TECHNISCHE UNIVERSITÄT MÜNCHEN
Fachgebiet Methoden der Signalverarbeitung

Location Determination in OFDM Based Mobile Radio Systems

Christian Mensing

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. techn. Dr. h. c. Josef A. Nossek
Prüfer der Dissertation:
1. Univ.-Prof. Dr.-Ing. Wolfgang Utschick
2. Prof. Dr. sc. techn. Bernard H. Fleury,
Aalborg University, Dänemark

Die Dissertation wurde am 14. Juni 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 17. Mai 2013 angenommen.

Acknowledgments

First of all, I would like to thank Univ.-Prof. Dr.-Ing. Wolfgang Utschick from the Munich University of Technology for supervising this thesis and for his encouraging and patient support during the past years. Furthermore, I would like to thank Prof. Dr. sc. techn. Bernard H. Fleury from Aalborg University for acting as co-examiner of this thesis.

I would like to thank all colleagues from the Department of Communications Systems at the Institute of Communications and Navigation of the German Aerospace Center (DLR) in Oberpfaffenhofen. I thank Prof. Dr.-Ing. Uwe-Carsten Fiebig for the opportunity to work as research assistant at his department. I am grateful to all members of DLR's Mobile Radio Transmission Group, especially Dr. Armin Dammann, Dr. Simon Plass, Dr. Ronald Raulefs, Dr. Stephan Sand, and Wei Wang, for their daily assistance and the illuminating discussions.

I thank my parents Theodor and Sabine Mensing, and my brother Stefan Mensing. They gave me the needful support from the family side for preparing this thesis. Finally, I would like to thank Dr. Hanne Borchmeyer for her encouragement during this endeavor.

Abstract

In this thesis, location determination algorithms for cellular mobile radio systems employing orthogonal frequency division multiplexing (OFDM) are investigated. For location estimation, usually timing measurements between the base stations and the mobile station are performed, where suitable algorithms for OFDM and corresponding performance bounds are derived. The timing estimation process is faced with different effects, e.g., limited number of pilot symbols or intercell interference, that decrease the accuracy of timing estimates, and hence, location estimates. Therefore, a feedback of decided data symbols is proposed, where particularly data-aided timing estimation and interference cancelation are considered. Simulation results for a Long Term Evolution (LTE) system show the ability of these approaches to improve the accuracy and reliability of location estimates for static and dynamic mobile stations.

Kurzzusammenfassung

In dieser Arbeit werden Algorithmen zur Positionsbestimmung in zellularen Mobilfunksystemen, die Orthogonal Frequency Division Multiplexing (OFDM) verwenden, untersucht. Zur Positionsschätzung werden üblicherweise Zeitmessungen zwischen den Basisstationen und der Mobilstation durchgeführt, wobei geeignete Algorithmen für OFDM und entsprechende Schranken hergeleitet werden. Verschiedene Effekte reduzieren dabei die Leistungsfähigkeit der Zeit- und damit Positionsschätzung, insbesondere die limitierte Anzahl von Pilotsymbolen und interzellulare Interferenz. Deshalb werden Ansätze mit Datenrückkopplung vorgeschlagen, die Verfahren für datenbasierte Zeitschätzung und Interferenzbeseitigung ermöglichen. Simulationsergebnisse für ein Long Term Evolution (LTE) System zeigen, daß diese Verfahren die Genauigkeit und Zuverlässigkeit der Positionsschätzung für statische und dynamische Mobilstationen verbessern.

Contents

1 Introduction

1.1 Location determination in mobile radio systems

Provision and exploitation of location information became very important features of mobile radio systems in recent years [1]. Moreover, services and applications based on accurate knowledge of the mobile station (MS) location will play a fundamental role in future wireless systems. Point-to-point navigation, fraud detection, or automated billing are well-known examples for these location based services on the user side [2]. Additionally, applications on the system side, e.g., location based mobility or radio resource management [3], can be further enablers for high precise positioning. Finally, regulatory requirements force service providers to implement location estimation enhancements in their network deployments. For instance, the United States Federal Communications Commission (FCC) has stated accuracy requirements on the location estimation process of enhanced 911 (E911) emergency callers [4]. Corresponding requirements for Europe are currently under development [5].

Global navigation satellite systems (GNSSs) like the Global Positioning System (GPS) [6] or the upcoming European Galileo system [7] are well-suited to meet such requirements. They deliver accurate location estimates under optimum conditions, where — in general — GNSS based location estimation needs access to at least four satellites. However, especially in case of severe multipath propagation or blocking of the line of sight (LOS) access to satellites, the performance loss can be unacceptable high [8]. Techniques like assisted GNSS (AGNSS) [9] can provide a faster and more reliable access to GNSS services and reduce the — still critical — power consumption. Nevertheless, a seamless outdoor and indoor coverage can not be realized by pure GNSS receivers today.

Therefore, it is reasonable to exploit an already available communications infrastructure for the estimation of the MS location [1]. Base stations (BSs) of mobile radio systems provide a high coverage especially in GNSS-critical environments like urban canyons or inside buildings. They can be used to provide location information as supplementation to GNSS or as stand-alone solution. As depicted in Figure 1.1, this location determination process is usually implemented in a

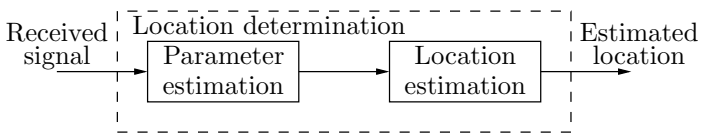


Figure 1.1: Two-step location determination process

two-step procedure (e.g., [2, 10]). In a first step, location dependent parameters are extracted from the received signal. In a second step — by using these estimates — the location is determined.

The choice of suitable location dependent parameters is a trade-off between desired accuracy and permitted complexity of the location enabling system. Some commonly used examples are parameters based on measurements in terms of time of arrival (TOA), time difference of arrival (TDOA), received signal strength (RSS), or angle of arrival (AOA) [10, 11]. They can be provided either by the MS in the downlink mode, using signals transmitted from the BSs exploited at the MS, or by the network in the uplink mode, using signals received from the MS.

TOA uses measurements of the signal propagation delay in order to calculate the distance between the BSs and the MS. It requires that all devices are synchronized or share the same clock. If MS and BSs have different clocks, which is usually the case for mobile radio systems or GNSSs, the unknown time offset between the clocks has to be estimated as additional parameter. In TDOA, time differences of arrival are estimated. Compared to TOA, the timing offset between MS and BSs is eliminated by a difference operation. Both TOA and TDOA require that the BSs are synchronized or the timing offsets between the BSs are known. With RSS measurements the estimated received power level is compared with the transmit power to derive distance information using adequate path loss models. AOA measurements require multiple antenna elements at BS or MS to obtain angle information about the incident waves.

From a theoretical point of view, timing based approaches — which are closely related to the timing synchronization problem — provide the most accurate location estimates. The corresponding Cramer-Rao lower bound (CRLB) for the estimation error [12] solely depends on the signal-to-noise ratio (SNR) at the receiver and the transmission bandwidth of the signal [10]. Nevertheless, in practical systems various propagation and system effects make it difficult to achieve this bound with computational efficient signal processing techniques. Well-known performance limiting factors in the context of location estimation are, e.g., multipath or non line of sight (NLOS) propagation. Timing estimates are affected with a bias under these conditions. Furthermore, communications networks are designed in a way that only one serving BS should be received with sufficient quality due to spectral efficiency demands. All other BSs are interfering the desired signal. Contrary to that, for location determination more than one BS has to be received with sufficient quality, e.g., for timing based approaches at least three BSs are required. Hence, aspects like pilot design and frequency re-use have a direct impact on the location estimation performance.

Timing based location estimation is part of several standards. In particular, TDOA methods are included in the standards for 2nd generation (2G) and 3rd generation (3G) systems [9, 13]. Nevertheless, the parameters of these systems were originally designed for communications purposes. Location functionalities were not inherently included in that process. However, it turned out that the communications design was not sufficient for providing an acceptable location accuracy or coverage. Hence, additional features were added to the standards afterwards to improve the location estimation performance. One example is the extension of the 3G standards with idle periods (cf. [13]). They allow a better reception of other BSs especially close to the serving cell yielding an improved overall reception, and hence, increased location estimation quality.

Currently, TDOA methods are in the standardization process for beyond 3G systems like 3rd Generation Partnership Project (3GPP) Long-Term Evolution (LTE) [14] and under investigation for future 4th generation (4G) systems as proposed, e.g., by the project Wireless World Initiative New Radio (WINNER) [15]. These systems will employ orthogonal frequency division multiplexing (OFDM). OFDM is well analyzed and understood for communications (e.g., [16]), whereas this field is rather new for location estimation purposes, especially in the cellular context of mobile

radio systems. First investigations in LTE have shown that the standard's communications signals do not allow precise location estimation with state-of-the-art approaches [17] — similar as observed for 3G systems some years ago. In particular LTE's targeted frequency re-use of one limits the performance due to inter-cell interference. Hence, also here an extension of the standard for more precise location estimation will be realized as optional feature. A second opportunity is the development of enhanced receiver algorithms to cope with performance limiting effects.

Once the location dependent parameters are available (e.g., TDOAs), the location of the MS can be estimated (cf. Figure 1.1). We distinguish between static and dynamic location estimation. In static location estimation no further a priori information about previous estimates is available, hence, it is a snap-shot approach using only the parameters available at a specific time instance. Besides the location dependent parameters, also their reliability and other side-information (e.g., LOS/NLOS status) can be provided to the location estimation entity. Several approaches exist to solve the underlying problem which belongs to the class of non-linear optimization problems (e.g., [18]). For its solution standard numerical optimization algorithms can straightforwardly be applied, taking into account solely the resulting cost function. Other approaches exploit specifically the geometric aspects of the problem.

If information about previous estimates is available, dynamic location estimation can be applied. In this approach, it is assumed that the MS follows a certain path and can not jump over an arbitrary distance in limited time. These constraints can be integrated directly in the estimation problem in terms of movement or mobility models. Well-known methods for the solution of this estimation problem are based on Kalman [19, 12] or particle [20] filtering.

A further extension is the fusion of estimates from different sources. Besides mobile radio systems as discussed before, also GNSSs or short-range systems based on ultra-wideband (UWB) [21] or radio frequency identification (RFID) [22] can provide location dependent parameters. Another source are sensors like inertial sensors, magnetometers, or barometers, that can provide information about acceleration, orientation, or height (e.g., [23, 24]). Finally, context-aware information can be integrated in the location estimation process. Especially the knowledge of map information can be beneficial for resolving ambiguous location estimates (e.g., [25, 26]).

1.2 Contributions

In this thesis, we will analyze location determination in the context of cellular OFDM based mobile radio systems and assess the performance with realistic parameters considering LTE systems. Chapter 2 presents an introductory overview of general location estimation principles for cellular mobile radio systems. Since the focus will be on timing based approaches, in Chapter 3 synchronization and timing estimation algorithms for cellular OFDM systems will be analyzed. In particular, we will

- Derive multi-link synchronization algorithms for timing estimation in cellular OFDM systems taking into account receiver imperfections.

- Propose extensions of classical synchronization approaches in the location estimation context.

- Propose new data-aided synchronization algorithms including schemes with interference cancelation enabling iterative location estimation techniques.

- Derive performance bounds for timing estimation in cellular OFDM systems.
- Assess the timing estimation performance under realistic system and channel environments.

In Chapter 4 the gained timing estimates are used to estimate the location of the MS in the static case. In particular, we will

- Describe numerical methods and geometric approaches to solve the static location estimation problem and corresponding performance bounds.
- Assess the static location estimation performance with generic models and under realistic system and channel environments.

Chapter 5 will include a priori information from previous location estimates to improve the performance by using tracking algorithms and mobility models. In particular, we will

- Describe Bayesian approaches for tracking algorithms to solve the dynamic location estimation problem and corresponding performance bounds.
- Assess the dynamic location estimation performance with generic models and under realistic system, channel, and mobility models.

Finally, Chapter 6 shows as application scenario the fusion of measurements from mobile radio systems with measurements from GPS. In particular, we will

- Propose a fusion process for location estimation as extension of the algorithms from the previous chapters.
- Assess the location estimation performance using GPS and LTE.

Note that parts of these topics have already been investigated by the author in [27, 28, 29, 30, 31, 32, 33, 34, 35] concerning Chapter 3, in [36, 37, 38, 39] concerning Chapter 4, and in [17, 40, 41, 42] concerning Chapter 5 and Chapter 6.

1.3 Notation

Throughout this thesis, vectors and matrices are denoted by lower and upper case bold letters. The matrix $\mathbf{I}_n$ is the $n \times n$ identity matrix, the matrix $\mathbf{0}_{n\times m}$ is the $n \times m$ matrix with zeros, and the matrix $\mathbf{1}_{n\times m}$ is the $n \times m$ matrix with ones. The operation $\otimes$ denotes the Kronecker product, $(\cdot)^*$ conjugate, $(\cdot)^{\mathrm{T}}$ transpose, $(\cdot)^{\mathrm{H}}$ Hermitian, $\mathrm{E}\{\cdot\}$ expectation, $\|\cdot\|_2$ the Euclidean norm, and $\lfloor\cdot\rfloor$ rounding towards minus infinity. The modulo operation between n and m is denoted as $a \bmod b$. With $\left[\boldsymbol{A}\right]_{n,m}$ we denote the element in the n-th row and m-th column of matrix $\boldsymbol{A}$.

2 Location estimation principles

In this chapter, we will describe commonly used location estimation principles and their advantages and constraints. We will start with timing based approaches measuring the propagation time between BS and MS. From that parameters simply distance or range information can be derived. Another discussed method will consider signal strength measurements of the received signals resulting in distance or range information by application of adequate pathloss models. Angle information is an additional relevant location dependent parameter, however, multiple antenna elements are required at BS or MS. A comprehensive overview of these principles can also be found in [10], [11], and [18].

Generally, the location dependent parameters can be estimated at BS or MS resulting in network based or MS based location determination, respectively (cf. [43]). In network based location determination solely uplink measurements from the MS to the BSs are exploited for deriving location information. In MS based location determination downlink measurements from the BSs to the MS will be used to determine its location. Additionally, side information (e.g., locations of the BSs) have to be further communicated in this mode. Moreover, a combination of both approaches is possible, where the relevant parameters are estimated at the MS. Then, this information is communicated to the network, where finally the location is determined. This is usually denoted as MS assisted location determination.

System capabilities and constraints will usually not allow a deployment of all modes in the individual systems. For instance, broadcasting systems like Digital Video Broadcasting (DVB) or Digital Audio Broadcasting (DAB) are unidirectional. Therefore, they can only support MS based approaches, where the MS receives information in the downlink. Also GNSSs belong to this class of systems as there is no uplink connection from the MS to the satellites. Classical mobile radio systems like Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS), or LTE are bidirectional, and hence, signals are transmitted between BSs and MS in both directions, i.e., in downlink and uplink. Therefore, in these systems also network oriented approaches can be integrated.

The subsequently presented location estimation principles will hold in general for all three modes (MS based, network based, MS assisted), independently of the entity, where the location of the MS is finally determined.

2.1 Propagation time

Electromagnetic waves propagate with the speed of light c_0 in free space. Therefore, the propagation time is a useful metric with a linear relation between the propagation time of the signal and the distance between BS and MS. The type of synchronization in the overall network — usually, not

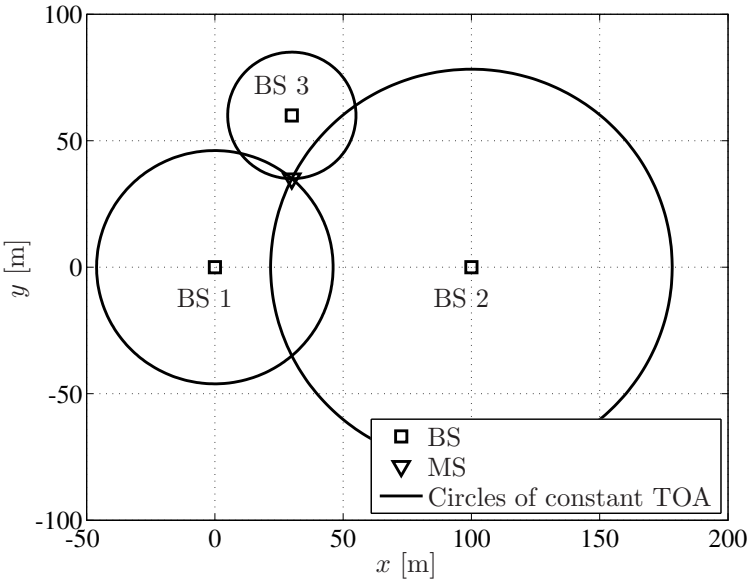


Figure 2.1: TOA location estimation principle

all entities share the same clock — requires a different processing of the resulting timing estimates. This will be outlined in the following subsections.

Time of arrival

TOA uses estimates of the signal propagation time to obtain the distance between BSs and MS. We denote the distance or range between BS μ located at (x_μ, y_μ, z_μ) and the MS located at (x, y, z) as

$$d_\mu = \sqrt{(x - x_\mu)^2 + (y - y_\mu)^2 + (z - z_\mu)^2}. \tag{2.1}$$

With the known transmission time of a signal at BS μ as $\tau_{\mu,0}$ and the arrival time of this signal at the MS as τ_μ, we easily obtain the relation

$$d_\mu = c_0 \left(\tau_\mu - \tau_{\mu,0}\right), \quad \mu = 1, 2, \ldots, N_{\text{BS}}, \tag{2.2}$$

for the considered N_{BS} BSs. If the BSs are perfectly synchronized and they are transmitting the relevant signals at the same time instance τ_0, this simplifies to

$$d_\mu = c_0 \left(\tau_\mu - \tau_0\right), \quad \mu = 1, 2, \ldots, N_{\text{BS}}. \tag{2.3}$$

Assuming LOS propagation, the distance d_μ determines points of equal distance from BS μ to MS. In the two-dimensional case this defines a circle around BS μ with radius d_μ. In the three-dimensional case it defines a sphere. For obtaining a unique location estimate, the distances to several BSs have to be determined. The intersection of the corresponding circles or spheres provides a unique location estimate of the MS. In Figure 2.1 the TOA principle for a two-dimensional example is shown. The intersection of two circles yields two possible solutions. Distance information to a third BS can resolve that ambiguity. Note that due to the geometric interpretation the TOA principle is often denoted as circular or spherical location estimation. Also the term range or distance based location estimation can be used.

With the N_{BS} TOA or distance estimates τ_μ we can form a system of N_{BS} non-linear equations with the unknown values (x, y, z) according to

$$\begin{aligned}
\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} &= c_0(\tau_1-\tau_0)\\
\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2} &= c_0(\tau_2-\tau_0)\\
&\vdots\\
\sqrt{(x-x_{N_{\mathrm{BS}}})^2+(y-y_{N_{\mathrm{BS}}})^2+(z-z_{N_{\mathrm{BS}}})^2} &= c_0(\tau_{N_{\mathrm{BS}}}-\tau_0),
\end{aligned} \tag{2.4}$$

where we assume that the locations of the BSs (x_μ, y_μ, z_μ) and the transmission time τ_0 are exactly known.

As already can be observed from the geometric interpretation of the TOA principle in Figure 2.1, the non-linear relation between the unknown values and the estimated TOAs results in the fact that more equations than unknown values are required to resolve ambiguous solutions. For two-dimensional processing at least three linear independent equations are required, whereas for three-dimensional processing we need at least four linear independent equations to get a unique location estimate. Furthermore, in practical systems the estimates of the TOAs or distances are noisy in general, i.e., the circles as depicted in Figure 2.1 do not intersect in one unique point. Therefore, the equation system (2.4) will not be suitable for providing a solution. Instead, the error ε_μ due to noisy estimates is included in the location determination process as object for minimization. Hence, it is a common procedure that in the restructured system of equations

$$(x-x_\mu)^2+(y-y_\mu)^2+(z-z_\mu)^2-c_0(\tau_\mu-\tau_0)=\varepsilon_\mu^2, \quad \mu=1,2,\ldots,N_{\mathrm{BS}}, \tag{2.5}$$

the overall quadratic error will be minimized, i.e., with

$$(\hat{x},\hat{y},\hat{z}) = \underset{(x,y,z)}{\operatorname{argmin}} \sum_{\mu=1}^{N_{\mathrm{BS}}} \varepsilon_\mu^2 \tag{2.6}$$

the location estimates are computed according to the least squares criterion (e.g., [12]).

The noise can be caused by manifold effects in the system and transmission chain. Thermal noise is usually modeled as a zero-mean additive white Gaussian distributed random process resulting in unbiased estimates of the TOA or distance measurements. Multipath propagation, where — besides the LOS path — reflected, refracted, or diffracted paths are received at the MS, results in biased estimates. Furthermore, if the LOS path is totally blocked, we measure a propagation time which is always larger than that of the direct LOS path. Hence, under NLOS propagation conditions the additional bias is always positive.

In Chapter 3 we will show in detail, how timing information can be gained in OFDM based mobile radio systems. Chapter 4 will then describe suitable algorithms for location estimation, i.e., approaches for solving the non-linear estimation problem in (2.6).

Time of arrival with unknown time offset

In TOA, the distances d_μ are calculated from the corresponding signal propagation times $\tau_\mu - \tau_0$, i.e., the time basis in all the BSs and in the MS must be the same. This requirement is hard to be met in practical systems. Certainly, on the network side the BSs can be synchronized by the network itself using wired connections or by integrating external entities like GPS receivers providing a

common time basis. Even if the network is not totally synchronized the asynchronism can be estimated and taken into account for the location estimation. For instance, in UMTS location measurement units (LMUs) are standardized for providing such information [44]. Nevertheless, the problem to have the same time basis for the MS still remains.

Therefore, a commonly used approach is to handle the time offset between the synchronized BS clocks and the MS clock as additional unknown value in the system of equations. We obtain

$$\begin{aligned}
\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} - c_0 b_{\text{clock}} &= c_0\tau_1 \\
\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2} - c_0 b_{\text{clock}} &= c_0\tau_2 \\
&\vdots \\
\sqrt{(x-x_{N_{\text{BS}}})^2+(y-y_{N_{\text{BS}}})^2+(z-z_{N_{\text{BS}}})^2} - c_0 b_{\text{clock}} &= c_0\tau_{N_{\text{BS}}},
\end{aligned} \tag{2.7}$$

but now, contrary to (2.4), the transmission time τ_0 is no longer assumed to be known and introduces an unknown clock offset or clock bias b_{clock} to all TOA measurements. This is the typical case in GNSSs, where the satellites are precisely synchronized by atomic clocks, whereas the time offset between satellites and MS has to be estimated explicitly. Hence, one more unknown has to be determined using an additional timing estimate. In general, for two-dimensional processing at least four linear independent equations are required, for three-dimensional processing we need at least five linear independent equations. These requirements can be weakened for certain systems or environments. For instance, in GNSSs the fact is exploited that the MS location is usually on the surface of the earth or at least close to it. With that side information unambiguous solutions can be resolved by only four available equations. Due to the clock offset or clock bias the measured propagation times do no longer reflect distance or ranges. Therefore, this method — especially in the GNSS community — is often denoted as location estimation with pseudo-ranges. Also the term circular (two-dimensional) or spherical (three-dimensional) location estimation with time offset or time bias can be used. The resulting optimization problem is similar to that for TOA and will be analyzed in detail in Chapter 4.

Time difference of arrival

Another method in systems, where BSs and MS do not share the same clock, relies on propagation time differences. Assuming that two signals are transmitted from two BSs $\mu = 1, 2, \ldots, N_{\text{BS}}$, and $\nu = 1, 2, \ldots, N_{\text{BS}}$, at the same time instant τ_0 assuming $\mu \neq \nu$. These signals arrive at the MS at the time instants τ_μ and τ_ν. Hence, the corresponding distance difference can directly be obtained from the TDOA

$$\tau_{\mu,\nu} = \tau_\mu - \tau_\nu \tag{2.8}$$

according to

$$d_{\mu,\nu} = d_\mu - d_\nu = c_0(\tau_\mu - \tau_0) - c_0(\tau_\nu - \tau_0) = c_0(\tau_\mu - \tau_\nu) = c_0\tau_{\mu,\nu}. \tag{2.9}$$

We observe, that the TDOAs are no longer depending on the absolute transmission time τ_0 and the only relevant time basis is that of the MS.

Contrary to TOA, where the propagation delays define circles around the BSs, a TDOA or distance difference estimate represents points of equal distance differences to the considered BSs. This defines a hyperbola in the two-dimensional case or a hyperboloid in the three-dimensional

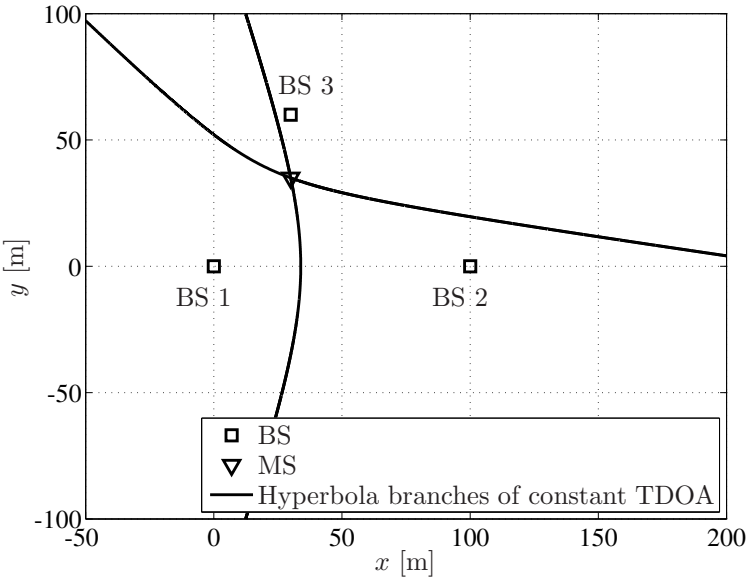


Figure 2.2: TDOA location estimation principle

case. Therefore, this method is often denoted as hyperbolic location estimation. Also the term range difference based or distance difference based location estimation is commonly used.

Figure 2.2 depicts the principle of TDOA based location estimation for a two-dimensional example. Two hyperbola branches (using the TDOAs $\tau_{2,1}$ and $\tau_{3,1}$) with foci at the locations of the involved BSs define points of equal distance difference to those BSs. The intersection of different hyperbola branches yields the MS location. A third possible hyperbola branch — using the TDOA $\tau_{3,2} = \tau_3 - \tau_2 = (\tau_3 - \tau_1) - (\tau_2 - \tau_1)$ — would directly be obtained from the other two measurements, and hence, would not provide any additional information. Usually, all TDOAs are computed with respect to one reference BS (without loss of any generality we choose BS 1). This provides $N_{\mathrm{BS}} - 1$ linear independent TDOAs or distance differences for N_{BS} available BSs. The resulting system of equations can be written as

$$\begin{aligned}
\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2} - \sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} &= c_0\tau_{2,1} \\
\sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2} - \sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} &= c_0\tau_{3,1} \\
&\vdots \\
\sqrt{(x-x_{N_{\mathrm{BS}}})^2+(y-y_{N_{\mathrm{BS}}})^2+(z-z_{N_{\mathrm{BS}}})^2} - \sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} &= c_0\tau_{N_{\mathrm{BS}},1}.
\end{aligned} \tag{2.10}$$

It is obvious that TDOA is closely related to the principle of TOA with unknown time offset and under certain assumptions (as it will be discussed in Chapter 4) both approaches yield the same results from the location estimation point of view. Nevertheless, in mobile radio systems usually the TDOA principle is favored. Under noisy estimates the resulting equation system can be solved using a least squares approach similar as shown for TOA before.

Round-trip time of arrival

The round-trip time of arrival (RTTOA) principle is sometimes also denoted as two-way ranging. The basic idea is, that the MS transmits a signal to a BS which acknowledges the reception by

transmitting another signal back to the MS. The overall round-trip time is estimated by the MS. Note that also the BS can trigger the overall process. It is obvious that — contrary to the approaches described before — measurements in both uplink and downlink have to be performed.

We assume that the RTTOA process between the MS and BS μ is initiated by the MS at time instance $\tau_{\text{init},\mu}$, where this is determined according to the internal time scale τ_{MS} of the MS. The signal is received at BS μ at time instance $\tau_{\text{rec},\mu}$. The internal time scales τ_{MS} and $\tau_{\text{BS},\mu}$ of the MS and BS μ are assumed to have a constant difference. The signal propagation distance between the MS and the BS can then be computed as

$$d_{\text{MS},\mu} = c_0 \left((\tau_{\text{rec},\mu} + \tau_{\text{BS},\mu} - \tau_{\text{MS}}) - \tau_{\text{init},\mu} \right), \tag{2.11}$$

where $\tau_{\text{rec},\mu} + \tau_{\text{BS},\mu} - \tau_{\text{MS}}$ is the time instance $\tau_{\text{rec},\mu}$ relative to the time scale of the MS. As next step, an acknowledgment signal is transmitted back from BS μ to the MS at time instance $\tau_{\text{ack},\mu}$ relative to the time scale of BS μ. The acknowledgment signal is received at the MS at time $\tau_{\text{rec, ack},\mu}$. The resulting propagation distance is

$$d_{\mu,\text{MS}} = c_0 \left(\tau_{\text{rec,ack},\mu} - (\tau_{\text{ack},\mu} + \tau_{\text{BS},\mu} - \tau_{\text{MS}}) \right). \tag{2.12}$$

Since we assume a reciprocal channel between MS and the BSs — this also includes that the geometry does not change during the processing time — both estimates have to be equal, i.e., $d_{\text{MS},\mu} = d_{\mu,\text{MS}}$. Therefore, we can compute the estimated distance applying the RTTOA principle as

$$d_\mu = \frac{d_{\text{MS},\mu} + d_{\mu,\text{MS}}}{2} = \frac{1}{2} c_0 \left(\tau_{\text{prop},\mu} - \tau_{\text{proc},\mu} \right). \tag{2.13}$$

Hence, the estimated distance depends on the propagation time $\tau_{\text{prop},\mu} = \tau_{\text{rec,ack},\mu} - \tau_{\text{init},\mu}$ and is calculated at the MS with respect to its internal time basis. The processing time $\tau_{\text{proc},\mu} = \tau_{\text{ack},\mu} - \tau_{\text{rec},\mu}$ is calculated at BS μ with respect to its internal time basis and has to be communicated to the MS. Therefore, the estimated propagation distance does not depend on an unknown time basis difference $\tau_{\text{BS},\mu} - \tau_{\text{MS}}$. In general, this time basis can not assumed to be constant over time and effects like clock drift or clock jitter have to be considered in a suitable way [45]. A popular example for RTTOA based location estimation is the ranging approach in the IEEE standard 802.15.4 for UWB systems (cf. [46]).

As resulting system of equations for computing the unknown values (x, y, z) we obtain

$$\begin{aligned}
\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} &= \frac{1}{2} c_0 \left(\tau_{\text{prop},1} - \tau_{\text{proc},1} \right) \\
\sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2} &= \frac{1}{2} c_0 \left(\tau_{\text{prop},2} - \tau_{\text{proc},2} \right) \\
&\vdots \\
\sqrt{(x - x_{N_{\text{BS}}})^2 + (y - y_{N_{\text{BS}}})^2 + (z - z_{N_{\text{BS}}})^2} &= \frac{1}{2} c_0 \left(\tau_{\text{prop},N_{\text{BS}}} - \tau_{\text{proc},N_{\text{BS}}} \right).
\end{aligned} \tag{2.14}$$

It is obvious, that the RTTOA approach results in the same geometric structure as TOA. Therefore, the solution strategies — at least on the level of location estimation — are identical. Since RTTOA approaches require measurements in both uplink and downlink, only dedicated location estimation systems like short-range systems based on UWB technology exploit them.

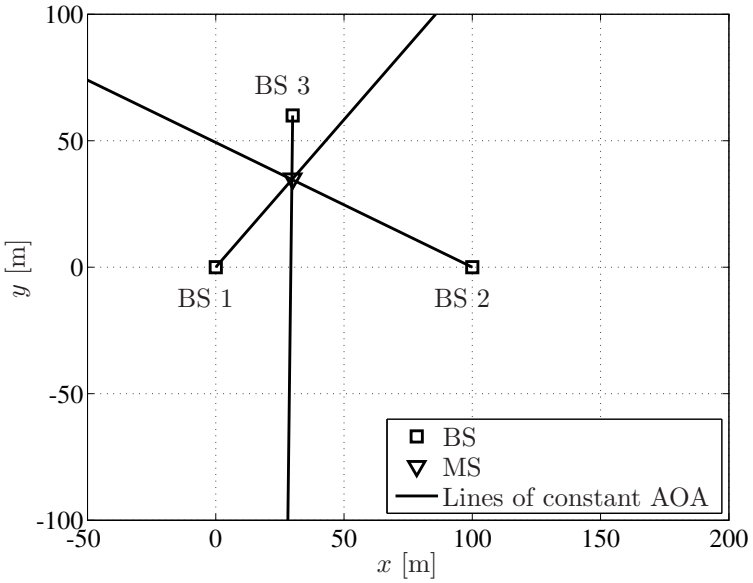


Figure 2.3: AOA location estimation principle

2.2 Angle of arrival

Another parameter depending on the location is the direction or angle from which a signal is arriving at the MS or BS. This results in AOA based location estimation suitable for uplink and downlink processing.

Assuming the two-dimensional case, the estimate of the azimuth angle or direction α_μ of the received signal and the BS location (x_μ, y_μ) determines a line with possible locations of the MS (cf. Figure 2.3). For an arbitrary BS μ we obtain the relations

$$\begin{aligned} x - x_\mu &= d_\mu \cos(\alpha_\mu) \\ y - y_\mu &= d_\mu \sin(\alpha_\mu) . \end{aligned} \tag{2.15}$$

This can be simplified to

$$y - y_\mu = \tan(\alpha_\mu)(x - x_\mu) , \tag{2.16}$$

i.e., the relation is no longer depending on the distance d_μ between MS and BS μ. Hence, we can setup the system of equations with the unknown values (x, y) according to

$$\begin{aligned} y - y_1 &= \tan(\alpha_1)(x - x_1) \\ y - y_2 &= \tan(\alpha_2)(x - x_2) \\ &\vdots \\ y - y_{N_{\text{BS}}} &= \tan(\alpha_{N_{\text{BS}}})(x - x_{N_{\text{BS}}}) \end{aligned} \tag{2.17}$$

assuming N_{BS} involved BSs. Since the equations are linear with respect to the location of the MS, for two-dimensional processing two linear independent equations are sufficient for obtaining a unique solution. In the three-dimensional case additionally the elevation angle needs to be available, however, measurements with two BSs are sufficient.

Clearly, the determination of the angles requires multiple well synchronized and calibrated antenna elements at MS or BS. Furthermore, especially in urban canyons and indoor environments

NLOS and multipath propagation make a reasonable estimation of these angles very challenging. Therefore, AOA is currently not used in mobile radio systems for location estimation.

2.3 Received signal strength

Another principle is based on estimating the signal strength of the received signal. In general, it can be observed that the average received power or RSS is depending inversely on the distance d_μ between BS μ and MS. For instance, in free space the RSS is reciprocally proportional to the square of that distance. However, discontinuities in the propagation medium cause changes in the propagation direction of electromagnetic waves. Such discontinuities, e.g., reflection, refraction, diffraction, or scattering, are omnipresent in typical mobile radio environments and cause multipath and NLOS propagation.

Commonly used channel models describe the relevant dependency as path loss. The proportionality of the received power $P_{\mathrm{RX},\mu}$ is usually described as

$$P_{\mathrm{RX},\mu} \sim P_{\mathrm{TX},\mu} G_{\mathrm{TX},\mu} G_{\mathrm{RX}} \left(\frac{d_\mu}{d_{\mathrm{ref}}} \right)^{-\beta_\mu}, \tag{2.18}$$

where β_μ is denoted as decay factor and depends strongly on the environment. Its value is 2 in case of free space propagation and greater than 2 for general multipath propagation environments. The variables $P_{\mathrm{TX},\mu}$, $G_{\mathrm{TX},\mu}$, G_{RX}, and d_{ref} describe the transmitted power, the antenna gains of transmitter and receiver, and a normalization constant. They all have to be known in order to determine absolute values of the RSS. Note that this dependency is empirical and describes solely an average. Nevertheless, this principle allows at least a rough estimation of the distance d_μ between MS and BS μ.

To determine a location estimate of the MS, several RSS values from measurements between BSs and the MS have to be obtained. The resulting geometric situation and the location estimation problem is the same as for the TOA principle. Note that — similar to AOA — RSS based location estimation is not part of the standards for current mobile radio systems like GSM, UMTS, or LTE.

2.4 Further approaches

The cell identity (e.g., [13]) is a commonly used location estimation principle. It derives MS location information from the locations of the available BSs and — if available — their coverage. Clearly, especially for systems with large cell sizes (e.g., GSM allows cell radii up to $35\,\mathrm{km}$) the obtained location information is not very accurate. Nevertheless, this principle is included in all state-of-the-art communications standards (e.g., GSM, UMTS, or LTE) and is used if other methods (based on, e.g., TDOA or AGNSS) fail. In addition, the sector identity can be provided if the considered BS serves several sectors by multiple antennas. The intersection of several cell or sector identities can further refine the location estimate.

The classical location estimation principles described in the previous sections usually require that several BSs are incorporated. Fingerprinting or pattern matching approaches have the advantage that unique location information can already be obtained with only one BS. The basic idea is the creation of a data-base, where location depending information is stored. Besides classical parameters like (TOAs, TDOAs, RSSs, etc.) the data-base could also include measured channel impulse responses or derived metrics like delay spread or angular spread. In the setup or calibration phase for the data-base, accurate location information of the positions, where the reference

measurements were taken, is required. In the measurement phase, the data-base can be exploited to compare actual measurements with the information stored in the data-base to retrieve location information of the MS. It is obvious that in practical systems it might be difficult to keep the data-base up-to-date in changing environments. Furthermore, the granularity of the data-base has a direct impact on the achievable location estimation accuracy. Examples how fingerprinting is used for location estimation of GSM devices can be found in [47]. A wireless local area network (WLAN) location estimation scheme using fingerprints is described in [25].

Besides location estimation principles that require a certain infrastructure (e.g., a cellular network with BSs) also sensors that deliver location dependent information integrated into the MS can be used. This class includes, e.g., inertial sensors, magnetometers, or barometers, yielding local location information about acceleration, orientation, or height directly to the MS (e.g., [23, 24]). In addition to that, also context-aware information — even not suitable for a stand-alone solution — can support the described principles. For instance, map information can help in the location determination process to resolve ambiguities or adding constraints in the estimation process (e.g., [25, 26]).

3 Timing estimation for location determination in OFDM

In this chapter, we will investigate timing estimation algorithms in OFDM based mobile radio systems as part of the location determination process depicted in Figure 3.1. We start with a description of the considered OFDM system model in Section 3.1, including transmitter, channel, and receiver. Since timing estimation is closely related to the MS synchronization task performed in cellular mobile radio systems, we outline general pilot-aided synchronization techniques in Section 3.2. They usually cope with receiver imperfections like timing offsets or carrier frequency offsets and allow the initial access to the system. These algorithms will be assessed and extended to provide reasonable timing estimates for the location determination process. Additionally, corresponding performance bounds will be derived. As simulations will show, a reduced number of available pilot symbols and inter-cell interference limit the timing estimation performance. Therefore, after a brief discussion about non-pilot-aided synchronization schemes in Section 3.3, the idea of data-aided synchronization in the context of location determination will be introduced in Section 3.4. In that manner, we propose to aid synchronization with estimated data symbols that are used as additional pilot symbols for more accurate timing estimation. Another derived scheme is based on inter-cell interference cancelation to allow a better reception of several BSs in interference limited networks which is essential for reliable location estimates. Both approaches pave the way for new iterative timing estimation schemes.

3.1 OFDM system model

3.1.1 Transmitter

We consider a general mobile radio system based on bit-interleaved coded modulation (BICM) [48] and OFDM [16]. The notation in this section was adapted from [49]. Figure 3.2 shows a general block diagram of the transmitter for one BS of a cellular system. According to the principle of BICM, a bit interleaver decouples the encoder and mapper at the transmitter side. The data bits $b_\eta \in \{0, 1\}$ from a binary source are encoded by, e.g., a convolutional encoder. Other coding schemes as like turbo or low-density parity-check codes could be employed if appropriate. The

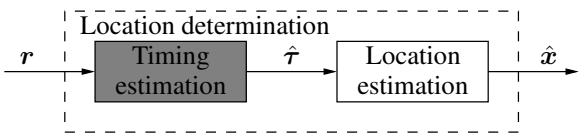


Figure 3.1: Two-step location determination process: timing estimation

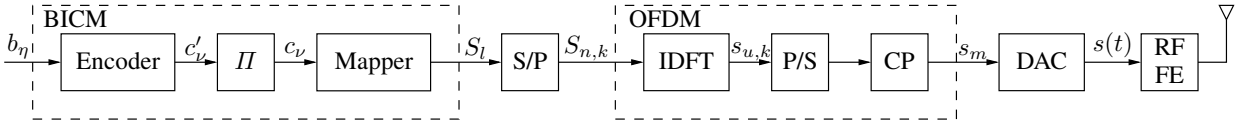


Figure 3.2: BICM-OFDM transmitter

code bits c'_ν are bit-wise interleaved by the interleaver (Π). The interleaved code bits c_ν are then mapped onto a discrete complex-valued symbol alphabet according to

$$S_l = f_{\text{mapper}}\left(c_\nu, c_{\nu+1}, \ldots, c_{\nu+\xi-1}\right) \in \mathcal{S}, \tag{3.1}$$

where the function $f_{\text{mapper}}\left(c_\nu, c_{\nu+1}, \ldots, c_{\nu+\xi-1}\right)$ denotes the mapping of a sequence of ξ bits to one of the 2^ξ complex-valued symbols from the alphabet $\mathcal{S}$. The serial-to-parallel (S/P) converter orders the symbols S_l to N_u used subcarriers and N_s OFDM symbols to form the OFDM frame yielding

$$\begin{aligned} S_{n,k} = S_{l \bmod(N_\text{u}+\lfloor(N_\text{u}-1)/2\rfloor),\lfloor l/N_\text{u}\rfloor}&, \\ n = \lfloor -\left(N_\text{u}-1\right)/2\rfloor, \lfloor -\left(N_\text{u}-1\right)/2\rfloor + 1, \ldots, \lfloor \left(N_\text{u}-1\right)/2\rfloor&, \\ k = 0, 1, \ldots, N_\text{s}-1.& \end{aligned} \tag{3.2}$$

The spacing between two neighboring subcarriers is given as

$$F_\text{s} = \frac{1}{T_\text{s}}, \tag{3.3}$$

where T_s is the duration of one OFDM symbol. Therefore, the corresponding bandwidth occupied by the OFDM transmit signal can be computed as

$$B = N_\text{u} F_\text{s}. \tag{3.4}$$

The frequency domain symbols are then transformed into the time domain by application of an inverse discrete Fourier transform (IDFT)

$$s_{u,k} = \frac{1}{\sqrt{N_\text{FFT}}} \sum_{n=\lfloor -(N_\text{u}-1)/2\rfloor}^{\lfloor (N_\text{u}-1)/2\rfloor} S_{n,k}\, \mathrm{e}^{\jmath \frac{2\pi n u}{N_\text{FFT}}}, \quad u = 0, 1, \ldots, N_\text{FFT}-1, \quad k = 0, 1, \ldots, N_\text{s}-1, \tag{3.5}$$

where in practical systems the IDFT size is often chosen as a power of 2, resulting in an inverse fast Fourier transform (IFFT) of size N_FFT. For pure frequency domain investigations of OFDM communications systems, the classical definition with a summation index $n = 0, 1, \ldots, N_\text{u}-1$, would have been sufficient due to the cyclic properties of OFDM. In our definition the summation index is symmetrically distributed around zero. As we later will see in the derivation of the timing synchronization algorithms (cf. Section 3.2), this choice is of fundamental importance. After parallel-to-serial (P/S) conversion to s_{u+kN_FFT}, a cyclic prefix (CP) of N_CP samples is periodically inserted between N_FFT consecutive samples belonging to one OFDM symbol. Hence, inter-symbol interference between consecutive OFDM symbols will be avoided if the length of the CP $T_\text{CP} = N_\text{CP} T_\text{samp}$, with the sampling time $T_\text{samp} = T_\text{s}/N_\text{FFT}$, is larger than the maximum excess

delay of the channel. The resulting signal is

$$\begin{aligned} s_m = s_{N_{\text{CP}}+i+k(N_{\text{FFT}}+N_{\text{CP}})} &= \frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{n=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} S_{n,k}\, \mathrm{e}^{\jmath \frac{2\pi n i}{N_{\text{FFT}}}}, \\ & i = -N_{\text{CP}}, -N_{\text{CP}}+1, \ldots, N_{\text{FFT}}-1, \\ & k = 0, 1, \ldots, N_{\text{s}}-1. \end{aligned} \tag{3.6}$$

Thus, one OFDM symbol is of duration $T'_{\text{s}} = T_{\text{s}} + T_{\text{CP}}$ and the complete OFDM frame is of duration $T_{\text{frame}} = N_{\text{s}} T'_{\text{s}}$. With a digital-to-analog converter (DAC) the discrete-time signal s_m is transformed to the continuous-time signal $s(t)$, where the waveform of a continuous-time pulse in OFDM is a rectangular pulse. Therefore, the DAC output is

$$s(t) = \sum_{m=-\infty}^{\infty} s_m \operatorname{rect}\left(\frac{t - mT_{\text{samp}} - T_{\text{samp}}/2}{T_{\text{samp}}} \right), \tag{3.7}$$

using

$$\operatorname{rect}(x) = \begin{cases} 1, & \text{for } -\dfrac{1}{2} \le x \le \dfrac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.8}$$

Finally, the continuous-time signal $s(t)$ is up-converted by the radio frequency (RF) frontend (FE) to the RF transmission band and is sent via the antenna of the transmitter (TX) over the mobile radio channel.

