UNIVERSITÄT DER BUNDESWEHR MÜNCHEN

FAKULTÄT FÜR ELEKTROTECHNIK UND INFORMATIONSTECHNIK

SYNCHRONIZATION AND EQUALIZATION FOR MULTI-TRANSCEIVER SYSTEMS WITH A FOCUS ON LINE-OF-SIGHT

Tim Hälsig

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

DOKTOR-INGENIEUR

(Dr.-Ing.)

genehmigten Dissertation.

Gutachter: Prof. Dr.-Ing. Berthold Lankl
Prof. Dr. techn. Dr. h.c. Josef A. Nossek

Die Dissertation wurde am 23. Juni 2020 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 21. November 2020 angenommen. Die mündliche Prüfung fand am 14. Dezember 2020 statt.

Kurzfassung

Diese Arbeit befasst sich mit Synchronisations- und Entzerrungsalgorithmen für frequenzselektive Multiple-Input Multiple-Output (MIMO) Systeme. Der Schwerpunkt liegt auf Line-of-Sight-MIMO-Übertragungen bei Millimeterwellenfrequenzen für Backhaul-ähnliche Szenarien. Der gewählte Ansatz ist ferner allgemein für MIMO-Systeme mit räumlichem Multiplexing gültig. Es wird gezeigt, dass in einem drahtlosen Übertragungssystem typischerweise zwei Synchronisationsbeeinträchtigungen vorhanden sind, nämlich Trägerfrequenzversatz und Symboltaktunterschiede. Beide existieren aufgrund der Nachfrage nach Hardware- und insbesondere Oszillator-Implementierungen mit geringer Komplexität. Sie können mit einem Phasenversatzprozess bezogen auf den jeweiligen idealen Sollwert charakterisiert werden. Die Beeinträchtigungen führen dazu, dass die beobachtete Kanalübertragungsfunktion zwischen Sender- und Empfängerbasisband im Allgemeinen zeitvariant ist. Für MIMO-Systeme, in denen mehrere Transceiver verwendet werden, wird gezeigt, dass je nach Hardware-Aufbau eine unterschiedliche Anzahl dieser Phasenprozesse zu beobachten ist. Wenn zum Beispiel die Transceiver auf beiden Seiten der Verbindung weit voneinander entfernt sind, ist es oft nur schwer möglich, eine Frequenzreferenz zwischen ihnen zu teilen. In diesem Fall existieren so viele unabhängige Phasenprozesse wie Transceiver im System.

Es werden zwei auf Trainingssignalen basierende Kanalschätzverfahren vorgeschlagen, um die Phasenprozesse und die zeitveränderlichen Kanaleigenschaften zu identifizieren. Eines davon ist eine Korrelationsmethode, die ein aus der Literatur bekanntes Verfahren ist, aber in dieser Arbeit erstmals für den Fall mehrerer Synchronisationsbeeinträchtigungen, insbesondere mehrerer Symboltaktunterschiede, in einem MIMO-System untersucht wird. Als Alternative wird die adaptive Filterung zur direkten Identifikation und Verfolgung des zeitvarianten Systems vorgeschlagen. Für beide Ansätze werden Empfehlungen zur Parameterauswahl gegeben. Insbesondere werden einige neue Ergebnisse für die Schrittweitenwahl des adaptiven Filters für MIMO-Systeme mit mehreren Synchronisationsbeeinträchtigungen abgeleitet. Standard und adaptive Entzerrung werden zur Kompensation der Kanalbeeinträchtigungen diskutiert. Es zeigt sich, dass die Komplexität der Entzerrung stark von der Oszillatoranordnung abhängt. Für den allgemeinsten Fall wird beispielsweise der komplexeste Entzerrer benötigt, welcher überabgetastet ausgeführt werden muss.

Simulations- und Messergebnisse bestätigen, dass die vorgeschlagenen Schätz- und Entzerrungsstrategien praktikabel sind und zeigen, dass eine drahtlose Übertragung mit mehreren Gigabit pro Sekunde mit einem räumlich-multiplexenden Line-of-Sight-MIMO-System möglich sind, selbst wenn für jede Transceiver-Kette unabhängige Oszillatoren verwendet werden.

Schlagwörter: Synchronisation, Entzerrung, MIMO, LoS, räumliches Multiplexing, Frequenzversätze, Symboltaktunterschiede, Parameterschätzung, adaptive Filter, Millimeterwellen, 60 GHz, Multi-Gigabit pro Sekunde, Messung

Abstract

This work considers the problem of synchronization and equalization for frequency-selective multiple-input and multiple-output (MIMO) systems. The focus is on line-of-sight MIMO transmission at millimeter-wave frequencies for backhaul-like scenarios, but the approach is generally applicable to MIMO systems using spatial multiplexing. It is shown that two timing impairments are typically present in a wireless transmission system, namely carrier frequency offset and sampling frequency offset. They exist due the desire for low-complexity hardware, and in particular oscillator, implementations. Both of them can be characterized with a phase offset process with respect to their ideal nominal value. These impairments cause the observed channel characteristic, between transmitter and receiver baseband, to be time varying in general. For MIMO systems, where multiple transceivers are used, it is discussed that a different number of these phase processes will be observed, depending on the hardware setup. For example, when the transceivers are widely separated on either side of the link, sharing a time reference between them may be infeasible, meaning that as many independent phase processes as transceivers exist in the system.

Two training signal based channel estimation techniques are proposed, in order to identify the phase processes, and the time-varying channel characteristics in general. One of them is a correlation method, which is a standard technique known from the literature, but will in this work be firstly investigated for the case of multiple timing impairments, especially multiple sampling frequency offset processes, in a MIMO system. As an alternative, adaptive filtering is proposed for direct identification and tracking of the time-varying system. Parameter selection recommendations for both approaches are given. In particular, some new results for the step size selection of the adaptive filter for MIMO systems with multiple timing impairments are derived. Standard and adaptive equalization are discussed for dealing with the channel impairments. It is seen that standard equalization greatly simplifies, depending on the oscillator setup. However, for the most general case, the equalizer needs to be fully fractionally spaced.

Simulation and measurement results corroborate that the proposed estimation and equalization strategies are viable, showing that multi-gigabit per second wireless transmission is feasible with a spatial-multiplexing line-of-sight MIMO system, even when independent oscillators are used for each transceiver chain.

Keywords: synchronization, equalization, MIMO, LoS, spatial multiplexing, frequency offsets, symbol timing, parameter estimation, adaptive filters, mmWave, 60 GHz, multi-gigabit per second, measurement

Preface

This thesis was written during the five and a half years I spent working at the Institute for Communications Engineering at the Universität der Bundeswehr München. During my time there, I had the pleasure of working and spending time with several insightful and interesting people. First and foremost, I would like to express my gratitude to my thesis advisor Prof. Berthold Lankl. He is one of the most knowledgeable and astute engineers I have met in the wireless communication industry so far. He was always very helpful and supportive, whether it was about getting some new hardware, or about going on a research stay of several months to Rio de Janeiro. I would also like to thank Prof. Josef A. Nossek for serving on my thesis committee, and for the few, yet always insightful and stimulating, discussions that we had during several conferences and at other places. Special thanks go to Prof. David V. Thiel and Prof. Lukas T. N. Landau for motivating me to enter the field of wireless communications research. I will always be indebted to Lukas, who got me started writing research papers, and who was a big part of making my research stay at PUC in Rio possible. At PUC, the idea for using adaptive filters for MIMO synchronization and equalization problems came up. I am very happy that I got to stay in Rio, and to have met a bunch of fascinating people there, like Prof. Rodrigo C. de Lamare, Marcelo Balisteri, Maboud F. Kaloorazi, and Dr. Joao T. Dias, just to name a few.

In Munich, I spent most of my time working on maximumMIMO, which was a great project to investigate and learn all kinds of different aspects of wireless communications, and I highly appreciate the funding by the German Research Foundation. During that project, I had the pleasure of working and publishing results with Dr. Xiaohang Song, Dr. Wolfgang Rave, Prof. Gerhard P. Fettweis, Darko Cvetkovski, and Prof. Eckhard Grass. I want to especially highlight Darkos contributions by spending long days with me during our measurement campaigns in Frankfurt (Oder). With no guarantee of completeness, I want to list some former and current colleagues, who helped me in one way or another to finish this journey: Dr. Mohammad Vatankhah, Dr. Stefan Ludwig, Bernhard Börner, Karl Heinrich Besthorn, Zifeng Wu, Vito Dantona, Anselm Karl, Makram El Chamaa, Kai-Uwe Storek, Florian Brauchle, and Andreas Lang. Specially with the last two I had many fun hours during and after work.

When you move to a new place for work, like from Dresden to Munich in my case, it is always a bit difficult in the beginning. Luckily, I found some like-minded people early on outside of work through playing volleyball. The first group with whom I spent quite some time includes Dr. Stephan Bansmer, Dr. Anja Kölzsch, Drs. Margit and Christian Egerer, and Prof. Stefan Hickel, to mention a few. Later on, that group of friends expanded and got to include

Thomas Zwanziger, Dr. Babette van Baarsel, Dr. Ralf Mörschel, Dr. Zhaohai Jiang, Dr. Maximilian Bradler, Alexander Amon, Roland Sadlek, Kai Habenreich, Prof. Tobias Harks, and (soon to be Dr.) Leonhard Binzenhöfer, most of whom I met through SW Volleyball München, and all of whom made living in Munich such a great delight.
When I left Dresden, I left behind a group of four people, with whom I spent most of my time on- and off-campus during my diploma studies. They are now Dr. Lukas Lorenz, Dr. Steve Bigalke, Dr. Sebastian Pech, and Dr. Martin Schubert. I am not very happy that I am the last one to finish this journey, but it still amazes me that all of us managed to complete it. Ours is certainly a special group of people and I hope that each of us finds his own way to continue after this milestone.
Furthermore, I would like to express my gratitude to my family and extended family for the support that I have gotten over the years, and without which this thesis would not have been possible. I particularly want to thank my grandparents, Gudrun and Klaus, and Evelyn and Gerald. I am also indebted to my parents and extensions thereof, Anke and Jens, and Sven. All of you always helped me when I needed it, and make life in general so much more fun. Finally, I want to appreciate the love and distraction from my more immediate family (or pack), Verena and Mikey. Verena, it always impresses me how positive, fun, and loving a person can be. I hope I can match this at least to some extent. I am so excited that our family will grow next year. To the little one(s), I can only hope that you encounter as many interesting, smart, fun, loving, and supportive people over your future years, as I have met along this journey.

