two-step DL channel prediction adaptive Kalman filter parameters, whereas the UL-DL channel transformation adaptive Kalman filter converges gradually. Furthermore, it has been observed from the simulation results that the rate of convergence for the two-step DL channel prediction and the UL-DL channel transformation adaptive Kalman filters increases considerably as the PSNR increases.

Even though the UL-DL channel transformation adaptive Kalman filter has more information available than the two-step DL channel prediction adaptive Kalman filter, the lack of sufficiently correct estimates of the DL to UL channel transformation matrix can preclude the proposed UL-DL transformation algorithm from exploiting the available information. Furthermore, in some cases the gap band between the UL and DL channels is so large, e.g., up to 680 MHz [3GP14], that the assumption of radio propagation path parameters being constant over the UL and DL channels might not be a sound assumption. Consequently, in such cases, the two-step DL channel prediction is probably a safe choice in order to guarantee a satisfactory performance.

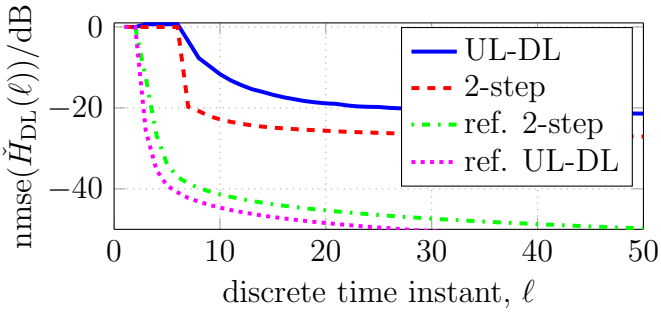



(a) The NMSE prediction performances for $P = 5$

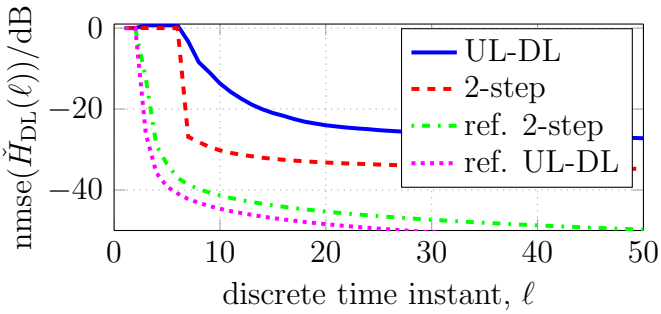



(b) The NMSE prediction performances for $P = 10$

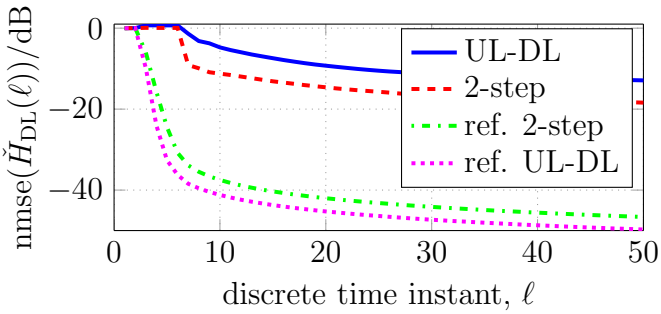
Figure 8.3. The NMSE prediction performances of the UL-DL channel transformation (UL-DL), the two-step DL channel prediction (2-step), the reference UL-DL channel transformation (ref. UL-DL) and the reference two-step DL channel prediction (ref. 2-step) Kalman filters



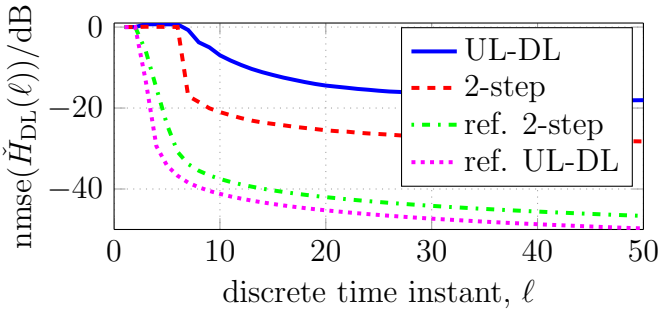

(a) The convergence behaviours for a $2 \times 2$ MIMO channel, $P = 5$




(b) The convergence behaviours for a $4 \times 4$ MIMO channel, $P = 5$




(c) The convergence behaviours for a $2 \times 2$ MIMO channel, $P = 10$




(d) The convergence behaviours for a $4 \times 4$ MIMO channel, $P = 10$

Figure 8.4. The convergence behaviours of the UL-DL channel transformation (UL-DL), the two-step DL channel prediction (2-step), the reference UL-DL channel transformation (ref. UL-DL) and the reference two-step DL channel prediction (ref. 2-step) Kalman filters, PSNR = 25 dB

# Chapter 9

# Summaries

## 9.1 Summary in English

In this thesis, the spatio-temporal channel properties of multiple-input multiple-output (MIMO) systems are exploited to obtain valuable insights and new approaches to localization and channel prediction in MIMO systems which yield elegant solutions to crucial problems in localization and channel prediction. The spatio-temporal channel properties of MIMO systems are studied using the double directional channel model which characterizes MIMO channels by the complex amplitudes, delays, Doppler shifts, angles-of-departure and angles-of-arrival of the propagation paths.

Exploitation of the spatio-temporal channel properties of MIMO systems enables localization of a mobile station (MS) in non-line-of-sight (NLOS) multipath environments under the explicit consideration of scatterers. It is assumed that the propagation paths which are considered for localization contain a maximum of one scattering. The parameters of each propagation path, i.e., the time-of-arrival, the angle-of-departure and the angle-of-arrival, define a linear relation between the positions of the BS, the corresponding scatterer and the MS. The linear relations obtained from all propagation paths are used to determine the position of the MS. Owing to the physical mobility constraints of a MS, there is a strong statistical dependency between the consecutive positions of the MS. By modeling the mobility of a MS using a first-order Gauss-Markov process, the dependency between the successive positions of the MS can be exploited to yield improved estimates of the positions of the MS.

The idea of MS localization in NLOS multipath environments under the explicit consideration of scatterers can be easily extended to network localization. In network localization, cooperation between the MSs can be used to obtain satisfactory performance even without the explicit consideration of scatterers, i.e., using only the times-of-arrival (TOA) measurements. In this case, the number of TOA measurements increases quadratically with the number of MSs, whereas the number of unknowns increases only linearly with the number of MSs. Hence the extra information from cooperation can be used to reduce the estimation error. In general, the measured TOA of the NLOS propagation paths are much larger than the true TOA of the line-of-sight (LOS) propagation paths. Hence the direct usage of the TOA measurements would result in a

severe estimation error. Consequently, nearly optimal estimates of the positions of the MSs are obtained by using techniques from robust estimation theory where the TOA measurements are systematically weighted so that the impact of outliers resulting from the NLOS propagation paths is minimized.

Valuable insights and new approaches to MIMO channel prediction can be obtained from exploitation of the spatio-temporal channel properties of MIMO systems. In the double directional channel model, each channel coefficient is represented as a weighted sum of products of complex exponentials which depend on the propagation path parameters. For small scale MS mobility, the propagation path parameters can be assumed to be constant. Consequently, each channel coefficient is a weighted sum of products of complex exponentials with the same weights for each channel coefficient. Thus the channel prediction is performed by a mere phase rotation of each of the products of the complex exponentials. Since each channel coefficient is a weighted sum of products of complex exponentials with the same weights for each channel coefficient, it is possible to predict any channel coefficient as a linear combination of other channel coefficients. It is proved that a time and frequency independent filter whose dimension is at least equal to the number of propagation paths can be used to weight the channel coefficients. The filter depends on the delays and Doppler shifts of the propagation paths. Owing to the narrowband assumption for array signal processing, the filter coefficients are the same for all single-input single-output (SISO) subchannels of a MIMO channel.

For small scale MS mobility, satisfactory channel prediction performance could be obtained by tracking a MIMO channel. For optimal MIMO channel prediction, estimates of the propagation path parameters are required to determine the filter coefficients. However, estimation of the propagation path parameters requires computationally intensive array signal processing algorithms. Thus initial estimates of the filter coefficients are determined using estimates of the channel correlation obtained from stochastic channel models. However, as most channels of interest are not ergodic, the considered stochastic channel models do not describe the statistics of the individual channel realizations. Consequently, the stochastic channel models are used for initialization only and then the estimates of the filter coefficients are periodically updated using the linear relation between the estimated channel coefficients. Further performance gains could be obtained by exploiting side information. Of particular importance for frequency-division-duplex systems is the side information obtained from the feedback channel. The delayed and hence outdated transmitter side channel state information (CSI) obtained from the feedback channel can be used independently or jointly with the receiver side CSI to obtain up-to-date transmitter side CSI.

## 9.2 Zusammenfassung in deutscher Sprache

In der Disssertation werden die raumzeitlichen Kanaleigenschaften von MIMO-Systemen (engl. multiple-input multiple-output) ausgenutzt, um wertvolle Einblicke und neue Ansätze zur Lokalisierung und Kanalprädiktion in MIMO-Systemen zu erhalten. Dadurch können elegante Lösungen für zentrale Probleme bei der Lokalisierung und Kanalprädiktion erhalten werden. Die raumzeitlichen Kanaleigenschaften von MIMO-Systemen werden mit dem doppelt-direktionalen Kanalmodell untersucht, das MIMO-Kanäle anhand der komplexen Amplituden, Laufzeiten, Dopplerfrequenzen, Einfallswinkel und Austrittswinkel der Ausbreitungspfade charakterisiert.

Das Ausnutzen der raumzeitlichen Kanaleigenschaften von MIMO-Systemen ermöglicht die Lokalisierung einer Mobilstation (MS) in NLOS-Mehrwegeumgebungen (engl. non-line-of-sight) unter der expliziter Berücksichtigung von Streuern. Es wird angenommen, dass die für die Lokalisierung genutzten Ausbreitungspfade maximal eine Streuung enthalten. Die Parameter eines Ausbreitungspfades, d.h. die Laufzeit, der Einfallswinkel und der Austrittswinkel, definieren eine lineare Beziehung zwischen den Positionen der BS, dem entsprechenden Streuer und der MS. Die Position der MS wird auf Basis der linearen Beziehungen aller Ausbreitungspfade bestimmt. Aufgrund der eingeschränkten physikalischen Mobilität einer MS gibt es eine starke statistische Abhängigkeit zwischen den aufeinanderfolgenden Positionen der MS. Durch Modellieren der Mobilität unter Verwendung eines Gauss-Markov-Prozesses erster Ordnung kann diese Abhängigkeit zur Verbesserung der Positionsschätzung genutzt werden.

Die Idee der MS Lokalisierung in NLOS-Mehrwegeumgebungen unter der expliziten Berücksichtigung von Streuern kann leicht für die Netzwerk-Lokalisierung erweitert werden. In der Netzwerk-Lokalisierung kann die Kooperation zwischen den MSen genutzt werden, um eine ausreichende Leistung auch ohne die explizite Berücksichtigung von Streuern zu erhalten, d.h. nur unter Verwendung der Laufzeitmessungen. In diesem Fall wächst die Anzahl der Laufzeitmessungen quadratisch mit der Anzahl der MSen, während die Anzahl der Unbekannten nur noch linear mit der Anzahl der MSen wächst. Somit kann die Zusatzinformation aus der Kooperation verwendet werden, um den Schätzfehler zu verringern. Die gemessenen Laufzeiten der NLOS-Ausbreitungspfade sind in der Regel viel größer als die tatsächlichen Laufzeiten der LOS-Ausbreitungspfade (engl. line-of-sight). Daher würde die direkte Verwendung der Messungen zu einem großen Schätzfehler führen. Folglich werden nahezu optimale Positionsschätzungen der MSen unter Verwendung von Techniken der robusten Schätztheorie erhalten. Dabei werden die Laufzeitmessungen systematisch gewichtet,

so dass die Auswirkung von Ausreißern resultierend aus den Messungen der NLOS-Ausbreitungspfade minimiert wird.