For compact notation, we introduce

$$\boldsymbol{s}_k^{\text{FD}} = \left[S_{\lfloor -(N_{\text{u}}-1)/2 \rfloor, k}, S_{\lfloor -(N_{\text{u}}-1)/2 \rfloor + 1, k}, \ldots, S_{\lfloor (N_{\text{u}}-1)/2 \rfloor, k} \right]^{\text{T}} \in \mathbb{C}^{N_{\text{u}}}, \quad k = 0, 1, \ldots, N_{\text{s}}-1, \tag{3.9}$$

as vectors composed of the N_{u} transmitted frequency domain symbols for the N_{s} OFDM symbols in the frame. The corresponding transmitted OFDM frame in the frequency domain is defined as

$$\boldsymbol{S}^{\text{FD}} = \left[\boldsymbol{s}_0^{\text{FD}}, \boldsymbol{s}_1^{\text{FD}}, \ldots, \boldsymbol{s}_{N_{\text{s}}-1}^{\text{FD}} \right] \in \mathbb{C}^{N_{\text{u}} \times N_{\text{s}}}. \tag{3.10}$$

Figure 3.3 depicts a general OFDM frame according to the previous definitions. With the Fourier matrix $\boldsymbol{F}_{\text{Fourier}} \in \mathbb{C}^{N_{\text{u}} \times N_{\text{FFT}}}$ according to

$$\left[\boldsymbol{F}_{\text{Fourier}} \right]_{n,u} = \frac{1}{\sqrt{N_{\text{FFT}}}}\, \mathrm{e}^{\jmath \frac{2\pi u n}{N_{\text{FFT}}}} \tag{3.11}$$

we can represent the time domain signal after the IFFT as

$$\boldsymbol{s}_k^{\text{TD}} = \boldsymbol{F}_{\text{Fourier}}^{\text{H}} \boldsymbol{s}_k^{\text{FD}} \in \mathbb{C}^{N_{\text{FFT}}}, \quad k = 0, 1, \ldots, N_{\text{s}}-1, \tag{3.12}$$

and the corresponding OFDM frame as

$$\boldsymbol{S}^{\text{TD}} = \left[\boldsymbol{s}_0^{\text{TD}}, \boldsymbol{s}_1^{\text{TD}}, \ldots, \boldsymbol{s}_{N_{\text{s}}-1}^{\text{TD}} \right] \in \mathbb{C}^{N_{\text{FFT}} \times N_{\text{s}}}. \tag{3.13}$$

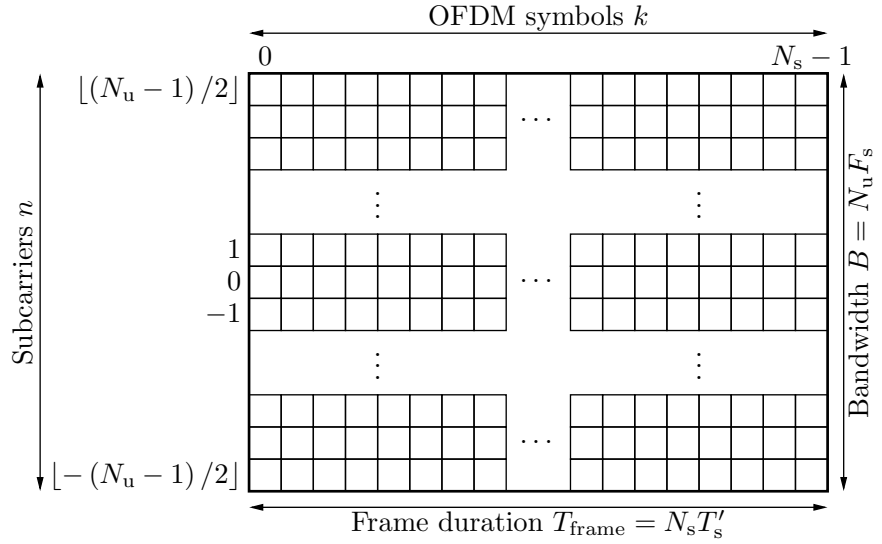


Figure 3.3: OFDM frame

3.1.2 Channel

After transmission over the mobile radio channel with the time-variant channel impulse response $h(\tau, t)$ and down-conversion into the baseband by the RF FE, the continuous-time baseband signal at the receiver (RX) can be written as

$$r(t) = \int_{\tau=-\infty}^{\infty} h(\tau, t)\, s(t-\tau)\, \mathrm{d}\tau + z(t)\,. \tag{3.14}$$

The signal $z(t)$ denotes complex-valued zero-mean additive white Gaussian noise (AWGN) of variance σ_z^2, where real and imaginary parts are independent and identically distributed. An equivalent frequency domain representation of the continuous-time baseband signal is

$$r(t) = \int_{f=-\infty}^{\infty} H(f, t)\, S(f)\, \mathrm{e}^{\jmath 2\pi f t}\, \mathrm{d}f + z(t)\,, \tag{3.15}$$

where the time-variant channel transfer function

$$H(f, t) = \int_{\tau=-\infty}^{\infty} h(\tau, t)\, \mathrm{e}^{-\jmath 2\pi f \tau}\, \mathrm{d}\tau \tag{3.16}$$

is the Fourier transform of the time-variant channel impulse response, and

$$S(f) = \int_{\tau=-\infty}^{\infty} s(\tau)\, \mathrm{e}^{-\jmath 2\pi f \tau}\, \mathrm{d}\tau \tag{3.17}$$

is the Fourier transform of $s(t)$.

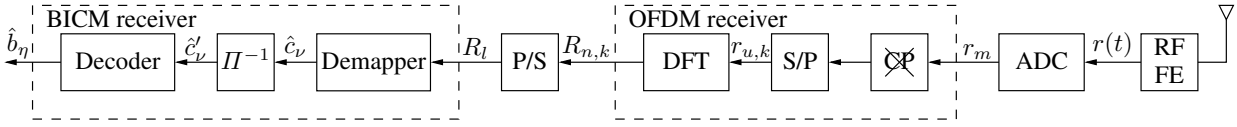


Figure 3.4: BICM-OFDM receiver

The samples of the channel impulse response for the OFDM frame are given as

$$h_{l,k} = h\left(lT_{\text{samp}}, kT'_{\text{s}}\right), \quad l = 0, 1, \ldots, L-1, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.18}$$

where it is assumed that the channel is constant during one OFDM symbol and $L < N_{\text{CP}}$ defines the considered length of the channel impulse response in terms of samples. With that we compose the channel vector

$$\boldsymbol{h}_k^{\text{TD}} = \left[h_{0,k}, h_{1,k}, \ldots, h_{L-1,k}\right]^{\text{T}} \in \mathbb{C}^L, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.19}$$

for the OFDM symbols, and the matrix

$$\boldsymbol{H}^{\text{TD}} = \left[\boldsymbol{h}_0^{\text{TD}}, \boldsymbol{h}_1^{\text{TD}}, \ldots, \boldsymbol{h}_{N_{\text{s}}-1}^{\text{TD}}\right] \in \mathbb{C}^{L \times N_{\text{s}}} \tag{3.20}$$

for the OFDM frame. Using the samples of the channel transfer function for the OFDM frame according to

$$\begin{aligned} H_{n,k} &= H\left(nF_{\text{s}}, kT'_{\text{s}}\right), \\ &\quad n = \lfloor -\left(N_{\text{u}} - 1\right)/2 \rfloor, \lfloor -\left(N_{\text{u}} - 1\right)/2 \rfloor + 1, \ldots, \lfloor \left(N_{\text{u}} - 1\right)/2 \rfloor, \\ &\quad k = 0, 1, \ldots, N_{\text{s}} - 1, \end{aligned} \tag{3.21}$$

we introduce

$$\boldsymbol{h}_k^{\text{FD}} = \left[H_{\lfloor -(N_{\text{u}}-1)/2 \rfloor, k}, H_{\lfloor -(N_{\text{u}}-1)/2 \rfloor + 1, k}, \ldots, H_{\lfloor (N_{\text{u}}-1)/2 \rfloor, k}\right]^{\text{T}} \in \mathbb{C}^{N_{\text{u}}}, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.22}$$

as vector composed of the N_{u} samples of the channel transfer function for the N_{s} OFDM symbols in the frame. The corresponding samples for the OFDM frame are arranged in the matrix

$$\boldsymbol{H}^{\text{FD}} = \left[\boldsymbol{h}_0^{\text{FD}}, \boldsymbol{h}_1^{\text{FD}}, \ldots, \boldsymbol{h}_{N_{\text{s}}-1}^{\text{FD}}\right] \in \mathbb{C}^{N_{\text{u}} \times N_{\text{s}}}. \tag{3.23}$$

For the frequency domain noise contributions $Z_{n,k}$, $\boldsymbol{z}_k^{\text{FD}}$, and $\boldsymbol{Z}^{\text{FD}}$, we assume an equivalent definition as in (3.21), (3.22), and (3.23).

3.1.3 Receiver

The discrete-time received signal (cf. Figure 3.4) is obtained as

$$r_m = r\left(t\right)\big|_{t = mT_{\text{samp}}} = \sum_{m=-\infty}^{\infty} r\left(t\right) \delta\left(t - mT_{\text{samp}}\right) \tag{3.24}$$

with the Dirac delta function $\delta\left(t\right)$. Timing and frequency synchronization take care that receiver imperfections as like timing and carrier frequency offsets between transmitted and received signals are compensated. This will be further discussed in Section 3.2. The N_{CP} samples from the CP are removed and the S/P converted remaining signal stream is

$$r_{u,k} = r_{u + N_{\text{CP}} + k(N_{\text{FFT}} + N_{\text{CP}})}, \quad u = 0, 1, \ldots, N_{\text{FFT}} - 1, \quad k = 0, 1, \ldots, N_{\text{s}} - 1. \tag{3.25}$$

Next, with an fast fourier transform (FFT) the time domain signal $r_{u,k}$ is transformed into the frequency domain resulting in the received frame (cf. Figure 3.3)

$$\begin{aligned} R_{n,k} = \frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{u=0}^{N_{\text{FFT}}-1} r_{u,k}\, \mathrm{e}^{-\jmath \frac{2\pi u n}{N_{\text{FFT}}}}, \\ n = \lfloor -(N_{\text{u}}-1)/2 \rfloor, \lfloor -(N_{\text{u}}-1)/2 \rfloor + 1, \ldots, \lfloor (N_{\text{u}}-1)/2 \rfloor, \\ k = 0, 1, \ldots, N_{\text{s}} - 1. \end{aligned} \tag{3.26}$$

If the subcarrier spacing F_{s} is much smaller than the coherence bandwidth of the channel and the duration of the CP T_{CP} is smaller than the maximum excess delay of the channel, there is frequency flat fading on each subcarrier and no inter-symbol interference between consecutive OFDM symbols is present. If further the duration of one OFDM symbol T_{s} is much smaller than the coherence time of the channel, there is no inter-carrier interference on the received signal. Then, the frequency domain received signal can simply be described as

$$\begin{aligned} R_{n,k} = H_{n,k} S_{n,k} + Z_{n,k}, \\ n = \lfloor -(N_{\text{u}}-1)/2 \rfloor, \lfloor -(N_{\text{u}}-1)/2 \rfloor + 1, \ldots, \lfloor (N_{\text{u}}-1)/2 \rfloor, \\ k = 0, 1, \ldots, N_{\text{s}} - 1. \end{aligned} \tag{3.27}$$

This frame is P/S converted according to

$$R_l = R_{l \bmod N_{\text{u}}, \lfloor l/N_{\text{u}} \rfloor}, \quad l = 0, 1, \ldots, N_{\text{u}} N_{\text{s}} - 1. \tag{3.28}$$

As final step of the receiver chain a BICM receiver is employed. The demapper computes hard decision or soft estimates $\hat{c}_\nu$ of the transmitted bits c_ν. Channel state information (CSI) is either taken into account by applying a one-tap frequency equalizer before the demapper or by directly including this information in the demapper. After deinterleaving the hard decision or soft estimates $\hat{c}'_\nu$ are decoded resulting in the estimated bits $\hat{b}_\eta$.

At the receiver side, we denote

$$\boldsymbol{r}_k^{\text{TD}} = \left[r_{0,k}, r_{1,k}, \ldots, r_{N_{\text{FFT}}-1,k} \right]^{\text{T}} \in \mathbb{C}^{N_{\text{FFT}}}, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.29}$$

as vector composed of the N_{FFT} time domain symbols for the N_{s} OFDM symbols, and

$$\boldsymbol{R}^{\text{TD}} = \left[\boldsymbol{r}_0^{\text{TD}}, \boldsymbol{r}_1^{\text{TD}}, \ldots, \boldsymbol{r}_{N_{\text{s}}-1}^{\text{TD}} \right] \in \mathbb{C}^{N_{\text{FFT}} \times N_{\text{s}}} \tag{3.30}$$

as the corresponding OFDM frame. After the IFFT, the resulting frequency domain symbols can be represented as

$$\boldsymbol{r}_k^{\text{FD}} = \boldsymbol{F}_{\text{Fourier}} \boldsymbol{r}_k^{\text{TD}} \in \mathbb{C}^{N_{\text{u}}}, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.31}$$

and the corresponding OFDM frame as

$$\boldsymbol{R}^{\text{FD}} = \left[\boldsymbol{r}_0^{\text{FD}}, \boldsymbol{r}_1^{\text{FD}}, \ldots, \boldsymbol{r}_{N_{\text{s}}-1}^{\text{FD}} \right] \in \mathbb{C}^{N_{\text{u}} \times N_{\text{s}}}. \tag{3.32}$$

Since the OFDM properties allow a simple computation of the received frequency domain symbols (cf. (3.27)), we can calculate them according to

$$\boldsymbol{r}_k^{\text{FD}} = \boldsymbol{H}_k^{\text{FD}} \boldsymbol{s}_k^{\text{FD}} + \boldsymbol{z}_k^{\text{FD}}, \quad k = 0, 1, \ldots, N_{\text{s}} - 1, \tag{3.33}$$

using the diagonal channel matrix

$$\boldsymbol{H}_k^{\text{FD}} = \mathbf{I}_{N_{\text{u}}} \boldsymbol{h}_k^{\text{FD}} \in \mathbb{C}^{N_{\text{u}} \times N_{\text{u}}}. \tag{3.34}$$

3.2 Pilot-aided synchronization

3.2.1 Receiver imperfections

Receivers are affected by imperfections which disturb the transmitted signal in addition to channel effects and noise. Thus, synchronization tasks are required in the receiver processing chain to estimate and compensate these effects for a reliable detection of the transmitted data [50, 51, 52]. In the following, we discuss the impact of timing offset and carrier frequency offset (CFO) on the received signal. We assume that there are no further RF FE related effects like non-linear power amplification or phase noise, and assume a quantizer in the ADC with an arbitrary floating point precision.

Timing offset

Processing time within the transmission chain and the propagation over the mobile radio channel introduce a timing offset in the received signal. Since transmitter and receiver usually do not share the same clock in mobile radio systems, the receiver needs to recover timing information about the received signal stream before the data can be processed. This includes symbol as well as frame timing information. The symbol timing is used to adjust the FFT window to an appropriate place. The frame timing is used to extract the correct OFDM symbols from the overall frame. Even if timing recovery or timing offset estimation is an important task for communications systems as part of the initial access, the requirements on the accuracy are comparatively relaxed. The OFDM system design allows, that the FFT window can start in the inter-symbol interference free region of the received samples, i.e., there are $N_{\text{CP}} - L$ possible samples, where the FFT window can be placed, without reducing the performance [53, 54]. Clearly, for location determination much more accurate timing estimates are required as it will be discussed during the description of the respective algorithms.

An arbitrary timing offset τ_0 can be modeled by including it in the channel according to

$$h(\tau, t) = \delta(t - \tau_0). \tag{3.35}$$

This timing offset can represent, e.g., a propagation delay or the start of a synchronization symbol relative to the internal MS clock. Hence, following (3.14), the received continuous-time signal can be written as

$$r(t) = s(t - \tau_0) + z(t). \tag{3.36}$$

The corresponding sampled discrete-time version of the signal according to (3.24) is then

$$r_m = s(mT_{\text{samp}} - \tau_0) + z(mT_{\text{samp}}), \tag{3.37}$$

or in a more compact notation

$$r_m = s_m(\tau_0) + z_m. \tag{3.38}$$

The index m denotes the samples with respect to the internal receiver sampling clock. Note that at this stage no frame or symbol timing is available. Due to the timing offset τ_0, the samples are no

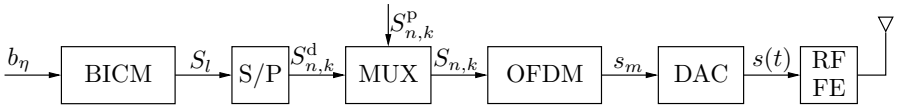


Figure 3.5: BICM-OFDM transmitter with MUX for pilot and data symbols

longer on the sampling grid and according to

$$\begin{aligned} s_m(\tau_0) &= \frac{1}{\sqrt{N_{\mathrm{FFT}}}} \sum_{n=\lfloor -(N_{\mathrm{u}}-1)/2 \rfloor}^{\lfloor (N_{\mathrm{u}}-1)/2 \rfloor} S_{n,k}\, \mathrm{e}^{\jmath 2\pi n F_{\mathrm{s}}(mT_{\mathrm{samp}}-\tau_0)} \\ &= \frac{1}{\sqrt{N_{\mathrm{FFT}}}} \sum_{n=\lfloor -(N_{\mathrm{u}}-1)/2 \rfloor}^{\lfloor (N_{\mathrm{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n F_{\mathrm{s}}\tau_0}\, S_{n,k}\, \mathrm{e}^{\jmath \frac{2\pi n m}{N_{\mathrm{FFT}}}} \end{aligned} \tag{3.39}$$

the timing offset introduces a linear phase shift in the frequency domain signal.

Carrier frequency offset

Due to different oscillator properties on transmitter and receiver side, a CFO occurs in the received signal. The CFO introduces a linear phase shift in the time domain signal. This results in a loss of the mutual orthogonality among the subcarriers, i.e., inter-carrier interference occurs (e.g., [53]) which reduces the performance. Contrary to the timing offset, OFDM is very sensitive with respect to CFOs. They directly have an impact on the achievable transmission performance (e.g., [55]).

We assume an arbitrary CFO ϕ_0 which is normalized to the subcarrier spacing F_{s}. Then, the received sampled discrete-time signal can be represented as

$$\begin{aligned} \tilde{r}_m &= \mathrm{e}^{\jmath \frac{2\pi\phi_0 m + \phi_{\mathrm{s}}}{N_{\mathrm{FFT}}}}\, r_m \\ &= \mathrm{e}^{\jmath \frac{2\pi\phi_0 m + \phi_{\mathrm{s}}}{N_{\mathrm{FFT}}}}\, \left(s_m(\tau_0) + z_m\right), \end{aligned} \tag{3.40}$$

including timing offset, CFO, and an arbitrary carrier phase ϕ_{s}.

3.2.2 Synchronization algorithms

To estimate timing offset and CFO at the MS of a mobile radio system, synchronization algorithms are employed. In pilot-aided synchronization, these algorithms rely on pilot or synchronization symbols that are included in the transmitted signal stream. These symbols — or at least their structure — have to be known at the MS. Usually, these pilot symbols $S^{\mathrm{p}}_{n,k}$ are combined by a multiplexer (MUX) with the data symbols $S^{\mathrm{d}}_{n,k}$ in frequency domain to form the transmitted frame $S_{n,k}$ (cf. (3.2) and Figure 3.5).

At the receiver side, the synchronization signals are exploited to recover timing information and to estimate the CFO (cf. Figure 3.6). The CFO is usually directly compensated on the received time domain signal. The timing information is used to adjust the FFT window to an appropriate place and to remove the CP. In the following, we describe commonly used OFDM synchronization strategies and discuss them with respect to their suitability to provide precise timing information for location determination in cellular mobile radio systems. For the derivation of the algorithms, we currently assume, that there is one OFDM symbol in the frame dedicated to synchronization

Figure 3.6: BICM-OFDM receiver with timing and CFO estimation

purposes. Information about this symbol has to be available at the receiver. This can either be the knowledge of the complete transmitted signal or only its structure. Latter one is usually used for differential correlation based synchronization being robust with respect to CFOs. These approaches allow timing and CFO synchronization without the explicit knowledge of the transmitted synchronization signal.

Synchronization using differential correlation

The idea of OFDM synchronization using differential correlation (DC) was introduced by Schmidl and Cox [56]. The approach requires a specific synchronization signal consisting of two identical halves in the time domain. Without loss of any generality, this synchronization signal

$$s\left(\tau_0 + kT_{\text{samp}}\right) = s\left(\tau_0 + \left(k + \frac{N_{\text{FFT}}}{2}\right) T_{\text{samp}}\right), \quad k = 0, 1, \ldots, \frac{N_{\text{FFT}}}{2} - 1, \tag{3.41}$$

starts at τ_0 — relative to the MS clock — and has a length of N_{FFT} samples. Such a signal with a repetitive structure in the time domain can easily be generated in frequency domain by solely occupying the even subcarriers and setting the odd subcarriers to zero.

To detect the starting point of the synchronization signal designed according to (3.41), a DC operation on the received signal is required, where for the first we assume that there is no CFO present. According to [56], the receiver computes the metric

$$C_m^{\text{DC}} = \frac{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} r_{m+k} r^*_{m+k+N_{\text{FFT}}/2}}{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} \left| r_{m+k+N_{\text{FFT}}/2} \right|^2}, \tag{3.42}$$

where the nominator is the correlation function using the sampled discrete-time received signal for all available samples m and the denominator is a normalization term. Practically, this operation will be implemented in terms of a sliding window operation. Even though the metric in (3.42) lacks of a theoretical justification, it provides reliable estimates [56]. A comparison with other metrics can be found in [57, 53]. A search for the correlation peak results in the estimated sample for the start of the synchronization symbol and the corresponding timing estimate, i.e.,

$$\begin{aligned} \hat{m}_0^{\text{DC, coh.}} &= \underset{m}{\operatorname{argmax}}\, \Re\left\{C_m^{\text{DC}}\right\} \\ \hat{\tau}_0^{\text{DC, coh.}} &= \underset{m}{\operatorname{argmax}}\, \Re\left\{C_m^{\text{DC}} T_{\text{samp}}\right\}. \end{aligned} \tag{3.43}$$

As the phase information is still available by this approach, we denote it as coherent timing estimation. If the received signal is affected by a CFO, the DC results in

$$\begin{aligned}\tilde{C}_m^{\text{DC}} &= \frac{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} \tilde{r}_{m+k}\tilde{r}^*_{m+k+N_{\text{FFT}}/2}}{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} \left|\tilde{r}_{m+k+N_{\text{FFT}}/2}\right|^2} \\ &= \frac{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} \mathrm{e}^{\jmath\frac{2\pi\phi_0(m+k)+\phi_{\text{s}}}{N_{\text{FFT}}}} r_{m+k}\, \mathrm{e}^{-\jmath\frac{2\pi\phi_0(m+k+N_{\text{FFT}}/2)+\phi_{\text{s}}}{N_{\text{FFT}}}} r^*_{m+k+N_{\text{FFT}}/2}}{\sum_{k=0}^{\frac{N_{\text{FFT}}}{2}-1} \left|\mathrm{e}^{\jmath\frac{2\pi\phi_0(m+k+N_{\text{FFT}}/2)+\phi_{\text{s}}}{N_{\text{FFT}}}} r_{m+k+N_{\text{FFT}}/2}\right|^2} \\ &= \mathrm{e}^{-\jmath\pi\phi_0}\, C_m^{\text{DC}}.\end{aligned} \tag{3.44}$$

Hence, it includes a constant phase rotation due to the CFO. For timing estimation, this can be eliminated by taking the absolute value before searching for the maximum of the correlation metric, i.e.,

$$\begin{aligned}\hat{m}_0^{\text{DC, non-coh.}} &= \underset{m}{\operatorname{argmax}} \left|\tilde{C}_m^{\text{DC}}\right| = \underset{m}{\operatorname{argmax}} \left|C_m^{\text{DC}}\right| \\ \hat{\tau}_0^{\text{DC, non-coh.}} &= \underset{m}{\operatorname{argmax}} \left|\tilde{C}_m^{\text{DC}} T_{\text{samp}}\right| = \underset{m}{\operatorname{argmax}} \left|C_m^{\text{DC}} T_{\text{samp}}\right|.\end{aligned} \tag{3.45}$$

Due to the absolute value operation which eliminates the phase information, we denote this approach as non-coherent timing estimation contrary to the coherent technique in (3.43). Clearly, also in channels affected by multipath fading the non-coherent timing estimation can provide a more robust approach.

It is obvious, that the repetitive synchronization sequence is suitable for CFO estimation as well. Observing (3.44) suggests that the CFO can be estimated by

$$\hat{\phi}_0^{\text{DC}} = \frac{\angle \tilde{C}^{\text{DC}}_{\hat{m}_0}}{\pi}, \tag{3.46}$$

i.e., it is the normalized argument of the correlation metric at the estimated sample of the timing offset. The properties of the complex exponential function limit the unambiguous range for the possible CFOs to $-1 < \phi_0 \leq 1$, i.e., for $|\phi_0| > 1$ solely the fractional part of the CFO can be estimated. For an estimation of the integer part a second synchronization symbol would be required (cf. [56]).

To be conform with all the other OFDM symbols in the frame, also the synchronization symbol usually includes a CP. This has the effect, that instead of a correlation peak a plateau with the length of the CP will be the outcome of the DC process. Hence, for estimating the start of the synchronization symbol the end of this plateau needs to be determined. This can be done, e.g., with the 90%-method [56].

Figure 3.9 depicts the real and absolute value of the correlation function over the sampling index, where it is assumed that the synchronization symbols starts at $\tau_0 = 0$ and neither noise nor CFO are present. For the generation of the synchronization sequences, LTE-specific parameters

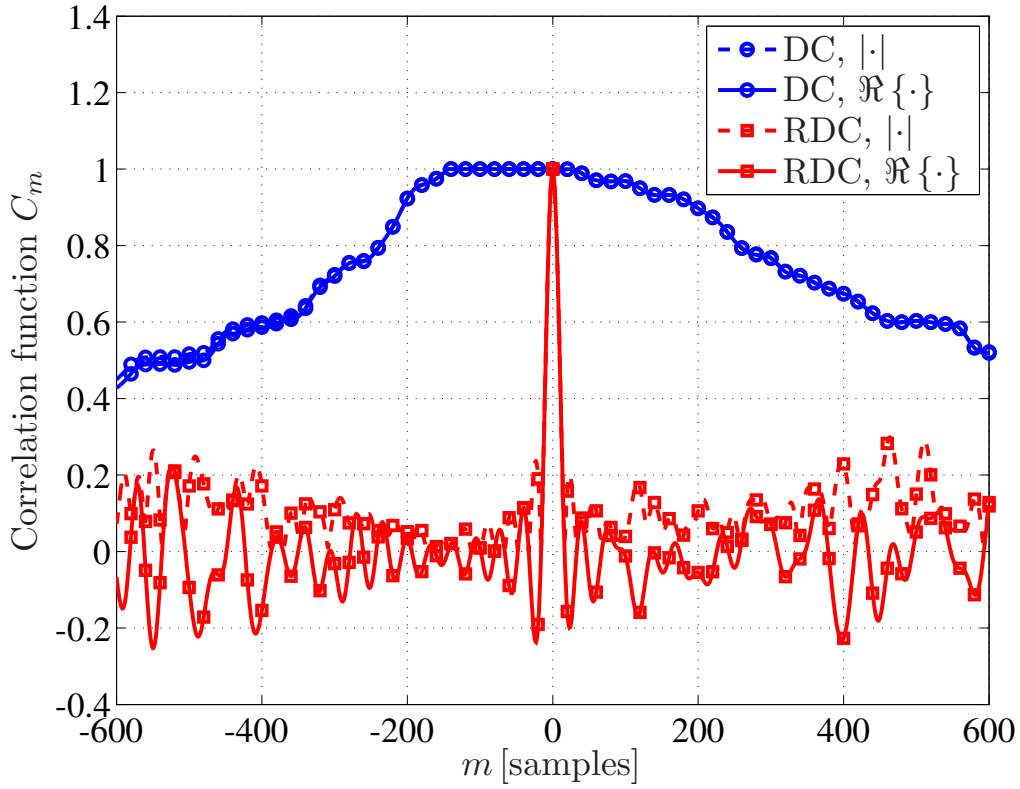


Figure 3.7: Correlation functions using differential correlation based synchronization

are employed, including an FFT length of $N_{\text{FFT}} = 2048$, CP length of $N_{\text{CP}} = 144$, and $N_{\text{u}} = 62$ occupied subcarriers [58]. Figure 3.9 shows clearly the plateau with the length of the CP. It is obvious that the detection of the end of this plateau — especially under noisy conditions — will be very challenging. Therefore, high accurate timing estimation, and hence, location determination cannot be expected with synchronization using DC. We further observe that there are no major differences between the real and absolute values of the correlation function, hence, coherent and non-coherent processing will almost give the same results for this setup.

For communications, where it is solely required to place the FFT window at an appropriate place of the received signal, DC is a suitable approach. The remaining phase shift in frequency domain due to an incorrect timing estimate will anyhow be compensated by a suitable channel estimation. Nevertheless, the DC method can be adapted to avoid the plateau in the correlation function for obtaining a clear correlation peak. This can be realized by increasing the number of signal repetitions in the synchronization signal as it is proposed in [59] and [60]. In [61] this approach is extended to enhance CFO estimation capabilities and to integrate channel estimation. In [62] and [63] further variations of repetitive structures for OFDM synchronization are studied. A dedicated application of DC to location determination purposes and extensions for performance improvements are investigated in [35].

Synchronization using reverse differential correlation

Another approach that overcomes the plateau-like output of the correlation function results in synchronization using reverse differential correlation (RDC) [64]. It is based on a reverse repetitive

structure of the synchronization signal according to

$$s\left(\tau_0 + kT_{\text{samp}}\right) = s^*\left(\tau_0 + \left(N_{\text{FFT}} - k - 1\right)T_{\text{samp}}\right), \quad k = 0, 1, \ldots, \frac{N_{\text{FFT}}}{2} - 1. \tag{3.47}$$

Thus, the second half of the synchronization symbol is the conjugate mirrored part of the first half. These sequences can easily be generated in the frequency domain by using solely real-valued symbols. To obtain timing information with such kind of sequences, RDC has to be applied, i.e., a reasonable metric is

$$C_m^{\text{RDC}} = \frac{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} r_{m+k} r_{N_{\text{FFT}}-k+m}}{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} \left|r_{N_{\text{FFT}}-k+m}\right|^2}. \tag{3.48}$$

As the conjugate operation is already included in the signal design in (3.47), it is not part of the calculation of the correlation sum. The estimation process for RDC is similar to that for DC, hence, with

$$\begin{aligned} \hat{m}_0^{\text{RDC, coh.}} &= \underset{m}{\operatorname{argmax}}\, \Re\left\{C_m^{\text{RDC}}\right\} \\ \hat{\tau}_0^{\text{RDC, coh.}} &= \underset{m}{\operatorname{argmax}}\, \Re\left\{C_m^{\text{RDC}} T_{\text{samp}}\right\} \end{aligned} \tag{3.49}$$

the sample index and timing for the start of the synchronization symbol is determined.

If a CFO is present in the signal, the resulting RDC gives

$$\begin{aligned} \tilde{C}_m^{\text{RDC}} &= \frac{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} \tilde{r}_{m+k} \tilde{r}_{N_{\text{FFT}}-k+m}}{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} \left|\tilde{r}_{N_{\text{FFT}}-k+m}\right|^2} \\ &= \frac{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} \mathrm{e}^{\jmath\frac{2\pi\phi_0(m+k)+\phi_{\text{s}}}{N_{\text{FFT}}}} r_{m+k}\, \mathrm{e}^{\jmath\frac{2\pi\phi_0\left(N_{\text{FFT}}-k+m\right)+\phi_{\text{s}}}{N_{\text{FFT}}}} r_{N_{\text{FFT}}-k+m}}{\sum\limits_{k=1}^{\frac{N_{\text{FFT}}}{2}-1} \left|\mathrm{e}^{\jmath\frac{2\pi\phi_0\left(N_{\text{FFT}}-k+m\right)+\phi_{\text{s}}}{N_{\text{FFT}}}} r_{N_{\text{FFT}}-k+m}\right|^2} \\ &= \mathrm{e}^{\jmath\frac{2\pi\phi_0 N_{\text{FFT}}(2m+N_{\text{FFT}})+2\phi_{\text{s}}}{N_{\text{FFT}}}} C_m. \end{aligned} \tag{3.50}$$

Compared to DC, the resulting correlation sum for RDC similarly depends on the carrier phase offset. Nevertheless, by taking the absolute value this dependency can be eliminated and we obtain for the non-coherent estimation

$$\begin{aligned} \hat{m}_0^{\text{RDC, non-coh.}} &= \underset{m}{\operatorname{argmax}} \left|\tilde{C}_m^{\text{RDC}}\right| = \underset{m}{\operatorname{argmax}} \left|C_m^{\text{RDC}}\right| \\ \hat{\tau}_0^{\text{RDC, non-coh.}} &= \underset{m}{\operatorname{argmax}} \left|\tilde{C}_m^{\text{RDC}} T_{\text{samp}}\right| = \underset{m}{\operatorname{argmax}} \left|C_m^{\text{RDC}} T_{\text{samp}}\right|. \end{aligned} \tag{3.51}$$

Figure 3.9 shows in addition to DC also the correlation functions for RDC. We observe the clear peak making a timing estimation, and hence, location estimation much more reliable compared to DC.

Synchronization using cross-correlation

Even if the differential based methods (DC and RDC) discussed before can provide information about the signal timing, they have certain constraints and limitations that should briefly be outlined. First of all, for cellular mobile radio systems like LTE it is desired that all BSs transmit in the same frequency band (frequency re-use of one) due to spectral efficiency reasons. This has the effect, that the signals from all BSs superpose and have to be processed jointly at the MS. With the differential based approaches it is certainly possible to detect different synchronization sequences in the received signal stream. However, as the structure of these sequences is the same for all BSs, it can not be distinguished at the MS, which correlation peak belongs to which BS. For communications purposes it is usually sufficient to solely detect the overall highest peak for the overall strongest BS. Since for location determination timing estimation with least three BSs has to be performed, the (R)DC methods are not suitable to provide such side information. Furthermore, the differential based approaches usually work on the sampling grid which restricts the resolution even if interpolation techniques can be applied afterwards.

Therefore, in this section cross-correlation (CC) approaches will be discussed. Compared to the differential methods outlined before, for CC explicit knowledge of the transmitted synchronization symbol is necessary. In a practical system, this approach would be applied after a coarse timing estimation with the classical differential correlation techniques. Without loss of generality, we assume that the relevant synchronization symbol is still transmitted at τ_0 and N_{FFT} consecutive received samples

$$\boldsymbol{r} = \left[r_0, r_1, \ldots, r_{N_{\text{FFT}}-1}\right] \mathbb{C}^{N_{\text{FFT}}} \tag{3.52}$$

according to (3.38) are available. The target is to estimate τ_0 exploiting the received signal. Optimization criterion for this approach is the a posteriori probability

$$\mathrm{p}\left(\tau_0|\boldsymbol{r}\right) = \frac{\mathrm{p}\left(\boldsymbol{r}|\tau_0\right)\mathrm{p}\left(\tau_0\right)}{\mathrm{p}\left(\boldsymbol{r}\right)} \tag{3.53}$$

for τ_0 given the received signal $\boldsymbol{r}$. For maximizing this probability we obtain the maximum a posteriori (MAP) estimate [12, 65]

$$\hat{\tau}_0 = \underset{\tau_0}{\operatorname{argmax}}\, \mathrm{p}\left(\tau_0|\boldsymbol{r}\right) = \underset{\tau_0}{\operatorname{argmax}}\, \mathrm{p}\left(\boldsymbol{r}|\tau_0\right)\mathrm{p}\left(\tau_0\right). \tag{3.54}$$

Since the a priori probability $\mathrm{p}\left(\tau_0\right)$ is not always known, we assume it to be uniform. This results in the well-known maximum likelihood (ML) estimate [12, 65]

$$\hat{\tau}_0 = \underset{\tau_0}{\operatorname{argmax}}\, \mathrm{p}\left(\boldsymbol{r}|\tau_0\right). \tag{3.55}$$

Considering (3.38), the likelihood function can be represented as

$$\mathrm{p}\left(\boldsymbol{r}|\tau_0\right) = \frac{1}{\left(\pi\sigma_z^2\right)^{N_{\text{FFT}}}} \mathrm{e}^{-\frac{1}{\sigma_z^2}\sum_{m=0}^{N_{\text{FFT}}-1}|r_m - s_m(\tau_0)|^2}. \tag{3.56}$$

Hence, the ML estimate can be computed as

$$\begin{aligned}\hat{\tau}_0 &= \underset{\tau_0}{\operatorname{argmax}}\, \mathrm{p}\left(\boldsymbol{r}|\tau_0\right) \\ &= \underset{\tau_0}{\operatorname{argmax}}\, \frac{1}{\left(\pi\sigma_z^2\right)^{N_{\mathrm{FFT}}}}\, \mathrm{e}^{-\frac{1}{\sigma_z^2}\sum_{m=0}^{N_{\mathrm{FFT}}-1}|r_m - s_m(\tau_0)|^2} \\ &= \underset{\tau_0}{\operatorname{argmax}}\, \ln \frac{1}{\left(\pi\sigma_z^2\right)^{N_{\mathrm{FFT}}}}\, \mathrm{e}^{-\frac{1}{\sigma_z^2}\sum_{m=0}^{N_{\mathrm{FFT}}-1}|r_m - s_m(\tau_0)|^2} \\ &= \underset{\tau_0}{\operatorname{argmin}} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \left|r_m - s_m\left(\tau_0\right)\right|^2 \\ &= \underset{\tau_0}{\operatorname{argmin}} \sum_{m=0}^{N_{\mathrm{FFT}}-1} |r_m|^2 + \left|s_m\left(\tau_0\right)\right|^2 - 2\Re\left\{r_m^* s_m\left(\tau_0\right)\right\}.\end{aligned} \tag{3.57}$$

Since $|r_m|^2$ does not depend on τ_0 and we further assume that $\left|s_m\left(\tau_0\right)\right|^2$ is constant with respect to τ_0, the ML estimate can be obtained by maximizing the real part of the CC

$$C^{\mathrm{CC}}\left(\tau_0\right) = \sum_{m=0}^{N_{\mathrm{FFT}}-1} r_m^* s_m\left(\tau_0\right), \tag{3.58}$$

i.e., we compute

$$\hat{\tau}_0^{\mathrm{CC,\,coh.}} = \underset{\tau_0}{\operatorname{argmax}}\, \Re\left\{C^{\mathrm{CC}}\left(\tau_0\right)\right\}. \tag{3.59}$$

Since the phase information is still included in this approach, we denote it as synchronization using coherent CC, similar as for the differential synchronization algorithms.