I dedicate this thesis to all of you.

T. H.
Dresden, December 2020

Contents

1 Introduction

1.1 Synchronization and Communication

Synchronization is one of the fundamental processes in nature and technology, ranging from nature spectacles like fireflies flashing in unison on the river banks of Mexico, in order to attract mating partners, over the pulsing behavior of neurons to process information in the brain, to the calibration of vast antenna arrays, in order to detect the faintest signals from the beginning of the universe, [1, 2]. In essence, the task is described by matching the time reference of two or more spatially-separated systems. This can be achieved in a number of ways, like sharing the time reference between the systems, using high precision references for each system, or adjusting the time references of the systems after measuring their differences. Examples for these approaches are: distribution of a laser reference through fiber optic cables in radio astronomy [2], using highly accurate atomic clocks to provide the time references for navigation satellites, and adjusting a wrist watch after noticing a difference with respect to (w.r.t.) the standard time. From these examples, we can already identify one of the main goals in the field of synchronization, which is to use as few highly accurate time references as possible, since they are typically bulky, complex and expensive. Going back, the current navigation satellite standards have had such an incredible success because billions of users can rely on the tight synchronization between the different satellites, and do not require such a precise time reference themselves. Likewise, billions of watches can be set and updated periodically, some of them automatically, to one highly precise standard time, lowering the cost per watch significantly.

For communication systems, synchronization is required to reliably detect the message that has been transmitted. This includes finding the beginning of the message, as well as matching transmitter (Tx) and receiver (Rx) frequencies, and transmission rates [3, 4, 5]. The difference in the frequencies and rates occurs because the time reference in Tx and Rx are, due to their spatial separation, subject to different environmental impacts, and also simply vary due to manufacturing variability. Only when they are matched can the maximum possible amount of information be transmitted across the link. To be more specific, the focus of this work is

on undesired timing impairments in these systems, which have to be resolved. For example, assume a radio broadcast station that is transmitting at 94.50 MHz, but a receiving radio that is tuned to 94.40 MHz. Although automatic retuning of the Rx to 94.50 MHz could be treated as a synchronization problem, the focus here is on the case where the receiver is set to 94.50 MHz, but the inherent impairments of the circuitry only allow frequency generation of, e.g., 94.49 MHz.

The more transmit and receive nodes exist in the system, the more timing impairments have to be corrected, when a joint processing, e.g., to cancel interference, is desired to achieve high throughputs. In radio terms, imagine two different radio stations transmitting on the same frequency, with two receivers each wanting to receive only one of the stations, but actually receiving a superposition of both. The benefit of such a system is immediately clear, the two radio stations do not need to be separated in frequency, but can transmit on the same frequency resource, given that the interference of the other station can be canceled at the corresponding receiver. This essentially doubles the throughput compared to the more common system setup, where the stations are separated in frequency. It will be seen in this work that in order to cancel the interference, it is required that some joint processing, having knowledge about all the timing impairments in the system, is carried out. Such a system with multiple transmitting nodes and multiple receiving nodes using some joint processing to cancel interference, and transmitting on the same frequency resource, is called a multiple-input and multiple-output (MIMO) system. For such systems, synchronization is relatively straightforward when a shared time reference is considered for the nodes on the transmitter and receiver side. The other configuration with independent time references for each node has only been treated to a limited extent in the literature, is significantly more complex, and will be the main focal point of this work.

1.2 Application Examples & mmWave LoS MIMO

This thesis is written with a main application in mind, namely high throughput backhaul in a line-of-sight (LoS) scenario. With the increase in mobile data demand, the need for flexible low power and high data rate backhaul solutions also rises [6]. One solution to the increased backhaul traffic demand is to use wireless LoS MIMO systems at millimeter-wave (mmWave) frequencies. These systems offer great deployment flexibility, and can achieve high data rates in the order of 100 Gbit/s [7, 8, 9] through their large available bandwidths [10, 11, 12], and good spatial-multiplexing capabilities [13]. Spatial multiplexing describes the technique mentioned in the previous section, where multiple independent data streams may be transmitted during the same time and on the same frequency, increasing the throughput of the link. For spatial multiplexing to work in LoS conditions, the different transmitting and receiving nodes need to be widely separated on each side of the link [14, 15, 16], making the sharing of a common time reference difficult. Using multiple independent time references can, thus, be a requirement for these systems. LoS MIMO as a technique is, furthermore, particularly suited for mmWave systems, since such systems often require LoS conditions and

highly directive antennas to start with, in order to generate sufficient link margin. Another benefit is that due to the short wavelength and relatively short link distances, the required node separation for LoS MIMO systems is in a practically feasible range.

The considered wireless LoS MIMO systems are also a feasible solution for other applications, such as a short range wireless bus for data centers, short range fixed indoor data kiosk systems, vehicle-to-vehicle transmissions, or satellite systems [17]. More generally, the synchronization task dealing with multiple independent time references also needs to be addressed for other spatially-separated MIMO configurations, e.g., spatially-multiplexed multi-user systems [18], cooperative basestations, or other distributed antenna systems [19]. The last topic also leads to beamforming systems, where synchronization of a potentially huge number of different time references may be required [20, 21, 22, 23]. If synchronization cannot be achieved on the array level for such systems, beamforming becomes infeasible, as the weights need to be adapted continuously to compensate the variations due to the time reference differences.

1.3 Contributions of the Thesis

This section will briefly summarize the main outcomes of the thesis, what is different compared to results from the literature, and which system cases will not be explicitly considered. The following list sums up the main contributions:

- Full carrier and sampling phase process modeling, including the case where the sampling phase offset process is continuously changing, and not just a fixed sample phase offset as it is often approximated as in the literature.
- System description for different sampling scenarios, and under different timing impairments.
- Consideration of parallel linear filters for each MIMO transceiver, to account for different linear distortions of each chain.
- System model including multiple independent carrier and sampling phase processes.
- Correlation-based estimators for frequency-selective MIMO channels with multiple independent carrier and sampling phase processes.
- Separation of transmitter and receiver contributions from the carrier and sampling phase difference process estimates.
- Adaptive filter for estimating and tracking the complete time-varying MIMO channel including the timing impairments.
- Selection of a well-performing step size for the adaptive filter in the MIMO system estimation case with multiple carrier phase offset processes.

- Selection of a well-performing step size for the adaptive filter in the MIMO system estimation case with multiple sampling phase offset processes.
- Discussion of different oscillator setups and their impact, in particular, for equalization.
- Investigation of the viability of adaptive equalization for LoS MIMO systems with multiple timing impairments, and step size selection guidelines for adaptive equalization in such systems.
- Measurement results characterizing LoS MIMO systems at 60 GHz.
- Measured adaptive equalization results for a two- and four-antenna spatial-multiplexing LoS MIMO system, achieving multi-gigabit per second data rates.

Compared to the literature, the main difference is that the two timing impairments, i.e., carrier frequency offsets (CFOs) and sampling frequency offsets (SFOs), as well as a frequency-selective MIMO channel, are considered at the same time. In general, the goal is to give a comprehensive overview of synchronization and equalization in spatial-multiplexing MIMO systems with multiple timing impairments. Especially the presence of multiple sampling phase processes is novel, and has only been treated scarcely in the literature. Frequently, synchronization is about computing an error signal and using it for compensating the timing impairments through some form of feedback. For MIMO systems with multiple timing impairments, the situation is more complicated as the error signals contain multiple frequency differences, which have to be matched to the corresponding transceiver chain. We will assume that the frequencies of the independent oscillators of the spatially-separated transceivers are approximately accurate, i.e., the residual frequency errors are in the parts-per-million (ppm) range. The focus will be solely on digital synchronization algorithms, where no feedback to the analog domain is assumed possible or necessary, because that has been the foregoing trend in the communications engineering literature and system designs. This is especially beneficial as only some common baseband processing is required, while the analog transceivers can be fully separated. Throughout this work it will, accordingly, be assumed that a joint processing, i.e., a central receiver collects the signals from all receivers, is possible. Note, however, that for some oscillator setups or channel conditions, separate processing directly follows from the shown results. Additionally, low-complexity adaptive filtering has not been considered in order to deal with the multiple timing impairments in such MIMO systems. It was noted in [24, 25, 26] that fractionally-spaced adaptive equalizers can deal with part of the timing impairments in single-input and single-output (SISO) systems. Particularly, Gardner [25] noted that symbol-timing adjustment is partly done by adaptive equalizers, but that this notion should be reconsidered for high-speed systems. We propose here that adaptive equalization is a viable solution especially for such systems, where data is abundant and the relative change in the channel is fairly slow, in order to deal with both timing impairments.

This work will not consider coding, or joint equalization and decoding, mainly due to the complexity and breadth that this topic brings, especially for multi-gigabit per second systems.

Transmit signal shaping, e.g., precoding, will not be considered explicitly, as it can change the correlation structure of the input signals, meaning that some of the assumptions, which are used in this work, do not hold anymore. Additionally, the impact of Doppler shifts will not be explicitly treated, as we are mostly interested in quasi-static backhaul-like transmission links. However, most of the effects due to Doppler create a similar behavior as the carrier and sampling frequency offsets, which will be treated extensively throughout the work. Finally, the sample drop and add problem, which occurs for long-time continuous transmission due to the difference between transmitter and receiver rate, will not be treated explicitly. We will, however, briefly explain the problem, where it is coming from, and point to solutions from the literature. For the results that are shown in this work, this problem was made negligible by choosing burst transmissions of convenient length.