Wertvolle Einblicke und neue Ansätze zur MIMO-Kanalprädiktion können aus der Ausnutzung der raumzeitlichen Kanaleigenschaften von MIMO-Systemen erhalten werden. In dem doppelt-direktionalen Kanalmodell wird jeder Kanalkoeffizient als eine gewichtete Summe von Produkten komplexer Exponentialfunktionen, die von den Ausbreitungspfadparametern abhängig sind, dargestellt. Bewegt sich die MS nur geringfügig (engl. smale scale MS mobility) dann können die Ausbreitungspfadparameter als konstant angenommen werden. Folglich ist jeder Kanalkoeffizient eine gewichtete Summe von Produkten komplexer Exponentialfunktionen mit den selben Gewichten für jeden Kanalkoeffizienten. Für die Kanalprädiktion muss dann nur eine Phasendrehung der einzelnen Produkte aus den komplexen Exponentialfunktionen durchgeführt werden. Da jeder Kanalkoeffizient eine gewichtete Summe von Produkten komplexer Exponentialfunktionen mit den selben Gewichten für jeden Kanalkoeffizient ist, kann ein beliebiger Kanalkoeffizient als lineare Kombination von anderen Kanalkoeffizienten dargestellt werden. Es wird nachgewiesen, dass ein zeit- und frequenzunabhängiges Filter zur Wichtung der Kanalkoeffizienten verwendet werden kann. Die Dimension des Filters muss mindestens der Anzahl der Ausbreitungspfade entsprechen. Das Filter ist abhängig von den Laufzeiten und Dopplerfrequenzen der Ausbreitungspfade. Aufgrund der in der Array-Signalverarbeitung genutzten Schmalbandannahme sind die Filterkoeffizienten für alle SISO (engl. single-input single-output) Subkanäle eines MIMO-Kanals gleich.

Für geringfügige Bewegungen der MS konnte eine ausreichende Prädiktionleistung durch das Nachführen des MIMO-Kanals erhalten werden. Zur optimalen MIMO-Kanalprädiktion sind Schätzungen der Ausbreitungspfadparameter erforderlich, um die Filterkoeffizienten zu bestimmen. Jedoch erfordert das Schätzen der Ausbreitungspfadparameter rechenintensive Signalverarbeitungsalgorithmen. Daher werden erste Schätzungen der Filterkoeffizienten unter Verwenden von Schätzungen der Kanalkorrelation auf Basis stochastischer Kanalmodelle bestimmt. Da jedoch die meisten interessierenden Kanäle nicht ergodisch sind, beschreiben die stochastischen Kanalmodelle nicht die Statistik der einzelnen Kanalrealisierungen. Folglich werden die stochastischen Kanalmodelle nur für die Initialisierung verwendet. Die Schätzwerte der Filterkoeffizienten werden dann unter Verwenden der linearen Beziehung zwischen den geschätzten Kanalkoeffizienten periodisch aktualisiert. Weitere Performanzgewinne konnten durch Ausnutzen von Seiteninformation erzielt werden. Für Frequenzduplexsysteme ist die gewonnene Seiteninformationen aus dem Rückkanal von besonderer Bedeutung. Die aus dem Rückkanal gewonnene, verzögerte und somit veraltete senderseitige CSI (engl. channel state information) kann unabhängig oder gemeinsam mit der empfängerseitigen CSI verwendet werden, um aktuelle senderseitige CSI zu erhalten.

# Appendix A

# Calculation of the Jacobian matrix

In the following, simpler expressions for evaluating the Jacobian matrix of the propagation path parameters are presented. Let the $2(P+1)\times 1$ vector $\mathbf{p}_{\mathrm{sm}} = (\mathbf{p}_{\mathrm{sc},1}^{\mathrm{T}}, \ldots, \mathbf{p}_{\mathrm{sc},P}^{\mathrm{T}}, \mathbf{p}_{\mathrm{ms}}^{\mathrm{T}})^{\mathrm{T}}$ denote the two dimensional positions of the MS and the scatterers. Furthermore, let the $4P\times 1$ vector

$$\mathbf{g}(\mathbf{p}_{\mathrm{sm}}) = \left( \underbrace{d_1, \ldots, d_P}_{\mathbf{d}}, \underbrace{d_{1,q}, \ldots, d_{P,q}}_{\mathbf{d}_q}, \underbrace{\phi_{\mathrm{bs},1}, \ldots, \phi_{\mathrm{bs},P}}_{\boldsymbol{\phi}_{\mathrm{bs}}}, \underbrace{\phi_{\mathrm{ms},1}, \ldots, \phi_{\mathrm{ms},P}}_{\boldsymbol{\phi}_{\mathrm{ms}}} \right)^{\mathrm{T}} \tag{A.1}$$

denote the path lengths, the path length differences, the angles-of-departure and the angles-of-arrival of the $P$ propagation paths. The $4P\times 2(P+1)$ Jacobian matrix

$$\begin{aligned}
\frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial \mathbf{p}_{\mathrm{sm}}} &= \begin{pmatrix} \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial x_{\mathrm{sc},1}} & \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial x_{\mathrm{sc},P}} & \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial y_{\mathrm{sc},P}} & \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial x_{\mathrm{ms}}} & \frac{\partial \mathbf{g}(\mathbf{p}_{\mathrm{sm}})}{\partial y_{\mathrm{ms}}} \end{pmatrix} \\
&= \begin{pmatrix}
\frac{\partial d_1}{\partial x_{\mathrm{sc},1}} & \frac{\partial d_1}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial d_1}{\partial x_{\mathrm{sc},P}} & \frac{\partial d_1}{\partial y_{\mathrm{sc},P}} & \frac{\partial d_1}{\partial x_{\mathrm{ms}}} & \frac{\partial d_1}{\partial y_{\mathrm{ms}}} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial d_P}{\partial x_{\mathrm{sc},1}} & \frac{\partial d_P}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial d_P}{\partial x_{\mathrm{sc},P}} & \frac{\partial d_P}{\partial y_{\mathrm{sc},P}} & \frac{\partial d_P}{\partial x_{\mathrm{ms}}} & \frac{\partial d_P}{\partial y_{\mathrm{ms}}} \\
\frac{\partial d_{1,q}}{\partial x_{\mathrm{sc},1}} & \frac{\partial d_{1,q}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial d_1}{\partial x_{\mathrm{sc},P}} & \frac{\partial d_{1,q}}{\partial y_{\mathrm{sc},P}} & \frac{\partial d_{1,q}}{\partial x_{\mathrm{ms}}} & \frac{\partial d_{1,q}}{\partial y_{\mathrm{ms}}} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial d_{P,q}}{\partial x_{\mathrm{sc},1}} & \frac{\partial d_{P,q}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial d_{P,q}}{\partial x_{\mathrm{sc},P}} & \frac{\partial d_{P,q}}{\partial y_{\mathrm{sc},P}} & \frac{\partial d_{P,q}}{\partial x_{\mathrm{ms}}} & \frac{\partial d_{P,q}}{\partial y_{\mathrm{ms}}} \\
\frac{\partial \phi_{\mathrm{bs},1}}{\partial x_{\mathrm{sc},1}} & \frac{\partial \phi_{\mathrm{bs},1}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial \phi_{\mathrm{bs},1}}{\partial x_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{bs},1}}{\partial y_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{bs},1}}{\partial x_{\mathrm{ms}}} & \frac{\partial \phi_{\mathrm{bs},1}}{\partial y_{\mathrm{ms}}} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \phi_{\mathrm{bs},P}}{\partial x_{\mathrm{sc},1}} & \frac{\partial \phi_{\mathrm{bs},P}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial \phi_{\mathrm{bs},P}}{\partial x_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{bs},P}}{\partial y_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{bs},P}}{\partial x_{\mathrm{ms}}} & \frac{\partial \phi_{\mathrm{bs},P}}{\partial y_{\mathrm{ms}}} \\
\frac{\partial \phi_{\mathrm{ms},1}}{\partial x_{\mathrm{sc},1}} & \frac{\partial \phi_{\mathrm{ms},1}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial \phi_{\mathrm{ms},1}}{\partial x_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{ms},1}}{\partial y_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{ms},1}}{\partial x_{\mathrm{ms}}} & \frac{\partial \phi_{\mathrm{ms},1}}{\partial y_{\mathrm{ms}}} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \phi_{\mathrm{ms},P}}{\partial x_{\mathrm{sc},1}} & \frac{\partial \phi_{\mathrm{ms},P}}{\partial y_{\mathrm{sc},1}} & \cdots & \frac{\partial \phi_{\mathrm{ms},P}}{\partial x_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{ms},P}}{\partial y_{\mathrm{sc},P}} & \frac{\partial \phi_{\mathrm{ms},P}}{\partial x_{\mathrm{ms}}} & \frac{\partial \phi_{\mathrm{ms},P}}{\partial y_{\mathrm{ms}}}
\end{pmatrix}
\end{aligned} \tag{A.2}$$

can be calculated using the derivatives

$$\frac{\partial \mathbf{d}}{\partial x_{\mathrm{ms}}} = -\left( \frac{x_{\mathrm{sc},1} - x_{\mathrm{ms}}}{d_{\mathrm{ms},1}}, \ldots, \frac{x_{\mathrm{sc},P} - x_{\mathrm{ms}}}{d_{\mathrm{ms},P}} \right)^{\mathrm{T}} \tag{A.3}$$

$$\frac{\partial \mathbf{d}}{\partial y_{\mathrm{ms}}} = -\left( \frac{y_{\mathrm{sc},1} - y_{\mathrm{ms}}}{d_{\mathrm{ms},1}}, \ldots, \frac{y_{\mathrm{sc},P} - y_{\mathrm{ms}}}{d_{\mathrm{ms},P}} \right)^{\mathrm{T}} \tag{A.4}$$

$$\frac{\partial \mathbf{d}}{\partial \mathbf{x}_{\mathrm{sc}}} = \mathrm{diag}\left( \frac{x_{\mathrm{sc},1} - x_{\mathrm{ms}}}{d_{\mathrm{ms},1}}, \ldots, \frac{x_{\mathrm{sc},P} - x_{\mathrm{ms}}}{d_{\mathrm{ms},P}} \right) + \mathrm{diag}\left( \frac{x_{\mathrm{sc},1} - x_{\mathrm{bs},1}}{d_{\mathrm{bs},1}}, \ldots, \frac{x_{\mathrm{sc},P} - x_{\mathrm{bs},P}}{d_{\mathrm{bs},P}} \right) \tag{A.5}$$

$$\frac{\partial \mathbf{d}}{\partial \mathbf{y}_{\mathrm{sc}}} = \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{ms}}}{d_{\mathrm{ms},P}}\right) + \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{bs},1}}{d_{\mathrm{bs},1}},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{bs},P}}{d_{\mathrm{bs},P}}\right) \tag{A.6}$$