If a CFO is present in the signal, the CC is replaced by

$$\begin{aligned}\tilde{C}^{\mathrm{CC}}\left(\tau_0\right) &= \sum_{m=0}^{N_{\mathrm{FFT}}-1} \mathrm{e}^{-\jmath\frac{2\pi\phi_0 m + \phi_{\mathrm{s}}}{N_{\mathrm{FFT}}}}\, r_m^* s_m\left(\tau_0\right) \\ &= \mathrm{e}^{-\jmath\frac{\phi_{\mathrm{s}}}{N_{\mathrm{FFT}}}} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \mathrm{e}^{-\jmath\frac{2\pi\phi_0 m}{N_{\mathrm{FFT}}}}\, r_m^* s_m\left(\tau_0\right).\end{aligned} \tag{3.60}$$

We observe, that the dependency of the CFO is still inside the summation, and hence, CC is not robust against CFOs. Therefore, it has to be ensured that the CFO is estimated in advance, e.g., by differential correlation techniques, before applying CC. Nevertheless, taking the absolute value of the correlation sum results in the non-coherent estimate

$$\hat{\tau}_0^{\mathrm{CC,\,non\text{-}coh.}} = \underset{\tau_0}{\operatorname{argmax}} \left|C^{\mathrm{CC}}\left(\tau_0\right)\right|. \tag{3.61}$$

Synchronization using cross-correlation in frequency domain

Obviously, the CC approach as presented before is limited for estimation of timing offsets on the sampling grid. However, a simple CC on the sampling grid between the received signal and the transmitted signal will at least give a rough estimate. Subsequent interpolation techniques can refine this result. Since in OFDM most of the signal processing steps are applied in the frequency

domain, also this interpolation can be very efficiently realized using the OFDM properties. Hence, with

$$\begin{aligned}
C^{\text{CC, FD}}(\tau_0) &= \sum_{m=0}^{N_{\text{FFT}}-1} r_m^* s_m(\tau_0) \\
&= \sum_{m=0}^{N_{\text{FFT}}-1} \left(\frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{n'=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} R_{n'}^* \, \mathrm{e}^{-\jmath \frac{2\pi n' m}{N_{\text{FFT}}}} \right) \\
&\qquad \left(\frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{n''=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n'' F_{\text{s}} \tau_0} \, S_{n''} \, \mathrm{e}^{\jmath \frac{2\pi n'' m}{N_{\text{FFT}}}} \right) \\
&= \frac{1}{N_{\text{FFT}}} \sum_{n'=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \sum_{n''=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n'' F_{\text{s}} \tau_0} \, R_{n'}^* S_{n''} \sum_{m=0}^{N_{\text{FFT}}-1} \mathrm{e}^{\jmath \frac{2\pi (n''-n') m}{N_{\text{FFT}}}} \\
&= \frac{1}{N_{\text{FFT}}} \sum_{n'=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \sum_{n''=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n'' F_{\text{s}} \tau_0} \, R_{n'}^* S_{n''} N_{\text{FFT}} \delta(n''-n') \\
&= \sum_{n=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n F_{\text{s}} \tau_0} \, R_n^* S_n
\end{aligned} \tag{3.62}$$

we obtain a frequency domain representation of the correlation function depending on the transmitted and received symbols in frequency domain. Contrary to the correlation function in time domain as discussed before, here the timing offset is not limited to the sampling grid allowing a timing estimation with higher resolution. The coherent timing estimation in the frequency domain can then be computed according to

$$\hat{\tau}_0^{\text{CC, FD, coh.}} = \underset{\tau_0}{\operatorname{argmax}} \, \Re \left\{ C^{\text{CC, FD}}(\tau_0) \right\}. \tag{3.63}$$

In case of a present CFO, the correlation function is

$$\begin{aligned}
\tilde{C}^{\text{CC, FD}}(\tau_0) &= \mathrm{e}^{-\jmath \frac{\phi_{\text{s}}}{N_{\text{FFT}}}} \sum_{m=0}^{N_{\text{FFT}}-1} r_m^* s_m(\tau_0) \\
&= \frac{\mathrm{e}^{-\jmath \frac{\phi_{\text{s}}}{N_{\text{FFT}}}}}{N_{\text{FFT}}} \sum_{n'=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \sum_{n''=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n'' F_{\text{s}} \tau_0} \, R_{n'}^* S_{n''} \sum_{m=0}^{N_{\text{FFT}}-1} \mathrm{e}^{\jmath \frac{2\pi (n''-n'-\phi_0) m}{N_{\text{FFT}}}} \\
&= \frac{\mathrm{e}^{-\jmath \frac{\phi_{\text{s}}}{N_{\text{FFT}}}}}{N_{\text{FFT}}} \sum_{n'=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \sum_{n''=\lfloor -(N_{\text{u}}-1)/2 \rfloor}^{\lfloor (N_{\text{u}}-1)/2 \rfloor} \mathrm{e}^{-\jmath 2\pi n'' F_{\text{s}} \tau_0} \, R_{n'}^* S_{n''} N_{\text{FFT}} \delta(n''-n'-\phi_0)
\end{aligned} \tag{3.64}$$

and the corresponding non-coherent timing estimation yields

$$\hat{\tau}_0^{\text{CC, FD, non-coh.}} = \underset{\tau_0}{\operatorname{argmax}} \left| C^{\text{CC, FD}}(\tau_0) \right|. \tag{3.65}$$

Obviously, the orthogonality among the subcarriers is destroyed by the CFO and, hence, inter-carrier interference is introduced.

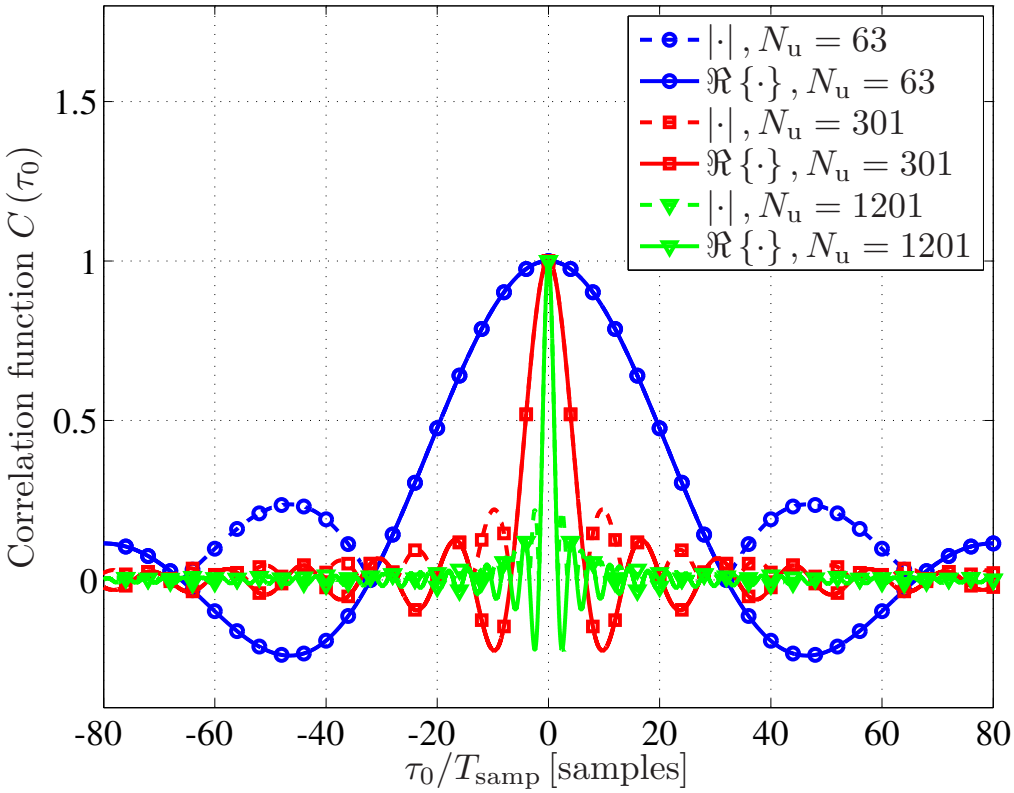


Figure 3.8: Correlation functions using cross-correlation based synchronization

Figure 3.8 depicts the correlation functions for CC, where different numbers of used subcarriers N_u are occupied. The LTE synchronization signals occupy 63 subcarriers corresponding to a bandwidth of $945\,\mathrm{kHz}$. In LTE's $5\,\mathrm{MHz}$-mode 301 subcarriers are occupied and in the $20\,\mathrm{MHz}$-mode 1201 subcarriers [58]. Note that in LTE the zero subcarrier at $n = 0$ is not occupied. We assume that one complete symbol is known to the receiver, and hence, CC with the described approaches can be applied. As expected, the higher the number of subcarriers, and hence the bandwidth, the sharper the correlation peak resulting in more accurate timing estimates. We clearly observe the difference between coherent and non-coherent processing. Finally, we can identify side-peaks which can cause outliers under noise.

The application of OFDM timing estimation using cross-correlation for LTE systems is also discussed in [27]. An analysis of this topic in the context of location determination for GNSSs can be found in [66]. In [67] the focus is on range estimation in multi-user WLAN systems.

Initial access and cell search in LTE

In this section, we briefly outline the initial access of the MS to an LTE system by exploiting the included synchronization signals. Figure 3.9 depicts a typical OFDM frame consisting of pilot symbols for synchronization and channel estimation as well as data symbols. In LTE the so-called primary synchronization signal (PSS) and the secondary synchronization signal (SSS) are specified [58] as deterministic signals exploited for synchronization and initial access [68]. They appear twice per frame and occupy the inner 63 subcarriers. Note that the zero subcarrier at $n = 0$ is not occupied by any data or control symbol in LTE. In this example, the data symbols occupy 181 subcarriers, where as a maximum up to 1201 subcarriers are specified in LTE. Reason for

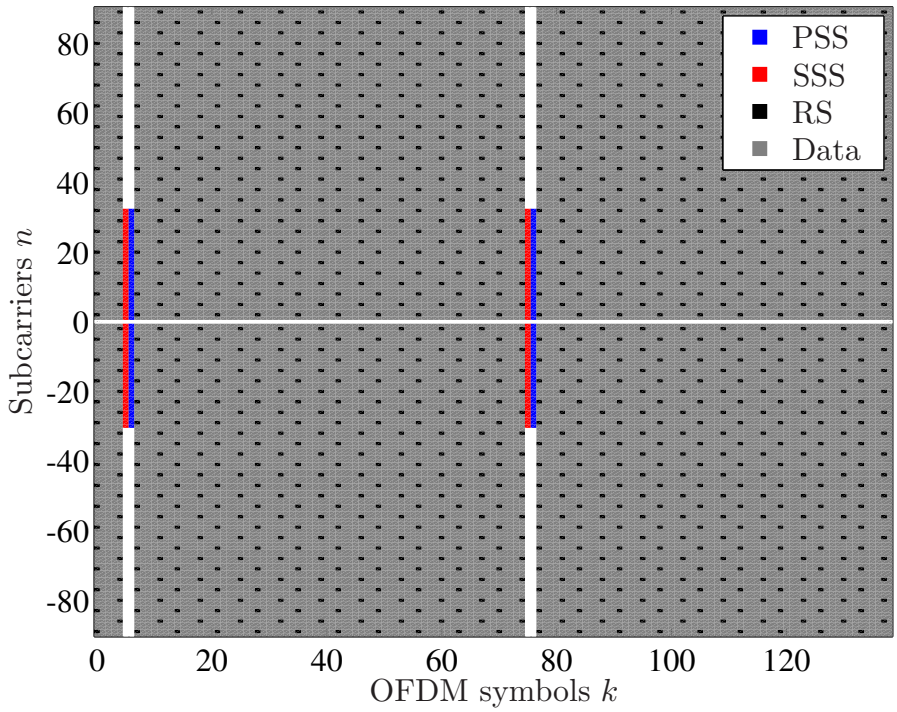


Figure 3.9: LTE frame including pilot and data symbols

the limited bandwidth of the synchronization signals is the flexible bandwidth support by LTE. In that manner, six different spectrum modes are specified in LTE (from $1.4\,\text{MHz}$ to $20\,\text{MHz}$) with the same configuration of the synchronization signals. This allows a flexible exploitation of the available spectrum. Note that for simplification we do not distinguish between information data and control data in Figure 3.9. It is obvious, that in practical mobile radio systems — besides information data symbols dedicated to the user — also control data symbols will be transmitted within the frame. The reference signal (RS) in Figure 3.9 consists of pilot symbols for channel estimation scattered over the whole frame in time and frequency.

The SSS in LTE has a reverse repetitive structure, and thus, can be detected by RDC techniques as discussed before or by non-pilot-aided synchronization algorithms as shown in Section 3.3. In LTE 504 different SSS sequences are specified, i.e., 504 different cell identities are supported. The SSSs consist of so-called Zadoff-Chu sequences with reasonably good auto- and cross-correlation properties. For the generation of the individual sequences we refer to [58]. After the rough timing estimation, e.g., with RDC, the PSS is exploited [64, 69]. For the PSS only three different sequences are specified. On the one hand, the PSS is applied to determine the CFO which is compensated on the time domain signal. On the other hand, with a CC in the frequency domain, the transmitted PSS sequence can be determined. This yields the so-called cell identity group in LTE. With that knowledge, the number of possible SSS sequences reduces to $504/3 = 168$ and by CC in the frequency domain the transmitted SSS can be determined providing the cell identity. Hence, cell identity, frame and symbol timing as well as CFO information is available which concludes the initial access or cell search in LTE.

3.2.3 Cramer-Rao lower bound

The CRLB is the fundamental lower bound on the variance of any estimator [12, 65]. For a scalar parameter τ_0, it states that the variance of any estimator for τ_0 has the lower bound

$$\Xi\left(\tau_0\right) \leq \operatorname{var}\left(\hat{\tau}_0\right) = \mathrm{E}\left\{\left(\mathrm{E}\left\{\hat{\tau}_0\right\} - \hat{\tau}_0\right)^2\right\}. \tag{3.66}$$

The CRLB can be calculated as

$$\Xi\left(\tau_0\right) = \frac{\left(1 + \dfrac{\partial b\left(\tau_0\right)}{\partial \tau_0}\right)^2}{\mathrm{E}\left\{\left(\dfrac{\partial \ln \mathrm{p}\left(\boldsymbol{r}|\tau_0\right)}{\partial \tau_0}\right)^2\right\}}, \tag{3.67}$$

where $b\left(\tau_0\right)$ is a bias function depending on τ_0 and $\mathrm{p}\left(\boldsymbol{r}|\tau_0\right)$ is a probability density function (PDF) for the received samples $\boldsymbol{r}$ conditioned on τ_0. Note that the CRLB is solely a function of the parameter of interest τ_0 and not of its estimate $\hat{\tau}_0$. For unbiased estimation the CRLB simplifies to

$$\Xi\left(\tau_0\right) = \frac{1}{\mathrm{E}\left\{\left(\dfrac{\partial \ln \mathrm{p}\left(\boldsymbol{r}|\tau_0\right)}{\partial \tau_0}\right)^2\right\}} = \frac{1}{-\mathrm{E}\left\{\left(\dfrac{\partial^2 \ln \mathrm{p}\left(\boldsymbol{r}|\tau_0\right)}{\partial^2 \tau_0}\right)\right\}}. \tag{3.68}$$

In the following, we will compute the CRLB for the considered timing estimation problem (cf. [27]) resulting in a performance bound for the derived synchronization algorithms in Section 3.2. Other contributions discuss this derivation in the context of UWB systems [70] and for a comparison between OFDM and pseudo-noise based transmission systems [71]. As we can observe from (3.68), we need to calculate the derivative of the natural logarithm of the conditioned PDF — which is the likelihood function as given in (3.56) — with respect to the delay τ_0, i.e.,

$$\begin{aligned}
\frac{\partial}{\partial \tau_0} \ln \mathrm{p}\left(\boldsymbol{r}|\tau_0\right) &= \frac{\partial}{\partial \tau_0} \ln \frac{1}{\left(\pi \sigma_z^2\right)^{N_{\mathrm{FFT}}}} \mathrm{e}^{-\frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \left|r_m - s_m\left(\tau_0\right)\right|^2} \\
&= -\frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \frac{\partial}{\partial \tau_0} \left|r_m - s_m\left(\tau_0\right)\right|^2 \\
&= -\frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \frac{\partial}{\partial \tau_0} \left(\left|r_m\right|^2 - r_m s_m^*\left(\tau_0\right) - r_m^* s_m\left(\tau_0\right) + \left|s_m\left(\tau_0\right)\right|^2\right) \\
&= \frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} r_m \frac{\partial}{\partial \tau_0} s_m^*\left(\tau_0\right) + r_m^* \frac{\partial}{\partial \tau_0} s_m\left(\tau_0\right) \\
&\qquad\qquad - s_m\left(\tau_0\right) \frac{\partial}{\partial \tau_0} s_m^*\left(\tau_0\right) - s_m^*\left(\tau_0\right) \frac{\partial}{\partial \tau_0} s_m\left(\tau_0\right) \\
&= \frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} \left(r_m - s_m\left(\tau_0\right)\right) \frac{\partial}{\partial \tau_0} s_m^*\left(\tau_0\right) + \left(r_m^* - s_m^*\left(\tau_0\right)\right) \frac{\partial}{\partial \tau_0} s_m\left(\tau_0\right) \\
&= \frac{1}{\sigma_z^2} \sum_{m=0}^{N_{\mathrm{FFT}}-1} z_m \frac{\partial}{\partial \tau_0} s_m^*\left(\tau_0\right) + z_m^* \frac{\partial}{\partial \tau_0} s_m\left(\tau_0\right).
\end{aligned} \tag{3.69}$$

Squaring this equation and taking the expectation with respect to the noise yields

$$\begin{aligned}
\mathrm{E}\left\{\left|\frac{\partial}{\partial\tau_0}\ln \mathrm{p}\left(\boldsymbol{r}|\tau_0\right)\right|^2\right\} &= \mathrm{E}\left\{\frac{1}{\sigma_z^4}\sum_{m=0}^{N_{\mathrm{FFT}}-1}\left|z_m\frac{\partial}{\partial\tau_0}s_m^*\left(\tau_0\right)+z_m^*\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\right|^2\right\} \\
&= \frac{2}{\sigma_z^4}\sum_{m=0}^{N_{\mathrm{FFT}}-1}\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\mathrm{E}\left\{z_m z_m^*\right\}\frac{\partial}{\partial\tau_0}s_m^*\left(\tau_0\right) \\
&= \frac{2}{\sigma_z^2}\sum_{m=0}^{N_{\mathrm{FFT}}-1}\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\frac{\partial}{\partial\tau_0}s_m^*\left(\tau_0\right) \\
&= \frac{2}{\sigma_z^2}\sum_{m=0}^{N_{\mathrm{FFT}}-1}\left|\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\right|^2 .
\end{aligned} \tag{3.70}$$

Using this result in (3.68) gives the CRLB for timing estimation

$$\Xi\left(\tau_0\right) = \frac{\sigma_z^2}{2\sum_{m=0}^{N_{\mathrm{FFT}}-1}\left|\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\right|^2}. \tag{3.71}$$

This expression includes the squared derivative of the time domain signal in the denominator. Hence, the CRLB will decrease for signals with higher frequency components or bandwidths. For an efficient calculation of this CRLB we can apply the OFDM properties — similar as for the CC algorithm in frequency domain described in Section 3.2. We start with the frequency domain representation of the delayed signal as defined in (3.39). Its derivative with respect to the timing offset τ_0 can simply be computed as

$$\begin{aligned}
\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right) &= \frac{\partial}{\partial\tau_0}\frac{1}{\sqrt{N_{\mathrm{FFT}}}}\sum_{n=\lfloor-(N_{\mathrm{u}}-1)/2\rfloor}^{\lfloor(N_{\mathrm{u}}-1)/2\rfloor}\mathrm{e}^{-\jmath 2\pi nF_{\mathrm{s}}\tau_0}\,S_n\,\mathrm{e}^{\jmath\frac{2\pi nm}{N_{\mathrm{FFT}}}} \\
&= \frac{1}{\sqrt{N_{\mathrm{FFT}}}}\sum_{n=\lfloor-(N_{\mathrm{u}}-1)/2\rfloor}^{\lfloor(N_{\mathrm{u}}-1)/2\rfloor}-\jmath 2\pi nF_{\mathrm{s}}\,\mathrm{e}^{-\jmath 2\pi nF_{\mathrm{s}}\tau_0}\,S_n\,\mathrm{e}^{\jmath\frac{2\pi nm}{N_{\mathrm{FFT}}}} .
\end{aligned} \tag{3.72}$$

Squaring this expression and a summation over N_{FFT} available consecutive samples yields

$$\begin{aligned}
\sum_{m=0}^{N_{\text{FFT}}-1}\left|\frac{\partial}{\partial\tau_0}s_m\left(\tau_0\right)\right|^2 &= \frac{4\pi^2F_{\text{s}}^2}{N_{\text{FFT}}}\sum_{m=0}^{N_{\text{FFT}}-1}\sum_{n'=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor}\sum_{n''=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} \\
&\qquad n'\,\mathrm{e}^{-\jmath2\pi n'F_{\text{s}}\tau_0}\,S_{n'}\,\mathrm{e}^{\jmath\frac{2\pi n'm}{N_{\text{FFT}}}}\,n''\,\mathrm{e}^{\jmath2\pi n''F_{\text{s}}\tau_0}\,S_{n''}^*\,\mathrm{e}^{-\jmath\frac{2\pi n''m}{N_{\text{FFT}}}} \\
&= \frac{4\pi^2F_{\text{s}}^2}{N_{\text{FFT}}}\sum_{n'=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor}\sum_{n''=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} n'n''\,\mathrm{e}^{\jmath2\pi(n''-n')F_{\text{s}}\tau_0}\,S_{n'}S_{n''}^* \\
&\qquad \sum_{m=0}^{N_{\text{FFT}}-1}\mathrm{e}^{-\jmath\frac{2\pi(n''-n')m}{N_{\text{FFT}}}} \\
&= \frac{4\pi^2F_{\text{s}}^2}{N_{\text{FFT}}}\sum_{n'=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor}\sum_{n''=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} n'n''\,\mathrm{e}^{\jmath2\pi(n''-n')F_{\text{s}}\tau_0}\,S_{n'}S_{n''}^*N_{\text{FFT}}\delta\left(n''-n'\right) \\
&= 4\pi^2F_{\text{s}}^2\sum_{n=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} n^2\left|S_n\right|^2.
\end{aligned} \tag{3.73}$$

Combining this result with (3.71) gives the CRLB for timing estimation in frequency domain representation

$$\Xi\left(\tau_0\right) = \frac{\sigma_z^2}{8\pi^2F_{\text{s}}^2\sum\limits_{n=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} n^2\left|S_n\right|^2}. \tag{3.74}$$

Note that the dependency of the CRLB on the timing offset τ_0 is implicitly included in the subcarrier contributions S_n. This is consistent as a timing offset directly relates to a linear phase shift in the frequency domain signal according to the FFT operation.

Observing (3.74), several interesting conclusions can be drawn as follows:

- If we define the mean subcarrier SNR as

$$\gamma = \frac{\sum\limits_{n=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor}\left|S_n\right|^2}{N_{\text{u}}\sigma_z^2} \tag{3.75}$$

  and the normalized mean squared bandwidth as

$$\beta^2 = \frac{F_{\text{s}}^2\sum\limits_{n=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor} n^2\left|S_n\right|^2}{\sum\limits_{n=\lfloor-(N_{\text{u}}-1)/2\rfloor}^{\lfloor(N_{\text{u}}-1)/2\rfloor}\left|S_n\right|^2}, \tag{3.76}$$

  the CRLB can be rewritten as

$$\Xi\left(\tau_0\right) = \frac{1}{8\pi^2N_{\text{u}}\gamma\,\beta^2}. \tag{3.77}$$

Hence, the derived CRLB decreases linearly with the SNR and quadratically with the occupied bandwidth of the signal. This is in line with the results from other investigations (e.g., [10, 12]).

- For the definition of the OFDM system model in Section 3.1 and for the derivation of the CRLB it was explicitly assumed, that the occupied subcarriers are symmetrically distributed around the zero subcarrier at $n = 0$. For communications tasks in the frequency domain, very often the circular OFDM properties are exploited. This has the effect, that — mainly due to notational convenience — the relevant subcarriers can be shifted to the positive domain, i.e., to $n = 0, 1, \ldots, N_\mathrm{u} - 1$, without loss of any generality. However, for the targeted synchronization problem and the calculation of the CRLB this symmetry around the zero subcarrier is of fundamental importance. In a practical realization it ensures that the received signal can be downconverted to the baseband within the dedicated receiver bandwidth. A shift to the positive domain would require a receiver with around a doubled bandwidth (cf. [72, 73]).

- If the symbols have equal energy in the frequency domain ($|S_n| = |S|$), e.g., for phase shift keying modulation schemes, the CRLB simplifies to

$$\Xi(\tau_0) = \begin{cases} \dfrac{3\sigma_z^2}{2\pi^2 F_\mathrm{s}^2 |S|^2 N_\mathrm{u} (N_\mathrm{u}^2 - 1)}, & \text{for } N_\mathrm{u} \text{ odd}, \\ \dfrac{3\sigma_z^2}{2\pi^2 F_\mathrm{s}^2 |S|^2 N_\mathrm{u} (N_\mathrm{u}^2 + 2)}, & \text{for } N_\mathrm{u} \text{ even}. \end{cases} \tag{3.78}$$

  This results in the fact, that all synchronization signals in LTE provide the same CRLB since all synchronization symbols have the same energy in the frequency domain. Hence, the CRLB is independent of the chosen synchronization signal. Note that this does not necessarily yield the same timing estimation results if synchronization algorithms are applied. As the CRLB provides only a local performance bound, effects like converging to side-lobes of the correlation function are not considered in its calculation.

- The zero subcarrier at $n = 0$ does not contribute to the CRLB. This is not surprising, since timing estimation with a constant signal would not produce any useful results.

- The spectral properties have a direct impact on the CRLB. Not only the number of used subcarriers, also their individual power and distribution affects the CRLB. The more power is available at the band edges, the smaller the CRLB will be. Hence, from a timing estimation point of view, it is beneficial to distribute the power to the band edges as much as possible. This was also part of the investigations in [72], where this distribution was optimized under power constraints.

- The result in (3.74) for the AWGN channel can easily be extended to arbitrary channels. To do so, the channel coefficients in frequency domain H_n can simply be multiplied with the transmitted symbols on each subcarrier, and hence, change the spectral properties of the transmitted signal. We obtain

$$\Xi(\tau_0) = \frac{\sigma_z^2}{8\pi^2 F_\mathrm{s}^2 \sum\limits_{n=\lfloor -(N_\mathrm{u}-1)/2 \rfloor}^{\lfloor (N_\mathrm{u}-1)/2 \rfloor} n^2 |H_n S_n|^2}. \tag{3.79}$$

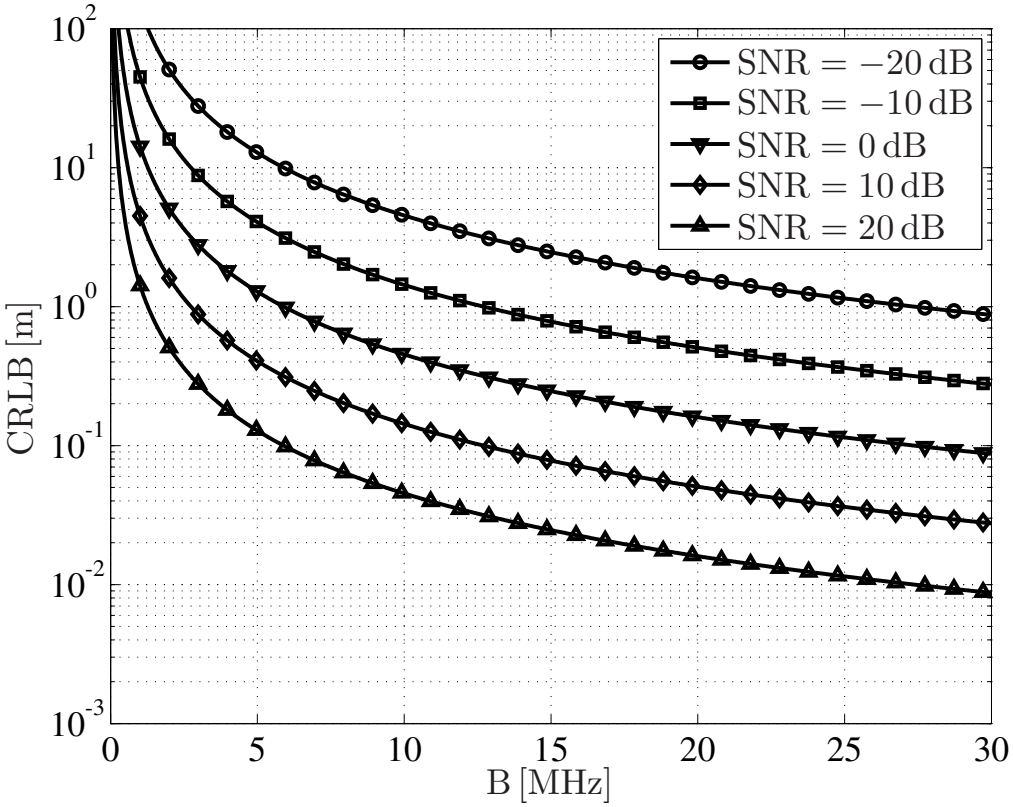


Figure 3.10: CRLB over bandwidth for different SNRs

- So far, we assumed that the subcarriers are allocated contiguously as a block in the spectrum. Beyond that, the result in (3.74) also holds for non-contiguously allocated subcarriers, e.g., for the scattered pilot symbols used for channel estimation in LTE.

In Figure 3.10 the CRLB is plotted over the occupied bandwidth B as defined in (3.4) for different SNR values between -20 dB and 20 dB. For convenience and without loss of any generality, we define the CRLB in meters as the square root of (3.74) multiplied with the speed of light c_0, i.e.,

$$\text{CRLB} = c_0 \sqrt{\Xi\left(\tau_0\right)}, \tag{3.80}$$

since location estimation is the target process. As discussed before, LTE's synchronization signals occupy 63 subcarriers, and hence, have a bandwidth of 945 kHz. This results in a CRLB of around 4.3 m for 10 dB and of around 1.5 m for 20 dB. Therefore, for reasonable SNRs the achievable location estimation accuracy is in the order of a few meters. If a complete OFDM symbol is known in the 20 MHz-mode of LTE, this accuracy improves to an order of a few centimeters.

3.2.4 Simulation results

Single-link synchronization

In the following, we investigate the single-link synchronization performance in different scenarios for LTE-conform parameters. We assume a downlink transmission from BS to MS of $N_\text{s} = 140$ symbols per frame, where the symbols are OFDM modulated with an FFT length of $N_\text{FFT} = 2048$. For the CP $N_\text{CP} = 144$ samples are used. As mentioned before, the synchronization channels

Parameter name	Parameter symbol	Value
FFT length	N_{FFT}	2048
Maximum number of used subcarriers for data	N_{u}^{d}	1201
Number of used subcarriers for PSS/SSS	N_{u}^{p}	63
Number of samples for CP	N_{CP}	144
Number of symbols per frame	N_{s}	140
Subcarrier spacing	F_{s}	$15\,\text{kHz}$
Convolutional code	$\mathcal{G}$	$\{(171)_8, (133)_8\}$
Interleaver	$\mathit{\Pi}$	Random
Modulation alphabet	$\mathcal{S}$	QPSK

Table 3.1: LTE parameters

occupy $N_{\text{u}}^{\text{p}} = 63$ subcarriers, whereas the data symbols occupy $N_{\text{u}}^{\text{d}} = 1201$ subcarriers. The subcarrier spacing is $15\,\text{kHz}$. For the generation of the data symbols we apply convolutional coding with the generator polynomial $\mathcal{G} = \{(171)_8, (133)_8\}$ in octal representation, random interleaving, and quadrature phase shift keying (QPSK) modulation. These LTE core parameters are concluded in Table 3.1. For synchronization with the algorithms derived in Section 3.2.2 we exploit the SSS as it provides 504 different sequences, and hence, will also be suited for later investigated multi-link approaches which are required for location estimation. Note that for DC based synchronization we flipped the second half of the SSS in time domain to obtain the required repetitive structure. As performance metric we choose the root mean squared error (RMSE)

$$\text{RMSE} = c_0 \sqrt{\text{E}\left\{(\tau_0 - \hat{\tau}_0)^2\right\}} \geq \text{CRLB}, \tag{3.81}$$

which is a measure for the average synchronization or timing estimation error given in meters, where it was averaged over several noise realizations, timing offsets, and different synchronization sequences. Performance bound for the RMSE is the CRLB as computed in Section 3.2.3.

Figure 3.11 shows the RMSE over SNR for the different synchronization schemes as outlined in Section 3.2.2 using an AWGN channel and no CFO for the simulations. We further assumed a search space in the interval of $[-N_{\text{CP}}, N_{\text{CP}}]$ samples around the timing offset. This reduced search space is reasonable as for an established communication an initial timing estimate inside the CP is required anyhow. We observe the following:

- The DC synchronization algorithms (coherent and non-coherent) are not suitable for precise timing, and hence, location estimation. There is an error floor for high SNRs at around $150\,\text{m}$ which corresponds to around 15 samples in the considered LTE setup. For communications purposes this accuracy is usually acceptable, however, for location estimation even for high SNRs the DC based estimates are not useful. Furthermore, we observe that there is no major difference between coherent and non-coherent processing. Main reason for the limited performance of DC is the difficult detection of the end of the correlation function's plateau.

- For low SNR values, the estimates for RDC, CC, and CC in frequency domain are uniformly distributed in the interval determined by the search space resulting in

$$\text{RMSE}|_{\gamma \to -\infty} = \frac{1}{\sqrt{3}} \frac{c_0 N_{\text{CP}}}{N_{\text{FFT}} F_{\text{s}}} \approx 811.34\,\text{m}. \tag{3.82}$$

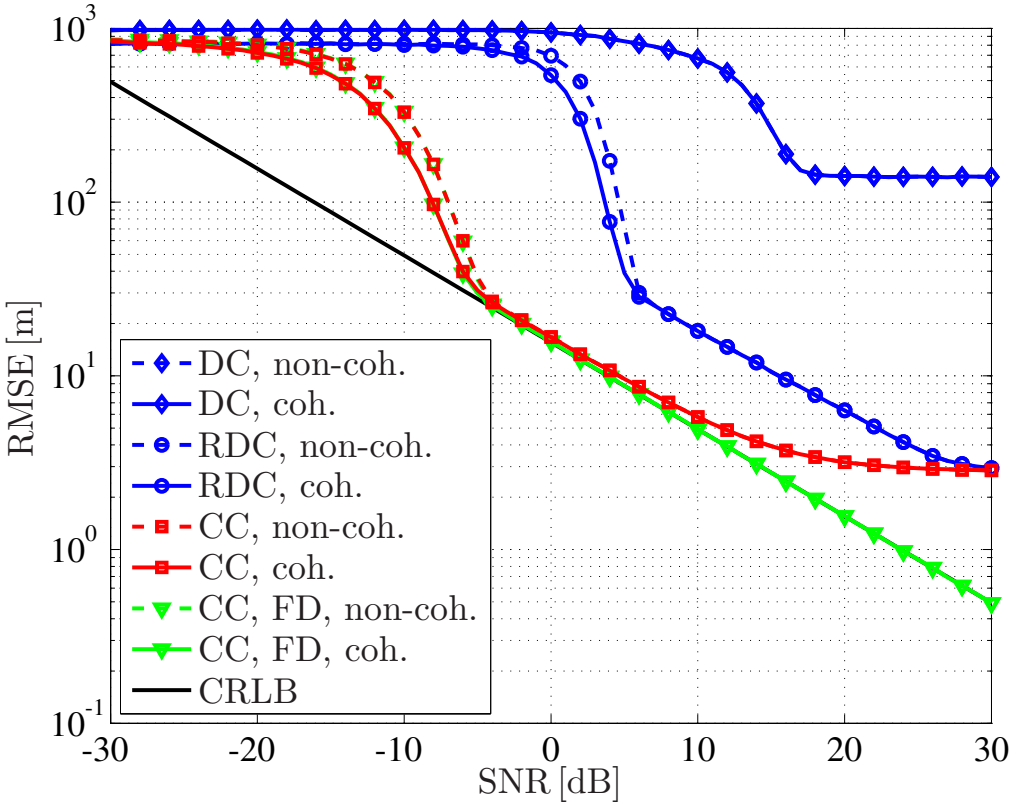


Figure 3.11: RMSE over SNR using different synchronization algorithms, AWGN channel

- For high SNR values, the RMSE of RDC and CC is limited by the sampling time as a resolution of sub-sample timing offsets is not possible with these algorithms. Hence, the residual error is given as the square root of the variance for a uniform distribution of the length of one sample, i.e.,

$$\mathrm{RMSE}|_{\gamma\to\infty} = \frac{1}{\sqrt{12}} \frac{c_0}{N_{\mathrm{FFT}} F_\mathrm{s}} \approx 2.82\,\mathrm{m}. \tag{3.83}$$

- As RDC provides a clear correlation peak compared to the plateau of DC (cf. Figure 3.7) it is the more accurate differential correlation algorithm. For an SNR smaller than 0 dB no precise timing estimation is possible. Between 0 dB and 6 dB there is a waterfall region, where the coherent approach outperforms the non-coherent approach by around 1 dB. For higher SNRs up to 30 dB there still is a gap of around 12 dB between the achieved performance and the CRLB.

- For the CC approaches in time domain the overall performance improves by around 10 dB compared to RDC, however, here the complete synchronization sequence has to be known in advance. We observe a similar behavior in the waterfall region (coherent outperforms non-coherent estimation). For SNRs between -5 dB and 5 dB the RMSE is very close to the CRLB, however, as time domain CC is similarly restricted to the sampling grid, also here an error floor is present.

- In case of CC in the frequency domain, the restriction of estimates on the sampling grid can be overcome. With this approach the CRLB can be achieved for SNRs larger than -5 dB

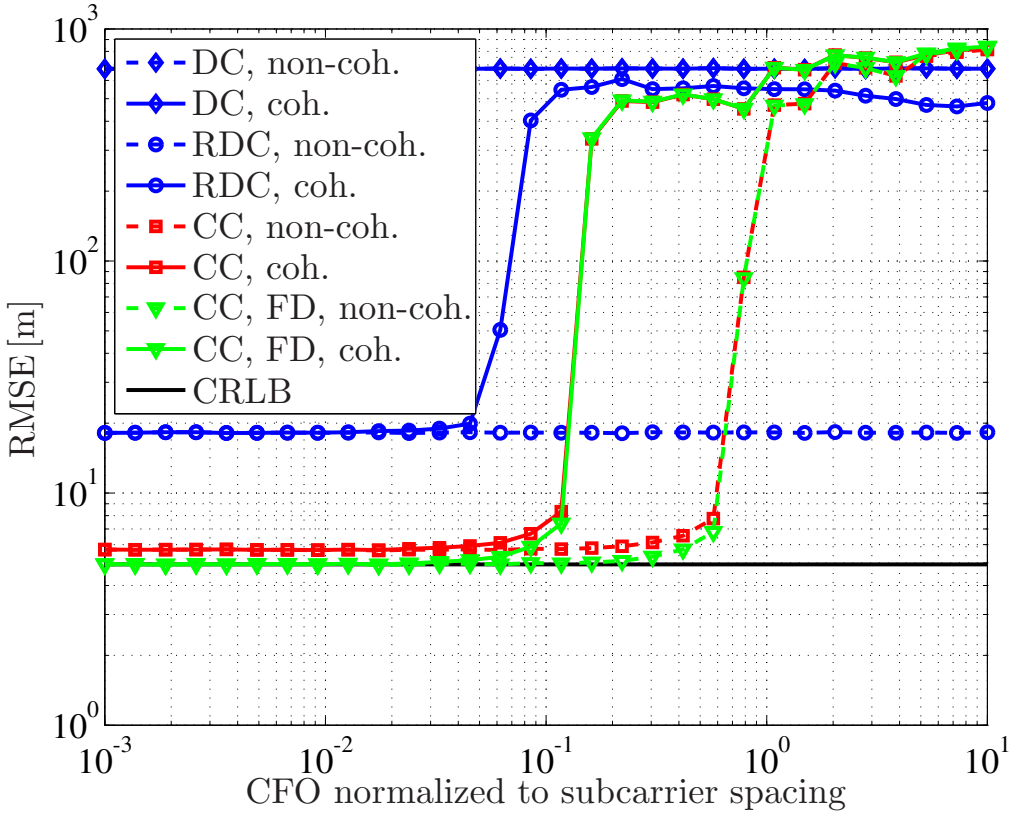


Figure 3.12: RMSE over CFO using different synchronization algorithms at SNR of 10 dB, AWGN channel

and, hence, it is efficient as expected for this ML based algorithm. For SNRs smaller than -5 dB, the performance corresponds to the CC approach in time domain.

- The correlation functions as shown in Figure 3.8 have clear side lobes besides the main lobe depending on, e.g., bandwidth and processing type (coherent or non-coherent). Therefore, for reasonably low SNR values it is possible to run into the side lobes during the maximum search. Hence, in the threshold or waterfall region between low and high SNR (between -5 dB and -20 dB for CC in frequency domain) the CRLB is no longer a reasonably good performance bound as CRLB solely holds locally around the main lobe. Better approximations or bounds for this region can be found in [74, 75, 76].

- The side lobes of the correlation function for non-coherent processing are higher and closer to the main lobe compared to coherent processing (cf. Figures 3.7 and 3.8). Therefore, for a reasonably low SNR the probability to run into these side lobes is higher, and hence, coherent processing outperforms non-coherent processing in this region. The performance for high SNR is equal as the correlation function is nearly identical around the main lobe of the correlation function for both procedures.

- The limited search space has the effect that a priori knowledge is introduced in the estimation process. Therefore, for low SNR values the RMSE can outperform the CRLB. To include this a priori information, the posterior CRLB or Bayesian CRLB could be applied [65].

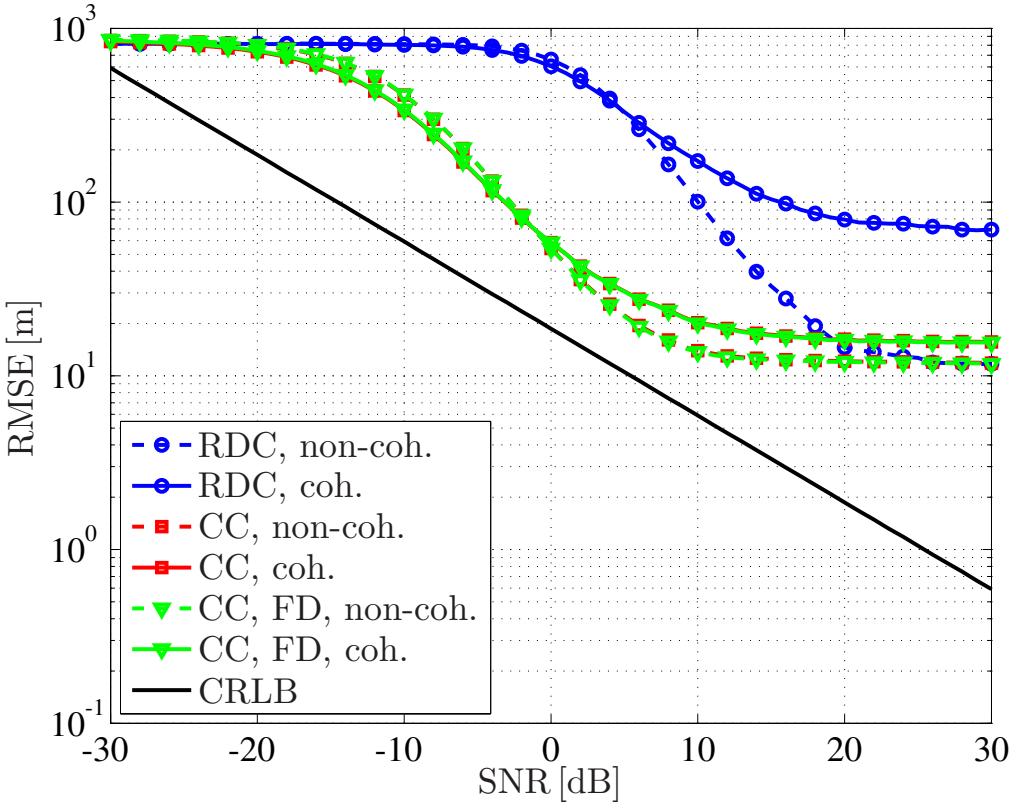


Figure 3.13: RMSE over SNR using different synchronization algorithms, multipath LOS channel

Next, we investigate the performance of the different synchronization algorithms under CFOs at an SNR of $10\,\mathrm{dB}$. Figure 3.12 depicts the RMSE over the maximum CFO, i.e., the CFO is uniformly distributed within the interval $[-\phi_0/F_\mathrm{s}, \phi_0/F_\mathrm{s}]$. First of all, we observe that the non-coherent differential approaches for DC and RDC are — as expected — not depending on the CFO, and hence, are well suited for initial timing estimation. We further observe, that also the CC approaches are suitable, if the residual CFO is not too large. For coherent processing a normalized CFO of 0.1 is still acceptable, whereas for non-coherent processing a CFO of 0.6 still allows reasonable timing estimates. Since for communications purposes (especially for channel estimation) a residual CFO of around 0.02 is targeted, the consideration of CFO effects is no longer in the main focus of this thesis. Therefore, in the following we assume that the CFO is perfectly compensated or the residual CFO is in a range, where CC approaches are not affected.