1.4 Thesis Organization

In chapter 2, the basic fundamentals for MIMO systems, and in particular LoS MIMO, are described. It is derived how the system model arises, and why and when spatial multiplexing in LoS conditions is feasible. The discrete-time models of different MIMO system configurations are laid out, introducing several of the basic signal vectors that are used throughout the work. A basic introduction into the influence and behavior of oscillators in communication systems is included, describing the problem of frequency differences, and introducing a phase process model that can be used to represent the frequency differences. Furthermore, the influence of oscillators on MIMO systems is briefly discussed. Finally, the chapter contains a summary of the most important results for MIMO systems with timing impairments.

Chapter 3 deals with parameter estimation for frequency-selective MIMO systems with multiple timing impairments. Some basic one-shot estimators, stemming from a maximum-a-posteriori (MAP) approach, are derived, and the need for training signals in MIMO systems is discussed. Further, one-shot estimators for different specific MIMO system models are developed, which simplify processing and separate the different parameters. As an alternative to these approaches, estimation based on adaptive filtering is proposed. The basics of adaptive filtering based on the least-mean-squares (LMS) principle are described, and the significant statistics for MIMO systems with timing impairments are derived. Convergence and tracking behavior for estimation are investigated, and several different suggestions on how to select the adaptive filter parameters are given. Some of them are optimal solutions under some conditions and require certain approximations. At the end of the chapter, simulation results for both estimation approaches are provided, showing their performance.

Equalization and synchronization for MIMO systems with multiple timing impairments is described in chapter 4. It is shown how the estimated parameters can be used for equalization, and how equalization and synchronization interact. Separation of the timing impairments from the MIMO channel is discussed, and simplified equalization schemes for different system configurations are derived. As an alternative, adaptive equalization, also based on

adaptive filtering, is introduced to directly deal with all channel influences simultaneously. Convergence and tracking for the equalization case are investigated, and some suggestions for selecting the adaptive equalizer parameters are given. Simulation results are provided for the adaptive equalizer in different system scenarios.

Chapter 5 provides experimental results with a practical mmWave setup. The basic measurement setup is described, and the LoS MIMO spatial-multiplexing capabilities are measured. Furthermore, some statistical variations of LoS MIMO channels and systems at mmWave frequencies are quantified. In particular, the change in carrier phase difference processes with different oscillators is shown. Finally, the adaptive equalization performance for a two- and four-antenna setup, with independent oscillators for carrier frequency generation, is investigated.

The thesis concludes with chapter 6, which summarizes the most important aspects of the work, and describes open topics that require further investigation.

Two appendices augment the thesis. In appendix A, additional information about oscillators, timing impairments, and their impacts, supplementing the material of chapter 2, can be found. Appendix B contains additional derivations for the parameter selection of the adaptive filters proposed for channel estimation in chapter 3.

1.5 Remarks on Notation and Normalizations

In this thesis, vectors and matrices are defined with bold letters, with vectors being, furthermore, defined as column vectors, e.g., $\mathbf{x}[k] = \begin{bmatrix} x_1[k] & x_2[k] & \cdots & x_N[k] \end{bmatrix}^{\mathrm{T}}$. The discrete-time index k will always be associated with the nominal receiver sampling rate. Accordingly, when the oversampling factor is chosen as $Q = 1$, this means that $y_m[k]$ will correspond to the kth received symbol from the mth antenna. For $Q > 1$, $\dot{y}_m[k]$ corresponds to the kth sample from the mth antenna. Oversampled quantities, like this sample, and variables related to the sampling process are marked with a small dot above, i.e., $\dot{(\cdot)}$. The discrete-time notation will sometimes be overloaded, in that we write $h_{mn}[l+\dot{\phi}]$ for a discrete-time impulse response. Rigorously, the discrete-time version is only defined for integer values in the squared brackets. When we write $h_{mn}[l+\dot{\phi}]$, we mean the discrete samples of $h_{mn}(t)$ at sampling phase $\dot{\phi}$, in order to make the difference compared to the nominal ideal sampling rate more visible. Most of the variables should be considered complex, especially when dealing with signals. It should be clear from the context when a variable is just real. Throughout this work, we tried to stick to the convention of transmitter-receiver space-time (TxRx-ST) stacking, meaning that first Tx/Rx space and then time are stacked, for vectors where this is applicable. For example, consider the discrete-time finite-length channel impulse responses in a MIMO system $h_{mn}[l]$ that are associated with the mth receive antenna. The corresponding channel vector is then stacked as $\mathbf{h}_m = \begin{bmatrix} h_{m1}[0] & \cdots & h_{mN}[0] & \cdots & h_{m1}[L-1] & \cdots & h_{mN}[L-1] \end{bmatrix}^{\mathrm{T}}$. The definition of probability density functions will be a bit lax in order to avoid a lengthy notation. In particular,

$p(x)$ will mean the probability density function (PDF) of the random variable X, which is mathematically and notationally more precisely defined as $p_X(x)$, in order to distinguish between the random variable itself and its realization. The same thing also applies to operators on random variables, for example $\mathrm{E}[x]$, which should be $\mathrm{E}[X]$. To shorten notation, the time index k will sometimes be omitted for quantities with time-varying statistical properties, e.g., the time-varying expected value $\mathrm{E}[X[k]] = \mathrm{E}[x[k]]$ may be simplified to $\mathrm{E}[x]$. It should be clear from the context, when this is the case.

For practical reasons[1], a per antenna average transmit power constraint will often be employed. Furthermore, the channel impulse response of any sub-path, i.e., each connection between the nth Tx and mth Rx, will be normalized to fulfill $\sum_l |h_{mn}[l]|^2 = 1$. This impulse response normalization, i.e., the same power for each sub-connection, implicitly means that there is equal sum gain between all Tx and Rx antennas assumed, while in a practical scenario sub-connection power differences, e.g., due to antenna pattern differences, can be observed. To summarize, the two assumptions mean that adding transmit antennas increases the sum transmit power of the complete system, while adding receiving antennas increases the amount of transmitted power that is captured at the receiver. Additionally, the fixed input power of one antenna is distributed over all multi-path components, according to the significance[2] of each path.

An overview of most of the used nomenclature and all abbreviations can be found at the end of this work.

[1] In most cases, we will use the maximum amount of efficiently available power in a transmitter, unless otherwise required, for example, through regulation, in order to maximize SNR.

[2] For example, when comparing a single-path with a multi-path channel, the complete transmitted power is concentrated in one path for the first, while it is distributed over all paths in a certain manner for the second. Practically, the power of one path does of course not reduce, when an additional one is observed. However, this normalization is mathematically convenient, because it means that the received power does not depend on the exact characteristics of the channel.

2 MIMO System Modeling & Synchronization Essentials

In this part, the main MIMO system modeling foundations are laid out, including generalized formulations for the effects of frequency-selective channels and independent oscillators for each of the MIMO antennas.

2.1 Continuous-Time Band-Pass MIMO Model

Before describing the more commonly used complex baseband notation for MIMO systems, a brief description of the signal in the actual transmission band is carried out, in order to outline why LoS MIMO can achieve spatial-multiplexing gains, and to describe some peculiar effects that occur due to high carrier frequencies and large bandwidths.

Consider, without loss of generality, the transmitted real band-pass signal from the nth transmit antenna to be given by

$$s_n(t) = \sqrt{2}\,\mathrm{Re}\left\{\bar{x}_n(t)\cdot\exp\left(j\left(2\pi f_{\mathrm{Tx},n}\cdot t+\phi_{\mathrm{Tx},n}\right)\right)\right\}, \tag{2.1}$$

where $\bar{x}_n(t)$ is the complex-valued baseband signal with symbol (or baud) rate $1/T_s$, $f_{\mathrm{Tx},n}$ is the carrier frequency from the local oscillator (LO) of that nth transmitter and $\phi_{\mathrm{Tx},n}$ is the phase of that LO[1]. Figure 2.1 shows a typical Tx system diagram for one antenna and the corresponding linear model, which will be used throughout the work, confer also [4]. The corresponding Rx diagram can be seen in Figure 2.2. The different components will be described in more detail in the following sections.

A linear time-invariant[2] band-limited wireless-transmission channel can be modeled as a

[1] For ease of exposition, the independent oscillators are assumed to generate a single frequency or spectral line with a fixed phase, both not necessarily the same, here. The practical case where deviations from those perfect frequencies and phases occur will be considered later on. Although different frequencies occur, it is assumed that the difference is very small, e.g., in the order of ppm, such that transmission still happens approximately in the same frequency band.

[2] This assumption is made, since the focus of this work is on backhaul systems with stationary transmitters and receivers. It will be seen later that the observed channel, including the hardware impairments, is time variant.