$$\frac{\partial \mathbf{d}_q}{\partial x_{\mathrm{ms}}} = -\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{ms}}}{d_{\mathrm{ms},P}}\right)^{\mathrm{T}} + \left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}}\right)^{\mathrm{T}} \tag{A.7}$$

$$\frac{\partial \mathbf{d}_q}{\partial y_{\mathrm{ms}}} = -\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{ms}}}{d_{\mathrm{ms},P}}\right)^{\mathrm{T}} + \left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}}\right)^{\mathrm{T}} \tag{A.8}$$

$$\begin{aligned}\frac{\partial \mathbf{d}_q}{\partial \mathbf{x}_{\mathrm{sc}}} = {} & \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{ms}}}{d_{\mathrm{ms},P}}\right) + \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{bs},1}}{d_{\mathrm{bs},1}},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{bs},P}}{d_{\mathrm{bs},P}}\right) \\ & - \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}}\right) - \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{bs},1}}{d_{\mathrm{bs},1}},\ldots,\frac{x_{\mathrm{sc},1}-x_{\mathrm{bs},1}}{d_{\mathrm{bs},1}}\right)\end{aligned} \tag{A.9}$$

$$\begin{aligned}\frac{\partial \mathbf{d}_q}{\partial \mathbf{y}_{\mathrm{sc}}} = {} & \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{ms}}}{d_{\mathrm{ms},P}}\right) + \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{bs},1}}{d_{\mathrm{bs},1}},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{bs},P}}{d_{\mathrm{bs},P}}\right) \\ & - \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}},\ldots,\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}}\right) - \mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{bs},1}}{d_{\mathrm{bs},1}},\ldots,\frac{y_{\mathrm{sc},1}-y_{\mathrm{bs},1}}{d_{\mathrm{bs},1}}\right)\end{aligned} \tag{A.10}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{bs}}}{\partial x_{\mathrm{ms}}} = \mathbf{0}_{P\times 1} \tag{A.11}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{bs}}}{\partial y_{\mathrm{ms}}} = \mathbf{0}_{P\times 1} \tag{A.12}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{bs}}}{\partial \mathbf{x}_{\mathrm{sc}}} = -\mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{bs},1}}{d_{\mathrm{bs},1}^2},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{bs},P}}{d_{\mathrm{bs},P}^2}\right) \tag{A.13}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{bs}}}{\partial \mathbf{y}_{\mathrm{sc}}} = \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{bs},1}}{d_{\mathrm{bs},1}^2},\ldots,\frac{x_{\mathrm{bs},P}-x_{\mathrm{bs},P}}{d_{\mathrm{bs},P}^2}\right) \tag{A.14}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{ms}}}{\partial x_{\mathrm{ms}}} = \left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}^2},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{ms}}}{d_{\mathrm{ms},P}^2}\right)^{\mathrm{T}} \tag{A.15}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{ms}}}{\partial y_{\mathrm{ms}}} = -\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}^2},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{ms}}}{d_{\mathrm{ms},P}^2}\right)^{\mathrm{T}} \tag{A.16}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{ms}}}{\partial \mathbf{x}_{\mathrm{sc}}} = -\mathrm{diag}\left(\frac{y_{\mathrm{sc},1}-y_{\mathrm{ms}}}{d_{\mathrm{ms},1}^2},\ldots,\frac{y_{\mathrm{sc},P}-y_{\mathrm{ms}}}{d_{\mathrm{ms},P}^2}\right) \tag{A.17}$$

$$\frac{\partial \boldsymbol{\phi}_{\mathrm{ms}}}{\partial \mathbf{y}_{\mathrm{sc}}} = \mathrm{diag}\left(\frac{x_{\mathrm{sc},1}-x_{\mathrm{ms}}}{d_{\mathrm{ms},1}^2},\ldots,\frac{x_{\mathrm{sc},P}-x_{\mathrm{ms}}}{d_{\mathrm{ms},P}^2}\right) \tag{A.18}$$

$$d_{\mathrm{bs},p} = \sqrt{(x_{\mathrm{sc},p}-y_{\mathrm{sc},p})^2 + (y_{\mathrm{sc},p}-y_{\mathrm{bs}})^2} \tag{A.19}$$

$$d_{\mathrm{ms},p} = \sqrt{(x_{\mathrm{sc},p}-x_{\mathrm{ms}})^2 + (y_{\mathrm{sc},p}-y_{\mathrm{ms}})^2}, \tag{A.20}$$

where $\mathbf{x}_{\mathrm{sc}} = (x_{\mathrm{sc},1},\ldots,x_{\mathrm{sc},P})$ and $\mathbf{y}_{\mathrm{sc}} = (y_{\mathrm{sc},1},\ldots,y_{\mathrm{sc},P})$.

# Appendix B

# Derivation of the posterior Cramér-Rao lower bound

## B.1 Posterior Cramér-Rao lower bound for a scalar parameter

**Theorem 1.** *Consider a random scalar parameter $\theta$ and an observation vector $\mathbf{z}$. The mean square error of any estimate $\hat{\theta}$ satisfies the inequality, which is often referred to as the Posterior Cramér-Rao lower bound (PCRLB) or the Van Trees bound [VT68],*

$$\begin{aligned}\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\{(\hat{\theta}-\theta)^2\} &\geq \left(-\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\frac{\partial^2 \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta^2}\right\}\right)^{-1} \\ &\geq \left(\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2\right\}\right)^{-1}, \end{aligned} \tag{B.1}$$

*where $\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\{\cdot\}$ denotes the expectation with respect to $\mathrm{p}\left(\boldsymbol{\theta},\mathbf{z}\right)$. The probability density $\mathrm{p}\left(\mathbf{z},\theta\right)$ is a joint probability density and the expectation is done over both $\theta$ and $\mathbf{z}$. The following conditions are assumed to hold:*

1. *$\frac{\partial \mathrm{p}(\mathbf{z},\theta)}{\partial\theta}$ exists and is absolutely integrable with respect to $\theta$ and $\mathbf{z}$.*
2. *$\frac{\partial^2 \mathrm{p}(\mathbf{z},\theta)}{\partial\theta^2}$ exists and is absolutely integrable with respect to $\theta$ and $\mathbf{z}$.*
3. *$\lim\limits_{\theta\to+\infty} \mathrm{b}(\theta)\mathrm{p}\left(\theta\right) = 0$ and $\lim\limits_{\theta\to-\infty} \mathrm{b}(\theta)\mathrm{p}\left(\theta\right) = 0$, where $\mathrm{b}(\theta)$ is the conditional expecation of the estimation bias.*
4. *$\lim\limits_{\theta\to+\infty} \frac{\mathrm{dp}(\theta)}{\mathrm{d}\theta} = 0$ and $\lim\limits_{\theta\to-\infty} \frac{\mathrm{dp}(\theta)}{\mathrm{d}\theta} = 0$.*

*Proof.* The conditional expectation of the estimation bias, given the true value $\theta$ of the parameter, is

$$\mathrm{b}(\theta) = \int_{-\infty}^{+\infty} (\hat{\theta}-\theta)\mathrm{p}\left(\theta|\mathbf{z}\right)\,\mathrm{d}\mathbf{z}. \tag{B.2}$$

From this one can obtain:

$$\mathrm{b}(\theta)\mathrm{p}\left(\theta\right) = \int_{-\infty}^{+\infty} (\hat{\theta}-\theta)\mathrm{p}\left(\theta|\mathbf{z}\right)\mathrm{p}\left(\theta\right)\,\mathrm{d}\mathbf{z} = \int_{-\infty}^{+\infty} (\hat{\theta}-\theta)\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z} \tag{B.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left\{\mathrm{b}(\theta)\mathrm{p}\left(\theta\right)\right\} = \frac{\mathrm{d}}{\mathrm{d}\theta}\left\{\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\right\}$$
$$= -\int_{-\infty}^{+\infty}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z} + \int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\frac{\partial\mathrm{p}\left(\mathbf{z},\theta\right)}{\partial\theta}\,\mathrm{d}\mathbf{z}. \tag{B.4}$$

Now integrate with respect to $\theta$:

$$\int_{-\infty}^{+\infty}\frac{\mathrm{d}}{\mathrm{d}\theta}\left\{\mathrm{b}(\theta)\mathrm{p}\left(\theta\right)\right\}\,\mathrm{d}\theta = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\frac{\partial\mathrm{p}\left(\mathbf{z},\theta\right)}{\partial\theta}\,\mathrm{d}\mathbf{z}\mathrm{d}\theta$$
$$\Big[\mathrm{b}(\theta)\mathrm{p}\left(\theta\right)\Big]_{-\infty}^{+\infty} = -1 + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\frac{\partial\mathrm{p}\left(\mathbf{z},\theta\right)}{\partial\theta}\,\mathrm{d}\mathbf{z}\mathrm{d}\theta. \tag{B.5}$$

Using the third assumption, the left hand side becomes zero:

$$0 = -1 + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\frac{\partial\mathrm{p}\left(\mathbf{z},\theta\right)}{\partial\theta}\,\mathrm{d}\mathbf{z}\mathrm{d}\theta. \tag{B.6}$$

Exploiting

$$\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta} = \frac{1}{\mathrm{p}\left(\mathbf{z},\theta\right)}\frac{\partial\mathrm{p}\left(\mathbf{z},\theta\right)}{\partial\theta}, \tag{B.7}$$

we obtain:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta = 1$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left((\hat{\theta}-\theta)\sqrt{\mathrm{p}\left(\mathbf{z},\theta\right)}\right)\left(\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\sqrt{\mathrm{p}\left(\mathbf{z},\theta\right)}\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta = 1. \tag{B.8}$$

Using the Schwarz inequality it follows:

$$\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\hat{\theta}-\theta)^2\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta\right)\cdot\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta\right)$$
$$\geq\left|\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left((\hat{\theta}-\theta)\sqrt{\mathrm{p}\left(\mathbf{z},\theta\right)}\right)\left(\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\sqrt{\mathrm{p}\left(\mathbf{z},\theta\right)}\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta\right|^2 = 1$$
$$\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{(\hat{\theta}-\theta)^2\right\}\cdot\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\left(\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2\right\}\geq 1$$
$$\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{(\hat{\theta}-\theta)^2\right\}\geq\left(\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\left(\frac{\partial\ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2\right\}\right)^{-1}. \tag{B.9}$$

Thus the theorem is proved. □

To prove that $\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\left(\frac{\partial \ln(\mathrm{p}(\mathbf{z},\theta))}{\partial\theta}\right)^2\right\} = -\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\frac{\partial^2 \ln(\mathrm{p}(\mathbf{z},\theta))}{\partial\theta^2}\right\}$, consider the following:

$$\begin{aligned}
\mathrm{p}(\theta) &= \int_{-\infty}^{+\infty} \mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z} \\
\frac{\mathrm{dp}(\theta)}{\mathrm{d}\theta} &= \frac{\partial}{\partial\theta}\left\{\int_{-\infty}^{+\infty} \mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\right\} \\
\frac{\mathrm{dp}(\theta)}{\mathrm{d}\theta} &= \int_{-\infty}^{+\infty} \frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z} \\
\frac{\mathrm{d}^2\mathrm{p}(\theta)}{\mathrm{d}\theta^2} &= \frac{\partial}{\partial\theta}\left\{\int_{-\infty}^{+\infty} \frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\right\} \\
\frac{\mathrm{d}^2\mathrm{p}(\theta)}{\mathrm{d}\theta^2} &= \int_{-\infty}^{+\infty} \frac{\partial^2 \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta^2}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z} + \int_{-\infty}^{+\infty}\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2 \mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}. \qquad \text{(B.10)}
\end{aligned}$$

In the third and in the last line the relation in (B.7) was used. Integrating with respect to $\theta$ yields:

$$\frac{\mathrm{dp}(\theta)}{\mathrm{d}\theta}\Bigg|_{-\infty}^{+\infty} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{\partial^2 \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta^2}\mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2 \mathrm{p}\left(\mathbf{z},\theta\right)\,\mathrm{d}\mathbf{z}\mathrm{d}\theta \qquad \text{(B.11)}$$

Using the fourth assumption, it follows that:

$$\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\left(\frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta}\right)^2\right\} = -\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\left\{\frac{\partial^2 \ln\left(\mathrm{p}\left(\mathbf{z},\theta\right)\right)}{\partial\theta^2}\right\} \qquad \text{(B.12)}$$

holds.