The AWGN channel considered so far gives a performance limit for the synchronization algorithms. For a realistic performance evaluation, we include more realistic multipath channel models in the simulation chain. We choose a LOS and a NLOS version of the WINNER channel models for urban scenarios (C2) as specified in [77]. The main parameters for these models can also be found in Appendix A.1. Figure 3.13 depicts the RMSE over SNR for different synchronization algorithms if the signal is transmitted over the LOS channel. First of all, we observe that the timing estimates are biased if the channel is not known. This results in an overall performance loss compared to the AWGN channel. We further can see that the non-coherent approaches outperform the coherent approaches for reasonably high SNRs, i.e., they are more robust against multipath fading channels. The coherent approaches exploit the real part of the correlation function which requires a matched phase relation between the received signal (affected by channel and noise) and

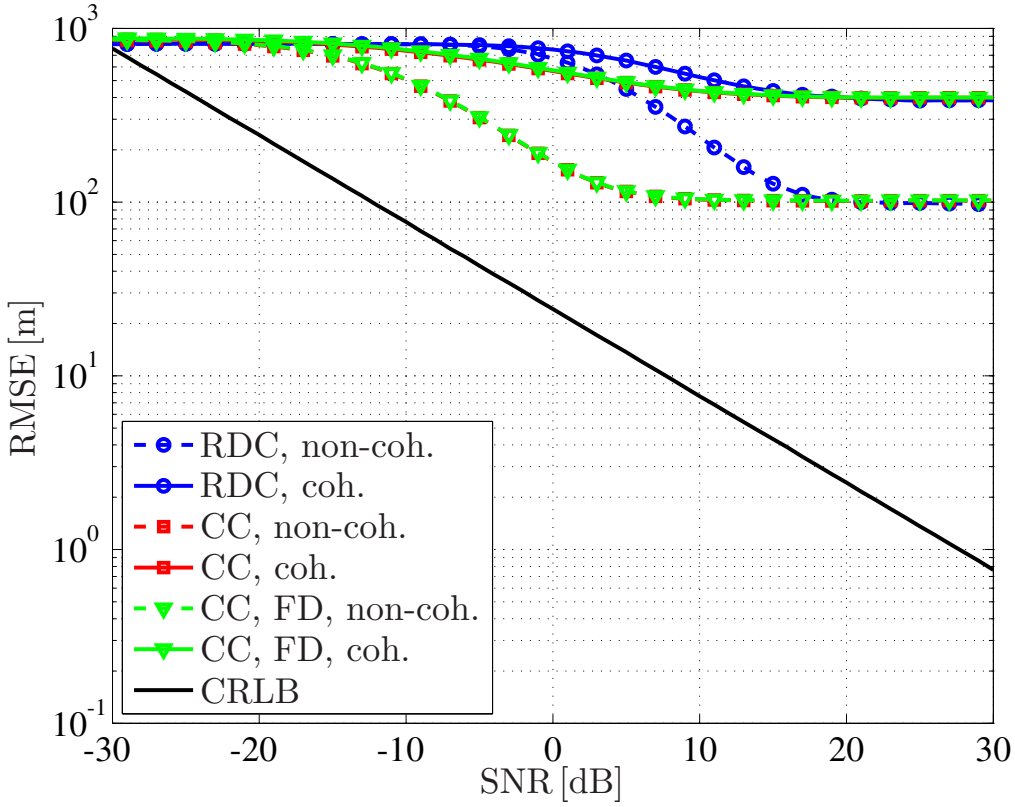


Figure 3.14: RMSE over SNR using different synchronization algorithms, multipath NLOS channel

the transmitted signal. If the channel is not known, the non-coherent approaches exploiting the absolute value of the correlation function yield more reliable results. Nevertheless, as the constant direct path in the LOS channel is comparatively strong (compared to the multipath components), the LOS channel is quite close to the AWGN channel. This has the effect that the performance loss of coherent processing is limited — at least for the CC algorithms. For high SNRs, non-coherent RDC and CC result in an average error of around 11 m. For the coherent approach, 18 m can be achieved by CC and 70 m for RDC. The performance of CC in time and frequency domain is nearly the same as a sub-sample resolution is not relevant for this setup. Comparing the CRLBs between the LOS channel and AWGN, we observe that for the LOS channel a 2 dB performance loss must be accepted.

Figure 3.14 depicts the corresponding curves for the NLOS channel model. Here, the performance losses compared to AWGN are even higher. With non-coherent synchronization algorithms an average error of around 100 m can be achieved for high SNRs. For coherent processing with an error of around 400 m no useful information for location estimation can be gained. Compared to AWGN, the CRLB for the NLOS channel shifts by around 4 dB.

Finally, Figure 3.15 depicts the RMSE performance over SNR for different channels, if perfect channel knowledge is available. In that case, the synchronization algorithms can be matched to the channel by a convolution of the transmitted signal with the channel impulse response (for time domain processing) or by multiplying the transmitted symbols with the samples of the channel transfer function (for frequency domain processing). Here, we consider coherent CC in the frequency domain. We observe, that also for the LOS and NLOS channels the CRLB can be achieved if appropriate channel knowledge is available.

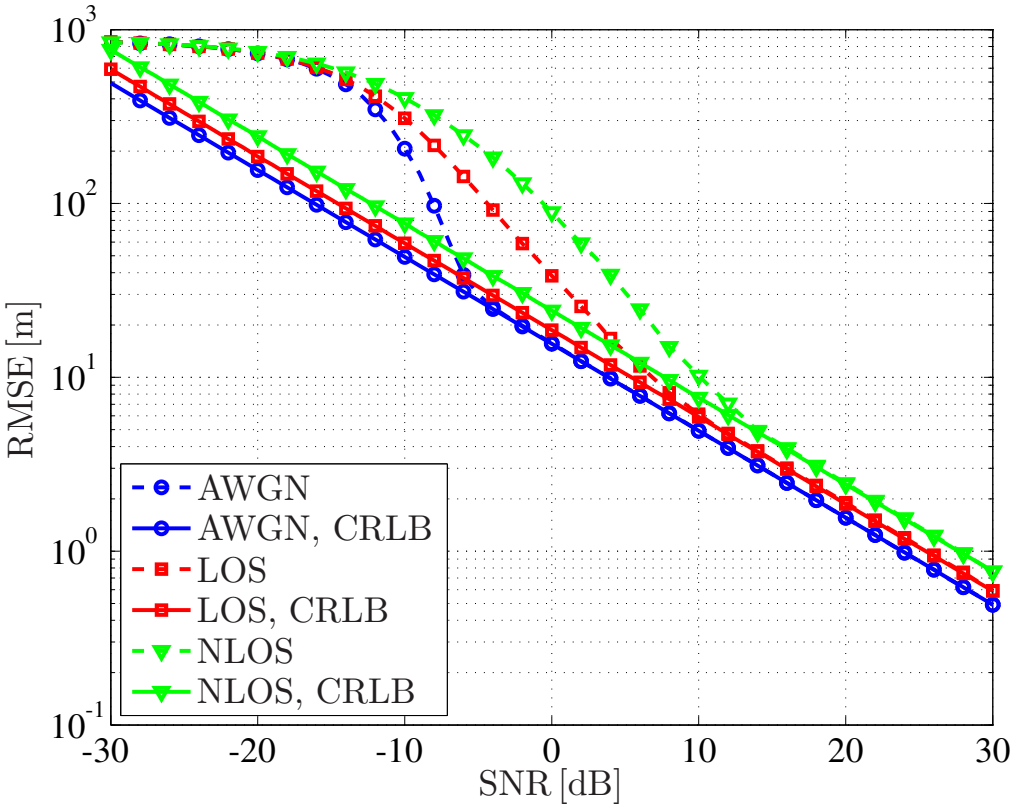


Figure 3.15: RMSE over SNR for different channels, perfect channel estimation, synchronization with CC in frequency domain

Parameter name	Value
Carrier frequency	2 GHz
Cell layout	Hexagonal, inter-BS distance of 750 m
BS transmit power	43 dBm
BS antenna model	Three sectors, 3 dB-beamwidth of 70 degree, 14 dBi
MS antenna model	Omnidirectional, 0 dBi
MS noise figure	7 dB
Doppler frequency	According to MS speed of 2 m/s

Table 3.2: Cellular system parameters

Multi-link synchronization

Considering single-link synchronization as shown before allows a comprehensive study of the various timing estimation algorithms and absolute performance bounds for different SNR values. For the targeted location determination process simultaneous timing estimation with at least three BSs has to be performed at the MS. Therefore, the geometric relation between MS and BS locations has to be taken into account to derive realistic values of the relevant parameters like SNRs or relative time offsets. To do so, in this section we investigate the performance of the OFDM timing estimation algorithms in a multi-link cellular environment. We assume a two-dimensional cell layout with a distance between the BSs of 750 m and a hexagonal structure of the cells. The BSs transmit with a transmit power of 43 dBm at a carrier frequency of 2 GHz. Each BS has three antenna elements with a 3 dB-beamwidth of 70 degree and an overall antenna gain of 14 dBi, i.e., three individual cells are served per BS. The MS has an omnidirectional antenna with 0 dBi, a noise figure of 7 dB, and moves with a speed of 2 m/s. This is a typical value for pedestrian applications. The described cellular system parameters are concluded in Table 3.2. The small scale and large scale channel parameters for AWGN, LOS, and NLOS conditions are outlined in Appendix A.1. They make use of the WINNER channel models for urban scenarios (C2) as specified in [77].

Figure 3.16 shows a geometric overview of the cellular network. We investigate the links between the BSs of interest (BS 1, BS 2, and BS 3) and the MS, where different locations of the MS are considered as depicted. Since all BSs can transmit in the same frequency band, a strong inter-cell interference is present in the network. A measure for the interference is the signal-to-interference-and-noise ratio (SINR) defined as

$$\gamma_\mu = \frac{P_\mu}{\sigma_z^2 + \sum\limits_{\substack{\nu=1 \\ \nu \neq \mu}}^{3N_{\text{BS}}} P_\nu}, \quad \mu = 1, 2, \ldots, 3N_{\text{BS}}, \tag{3.84}$$

where three cells per BS are assumed. The received power from a BS with cell μ (including transmit power, transmit antenna gain, and large scale channel effects as defined in Table 3.2 and Appendix A.1) is denoted as P_μ. The noise power σ_z^2 is determined by the noise figure of the MS and the sampling time T_{samp}.

Figure 3.17 shows the resulting SINR for the AWGN channel over the individual MS locations as depicted in Figure 3.16. These locations are determined by the distance between MS and BS 1. Note that we always choose the strongest cell of a specific BS as the desired signal. BS 1 as the overall strongest BS will be the serving BS of the MS in this setup. First of all, we can observe

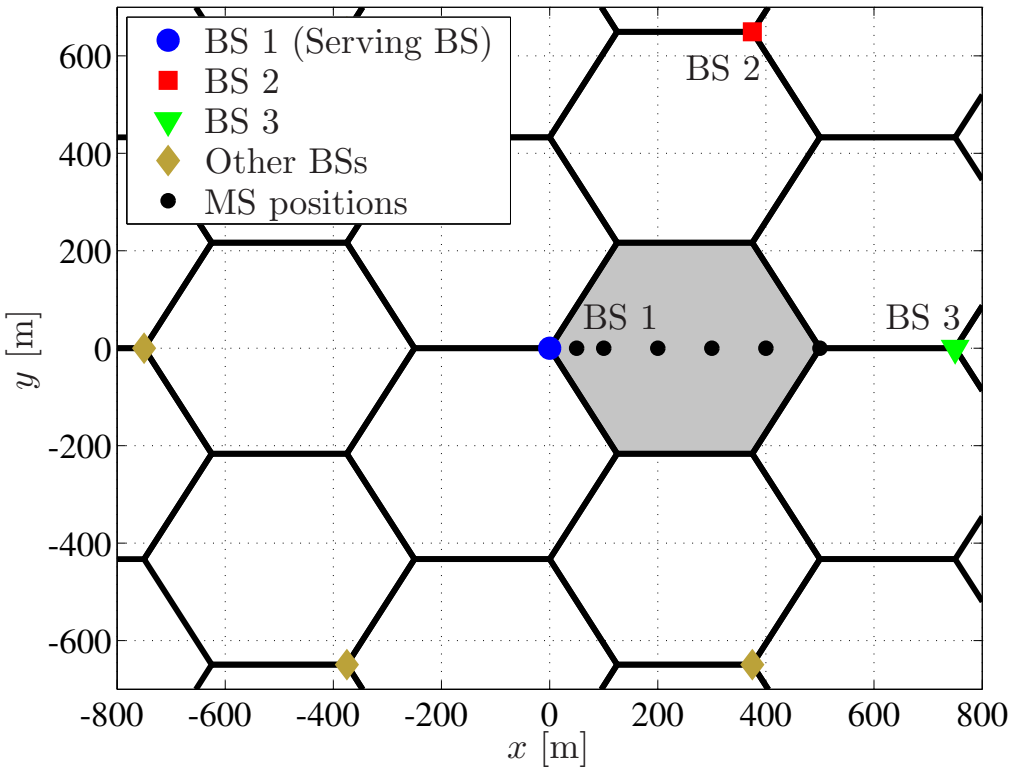


Figure 3.16: Cellular network

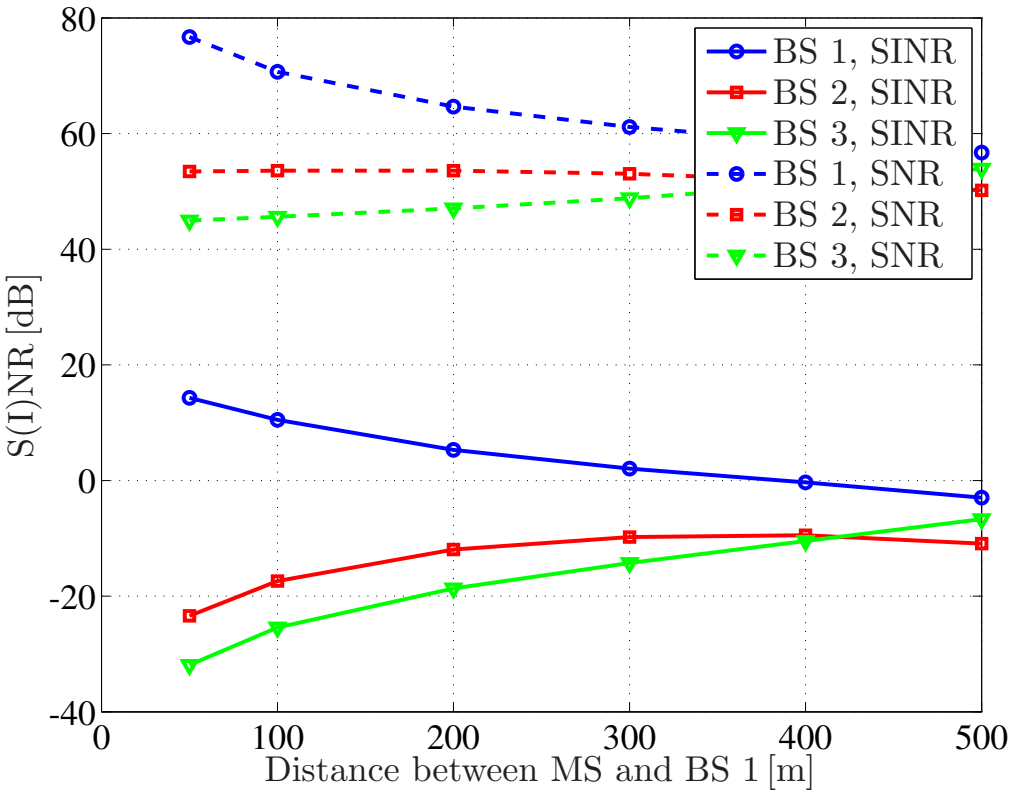


Figure 3.17: SINR and SNR over distance for a cellular network, AWGN channel

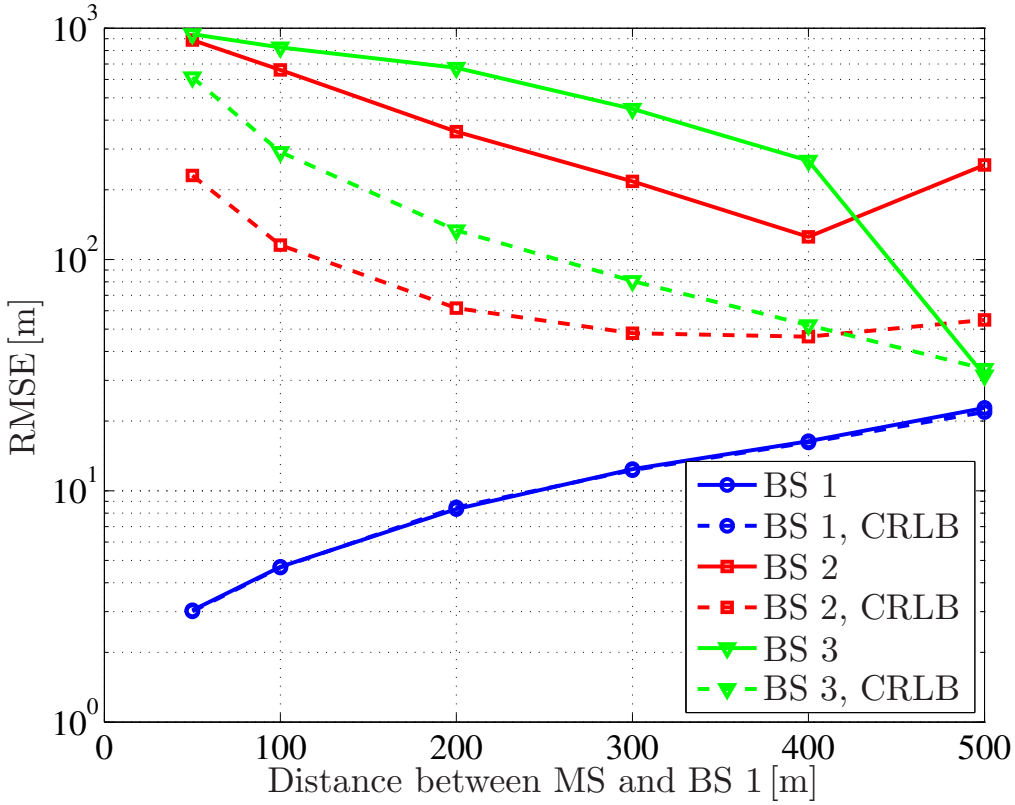


Figure 3.18: RMSE over distance for multi-link synchronization with different BSs, AWGN channel

the dependency of the SINR from the various MS locations in the network. For instance, the SINR difference between a MS location close to BS 1 (distance of 50 m) and at the cell edge (distance of 500 m) is around 18 dB for considering the serving BS 1. The synchronization channel of this BS can be received with an SINR between 16 dB and −2 dB which should allow a reasonably good timing estimation. The reception of BS 2 and BS 3 will be more challenging as the interference from the serving BS makes a reasonably good timing estimation with the out-of-cell BSs difficult. Especially close to the serving BS, the SINR can be below −30 dB for BS 3, making a detection of this BS with the described methods impossible. At the cell edge (distance of 500 m) the SINR values of the three considered BSs have — as expected — a similar level. Additionally depicted in Figure 3.17 are the corresponding SNR values, i.e., the interference-free situation assuming that all BSs transmit in different frequency bands. Here, the overall signal level is much higher and, e.g., BS 1 can be received with an SNR between 77 dB and 57 dB. Also the out-of-cell reception of BS 2 and BS 3 is possible with SNR values higher than 40 dB. Hence, the network is strongly interference-limited and only the serving BS can be received with sufficient quality for the desired LTE setup with a frequency reuse of one.

In Figure 3.18 the corresponding RMSE of the timing estimation is depicted over the distance between MS and BS 1 for transmission over an AWGN channel. As synchronization algorithm the coherent cross-correlation approach in the frequency domain was exploited. For timing estimation with BS 1 an acceptable performance can be achieved. The resulting RMSE is around 2 m close to BS 1 and around 20 m at the cell edge. Additionally, the corresponding CRLB is plotted. For computing the CRLB, the SNR γ in (3.77) was replaced by the SINR γ_μ as defined in (3.84). Instead of AWGN, the impairment in the considered cellular network also includes a superposition

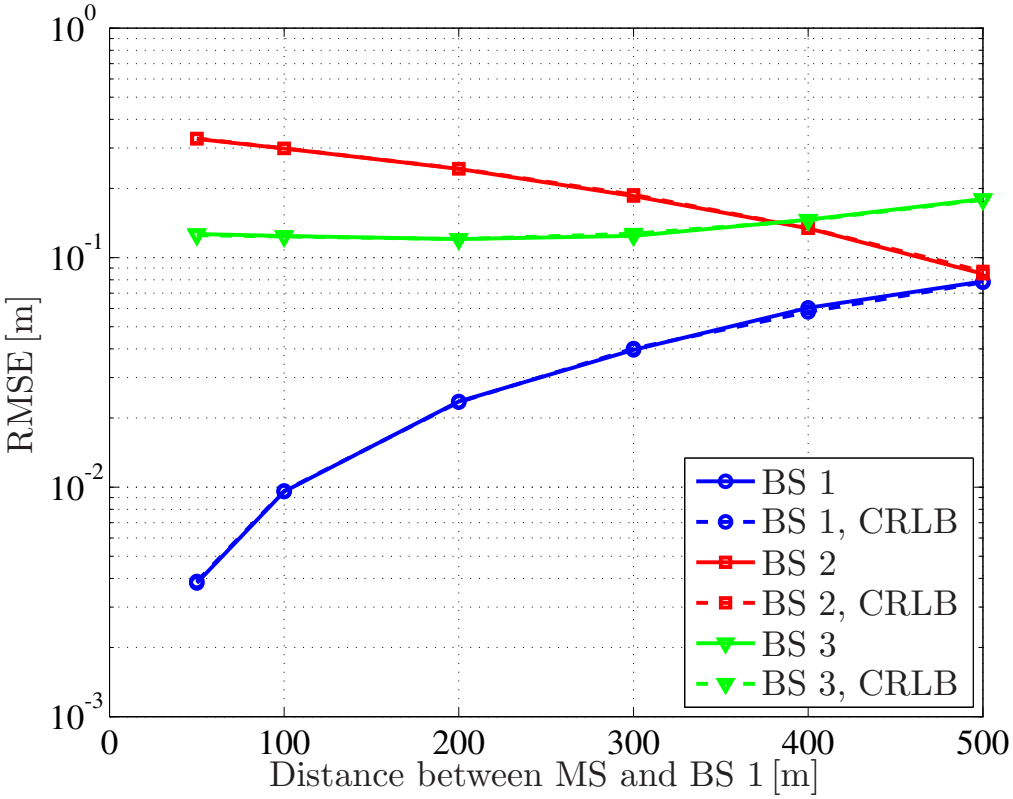


Figure 3.19: RMSE over distance for multi-link synchronization with different BSs, no interference, AWGN channel

of synchronization signals as interference. Even though the requirement for the CRLB computation on AWGN is not met here, the replacement of the SNR by the SINR gives a reliable approximation. Furthermore, we can observe that timing estimation with BS 2 and BS 3 is not possible with a reasonable accuracy. For instance, close to BS 1 solely an RMSE of several hundred meters is possible. This is in the order of the inter-BS distance, and hence, these estimates are not useful for a reliable location estimation. We further can observe that the CRLB deviates significantly from the RMSE for BS 2 and BS 3. This is due to the fact that the corresponding SINRs are smaller than around -5 dB, where according to Figure 3.11 the CRLB is no longer a good approximation for the RMSE. Only at the cell edge (distance of 400 m to 500 m) the timing estimates with three BSs give useful results for location estimation.

Figure 3.19 shows the corresponding results for an interference-free network, i.e., all BSs transmit in different frequency bands. Since the SNRs for the AWGN channel are comparatively high, the timing estimation performance is very accurate, i.e., in the order of centimeters even for the out-of-cell BSs. Furthermore, we observe that the CRLB is achieved over the complete range.

Hence, we can conclude that the interference limits the overall performance in the network. Even for an AWGN channel no accurate timing estimation with three BSs can be achieved to allow precise location estimation. In an interference-free network the timing estimates are of sufficiently high quality. A deeper analysis for the multipath LOS and NLOS channels is not presented here, as already the performance in AWGN is not acceptable for location determination purposes.

3.3 Non-pilot-aided synchronization

The pilot-aided synchronization algorithms as discussed in the previous section exploit dedicated signals for timing and CFO estimation. The required pilot symbols or there structure need to be known a priori at the MS. Another class of synchronization algorithms denoted as non-pilot-aided or blind synchronization does not require these specific signals. The respective algorithms exploit the signal itself by application of certain OFDM properties.

A practical example for non-pilot-aided synchronization can be found in [78]. It exploits the CP of the OFDM system and correlates the CP with the respective part of the symbol to derive timing information. Also the CFO can be estimated by observing the phase of the correlation function at the estimated timing sample. This is similar to the differential correlation approaches discussed in Section 3.2.2, however, here every received symbol can be used for synchronization as long as the CP and symbol lengths do not change. This approach is not well suited for location determination as a multi-link synchronization in a cellular network with frequency re-use of one is not a simple task. Since CP and symbol lengths will usually be identical for signals transmitted from different BSs, besides timing and CFO also the transmitting BS would need to be identified from the received signal. Nevertheless, due to its low computational complexity the CP based synchronization algorithm is of high interest for hardware-oriented implementations as discussed in [79] or to aid more sophisticated synchronization algorithms by exclude ambiguous estimates [80].

The approach proposed in [81] exploits the cyclostationarity of OFDM signals. It can provide timing estimates with a higher resolution than the sampling grid and is as well suited for CFO estimation. The timing estimation algorithm proposed in [82] exploits interference from the loss of orthogonality between subcarriers in case of incorrect timing information. Also these two approaches can not directly be applied for multi-link synchronization in cellular environments.

3.4 Data-aided synchronization

As we have seen in Section 3.2.4, the performance of the timing estimation algorithms is mainly limited by the reduced number of available pilot symbols (compared to the complete transmission bandwidth) and by inter-cell interference. In the following, we will outline new concepts to cope with the aforementioned problems based on data feedback. Note that the effects of multipath and NLOS propagation and respective mitigation algorithms are not directly studied in this thesis. They will further reduce the overall performance in a practical implementation. Hence, we assume a reasonably good channel estimation to compensate these effects.

3.4.1 Synchronization with estimated data

Synchronization with estimated data is a well-studied topic in the context of mobile communications algorithms. Sometimes denoted as turbo synchronization, a feedback of already estimated and decided data back to the synchronization entity allows a refined estimation of the desired parameters in an iterative way. A comprehensive overview about the general principle and various frameworks can be found in [83]. It includes a discussion of the concept proposed in [84], where the ML synchronization problem is iteratively solved by application of the expectation maximization (EM) algorithm, the approach in [85], where for the same problem the sum-product algorithm is exploited, and the gradient based method for iterative synchronization derived in [86].

In the context of OFDM based communications systems, the focus of the proposed iterative synchronization algorithms is mainly on improving the CFO estimation performance. This is rea-

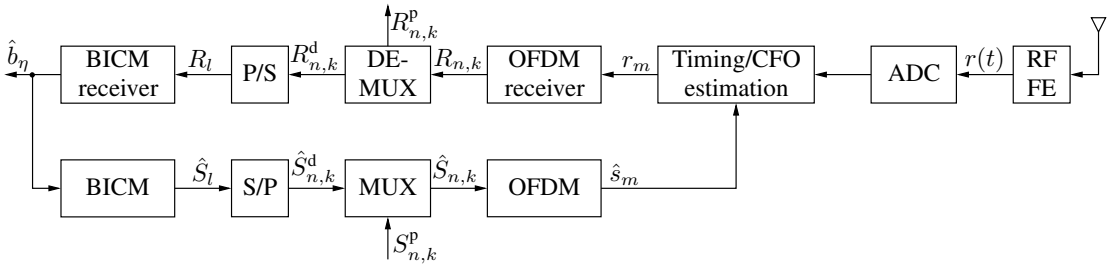


Figure 3.20: BICM-OFDM receiver for synchronization with estimated data

sonable, since the required timing estimation accuracy is comparatively small (as discussed in Section 3.2.2, the timing estimation solely has to ensure that the FFT window starts in the region of the CP without inter-symbol interference). Therefore, in a communications system the timing estimation is not that critical and the main focus is on iterative estimation of the CFO. For instance, in [87] a suitable EM based ML estimator is proposed for improving the bit error rate (BER). A similar approach is discussed in [88]. In [89], a joint data-aided estimation of CFO and sampling clock frequency offset is derived. Nevertheless, all described techniques do not consider an improved timing estimation by using estimated data.

Therefore, we propose to exploit already estimated data for an improved data-aided timing estimation as depicted in Figure 3.20 (cf. [30]). In the considered LTE system the synchronization symbols solely occupy 63 subcarriers, i.e., a bandwidth of 945 MHz, whereas the maximum number of subcarriers for the data symbols is 1201 in the 20 MHz mode. According to the CRLB investigation in Figure 3.10 large performance gains would be possible by increasing the available bandwidth with the proposed method.

To do so, the complete signal stream from the serving BS is demodulated and decoded at the MS. In a first step, the dedicated data symbols for this MS are determined as required anyhow for communications, i.e., they are OFDM demodulated and the bit stream is estimated and decided by the BICM receiver. These bits can then be used for further processing in the communications part. Here, we feed this bit stream back and reconstruct an estimate of the transmitted signal. This includes BICM resulting in the reconstructed data stream $\hat{S}^{\mathrm{d}}_{n,k}$ after S/P conversion. For the pilot part we can use the a priori known pilot symbols $S^{\mathrm{p}}_{n,k}$ in the OFDM frame. The overall reconstructed frame $\hat{S}_{n,k}$ is then OFDM modulated. Instead of the limited number of pilot signals in the original approach, now the complete estimated and reconstructed frame — or at least parts of it — is used for performing an improved timing estimation with respect to the serving BS. This can especially improve the estimates close to the serving BS, i.e., when near-optimum reconstruction is possible. At the cell edge the performance gains are limited due to an increased BER.

3.4.2 Synchronization with interference cancelation

Synchronization with estimated data as discussed before can be very useful for improving the performance of the timing estimates obtained from synchronization with the serving cell. Nevertheless, the inter-cell interference as liming factor remains (cf. Figure 3.18). Therefore, we outline principles to enhance timing estimation algorithms by including approaches that are aware of the interference.

The first idea is the complete avoidance of interference for the dedicated signals. This could be realized, e.g., by applying classical re-use partitioning schemes. This has the effect that —

at least — neighboring BSs transmit on different frequency bands resulting in an effective situation without inter-cell interference. Since LTE and other upcoming mobile radio systems target a frequency re-use of one, this is not a feasible approach as all BSs are desired to transmit in the same frequency band due to spectral efficiency demands. Nevertheless, the OFDM concept allows a smart partitioning on the subcarrier basis. This ensures that neighboring cells are coordinated in a way that inter-cell interference for the data symbols is avoided. A comprehensive overview about these principles can be found in [90, 91, 92, 93]. In a similar context the discussion in LTE on a so-called positioning reference signal could be seen [58]. This should consist of additional pilot symbols for improved location determination and further includes idle periods of neighboring BSs to reduce or even avoid the inter-cell interference on these pilot symbols. First investigations with these signals can be found in [94], however, these dedicated signals would certainly cause a signaling overhead in the OFDM frame.

The interference avoidance principles can not be used for the investigated setup as in LTE the synchronization and pilot signals are time and frequency aligned, i.e., on these signals inter-cell interference can not be avoided. The idea to reduce the interference by averaging over several frames is also not a reasonable solution as the synchronization and pilot signals usually do not change over time. Therefore, interference cancelation approaches need to be applied to reduce the interference on the desired signal. In communications systems inter-cell interference cancelation is especially applied to improve the system performance at the cell edge. At the cell edge it is difficult to estimate the data symbols from the serving BS as the interference from the other BS is comparatively high. Close to the serving BS, the received power is usually strong enough that the interference from other BS can be neglected. This is the major difference to the considered location determination application, where it is required to obtain timing information from at least three BSs. This is especially problematic close to the serving BS, where a reception of the out-of-cell BSs is very difficult as simulation results have already shown in Figure 3.19. Contrary to that, at the cell edge we have a beneficial situation for location determination, as there the received power from various BSs is similar.

Therefore, we propose to cancel the interference of the (strongest) serving BS before timing estimation with the out-of-cell BSs is performed. Figure 3.21 shows the resulting overview for data-aided synchronization with interference cancelation (cf. [31]). A similar approach was investigated for UMTS in [95]. As we can see, the complete signal stream from the serving BS is demodulated and decoded at the MS — in the same way as for synchronization with estimated data described before. Then, these estimated bits are used for obtaining the reconstructed frame $\tilde{S}_{n,k}$. This frame is OFDM modulated and pre-distorted with the estimated channel. Finally, the reconstructed time-domain samples are subtracted from the received samples at the respective time index which is determined by the timing estimate for BS 1. After canceling the interference, the timing estimates for the out-of-cell BSs are determined by pilot-aided synchronization algorithms.

In case of perfect reconstruction the complete interference from the serving BS is eliminated and only a superposition of the signals from the out-of-cell BSs is present at the receiver. However, in a real system detection and estimation errors will degrade the performance of this approach. Nevertheless, especially in situations which are critical for location determination, i.e., close to a BS, we have a favorable situation for communications. In these situations, the signal can usually be reconstructed near optimum, and hence, interference cancelation is very beneficial for the overall location determination.

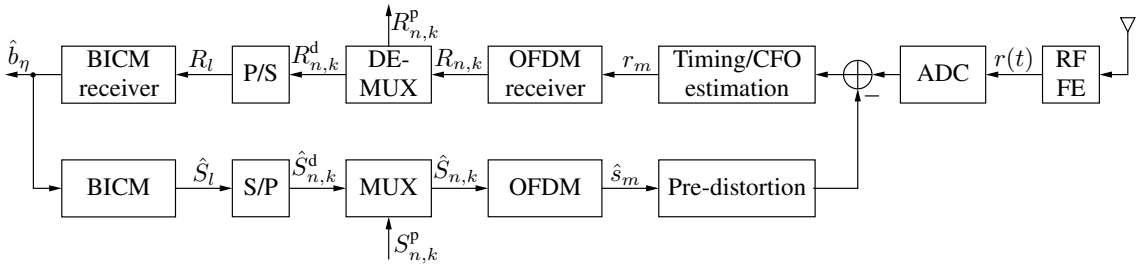


Figure 3.21: BICM-OFDM receiver for synchronization with interference cancelation

3.4.3 Iterative synchronization

So far, we just have considered a one-step iteration for the data-aided synchronization and synchronization with interference cancelation. Clearly, the highest performance gain can be obtained for considering the strongest BS which is the serving BS. In data-aided synchronization it allows a good estimation of the data symbols, and hence, a good reconstruction of the signal for an improved timing estimation. This holds in a similarly way for interference cancelation approaches, since the cancelation of this BS will be nearly ideal, especially in the most interesting case close to the serving BS. Nevertheless, after a cancelation of the serving BS this concept can be applied again for data-aided synchronization and synchronization with interference cancelation for the second-strongest BS and so on in an iterative way.

3.4.4 Simulation results

Figure 3.22 depicts the RMSE over SNR for data-aided synchronization with estimated data. For simplicity, only one active BS is considered so far. As synchronization algorithm we exploit the coherent cross-correlation approach in frequency domain according to Section 3.2.2. For pilot-aided (PA) synchronization we can achieve the CRLB for SNRs higher than $-5\,\mathrm{dB}$, however, the performance is limited by the reduced bandwidth of the synchronization signal. For data-aided (DA) synchronization with estimated data (ED), we assume that we exploit one OFDM symbol that includes 1201 subcarriers, i.e., over the maximum transmission bandwidth of LTE. The respective CRLB shows a performance gain of around $20\,\mathrm{dB}$ for the data-aided approach. We further can see that we can achieve this CRLB for SNRs higher than around $4\,\mathrm{dB}$. Note that for SNRs smaller than around $-3\,\mathrm{dB}$ the pilot-aided approach outperforms the data-aided approach since a reasonably reconstruction can not be ensured at these low SNR values, and hence, the overall performance decreases. We further observe that we loose around $10\,\mathrm{dB}$ with the data-aided synchronization with estimated data compared to the curve with perfect reconstruction. For reasonable SNR values we then can achieve RMSEs in the sub-meter region.

Figure 3.23 shows the RMSE performance in the cellular network according to Figure 3.16 under interference at the respective MS locations. Compared to the pilot-aided approach which yields RMSEs between around $2\,\mathrm{m}$ and $22\,\mathrm{m}$, the performance can be improved to RMSEs between $20\,\mathrm{cm}$ and $18\,\mathrm{m}$ for the data-aided synchronization with estimated data. The comparison with the curve for perfect reconstruction shows that the estimated reconstruction performs near optimum. This is in line with the expectations for timing estimation with the reasonably strong serving BS.

In Figure 3.24 the RMSE performance in the cellular network is evaluated for applied data-aided synchronization with interference-cancelation (IC), i.e., the interference from BS 1 is canceled before (pilot-aided) timing estimation with the out-of-cell BSs is performed. We observe a

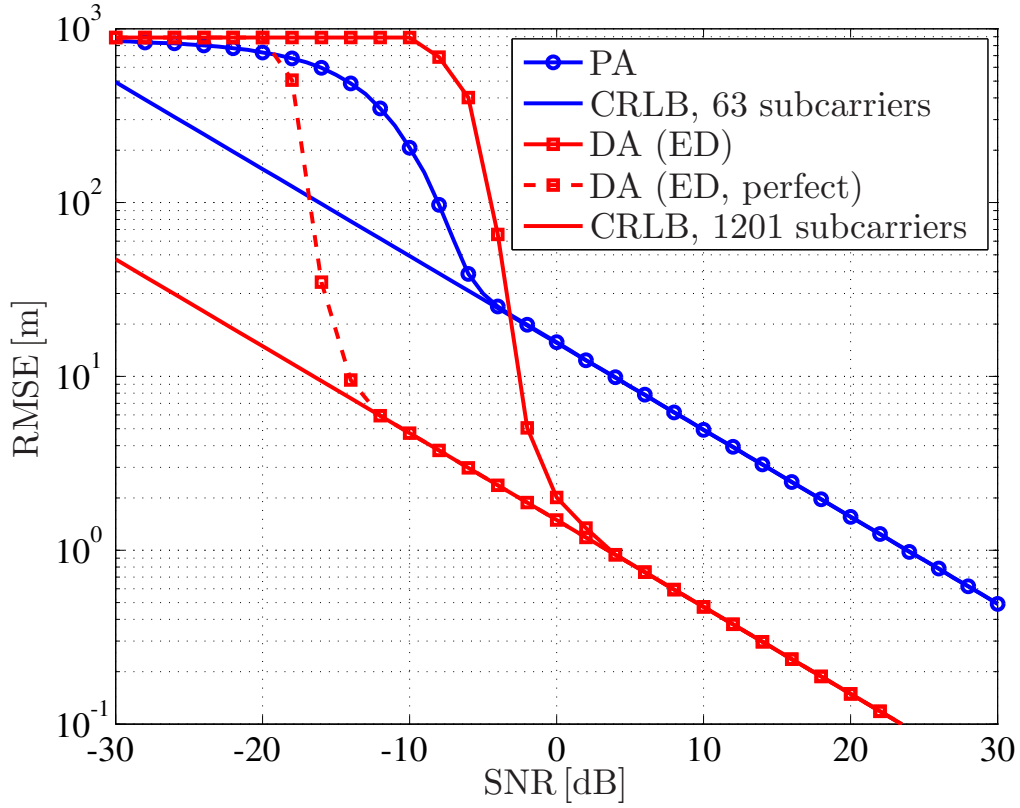


Figure 3.22: RMSE over SNR for synchronization with estimated data, AWGN channel

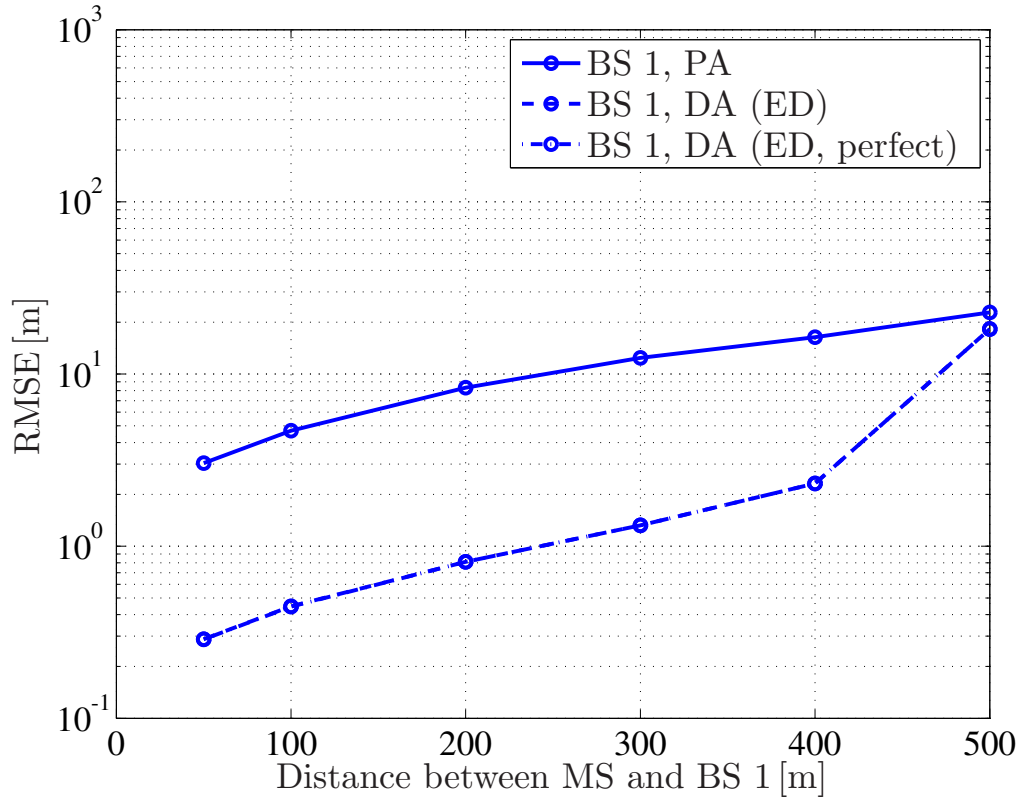


Figure 3.23: RMSE over distance for synchronization with estimated data, AWGN channel

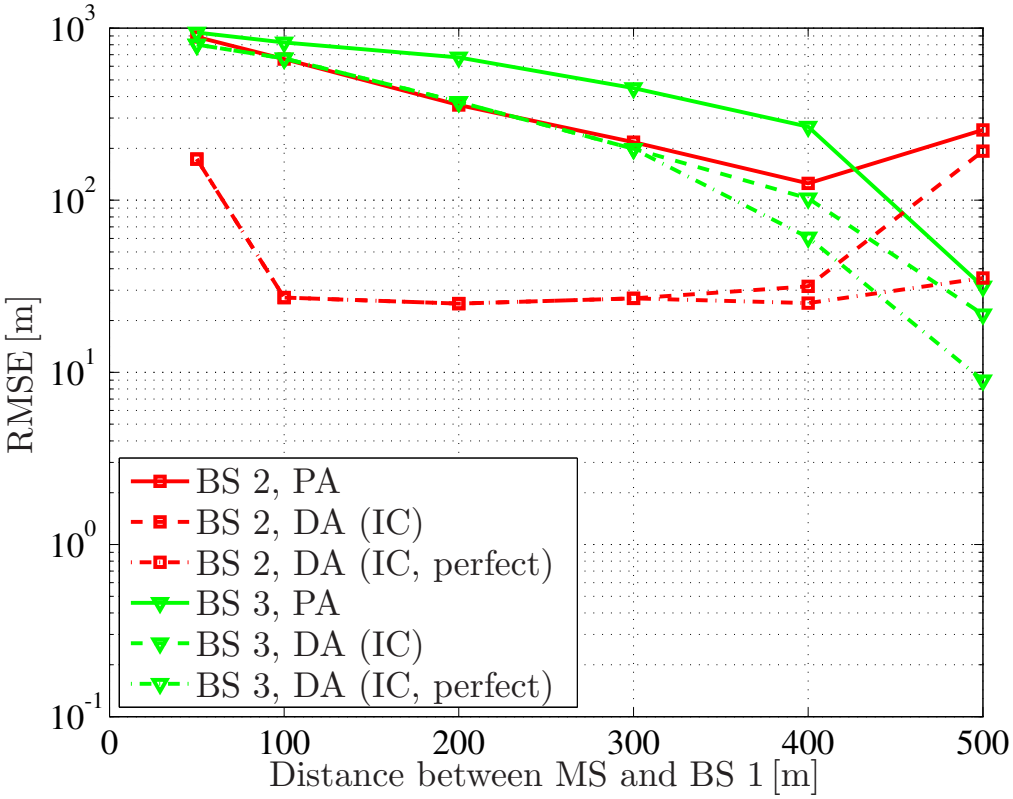


Figure 3.24: RMSE over distance for synchronization with interference cancelation, AWGN channel

remarkable performance improvement especially for BS 2 if the MS is close to BS 1. In some situations the RMSE can be reduced from several hundred meters to below 30 m. At the cell edge, the performance improvement is not that drastically, however, still beneficial compared to the standard pilot-aided approach. The performance improvement for BS 3 is not that high as BS 2 (as second-strongest BS at the respective MS locations) introduces interference which limits the timing estimation accuracy with BS 3. At the cell edge, we further observe a performance reduction compared to perfect interference cancelation. As the reconstruction is near optimum (as could be seen in Figure 3.23), reason for this deviation is the incorrect timing estimate for BS 1 and the resulting non-ideal interference cancelation.