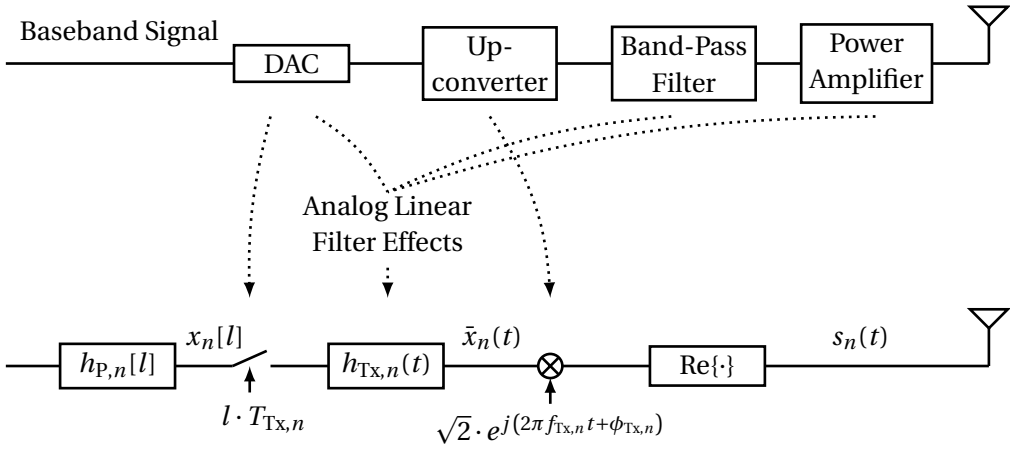


Figure 2.1: Exemplary transmitter system diagram, and the corresponding linear system model components, for one antenna of a MIMO system.

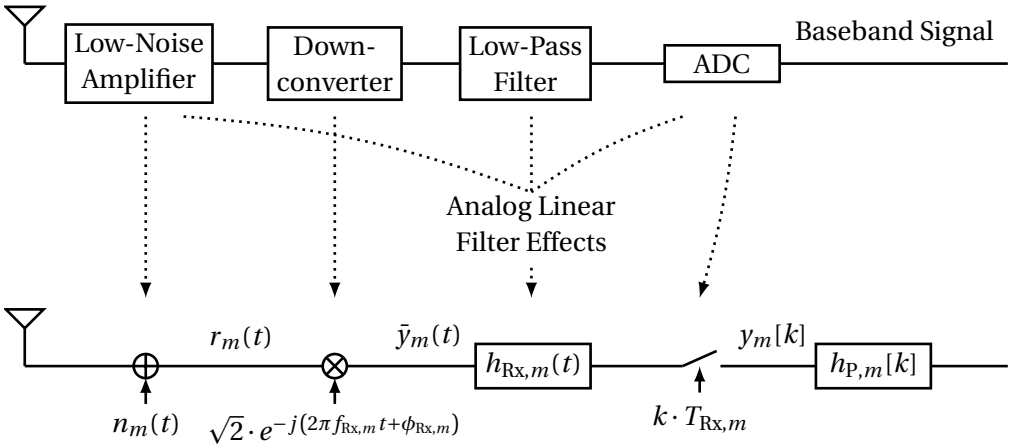


Figure 2.2: Exemplary receiver system diagram, and the corresponding linear system model components, for one antenna of a MIMO system.

number of paths arriving at different times and with different gains [4, 27, 28] at the various receive antennas, i.e.,

$$h_{\mathrm{BP},mn}(t) = 2\,\mathrm{Re}\left\{ \sum_{l=0}^{L_{mn}-1} a_{mn,l}\,\mathrm{sinc}\left(\pi \frac{t-\tau_{mn,l}}{T_{\mathrm{nom}}}\right) \cdot \exp\left(j2\pi f_{\mathrm{nom}} \cdot t\right) \right\}, \tag{2.2}$$

where $h_{\mathrm{BP},mn}(t)$ is the band-pass channel impulse response between nth transmit and mth receive antenna. Furthermore, L_{mn} is the number of different paths for that antenna pair, $a_{\cdot}$ is an amplitude coefficient, $\mathrm{sinc}(\cdot)$ is the sinc function, f_{nom} is the nominal carrier frequency, and T_{nom} is the nominal receiver sampling interval. The last two quantities will be discussed in more detail later on. The parameter $\tau_{\cdot}$ is the respective propagation delay [4, 27] of the different paths, which is given by the length of the particular path $r_{mn,l}$ and the speed of light

c with

$$\tau_{mn,l} = \frac{r_{mn,l}}{c}. \tag{2.3}$$

The band-pass signal at the mth receive antenna before downconversion is then given by

$$\begin{aligned} r_m(t) &= \sum_{n=1}^{N} h_{\mathrm{BP},mn}(t) * s_n(t) + n_m(t) && (2.4)\\ &\approx \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} a_{mn,l} \cdot s_n(t-\tau_{mn,l}) + n_m(t) && (2.5)\\ &= \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} a_{mn,l} \cdot \sqrt{2}\,\mathrm{Re}\left\{\bar{x}_n(t-\tau_{mn,l}) \cdot \exp\left(j\left(2\pi f_{\mathrm{Tx},n} \cdot \left(t-\tau_{mn,l}\right) + \phi_{\mathrm{Tx},n}\right)\right)\right\} && (2.6)\\ &\quad + n_m(t), \end{aligned}$$

where $*$ denotes convolution[3], and $n_m(t)$ is the noise process at the receiver, assumed to be white and Gaussian. Additionally, it was assumed that $T_{\mathrm{nom}} \ll T_s$, meaning $s_n(t) * \mathrm{sinc}\left(\pi \frac{t-\tau_{mn,l}}{T_{\mathrm{nom}}}\right) \approx s_n(t-\tau_{mn,l})$. Finally, downconverting the mth received signal to baseband, using suitable low-pass filters $h_{\mathrm{L},m}(t)$[4], see [4] for details, and omitting the impact of noise yields

$$\begin{aligned} \bar{y}_m(t) &= \left[r_m(t) \cdot \sqrt{2}\exp\left(-j\left(2\pi f_{\mathrm{Rx},m} \cdot t + \phi_{\mathrm{Rx},m}\right)\right)\right] * h_{\mathrm{L},m}(t) && (2.7)\\ &= \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} a_{mn,l} \cdot \bar{x}_n(t-\tau_{mn,l}) \cdot \exp\left(j\left(2\pi\Delta f_{mn} \cdot t - 2\pi f_{\mathrm{Tx},n} \cdot \tau_{mn,l} + \Delta\phi_{mn}\right)\right) && (2.8)\\ &= \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} a_{mn,l} \cdot \bar{x}_n(t-\tau_{mn,l}) \cdot e^{j2\pi\Delta f_{mn}\cdot t} \cdot e^{j\Delta\phi_{mn}} \cdot e^{-j2\pi f_{\mathrm{Tx},n}\cdot\tau_{mn,l}}, && (2.9) \end{aligned}$$

where $\bar{y}_m(t)$ is the complex baseband signal at the mth receive antenna, $\Delta f_{mn} = f_{\mathrm{Tx},n} - f_{\mathrm{Rx},m}$ is the frequency offset between the nth transmitter and mth receiver LO, and $\Delta\phi_{mn} = \phi_{\mathrm{Tx},n} - \phi_{\mathrm{Rx},m}$ is the corresponding phase offset. Special focus should be put on two terms of equation (2.9). First, $\bar{x}_n(t-\tau_{mn,l})$ shows that depending on the symbol rate $1/T_s$ of the transmitted baseband signal, and the propagation conditions in the channel, intersymbol interference (ISI) can occur, which needs to be dealt with by using a suitable equalization structure. Secondly, the last two terms in (2.9) determine the channel phases and, therefore, capabilities of the MIMO system. If the channel phases are sufficiently independent between the different antennas, spatial multiplexing can be used yielding maximum MIMO gain. Note that the impact of the first important term of (2.9), $\bar{x}_n(t-\tau_{mn,l})$, depends on the symbol rate $1/T_s$, while the impact of the second important term depends on transmit carrier frequency $f_{\mathrm{Tx},n}$, which will be an important fact for LoS MIMO as considered in the following.

[3] Here, it is implicitly assumed that the system is linear.

[4] The low-pass filter $h_{\mathrm{L},m}(t)$ is not shown explicitly in the system model of Figure 2.2, but is contained in the linear filter $h_{\mathrm{Rx},m}(t)$ expressing the combined linear-filtering characteristics of a particular receive chain.

2.2 Properties of LoS MIMO

2.2.1 Spatial Multiplexing

Assume a linear time-invariant pure LoS transmission, i.e., $L_{mn} = 1$, meaning that only a LoS propagation path is present. In that case, the propagation times $\tau_{mn,0}$ depend only on the path length between the nth transmit and mth receive antenna, as described by equation (2.3) and shown in Figure 2.3, in a simple geometric way. Furthermore, the coefficients $a_{mn,0}$ are purely real and have no phase contribution[5], as they only contain the power losses in the channel from the finite directivity of the antennas.

What generally matters is not the absolute time of $\tau_\cdot$, but rather the differences $\Delta\tau_\cdot = \tau_{mn,0} - \tau_{ij,0}$, with $mn \neq ij$, between all the different paths. Typically, it has been assumed that the term $f_{\mathrm{Tx},n}\Delta\tau_\cdot = \Delta r_\cdot / \lambda_{\mathrm{Tx},n}$, where $\Delta r_\cdot = r_{mn,0} - r_{ij,0}$ is the path length difference and $\lambda_{\mathrm{Tx},n}$ is the wavelength, is negligible or can be well approximated by a planar wavefront for LoS transmission. While physically not correct, this is a valid approximation for closely-spaced arrays at lower frequencies transmitting over long distances [33]. The effect of the approximation is that all receiving antennas exhibit the same channel phase with respect to one transmitting antenna, which obviates the possibility of spatial multiplexing because the rows of the channel matrix $\mathbf{H}_0$, to be introduced later on, are always linearly dependent irrespective of the antenna array geometry.