## B.2 Posterior Cramér-Rao lower bound for a vector parameter

**Theorem 2.** *For a random $r$-dimensional vector parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r)$ and an observation vector $\mathbf{z}$, the mean square error of any estimate $\hat{\theta}_i$ satisfies the inequality*

$$\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\{(\hat{\theta}_i - \theta_i)^2\} \geq [\mathbf{J}^{-1}]_{ii}, \qquad \text{(B.13)}$$

*where* $[\mathbf{J}^{-1}]_{ii}$ *is the* $i^{\text{th}}$ *row and the* $i^{\text{th}}$ *column element of the inverse of the* $r \times r$ *matrix* $\mathbf{J}$ *with the elements*

$$\begin{aligned}[\mathbf{J}]_{ij} &= -\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}} \left\{ \frac{\partial^2 \ln\left(\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\right)}{\partial\theta_i \partial\theta_j} \right\} \\ &= \mathbb{E}_{\mathbf{z},\boldsymbol{\theta}} \left\{ \frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\right)}{\partial\theta_i} \frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\right)}{\partial\theta_j} \right\}, \qquad i,j = 1,\ldots,r. \end{aligned} \tag{B.14}$$

*Proof.* The proof follows the same conditions and procedures as for the scalar PCRLB case. Consider the vector parameter version of the product of the conditional expectation of the estimation bias and the a priori probability defined in equation (B.3)

$$\mathrm{b}(\theta_i)\mathrm{p}\left(\boldsymbol{\theta}\right) = \int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i)\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\, \mathrm{d}\mathbf{z}. \tag{B.15}$$

Differentiating with respect to $\theta_j$ yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta_j}\left\{\mathrm{b}(\theta_i)\mathrm{p}\left(\boldsymbol{\theta}\right)\right\} &= \frac{\mathrm{d}}{\mathrm{d}\theta_j}\left\{ \int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i)\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\, \mathrm{d}\mathbf{z} \right\} \\ &= -\delta_{ij} \int_{-\infty}^{+\infty} \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\, \mathrm{d}\mathbf{z} + \int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i) \frac{\partial \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)}{\partial\theta_j}\, \mathrm{d}\mathbf{z}, \end{aligned} \tag{B.16}$$

where $\delta_{ij}$ is a Kronecker delta function which equals to 1 when $i = j$ and which equals to 0 when $i \neq j$. Now performing a multidimensional integration with respect to $\boldsymbol{\theta}$:

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{\mathrm{d}}{\mathrm{d}\theta_j}\left\{\mathrm{b}(\theta_i)\mathrm{p}\left(\boldsymbol{\theta}\right)\right\}\, \mathrm{d}\boldsymbol{\theta} &= -\delta_{ij} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\, \mathrm{d}\mathbf{z}\mathrm{d}\boldsymbol{\theta} + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i) \frac{\partial \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)}{\partial\theta_j}\, \mathrm{d}\mathbf{z}\mathrm{d}\boldsymbol{\theta} \\ \Big[\mathrm{b}(\theta_i)\mathrm{p}\left(\boldsymbol{\theta}\right)\Big]_{-\infty}^{+\infty} &= -\delta_{ij} + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i) \frac{\partial \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)}{\partial\theta_j}\, \mathrm{d}\mathbf{z}\mathrm{d}\boldsymbol{\theta}. \end{aligned} \tag{B.17}$$

Using the vector form of the third assumption, the left hand side of the equation becomes zero:

$$0 = -\delta_{ij} + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i) \frac{\partial \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)}{\partial\theta_j}\, \mathrm{d}\mathbf{z}\mathrm{d}\boldsymbol{\theta}. \tag{B.18}$$

Exploiting the relation shown in equation (B.7), we obtain

$$\delta_{ij} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\hat{\theta}_i - \theta_i) \frac{\partial \ln\left(\mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\right)}{\partial\theta_j} \mathrm{p}\left(\mathbf{z},\boldsymbol{\theta}\right)\, \mathrm{d}\mathbf{z}\mathrm{d}\boldsymbol{\theta}. \tag{B.19}$$

Consider the case where $i = 1$ for simplicity. Define an $r + 1$ dimensional vector

$$\boldsymbol{\zeta} = \begin{pmatrix} \hat{\theta}_1 - \theta_1 \\ \frac{\partial \ln(\mathrm{p}(\mathbf{z},\boldsymbol{\theta}))}{\partial \theta_1} \\ \vdots \\ \frac{\partial \ln(\mathrm{p}(\mathbf{z},\boldsymbol{\theta}))}{\partial \theta_r} \end{pmatrix}. \tag{B.20}$$

The covariance matrix of the vector $\boldsymbol{\zeta}$ is

$$\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\{\boldsymbol{\zeta}\boldsymbol{\zeta}^{\mathrm{T}}\} = \begin{pmatrix} \mathrm{var}(\hat{\theta}_1 - \theta_1) & 1 & 0 & \dots & 0 \\ 1 & J_{11} & J_{12} & \dots & J_{1r} \\ 0 & J_{12} & J_{22} & \dots & J_{2r} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & J_{r1} & J_{r2} & \dots & J_{rr} \end{pmatrix} \tag{B.21}$$

where the ones and zeros in the matrix come from the Kronecker delta function $\delta_{ij}$ according to equation (B.19). Since it is a covariance matrix, it is positive semidefinite, which implies that the determinant of the entire matrix is greater than or equal to zero. Evaluating the determinant of the covariance matrix $\mathbb{E}_{\mathbf{z},\boldsymbol{\theta}}\{\boldsymbol{\zeta}\boldsymbol{\zeta}^{\mathrm{T}}\}$ using a cofactor (adjunct) expansion yields:

$$\mathrm{var}(\hat{\theta}_1 - \theta_1) \det(\mathbf{J}) - \det \begin{pmatrix} 1 & J_{12} & \dots & J_{1r} \\ 0 & J_{22} & \dots & J_{2r} \\ \vdots & \vdots & & \vdots \\ 0 & J_{r2} & \dots & J_{rr} \end{pmatrix} \geq 0. \tag{B.22}$$

The determinant can be further simplified as:

$$\det \begin{pmatrix} 1 & J_{12} & \dots & J_{1r} \\ 0 & J_{22} & \dots & J_{2r} \\ \vdots & \vdots & & \vdots \\ 0 & J_{r2} & \dots & J_{rr} \end{pmatrix} = \det \begin{pmatrix} J_{22} & \dots & J_{2r} \\ \vdots & & \vdots \\ J_{r2} & \dots & J_{rr} \end{pmatrix} = \mathrm{cofactor}\{[\mathbf{J}]_{11}\}. \tag{B.23}$$

Equation (B.22) can thus be rewritten as:

$$\mathrm{var}(\hat{\theta}_1 - \theta_1) \det(\mathbf{J}) - \mathrm{cofactor}\{[\mathbf{J}]_{11}\} \geq 0. \tag{B.24}$$

If we assume that $\mathbf{J}$ is nonsingular, then

$$\mathrm{var}(\hat{\theta}_1 - \theta_1) \geq \frac{\mathrm{cofactor}\{[\mathbf{J}]_{11}\}}{\det(\mathbf{J})} = [\mathbf{J}^{-1}]_{11}, \tag{B.25}$$

which is the desired result. □

The modifications for singular $\mathbf{J}$ follow easily by considering the nonsingular part of the matrix $\mathbf{J}$. The proof for $i \neq 1$ is also an obvious modification.

# Appendix C

# Modeling channels as an autoregressive process

The mobile radio channel has been modelled as an autoregressive (AR) process [Ekm02, CZ04, SM05, Aro11, AVWS12]. In the following, a SISO channel is considered for simplicity. The SISO channel is assumed to be a realization of a random process. Let the column vector $\mathbf{h}(\ell)$ denote the vector of $K$ channel coefficients at the $\ell^{\text{th}}$ discrete time instant. The $\acute{L}^{\text{th}}$ order vector AR model of the channel is given as

$$\mathbf{h}(\ell) = \sum_{\acute{\ell}=1}^{\acute{L}} \mathbf{\Omega}(\acute{\ell}) \cdot \mathbf{h}(\ell - \acute{\ell}) + \mathbf{u}(\ell), \tag{C.1}$$

where $\mathbf{\Omega}(\acute{\ell})$ are fixed coefficient matrices and $\mathbf{u}(\ell)$ is assumed to be a multivariate white Gaussian noise. Equation (C.1) can be re-written as

$$\begin{pmatrix} \mathbf{h}(\ell) \\ \mathbf{h}(\ell-1) \\ \vdots \\ \mathbf{h}(\ell-\acute{L}+1) \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{\Omega}(1) & \cdots & \mathbf{\Omega}(\acute{L}-1) & \mathbf{\Omega}(\acute{L}) \\ \mathbf{I}_K & \cdots & \mathbf{0}_K & \mathbf{0}_K \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_K & \cdots & \mathbf{I}_K & \mathbf{0}_K \end{pmatrix}}_{\mathbf{\Lambda}} \cdot \begin{pmatrix} \mathbf{h}(\ell-1) \\ \mathbf{h}(\ell-2) \\ \vdots \\ \mathbf{h}(\ell-\acute{L}) \end{pmatrix} + \begin{pmatrix} \mathbf{u}(\ell) \\ \mathbf{0}_{K\times 1} \\ \vdots \\ \mathbf{0}_{K\times 1} \end{pmatrix}. \tag{C.2}$$

The above vector autoregressive process is stable if all the eigenvalues of the matrix $\mathbf{\Lambda}$ are less than one in magnitude [Kay93, Lüt07].