Interesting in this context is also an assessment of the resulting SINR in the cellular network. Figure 3.25 depicts the cumulative distribution function (CDF) for the SINR

$$\mathrm{CDF}\left(x_{\mathrm{SINR}}\right) = \mathrm{P}\left(\mathrm{SINR} < x_{\mathrm{SINR}}\right), \tag{3.85}$$

i.e., the probability that the SINR is smaller than the value of the abscissae x_{SINR}, averaged over all MS locations in the network. We can see that for BS 1 in nearly all situations we obtain an SINR higher than $-5\,\mathrm{dB}$, i.e., in nearly all situations we achieve the CRLB with the timing synchronization algorithm according to Figure 3.11. For the out-of-cell BSs an SINR higher than $-5\,\mathrm{dB}$ is obtained in around 35% of the situations for BS 2 and in around 2% of the situations for BS 3. In case that BS 1 is not present in the system (perfect interference cancelation) these values can be increased to around 99% for BS 2 and 70% for BS 3.

Finally, Figure 3.26 depicts the performance for all considered pilot-aided and data-aided approaches. This includes also the iterative procedures, where synchronization with estimated data

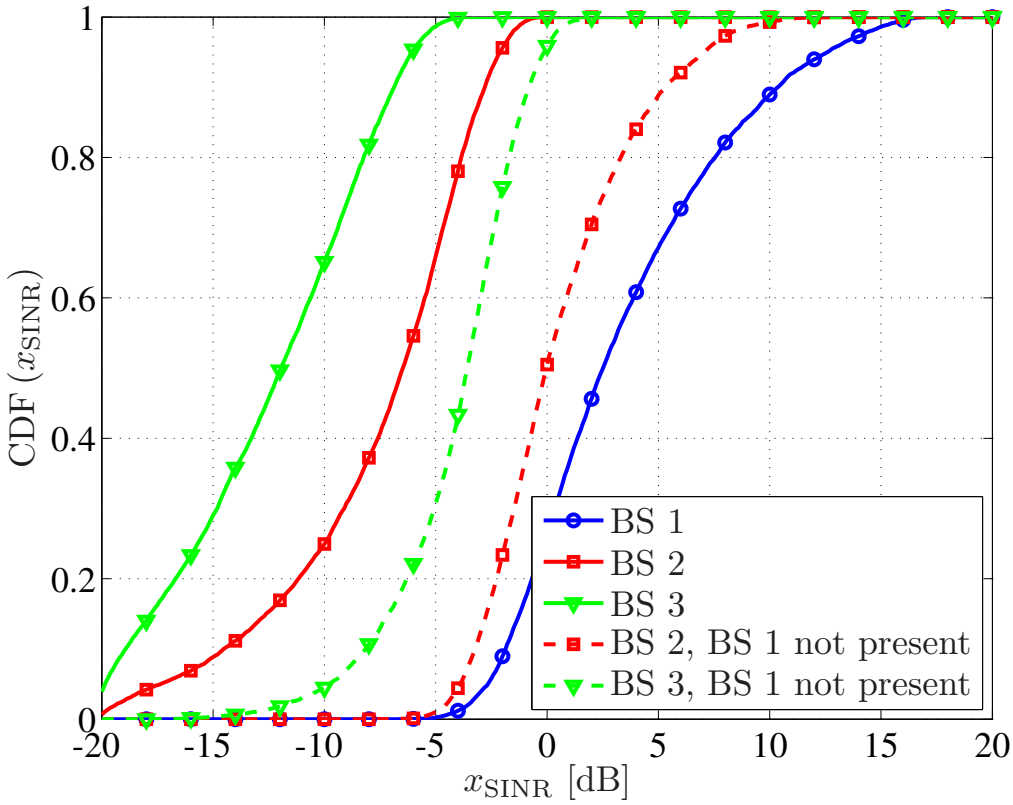


Figure 3.25: CDF of SINR in cellular network, AWGN channel

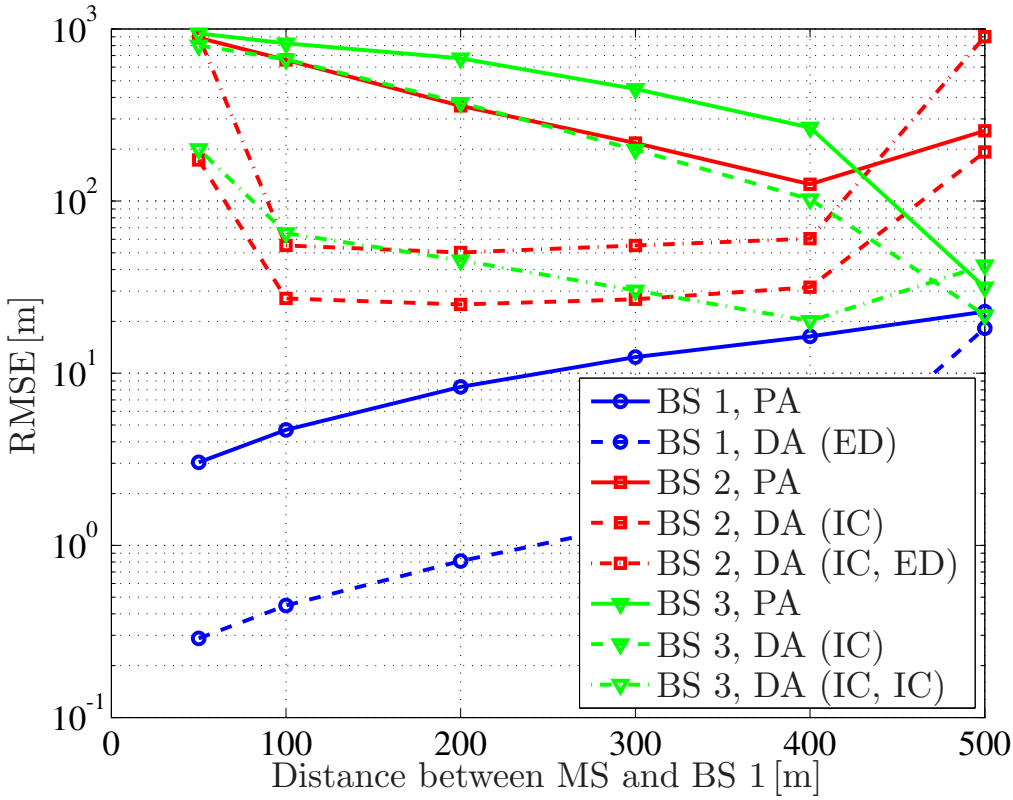


Figure 3.26: RMSE over distance for iterative synchronization, AWGN channel

or interference cancelation is applied a second time. Here, after the interference cancelation of BS 1 a synchronization with estimated data was performed for BS 2 (IC, ED). We observe that the additional synchronization reduces the overall performance due to the fact that in the considered situation the reconstruction is not optimum. Additionally, before timing estimation with BS 3 is performed, the interference from BS 1 and BS 2 is canceled (IC, IC). This results in a high performance gain compared to a single interference cancelation of only BS 1. We can achieve RMSEs between 65 m and 20 m at MS locations at a distance from BS 1 between 100 m and 400 m.

4 Static location estimation with timing information

In Chapter 3 it was discussed, how timing information of signals from various BSs can be obtained at the MS. In this chapter, we investigate how this timing information can be exploited to estimate the location of the MS. This process has to be seen in the context of the two-step location determination approach shown in Figure 4.1. Section 4.1 introduces the system model considered in this chapter. In Section 4.2, relevant estimation criteria for location estimation are described and discussed. It turns out that the underlying estimation problem results in a non-linear optimization problem. Therefore, Section 4.3 outlines suitable optimization algorithms for its solution, where it is distinguished between general numerical methods and dedicated approaches taking into account the geometric structure of the location estimation problem. The fundamental performance bound for these techniques is given by the CRLB being presented in Section 4.4. Finally, simulation results evaluate the performance of the algorithms in Section 4.5.

4.1 System model

For static location estimation it is assumed that the MS is at a constant location during the estimation process and no a priori information about this location is available. Hence, the location of the MS is a constant and deterministic parameter. The available timing information is given in terms of the timing estimates obtained from N_{BS} BSs, i.e.,

$$\hat{\boldsymbol{\tau}} = \left[\hat{\tau}_{0,1}, \hat{\tau}_{0,2}, \ldots, \hat{\tau}_{0,N_{\mathrm{BS}}}\right]^{\mathrm{T}} \in \mathbb{R}^{N_{\mathrm{BS}}}, \tag{4.1}$$

where $\hat{\tau}_{0,\mu}, \mu = 1, 2, \ldots, N_{\mathrm{BS}}$, can be provided by the respective timing estimation algorithms discussed in Chapter 3.

The processing of this information strongly depends on the structure of the network as described in Section 2.1. If BSs and MS share the same clock or do not depend on a synchronized network, i.e., if knowledge of the absolute signal transmit time is available, the timing estimates $\hat{\boldsymbol{\tau}}$ directly include absolute propagation time information between BSs and MS (cf. TOA and RT-TOA principles). Due to the underlying geometric structure we refer to this procedure as spherical

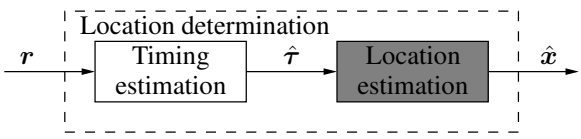


Figure 4.1: Two-step location determination process: static location estimation

approach. If there is a time offset between BSs and MS clock, the timing estimates $\hat{\boldsymbol{\tau}}$ still contain propagation time information between BSs and MS, however, biased by the constant offset between BSs and MS clocks. This clock offset can be treated as additional parameter that has to be estimated besides the MS location (cf. TOA with unknown time offset principle). We refer to this class as spherical with unknown time offset in the following. A third option is the use of difference operations between various timing estimates which eliminate the clock offset (cf. TDOA principle). This will be denoted as hyperbolic approach. This is the typical procedure to process timing estimates in a cellular network. Note that we further assume either synchronized BSs or knowledge of their asynchronism which can be taken into account when generating $\hat{\boldsymbol{\tau}}$. Therefore, we can base our analysis on three different classifications [96, 97]:

- Spherical with knowledge of the time offset between BSs and MS clocks (SPH).
- Spherical without knowledge of the time offset between BSs and MS clocks (SPHT).
- Hyperbolic (HYP).

Note that we focus on timing estimates in this thesis. Since time and range are related to each other by the speed of light, time and range information can be considered as an equivalent source of information. Hence, range information gained from RSS estimates could also be used for location estimation with the methods derived in this chapter. Furthermore, all presented approaches can be applied for both MS and network based techniques, even though the analysis will be in the context of timing estimation at the MS side as implied by Chapter 3.

We assume that the MS is located at

$$\boldsymbol{x} = [x, y, z]^{\mathrm{T}} \in \mathbb{R}^3, \tag{4.2}$$

where we presume a three-dimensional cartesian coordinate system. The BSs in the cellular network that are included in the location determination process are located at

$$\boldsymbol{x}_\mu = [x_\mu, y_\mu, z_\mu]^{\mathrm{T}} \in \mathbb{R}^3, \quad \mu = 1, 2, \ldots, N_{\mathrm{BS}}. \tag{4.3}$$

Hence, the distances between the BSs and the MS are given as

$$\boldsymbol{d} = [d_1, d_2, \ldots, d_{N_{\mathrm{BS}}}]^{\mathrm{T}} \in \mathbb{R}^{N_{\mathrm{BS}}}, \tag{4.4}$$

with

$$d_\mu = \|\boldsymbol{x}_\mu - \boldsymbol{x}\|_2 = \sqrt{(x_\mu - x)^2 + (y_\mu - y)^2 + (z_\mu - z)^2}, \quad \mu = 1, 2, \ldots, N_{\mathrm{BS}}. \tag{4.5}$$

We assume that the propagation speed is constant, and hence, the distance or range is proportional to the propagation time with a constant propagation factor. For the considered scenarios this is the speed of light c_0. Thus, the corresponding propagation times between the BSs and the MS are given as

$$\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_{N_{\mathrm{BS}}}]^{\mathrm{T}} \in \mathbb{R}^{N_{\mathrm{BS}}}, \tag{4.6}$$

with

$$\tau_\mu = \frac{d_\mu}{c_0}, \quad \mu = 1, 2, \ldots, N_{\mathrm{BS}}. \tag{4.7}$$

	Classification		
Parameter symbol	SPH	SPHT	HYP
N_{obs}	N_{BS}	N_{BS}	$N_{\text{BS}}-1$
$\hat{\boldsymbol{f}}$	$\hat{\boldsymbol{\tau}}$	$\hat{\boldsymbol{\tau}}$	$\boldsymbol{D}\hat{\boldsymbol{\tau}}$
N_{par}	3	4	3
$\boldsymbol{p}$	$\boldsymbol{x}$	$\left[\boldsymbol{x}^{\text{T}}, b_{\text{clock}}\right]^{\text{T}}$	$\boldsymbol{x}$
$\boldsymbol{f}(\boldsymbol{p})$	$\boldsymbol{\tau}(\boldsymbol{x})$	$\boldsymbol{\tau}(\boldsymbol{x}) + b_{\text{clock}}\mathbf{1}_{N_{\text{BS}}\times 1}$	$\boldsymbol{D}\boldsymbol{\tau}(\boldsymbol{x})$
$\boldsymbol{n}_f$	$\boldsymbol{n}_\tau$	$\boldsymbol{n}_\tau$	$\boldsymbol{D}\boldsymbol{n}_\tau$
$\boldsymbol{\Sigma}_{\boldsymbol{n}_f}$	$\boldsymbol{\Sigma}_{\boldsymbol{n}_\tau}$	$\boldsymbol{\Sigma}_{\boldsymbol{n}_\tau}$	$\boldsymbol{D}\boldsymbol{\Sigma}_{\boldsymbol{n}_\tau}\boldsymbol{D}^{\text{T}}$

Table 4.1: Mappings to generalized parameters for location estimation with timing information

In case of spherical processing, the parameter vector that has to be estimated is $\boldsymbol{x}$, i.e., the three-dimensional location of the MS. The timing estimates are modeled according to

$$\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}(\boldsymbol{x}) + \boldsymbol{n}_\tau, \tag{4.8}$$

where

$$\boldsymbol{n}_\tau = \left[n_{\tau,1}, n_{\tau,2}, \ldots, n_{\tau,N_{\text{BS}}}\right]^{\text{T}} \in \mathbb{R}^{N_{\text{BS}}} \tag{4.9}$$

denotes noise with the PDF $\text{p}(\boldsymbol{n}_\tau)$. Note that the dependency of the propagation time $\boldsymbol{\tau}$ on the deterministic location $\boldsymbol{x}$ is explicitly denoted in (4.8).

In case of spherical processing with unknown time offset, the parameter vector has to be extended by the time offset b_{clock} between the clocks of BSs and MS, i.e., to $\left[\boldsymbol{x}^{\text{T}}, b_{\text{clock}}\right]^{\text{T}}$. The timing estimates are then modeled according to

$$\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}(\boldsymbol{x}) + b_{\text{clock}}\mathbf{1}_{N_{\text{BS}}\times 1} + \boldsymbol{n}_\tau, \tag{4.10}$$

with the same noise properties of $\boldsymbol{n}_\tau$ as for spherical processing.

In case of hyperbolic processing, the time offset b_{clock} is removed by processing suitable time differences. Hence, the timing estimates are modeled according to

$$\boldsymbol{D}\hat{\boldsymbol{\tau}} = \boldsymbol{D}\left(\boldsymbol{\tau}(\boldsymbol{x}) + b_{\text{clock}}\mathbf{1}_{N_{\text{BS}}\times 1}\right) + \boldsymbol{D}\boldsymbol{n}_\tau, \tag{4.11}$$

where the matrix $\boldsymbol{D}$ is a full rank matrix ensuring that the dependency on b_{clock} is eliminated. A common definition, e.g., for TDOA positioning is

$$\boldsymbol{D} = \begin{bmatrix} -\mathbf{1}_{(N_{\text{BS}}-1)\times 1} & \mathbf{I}_{N_{\text{BS}}-1} \end{bmatrix} \in \mathbb{R}^{(N_{\text{BS}}-1)\times N_{\text{BS}}}, \tag{4.12}$$

where the reference BS for the time differences is chosen as BS 1.

To allow a generalized derivation of the optimization algorithms, we choose a system model according to

$$\hat{\boldsymbol{f}} = \boldsymbol{f}(\boldsymbol{p}) + \boldsymbol{n}_f \in \mathbb{R}^{N_{\text{obs}}}, \tag{4.13}$$

where all corresponding mappings for SPH, SPHT, and HYP, to the generalized parameters are concluded in Table 4.1. Hence, the estimation problem is to determine the N_{par} parameters of the parameter vector $\boldsymbol{p}$ by using an observation vector $\hat{\boldsymbol{f}}$ of dimension N_{obs} under the system model given in (4.13).

4.2 Estimation criteria

Obviously, the system model generalized in (4.13) is non-linear with respect to the parameter vector $\boldsymbol{p}$ due to the non-linear relation between the location and the timing or range information. If no a priori information about the noise is available, a straightforward criterion for the solution of the resulting estimation problem is based on non-linear least squares optimization (e.g., [12]), where the squared differences between the observations and all possible parameter sets for the given system model are minimized, i.e.,

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}}{\operatorname{argmin}} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right)^{\mathrm{T}} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right). \tag{4.14}$$

It is the optimum approach if the noise is Gaussian distributed according to $\boldsymbol{n}_f \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{n}_f}, \boldsymbol{\Sigma}_{\boldsymbol{n}_f}\right)$ with mean $\boldsymbol{\mu}_{\boldsymbol{n}_f} = \mathbf{0}_{N_{\text{obs}}\times 1}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_f} = \sigma^2_{\boldsymbol{n}_f}\mathbf{I}_{N_{\text{obs}}}$, i.e., zero-mean with a scaled identity matrix as covariance matrix. In this case, the non-linear least squares estimation criterion is identical to ML estimation [12]. The ML estimator is defined as the value that maximizes the likelihood function $\mathrm{p}\left(\hat{\boldsymbol{f}}; \boldsymbol{p}\right)$ over the allowed domain of $\boldsymbol{p}$.

If a priori information about the noise properties is available, the weighted non-linear least squares criterion yields a more appropriate approach. Assuming zero-mean Gaussian distributed noise with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_f}$, the optimum estimator results in

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}}{\operatorname{argmin}} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right). \tag{4.15}$$

In this approach, the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_f}$ includes reliability and correlation information of the estimates. For the location estimation problem with timing information, this matrix is identical to $\boldsymbol{\Sigma}_{\boldsymbol{n}_t}$ for both spherical approaches. For hyperbolic processing it results in $\boldsymbol{\Sigma}_{\boldsymbol{n}_f} = \boldsymbol{D}\boldsymbol{\Sigma}_{\boldsymbol{n}_t}\boldsymbol{D}^{\mathrm{T}}$, hence, the estimates used for location determination are correlated (cf. Table 4.1). Under the given constraints, the weighted non-linear least squares estimation criterion is identical to ML estimation.

4.3 Optimization algorithms

The weighted non-linear least squares approach is commonly used in the context of location estimation applications (e.g., [18, 21]), and hence, it will be the baseline during the following derivation of the optimization algorithms. The underlying estimation problem given in (4.15) usually does not have any closed-form solution. Therefore, we consider two different approaches for the solution of such kind of estimation problem. In Section 4.3.1, classical numerical algorithms for non-linear optimization problems are described. They certainly can solve the underlying location estimation problem by generally minimizing a given cost function. Nevertheless, they usually do not take into account the respective structure of the problem explicitly. Contrary to that, in Section 4.3.2 dedicated location estimation algorithms are introduced. They exploit, e.g., the geometric structure of the problem or specific location estimation constraints.

4.3.1 Numerical methods

With the system model introduced in Section 4.1 and following the weighted non-linear least squares approach, the target is to minimize the resulting cost function

$$\varepsilon\left(\boldsymbol{p}\right) = \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right) \tag{4.16}$$

with respect to the unknown parameter vector $\boldsymbol{p}$, yielding

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}}{\operatorname{argmin}}\, \varepsilon\left(\boldsymbol{p}\right). \tag{4.17}$$

Note that contrary to linear least squares, the considered non-linear least squares cost function $\varepsilon\left(\boldsymbol{p}\right)$ is in general non-convex.

In the following, we outline iterative optimization techniques to find a suitable estimate (e.g., [98]). Iterative in this sense means that from a starting or initial value $\boldsymbol{p}_0$ the methods produce a series of estimates $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots$, that should converge to a local minimizer

$$\hat{\boldsymbol{p}} = \lim_{k\to\infty} \boldsymbol{p}_k \tag{4.18}$$

of the cost function $\varepsilon\left(\boldsymbol{p}\right)$. Hence, the iterations can be written as

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \Delta\boldsymbol{p}_k, \tag{4.19}$$

where $\Delta\boldsymbol{p}_k$ is the correction term at iteration step k.

A necessary condition for a value $\hat{\boldsymbol{p}}$ to be a local minimum, maximum, or inflection point, for an — for the considered case differentiable — function $\varepsilon\left(\boldsymbol{p}\right)$, is that the gradient of this function evaluated at $\hat{\boldsymbol{p}}$ vanishes, i.e.,

$$\nabla_{\boldsymbol{p}}\varepsilon\left(\hat{\boldsymbol{p}}\right) = \mathbf{0}_{N_{\text{par}}\times 1}, \tag{4.20}$$

where the gradient of a scalar-valued function $\varepsilon\left(\boldsymbol{p}\right)$ provides the local direction of steepest ascent and

$$\nabla_{\boldsymbol{p}} = \left[\frac{\partial}{\partial p_1}, \frac{\partial}{\partial p_2}, \ldots, \frac{\partial}{\partial p_{N_{\text{par}}}}\right]^{\mathrm{T}} \in \mathbb{R}^{N_{\text{par}}} \tag{4.21}$$

includes the partial derivatives with respect to the elements of the parameter vector $\boldsymbol{p}$. For the particular weighted non-linear least squares cost function in (4.16), the gradient can be calculated as

$$\nabla_{\boldsymbol{p}}\varepsilon\left(\boldsymbol{p}\right) = -\boldsymbol{\Phi}\left(\boldsymbol{p}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}\right)\right), \tag{4.22}$$

using the Jacobian matrix

$$\boldsymbol{\Phi}\left(\boldsymbol{p}\right) = \nabla_{\boldsymbol{p}}^{\mathrm{T}} \otimes \boldsymbol{f}\left(\boldsymbol{p}\right) \in \mathbb{R}^{N_{\text{obs}}\times N_{\text{par}}}. \tag{4.23}$$

The respective Jacobian matrices for SPH, SPHT, and HYP are shown in Appendix A.2. Hence, the necessary condition for a local minimum, maximum, or inflection point can be written as

$$-\boldsymbol{\Phi}\left(\hat{\boldsymbol{p}}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\hat{\boldsymbol{p}}\right)\right) = \mathbf{0}_{N_{\text{par}}\times 1}. \tag{4.24}$$

The sufficient condition for $\hat{\boldsymbol{p}}$ to be a local minimum for the function $\varepsilon\left(\boldsymbol{p}\right)$ additionally requires that the Hessian matrix of $\varepsilon\left(\boldsymbol{p}\right)$ is positive definite at $\hat{\boldsymbol{p}}$, i.e.,

$$\boldsymbol{H}\left(\hat{\boldsymbol{p}}\right) > \mathbf{0}_{N_{\text{par}}\times N_{\text{par}}}, \tag{4.25}$$

where the Hessian matrix is defined as

$$\boldsymbol{H}\left(\boldsymbol{p}\right) = \nabla_{\boldsymbol{p}}\nabla_{\boldsymbol{p}}^{\mathrm{T}}\varepsilon\left(\boldsymbol{p}\right) \in \mathbb{R}^{N_{\text{par}}\times N_{\text{par}}}. \tag{4.26}$$

Both necessary and sufficient conditions indicate that a local optimum is found.

Since there are available numerous references for iterative solutions for non-linear least squares problems (e.g., [98]), we focus here on outlining a set of suitable and commonly used algorithms in the context of location estimation (cf. [18, 11]).

Steepest descent algorithm

An intuitive choice for a correction term $\Delta\boldsymbol{p}_k$ in (4.19) is based on the gradient at the estimate of the current iteration step (e.g., [98, 99, 11]). Often, the resulting search direction is weighted by a scalar parameter $\alpha_k > 0$ in order to guarantee stability and convergence of this method. The resulting algorithm can then be written as

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \alpha_k \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_k)\, \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k)\right). \tag{4.27}$$

The parameter α_k can either be determined by line-search methods [98] in each iteration step or set to a constant. This approach is denoted as gradient or steepest descent method. The choice of the descent direction is locally optimum, however, the final convergence is linear and can be very slow.

Newton algorithm

The Newton algorithm (e.g., [98]) is based on minimizing the quadratic approximation of the cost function around the current estimate $\boldsymbol{p}_k$ which is given as

$$\varepsilon(\boldsymbol{p}) \approx \varepsilon(\boldsymbol{p}_k) + \nabla_{\boldsymbol{p}}\varepsilon(\boldsymbol{p}_k)(\boldsymbol{p} - \boldsymbol{p}_k) + \frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}_k)^{\mathrm{T}} \boldsymbol{H}(\boldsymbol{p}_k)(\boldsymbol{p} - \boldsymbol{p}_k). \tag{4.28}$$

The solution of the necessary condition in (4.20) with this approximation yields

$$\Delta\boldsymbol{p}_k = \boldsymbol{H}^{-1}(\boldsymbol{p}_k)\, \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_k)\, \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k)\right) \tag{4.29}$$

as correction term of iteration step k. The iterated solution can then be written as

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \alpha_k \boldsymbol{H}^{-1}(\boldsymbol{p}_k)\, \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_k)\, \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k)\right). \tag{4.30}$$

Similar as for the steepest descent algorithm, the parameter α_k can be used to control the convergence of the algorithm. However, in the classical Newton algorithm this parameter is set to $\alpha_k = \alpha = 1$. Even though this technique can be applied to general cost functions, the method is usually difficult to implement since the construction of the Hessian matrix and its inversion in each iteration step is computational complex. Nevertheless, the Newton algorithm has robust and fast convergence properties.

Gauss-Newton algorithm

The Gauss-Newton algorithm [12, 100, 11] linearizes the non-linear term $\boldsymbol{f}(p)$ in the system model (4.13) around the current estimate $\boldsymbol{p}_k$ yielding

$$\boldsymbol{f}(\boldsymbol{p}) \approx \boldsymbol{f}(\boldsymbol{p}_k) + \boldsymbol{\Phi}(\boldsymbol{p}_k)(\boldsymbol{p} - \boldsymbol{p}_k). \tag{4.31}$$

Using this linearization, the cost function in (4.16) can be rewritten as

$$\varepsilon(\boldsymbol{p}) \approx \left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k) - \boldsymbol{\Phi}(\boldsymbol{p}_k)(\boldsymbol{p} - \boldsymbol{p}_k)\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1} \left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k) - \boldsymbol{\Phi}(\boldsymbol{p}_k)(\boldsymbol{p} - \boldsymbol{p}_k)\right), \tag{4.32}$$

i.e., it is only linear with respect to $\boldsymbol{p}$. Hence, the respective optimization problem can be solved by classical linear least squares techniques [12]. The resulting correction term at iteration step k is

$$\Delta\boldsymbol{p}_k = \left(\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{p}_k\right)\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}_k\right)\right). \tag{4.33}$$

It can be used for the iterated solution

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \alpha_k\left(\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{p}_k\right)\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}_k\right)\right). \tag{4.34}$$

Again, a parameter α_k can be used to control the convergence of the algorithm, where in the classical Gauss-Newton algorithm this parameter is set to $\alpha_k = \alpha = 1$. Compared to the Newton algorithm, where the derivative of the cost function is linearized about the current estimate, in the Gauss-Newton algorithm the system model is linearized directly with a following application of linear least squares. Another interpretation when comparing the results in (4.30) and (4.34) is that the Gauss-Newton algorithm can be seen as an approximation of the Newton algorithm without using the second-order terms. From a computational complexity point of view the Gauss-Newton algorithm avoids the calculation of the Hessian matrix compared to the Newton algorithm, nevertheless, a matrix inversion is required in each iteration step. The convergence is similar to the Newton algorithm, however, much better compared to the — computational less complex — steepest descent algorithm.

Levenberg-Marquardt algorithm

Levenberg [101] and Marquardt [102] proposed to use a so-called damped Gauss-Newton method for the solution of non-linear least squares problems. Basic idea is to modify the correction term of the Gauss-Newton algorithm to

$$\Delta\boldsymbol{p}_k = \left(\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{p}_k\right) + \beta_k\mathbf{I}_{N_{\mathrm{par}}}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}_k\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}\left(\boldsymbol{p}_k\right)\right), \tag{4.35}$$

where β_k is denoted as damping parameter. This damping parameter has several effects on the algorithm: first of all, with $\beta_k > 0$ the coefficient matrix can be kept positive definite which ensures that $\Delta\boldsymbol{p}_k$ is a descent direction. Furthermore, for large values of β_k the correction step results in the correction step of the steepest descent algorithm which is favorable if the current estimate is far from the optimum. If β_k is very small, the correction step results in the correction step of the Gauss-Newton algorithm which is favorable in the final stages of iteration, when the current estimate is close to the optimum. The Levenberg-Marquardt algorithm can be extended by implicit line-search procedures (e.g., [103]). The resulting algorithm is concluded in Algorithm 4.1.

4.3.2 Geometric approaches

The previously presented numerical methods provide a solution for (4.15) without consideration of the underlying location estimation problem. This is done in an iterative way by starting from an initial value. As simulation results will show in Section 4.5, they can provide reliable and accurate estimates. Nevertheless, there were developed numerous closed-form solutions for this problem that exploit the geometric structure of the problem by taking into account, e.g., spherical or hyperbolical properties or constraints derived from realistic environments. The performance of these closed-form algorithms was extensively investigated in [104] and compared with iterative

Algorithm 4.1 Levenberg-Marquardt algorithm

$\nu \leftarrow 2$
$\boldsymbol{A}_0 \leftarrow \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_0)\,\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}(\boldsymbol{p}_0)$
$\boldsymbol{g}_0 \leftarrow \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_0)\,\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_0)\right)$
$\beta_0 \leftarrow \max\left\{[\boldsymbol{A}_0]_{i,i}\right\}$
$k \leftarrow 0$
repeat
  $k \leftarrow k + 1$
  $\boldsymbol{h}_{k-1} \leftarrow -\left(\boldsymbol{A}_{k-1} + \beta_{k-1}\mathbf{I}_{N_{\text{par}}}\right)^{-1}\boldsymbol{g}_{k-1}$
  $\boldsymbol{p}_k \leftarrow \boldsymbol{p}_{k-1} + \boldsymbol{h}_{k-1}$
  $\rho \leftarrow \dfrac{\varepsilon(\boldsymbol{p}_{k-1}) - \varepsilon(\boldsymbol{p}_k)}{\boldsymbol{h}_{k-1}^{\mathrm{T}}\left(\beta_{k-1}\boldsymbol{h}_{k-1} + \boldsymbol{g}_{k-1}\right)}$
  if $\rho > 0$ **then**
    $\boldsymbol{A}_k \leftarrow \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_k)\,\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}(\boldsymbol{p}_k)$
    $\boldsymbol{g}_k \leftarrow \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{p}_k)\,\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\left(\hat{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{p}_k)\right)$
    $\beta_k \leftarrow \beta_{k-1}\max\left\{\frac{1}{3}, 1 - (2\rho - 1)^3\right\}$
    $\nu \leftarrow 2$
  else
    $\beta_k \leftarrow \beta_{k-1}\nu$
    $\nu \leftarrow 2\nu$
  end if
until convergence

approaches for the spherical class. The outcome of this analysis was, that properly designed iterative approaches are more robust and accurate than respective closed-form solutions at a similar computational complexity. The investigation in [39] came to a similar conclusion for hyperbolic approaches.

Nevertheless, to have a benchmark for closed-form estimators, in the following we present a competing technique to derive location information from timing estimates in a non-iterative way. We limit a detailed presentation to the algorithm developed by Chan and Ho [105, 106] providing a closed-form solution for hyperbolic location estimation. Note that at this stage we focus on a hyperbolic estimator, as the structure of the obtained timing estimates (cf. Chapter 3) in the cellular network implies hyperbolic processing.

Chan-Ho algorithm

In the Chan-Ho algorithm [105, 106] it was shown that the system model for hyperbolic location estimation (4.11) — assuming the difference operations defined in (4.12) — can be rewritten as linear equation system

$$\boldsymbol{G}_{\mathrm{a}}\boldsymbol{x}_{\mathrm{a}} + \boldsymbol{h}_{\mathrm{a}} = \boldsymbol{n}_{\mathrm{a}} \tag{4.36}$$

for the unknown vector

$$\boldsymbol{x}_{\mathrm{a}} = \begin{bmatrix} \boldsymbol{x} \\ d_1 \end{bmatrix} \in \mathbb{R}^4, \tag{4.37}$$

using the matrix

$$\boldsymbol{G}_{\mathrm{a}} = -\begin{bmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 & c_0\left(\hat{\tau}_2 - \hat{\tau}_1\right) \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 & c_0\left(\hat{\tau}_3 - \hat{\tau}_1\right) \\ \vdots & \vdots & \vdots & \vdots \\ x_{N_{\mathrm{BS}}} - x_1 & y_{N_{\mathrm{BS}}} - y_1 & z_{N_{\mathrm{BS}}} - z_1 & c_0\left(\hat{\tau}_{N_{\mathrm{BS}}} - \hat{\tau}_1\right) \end{bmatrix} \in \mathbb{R}^{(N_{\mathrm{BS}}-1)\times 4} \tag{4.38}$$

and the vector

$$\boldsymbol{h}_{\mathrm{a}} = \frac{1}{2}\begin{bmatrix} c_0^2\left(\hat{\tau}_2 - \hat{\tau}_1\right)^2 - x_2^2 - y_2^2 - z_2^2 + x_1^2 + y_1^2 + z_1^2 \\ c_0^2\left(\hat{\tau}_3 - \hat{\tau}_1\right)^2 - x_3^2 - y_3^2 - z_3^2 + x_1^2 + y_1^2 + z_1^2 \\ \vdots \\ c_0^2\left(\hat{\tau}_{N_{\mathrm{BS}}} - \hat{\tau}_1\right)^2 - x_{N_{\mathrm{BS}}}^2 - y_{N_{\mathrm{BS}}}^2 - z_{N_{\mathrm{BS}}}^2 + x_1^2 + y_1^2 + z_1^2 \end{bmatrix} \in \mathbb{R}^{N_{\mathrm{BS}}-1}. \tag{4.39}$$

The noise term in (4.36) can be computed as

$$\boldsymbol{n}_{\mathrm{a}} = \begin{bmatrix} c_0 d_2\left(n_{\tau,2} - n_{\tau,1}\right) + \frac{1}{2}c_0^2\left(n_{\tau,2} - n_{\tau,1}\right)^2 \\ c_0 d_3\left(n_{\tau,3} - n_{\tau,1}\right) + \frac{1}{2}c_0^2\left(n_{\tau,3} - n_{\tau,1}\right)^2 \\ \vdots \\ c_0 d_{N_{\mathrm{BS}}}\left(n_{\tau,N_{\mathrm{BS}}} - n_{\tau,1}\right) + \frac{1}{2}c_0^2\left(n_{\tau,N_{\mathrm{BS}}} - n_{\tau,1}\right)^2 \end{bmatrix} \in \mathbb{R}^{N_{\mathrm{BS}}-1}, \tag{4.40}$$

with covariance matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{a}}} = c_0^2 \boldsymbol{B}_{\mathrm{a}} \boldsymbol{\Sigma}_{\boldsymbol{n}_\tau} \boldsymbol{B}_{\mathrm{a}} \in \mathbb{R}^{(N_{\mathrm{BS}}-1)\times(N_{\mathrm{BS}}-1)}, \tag{4.41}$$

using

$$\boldsymbol{B}_{\mathrm{a}} = \begin{bmatrix} d_2 & & & \\ & d_3 & & \\ & & \ddots & \\ & & & d_{N_{\mathrm{BS}}-1} \end{bmatrix} \in \mathbb{R}^{(N_{\mathrm{BS}}-1)\times(N_{\mathrm{BS}}-1)}. \tag{4.42}$$

It is obvious that the equation system (4.36) is still non-linear in x, y, and z, since the components of the unknown vector (4.37) are non-linearly constraint according to (4.5). Nevertheless, for the ongoing derivation independency is assumed. Therefore, the system of equations (4.36) can simply be solved by applying standard least squares techniques resulting in

$$\hat{\boldsymbol{x}}_{\mathrm{a}} = \begin{bmatrix} \hat{x}_{\mathrm{a},1} \\ \hat{x}_{\mathrm{a},2} \\ \hat{x}_{\mathrm{a},3} \\ \hat{x}_{\mathrm{a},4} \end{bmatrix} = \left(\boldsymbol{G}_{\mathrm{a}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{a}}}^{-1} \boldsymbol{G}_{\mathrm{a}} \right)^{-1} \boldsymbol{G}_{\mathrm{a}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{a}}}^{-1} \boldsymbol{h}_{\mathrm{a}}. \tag{4.43}$$

Note that $\boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{a}}}$ is not exactly known since (4.42) depends on the true distances between MS and BSs. Therefore, an appropriate approximation based on intermediate estimates of $\hat{\boldsymbol{x}}_{\mathrm{a}}$ has to be applied [105, 106].

In a second step of the Chan-Ho algorithm the dependency between the variables in $\boldsymbol{x}_{\mathrm{a}}$ is taken into account by solving

$$\boldsymbol{G}_{\mathrm{b}} \boldsymbol{x}_{\mathrm{b}} + \boldsymbol{h}_{\mathrm{b}} = \boldsymbol{n}_{\mathrm{b}} \tag{4.44}$$

for the unknown vector

$$\boldsymbol{x}_{\mathrm{b}} = \begin{bmatrix} (x - x_1)^2 \\ (y - y_1)^2 \\ (z - z_1)^2 \end{bmatrix} \in \mathbb{R}^3, \tag{4.45}$$

using the matrix

$$\boldsymbol{G}_{\mathrm{b}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 3} \tag{4.46}$$

and the vector

$$\boldsymbol{h}_{\mathrm{b}} = \begin{bmatrix} (\hat{x}_{\mathrm{a},1} - x_1)^2 \\ (\hat{x}_{\mathrm{a},2} - y_1)^2 \\ (\hat{x}_{\mathrm{a},3} - z_1)^2 \\ \hat{x}_{\mathrm{a},4}^2 \end{bmatrix} \in \mathbb{R}^4. \tag{4.47}$$

It results in the linear least squares solution

$$\hat{\boldsymbol{x}}_{\mathrm{b}} = \begin{bmatrix} \hat{x}_{\mathrm{b},1} \\ \hat{x}_{\mathrm{b},2} \\ \hat{x}_{\mathrm{b},3} \end{bmatrix} = \left(\boldsymbol{G}_{\mathrm{b}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{b}}}^{-1} \boldsymbol{G}_{\mathrm{b}} \right)^{-1} \boldsymbol{G}_{\mathrm{b}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{b}}}^{-1} \boldsymbol{h}_{\mathrm{b}}, \tag{4.48}$$

however, also here appropriate approximations to obtain the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_{\mathrm{b}}}$ of the noise vector $\boldsymbol{n}_{\mathrm{b}}$ are required [105, 106].

In a third step, the final location estimate can be computed from the intermediate estimate $\hat{\boldsymbol{x}}_{\mathrm{b}}$ according to

$$\hat{\boldsymbol{x}} = \pm \begin{bmatrix} \sqrt{\hat{x}_{\mathrm{b},1}} \\ \sqrt{\hat{x}_{\mathrm{b},2}} \\ \sqrt{\hat{x}_{\mathrm{b},3}} \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{4.49}$$

The appropriate solution can be found by choosing the location estimate in the region of interest [105, 106]. In case of negative arguments of the square roots, the resulting imaginary parts should be set to zero. It was further shown in [105, 106] that the this method can achieve the CRLB for reasonably low noise values.

Fallback approach

A fallback approach can use — as very simple location estimate — the geometric mean value of the locations of all involved BSs, i.e.

$$\hat{\boldsymbol{x}} = \frac{1}{N_{\mathrm{BS}}} \sum_{\mu=1}^{N_{\mathrm{BS}}} \boldsymbol{x}_{\mu}. \tag{4.50}$$

This estimate will also be used as initial value for the numerical methods presented in Section 4.3.1 and might be used as fallback solution if no reliable timing information is available.

4.3.3 Further approaches

In the following, we briefly outline other concepts of location estimation with a focus on hyperbolic processing of timing estimates as the most relevant case in this thesis.