However, for a number of cases the term $f_{\mathrm{Tx},n}\Delta\tau_\cdot$ is significant and should be accounted for. With some reordering and a slight change of notation from (2.9), we have first

$$\bar{y}_m(t) = \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} \bar{x}_n(t) \cdot e^{j2\pi\Delta f_{mn} \cdot t} \cdot e^{j\Delta\phi_{mn}} * \underbrace{\left(a_{mn,l} \cdot e^{-j2\pi f_{\mathrm{Tx},n} \cdot \tau_{mn,l}} \cdot \mathrm{sinc}\left(\pi \frac{t - \tau_{mn,l}}{T_{\mathrm{nom}}} \right) \right)}_{h_{mn}(t)}, \quad (2.10)$$

where $h_{mn}(t)$ denotes the channel impulse response transformed to baseband. Then, considering the pure LoS case with $L_{mn} = 1$, and the additional assumption that the symbol duration is much higher than all propagation time differences, i.e., $\forall\Delta\tau_\cdot \ll T_s$, the LoS MIMO channel can be modeled[6] by a fixed complex constant with a common propagation delay τ for all paths as

$$h_{mn}(t) = a_{mn,0} \cdot e^{-j2\pi f_{\mathrm{Tx},n} \cdot \tau_{mn,0}} \cdot \mathrm{sinc}\left(\pi \frac{t - \tau_{mn,l}}{T_{\mathrm{nom}}} \right) \quad (2.11)$$

$$\approx a_{mn,0} \cdot e^{-j2\pi f_{\mathrm{Tx},n} \cdot \tau_{mn,0}} \cdot \delta(t - \tau) = a_{mn,0} \cdot e^{-j2\pi \frac{r_{mn,0}}{\lambda_{\mathrm{Tx},n}}} \cdot \delta(t - \tau). \quad (2.12)$$

It is seen that depending on the antenna arrangement on the transmitter and receiver side, different values for $r_{mn,0}$ and, therefore, $\Delta r_\cdot$ generating different channel phases, are obtained.

[5] Physically, it can happen that additional phase shifts aside from the propagation delay occur when the dielectric properties of the propagation medium, e.g., due to rain [29, 30, 31], are not equal for all paths. However, this can also be included in the last term of (2.9) by using the effective wavelength of that path [32].

[6] For the approximation in the second line it was again assumed that $T_{\mathrm{nom}} \ll T_s$.

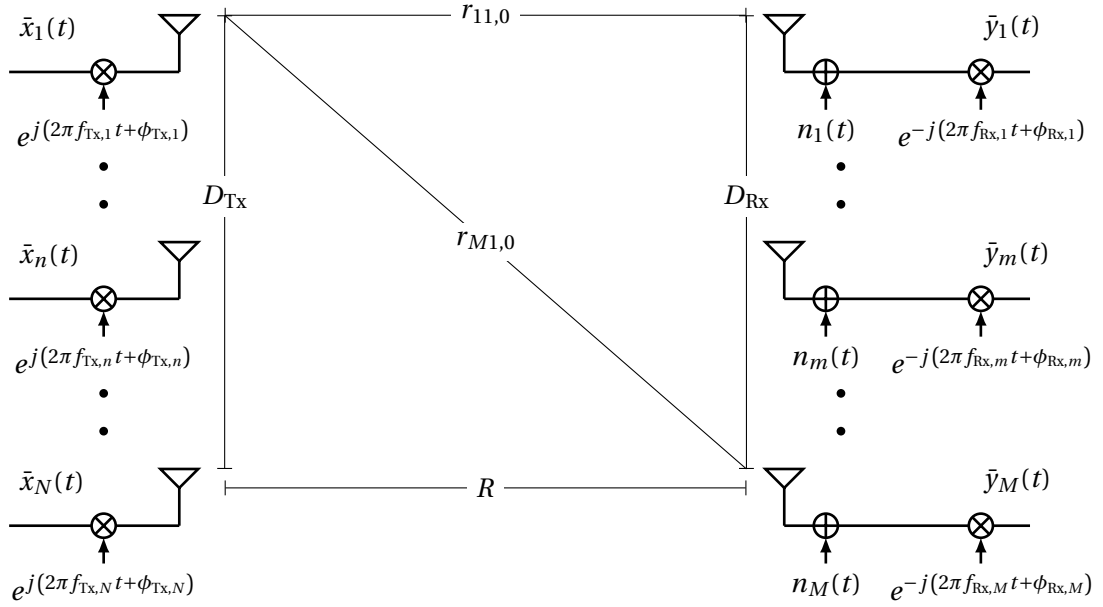


Figure 2.3: Simplified band-pass model of a one-dimensional LoS MIMO system with independent and ideal carrier frequencies for each antenna. Taking only the real part of the signals on the transmitter side, and some normalization constants are omitted for brevity.

In order to gain some insight for when the term becomes significant, consider the example from Figure 2.3, a thorough treatment of the topic can be found in [33]. If the maximum path length difference $\Delta r_{\max} = \max_{\forall mn} r_{mn,0} - \min_{\forall mn} r_{mn,0}$ is negligible compared to $\lambda_{\mathrm{Tx},n}$, all the smaller ones are too, and the phase shift between the different paths is insignificant. For the example, this term can be written as

$$\Delta r_{\max}/\lambda_{\mathrm{Tx},n} = \frac{r_{M1,0} - r_{11,0}}{\lambda_{\mathrm{Tx},1}} = \Delta\tau_{\max} f_{\mathrm{Tx},n} \tag{2.13}$$

$$= \frac{\sqrt{R^2 + D^2}}{\lambda_{\mathrm{Tx},1}} - \frac{R}{\lambda_{\mathrm{Tx},1}} \tag{2.14}$$

$$\approx \frac{\sqrt{R^2}}{\lambda_{\mathrm{Tx},1}} + \frac{1}{\lambda_{\mathrm{Tx},1}} \frac{D^2}{2\sqrt{R^2}} - \frac{R}{\lambda_{\mathrm{Tx},1}} \tag{2.15}$$

$$= \frac{D^2}{2\lambda_{\mathrm{Tx},1} R}, \tag{2.16}$$

where R is the propagation distance between the arrays, $D = D_{\mathrm{Tx}} = D_{\mathrm{Rx}}$ is the largest dimension of the array, and the root approximation in (2.15) can be made if R is much larger than D. The ratio $\Delta r_{\max}/\lambda_{\mathrm{Tx},n}$ gives an indication how much phase shift between the shortest and longest path in the system occurs. Thus, in order to have a significant phase shift between these paths, equation (2.16) shows that, for the considered example, the terms D^2 and $2\lambda_{\mathrm{Tx},1} R$ should be of similar magnitude.

This observation is in line with earlier results from the literature, e.g., [34] for one-dimensional

antenna arrays, where it was shown that the possible number of spatially-multiplexed streams, or spatial degrees of freedom (SDoF), for given array dimensions D_{Tx}, D_{Rx}, and wavelength-distance product $\lambda_{\mathrm{Tx},n}R$, is obtained by

$$\mathrm{SDoF} = \frac{D_{\mathrm{Tx}} D_{\mathrm{Rx}}}{\lambda_{\mathrm{Tx},n} R} + 1. \tag{2.17}$$

Let us briefly consider the impact of different transmit wavelengths $\lambda_{\mathrm{Tx},n}$. While the different values, due to independent LOs, affect the LoS MIMO channel phases, the effect is in the ppm range, proportional to the frequency difference of LOs, and does, therefore, not induce a significant difference. An active control of the channel phases would be possible by changing the wavelength $\lambda_{\mathrm{Tx},n}$ of each transmitter. This seems practically less relevant because only substantial changes of $\lambda_{\mathrm{Tx},n}$, in the order of several percent, would surmount to significant variations. However, changing the wavelength by those orders of magnitude also changes the carrier frequencies by the same amount, such that transmission happens in an entirely different band, which the system would need to be able to operate in.

2.2.2 Frequency Selectivity

In this section, it will be shown that frequency-selective channel behavior can be exhibited for certain systems, even in a pure LoS propagation channel. Part of equation (2.9) is the term $\bar{x}_n(t-\tau_{mn,l})$ which accounts for the time shift of the baseband signal due to the propagation time differences $\tau_{mn,l}$ of the different paths. Similar to the previous section the time differences or spreads $\Delta\tau_\cdot$ of the different paths matter most. However, compared to the previous considerations the magnitude of $\Delta\tau_\cdot$ now needs to be compared to the average symbol rate $1/T_s$ of the baseband signal $\bar{x}_n(t)$, instead of the carrier frequency.

Considering the same example as above and given in Figure 2.3, and checking again just the maximum difference $\Delta\tau_{\max} = \max_{\forall mn}\tau_{mn,0} - \min_{\forall mn}\tau_{mn,0}$ compared to the symbol rate, gives similarly

$$\Delta\tau_{\max}/T_s = \frac{\sqrt{R^2+D^2}}{cT_s} - \frac{R}{cT_s} \tag{2.18}$$

$$\approx \frac{D^2}{2cT_sR}. \tag{2.19}$$

The ratio $\Delta\tau_{\max}/T_s$ thus indicates how much ISI[7], in terms of symbols, between the shortest and longest path in a pure LoS MIMO system occurs. As in the previous section, D^2 thus needs to be comparable to the denominator, in this case $2cT_sR$. Since cR is typically a very large quantity, the effect only becomes significant for very high symbol rates and widely-spaced antenna arrays. Some example results for a wide range of practical values can be found in Figure 2.4. The results show that delays in the symbol range occur only for arrays with a largest