Given estimates of the channel coefficient vectors $\mathbf{h}(\ell)$, the fixed coefficient matrices $\mathbf{\Omega}(\acute{\ell})$ can be estimated from (C.1) using the method of least squares. However, the least squares estimate of the fixed coefficient matrices $\mathbf{\Omega}(\acute{\ell})$ is often derived, without estimates of the channel coefficient vectors, using the so-called Yule-Walker equations. The Yule-Walker equations are derived based on the stationary autocorrelation function of the channel which satisfy [Kay93]

$$\mathbf{R}_{\text{hh}}(\ell) = \begin{cases} \displaystyle\sum_{\acute{\ell}=1}^{\acute{L}} \mathbf{\Omega}(\acute{\ell}) \cdot \mathbf{R}_{\text{hh}}(\ell - \acute{\ell}) & \ell \geq 1 \\ \displaystyle\sum_{\acute{\ell}=1}^{\acute{L}} \mathbf{\Omega}(\acute{\ell}) \cdot \mathbf{R}_{\text{hh}}(-\acute{\ell}) + \mathbf{R}_{\text{uu}} & \ell = 0 \end{cases}, \tag{C.3}$$

where

$$\mathbf{R}_{\text{hh}}(\acute{\ell}) = \mathbb{E}\{\mathbf{h}(\ell) \cdot \mathbf{h}^{*\text{T}}(\ell - \acute{\ell})\} \tag{C.4}$$

is the autocorrelation matrix of the channel and $\mathbf{R}_{\mathrm{uu}}$ is the covariance matrix of the noise $\mathbf{u}(\ell)$. The above equation can be set up to a system of linear equations

$$\begin{pmatrix} \mathbf{R}_{\mathrm{hh}}(-1) \\ \mathbf{R}_{\mathrm{hh}}(-2) \\ \vdots \\ \mathbf{R}_{\mathrm{hh}}(-\acute{L}) \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\mathrm{hh}}(0) & \mathbf{R}_{\mathrm{hh}}(1) & \dots & \mathbf{R}_{\mathrm{hh}}(\acute{L}-1) \\ \mathbf{R}_{\mathrm{hh}}(-1) & \mathbf{R}_{\mathrm{hh}}(0) & \dots & \mathbf{R}_{\mathrm{hh}}(\acute{L}-2) \\ \vdots & \vdots & & \vdots \\ \mathbf{R}_{\mathrm{hh}}(-(\acute{L}-1)) & \mathbf{R}_{\mathrm{hh}}(-(\acute{L}-2)) & \dots & \mathbf{R}_{\mathrm{hh}}(0) \end{pmatrix} \cdot \begin{pmatrix} \mathbf{\Omega}^{*\mathrm{T}}(1) \\ \mathbf{\Omega}^{*\mathrm{T}}(2) \\ \vdots \\ \mathbf{\Omega}^{*\mathrm{T}}(\acute{L}) \end{pmatrix}, \tag{C.5}$$

using the relation $\mathbf{R}_{\mathrm{hh}}^{*\mathrm{T}}(\acute{\ell}) = \mathbf{R}_{\mathrm{hh}}(-\acute{\ell})$. Similarly, an equation can be obtained containing the covariance matrix of the noise, i.e.,

$$\mathbf{R}_{\mathrm{hh}}(0) = \sum_{\acute{\ell}=1}^{\acute{L}} \mathbf{\Omega}(\acute{\ell}) \cdot \mathbf{R}_{\mathrm{hh}}(-\acute{\ell}) + \mathbf{R}_{\mathrm{uu}}. \tag{C.6}$$

These equations are known as the Yule-Walker equations. Estimates of the autocorrelation function of the channel can be obtained from stochastic channel models under the WSSUS assumption. Thus given estimates of the autocorrelation function of the channel, the fixed coefficient matrices $\mathbf{\Omega}(\ell)$ can be determined by solving the Yule-Walker equations. The covariance matrix $\mathbf{R}_{\mathrm{uu}}$ of the noise $\mathbf{u}(\ell)$ can be determined using estimates of the fixed coefficient matrices $\mathbf{\Omega}(\ell)$ and the autocorrelation function of the channel.

It is often assumed that the autocorrelation function of the channel is separable in to time-domain and frequency-domain correlations. For this case the autocorrelation matrix of the channel can be simplified as

$$\mathbf{R}_{\mathrm{hh}}(\acute{\ell}) = r_{\mathrm{t}}(\acute{\ell}) \cdot \mathbf{R}_{\mathrm{f}}, \tag{C.7}$$

where $r_{\mathrm{t}}(\acute{\ell})$ is the temporal autocorrelation function and $\mathbf{R}_{\mathrm{f}}$ is the spectral autocorrelation matrix of the channel. The Yule-Walker equation (C.5) can be re-written as

$$\underbrace{\begin{pmatrix} r_{\mathrm{t}}(-1) \\ r_{\mathrm{t}}(-2) \\ \vdots \\ r_{\mathrm{t}}(-\acute{L}) \end{pmatrix}}_{\mathbf{r}_{\mathrm{t}}} \otimes \mathbf{R}_{\mathrm{f}} = \underbrace{\begin{pmatrix} r_{\mathrm{t}}(0) & r_{\mathrm{t}}(1) & \dots & r_{\mathrm{t}}(\acute{L}-1) \\ r_{\mathrm{t}}(-1) & r_{\mathrm{t}}(0) & \dots & r_{\mathrm{t}}(\acute{L}-2) \\ \vdots & \vdots & & \vdots \\ r_{\mathrm{t}}(-(\acute{L}-1)) & r_{\mathrm{t}}(-(\acute{L}-2)) & \dots & r_{\mathrm{t}}(0) \end{pmatrix}}_{\mathbf{R}_{\mathrm{t}}} \otimes \mathbf{R}_{\mathrm{f}} \cdot \underbrace{\begin{pmatrix} \mathbf{\Omega}^{*\mathrm{T}}(1) \\ \mathbf{\Omega}^{*\mathrm{T}}(2) \\ \vdots \\ \mathbf{\Omega}^{*\mathrm{T}}(\acute{L}) \end{pmatrix}}_{\mathbf{\Omega}^{*\mathrm{T}}}. \tag{C.8}$$

Solving for $\mathbf{\Omega}^{*\mathrm{T}}$ using the method of least squares results in

$$\begin{aligned} \hat{\mathbf{\Omega}}^{*\mathrm{T}} &= ((\mathbf{R}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}})^{*\mathrm{T}} \cdot (\mathbf{R}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}}))^{-1} \cdot (\mathbf{R}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}})^{*\mathrm{T}} \cdot (\mathbf{r}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}}), \\ &= ((\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{t}}) \otimes (\mathbf{R}_{\mathrm{f}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{f}}))^{-1} \cdot (\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \otimes \mathbf{R}_{\mathrm{f}}^{*\mathrm{T}}) \cdot (\mathbf{r}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}}), \\ &= ((\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{t}})^{-1} \otimes (\mathbf{R}_{\mathrm{f}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{f}})^{-1}) \cdot (\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \otimes \mathbf{R}_{\mathrm{f}}^{*\mathrm{T}}) \cdot (\mathbf{r}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}}), \end{aligned}$$

$$
\begin{aligned}
&= ((\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{t}})^{-1} \cdot \mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \otimes (\mathbf{R}_{\mathrm{f}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{f}})^{-1} \cdot \mathbf{R}_{\mathrm{f}}^{*\mathrm{T}}) \cdot (\mathbf{r}_{\mathrm{t}} \otimes \mathbf{R}_{\mathrm{f}}), \\
&= (\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{t}})^{-1} \cdot \mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{r}_{\mathrm{t}} \otimes (\mathbf{R}_{\mathrm{f}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{f}})^{-1} \cdot \mathbf{R}_{\mathrm{f}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{f}}, \\
&= (\mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{R}_{\mathrm{t}})^{-1} \cdot \mathbf{R}_{\mathrm{t}}^{*\mathrm{T}} \cdot \mathbf{r}_{\mathrm{t}} \otimes \mathbf{I}_K. \qquad \text{(C.9)}
\end{aligned}
$$

Thus if the autocorrelation function of the channel is separable in to time-domain and frequency-domain correlations, then the frequency domain correlation matrix $\mathbf{R}_{\mathrm{f}}$ has no influence on the estimated fixed coefficient matrix $\hat{\mathbf{\Omega}}$.

# Appendix D
# Glossary of symbols and abbreviations

## D.1 Symbols

(Boldface lower case characters denote vectors, whereas boldface upper case characters denote matrices.)

| | |
|---|---|
| $\hat{(\cdot)}$ | denotes estimate |
| $(\cdot)^{\mathrm{T}}$ | denotes transpose |
| $(\cdot)^*$ | denotes conjugate |
| $(\cdot)^{-1}$ | denotes inverse |
| $(\cdot)^+$ | denotes the Moore-Penrose pseudoinverse of a matrix |
| $\odot$ | denotes the Hadamard (element-wise) matrix product |
| $*$ | denotes the Khatri-Rao (row-wise Kronecker) matrix product |
| $\sim$ | denotes is distributed according to |
| $\lvert\cdot\rvert$ | denotes absolute value |
| $\lVert\cdot\rVert_2$ | denotes the Euclidean norm |
| $\lVert\cdot\rVert_{\mathrm{F}}$ | denotes the Frobenius norm |
| $\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ | Jacobian matrix of vector-valued function $\mathbf{g}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ |
| $\mathbf{0}$ | zero matrix |
| $\mathbf{1}$ | vector of ones |
| $\arg\max_{\boldsymbol{\theta}} \mathrm{p}(\boldsymbol{\theta})$ | denotes the value of $\boldsymbol{\theta}$ that maximizes $\mathrm{p}(\boldsymbol{\theta})$ |
| $\arg\min_{\boldsymbol{\theta}} \mathrm{p}(\boldsymbol{\theta})$ | denotes the value of $\boldsymbol{\theta}$ that minimizes $\mathrm{p}(\boldsymbol{\theta})$ |
| $a_i$ or $a(i)$ | the $i^{\text{th}}$ element of vector $\mathbf{a}$ |
| $a_{ij}$ or $a(i,j)$ | the $(i,j)^{\text{th}}$ entry of matrix $\mathbf{A}$ |
| $B$ | bandwidth |
| $\mathbf{B}$ | state transition matrix |
| $\mathbf{D}$ | filter coefficient matrix |
| $\det(\mathbf{A})$ | determinant of matrix $\mathbf{A}$ |
| $\mathrm{diag}(\mathbf{a})$ | diagonal matrix with elements $\mathbf{a}$ on main diagonal |
| $d_p$ | length of the $p^{\text{th}}$ propagation path |
| $d_{p,q}$ | path length difference of the $p^{\text{th}}$ and the $q^{\text{th}}$ propagation paths |
| $\mathbf{e}$ | vector of received pilot symbols |
| $\mathbb{E}_{\boldsymbol{\theta}}\{\cdot\}$ | denotes expectation with respect to the pdf $\mathrm{p}(\boldsymbol{\theta})$ |
| $F$ | subcarrier spacing |