Spherical intersection based approaches

The basic concept of the Chan-Ho algorithm, i.e., the reorganization of the non-linear optimization problem into a simpler linear optimization problem by introducing a dependent intermediate variable according to (4.37), was already assessed in [107]. This method — later denoted as spherical intersection algorithm — can simply be derived from the more sophisticated Chan-Ho algorithm by skipping its second and third steps. The spherical intersection method was further extended by the spherical interpolation method [108, 109, 110] which takes the dependency of the intermediate variable into account by appropriate weighting functions. Spherical intersection and spherical interpolation methods, and also the closed-form solution developed for three available BSs according to [111], can be seen as special cases of the Chan-Ho algorithm with reduced accuracy and less complexity. Other variations of this approach can be found in [112], [113], and [114], where the principle of Lagrangian multipliers is exploited to obtain appropriate correction terms for considering the constraints introduced by the dependent intermediate variable.

Squared differences based approaches

The concept of processing squared timing differences instead of hyperbolic timing differences was proposed in [115, 116] and extensively discussed in [117, 118]. The basic idea of this approach is that by squaring the time differences the non-convex optimization problem can be turned into a convex optimization problem. Also here, a dependency of the introduced intermediate variables due to the squaring operation limits the performance.

Factor graphs based approaches

In [119, 120, 121] a factor graph based location estimation approach was proposed for spherical location estimation. According to the sum-product algorithm, mean values and variances of intermediate estimates were passed between the respective factor and variable nodes. In the factor nodes the various estimates are then processed and combined with each other by reflecting the spherical properties in (4.5). In [36], this approach was extended for supporting hyperbolic location estimation.

Bias mitigation approaches

The class of bias mitigation approaches includes algorithms that limit the influence of biased timing estimates coming from, e.g., NLOS or multipath propagation conditions. If it is not possible to mitigate these influences inside the timing estimation algorithms (e.g., by appropriate channel estimation algorithms), the resulting timing estimates have a positive or negative bias for multipath propagation and might have an additional positive bias under NLOS conditions. Two principles can be identified to cope with this problem. In the first, a priori knowledge of the bias statistics is required. This might also include the knowledge of the LOS/NLOS condition of the individual links between MS and BSs. It allows an appropriate weighting or setting of constraints to improve the overall location estimate [122, 123, 124]. A second idea is the exploitation of redundancy inside the available measurements. If more timing estimates are available than required, intermediate location estimates using only a subset of the available timing estimates can be computed. These intermediate estimates can then be combined to drop out outliers and to identify biased links [125, 126]. Note that this concept is denoted as receiver autonomous integrity monitoring (RAIM) in the context of GNSS (e.g., [6]). Even though the outlined approaches are developed for spherical location estimation, an extension to hyperbolic approaches should be straightforward. A comprehensive overview about this topic can be found in [127].

Direct location estimation

The basic problem of the discussed classical location estimation techniques is the requirement that timing information from at least three BSs is needed. Direct location estimation techniques rely solely on estimates obtained from a single BS. To do so, usually accurate knowledge about the propagation channel is required. The received signals are then directly processed to obtain a location estimate, without the intermediate estimate using an appropriate location dependent parameter like timing information. The concept was generally introduced in [128]. A concise evaluation of the direct location estimation concept and the extension to multiple antenna systems can be found in [129, 130].

Extension of iterative least squares approaches

Certainly, also the numerical methods described in Section 4.3.1 can be extended and improved for dedicated scenarios. Mentioned at this point should be the contribution in [131], where the Hessian matrix in the Newton algorithm is approximated by a less complex pseudo-Jacobian matrix, the contribution in [132], where semi-definite programming is applied to solve the non-linear optimization problem taking into account location constraints for the MS determined by the BS locations, and the approach discussed in [133, 134], where a total least squares approach is proposed to improve the performance in ill-conditioned geometric situations.

4.4 Cramer-Rao lower bound

The achievable accuracy for static location estimation depends on the geometric relation between MS and BSs as well as the quality of the timing estimates. The geometric relation is included in the Jacobian matrix $\boldsymbol{\Phi}(\boldsymbol{p})$ that — according to (4.23) and Appendix A.2 — solely depends on the locations of MS and BSs. The quality of the timing estimates is included in the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_f}$ (cf. Table 4.1). The resulting CRLB for the parameter vector $\boldsymbol{p}$ can then be computed

(cf. [12]) as the matrix

$$\boldsymbol{\Xi}\left(\boldsymbol{p}\right) = \left(\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{p}\right)\right)^{-1} \in \mathbb{R}^{N_{\mathrm{par}}\times N_{\mathrm{par}}}, \tag{4.51}$$

where the diagonal elements of $\boldsymbol{\Xi}\left(\boldsymbol{p}\right)$ includes the achievable accuracy of the individual components of $\boldsymbol{p}$. It was shown in [96, 97] that the CRLBs for spherical processing with unknown time offset and hyperbolic processing are related by

$$\boldsymbol{\Xi}^{\mathrm{HYP}}\left(\boldsymbol{p}\right) = \begin{bmatrix}\mathbf{I}_3 & \mathbf{0}_{3\times 1}\end{bmatrix}\boldsymbol{\Xi}^{\mathrm{SPHT}}\left(\boldsymbol{p}\right)\begin{bmatrix}\mathbf{I}_3 \\ \mathbf{0}_{1\times 3}\end{bmatrix}. \tag{4.52}$$

Hence, there is no difference in the x-, y-, and z-component of the CRLB for both classifications, i.e., the location estimation accuracy is identical for spherical processing with unknown time offset and hyperbolic processing. This is in line with the expectations as both approaches are based on the same timing information. Contrary to that, the achievable accuracy for spherical processing is always higher than the accuracy for hyperbolic processing as the time offset between MS and BS clocks is known for spherical processing, i.e., a priori knowledge is available [96, 97]. Therefore, in the following we solely focus on spherical and hyperbolic processing. A suitable performance bound for the absolute location estimation error in meters is then the definition

$$\mathrm{CRLB} = c_0\sqrt{\mathrm{trace}\left(\left(\boldsymbol{\Phi}^{\mathrm{T}}\left(\boldsymbol{p}\right)\boldsymbol{\Sigma}_{\boldsymbol{n}_f}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{p}\right)\right)^{-1}\right)}. \tag{4.53}$$

Figure 4.2 shows the CRLB for two-dimensional spherical location estimation exploiting timing information obtained from the nearest $N_{\mathrm{BS}} = 3$ BSs. The chosen scenario corresponds to the introduced cellular network according to Figure 3.16. For simplification, we assume that the noise on the timing estimates on each MS-BS-link corresponds to a standard deviation of $50\,\mathrm{m}$, i.e., $\boldsymbol{\Sigma}_{\boldsymbol{n}_\tau} = \sigma_{\boldsymbol{n}_\tau}^2\mathbf{I}_3$ with $\sigma_{\boldsymbol{n}_\tau} = 50\,\mathrm{m}/c_0$. Figure 4.2 clearly shows the dependency of the achievable location estimation accuracy on the location of the MS in the network. At the cell edge the CRLB is much smaller compared to the situation close to a BS. Furthermore, we observe that the optimum at the cell edge (around $55\,\mathrm{m}$) is smaller than the noise standard deviation corresponding to $50\,\mathrm{m}$, i.e., a dilution of precision by the geometric constellation takes place.

Figure 4.3 shows the resulting CRLB for two-dimensional hyperbolic location estimation. We observe that the overall achievable accuracy in all possible locations in the network is smaller compared to spherical location estimation as already shown in [96, 97]. More investigations on the difference between spherical and hyperbolic processing under geometric aspects can be found in [135, 136, 137, 138].

4.5 Simulation results

We start with an investigation of the convergence behavior for the numerical methods as described in Section 4.3.1. As simplified simulation model for the timing estimates we use a fixed standard deviation of the noise according to $\sigma_{\boldsymbol{n}_\tau} = 50\,\mathrm{m}/c_0$ (cf. Section 4.4). As initial estimate, i.e., starting point for the iterations, we use the fallback approach yielding the geometric mean value of the $N_{\mathrm{BS}} = 3$ considered nearest BSs. Performance criterion is the RMSE of the static location estimate, i.e.,

$$\mathrm{RMSE} = \sqrt{\left\|\boldsymbol{x} - \hat{\boldsymbol{x}}\right\|_2} \geq \mathrm{CRLB} \tag{4.54}$$

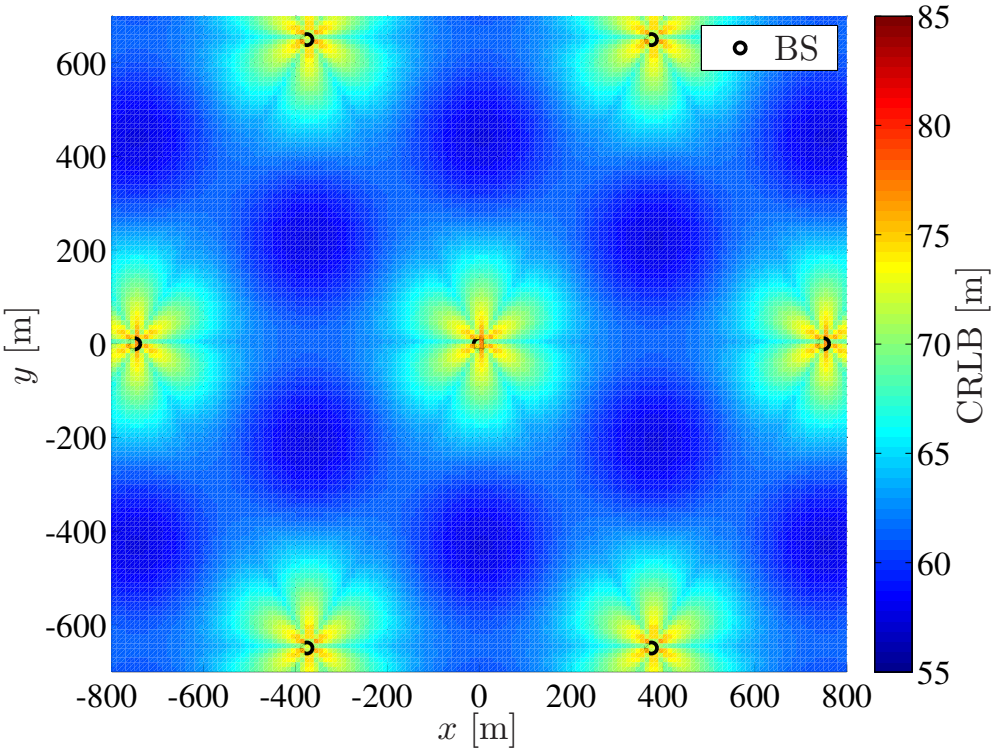


Figure 4.2: CRLB for spherical location estimation

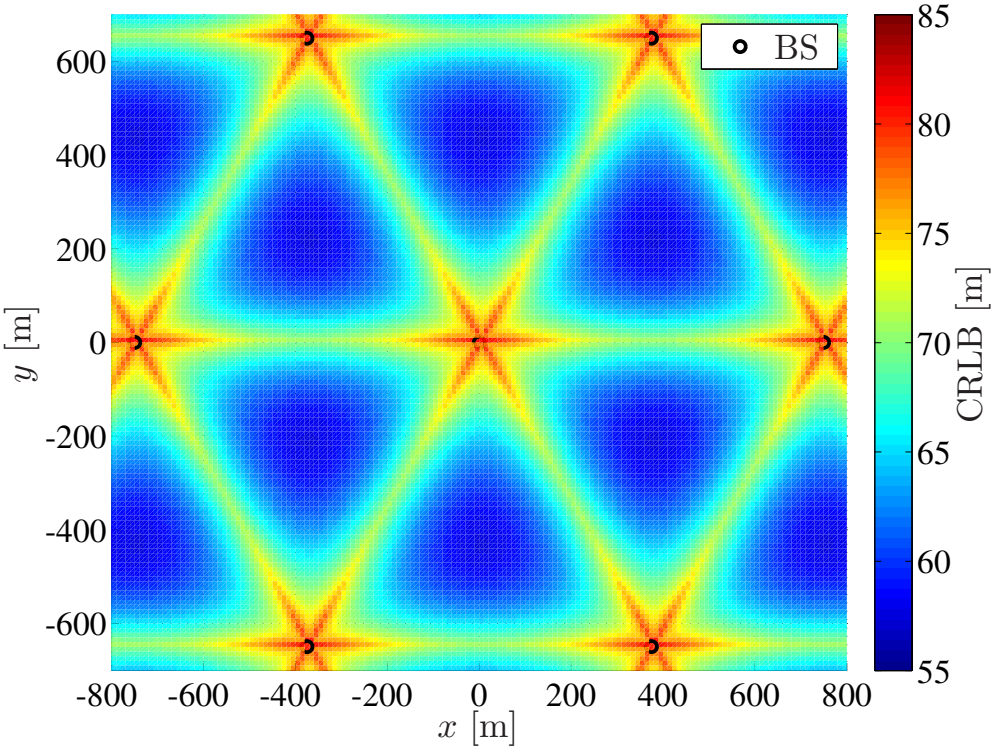


Figure 4.3: CRLB for hyperbolic location estimation

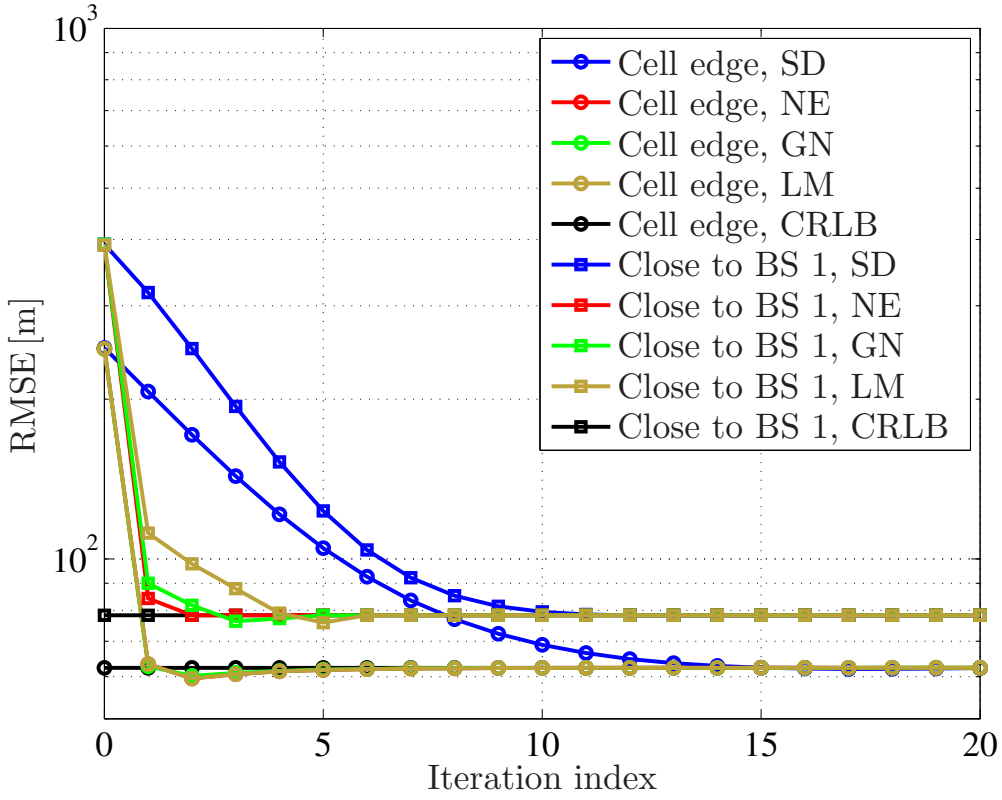


Figure 4.4: RMSE over number of iterations for different numerical methods, spherical location estimation

Figure 4.4 shows the RMSE over the number of iterations for two-dimensional spherical location estimation, where it was averaged over several noise realizations. The MS was located at the cell edge ($x = 500\,\text{m}$) and close to BS 1 ($x = 50\,\text{m}$), corresponding to the cellular network depicted in Figure 3.16. For a MS location at the cell edge, Newton (NE), Gauss-Newton (GN), and Levenberg-Marquardt (LM) methods converge after around 3 iterations, where the CRLB can be achieved. The steepest descent (SD) method shows a much slower convergence and it achieves CRLB after around 14 iterations. Contrary to that, close to BS 1 a more diverse behavior of the methods can be observed, however, also here all methods achieve the CRLB. We further can observe the difference in the achievable performance at the two considered MS locations as already investigated in Section 4.4.

Figure 4.5 depicts the corresponding curves for two-dimensional hyperbolic location estimation. Here, the different convergence behavior becomes much more obvious. At the cell edge, all methods achieve CRLB after a certain number of iterations, however, in average more iterations compared to spherical processing are required. Close to BS 1 we have the situation that in the Gauss-Newton algorithm the matrix to be inverted becomes close to singular resulting in numerical instabilities, and hence, no reliable location estimates for certain noise realizations. In such a situation the final location estimate is set to the initial value which results in a deviation from the CRLB using the Gauss-Newton method. Hence, the Gauss-Newton method — even though exploited in many practical implementations — is not a robust approach especially in geometric ill-conditioned situations (e.g., close to a BS). The other algorithms outperform Gauss-Newton in such situation. Note that the algorithm developed by Chan and Ho (not shown here for the sake of clarity) achieves the CRLB at the cell edge and achieves an RMSE of around 98 m close to BS 1.

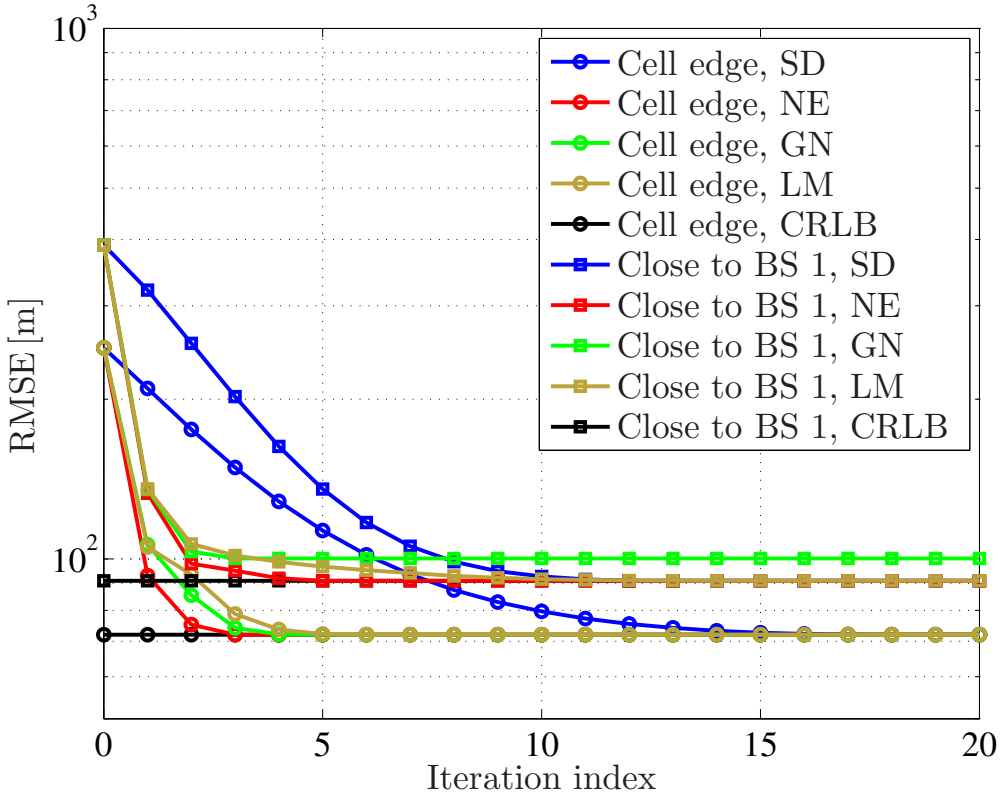


Figure 4.5: RMSE over number of iterations for different numerical methods, hyperbolic location estimation

In the following we investigate the average performance of the different algorithms at MS locations in the whole network, i.e., we average over all MS locations in the grey shaded area depicted in Figure 3.16 and determine the CDF of the location estimation error

$$\mathrm{CDF}\left(x_{\mathrm{error}}\right) = \mathrm{P}\left(\left\|\hat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 < x_{\mathrm{error}}\right). \tag{4.55}$$

Hence, this CDF is the probability that the location estimation error is smaller than the value of the abscissae x_{error}. For spherical processing the results are shown in Figure 4.6, where still a fixed standard deviation of the noise according to $\sigma_{\boldsymbol{n}_\tau} = 50\,\mathrm{m}/c_0$ is used. We observe no fundamental difference between the different algorithms and all are very close to the CRLB. This is in line with the results of the convergence behavior in Figure 4.4, where all algorithms converge to the CRLB. In around 92% of the situations we have a location estimation error of $100\,\mathrm{m}$ or less. For comparison, also the results for the fallback (FB) solution are depicted. It is obvious that this approach can not be used for accurate location estimates.

Figure 4.7 shows the corresponding results for hyperbolic processing. Here, we can observe a small degradation compared to spherical processing and only in around 90% of the situations a location estimation error smaller than $100\mathrm{m}$ can be achieved. We can see that steepest descent, Newton, and Levenberg-Marquardt algorithm perform very similar for this setup and are very close to the CRLB. The Gauss-Newton algorithm shows a small degradation (around $15\,\mathrm{m}$ for the 90%-error) and is even outperformed by the closed-form approach by Chan-Ho.

In the following, we assess the location estimation performance with the synchronization algorithms as studied in Chapter 3 for hyperbolic processing of the timing estimates. In that manner, the cellular network according to Figure 3.16 was simulated at the respective MS locations and the

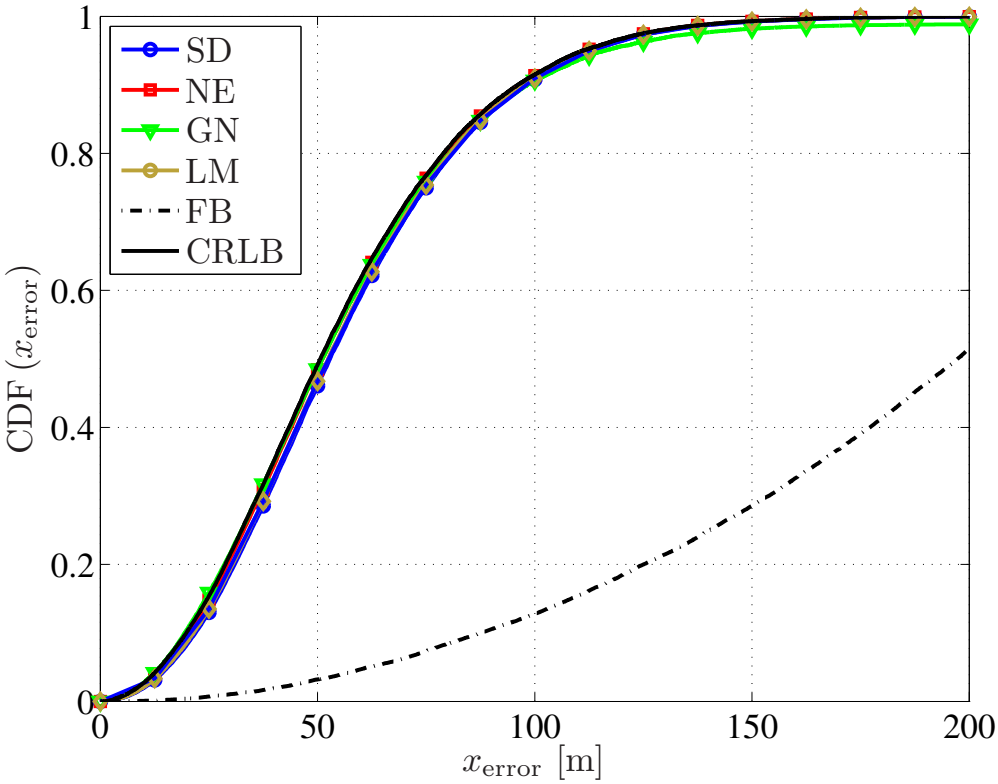


Figure 4.6: CDF of spherical location estimation error, different numerical methods

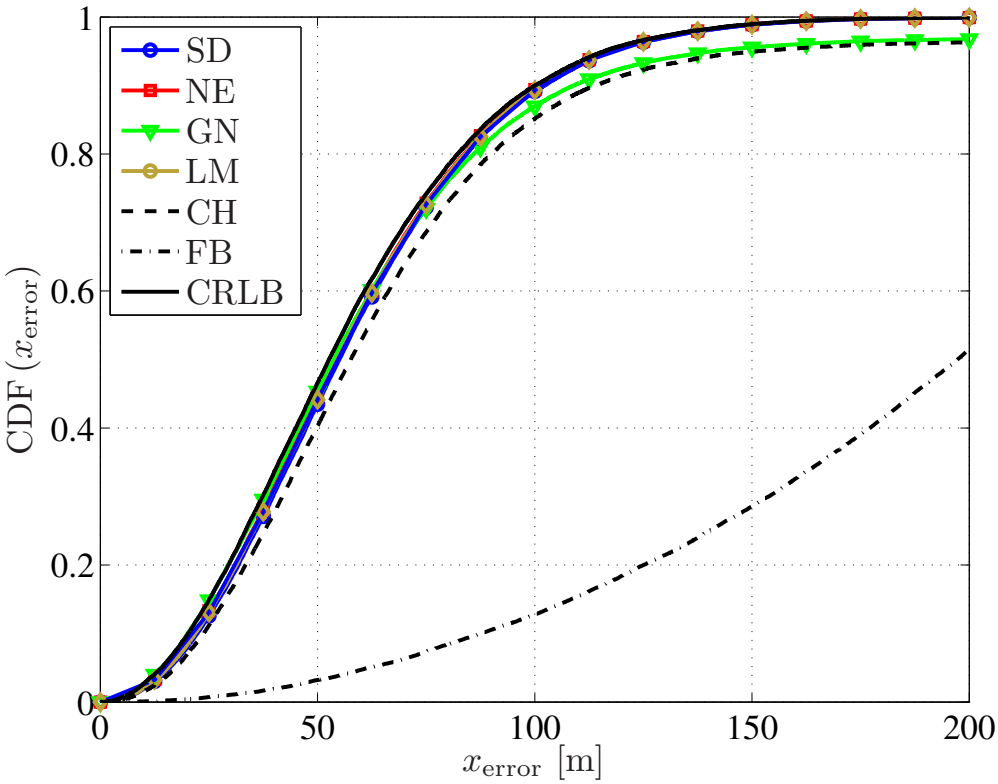


Figure 4.7: CDF of hyperbolic location estimation error, different numerical methods

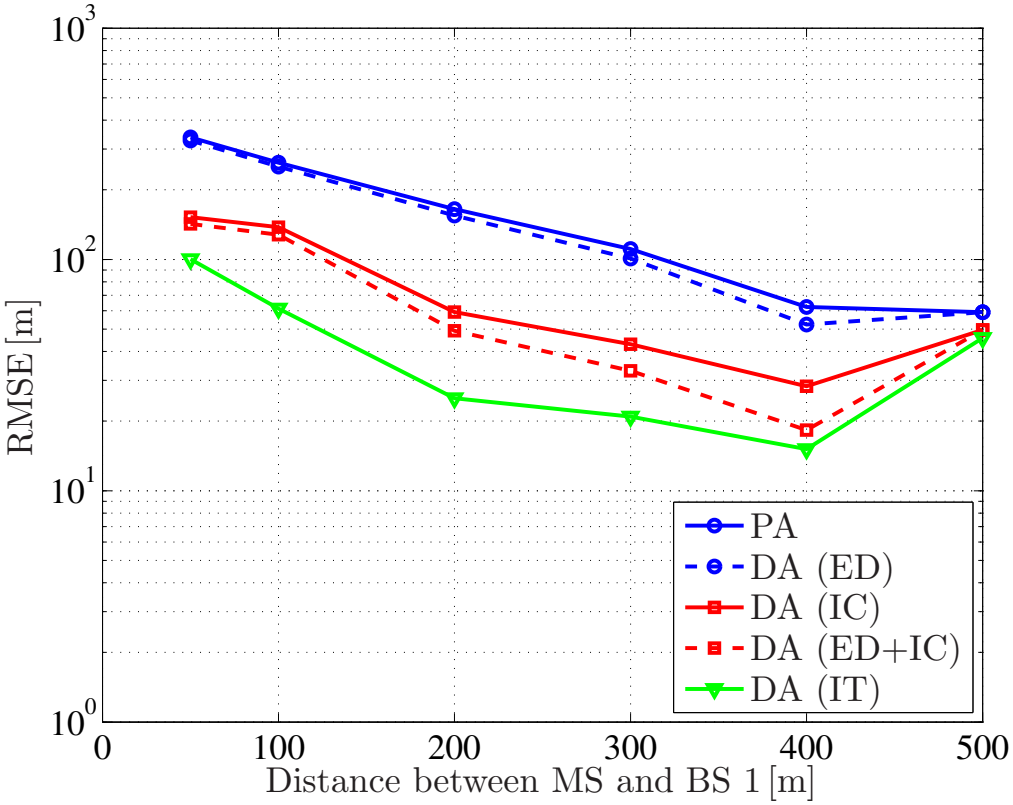


Figure 4.8: RMSE over distance for static hyperbolic location estimation, AWGN channel

average location RMSE was determined depending on the different synchronization approaches (cf. Figure 3.26) under AWGN and exploiting Levenberg-Marquardt as static location estimation approach. Note that from now on hyperbolic processing and Levenberg-Marquardt as solution for the static location estimation problem will be the baseline approaches in the rest of this thesis. As expected, we see that with the standard pilot-aided synchronization (using coherent cross-correlation in the frequency domain) only at the cell edge reliable location estimates are possible (around 60 m). Close to BS 1 the location error is higher than 300 m. This performance can slightly be improved if data-aided synchronization using estimated data (ED) is applied for timing estimation with BS 1. Nevertheless, no more than 10 m of performance improvement can be achieved by this method. A much higher performance gain can be obtained by application of data-aided synchronization with interference cancelation (IC) and a combination of both approaches (ED+IC). Whereas the RMSE close to BS 1 is still higher than 100 m the overall error in the cell core could be reduced remarkably. When applying data-aided interference cancelation in an iterative (IT) way another performance improvement of up to 50 m is possible.

For the results in Figure 4.9 we averaged over several MS locations in the network and determined the location estimation error CDF. As channel model we used a combination of the multipath LOS and NLOS channel models, where a distance dependent LOS probability was assumed (cf. Appendix A.1). For pilot-aided synchronization, the 90%-error is higher than 300 m in the considered setup. For the same algorithm, in 40% of the situations a location error smaller than 50 m could be achieved. The performance improvement by data-aided synchronization with estimated data is only negligible small. When we apply data-aided synchronization with interference cancelation of BS 1, in 62% of the situations the error is smaller than 50 m. For a combination

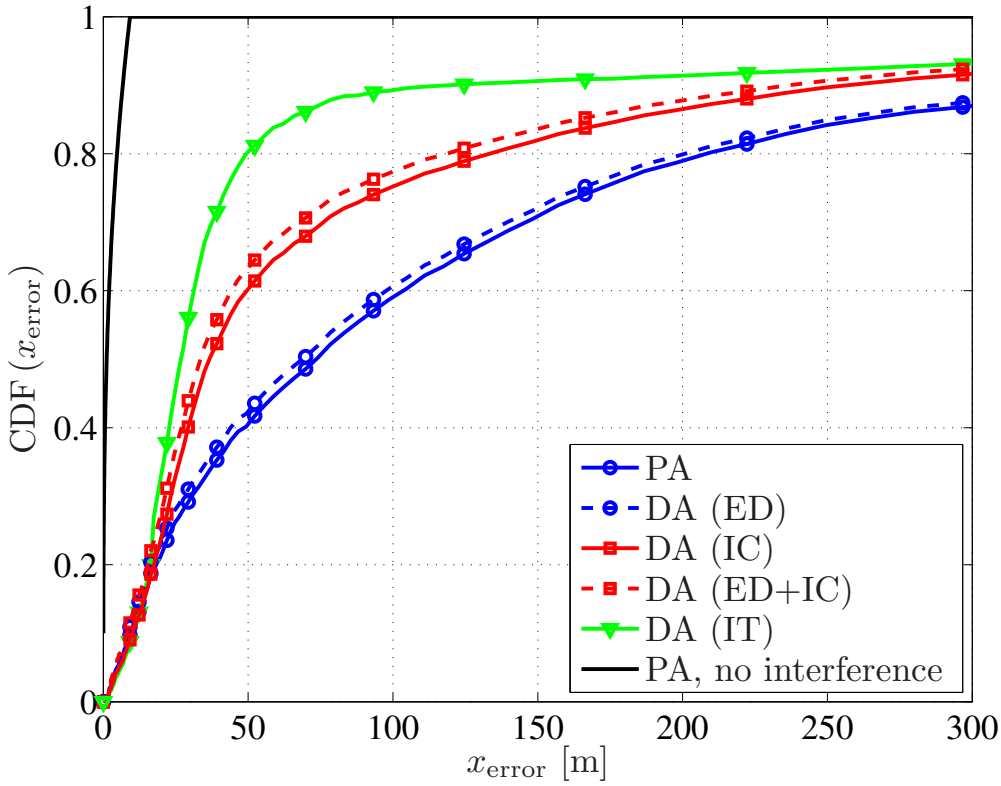


Figure 4.9: CDF of static hyperbolic location estimation error, multipath LOS/NLOS channel

of estimated data and interference cancelation this can slightly be improved. For a subsequent interference cancelation in an iterative manner this can finally be improved to 80%. For comparison, also the interference-free situation was simulated. The resulting error is always smaller than around $10\,\text{m}$. We further can see, that all CDFs do not achieve 100%, i.e., in some situations (even for the best approach the gap is around 10%) the location estimation error is higher than $300\,\text{m}$. Reason for that are mainly outliers of the timing and/or location estimation process. They can be compensated by the application of suitable dynamic location estimation algorithms as discussed in the next chapter.

5 Dynamic location estimation with timing information

The location estimation techniques presented in the previous chapter consider a static solution. It was assumed, that the MS is not moving during the location estimation process, and therefore, the MS location was treated as a deterministic parameter. In practical systems, the MS locations are usually correlated over time. For instance — considering a pedestrian user or a driving car — certain information about the actual location can be derived by exploiting the history of past estimates and suitable movement or mobility models. This especially includes a restricted movement behavior of the MS, e.g., a MS belonging to a pedestrian can not jump from one location to another in limited time or a MS belonging to a car can change its direction only smoothly. This behavior can be used as side-information for location tracking algorithms. The overall approach has to be seen in the context of the two-step location determination approach as shown in Figure 4.1. Section 5.1 introduces the system model applied for dynamic location estimation in this chapter. In Section 5.2 suitable estimation criteria are presented. Section 5.3 derives the respective tracking algorithms, in particular the Kalman filter, the extended Kalman filter, and the particle filter are discussed with a focus on dynamic location estimation. Simulation results evaluate the overall performance in Section 5.4.

5.1 System model

For the derivation of the algorithms, we assume that the time axis is divided in discrete time intervals. Further, we presume a causal system, i.e., future states (e.g., defined by location or velocity) can not impact current and past estimates. However, since the past states can impact the current and future states, this property has to be reflected in the chosen model. A commonly used model in the considered dynamic location estimation or location tracking context is a first order hidden Markov model.

Figure 5.2 depicts such a Markov model (e.g., [139]) with unknown states $\boldsymbol{s}_k \in \mathbb{R}^{N_s}$ that have to be estimated in each time-step $k \in \mathbb{N}$. It is a hidden Markov model since the states can only be observed implicitly in terms of the available observations or measurements. The estimation

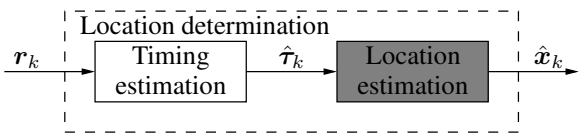


Figure 5.1: Two-step location determination process: dynamic location estimation

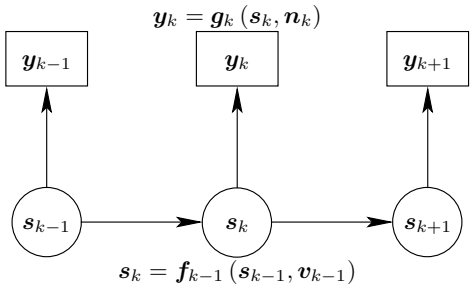


Figure 5.2: First order hidden Markov model

process takes into account these observations $\boldsymbol{y}_k \in \mathbb{R}^{N_y}$ in each time-step k in addition to the model parameters.

According to Figure 5.2, the observations $\boldsymbol{y}_k$ depend only on the state vector $\boldsymbol{s}_k$ at the current time-step. This dependency is defined by the so-called observation model

$$\boldsymbol{y}_k = \boldsymbol{g}_k(\boldsymbol{s}_k, \boldsymbol{n}_k). \tag{5.1}$$

The function

$$\boldsymbol{g}_k : \mathbb{R}^{N_s} \times \mathbb{R}^{N_n} \to \mathbb{R}^{N_y} \tag{5.2}$$

is a possibly non-linear function of the state $\boldsymbol{s}_k$ and the observation noise $\boldsymbol{n}_k \in \mathbb{R}^{N_n}$ (cf. [140]). The properties of the observation noise $\boldsymbol{n}_k$ define the observation uncertainties. Another equivalent representation of the observation model is based on the conditioned PDF of the observations given the states, i.e., $\mathrm{p}(\boldsymbol{y}_k|\boldsymbol{s}_k)$.

The state model defines a relation between the previous state $\boldsymbol{s}_{k-1}$ and the current state $\boldsymbol{s}_k$. It is given as

$$\boldsymbol{s}_k = \boldsymbol{f}_{k-1}(\boldsymbol{s}_{k-1}, \boldsymbol{v}_{k-1}), \tag{5.3}$$

where the function

$$\boldsymbol{f}_{k-1} : \mathbb{R}^{N_s} \times \mathbb{R}^{N_v} \to \mathbb{R}^{N_s} \tag{5.4}$$

is a possibly non-linear function of the state $\boldsymbol{s}_{k-1}$ and the state process noise $\boldsymbol{v}_{k-1} \in \mathbb{R}^{N_v}$ (cf. [140]). The properties of the state process noise $\boldsymbol{v}_{k-1}$ define how correlated the state changes can be. The equivalent representation of the state model is based on the conditioned PDF $\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{s}_{k-1})$. In the location estimation context the state vector can include information about the MS location or its velocity. The corresponding state model is determined by the mobility or movement behavior of the MS. Therefore, it is often denoted as mobility model.

Following the Bayesian approach (e.g., [12, 20, 140]), it is required that the PDF of the current state is estimated by considering all previous and the current observations, i.e., the PDF $\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k)$ has to be constructed. This is done recursively by assuming that the a priori distribution of the state $\boldsymbol{s}_0$ is known. To do so, in the first step of Bayesian estimation the state model is used to obtain the a priori PDF of the state at time-step k by

$$\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}) = \int_{\boldsymbol{s}_{k-1}} \mathrm{p}(\boldsymbol{s}_k|\boldsymbol{s}_{k-1})\, \mathrm{p}(\boldsymbol{s}_{k-1}|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1})\, \mathrm{d}\boldsymbol{s}_{k-1}. \tag{5.5}$$

The state PDF $\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{s}_{k-1})$ is defined by the state equation and the known statistics of the state noise $\boldsymbol{v}_{k-1}$. This step is denoted as prediction step since the new state is estimated as a prediction of the old state. At this stage the current observations are not yet used.

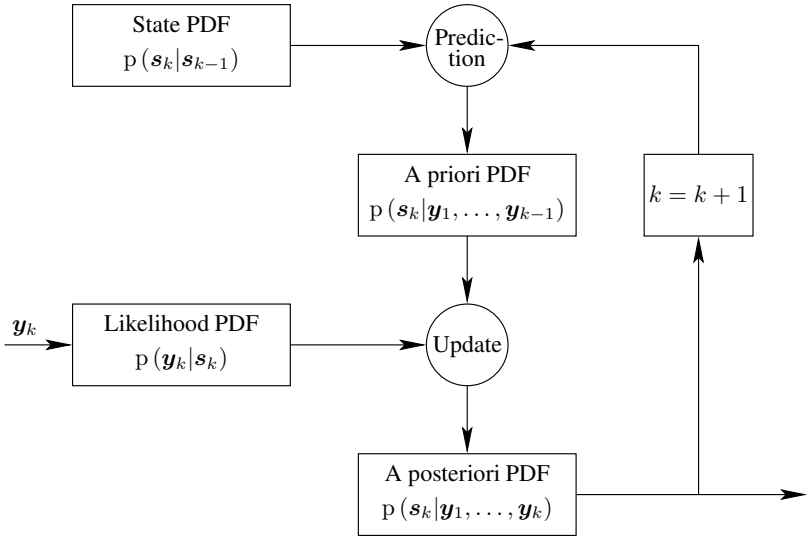


Figure 5.3: Recursive Bayesian estimation

For the second step it is required that at time-step k the observations $\boldsymbol{y}_k$ become available. They can be used to update the a priori PDF by the Bayesian rule resulting in a normalized product of the likelihood PDF $\mathrm{p}(\boldsymbol{y}_k|\boldsymbol{s}_k)$ and the a priori PDF, i.e.,

$$\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_k) = \frac{\mathrm{p}(\boldsymbol{y}_k|\boldsymbol{s}_k)\,\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_{k-1})}{\mathrm{p}(\boldsymbol{y}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_{k-1})}, \tag{5.6}$$

with the normalization constant

$$\mathrm{p}(\boldsymbol{y}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_{k-1}) = \int_{\boldsymbol{s}_k} \mathrm{p}(\boldsymbol{y}_k|\boldsymbol{s}_k)\,\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_{k-1})\,\mathrm{d}\boldsymbol{s}_k. \tag{5.7}$$

Therefore, the a posteriori PDF can be calculated by applying the observation model and the known statistics of the observation noise $\boldsymbol{n}_k$. Since the observations of time-step k are used to modify the a priori PDF for obtaining the a posteriori PDF, this step is called update step. The complete principle of the recursive Bayesian estimation is concluded in Figure 5.3 (cf. [139]).

5.2 Estimation criteria

The solution that maximizes the a posteriori PDF is the MAP estimator

$$\hat{\boldsymbol{s}}_{\mathrm{MAP},k} = \underset{\boldsymbol{s}_k}{\operatorname{argmax}}\, \mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_k)\,. \tag{5.8}$$

Contrary to that, the minimum mean squared error (MMSE) estimator computes the expectation of the PDF, i.e.,

$$\hat{\boldsymbol{s}}_{\mathrm{MMSE},k} = \int_{\boldsymbol{s}_k} \boldsymbol{s}_k\,\mathrm{p}(\boldsymbol{s}_k|\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_k)\,\mathrm{d}\boldsymbol{s}_k, \tag{5.9}$$

where for Gaussian noise distributions both estimators yield the same result.

Generally, there is no closed-form representation of the required PDFs in Bayesian estimation [140]. One option to cope with this problem are certain assumptions for the models or approximations: the classical Kalman filter approach presumes linear models and Gaussian noise distributions. It is described in Section 5.3.1. The extended Kalman filter approximates non-linear models in a linearization step which will be shown in Section 5.3.2. Another option to compute the PDFs is numerical integration. An approximate solution by Monte-Carlo methods is the particle filter being described in Section 5.3.3. All filters will be examined with a special focus on dynamic location estimation or location tracking applications.