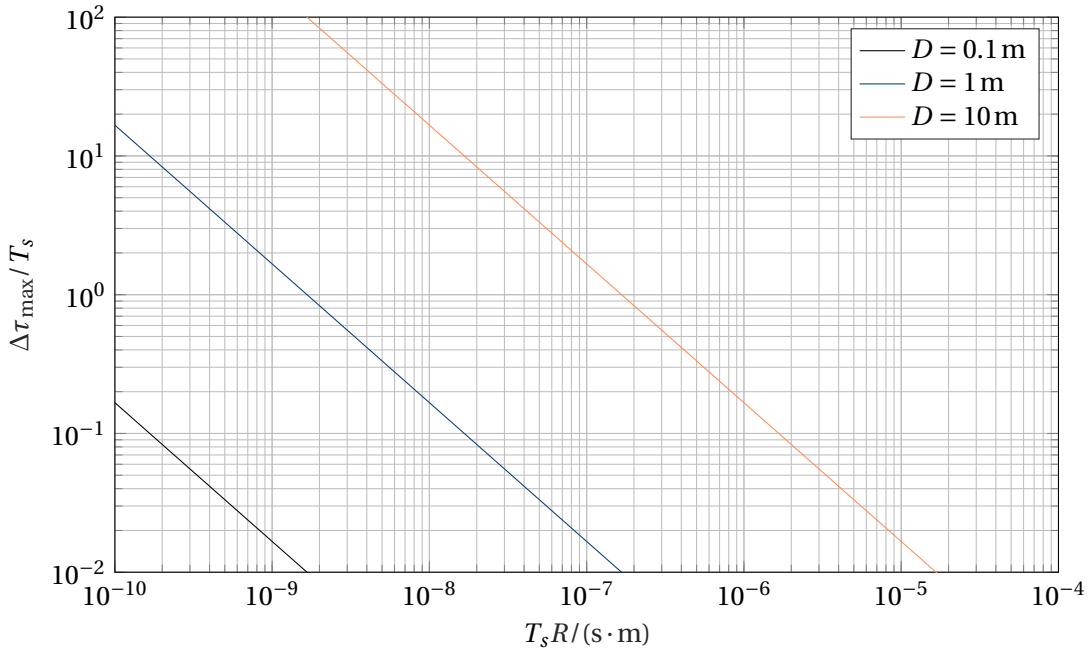


Figure 2.4: Example results for the level of ISI in pure LoS multiple antenna systems, expressed by the spread $\Delta\tau_{\max}/T_s$ in number of symbols, for different array sizes D and symbol-rate-distance products T_sR.

dimension of 1 m or bigger, in this example case.

Rotations due to misalignment between the transmitter and receiver array can further increase the impact of the effect, as observed in [35]. In the rest of this work, this kind of interference will be treated as part of the random frequency-selective channel, where no a-priori information about the delay of the taps is available. Nevertheless, note that this interference has some structure, which could be used for estimating it, as it is determined by the geometric setup of the system.

2.2.3 Antenna Array Design

From section 2.2, it should be clear that the antenna array setup plays an important role for the spatial-multiplexing capabilities of a LoS MIMO system. For the sake of completeness and future reference, some of the design equations derived in the literature will be repeated here. The earliest derivation of an optimal-spacing criterion, supporting full spatial multiplexing with $N = N_1 \times N_2$ transmit and $M = M_1 \times M_2$ receive antennas for a given wavelength-distance product λR, using uniformly-spaced one- and two-dimensional arrangements can be found

[7] A typical assumption is that a channel can be considered frequency non-selective or frequency flat, when $\Delta\tau_{\max}/T_s < 0.1$ [4].

in [36, 37, 38]. For the two-dimensional case on transmitter and receiver side, the smallest[8] optimal spacing in the first dimension of the array is derived by

$$d_{\mathrm{Tx},1}\, d_{\mathrm{Rx},1} = \frac{\lambda R}{\max(N_1, M_1)}, \tag{2.20}$$

and in the second array dimension it is given by

$$d_{\mathrm{Tx},2}\, d_{\mathrm{Rx},2} = \frac{\lambda R}{\max(N_2, M_2)}, \tag{2.21}$$

where $d_{\mathrm{Tx},\cdot}$, $d_{\mathrm{Rx},\cdot}$ is the antenna spacing at the respective side of the link for that array dimension. Thus, the total size of the antenna array in the respective dimension is defined by $D_{\mathrm{Tx},\cdot} = (N_\cdot - 1) \cdot d_{\mathrm{Tx},\cdot}$ and $D_{\mathrm{Rx},\cdot} = (M_\cdot - 1) \cdot d_{\mathrm{Rx},\cdot}$. Furthermore, $N_\cdot$ and $M_\cdot$ denote the number of antennas in the corresponding array dimension. Note that the possible number of spatially-multiplexed streams is limited by SDoF = $\min(M, N)$ and, hence, the most efficient system designs, in terms of array size, occur when $M = N$. Additional antennas between the optimally-spaced ones may still be used, but do not provide additional multiplexing gain in the pure LoS case. For the one-dimensional case on both sides, one can simply use one of the two equations.

The equations show that when the propagation distance R changes, the spacing should change with it in order to preserve the optimal channel. As in the previous section, also rotations of the array have an impact on the optimal spacing. Sensitivity analysis, e.g., with respect to rotations and offsets, as well as modified-spacing equations for such cases, and optimal equations when mixtures between one- and two-dimensional arrays occur can also be found in the mentioned papers [36, 37, 38]. A way of designing more compact one-dimensional arrays was suggested in [32] by using specially designed dielectric media in the propagation paths. Since it may not be feasible to adapt the spacing for every wavelength-distance product in practice, more involved antenna array designs have been proposed in [39, 40, 41], which have a more stable performance over a wider λR range, but do not achieve optimal channel conditions. Three-dimensional array architectures, i.e., arrays that also expand in the transmission direction, have also been considered in the literature [42], but were shown to yield only small benefits in pure LoS scenarios.

2.3 From Discrete Time to Continuous Time and Back

In this section, the sampling operations that are used to convert to and from the continuous-time signals $\bar{x}_n(t)$ and $\bar{y}_m(t)$ in equation (2.9) at transmitter and receiver, are briefly described. A block diagram of the conversion process is given in Figure 2.5. The most important aspect is that a representation of the continuous-time signals in the discrete-time domain is found with as little information loss as possible. This can be done by fulfilling the Nyquist-Shannon

[8] It should be noted that for each wavelength-distance product λR, there are multiple optimal spacings, as can be inferred from section 2.2, and will also be seen in chapter 5.

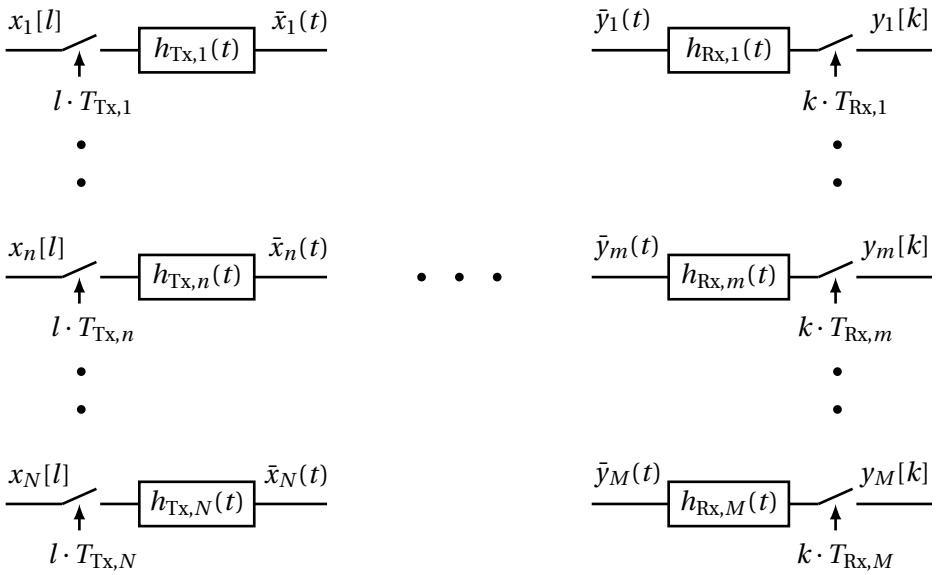


Figure 2.5: Simplified baseband sampling model for a MIMO system with independent and ideal sampling frequencies.

sampling theorem, see [4] for details, whereas more sophisticated methods based on sub-Nyquist sampling are also available, e.g., [43, 44]. Consider the idealized, in terms of memory and amplitude resolution, translation from discrete to continuous time at the transmitter to be given by

$$\bar{x}_n(t) = h_{\mathrm{Tx},n}(t) * \sum_{l=-\infty}^{\infty} x_n[l] \cdot \delta(t - l \cdot T_{\mathrm{Tx},n}) \tag{2.22}$$

$$= \sum_{l=-\infty}^{\infty} x_n[l] \cdot h_{\mathrm{Tx},n}(t - l \cdot T_{\mathrm{Tx},n}), \tag{2.23}$$

where $x_n[l]$ are the samples of the transmit signal at discrete-time index[9] k, $1/T_{\mathrm{Tx},n}$ is the sampling rate of the nth transmitter, and $h_{\mathrm{Tx},n}(t)$ are analog low-pass filters performing the interpolation[10] between the discrete-time samples to a continuous-time signal. Note that as for the carrier frequencies, this notation also models independent LOs for the generation of the sampling rates.

Similarly, the sampled representation $y_m[k]$ of the continuous-time received signal $\bar{y}_m(t)$ is

[9] $x_n[l]$ is not necessarily symbol spaced and $1/T_{\mathrm{Tx},n}$ is not necessarily the symbol rate $1/T_s$. For example, for standard linear modulations using a digital raised-cosine pulse shape with Q samples per symbol, the symbol rate is $1/(T_{\mathrm{Tx},n}Q)$, and is thus also slightly different for each transmitter.