| | |
|---|---|
| $f$ | continuous frequency |
| $f_0$ | carrier frequency |
| $\mathbf{G}$ | measurement matrix |
| $\mathbf{g}(\cdot)$ | vector-valued measurement function |
| $\mathbf{H}(k,\ell)$ | MIMO channel coefficient matrix of subcarrier $k$ at time instant $\ell$ |
| $\mathbf{h}(\ell)$ | vector of MIMO channel coefficients at time instant $\ell$ |
| $H_{m,n}(f,t)$ | channel transfer function of the channel between the $n^{\text{th}}$ transmitter and the $m^{\text{th}}$ receiver antenna at frequency $f$ and time $t$ |
| $H_{m,n}(k,\ell)$ | channel coefficient of the channel between the $n^{\text{th}}$ transmitter and the $m^{\text{th}}$ receiver antenna for the $k^{\text{th}}$ subcarrier at time instant $\ell$ |
| $\mathbf{h}_{m,n}(\ell)$ | vector of channel coefficients of the channel between the $n^{\text{th}}$ transmitter and the $m^{\text{th}}$ receiver antenna at time instant $\ell$ |
| $\mathbf{h}_{\text{DL}}(\ell)$ | vector of DL MIMO channel coefficients at time instant $\ell$ |
| $\mathbf{h}_{\text{UL}}(\ell)$ | vector of UL MIMO channel coefficients at time instant $\ell$ |
| $\mathbf{I}$ | identity matrix |
| $\mathbf{J}$ | Fisher information matrix |
| $K$ | number of subcarriers |
| $\mathbf{K}$ | Kalman gain matrix |
| $k$ | subcarrier index |
| $\acute{K}$ | number of channel coefficient samples in frequency direction |
| $L$ | number of time instants |
| $\ell$ | time instant |
| $\acute{L}$ | number of channel coefficient samples in time direction |
| $M$ | number of receiver antennas |
| $\mathbf{M}$ | mean square error matrix of $\hat{\boldsymbol{\theta}}$ |
| $\mathcal{N}(\mu,\sigma^2)$ | Gaussian distribution with mean $\mu$ and variance $\sigma^2$ |
| $\mathcal{N}(\boldsymbol{\mu},\mathbf{R})$ | multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\mathbf{R}$ |
| $N$ | number of transmitter antennas |
| $\text{nmse}(\hat{\boldsymbol{\theta}})$ | normalized mean square error of $\hat{\boldsymbol{\theta}}$ |
| $\mathbf{n}$ | vector of received pilot symbol measurement noise |
| $P$ | number of propagation paths |
| $p$ | index of the propagation paths |
| $\text{p}(\cdot)$ | probability density function or probability mass function |
| $\text{p}(\mathbf{z}\|\boldsymbol{\theta})$ | conditional probability density function of $\mathbf{z}$ conditioned on $\boldsymbol{\theta}$ |
| $\text{p}(\boldsymbol{\theta},\mathbf{z})$ | joint probability density function of $\boldsymbol{\theta}$ and $\mathbf{z}$ |
| $\mathbf{p}_{\text{bs}}$ | position of the BS |
| $\mathbf{p}_{\text{ms}}$ | position of the MS |
| $\mathbf{p}_{\text{mss}}$ | positions of the MSs in network localization |
| $\mathbf{p}_{\text{sc},p}$ | position of the $p^{\text{th}}$ scatter |

| | |
|---|---|
| $\mathbf{p}_{\mathrm{sm}}$ | positions of the scatterers and the MS |
| $\mathrm{rmse}(\hat{\boldsymbol{\theta}})$ | root mean square error of $\hat{\boldsymbol{\theta}}$ |
| $\mathbf{R}_{\mathrm{DL,DL}}$ | covariance matrix of the MIMO DL channel |
| $\mathbf{R}_{\mathrm{nn}}$ | covariance matrix of the measurement noise $\mathbf{n}$ |
| $\mathbf{R}_{\mathrm{UL,DL}}$ | covariance matrix of the MIMO UL and DL channels |
| $\mathbf{R}_{\mathrm{uu}}$ | covariance matrix of the state transition noise $\mathbf{u}$ |
| $\mathbf{R}_{\mathrm{ww}}$ | covariance matrix of the measurement noise $\mathbf{w}$ |
| $\mathbf{R}_{\theta\theta}(0)$ | covariance matrix of the initial state vector $\boldsymbol{\theta}(0)$ |
| $r_{\mathrm{f}}(k)$ | spectral autocorrelation function of the channel |
| $r_{\mathrm{t}}(\ell)$ | temporal autocorrelation function of the channel |
| $\mathbf{S}(\ell)$ | diagonal transmit pilot symbol matrix with diagonal elements $s(k,\ell)$ |
| $s(k,\ell)$ | transmit pilot symbol of the $k^{\mathrm{th}}$ subcarrier at time instant $\ell$ |
| $T$ | transmit symbol duration |
| $\mathbf{T}$ | DL to UL channel transformation matrix |
| $t$ | continuous time |
| $\mathbf{u}$ | state transition noise |
| $\mathrm{vec}(\cdot)$ | matrix vectorization operator |
| $\mathbf{w}$ | propagation path parameters measurement noise |
| $\mathbf{z}$ | measured propagation path parameters |
| $\alpha_p$ | complex amplitude of the $p^{\mathrm{th}}$ propagation path |
| $\Delta k$ | prediction width in frequency |
| $\Delta \ell$ | prediction width in time |
| $\Delta t$ | MS tracking interval |
| $\delta_{ij}$ | Kronecker delta |
| $\varepsilon$ | fraction of the propagation paths which are NLOS |
| $\boldsymbol{\theta}$ | state vector |
| $\hat{\boldsymbol{\theta}}(\ell\|\ell)$ | estimated state vector at time instant $\ell$ |
| $\hat{\boldsymbol{\theta}}(\ell\|\ell-1)$ | predicted state vector at time instant $\ell$ |
| $\boldsymbol{\theta}_{\mathrm{DL}}$ | state vector of the DL MIMO channel |
| $\theta_{\mathrm{bs},p}$ | elevation angle-of-departure of the $p^{\mathrm{th}}$ propagation path |
| $\theta_{\mathrm{ms},p}$ | elevation angle-of-arrival of the $p^{\mathrm{th}}$ propagation path |
| $\nu_p$ | Doppler shift of the $p^{\mathrm{th}}$ propagation path |
| $\nu_{\max}$ | maximum Doppler shift |
| $\sigma^2$ | variance |
| $\tau_{m,p}$ | delay of the $p^{\mathrm{th}}$ propagation path from the receiver antenna array RP to the $m^{\mathrm{th}}$ receiver antenna |
| $\tau_{n,p}$ | delay of the $p^{\mathrm{th}}$ propagation path from the transmitter antenna array RP to the $n^{\mathrm{th}}$ transmitter antenna |
| $\tau_p$ | delay of the $p^{\mathrm{th}}$ propagation path |

| | |
|---|---|
| $\tau_{\max}$ | maximum delay |
| $\mathbf{\Phi}_{\Delta k,\Delta \ell}$ | prediction filter coefficient matrix |
| $\phi_{\mathrm{bs},p}$ | azimuth angle-of-departure of the $p^{\mathrm{th}}$ propagation path |
| $\phi_{\mathrm{ms},p}$ | azimuth angle-of-arrival of the $p^{\mathrm{th}}$ propagation path |
| $\varphi_{\Delta k,\Delta \ell}$ | prediction filter coefficient |
| $\boldsymbol{\varphi}_{\Delta k,\Delta \ell}$ | prediction filter coefficient vector |

## D.2 Abbreviations

| | |
|---|---|
| **AKF** | adaptive Kalman filter |
| **AOA** | angle-of-arrival |
| **AOD** | angle-of-departure |
| **AR** | autoregressive |
| **BS** | base station |
| **CRLB** | Cramér-Rao lower bound |
| **CSI** | channel state information |
| **DL** | downlink |
| **EKF** | extended Kalman filter |
| **EM** | electromagnetic |
| **FDD** | frequency-division-duplex |
| **FIM** | Fisher information matrix |
| **i.i.d.** | independent and identically distributed |
| **IRLS** | iterative re-weighted least squares |
| **KF** | Kalman filter |
| **LOS** | line-of-sight |
| **LS** | least squares |
| **MAD** | median absolute deviation |
| **MIMO** | multiple-input multiple-output |
| **ML** | maximum likelihood |
| **MMSE** | minimum mean square error |
| **MS** | mobile station |
| **MSE** | mean square error |
| **NLOS** | non-line-of-sight |
| **NMSE** | normalized mean square error |
| **PCRLB** | posterior Cramér-Rao lower bound |
| **pdf** | probability density function |
| **PDP** | power delay profile |
| **pmf** | probability mass function |

| | |
|---|---|
| **PSNR** | pseudo signal-to-noise-ratio |
| **RMSE** | root mean square error |
| **RP** | reference point |
| **SISO** | single-input single-output |
| **SVD** | singular value decomposition |
| **TDD** | time-division-duplex |
| **TDOA** | time-difference-of-arrival |
| **TLS** | total least squares |
| **TOA** | time-of-arrival |
| **UL** | uplink |
| **WSSUS** | wide-sense stationary uncorrelated scattering |

# Bibliography


[3GP14] 3GPP: *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 12)*. Technical Report, 3GPP, 2014.

[ABB+07] Almers, P.; Bonek, E.; Burr, A.; Czink, N.; Debbah, M.; Degli-Esposti, V.; Hofstetter, H.; Kyösti, P.; Laurenson, D.; Matz, G.; Molisch, A. F.; Oestges, C.; Özcelik, H.: Survey of channel and radio propagation models for wireless MIMO systems. *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, 2007, pp. 1–19.

[AJJF99] Andersen, J. B.; Jensen, J.; Jensen, S. H.; Frederiksen, F.: Prediction of future fading based on past measurements. *Proc. IEEE 50th Vehicular Technology Conference (VTC'99-Fall)*, 1999, pp. 151–155.

[Ala98] Alamouti, S. M.: A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, vol. 16, 1998, pp. 1451–1458.

[AM08] Ananthasubramaniam, B.; Madhow, U.: Cooperative localization using angle of arrival measurements in non-line-of-sight environments. *Proceedings of the first ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, 2008, pp. 117–122.

[AMGC02] Arulampalam, M. S.; Maskell, S.; Gordon, N.; Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, vol. 50, 2002, pp. 174–188.

[Aro11] Aronsson, D.: *Channel Estimation and Prediction for MIMO OFDM Systems: Key Design and Performance Aspects of Kalman-based Algorithms*. PhD thesis, Uppsala Universitet, 2011.

[ATD14a] Adeogun, R.; Teal, P.; Dmochowski, P.: Extrapolation of MIMO mobile-to-mobile wireless channels using parametric model based prediction. *IEEE Transactions on Vehicular Technology*, vol. PP, 2014, pp. 1–12.

[ATD14b] Adeogun, R.; Teal, P.; Dmochowski, P.: Parametric channel prediction for narrowband MIMO systems using polarized antenna arrays. *IEEE 79th Vehicular Technology Conference (VTC'14-Spring)*, 2014, pp. 1–5.

[AVWS12] Akl, R.; Valentin, S.; Wunder, G.; Stanczak, S.: Compensating for CQI aging by channel prediction: The LTE downlink. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'12)*, 2012, pp. 4821–4827.

[BBT06] Beck, A.; Ben-Tal, A.: On the solution of the Tikhonov regularization of the total least squares problem. *SIAM Journal on Optimization*, vol. 17, 2006, pp. 98–118.

[Bel63] Bello, P. A.: Characterization of randomly time-variant linear channels. *IEEE Transactions on Communications Systems*, vol. 11, 1963, pp. 360–393.

[BKR97] Blankenship, T.; Kriztman, D.; Rappaport, T.: Measurements and simulation of radio frequency impulsive noise in hospitals and clinics. *IEEE Vehicular Technology Conference (VTC'97)*, vol. 3, 1997, pp. 1942–1946 vol.3.

[BP00] Bahl, P.; Padmanabhan, V. N.: RADAR: An in-building RF-based user location and tracking system. *Proc. IEEE Conference on Computer Communications (INFOCOM'00)*, vol. 2, 2000, pp. 775–784.

[BSKL01] Bar-Shalom, Y.; Kirubarajan, T.; Li, X.-R.: *Estimation with Applications to Tracking and Navigation*. New York, N.Y., USA: John Wiley & Sons, Inc., 2001.

[Cav91] Cavers, J.: An analysis of pilot symbol assisted modulation for Rayleigh fading channels. *IEEE Transactions on Vehicular Technology*, vol. 40, 1991, pp. 686–693.