5.3 Estimation algorithms

5.3.1 Kalman filter

The Kalman filter (cf. [19, 12]) is one of the most widely used implementations of Bayesian filters. It is a generalization of the Wiener filter, where the restriction that signal and noise have to be stationary is no longer necessary. It is a Bayesian sequential MMSE estimator of a signal embedded in noise, where the signal is characterized by a state model. For that model we focus on the first order hidden Markov model described before. One of the main advantages of Kalman filters is the computational efficiency in the implementation and its inherent robustness against model inaccuracies.

To perform optimum, the following assumptions must hold for the state model in (5.3) and the observation model in (5.1):

- State process noise $\boldsymbol{v}_k \sim \mathcal{N}\left(\mathbf{0}_{N_s}, \boldsymbol{Q}_k\right)$ and observation noise $\boldsymbol{n}_k \sim \mathcal{N}\left(\mathbf{0}_{N_y}, \boldsymbol{C}_k\right)$ are drawn from a zero-mean Gaussian distribution with known covariances. Note that for a general derivation it would not even be required that the noise is zero-mean (e.g., [12]), however, this is assumed here for simplification.
- Function $\boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}, \boldsymbol{v}_{k-1}\right)$ is a known linear function of $\boldsymbol{s}_{k-1}$ and $\boldsymbol{v}_{k-1}$.
- Function $\boldsymbol{g}_k\left(\boldsymbol{s}_k, \boldsymbol{n}_k\right)$ is a known linear function of $\boldsymbol{s}_k$ and $\boldsymbol{n}_k$.

If these requirements are met, the Kalman filter is the optimum MMSE estimator [12]. If these assumptions do not hold, it is still the optimum linear MMSE estimator.

For the fulfilled assumptions, we can rewrite (5.3) and (5.1) as

$$\boldsymbol{s}_k = \boldsymbol{F}_{k-1}\boldsymbol{s}_{k-1} + \boldsymbol{v}_{k-1} \tag{5.10}$$

and

$$\boldsymbol{y}_k = \boldsymbol{G}_k\boldsymbol{s}_k + \boldsymbol{n}_k. \tag{5.11}$$

The matrix $\boldsymbol{F}_k \in \mathbb{R}^{N_s \times N_s}$ is denoted as state matrix and includes the linear dependencies between the states of time-steps k and $k-1$. The observation matrix $\boldsymbol{G}_k \in \mathbb{R}^{N_y \times N_s}$ reflects the linear relation between the observations and the state at time-step k. In general, all matrices can be time-variant. In the context of location estimation applications this could reflect, e.g., changing mobility models over time or a changing number of available observations. Since all PDFs (including a priori, a posteriori, and likelihood) are Gaussian, they can be represented by mean values and covariance matrices. This allows a simple derivation of the estimates in terms of matrix-vector notation. The optimum filter equations can then be written as follows (cf. [12]):

In a first step (prediction) the state of the current time-step is calculated taking into account the state of the previous time-step and the knowledge of the state matrix $\boldsymbol{F}_{k-1}$. Then, the estimate of the state after prediction is

$$\hat{\boldsymbol{s}}_{k|k-1} = \boldsymbol{F}_{k-1}\hat{\boldsymbol{s}}_{k-1|k-1}, \tag{5.12}$$

with the estimate of the previous time-step $\hat{\boldsymbol{s}}_{k-1|k-1}$. Additionally, the corresponding MMSE or covariance matrix after that prediction step can be calculated as

$$\boldsymbol{M}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{M}_{k-1|k-1}\boldsymbol{F}_{k-1}^{\mathrm{T}} + \boldsymbol{Q}_k \in \mathbb{R}^{N_s \times N_s}, \tag{5.13}$$

where $\boldsymbol{M}_{k-1|k-1}$ is the MMSE matrix of the previous time-step. From the Bayesian PDF point of view, the a priori PDF in (5.5) can be represented as a Gaussian distribution according to

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\right) \sim \mathcal{N}\left(\hat{\boldsymbol{s}}_{k|k-1}, \boldsymbol{M}_{k|k-1}\right). \tag{5.14}$$

The Kalman gain matrix includes a weighting between the predicted estimate and the current observations. It is given as

$$\boldsymbol{K}_k = \boldsymbol{M}_{k|k-1}\boldsymbol{G}_k^{\mathrm{T}}\left(\boldsymbol{C}_k + \boldsymbol{G}_k\boldsymbol{M}_{k|k-1}\boldsymbol{G}_k^{\mathrm{T}}\right)^{-1} \in \mathbb{R}^{N_s \times N_y}. \tag{5.15}$$

Finally, the correction step combines the predicted estimates with the current observations weighted with the Kalman gain matrix. This results in the final estimate of the state vector

$$\hat{\boldsymbol{s}}_{k|k} = \hat{\boldsymbol{s}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{G}_k\hat{\boldsymbol{s}}_{k|k-1}\right). \tag{5.16}$$

The corresponding MMSE or covariance matrix after the correction step is obtained as

$$\boldsymbol{M}_{k|k} = \left(\mathbf{I}_{N_s} - \boldsymbol{K}_k\boldsymbol{G}_k\right)\boldsymbol{M}_{k|k-1}. \tag{5.17}$$

The resulting a posteriori PDF can then be written as Gaussian distribution according to

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) \sim \mathcal{N}\left(\hat{\boldsymbol{s}}_{k|k}, \boldsymbol{M}_{k|k}\right). \tag{5.18}$$

The Kalman filter is initialized with $\boldsymbol{s}_{0|0}$ and $\boldsymbol{M}_{0|0}$ determined by the a priori distribution of the initial state. We observe that the MMSE matrix can be computed independently of the state estimates. Furthermore, it only depends on model parameters and not on the actual observations. Hence, it can be calculated in advance or off-line and provides the expected accuracy for the state estimates over time without processing the observations.

In the following, we apply the Kalman filter described before explicitly for dynamic location estimation. To do so, we assume that the state vector that has to be estimated consists of the three-dimensional location and velocity of the MS at time-step k, i.e.,

$$\boldsymbol{s}_k = \left[x_k, y_k, z_k, v_{x,k}, v_{y,k}, v_{z,k}\right]^{\mathrm{T}}. \tag{5.19}$$

As mobility model, we choose an approach corresponding to random force [141]. For that, the resulting state matrix is given as

$$\boldsymbol{F}_k = \boldsymbol{F} = \begin{bmatrix} \mathbf{I}_3 & T_{\text{samp, dyn}}\mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}. \tag{5.20}$$

The sampling time $T_{\text{samp, dyn}}$ highly depends on the application and describes the update rate of the filter. This includes the observation period as well as the provision of location estimates. For pedestrian applications sampling times of around 1 s are usually sufficient. The covariance matrix of the state noise is the diagonal matrix

$$\boldsymbol{Q}_k = \boldsymbol{Q} = \sigma_Q^2 \begin{bmatrix} \left(T_{\text{samp, dyn}}^2/2\right) \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & T_{\text{samp, dyn}}^2 \mathbf{I}_3 \end{bmatrix}. \tag{5.21}$$

It includes the variance of the mobility (process drift) in x-, y-, and z-direction for location and velocity. This model implies that the change of the MS location is controlled by state noise with a certain variance σ_Q^2. This mobility model is often applied because of its simplicity. Nevertheless, it is inadequate to reflect MSs with high mobility, i.e., MSs with rapidly changing directions or velocities. A comprehensive overview of more sophisticated state mobility models can be found in [141].

For the observation model, we assume that in every time-step a location estimate is available. This could be realized by the algorithms for static location estimation as derived in Chapter 4. Hence, the observations are given in terms of location estimates and have a linear dependency with respect to the state vector which is reflected in the observation matrix

$$\boldsymbol{G}_k = \boldsymbol{G} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_3 \end{bmatrix}. \tag{5.22}$$

We do not have available any velocity estimates as the algorithms in Chapter 4 solely provide location estimates. Therefore, the velocity is handled as hidden state and is estimated implicitly in the filter equations without using any direct velocity estimates. Finally, the covariance matrix of the observation noise $\boldsymbol{C}_k$ consists of the variance of the static location estimates and can be approximated by exploiting the static CRLB as described in Section 4.4.

5.3.2 Extended Kalman filter

The performance of the Kalman filter is optimum if the conditions on Gaussianity and linearity are fulfilled completely. However, simulation results for static location estimation (e.g., in Figure 4.5) have shown, that in certain situations (e.g., if the MS is close to a BS) the performance is limited. Further, the Kalman filter requires that the underlying entity which provides the static location estimates performs optimum, i.e., in every time-step reasonably good timing estimates have to be available and the resulting estimates have further to fulfill the Gaussian assumption. Especially in critical location estimation situations (e.g., in an urban canyon or indoor) it may happen quite often that only less then the required number of timing estimates are available for a certain time. Then, the Kalman filter would totally fail since the static solution can not provide any reliable estimate.

The extended Kalman filter (e.g., [142, 12]) is a much more flexible tool being able to handle directly non-linear models. We assume that the state model is given as

$$\boldsymbol{s}_k = \boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}\right) + \boldsymbol{v}_{k-1} \tag{5.23}$$

and the observation model can be written as

$$\boldsymbol{y}_k = \boldsymbol{g}_k\left(\boldsymbol{s}_k\right) + \boldsymbol{n}_k. \tag{5.24}$$

Basic idea of the extended Kalman filter is a linearization of $\boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}\right)$ around the estimate of $\boldsymbol{s}_{k-1}$. We obtain

$$\boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}\right) \approx \boldsymbol{f}_{k-1}\left(\hat{\boldsymbol{s}}_{k-1|k-1}\right) + \boldsymbol{F}_{k-1}\left(\boldsymbol{s}_{k-1} - \hat{\boldsymbol{s}}_{k-1|k-1}\right), \tag{5.25}$$

with the Jacobian matrix

$$\boldsymbol{F}_{k-1} = \left.\frac{\partial \boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}\right)}{\partial \boldsymbol{s}_{k-1}}\right|_{\boldsymbol{s}_{k-1}=\hat{\boldsymbol{s}}_{k-1|k-1}} \in \mathbb{R}^{N_s \times N_s}. \tag{5.26}$$

Equivalently, we linearize $\boldsymbol{g}_k\left(\boldsymbol{s}_k\right)$ about the estimate of $\boldsymbol{s}_k$, i.e.,

$$\boldsymbol{g}_k\left(\boldsymbol{s}_k\right) \approx \boldsymbol{g}_k\left(\hat{\boldsymbol{s}}_{k|k-1}\right) + \boldsymbol{G}_k\left(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_{k|k-1}\right), \tag{5.27}$$

with the Jacobian matrix

$$\boldsymbol{G}_k = \left.\frac{\partial \boldsymbol{g}_k\left(\boldsymbol{s}_k\right)}{\partial \boldsymbol{s}_k}\right|_{\boldsymbol{s}_k=\hat{\boldsymbol{s}}_{k|k-1}} \in \mathbb{R}^{N_y \times N_s}. \tag{5.28}$$

Obviously, the Jacobian matrices have to be re-calculated in every time-step since they depend on the estimates of the previous time-steps. However, the resulting structure of the extended Kalman filter as pointed out in the following is very similar to the Kalman filter solution [12]:

It starts with the prediction, where knowledge of the state model is applied to obtain

$$\hat{\boldsymbol{s}}_{k|k-1} = \boldsymbol{f}_k\left(\hat{\boldsymbol{s}}_{k-1|k-1}\right), \tag{5.29}$$

with the estimate of the previous time-step $\hat{\boldsymbol{s}}_{k-1|k-1}$. Similarly, the corresponding MMSE or covariance matrix after that prediction step is

$$\boldsymbol{M}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{M}_{k-1|k-1}\boldsymbol{F}_{k-1}^{\mathrm{T}} + \boldsymbol{Q}_k. \tag{5.30}$$

Due to the linearization step the resulting estimated a priori PDF in the Bayesian sense is a Gaussian approximation of the true a priori PDF. Hence, the estimated a priori PDF is given as

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\right) \approx \mathcal{N}\left(\hat{\boldsymbol{s}}_{k|k-1}, \boldsymbol{M}_{k|k-1}\right). \tag{5.31}$$

The Kalman gain matrix can be obtained by

$$\boldsymbol{K}_k = \boldsymbol{M}_{k|k-1}\boldsymbol{G}_k^{\mathrm{T}}\left(\boldsymbol{C}_k + \boldsymbol{G}_k\boldsymbol{M}_{k|k-1}\boldsymbol{G}_k^{\mathrm{T}}\right)^{-1}, \tag{5.32}$$

Finally, the correction step combines the predicted estimates with the current observations weighted with the Kalman gain matrix. This results in the final estimate of the state vector being computed as

$$\hat{\boldsymbol{s}}_{k|k} = \hat{\boldsymbol{s}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{g}_k\left(\hat{\boldsymbol{s}}_{k|k-1}\right)\right). \tag{5.33}$$

The corresponding MMSE or covariance matrix after correction is obtained as

$$\boldsymbol{M}_{k|k} = \left(\mathbf{I}_{N_s} - \boldsymbol{K}_k\boldsymbol{G}_k\right)\boldsymbol{M}_{k|k-1}. \tag{5.34}$$

Also the resulting a posteriori PDF is a Gaussian distribution of the true a posteriori PDF. It is given as

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) \approx \mathcal{N}\left(\hat{\boldsymbol{s}}_{k|k}, \boldsymbol{M}_{k|k}\right). \tag{5.35}$$

Since the MMSE matrix depends on the Jacobian matrices (we perform a dynamic linearization for the extended Kalman filter), an off-line calculation similar as for the Kalman filter is no longer possible. Further, the extended Kalman filter has no optimality properties, i.e., its accuracy depends

on the actual quality of the linearization. Nevertheless, the extended Kalman filter turns out to be a flexible and robust approach widely used for dynamic location estimation applications.

Considering the location estimation example (as started for the Kalman filter in Section 5.3.1), the state vector $\boldsymbol{s}_k$ is defined in the same way for spherical and hyperbolic processing. For spherical processing with unknown time offset it is defined as $\left[x_k, y_k, z_k, v_{x,k}, v_{y,k}, v_{z,k}, b_{\text{clock}}\right]^{\mathrm{T}}$ since the time offset needs to be estimated as an additional parameter. Further, we assume the same state model, i.e., a linearization of the state equation is not necessary. Therefore, the prediction step is the same for extended Kalman filter and Kalman filter. As observation we process directly the timing estimates — contrary to the Kalman filter, where these timing estimates were processed beforehand by a static location estimation. Since the timing estimates are non-linear with respect to the location, for the update step a linearization of the observation model is necessary. The resulting Jacobian matrices that are required for the extended Kalman filter computation are concluded in Appendix A.3. Note that these Jacobian matrices can be seen as time-variant versions of the Jacobian matrices for the static location estimation. Finally, the covariance matrix of the observation noise $\boldsymbol{C}_k$ consists of the variance of the timing estimates and can be approximated by exploiting the respective CRLB as described in Section 3.2.3.

5.3.3 Particle filter

Another important class of Bayesian filters is based on approximation of integrals by numerical integration. These methods are commonly denoted as particle filters (PFs) [20, 140, 143] and became quite popular for location tracking applications (e.g., [144, 145]). PFs are based on a sequential Monte-Carlo methodology (cf. [146]) and calculate recursively the relevant PDFs by importance sampling and approximation of PDFs with discrete random measures. The basic principle of particle filtering is the representation of the state PDF by a defined number of hypotheses, hence, it does not implement an analytical function. The PF approximates the optimum solution numerically based on the state model, rather than applying an optimum filter to an approximate model as it can be seen for the Kalman filter. Compared to Kalman filters the PFs have usually a much higher complexity depending on the number of particles that have to be generated to model the PDF. In addition, they can suffer from phenomena like sample degeneracy or sample impoverishment causing unstable behavior.

In PFs (cf. [145]), the a posteriori PDF is represented as the weighted sum

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) = \sum_{i=1}^{N_{\mathrm{p}}} w_k^i \delta\left(\boldsymbol{s}_k - \boldsymbol{s}_k^i\right), \tag{5.36}$$

where each particle consists of a state $\boldsymbol{s}_k^i$ and a weight w_k^i. The particles are drawn according to the principle of importance sampling from a proposal density $\mathrm{q}\left(\boldsymbol{s}_k|\boldsymbol{s}_k^i, \boldsymbol{y}_k\right)$. The corresponding weights can then be calculated by

$$w_k^i \sim w_{k-1}^i \frac{\mathrm{p}\left(\boldsymbol{y}_k|\boldsymbol{s}_k^i\right)\mathrm{p}\left(\boldsymbol{s}_k^i|\boldsymbol{s}_{k-1}^i\right)}{\mathrm{q}\left(\boldsymbol{s}_k|\boldsymbol{s}_k^i, \boldsymbol{y}_k\right)}. \tag{5.37}$$

The generic PF applies the optimum proposal density which in practice is difficult to determine. Therefore, often the so-called sampling importance resampling PF (SIR-PF) is implemented (cf. [147, 20, 140]). It only requires that the state and observation functions $\boldsymbol{f}_k\left(\cdot\right)$ and $\boldsymbol{g}_k\left(\cdot\right)$ are known, and that sampling of realizations from the state noise distribution of $\boldsymbol{v}_{k-1}$ as well as the

a priori distribution is possible. In addition, the likelihood function $\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_k\right)$ has to be available for pointwise evaluation. Hence, compared to the generic PF it can be said that for the SIR-PF the proposal density is chosen to be the a priori density according to $\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{s}_{k-1}^i\right)$.

In the first step of SIR-PF, for each particle $i = 1, 2, \ldots, N_\mathrm{p}$, a sample from the proposal density has to be drawn, i.e.,

$$\boldsymbol{s}_k^i \sim \mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{s}_{k-1}^i\right). \tag{5.38}$$

This can be realized by generating a state noise sample $\boldsymbol{v}_{k-1}^i$ with the corresponding PDF $\mathrm{p}\left(\boldsymbol{v}_{k-1}\right)$ and setting

$$\boldsymbol{s}_k^i = \boldsymbol{f}_{k-1}\left(\boldsymbol{s}_{k-1}^i, \boldsymbol{v}_{k-1}^i\right). \tag{5.39}$$

In a second step, for each particle the weights have to be calculated. With the chosen proposal density, this step reduces to

$$w_k^i = \mathrm{p}\left(\boldsymbol{y}_k|\boldsymbol{s}_k^i\right). \tag{5.40}$$

Finally, all weights have to be normalized by

$$w_k^i = \frac{w_k^i}{W}, \tag{5.41}$$

using

$$W = \sum_{i=1}^{N_\mathrm{p}} w_k^i. \tag{5.42}$$

A crucial problem of the PF is the degeneracy phenomenon [20, 140]. It points out that after a few iterations, all but one particle will have weights very close to zero. It was shown that the variance of the weights can only increase over time, and hence, it is not possible to avoid the degeneracy problem. One approach to reduce this effect is simply to use a very large number of particles. However, this is often too inefficient from a computational complexity point of view. A much smarter method is the application of resampling where degeneracy can be reduced remarkably. Idea is an elimination of particles with low weights to concentrate on particles having large weights. In this manner, a new set of states $\tilde{\boldsymbol{s}}_k^i, i = 1, 2, \ldots, N_\mathrm{p}$, is created by resampling N_p times from an approximate discrete representation of

$$\mathrm{p}\left(\boldsymbol{s}_k|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) \sim \sum_{i=1}^{N_\mathrm{p}} w_k^i \delta\left(\boldsymbol{s}_k - \boldsymbol{s}_k^i\right). \tag{5.43}$$

Given

$$\mathrm{P}\left(\tilde{\boldsymbol{s}}_k^j = \boldsymbol{s}_k^i\right) = w_k^j, \tag{5.44}$$

the resulting sample is an independent and identically distributed sample from the discrete density. Even though the degeneracy can be reduced by resampling, another effect denoted as sample impoverishment is introduced in practical implementations. Besides the problem of limited parallelization due to the fact that the particles have to be combined, particles with large weights are statistically selected much more often then the other particles. So the diversity among the particles is reduced since the resulting sample will contain many repeated points. Especially for systems with small state noise the sample impoverishment can be a serious problems and all particles can be concentrated to a single state after a few iterations.

In addition to the SIR-PF, there exist several other PF approaches. Briefly mentioned at this point should be the auxiliary sampling importance resampling PF (ASIR-PF) [148, 140]. It can

be seen as an implementation of the SIR-PF with resampling at the previous time-step, i.e., the ASIR-PF generates points from the samples at time-step $k-1$. They are usually much closer to the true state if conditioned on the current observations. Compared to the SIR-PF, the ASIR-PF is not so sensitive against outliers if the state noise is small. Further, the weights are distributed more evenly. However, for large state noise the performance of the ASIR-PF can be worse compared to the SIR-PF.

Another PF implementation is the regularized PF (R-PF) [149, 140]. Compared to the SIR-PF, the R-PF has a different resampling stage. Whereas in the SIR-PF the resampling is done based on a discrete approximation, the R-PF resamples from a continuous approximation of the a posteriori PDF where a Kernel approach is applied. The R-PF outperforms the SIR-PF particularly in situations where sample impoverishment limits the performance. This could be situations with, e.g., low state noise.

5.3.4 Further approaches

Grid based methods

If the state space is discrete, and hence, does only include a finite number of states, grid based approaches can provide the optimum solution [140, 20]. Assuming that the state vector at time-step $k-1$ consists of N_g states $\boldsymbol{s}_{k-1}^i, i = 1, 2, \ldots, N_\mathrm{g}$. Then, the a posteriori PDF can be represented as a weighted sum of the discrete states, i.e.,

$$\mathrm{p}\left(\boldsymbol{s}_{k-1}|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\right) = \sum_{i=1}^{N_\mathrm{g}} w_{k-1|k-1}^i \delta\left(\boldsymbol{s}_{k-1} - \boldsymbol{s}_{k-1}^i\right). \tag{5.45}$$

Using (5.45) with (5.5), the prediction equation can be written as

$$\mathrm{p}\left(\boldsymbol{s}_{k}|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\right) = \sum_{i=1}^{N_\mathrm{g}} w_{k|k-1}^i \delta\left(\boldsymbol{s}_{k} - \boldsymbol{s}_{k}^i\right), \tag{5.46}$$

with the predicted weights

$$w_{k|k-1}^i = \sum_{j=1}^{N_\mathrm{g}} w_{k-1|k-1}^j \,\mathrm{p}\left(\boldsymbol{s}_k^i|\boldsymbol{s}_{k-1}^j\right). \tag{5.47}$$

Substituting (5.45) into (5.6), the update equation is given as

$$\mathrm{p}\left(\boldsymbol{s}_{k}|\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k}\right) = \sum_{i=1}^{N_\mathrm{g}} w_{k|k}^i \delta\left(\boldsymbol{s}_{k} - \boldsymbol{s}_{k}^i\right), \tag{5.48}$$

with the updated weights

$$w_{k|k}^i = \frac{\sum_{j=1}^{N_\mathrm{g}} w_{k|k-1}^j \,\mathrm{p}\left(\boldsymbol{y}_k|\boldsymbol{s}_k^j\right)}{w_{k|k-1}^i \,\mathrm{p}\left(\boldsymbol{y}_k|\boldsymbol{s}_k^i\right)}. \tag{5.49}$$

Second order extended Kalman filter

The extended Kalman filter does only take into account a linearization around the current state estimate. To approximate the non-linearities in a better way, besides the Jacobian matrices also Hessian matrices including the second derivatives can be included. These filters are denoted as second order extended Kalman filters or modified Gaussian second order filters. Derivations in the context of dynamic location estimation applications can be found in [141] and [150]. Even more sophisticated approaches including robust extended Kalman filtering are proposed in [151].

Unscented Kalman filter

The extended Kalman filter has two important drawbacks [152]. On the one hand, the derivation of the Jacobian matrices, i.e., the linear approximations of the non-linear functions, may be complex and can cause implementation difficulties. On the other hand, these linearizations can lead to filter instabilities if the time-step intervals are not sufficiently small, especially in highly non-linear environments. To address these limitations the unscented Kalman filter can be an alternative [153]. The philosophy of the unscented Kalman filter is that it uses the premise, that it is easier to approximate a Gaussian distribution than it is to approximate an arbitrary non-linear function. Hence, instead of linearizing the system by Jacobian matrices the unscented Kalman filter uses a deterministic sampling approach to capture the mean and covariance estimates with a minimum set of sample points (minimum but sufficient). These points are chosen that their mean, covariance, and possibly also higher order moments, match the corresponding Gaussian random variables.

The idea of unscented Kalman filter is based on the unscented transform that was proposed in [154] with the approach, that the non-linear function is applied to a set of points and the statistics of the transformed points can be used to estimate the transformed mean and covariance. Contrary to PFs the chosen points are not drawn randomly. These so-called sigma points are chosen deterministically with respect to certain given mean and covariance properties. These sigma points are then propagated through the non-linearities determined by the state or observation model. Afterwards, the sigma-points are weighted and finally recombined to produce the estimated mean and covariance. Different approaches to obtain a valid set of sigma points are provided in [155].

Gaussian mixture filter

Gaussian mixture filters or Gaussian sum filters approximate a priori and a posteriori PDFs by Gaussian mixtures which is a convex combination of Gaussian PDFs. It can be seen as an extension or generalization of the Kalman filtering approach. Limiting factor of Gaussian mixture filters is the number of included PDFs (similar as the number of particles in PF approaches). A comprehensive overview in the context of dynamic location estimation can be found in [141].

Rao-Blackwellization

With increasing number of states that have to be estimated, also the number of required particles needs to be increased. For instance, a state vector in inertial navigation requires a very high number of particles. The idea of Rao-Blackwellization (cf. [156]) which is sometimes denoted as marginalization is to reduce the number of particles by using a Kalman filter for the part of the system model that is linear. The non-linear part of the system model is still be treated by a PF. Therefore, we split the state vector into a linear and a non-linear part according to

$$\boldsymbol{s}_k = \left[\boldsymbol{s}_{\mathrm{l},k}, \boldsymbol{s}_{\mathrm{nl},k}\right]^{\mathrm{T}}. \tag{5.50}$$

Then, the Rao-Blackwellized PF factorizes the a posteriori PDF into

$$\mathrm{p}\left(\boldsymbol{s}_{\mathrm{l},k}, \boldsymbol{s}_{\mathrm{nl},k} | \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) = \mathrm{p}\left(\boldsymbol{s}_{\mathrm{l}} | \boldsymbol{s}_{\mathrm{nl}}, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right) \mathrm{p}\left(\boldsymbol{s}_{\mathrm{nl}} | \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\right), \tag{5.51}$$

where each of the conditional PDFs is handled by different filters. For the example of linear and non-linear separation this could be a Kalman filter and PF. But this is not a limitation of the Rao-Blackwellized filter since it allows arbitrary combinations of filters (e.g., Kalman filter and unscented Kalman filter as analyzed in [157]). It can be shown that for a Rao-Blackwellized PF with a fixed number of particles the performance is always better than for a standard PF. Also nested combinations handling more than two different sub-states are possible. For instance, in [158] an implementation that combines Kalman filter, PF, and grid based methods is investigated.

Map matching

In the context of particle filtering often map matching is discussed. The idea of map matching is to use map information as additional or side information in the filtering algorithms. For instance, if a street map is available it can be easily included in car tracking algorithms. Since it is usually only allowed to drive on the streets, hypotheses or estimates being off-road can be dropped out. In the same way, floor plans of buildings can be used to improve pedestrian tracking. In this context, it is used that pedestrians can not go through walls which can be considered in the algorithms. One straightforward approach can be implemented in a PF. For instance, particles that proceed through walls can be weighted quite low or even dropped out. Also for extended Kalman filter implementations map matching solutions are proposed [159].

5.4 Simulation results

In the following, we evaluate Kalman filter, extended Kalman filter, and particle filter in the dynamic location estimation context. As scenario we exploit the cellular network situation as depicted in Figure 5.4. Here, the MS is moving along a track with a constant speed of $2\,\mathrm{m/s}$. For the timing information we use the same parameters as already used for static location estimation described in Section 4.5. With the $N_{\mathrm{BS}} = 3$ nearest BSs timing estimation is performed, where the timing estimates are impacted by AWGN with a fixed standard deviation according to $\sigma_{\boldsymbol{n}_\tau} = 50\,\mathrm{m}/c_0$. For the static location estimation we exploit the Levenberg-Marquardt algorithm. The (linear) Kalman filter approach then smoothes the static location estimates by considering the mobility model, where we assume $\sigma_Q^2 = 0.1$ for the covariance matrix of the state model according to (5.21). This parameter controls, how much we trust the mobility model, i.e., it controls how new observations can lead to a deviation from the current track. The extended Kalman filter exploits directly the timing information — without intermediate step of static location estimation. For the particle filter we implemented the SIR-PF with $N_{\mathrm{p}} = 1000$ particles. This is computational much more complex compared to Kalman and extended Kalman filter, however, will give us a performance bound for the considered setup.

Figure 5.5 shows the time-variant RMSE for hyperbolic processing over the time-step k for the simulated track, where it was averaged over several noise realizations. First of all, we observe that there is a high dependency of the static solution on the actual location within the cellular network. For instance, close to BS 1 (around time-steps 200 to 300) there is a high performance loss compared to the other locations and the resulting RMSE is higher than $100\,\mathrm{m}$. When these static location estimates are smoothed by the Kalman filter (KF), the overall performance can

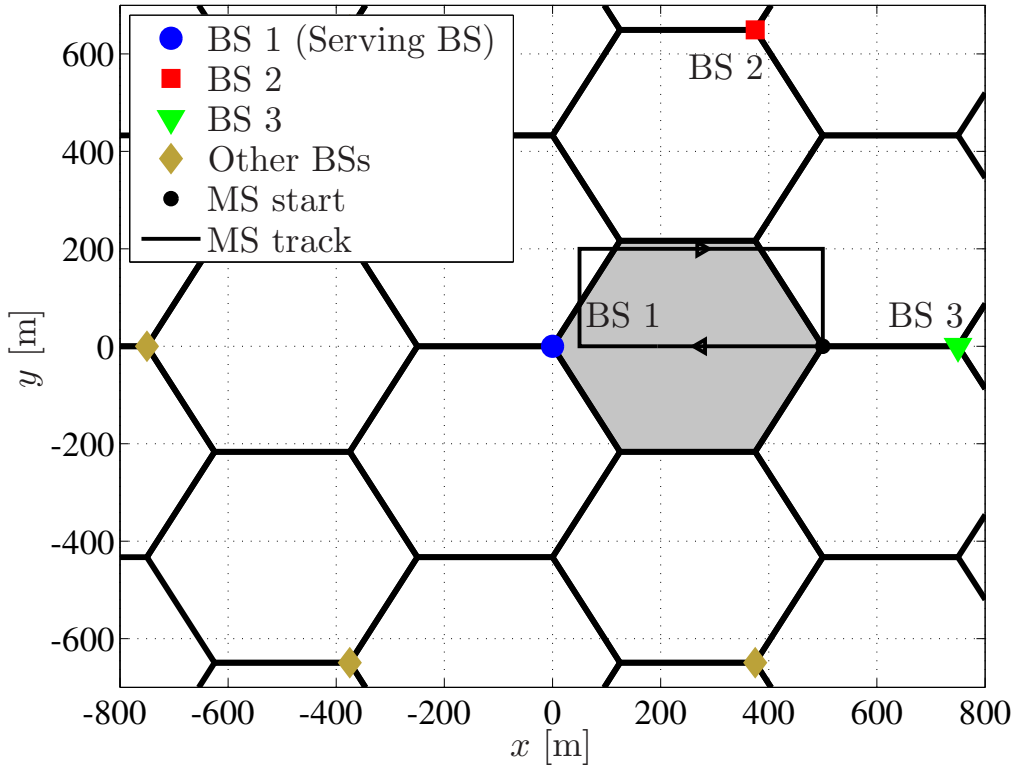


Figure 5.4: Cellular network with MS track

be improved to around 20 m. Nevertheless, also here a slight performance degradation around BS 1 can be observed. The extended Kalman filter (EKF) slightly outperforms the Kalman filter, especially close to BS 1. The particle filter (PF) gives no further performance improvement, hence, for the considered scenario the extended Kalman filter is a reasonably reliable approach.

The resulting CDF of the dynamic location estimation error is depicted in Figure 5.6. For these simulations random tracks with a duration of 100 seconds were generated within the grey depicted area of Figure 5.4. For static location estimation we can achieve a location estimation error which is smaller than 50 m in around 45% of the situations. With a Kalman filter this can be improved to around 92%. For the application of an extended Kalman filter we achieve around 99%, however, also here no further performance improvement can be realized with a particle filter.

Finally, we evaluate the CDF performance when the timing estimation algorithms according to Chapter 3 are applied (cf. Figure 5.7). We focus here on the best-performing iterative timing estimation approach under a combined multipath LOS/NLOS channel model (cf. Appendix A.1). As already could be seen in Figure 4.9, with static location estimation in 80% of the situations an error smaller than 50 m was possible. For comparison also the results for the pilot-aided approach is depicted which solely achieved 40%. With the Kalman filter the overall performance could be improved to around 90% and with the extended Kalman filter — which shows a similar performance compared to the particle filter — to around 97%. Even more impressive is the gained coverage by application of dynamic location estimation. Whereas for the static approach there still was a gap of the CDF to 100%, i.e., there were location estimation errors higher than 300 m in around 10% of the situations, this could be reduced to only a few outliers (below 1%) when applying extended Kalman filter or particle filter. This is a remarkable improvement compared to the state-of-the-art approaches.

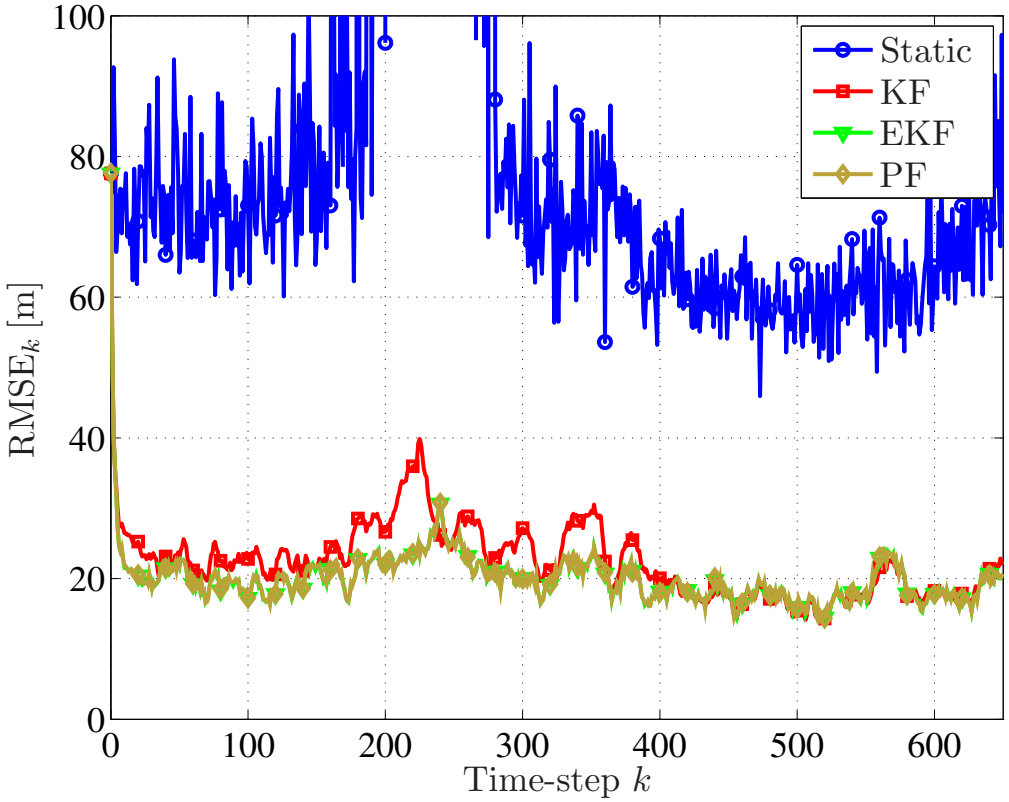


Figure 5.5: RMSE of MS track for dynamic hyperbolic location estimation

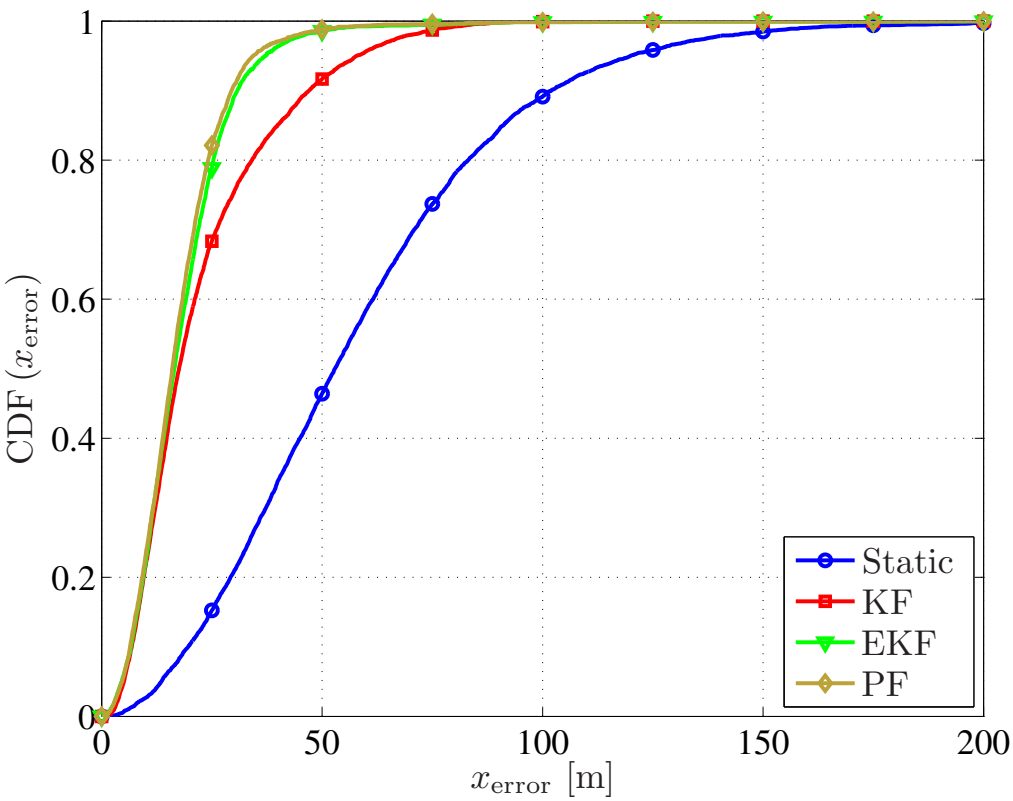


Figure 5.6: CDF of dynamic hyperbolic location estimation error

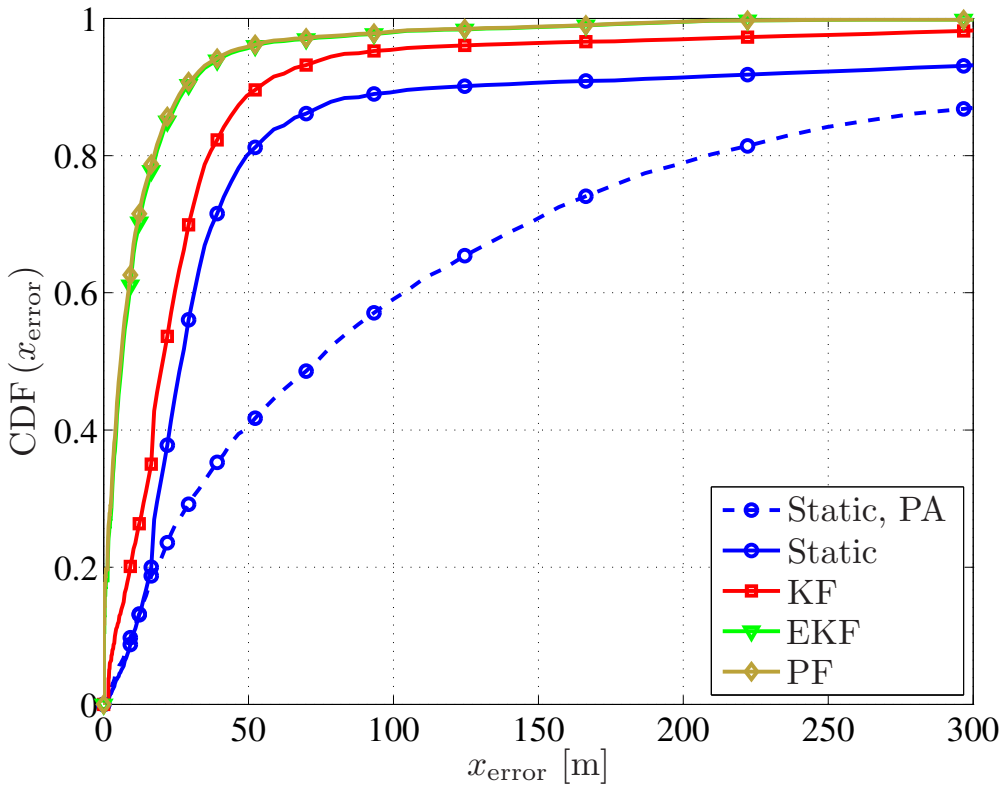


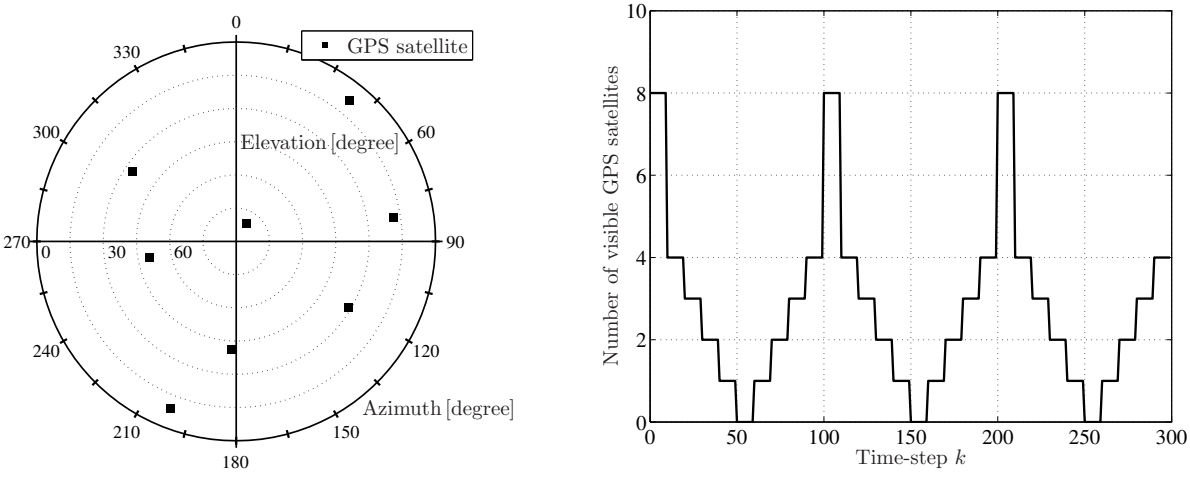
Figure 5.7: CDF of dynamic hyperbolic location estimation error, multipath LOS/NLOS channel

6 Data fusion with GNSS signals

As a final step we will show that the derived timing estimation and location estimation algorithms are suited for data fusion as well. As application scenario, we will assess the fusion process of signals received from GNSS and LTE. For GNSS we are choosing GPS, where a critical location estimation environment is assumed. As mentioned already in Chapter 1.1, LOS access to at least four satellites is required to get reliable location estimates. However, in critical environments like urban canyons or indoor this can not be guaranteed. In these situations we propose the exploitation of LTE to support GPS, i.e., missing GPS satellites should be compensated by timing estimates derived from LTE BSs.