[10] Recall from section 2.1 that these filters actually contain the complete linear-filtering behavior of the transmitter chain. Furthermore, due to this linear filter, $\bar{x}_n(l \cdot T_{\mathrm{Tx},n}) \neq x_n(l \cdot T_{\mathrm{Tx},n})$.

given by

$$\sum_{k=-\infty}^{\infty} y_m[k] \cdot \delta(t - k \cdot T_{\mathrm{Rx},m}) = \left(\bar{y}_m(t) * h_{\mathrm{Rx},m}(t)\right) \cdot \sum_{k=-\infty}^{\infty} \delta(t - k \cdot T_{\mathrm{Rx},m}) \tag{2.24}$$

$$= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \bar{y}_m([k-l] \cdot T_{\mathrm{Rx},m}) \cdot h_{\mathrm{Rx},m}(l \cdot T_{\mathrm{Rx},m}) \cdot \delta(t - k \cdot T_{\mathrm{Rx},m}), \tag{2.25}$$

where the analog filters $h_{\mathrm{Rx},m}(t)$ contain low-pass filters, which are suitably chosen in order to avoid aliasing, and $1/T_{\mathrm{Rx},m}$ is the sampling rate of the mth receiver. Generally, as for the Tx side, this filter contains all linear-filtering effects of a particular receiver chain, e.g., the combined characteristic of the anti-aliasing filters and the low-pass filters used in equation (2.7). Thus, $h_{\mathrm{Rx},m}(t)$ is assumed to contain the complete linear analog filter characteristic of the mth receiver chain. The kth received sample is then given by

$$y_m[k] = \sum_{l=-\infty}^{\infty} \bar{y}_m([k-l] \cdot T_{\mathrm{Rx},m}) \cdot h_{\mathrm{Rx},m}(l \cdot T_{\mathrm{Rx},m}) \tag{2.26}$$

$$= \bar{y}_m(k \cdot T_{\mathrm{Rx},m}) \star h_{\mathrm{Rx},m}(k \cdot T_{\mathrm{Rx},m}), \tag{2.27}$$

where $\star$ indicates discrete convolution. This equation, more generally, can be used to specify the complete linear chain in discrete notation. Using a slight abuse of notation, and additionally (2.1), (2.4), and (2.7), yields

$$y_m[k] = e^{-j(2\pi\Delta f_{\mathrm{Rx},m} T_{\mathrm{Rx},m} k + \phi_{\mathrm{Rx},m})} \cdot h_{\mathrm{Rx},m}(kT_{\mathrm{Rx},m}) \star \left(\sum_{n=1}^{N} h_{mn}(kT_{\mathrm{Rx},m}) \star \bar{x}_n(kT_{\mathrm{Rx},m}) \cdot e^{j(2\pi\Delta f_{\mathrm{Tx},n} T_{\mathrm{Rx},n} k + \phi_{\mathrm{Tx},n})} + n_m(kT_{\mathrm{Rx},m}) \right) \tag{2.28}$$

$$= h_{\mathrm{Rx},m}[k] \star \sum_{n=1}^{N} h_{mn}(kT_{\mathrm{Rx},m}) \star \bar{x}_n(kT_{\mathrm{Rx},m}) \cdot e^{j(2\pi\Delta f_{mn} T_{\mathrm{Rx},m} k + \Delta\phi_{mn})} + h_{\mathrm{Rx},m}[k] \star n_m[k] \tag{2.29}$$

$$= h_{\mathrm{Rx},m}[k] \star \sum_{n=1}^{N} e^{j(2\pi\Delta f_{mn} T_{\mathrm{Rx},m} k + \Delta\phi_{mn})} \cdot h_{mn}[k] \star \sum_{l=-\infty}^{\infty} x_n[l] h_{\mathrm{Tx},n}\left(kT_{\mathrm{Rx},m} - lT_{\mathrm{Tx},n}\right) + h_{\mathrm{Rx},m}[k] \star n_m[k] \tag{2.30}$$

$$= \sum_{n=1}^{N} e^{j(2\pi\Delta f_{mn} T_{\mathrm{Rx},m} k + \Delta\phi_{mn})} \cdot \bar{h}_{mn}[k] \star \sum_{l=-\infty}^{\infty} x_n[l] h_{\mathrm{Tx},n}\left[k - l\frac{T_{\mathrm{Tx},n}}{T_{\mathrm{Rx},m}}\right] + h_{\mathrm{Rx},m}[k] \star n_m[k], \tag{2.31}$$

where $\Delta f_{\mathrm{Rx},m}$, $\Delta f_{\mathrm{Tx},n}$ are the deviations from the nominal desired carrier frequency f_{nom}[11] of the oscillators on the transmitter and receiver side, i.e., $\Delta f_{\mathrm{Rx},m} = f_{\mathrm{Rx},m} - f_{\mathrm{nom}}$ and $\Delta f_{\mathrm{Tx},n} = f_{\mathrm{Tx},n} - f_{\mathrm{nom}}$. Furthermore, we use $\bar{h}_{mn}[k] = h_{\mathrm{Rx},m}(kT_{\mathrm{Rx},m}) \star h_{mn}(kT_{\mathrm{Rx},m}) = h_{\mathrm{Rx},m}[k] \star h_{mn}[k]$, $h_{mn}(t) = \sum_{l=0}^{L_{mn}-1} a_{mn,l} \cdot e^{-j2\pi f_{\mathrm{Tx},n} \cdot \tau_{mn,l}} \cdot \mathrm{sinc}\left(\pi \frac{t-\tau_{mn,l}}{T_{\mathrm{nom,Rx}}}\right)$, and the circularity of the noise, i.e., the stochastic invariance of the complex noise process to complex phase rotations. Equation (2.31) is the general linear discrete-time baseband frequency-selective input/output model for MIMO systems, impaired by independent sampling and carrier frequency offsets for each transmit and receive chain. Its properties will be described later on, for now the focus will just be on the impact of the sampling processes.

The received discrete-time baseband signal, consisting of the multiple superimposed transmitted signals, exhibits the normalized, w.r.t. the sampling rate, angular rotations, or frequency shifts, of $\Delta f_{mn} T_{\mathrm{Rx},m}$, which will in this work be estimated and compensated by suitable digital processing algorithms. Likewise, the ISI generated from the sampling phase shifts due to the term $T_{\mathrm{Tx},n}/T_{\mathrm{Rx},m}$ can be be estimated by capturing an oversampled version of the received signal, and compensated digitally with some form of interpolation [4, 25].

The previous discussion is somewhat simplified, since $T_{\mathrm{Rx},m}$ and $T_{\mathrm{Tx},n}$ are, in practice, not constant over time, i.e., there is no perfectly periodic uniform sampling, but they slowly vary around some mean value. Due to this, and the fact that the fraction between the two is practically never an integer, it has been noted in the literature that in a continuous transmission there will eventually a sample over- or underflow occur, see also appendix A, which needs to be appropriately dealt with. Additionally, especially in distributed systems it can happen that the sampling points in time, i.e., the time instants where a new sample is generated, are not synchronized between different front ends. Then, an additional time shift between the different front-ends is observed, which has a different impact depending on whether it appears on the transmitter or receiver side. Even more specifically, it can happen in I/Q sampling systems that the in- and quadrature phase sampling instants are not perfectly aligned, generating interference between the two branches. While the former effect can be treated as different delays in the complete channel impulse response, the latter can be treated as part of an I/Q imbalance compensation scheme, to be discussed later on. Since those effects depend on the oscillator setup that is used for generating the sampling frequencies, they will be discussed in more detail at a later stage. Other non-ideal effects of the sampling circuitry, e.g., sinc shaping of the output spectrum due to finite time duration of the sampling pulses, can be assumed as part of the analog filters $h_{\mathrm{Tx},n}(t)$ and $\bar{h}_{\mathrm{Rx},m}(t)$ at the transmitter and receiver, respectively. Whether or not sufficient statistics are generated by the sampling processes in general, i.e., $y_m[k]$ contains all information that $\bar{y}_m(t)$ contains, depends on the roll-off of the analog filter $\bar{h}_{\mathrm{Rx},m}(t)$ and the amount of oversampling used in the system [4]. Finally, as is common in the sampling literature, the effect of finite amplitude resolution, or quantization,

[11] Throughout this work, we try to stick to the notation that f_{nom} values are related to the carrier frequency, and T_{nom} values are related to the sampling interval. This leads to the dilemma that even though f_{nom} and T_{nom} seem synonymous, they are fundamentally different as one relates to the carrier process, while the other relates to the sampling process.

has not been discussed. It is thus inherently assumed that the effects due to finite resolution are negligible or, in other words, moderate resolution digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) are used. In general, however, taking the influence properly into account requires a fundamentally different system design approach [45].

2.4 A Primer on Oscillators and Frequency Accuracy

Oscillators are the most crucial component when it comes to synchronization in communications systems. In the previous section, it was seen that oscillators are necessary for generating the carrier frequency for up- and downconversion, and for generating the sample timing reference at transmitter and receiver, respectively. Ideally, an oscillator only creates a single spectral[12] line at a desired nominal frequency, e.g., the left plot in Figure 2.6 or the complex exponential seen in equation (2.7). If such oscillators were realizable, no synchronization would be needed. However, due to temperature variations, mechanical variations, manufacturing mismatches [12, 47], and other small scale effects, e.g., variations in the power supply, two independent oscillators will never have exactly the same frequency. A lot of effort can be spent on stabilizing the frequency by, e.g., controlling the temperature or using a laser reference [2]. In most communications systems, however, low-complexity solutions are preferred in order to reduce size and cost. Then, some form of synchronization that can cope with the frequency difference and its variations, preferably through digital signal processing techniques, is always required.

In section 2.1, ideal oscillators have been assumed, which generate only a single frequency with a fixed phase, albeit not the same for the different front ends. For example, the offsets from the nominal carrier frequency on the receiver side were characterized with $e^{-j(2\pi\Delta f_{\mathrm{Rx},m} T_{\mathrm{Rx},m} k + \phi_{\mathrm{Rx},m})}$, where $\Delta f_{\mathrm{Rx},m} = f_{\mathrm{Rx},m} - f_{\mathrm{nom}}$. A general description of oscillator behavior can be obtained by considering the progression of its phase over (discrete) time, i.e., the phase process[13] $\phi[k]$ [51]. In this work, phase process will mean the difference compared to the nominal desired frequency, i.e., the part that needs to be estimated and compensated in order to achieve synchronization. In the previous sections, the phase process was simply a linearly-increasing function with a fixed initial offset, and where the slope depends on the frequency offset. For the previous example this would be denoted as $e^{-j\phi_{\mathrm{Rx},m}[k]}$ with $\phi_{\mathrm{Rx},m}[k] = 2\pi\Delta f_{\mathrm{Rx},m} T_{\mathrm{Rx},m} k + \phi_{\mathrm{Rx},m}[0]$.