[CEV07] Chen, M.; Ekman, T.; Viberg, M.: New approaches for channel prediction based on sinusoidal modeling. *EURASIP Journal on Advances in Signal Processing*, vol. 2007, 2007, pp. 1–13.

[CH94] Chan, Y. T.; Ho, K. C.: A simple and efficient estimator for hyperbolic location. *IEEE Transactions on Signal Processing*, vol. 42, 1994, pp. 1905–1915.

[Che03] Chen, Z.: *Bayesian filtering: From Kalman filters to particle filters, and beyond*. Technical Report, McMaster University, 2003.

[CSMC04] Cheung, K.; So, H.; Ma, W.-K.; Chan, Y. T.: Least squares algorithms for time-of-arrival-based mobile location. *IEEE Transactions on Signal Processing*, vol. 52, 2004, pp. 1121–1130.

[CZ04] Cong, L.; Zhuang, W.: Non-line-of-sight error mitigation in mobile location. *IEEE Conference on Computer Communications (INFOCOM'04)*, vol. 1, 2004, pp. 650–659.

[DAB+13] Dammann, A.; Agapiou, G.; Bastos, J.; Brunelk, L.; Garcia, M.; Guillet, J.; Ma, Y.; Ma, J.; Nielsen, J. J.; Ping, L.; Raulefs, R.; Rodriguez, J.; Slock, D. T. M.; Yang, D.; Yi, N.: WHERE2 Location aided communications. *Proc. 19th European Wireless Conference (EW'13)*, Guildford, United Kingdom, 2013, pp. 1–8.

[DH07] Duel-Hallen, A.: Fading channel prediction for mobile radio adaptive transmission systems. *Proceedings of the IEEE*, vol. 95, 2007, pp. 2299–2313.

[DHHH00] Duel-Hallen, A.; Hu, S.; Hallen, H.: Long-range prediction of fading signals: Enabling adaptive transmission for mobile radio channels. *IEEE Signal Processing Magazine*, vol. 17, 2000, pp. 62–75.

[ECS+98] Ertel, R.; Cardieri, P.; Sowerby, K.; Rappaport, T.; Reed, J.: Overview of spatial channel models for antenna array communication systems. *IEEE Personal Communications*, vol. 5, 1998, pp. 10–22.

[EDHH98] Eyceoz, T.; Duel-Hallen, A.; Hallen, H.: Deterministic channel modeling and long range prediction of fast fading mobile radio channels. *IEEE Communications Letter*, vol. 2, 1998, pp. 254–256.

[EK99] Ekman, T.; Kubin, G.: Nonlinear prediction of mobile radio channels: measurements and MARS model designs. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, vol. 5, 1999, pp. 2667–2670.

[Ekm02] Ekman, T.: *Prediction of Mobile Radio Channels: Modeling and Design.* PhD thesis, Uppsala University, 2002.

[ER10] Ekambaram, V.; Ramchandran, K.: Distributed high accuracy peer-to-peer localization in mobile multipath environments. *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, Miami, 2010, pp. 1–5.

[FG98] Foschini, G.; Gans, M.: On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, vol. 6, 1998, pp. 311–335.

[FGHO97] Fierro, R. D.; Golub, G. H.; Hansen, P. C.; O'leary, D. P.: Regularization by truncated total least squares. *SIAM J. Sci. Comp*, vol. 18, 1997, pp. 1223–1241.

[FH94] Fessler, J. A.; Hero, A. O.: Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, vol. 42, 1994, pp. 2664–2677.

[GAMD04] Grosicki, E.; Abed-Meriam, K.; Dehak, R.: A novel method to fight the non-line-of-sight error in AOA measurements for mobile location. *IEEE International Conference on Communications*, vol. 5, 2004, pp. 2794–2798.

[GC09] Güvenç, I.; Chong, C.: A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Communications Surveys & Tutorials*, vol. 11, 2009, pp. 107–124.

[GKN+74] Gelb, A.; Kasper, J. F.; Nash, R. A.; Price, C. F.; Sutherland, A. A. (Eds.): *Applied Optimal Estimation.* Cambridge, MA: MIT Press, 1974.

[God97] Godara, L. C.: Applications of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, vol. 85, 1997, pp. 1195–1245.

[Gol05] Goldsmith, A.: *Wireless Communications.* New York, NY, USA: Cambridge University Press, 2005.

[GS04] Guillaud, M.; Slock, D.: A specular approach to MIMO frequency-selective channel tracking and prediction. *Proc. 7th Workshop on Signal Processing Advances in Wireless Communications (SPAWC'04)*, 2004, pp. 59–63.

[GVL80] Golub, G. H.; Van Loan, C.: An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, vol. 17, 1980, pp. 883–893.

[GVL13] Golub, G. H.; Van Loan, C. F.: *Matrix Computations*. 4th edition. Baltimore, USA: Johns Hopkins University Press, 2013.

[Hat80] Hata, M.: Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, vol. 29, 1980, pp. 317–325.

[Hay02] Haykin, S.: *Adaptive Filter Theory*. 4th edition. Upper Saddle River: Prentice Hall, 2002.

[HHDH99] Hu, S.; Hallen, H.; Duel-Hallen, A.: Physical channel modeling, adaptive prediction and transmitter diversity for flat fading mobile channel. *Proc. 2nd Workshop on Signal Processing Advances in Wireless Communications (SPAWC'99)*, 1999, pp. 387–390.

[HRS+99] Hampicke, D.; Richter, A.; Schneider, A.; Sommerkorn, G.; Thomä, R. S.: Statistical analysis of time-variant directional mobile radio channels, based on wide-band measurements. *Proc. European Wireless (EW'99)*, 1999, pp. 273–278.

[Hub64] Huber, P. J.: Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, vol. 35, 1964, pp. 73–101.

[HW98] Hwang, J.-K.; Winters, J. H.: Sinusoidal modeling and prediction of fast fading processes. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'05)*, 1998, pp. 892–897.

[Jak74] Jakes, W.: *Microwave Mobile Communications*. Wiley-IEEE Press, 1974.

[JStB01] Jiang, T.; Sidiropoulos, N.; ten Berge, J.: Almost-sure identifiability of multidimensional harmonic retrieval. *IEEE Transactions on Signal Processing*, vol. 49, 2001, pp. 1849–1859.

[Kal60] Kalman, R. E.: A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, 1960, pp. 35–45.

[Kay93] Kay, S. M.: *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River: Prentice Hall, 1993.

[KBF+05] Kaiser, T. Bourdoux, A.; Boche, H.; Fonollosa, J. R.; Andersen, J. B.; Utschick, W. (Eds.): *Smart Antennas – State of the Art*. Hindawi, 2005.

[KH06] Kaplan, E. D.; Hegarty, C.: *Understanding GPS: Principles and Applications*. 2nd edition. Boston: Artech House, 2006.

[LC98] Lehmann, E.; Casella, G.: *Theory of Point Estimation*. Springer Texts in Statistics. Springer, 1998.

[LGLG04] Luo, Z.; Gao, H.; Liu, Y.; Gao, J.: Robust pilot-symbol-aided MIMO channel estimation and prediction. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'04)*, vol. 6, 2004, pp. 3646–3650.

[LHSH04] Love, D.; Heath, R.; Santipach, W.; Honig, M.: What is the value of limited feedback for MIMO channels? *IEEE Communications Magazine*, vol. 42, 2004, pp. 54–59.

[LJ05] Lin, D.-B.; Juang, R.-T.: Mobile location estimation based on differences of signal attenuations for GSM systems. *IEEE Transactions on Vehicular Technology*, vol. 54, 2005, pp. 1447–1454.

[LR96] Liberti, J. C.; Rappaport, T. S.: A geometrically based model for line-of-sight multipath channels. *Proc. IEEE 46th Vehicular Technology Conference (VTC'96)*, vol. 2, Atlanta, 1996, pp. 844–848.

[LR99] Liberti, J. C.; Rappaport, T. S.: *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Upper Saddle River: Prentice Hall, 1999.

[LSS09] Larsen, M.; Swindlehurst, A.; Svantesson, T.: Performance bounds for MIMO-OFDM channel estimation. *IEEE Transactions on Signal Processing*, vol. 57, 2009, pp. 1901–1916.

[LTZN05] Li, Z.; Trappe, W.; Zhang, Y.; Nath, B.: Robust statistical methods for securing wireless localization in sensor networks. *International Symposium on Information Processing in Sensor Networks (IPSN'05)*, 2005, pp. 91–98.

[Lüt07] Lütkepohl, H.: *New Introduction to Multiple Time Series Analysis*. Springer, 2007.

[Mak75] Makhoul, J.: Linear prediction: A tutorial review. *Proceedings of the IEEE*, vol. 63, 1975, pp. 561–580.

[MBW+04] Meurer, M.; Baier, P. W.; Weber, T.; Jötten, C. A.; Heilmann, S.: SPIDER: Enhanced distance based localization of mobile radio terminals. *Proc. IEEE 60th Vehicular Technology Conference (VTC'04-Fall)*, vol. 5, Los Angeles, 2004, pp. 3521–3525.

[MCCK07] Min, C.; Chang, N.; Cha, J.; Kang, J.: MIMO-OFDM downlink channel prediction for IEEE802.16e systems using Kalman filter. *Proc. IEEE Wireless Communications and Networking Conference (WCNC'07)*, 2007, pp. 942–946.

[Mid99] Middleton, D.: Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models. *IEEE Transactions on Information Theory*, vol. 45, 1999, pp. 1129–1149.

[MMY06] Maronna, R. A.; Martin, R. D.; Yohai, V. J.: *Robust Statistics: Theory and Methods*. John Wiley & Sons, 2006.

[Mol04] Molisch, A. F.: A generic model for MIMO wireless propagation channels in macro- and microcells. *IEEE Transactions on Signal Processing*, vol. 52, 2004, pp. 61–71.

[Mol05] Molisch, A. F.: *Wireless Communications*. 2nd edition. John Wiley & Sons, 2005.

[MS00] Moon, T. K.; Stirling, W. C.: *Mathematical methods and algorithms for signal processing*. Upper Saddle River, NJ: Prentice Hall, 2000.

[MSL+09] Mietzner, J.; Schober, R.; Lampe, L.; Gerstacker, W.; Hoeher, P.: Multiple-antenna techniques for wireless communications - a comprehensive literature survey. *Communications Surveys Tutorials, IEEE*, vol. 11, 2009, pp. 87–105.

[MVH07] Markovsky, I.; Van Huffel, S.: Overview of total least squares methods. *Signal Processing*, vol. 87, 2007, pp. 2283–2302. Special Section: Total Least Squares and Errors-in-Variables Modeling.

[MYJ07] Miao, H.; Yu, K.; Juntti, M. J.: Positioning for NLOS propagation: Algorithm derivations and Cramér-Rao bounds. *IEEE Transactions on Vehicular Technology*, vol. 56, 2007, pp. 2568–2580.

[ONJ+08] Okino, K.; Nakayama, T.; Joko, S.; Kusano, Y.; Kimura, S.: Direction based beamspace MIMO channel prediction with ray cancelling. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'08)*, 2008, pp. 1–5.

[OOF68] Okumura, Y.; Omori, E.; Fukuda, K.: Field strength and its variability in VHF and UHF land mobile service. *Review Electrical Communication Laboratory*, vol. 16, 1968, pp. 825–873.

[Pal11] Palleit, N.: *Kanalprädiktion in Mehrantennensystemen*. PhD thesis, University of Rostock, 2011.