The GPS constellation depends on the location of the MS on earth and on the time. Figure 6.1a shows an example of such a constellation by depicting a so-called skyplot. Here, eight GPS satellites can be observed at different azimuth and elevation angles with respect to the actual MS location. Certainly, the satellites change their location over time, however, for the considered setup we assume them to be constant for the duration of one simulated MS track.

To simulate a critical environment for location estimation, we emulate an urban canyon situation. To do so, we assume that not all satellites are visible during the MS track by applying the satellite visibility as depicted in Figure 6.1b. We observe, that at the begin of the MS track all eight satellites are visible. When the MS enters the urban canyon, more and more satellites are dropped out (e.g., due to blocking of the LOS signal by buildings) until — for a period of 10 seconds no satellite is visible. Then, more and more satellites can be detected again. We further



(a) GPS constellation

(b) GPS visibility

Figure 6.1: GPS constellation and visibility

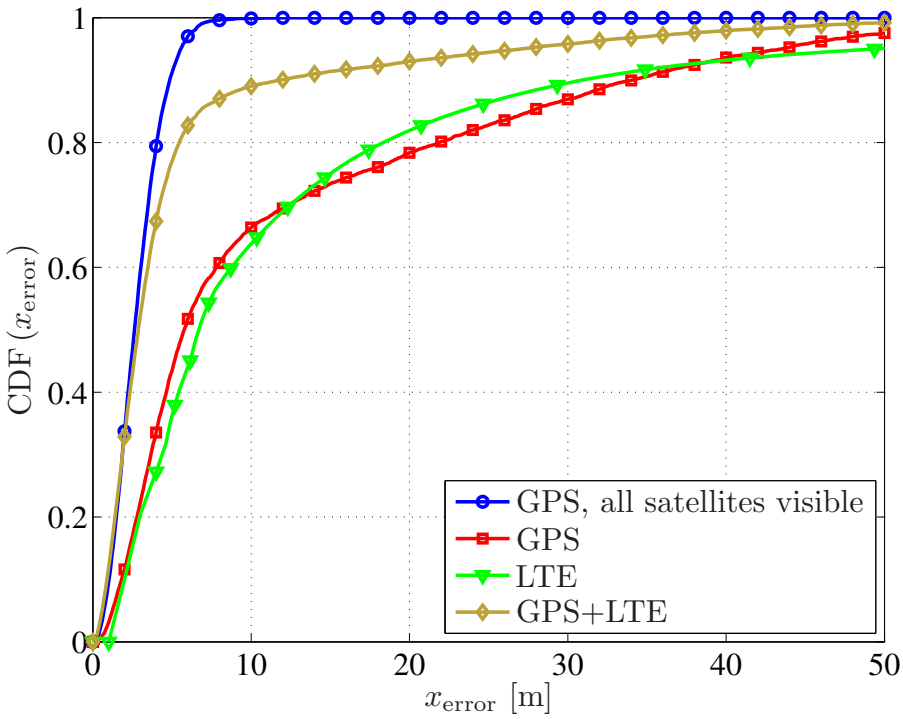


Figure 6.2: CDF of dynamic location estimation error for data fusion of GPS and LTE

assume that only the satellites with the highest elevations can be detected according to Figure 6.1a. This reflects the reality considering a typical urban canyon situation.

The timing estimates of the satellite links are assumed to be impacted by AWGN with a fixed standard deviation according to $\sigma_{\boldsymbol{n}_\tau} = 10\,\text{m}/c_0$. As discussed already in Section 4.1, GNSS, i.e., also GPS, requires a spherical processing of the timing estimates with unknown time offset. However, as we have seen in Section 4.4, an equivalent hyperbolic processing is possible without loss of any generality. This makes the data fusion process between GPS and LTE very simple, since the LTE based timing estimates are also processed in a hyperbolic way. Therefore, there is no difference if the timing estimates are obtained by GPS or LTE, i.e., the respective observations can simply be combined. The overall estimation process (e.g., using a static or dynamic solution) is then the same as for separated processing.

Figure 6.2 shows finally the CDF results for dynamic location estimation using an extended Kalman filter for data fusion of GPS and LTE under the assumptions as outlined before. For the LTE system we apply the best-performing iterative approach and assume a combined multipath LOS/NLOS channel model according to Appendix A.1. For stand-alone LTE, we observe that the depicted location estimation CDF results in an estimation error smaller than $10\,\text{m}$ in around 63% of the situations. For stand-alone GPS, we apply the urban canyon situation as described before and additionally average over several locations on earth for obtaining various satellite constellations. The resulting CDF is at around 65% for location estimation errors smaller than $10\,\text{m}$, i.e., only slightly better compared to stand-alone LTE. If we combine both approaches we can achieve around 90%. For comparison also the optimum GPS performance is depicted, where it is assumed that access to all satellites is possible. Hence, we observe that — at least partly — missing satellites can be compensated by timing estimates coming from an LTE system and improve the overall performance. The extension to more fusion sources is straightforward, e.g., in [17] additionally the European Galileo system is fused with GPS and LTE.

7 Conclusions

In this thesis, location determination techniques were proposed and assessed. A focus has been set to OFDM based mobile radio systems and to processing of timing information which is the most accurate location dependent parameter in the context of location determination approaches. In Chapter 3, timing estimation algorithms for OFDM were investigated and assessed. By an analysis of the system parameters it could be shown, that OFDM is in general suitable to provide reliable timing estimates. For instance, for the considered LTE system at a bandwidth of $20\,\mathrm{MHz}$ the achievable accuracy is in the centimeter-region if one OFDM symbol is exploited for timing estimation. Pilot-aided synchronization algorithms were investigated and assessed with respect to their timing estimation capabilities and the corresponding CRLB was derived. This absolute performance bound is difficult to achieve in practical systems. Besides the well-known performance limiting factors like multipath and NLOS propagation (which are not explicitly studied in this thesis), in a cellular LTE system other effects limit the performance. As simulations have shown, one limiting factor is the reduced number of pilots and synchronization symbols in an OFDM frame that can be exploited for timing estimation. From a communications point of view the number of pilots should be as small as possible for an increased spectral efficiency, however, for timing estimation, and hence, location estimation, as many pilots as possible should be available in an OFDM frame. Therefore, a data-aided approach was proposed that exploit already decided data as pilot symbols for an improved timing estimation. With this technique especially the performance of the serving BS could be improved. A second limiting factor is inter-cell interference, that limits the performance close to BSs. To overcome this problem, an interference cancelation procedure was proposed to reduce the interference of the strongest (serving) cell in the network for a better reception of the out-of-cell BSs. With that idea, the performance can be improved remarkably. Both approaches can be combined and extended by applying them in an iterative way. With this procedure a performance improvement of the timing estimation RMSE of nearly one magnitude could be realized.

The timing estimates were then processed by a static location estimation technique that was discussed in Chapter 4. Several approaches for processing these estimates were considered, including general numerical approaches as well as procedures that take into account the geometric properties of the location estimation problem. It turned out, that the Levenberg-Marquardt algorithm provides the most reliable and computational efficient estimates in the considered context and was chosen as the baseline approach during this thesis. The assessment of the different algorithms was performed by using a simplified noise model with a fixed standard deviation. Then, the performance of location determination using the timing estimates obtained with the algorithms in Chapter 3 was analyzed. With the proposed best-performing iterative timing estimation approach, e.g., a location

estimation error smaller than $50\,\mathrm{m}$ could be realized in 80% of the situations, whereas with the state-of-the-art pilot-aided approach solely 40% could be achieved.

To further include the mobility of the MS, dynamic location estimation schemes were analyzed in Chapter 5. The focus of these investigations was on Bayesian approaches, where especially the Kalman filter, the extended Kalman filter and the particle filter were part of the simulations. It turned out, that — although not the optimum approach for the considered scenario — the extended Kalman filter provided a robust, reliable, and computational efficient solution outperforming the linear Kalman filter and close to the near-optimum Particle filter. With the extended Kalman filter, the performance could be improved in a way, that in 97% of the situations a location estimation error smaller than $50\,\mathrm{m}$ was achieved.

Finally, in Chapter 6 it was shown that the general framework presented in this thesis allows a simple extension of the proposed algorithms to fuse timing estimates from a mobile radio systems also with observations from other sources. As example the fusion of LTE and GPS was evaluated. It was shown that timing estimates from a mobile radio system can compensate missing satellites in critical location estimation situations like urban canyons.

A Appendix

A.1 Channel models

The channel models are represented in terms of their small scale parameters defined by the average power delay profile including the path index l, the average path power $\mathrm{E}\left\{|h_l|^2\right\}$, and the fading type of the path defined as constant (C) or Jakes (J). The number of sub-paths for each fading path is 20. The large scale parameters include models for path loss and shadow fading. The multipath LOS and NLOS channel models are adapted from the WINNER C2 channel models for urban scenarios [77] and are computed assuming a height of the BS of $25\,\mathrm{m}$ and a height of the MS of $1.5\,\mathrm{m}$ as well as the parameters included in Tables 3.1 and 3.2. Tables A.1.1 to A.1.6 describe the small and large scale parameters for the AWGN, LOS, and NLOS channel models as used in this thesis, Figure A.1.1 shows the corresponding visualization for LOS and NLOS. In case of a mixed LOS/NLOS channel model, the LOS probability [77] depends on the distance between MS and BS and is computed according to

$$\mathrm{P}_{\mathrm{LOS},\mu} = \min\left(18, d_\mu\right)\left(1 - \mathrm{e}^{d_\mu/63}\right) + \mathrm{e}^{d_\mu/63}, \quad \mu = 1, 2, \ldots, N_{\mathrm{BS}}. \tag{A.1.1}$$

AWGN

l	0
$\mathrm{E}\left\{\lvert h_l\rvert^2\right\}$	1
Path type	C

Table A.1.1: Power delay profile for AWGN channel model

Path loss model	$20\log_{10}(d_\mu/1\,\mathrm{m})\,\mathrm{dB} + 38.44\,\mathrm{dB}$
Shadow fading model	Not applicable

Table A.1.2: Large scale parameters for AWGN channel model

LOS

l	0	0	1	3	4	5	7
$\mathrm{E}\left\{\lvert h_l\rvert^2\right\}$	0.846	0.041	0.025	0.018	0.026	0.038	0.004
Path type	C	J	J	J	J	J	J

Table A.1.3: Power delay profile for LOS channel model

Path loss model	$26\log_{10}(d_\mu/1\,\mathrm{m})\,\mathrm{dB} + 31.04\,\mathrm{dB}$
Shadow fading model	Log normal fading, $4\,\mathrm{dB}$ standard deviation

Table A.1.4: Large scale parameters for LOS channel model

NLOS

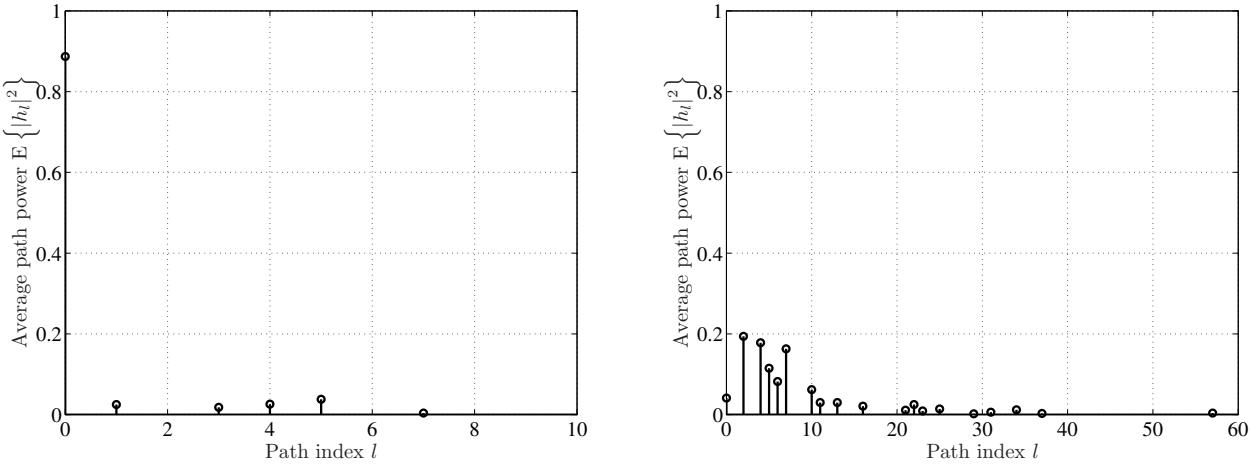
l	0	2	4	5	6	7	10	11	13
$\mathrm{E}\left\{\lvert h_l\rvert^2\right\}$	0.041	0.194	0.178	0.115	0.082	0.163	0.062	0.030	0.030
Path type	J	J	J	J	J	J	J	J	J

16	21	22	23	25	29	31	34	37	57
0.021	0.011	0.025	0.009	0.014	0.002	0.006	0.012	0.003	0.004
J	J	J	J	J	J	J	J	J	J

Table A.1.5: Power delay profile for NLOS channel model

Path loss model	$35.74\log_{10}(d_\mu/1\,\mathrm{m})\,\mathrm{dB} + 33.46\,\mathrm{dB}$
Shadow fading model	Log normal fading, $8\,\mathrm{dB}$ standard deviation

Table A.1.6: Large scale parameters for NLOS channel model



(a) LOS (b) NLOS

Figure A.1.1: Average power delay profiles for LOS and NLOS channel models

A.2 Jacobian matrices for static location estimation

This section includes the Jacobian matrices for static location estimation using spherical processing, spherical processing with unknown time offset, and hyperbolic processing.

SPH

$$\boldsymbol{\Phi}^{\text{SPH}}(\boldsymbol{p}) = \boldsymbol{\Phi}^{\text{SPH}}(\boldsymbol{x}) = \begin{bmatrix} \frac{x-x_1}{d_1} & \frac{y-y_1}{d_1} & \frac{z-z_1}{d_1} \\ \frac{x-x_2}{d_2} & \frac{y-y_2}{d_2} & \frac{z-z_2}{d_2} \\ \vdots & \vdots & \vdots \\ \frac{x-x_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} & \frac{y-y_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} & \frac{z-z_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} \end{bmatrix} \in \mathbb{R}^{(N_{\text{BS}}\times 3)} \tag{A.2.1}$$

SPHT

$$\begin{aligned} \boldsymbol{\Phi}^{\text{SPHT}}(\boldsymbol{p}) &= \boldsymbol{\Phi}^{\text{SPHT}}\left(\left[\boldsymbol{x}^{\text{T}}, b_{\text{clock}}\right]^{\text{T}}\right) \\ &= \begin{bmatrix} \frac{x-x_1}{d_1} & \frac{y-y_1}{d_1} & \frac{z-z_1}{d_1} & 1 \\ \frac{x-x_2}{d_2} & \frac{y-y_2}{d_2} & \frac{z-z_2}{d_2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x-x_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} & \frac{y-y_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} & \frac{z-z_{N_{\text{BS}}}}{d_{N_{\text{BS}}}} & 1 \end{bmatrix} \in \mathbb{R}^{(N_{\text{BS}}\times 4)} \end{aligned} \tag{A.2.2}$$

HYP

$$\boldsymbol{\Phi}^{\text{HYP}}(\boldsymbol{p}) = \boldsymbol{\Phi}^{\text{HYP}}(\boldsymbol{x}) = \boldsymbol{D}\boldsymbol{\Phi}^{\text{SPH}}(\boldsymbol{x}) \in \mathbb{R}^{((N_{\text{BS}}-1)\times 3)} \tag{A.2.3}$$

A.3 Jacobian matrices for dynamic location estimation

This section includes the Jacobian matrices for dynamic location estimation using spherical processing, spherical processing with unknown time offset, and hyperbolic processing.

SPH

$$\boldsymbol{G}_k^{\text{SPH}}(\boldsymbol{s}_k) = \begin{bmatrix} \frac{x_k-x_{1,k}}{d_{1,k}} & \frac{y_k-y_{1,k}}{d_{1,k}} & \frac{z_k-z_{1,k}}{d_{1,k}} & 0 & 0 & 0 \\ \frac{x_k-x_{2,k}}{d_{2,k}} & \frac{y_k-y_{2,k}}{d_{2,k}} & \frac{z_k-z_{2,k}}{d_{2,k}} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{x_k-x_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & \frac{y_k-y_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & \frac{z_k-z_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(N_{\text{BS}}\times 6)} \tag{A.3.1}$$

SPHT

$$\boldsymbol{G}_k^{\text{SPHT}}\left(\boldsymbol{s}_k\right) = \begin{bmatrix} \frac{x_k - x_{1,k}}{d_{1,k}} & \frac{y_k - y_{1,k}}{d_{1,k}} & \frac{z_k - z_{1,k}}{d_{1,k}} & 0 & 0 & 0 & 1 \\ \frac{x_k - x_{2,k}}{d_{2,k}} & \frac{y_k - y_{2,k}}{d_{2,k}} & \frac{z_k - z_{2,k}}{d_{2,k}} & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \frac{x_k - x_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & \frac{y_k - y_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & \frac{z_k - z_{N_{\text{BS}},k}}{d_{N_{\text{BS}},k}} & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{(N_{\text{BS}} \times 7)} \tag{A.3.2}$$

HYP

$$\boldsymbol{G}_k^{\text{HYP}}\left(\boldsymbol{s}_k\right) = \boldsymbol{D}\boldsymbol{G}_k^{\text{SPH}}\left(\boldsymbol{s}_k\right) \in \mathbb{R}^{((N_{\text{BS}}-1) \times 6)} \tag{A.3.3}$$

List of acronyms

2G	2nd generation
3G	3rd generation
3GPP	3rd Generation Partnership Project
4G	4th generation
AGNSS	Assisted global navigation satellite system
AOA	Angle of arrival
ASIR	Auxiliary sampling importance resampling
AWGN	Additive white Gaussian noise
BER	Bit error rate
BICM	Bit-interleaved coded modulation
BS	Base station
CC	Cross-correlation
CDF	Cumulative distribution function
CFO	Carrier frequency offset
CH	Chan-Ho
CP	Cyclic prefix
CRLB	Cramer-Rao lower bound
CSI	Channel state information
DA	Data-aided
DAC	Digital-to-analog converter
DC	Differential correlation
DVB	Digital Video Broadcasting
DAB	Digital Audio Broadcasting
E911	Enhanced 911
ED	Estimated data
EKF	Extended Kalman filter
EM	Expectation maximization
FB	Fallback

FCC	Federal Communications Commission
FD	Frequency domain
FE	Frontend
FFT	Fast Fourier transform
GN	Gauss-Newton
GNSS	Global navigation satellite system
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HYP	Hyperbolic
IC	Interference cancelation
IDFT	Inverse discrete Fourier transform
IFFT	Inverse fast Fourier transform
IT	Iterative
KF	Kalman filter
LM	Levenberg-Marquardt
LMU	Location measurement unit
LOS	Line-of-sight
LTE	Long Term Evolution
MAP	Maximum a posteriori
ML	Maximum likelihood
MMSE	Minimum mean squared error
MS	Mobile station
MSE	Mean squared error
MUX	Multiplexer
NE	Newton
NLOS	Non-line-of-sight
OFDM	Orthogonal frequency division multiplexing
PA	Pilot-aided
PDF	Probability density function
PF	Particle filter
P/S	Parallel-to-serial
PSS	Primary synchronization signal
QPSK	Quadrature phase shift keying
R	Regularized
RAIM	Receiver autonomous integrity monitoring
RDC	Reverse differential correlation
RF	Radio frequency
RFID	Radio frequency identification

RMSE	Root mean squared error
RS	Reference signal
RSS	Received signal strength
RTTOA	Round-trip time of arrival
RX	Receiver
SD	Steepest descent
SIR	Sampling importance resampling
SINR	Signal-to-interference-and-noise ratio
SNR	Signal-to-noise ratio
S/P	Serial-to-parallel
SPH	Spherical
SPHT	Spherical with unknown time offset
SSS	Secondary synchronization signal
TD	Time domain
TDOA	Time difference of arrival
TOA	Time of arrival
TX	Transmitter
UMTS	Universal Mobile Telecommunications System
UWB	Ultra-wideband
WINNER	Wireless World Initiative New Radio
WLAN	Wireless local area network

Bibliography

[1] G. Sun, J. Chen, W. Guo, and K. J. Ray Liu, "Signal processing techniques in network-aided positioning," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 12–23, July 2005.

[2] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 24–40, July 2005.

[3] J. J. Nielsen, "Location based network optimizations for mobile wireless networks," Ph.D. dissertation, Aalborg University, Aalborg, Denmark, 2011.

[4] Federal Communications Commission (FCC), "FCC 99-245: Third report and order," *http://www.fcc.gov/911/enhanced/*, October 1999.

[5] W. Michalski, "Technical and regulatory issues of emergency call handling," *Journal of Telecommunications and Information Technology*, no. 3, pp. 123–129, 2009.

[6] B. W. Parkinson and J. J. Spilker Jr., *Global Positioning System: Theory and Applications, Volume 1*. Progress in Astronautics and Aeronautics, Volume 163, 1996.

[7] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements and Performance*. Ganga-Jamuna Press, 2004.

[8] R. Ercek, P. D. Doncker, and F. Grenez, "Study of pseudo-range error due to non-line-of-sight-multipath in urban canyons," *Proceedings of the ION GNSS*, September 2005.

[9] Y. Zhao, "Mobile phone location determination and its impact on intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 1, pp. 55–67, March 2000.

[10] S. Gezici, "A survey on wireless position estimation," *Journal on Wireless Personal Communications*, vol. 44, pp. 263–282, 2008.

[11] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, July 2005.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[13] Y. Zhao, "Standardization of mobile phone positioning for 3G systems," *IEEE Communications Magazine*, vol. 40, no. 7, pp. 108–116, July 2002.

[14] Third Generation Partnership Project (3GPP), *http://www.3gpp.org*, March 2011.

[15] Wireless World Initiative New Radio (WINNER) Project, *https://www.ist-winner.org*, March 2011.

[16] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House, 2000.

[17] C. Mensing, S. Sand, and A. Dammann, "Hybrid data fusion and tracking for positioning with GNSS and 3GPP-LTE," *International Journal of Navigation and Observation, Special Issue on Integrating Radio Positioning and Communications: New Synergies*, August 2010.

[18] J. J. Caffery, *Wireless Location in CDMA Cellular Radio Systems*. Kluwer Academic Publishers, 2000.

[19] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, no. D, pp. 35–45, 1960.

[20] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter - Particle Filters for Tracking Applications*. Artech House, 2004.

[21] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, July 2005.

[22] K. Wendlandt, P. Robertson, M. Khider, M. Angermann, and K. Sukchaya, "Demonstration of a realtime active-tag RFID, Java based indoor localization system using particle filtering," *Proceedings of the International Conference on Ubiquitous Computing (UbiComp)*, September 2000.

[23] M. Grewal, L. Weill, and A. Andrews, *Global Positioning Systems, Inertial Navigation, and Integration*. John Wiley & Sons, 2001.

[24] Y. Hao, Z. Zhang, and Q. Xia, "Research on data fusion for SINS/GPS/magnetometer integrated navigation based on modified CDKF," *Proceedings of the IEEE International Conference on Progress in Informatics and Computing*, December 2010.

[25] H. Leppäkoski, S. Tikkinen, and J. Takala, "Optimizing radio map for WLAN fingerprinting," *Proceedings of the International Conference on Ubiquitous Positioning, Indoor Navigation and Location-Based Service (UPINLBS)*, October 2010.

[26] K. Frank, B. Krach, N. Catterall, and P. Robertson, "Development and evaluation of a combined WLAN & inertial indoor pedestrian positioning system," *Proceedings of the ION GNSS*, September 2009.

[27] A. Dammann, C. Mensing, and S. Sand, "Using CoMP for terminal localization," in *Coordinated Multi-Point in Mobile Communications*. P. Marsch and G. P. Fettweis (Editors), Cambridge University Press, 2011.

[28] A. Dammann, C. Mensing, and S. Sand, "On the benefit of location and channel state information for synchronization in 3GPP-LTE," April 2010.

[29] L. Brunel, M. Plainchault, N. Gresset, A. Dammann, C. Mensing, and R. Raulefs, "Inter-cell interference coordination and synchronization based on location information," *Proceedings of the Workshop on Positioning, Navigation, and Communication (WPNC)*, March 2010.

[30] C. Mensing, S. Sand, A. Dammann, and W. Utschick, "Data-aided location estimation in cellular OFDM communications systems," *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, November 2009.

[31] C. Mensing, S. Sand, A. Dammann, and W. Utschick, "Interference-aware location estimation in cellular OFDM communications systems," *Proceedings of the IEEE International Conference on Communications (ICC)*, June 2009.

[32] C. Mensing, S. Sand, and A. Dammann, "Cellular systems for positioning: Capabilities and challenges," *Proceedings of the Wireless World Research Forum (WWRF) Meeting*, April 2008.

[33] C. Mensing, E. Tragos, J. Luo, and E. Mino, "Location determination using in-band signaling for mobility management in future networks," *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2007.

[34] C. Mensing, S. Plass, and A. Dammann, "Positioning with generalized multi-carrier communications signals," *Proceedings of the International Workshop on Multi-Carrier Spread-Spectrum (MCSS)*, May 2007.

[35] C. Mensing, S. Plass, and A. Dammann, "Synchronization algorithms for positioning with OFDM communications signals," *Proceedings of the Workshop on Positioning, Navigation, and Communication (WPNC)*, March 2007.

[36] C. Mensing and S. Plass, "Positioning based on factor graphs," *EURASIP Journal on Advances in Signal Processing*, June 2007.

[37] C. Mensing and S. Plass, "TDOA positioning based on factor graphs," *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2006.

[38] C. Mensing and S. Plass, "Location determination with factor graphs for TDOA," *Proceedings of the NEWCOM-ACoRN Joint Workshop (NAW)*, September 2006.

[39] C. Mensing and S. Plass, "Positioning algorithms for cellular networks using TDOA," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2006.

[40] C. Mensing, S. Sand, and A. Dammann, "GNSS positioning in critical scenarios: Hybrid data fusion with communications signals," *Proceedings of the International Workshop on Synergies in Communications and Localization (SyCoLo)*, June 2009.

[41] C. Mensing and S. Sand, "Performance enhancement of GNSS positioning in critical scenarios by wireless communications systems," *Proceedings of the IEEE/ION Position Location and Navigation Symposium (PLANS)*, May 2008.

[42] C. Mensing and A. Dammann, "Positioning with OFDM based communications systems and GNSS in critical scenarios," *Proceedings of the Workshop on Positioning, Navigation, and Communication (WPNC)*, March 2008.

[43] I. K. Adusei, K. Kyamakya, and K. Jobmann, "Mobile positioning technologies in cellular networks: An evaluation of their performance metrics," *Proceedings of the IEEE International Conference for Military Communications*, October 2002.

[44] 3GPP TS 25.111, "Location measurement unit (LMU) performance specification; user equipment (UE) positioning in UTRAN, V10.0.0," *http://www.3gpp.org*, April 2011.

[45] Y. Shimizu and Y. Sanada, "Accuracy of relative distance measurement with ultra wideband system (uwbst)," *Proceedings of the IEEE Conference on Ultra Wideband Systems and Technologies*, November 2003.

[46] Z. Sahinoglu and S. Gezici, "Ranging in the ieee 802.15.4a standard," *Proceedings of the IEEE Wireless and Microwave Technology Conference (WAMICON)*, December 2006.

[47] B. Denby, Y. Oussar, I. Ahriz, and G. Dreyfus, "High-performance indoor localization with full-band GSM fingerprints," *Proceedings of the International Workshop on Synergies in Communications and Localization (SyCoLo)*, June 2009.

[48] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[49] S. Sand, "Joint iterative channel and data estimation in high mobility mimo-ofdm systems," Ph.D. dissertation, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 2009.

[50] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communications Receivers: Synchronization, Channel Estimation, and Signal Processing*. Wiley Series in Telecommunications and Signal Processing, 1998.

[51] G. L. Stüber, J. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 271–294, April 2004.

[52] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial overview," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, July 2007.

[53] S. Müller-Weinfurtner, "OFDM for wireless communications: Nyquist windowing, peak-power reduction, and synchronization," Ph.D. dissertation, University of Erlangen-Nürnberg, Erlangen, Germany, 2000.

[54] Y. Mostofi and D. C. Cox, "Mathematical anaylsis of the impact of timing synchronization errors on the performance of an OFDM system," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 226–230, February 2006.

[55] P. Dharmawansa, N. Rajatheva, and H. Minn, "An exact error probability analysis of OFDM systems with frequency offset," *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 26–31, January 2009.

[56] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, December 1997.

[57] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "On the optimality of metrics for coarse frame synchronization in OFDM: A comparison," *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 1998.

[58] 3GPP TS 36.211, "Evolved universal terrestrial radio access (E-UTRA): Physical channels and modulation, v10.3.0," *http://www.3gpp.org*, September 2011.

[59] H. Minn, M. Zeng, and V. K. Bhargava, "On timing offset estimation for OFDM systems," *IEEE Communications Letters*, vol. 4, no. 7, pp. 242–244, July 2000.

[60] H. Minn, V. K. Bhargava, and K. B. Letaief, "A robust timing and frequency synchronization for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 822–839, July 2003.

[61] H. Minn, V. K. Bhargava, and K. B. Letaief, "A combined timing and frequency synchronization and channel estimation for OFDM," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 416–422, March 2006.

[62] Z. Zhang, K. Long, M. Zhao, and Y. Liu, "Joint frame synchronization and frequency offset estimation in ofdm systems," *IEEE Transactions on Broadcasting*, vol. 51, no. 3, pp. 389–394, September 2005.

[63] Y.-H. Cho and D.-J. Park, "A new preamble design for synchronization and cell searching algorithms in OFDM cellular systems," *Proceedings of the IEEE Vehicular Technology Conference (VTC) Spring*, May 2008.

[64] F. Berggren and B. M. Popović, "A non-hierarchical cell search scheme," *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, March 2007.

[65] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation*. John Wiley & Sons, 2001.

[66] L. Dai, Z. Wang, J. Wang, and Z. Yang, "Positioning with OFDM signals for the next-generation GNSS," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 374–379, May 2010.

[67] M. Morelli, L. Sanguinetti, and H. V. Poor, "A robust ranging scheme for OFDMA-based networks," *IEEE Transactions on Communications*, vol. 57, no. 8, pp. 2441–2452, August 2009.

[68] Y. Tsai, G. Zhang, D. Grieca, F. Ozluturk, and X. Wang, "Cell search in 3GPP long term evolution systems," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 23–29, June 2007.

[69] I. Kim, Y. Han, Y. H. Kim, and S. C. Bang, "Sequence hopping cell search scheme for OFDM cellular systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1483–1489, May 2008.

[70] Y. Wang, G. Leus, and A.-J. van der Veen, "Cramer-Rao bound for range estimation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, April 2009.

[71] D. Wang and M. Fattouche, "OFDM transmission for time-based range estimation," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 571–574, June 2010.

[72] M. Luise and F. Zanier, "Multicarrier signals: A natural enabler for cognitive positioning systems," *Proceedings of the International Workshop on Multi-Carrier Systems & Solutions (MCSS)*, May 2009.

[73] D. Dardari, Y. Karisan, S. Gezici, A. A. D'Amico, and U. Mengali, "Performance limits on ranging with cognitive radio," *Proceedings of the IEEE International Conference on Communications (ICC) Workshops*, June 2009.

[74] F. Athley, "Threshold region performance of maximum likelihood direction of arrival estimators," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1359–1373, April 2005.

[75] D. Dardari, C.-C. Chong, and M. Z. Win, "Improved lower bounds on time-of-arrival estimation error in realistic UWB channels," *Proceedings of the IEEE International Conference on Ultra-Wideband*, September 2006.

[76] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 338–342, March 1988.

[77] IST-2003-507581: WINNER II Deliverable D1.1.2, "WINNER II channel models," *https://www.ist-winner.org*, September 2007.

[78] J.-J. van de Beek, M. Sandell, and P. O. Börjesson, "Ml estimation of time and frequency offset in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1800–1815, July 1997.

[79] E. del Castillo-Sanchez, F. J. Lopez-Martinez, E. Martos-Naya, and J. T. Entrambasaguas, "Joint time, frequency and sampling clock synchronization for OFDM-based systems," *Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC)*, March 2009.

[80] E. Staudinger and C. Gentner, "TDOA subsample delay estimator with multiple access interference mitigation and carrier frequency offset compensation for OFDM based systems," *Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC)*, March 2011.

[81] H. Bölcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Transactions on Communications*, vol. 49, no. 6, pp. 988–999, June 2001.

[82] A. J. Al-Dweik, "A novel non-data-aided symbol timing recovery technique for OFDM systems," *IEEE Transactions on Communications*, vol. 54, no. 1, pp. 37–40, January 2006.

[83] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, and L. Vandendorpe, "Code-aided turbo synchronization," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1255–1271, June 2007.

[84] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP Journal on Wireless Communications and Networking*, April 2005.

[85] C. Herzet, V. Ramon, and L. Vandendorpe, "A theoretical framework for iterative synchronization based on the sum-product and the expectation-maximization algorithms," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1644–1658, May 2007.

[86] X. Wu and H. Xiang, "Iterative carrier phase recovery methods in turbo receivers," *IEEE Communications Letters*, vol. 9, no. 8, pp. 735–737, August 2005.

[87] S. Godtmann, N. Hadaschik, W. Steinert, and G. Ascheid, "A concept for data-aided carrier frequency estimation at low signal-to-noise ratios," May 2008.

[88] M. Marey, M. Guencah, and H. Steendam, "Iterative residual frequency offset correction for OFDM systems," *Proceedings of the IEEE Vehicular Technology Conference (VTC) Spring*, May 2008.

[89] K. Shi, E. Serpedin, and P. Ciblat, "Decision-directed fine synchronization in OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 408–412, March 2005.

[90] Z. Xie and B. Walke, "Resource allocation and reuse for inter-cell interference mitigation in OFDMA based communication networks," *International Wireless Internet Conference (WICON)*, March 2010.

[91] S. Plass, "Cellular MC-CDMA downlink systems - coordination, cancellation, and use of inter-cell interference," Ph.D. dissertation, University of Ulm, August 2008.

[92] S. Plass and S. Kaiser, "MC-CDMA versus OFDMA in cellular environments," *Proceedings of the European Signal Processing Conference (EUSIPCO)*, September 2005.

[93] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 56–61, August 2006.

[94] J. A. del Peral-Rosado, J. A. Lopez-Salcedo, G. Seco-Granados, F. Zanier, and M. Crisci, "Achievable localization accuracy of the positioning reference signal of 3GPP LTE," *Proceedings of the International Conference on Localization and GNSS (ICL-GNSS)*, June 2012.

[95] S. Kim, Y. Jeong, and C. Lee, "Interference-cancellation-based IPDL method for position location in WCDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 1, pp. 117–126, January 2005.

[96] A. Urruela, J. Sala, and J. Rima, "Average performance analysis of circular and hyperbolic geolocation," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 52–66, January 2006.

[97] A. Urruela, "Signal processing techniques for wireless locationing," Ph.D. dissertation, Technical University of Catalonia, Barcelona, Spain, April 2006.

[98] R. Fletcher, *Practical Methods of Optimization*. Wiley, 1991.

[99] J. J. Caffery and G. L. Stüber, "Subscriber location in cdma cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 2, pp. 406–416, May 1998.

[100] W. Foy, "Position-location solutions by taylor-series estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 12, pp. 187–193, March 1976.

[101] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[102] D. Marquardt, "An algorithm for least squares estimation on nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, pp. 431–441, 1963.

[103] P. E. Frandsen, K. Jonasson, H. B. Nielsen, and O. Tingleff, "Unconstrained optimization," *Lecture notes, Informatics and Mathematical Modellingt, Technical University of Denmark*, March 2004.

[104] N. Sirola, "Closed-form algorithms in mobile positioning: myths and misconceptions," *Proceedings of the Workshop on Positioning, Navigation, and Communication (WPNC)*, March 2010.

[105] Y. T. Chan and K. C. Ho, "An efficient closed-form localization solution from time difference of arrival measurements," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, April 1994.

[106] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, August 1994.

[107] H. C. Schau and A. Z. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 8, pp. 1223–1225, August 1987.

[108] J. O. Smith and J. S. Abel, "The spherical interpolation method of source localization," *IEEE Journal of Oceanic Engineering*, vol. 12, pp. 246–252, January 1987.

[109] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 12, pp. 1661–1669, December 1987.

[110] B. Friedlander, "A passive location algorithm and its accuracy analysis," *IEEE Journal of Oceanic Engineering*, vol. 12, pp. 234–244, January 1987.

[111] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748–753, September 1990.

[112] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, "Real-time passive source localization: A practical linear-correction least-squares approach," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 8, pp. 943–956, November 2001.

[113] P. Stoica and J. Li, "Source localization from range-difference measurements," *IEEE Signal Processing Magazine*, vol. 23, no. 11, pp. 63–65, November 2006.

[114] Y. Zhou and L. Lamont, "Constrained linear least squares approach for TDOA localization: A global optimum solution," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2008.

[115] D. Li and H. Hu, "Least square solutions of energy based acoustic source localization problems," *Proceedings of the International Conference on Parallel Processing (ICPP) Workshops*, August 2004.

[116] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1121–1130, April 2004.

[117] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1770–1778, May 2008.

[118] E. G. Larsson and D. Danev, "Accuracy comparison of LS and squared-range LS for source localization," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 916–923, February 2010.

[119] J.-C. Chen, C.-S. Maa, Y.-C. Wang, and J.-T. Chen, "Mobile position location using factor graphs," *IEEE Communications Letters*, vol. 7, no. 9, pp. 431–433, September 2003.

[120] J.-C. Chen, P. Ting, C.-S. Maa, and J.-T. Chen, "Wireless geolocation with TOA/AOA measurements using factor graphs and sum-product algorithm," *Proceedings of the IEEE Vehicular Technology Conference (VTC) Fall*, vol. 5, pp. 3526–3529, September 2004.

[121] J.-C. Chen, Y.-C. Wang, C.-S. Maa, and J.-T. Chen, "Network-side mobile position location using factor graphs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2696–2704, October 2006.

[122] L. Cong and W. Zhuang, "Nonline-of-sight error mitigation in mobile location," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 560–573, March 2005.

[123] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 672–681, March 2006.

[124] I. Guvenc, S. Gezici, and Z. Sahinoglu, "Fundamental limits and improved algorithms for linear least-squares wireless position estimation," *Wireless Communications and Mobile Computing*, 2010.

[125] P.-C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, September 1999.

[126] J. Riba and A. Urruela, "A non-line-of-sight mitigation technique based on ML-detection," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2004.

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

[128] A. J. Weiss, "Direct position determination of narrowband radio frequency transmitters," *IEEE Signal Processing Letters*, vol. 11, no. 5, pp. 513–516, May 2004.

[129] K. Papakonstantinou and D. Slock, "Direct location estimation using single-bounce NLOS time-varying channel models," *Proceedings of the IEEE Vehicular Technology Conference (VTC) Fall*, September 2008.

[130] K. Papakonstantinou, "Applications of statistical signal processing in mobile terminal localization," Ph.D. dissertation, Telecom ParisTech, July 2010.

[131] P. Guillaume and R. Pintelon, "A gauss-newton-like optimization algorithm for weighted nonlinear least-squares problems," *IEEE Transactions on Signal Processing*, vol. 44, no. 9, pp. 2222–2228, September 1996.

[132] K. W. K. Lui, F. K. W. Chan, and H. C. So, "Semidefinite programming approach for range-difference based source localization," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1630–1633, April 2009.

[133] M. Laaraiedh, S. Avrillon, and B. Uguen, "Overcoming singularities in TDOA based location estimation using total least square," *Proceedings of the International Conference on Signals, Circuits & Systems (SCS)*, November 2009.

[134] M. Laaraiedh, "Contributions on hybrid localization techniques for heterogeneous wireless networks," Ph.D. dissertation, University of Rennes 1, December 2010.

[135] M. A. Spirito, "On the accuracy of cellular mobile station location estimation," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 3, pp. 674–685, May 2001.

[136] I. Sharp, K. Yu, and Y. J. Guo, "GDOP analysis for positioning system design," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3371–3382, September 2009.

[137] Y. Qi, T. Asai, H. Yoshino, and N. Nakajima, "On geolocation in ill-conditioned BS-MS layouts," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2005.

[138] A. G. Guimaraes and M. A. Grivet, "A study of ambiguity in hyperbolic mobile position location," *Proceedings of the European Personal Mobile Communications Conference*, April 2003.

[139] B. Krach, M. Lentmaier, and P. Robertson, "Bayesian detection and tracking for joint positioning and multipath mitigation in GNSS," *Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC)*, March 2008.

[140] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.

[141] S. Ali-Löytty, "Gaussian mixture filters in hybrid positioning," Ph.D. dissertation, Tampere University of Technology, August 2009.

[142] P. Maybeck, *Stochastic Models, Estimation and Control, Volume I*. Academic Press, 1979.

[143] P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, pp. 19–38, September 2003.

[144] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 425–437, February 2002.

[145] B. Krach, "Sensor fusion by bayesian filtering for seamless pedestrian navigation," Ph.D. dissertation, University of Erlangen-Nürnberg, March 2010.

[146] A. Doucet, J. F. G. de Freitas, and N. J. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. de Freitas, and N. J. Gordon, Eds. Springer Verlag, 2001.

[147] N. Gordon, D. Salmond, and A. F. M. Smith, "Novel approach to non-linear and non-Gaussian Bayesian state estimation," *Proceedings of the IEE*, vol. 140, pp. 107–113, 1993.

[148] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–599, June 1999.

[149] C. Musso, N. Oudjane, and F. LeGland, "Improving regularized particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. de Freitas, and N. J. Gordon, Eds. Springer Verlag, 2001.

[150] Y. Bar-Shalom, R. X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2001.

[151] T. Perälä and R. Piché, "Robust extended Kalman filtering in hybrid positioning applications," *Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC)*, March 2007.

[152] E. A. Wan and R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," *Proceedings of the Adaptive Systems for Signal Processing, Communications, and Control Symposium*, October 2000.

[153] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," *Proceedings of AeroSense*, 1997.

[154] J. K. Uhlmann, "Simultaneous map building and localization for real time applications," University of Oxford, Technical Report, 1994.

[155] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, March 2004.

[156] T. Schön, F. Gustafsson, and P. J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, July 2005.

[157] M. Briers, S. Maskell, and R. Wright, "A Rao-Blackwellized unscented Kalman filter," *Proceedings of the International Conference on Information Fusion*, July 2003.

[158] B. Krach and R. Weigel, "Markovian channel modeling for multipath mitigation in navigation receivers," *Proceedings of the European Conference on Antennas and Propagation (EuCAP)*, March 2009.

[159] T. Perälä and S. Ali-Löytty, "Kalman-type positioning filters with floor plan information," *Proceedings of the International Conference on Advances in Mobile Computing & Multimedia (MoMM)*, November 2008.