[Pät02] Pätzold, M.: *Mobile Fading Channels*. John Wiley & Sons, 2002.

[PP12] Petersen, K. B.; Pedersen, M. S.: *The Matrix Cookbook*. Technical University of Denmark, 2012.

[Pro95] Proakis, J. G.: *Digital Communications*. 3. edition. New York: McGraw-Hill, 1995.

[PS09] Papakonstantinou, K.; Slock, D.: Hybrid TOA/AOD/Doppler-shift localization algorithm for NLOS environments. *IEEE 20th Personal, Indoor and Mobile Radio Communications Symposium (PIMRC'09)*, 2009, pp. 1948–1952.

[PW09a] Palleit, N.; Weber, T.: Obtaining transmitter side channel state information in MIMO FDD systems. *Proc. IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'09)*, 2009, pp. 2439–2443.

[PW09b] Palleit, N.; Weber, T.: Prediction of the channel transfer function. *Workshop Selbstorganisierende Sensor- und Datenfunknetze*, 2009, pp. 1–4.

[QKS06] Qi, Y.; Kobayashi, H.; Suda, H.: Analysis of wireless geolocation in a non-line-of-sight environment. *IEEE Transactions on Wireless Communications*, vol. 5, 2006, pp. 672–681.

[RK89] Roy, R.; Kailath, T.: ESPRIT – estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, 1989, pp. 984–995.

[Rot71] Rothenberg, T. J.: Identification in parametric models. *Econometrica*, vol. 39, 1971, pp. 577–591.

[SAK+92] Sakagami, S.; Aoyama, S.; Kuboi, K.; Shirota, S.; Akeyama, A.: Vehicle position estimates by multibeam antennas in multipath environments. *IEEE Transactions on Vehicular Technology*, vol. 41, 1992, pp. 63–68.

[SCC+14] Shi, Y.; Cui, Q.; Cao, S.; Zhang, X.; Tao, X.: Performance relationship between distributed and centralised cooperative localisations. *Electronics Letters*, vol. 50, 2014, pp. 127–128.

[Sch86] Schmidt, R. O.: Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, vol. 34, 1986, pp. 276–280.

[Ser93] Sernesi, E.: *Linear algebra. A geometric approach. Transl. by J. Montaldi.* London: Chapman & Hall, 1993.

[SFW12] Shikur, B. Y.; Farmani, M.; Weber, T.: AOD/AOA/TOA-based 3-D mobile terminal tracking in NLOS multipath environments. *Proc. 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12)*, Dresden, 2012, pp. 201–205.

[SG04] Sun, G.; Guo, W.: Bootstrapping $M$-estimators for reducing errors due to non-line-of-sight (NLOS) propagation. *IEEE Communications Letters*, vol. 8, 2004, pp. 509–510.

[SJJS00] Spencer, Q.; Jeffs, B.; Jensen, M.; Swindlehurst, A.: Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel. *IEEE Journal on Selected Areas in Communications*, vol. 18, 2000, pp. 347–360.

[SK02] Semmelrodt, S.; Kattenbach, R.: A 2-D fading forecast of time-variant channels based on parametric modeling techniques. *Proc. 13th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'02)*, 2002, pp. 1640–1644.

[Slo12] Slock, D. T. M.: Location aided wireless communications. *Proc. 5th International Symposium on Communications Control and Signal Processing (ISCCSP'12)*, 2012, pp. 1–6.

[SM05] Schafhuber, D.; Matz, G.: MMSE and adaptive prediction of time-varying channels for OFDM systems. *IEEE Transactions on Wireless Communications*, vol. 4, 2005, pp. 593–602.

[SMB01] Steinbauer, M.; Molisch, A.; Bonek, E.: The double-directional radio channel. *IEEE Antennas and Propagation Magazine*, vol. 43, 2001, pp. 51–63.

[SS06] Svantesson, T.; Swindlehurst, A.: A performance bound for prediction of MIMO channels. *IEEE Transactions on Signal Processing*, vol. 54, 2006, pp. 520–529.

[ST08] Seow, C. K.; Tan, S. Y.: Non-line-of-sight localization in multipath environments. *IEEE Transactions on Mobile Computing*, vol. 7, 2008, pp. 647–660.

[Ste01] Steinbauer, M.: *The Radio Propagation Channel – A Non-Directional, Directional, and Double-Directional Point-of-View*. PhD thesis, Technische Universität Wien, 2001.

[SVHG04] Sima, D. M.; Van Huffel, S.; Golub, G. H.: Regularized total least squares based on quadratic eigenvalue problem solvers. *BIT Numerical Mathematics*, vol. 44, 2004, pp. 793–812.

[SW12] Shikur, B. Y.; Weber, T.: Discrete position processing techniques for indoor NLOS localization. *The 9th International Symposium on Wireless Communication Systems (ISWCS12)*, Paris, 2012, pp. 631–635.

[SW13] Shikur, B. Y.; Weber, T.: Posterior CRLB for tracking a mobile station in NLOS multipath environments. *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'13)*, Vancouver, 2013, pp. 5175–5179.

[SW14a] Shikur, B. Y.; Weber, T.: Localization in NLOS environments using TOA, AOD, and Doppler-shift. *11th Workshop on Positioning Navigation and Communication (WPNC'14)*, Dresden, 2014, pp. 1–6.

[SW14b] Shikur, B. Y.; Weber, T.: Robust cooperative localization in mixed LOS and NLOS environments using TOA. *11th Workshop on Positioning Navigation and Communication (WPNC'14)*, Dresden, 2014, pp. 1–6.

[SW14c] Shikur, B. Y.; Weber, T.: TDOA/AOD/AOA localization in NLOS environments. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14)*, Florence, 2014, pp. 6518–6522.

[SW14d] Shikur, B. Y.; Weber, T.: Uplink-downlink transformation of the channel transfer function for OFDM systems. *18th International OFDM Workshop (InOWo'14)*, Essen, 2014, pp. 185–190.

[SW15a] Shikur, B. Y.; Weber, T.: Channel prediction using an adaptive Kalman filter. *International ITG Workshop on Smart Antennas (WSA'15)*, Ilmenau, 2015.

[SW15b] Shikur, B. Y.; Weber, T.: Uplink-downlink channel transformation using an adaptive Kalman filter for multicarrier systems. *16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC'15)*, Stockholm, 2015, pp. 465–469.

[SY03] Sayed, A. H.; Yousef, N. R.: Wireless location. Proakis, J. (Ed.): *Wiley Encyclopedia of Telecommunications*. New York: John Wiley & Sons, 2003.

[TA77] Tikhonov, A. N.; Arsenin, V. Y.: *Solutions of Ill-posed problems*. W.H. Winston, 1977.

[TCL01] Thomas, N. J.; Cruickshank, D. G. M.; Laurenson, D. I.: Calculation of mobile location using scatterer information. *IEE Electronics Letters*, vol. 37, 2001, pp. 1193–1194.

[Tel99] Telatar, E.: Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, vol. 10, 1999, pp. 585–595.

[TMN98] Tichavsky, P.; Muravchik, C.; Nehorai, A.: Posterior Cramer-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, vol. 46, 1998, pp. 1386–1396.

[TO05] Takei, Y.; Ohtsuki, T.: Throughput maximization transmission control scheme using channel prediction for MIMO systems. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'05)*, vol. 4, 2005, pp. 2344–2348.

[TQ08] Tang, H. Park, Y.; Qiu, T.: A TOA-AOA-based NLOS error mitigation method for location estimation. *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008, pp. 1–14.

[TSC98] Tarokh, V.; Seshadri, N.; Calderbank, A. R.: Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Transactions on Information Theory*, vol. 44, 1998, pp. 744–765.

[TV01] Teal, P.; Vaughan, R.: Simulation and performance bounds for real-time prediction of the mobile multipath channel. *IEEE Signal Processing Workshop on Statistical Signal Processing*, 2001, pp. 548–551.

[VB12] Vaghefi, R.; Buehrer, R.: Cooperative sensor localization with NLOS mitigation using semidefinite programming. *Proc. 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12)*, Dresden, 2012, pp. 13–18.

[VHV91] Van Huffel, S.; Vandewalle, J.: *The Total Least Squares Problem*. Philadelphia: Society for Industrial and Applied Mathematics, 1991.

[VT68] Van Trees, H. L.: *Detection, Estimation, and Modulation Theory: Part I.* Wiley, 1968.

[VT02] Van Trees, H. L.: *Optimum Array Processing Part IV of Detection, Estimation, and Modulation Theory.* John Wiley & Sons, 2002.

[VTR00] Vaughan, R.; Teal, P.; Raich, R.: Short-term mobile channel prediction using discrete scatterer propagation model and subspace signal processing algorithms. *Proc. IEEE 52nd Vehicular Technology Conference (VTC'00-Fall)*, vol. 2, 2000, pp. 751–758.

[WABD06] Weber, T.; Ahrens, A.; Buch, T.; Deng, S.: Exploiting second order statistics in MIMO channel estimation. *Proc. 2006 IST Summit*, Myconos, 2006.

[WE05] Wong, I.; Evans, B.: Joint channel estimation and prediction for OFDM systems. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'05)*, vol. 4, 2005, pp. 2255–2259.

[WE06a] Wong, I.; Evans, B.: Exploiting spatio-temporal correlations in MIMO wireless channel prediction. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'06)*, 2006, pp. 1–5.

[WE06b] Wong, I.; Evans, B.: Low-complexity adaptive high-resolution channel prediction for OFDM systems. *Proc. IEEE Global Telecommunications Conference (GLOBECOM'06)*, 2006, pp. 1–5.

[WLW09] Wymeersch, H.; Lien, J.; Win, M.: Cooperative localization in wireless networks. *Proceedings of the IEEE*, vol. 97, 2009, pp. 427–450.

[WM05] Weber, T.; Meurer, M.: Temporal post-processing of position estimates exploiting Markoff models. *Proc. Int. Symposium on personal, Indoor and Mobile Radio Communications (PIMRC'05)*, vol. 3, Berlin, 2005, pp. 1979–1983.

[WPW11] Wei, X.; Palleit, N.; Weber, T.: AOD/AOA/TOA-based 3D Positioning in NLOS Multipath Environments. *Proc. 22nd IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'11)*, Toronto, Canada, 2011, pp. 1294–1298.

[WPY07] Wang, C.-X.; Pätzold, M.; Ya0, Q.: Stochastic modeling and simulation of frequency-correlated wideband fading channels. *IEEE Transactions on Vehicular Technology*, vol. 56, 2007, pp. 1050–1063.

[Xio98] Xiong, L.: A selective model to suppress NLOS signals in angle-of-arrival (AOA) location estimation. *The Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'98)*, vol. 1, 1998, pp. 461–465.

[YES09] Youssef, M.; El-Sheimy, N.: Robust cooperative localization technique for wireless sensor networks. *International Conference on New Technologies, Mobility and Security (NTMS'09)*, 2009, pp. 1–4.

[YO02] Yu, K.; Ottersten, B.: Models for MIMO propagation channels: a review. *Wireless Communications and Mobile Computing*, vol. 2, 2002, pp. 653–666.

[ZG04] Zhou, S.; Giannakis, G.: How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels? *IEEE Transactions on Wireless Communications*, vol. 3, 2004, pp. 1285–1294.

[ZKCM12] Zoubir, A.; Koivunen, V.; Chakhchoukh, Y.; Muma, M.: Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts. *IEEE Signal Processing Magazine*, vol. 29, 2012, pp. 61–80.