In the rest of this work, the following phase process model will be considered for both oscillators used for carrier frequency generation, and oscillators used to derive the sampling instants.

[12] Likewise, oscillators can also be evaluated in terms of their time domain stability using, for example, the Allan variance. Since phase noise data is often more readily available for system components, we will not treat time domain stability measures further. However, it should be noted that one can convert between time and frequency domain stability measures [46].

[13] Physically, oscillators also experience variations in the amplitude due to noise, which also affect the system [12, 47, 48, 49], but are typically assumed less severe than the phase variations. On the other hand, [50] argues that especially for mmWave systems, the amplitude noise should be taken into account.

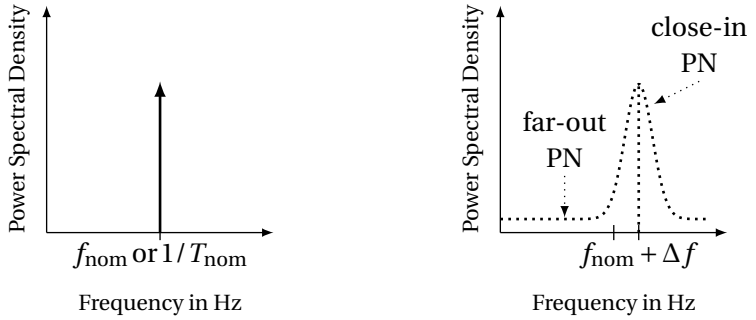


Figure 2.6: Qualitative power spectral density of an ideal and a real oscillator.

It is a variant of the model derived in [49]. For the carrier frequency, the corresponding carrier phase process is given by

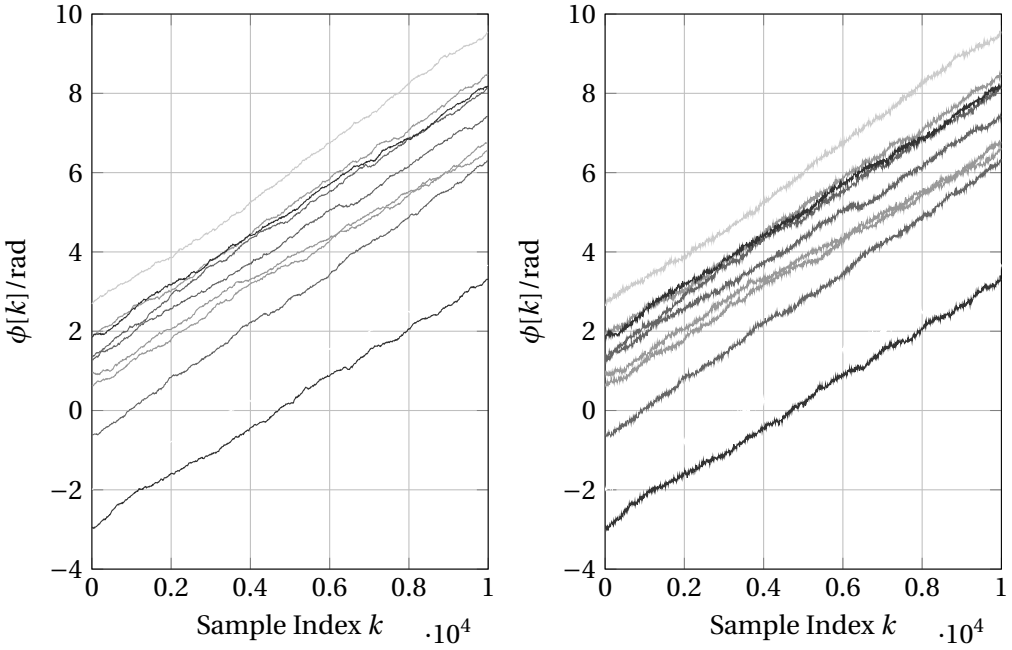
$$\phi[k] = \phi_{\text{w}}[k] + \varphi_{\text{n}}[k] \tag{2.32}$$

$$= \underbrace{\phi_{\text{w}}[k-1] + \varphi_{\text{w}}[k]}_{\text{Wiener process with drift}} + \underbrace{\varphi_{\text{n}}[k]}_{\text{white noise process}} , \tag{2.33}$$

where $\phi_{\text{w}}[0]$ is the initial phase state of the oscillator, assumed to be distributed with $\phi_{\text{w}}[0] \sim \mathcal{U}(-\pi, \pi)$. Furthermore, all random processes are assumed independent. The random process $\varphi_{\text{w}}[k]$ expresses a fixed mean drift and a random variation per sample. It will be assumed normally distributed with $\varphi_{\text{w}}[k] \sim \mathcal{N}(\mu_{\varphi_{\text{w}}}, \sigma^2_{\varphi_{\text{w}}})$, where $\mu_{\varphi_{\text{w}}} = 2\pi\Delta f\, T_{\text{nom}}$ is the mean phase change from sample to sample, related to the frequency offset with T_{nom} being the sampling interval of the discrete process, and where $\sigma^2_{\varphi_{\text{w}}}$ specifies the level of random variation per sample, related to the close-in phase noise (PN) of the oscillator [49]. For example, for 60 GHz oscillators $\Delta f = 100\,\text{kHz}$ (1.7 ppm) [52, 53] and one sided 3 dB PN bandwidths of 1 kHz [54, 55] are reasonable values. These numbers yield with $T_{\text{nom}} = 1\,\text{ns}$, $\mu_{\varphi_{\text{w}}} = 2\pi \cdot 10^{-4}\,\text{rad/ns}$ and $\sigma^2_{\varphi_{\text{w}}} = 4\pi \cdot 10^{-6}\,\text{rad}^2/\text{ns}^2$, respectively. Further, with these definitions, $\phi_{\text{w}}[k]$ can be seen as a Wiener process with drift. The white noise region far from the carrier, also called far-out PN, is modeled with the random process $\varphi_{\text{n}}[k] \sim \mathcal{N}(0, \sigma^2_{\varphi_{\text{n}}})$, where $\sigma^2_{\varphi_{\text{n}}} = \frac{K_0}{T_{\text{nom}}}$ with K_0 being the white noise floor level of the oscillator, which can be read directly from the measured PN spectrum. For example, for a 60 GHz LO a typical value is $K_0 = 10^{\frac{-120\,\text{dBc/Hz}}{10}}$ [9, 12, 50], yielding with $T_{\text{nom}} = 1\,\text{ns}$ a variance of $\sigma^2_{\varphi_{\text{n}}} = 10^{-3}\,\text{rad}^2/\text{ns}^2$. The conversion between datasheet values and model parameters will be further explained in appendix A, see also [49]. Example realizations of the phase process model with the just mentioned parameters can be found in Figure 2.7.

The motivation of the model[14] in (2.33) is as follows. The drifting of the oscillator from the nominal frequency f_{nom} and around the mean frequency difference Δf is modeled through the Wiener process $\phi_{\text{w}}[k]$, while fast variations are modeled through the white noise process

[14] Note that this model covers the steady-state behavior of an oscillator. It does not accurately model erratic behavior, such as aperiodic mechanical vibrations.



(a) Wiener process with drift part

(b) Complete process

Figure 2.7: Ten example realizations of the phase process model given in (2.33), with parameters relevant for carrier phase processes in mmWave systems: $\phi_{\mathrm{w}}[0] \sim \mathcal{U}(-\pi,\pi)$, $\mu_{\varphi_{\mathrm{w}}} = 2\pi \cdot 10^{-4}\,\mathrm{rad}$, $\sigma^2_{\varphi_{\mathrm{w}}} = 10^{-5}\,\mathrm{rad}^2$, and $\sigma^2_{\varphi_{\mathrm{n}}} = 10^{-3}\,\mathrm{rad}^2$.

$\varphi_{\mathrm{n}}[k]$. By changing the parameters $\mu_{\varphi_{\mathrm{w}}}$, $\sigma^2_{\varphi_{\mathrm{w}}}$, and $\sigma^2_{\varphi_{\mathrm{n}}}$, this model can cover a wide range of oscillator behaviors. For example, it has been noted in the literature [7, 49, 50] that for high symbol rate systems, e.g., mmWave communications, the far-out PN is dominant, and can even be the performance limiting factor. This is inherently included in the model through the scaling of the variance $\sigma^2_{\varphi_{\mathrm{n}}}$ with $1/T_{\mathrm{nom}}$, and can also be seen from the example values calculated in the previous paragraph. Note, however, that in the present form no correlation across time, except for the mean part $\mu_{\varphi_{\mathrm{w}}}$, among consecutive samples of the phase process is considered in the model. For certain system setups, e.g., when using a phase-locked loop (PLL), the samples will be correlated in a more complicated way and the model can be modified to take that into account [54]. In the remainder of this work, this type of correlation across time will not be considered, as it is assumed that the drifting of the phase process model sufficiently describes the behavior of most oscillators. Nevertheless, better results might be achieved by taking this circuit-dependent knowledge about the correlation w.r.t. time into account. The goal will be to find suitable estimation and compensation algorithms for multiple of these slowly varying phase processes in MIMO systems.

For the sampling processes, the positive zero crossings or, similarly, edges of a periodic signal determine the sample instants [51]. Consider, for example, the ideal sine wave used at a receiver with $\sin\left(2\pi \frac{t}{T_{\mathrm{nom,Rx}}}\right)$, where a sample of the continuous-time signal is taken with ideal equal spacing every nominal period $T_{\mathrm{nom,Rx}} = T_{\mathrm{nom}}$, see Figure 2.8. In practice, as for the

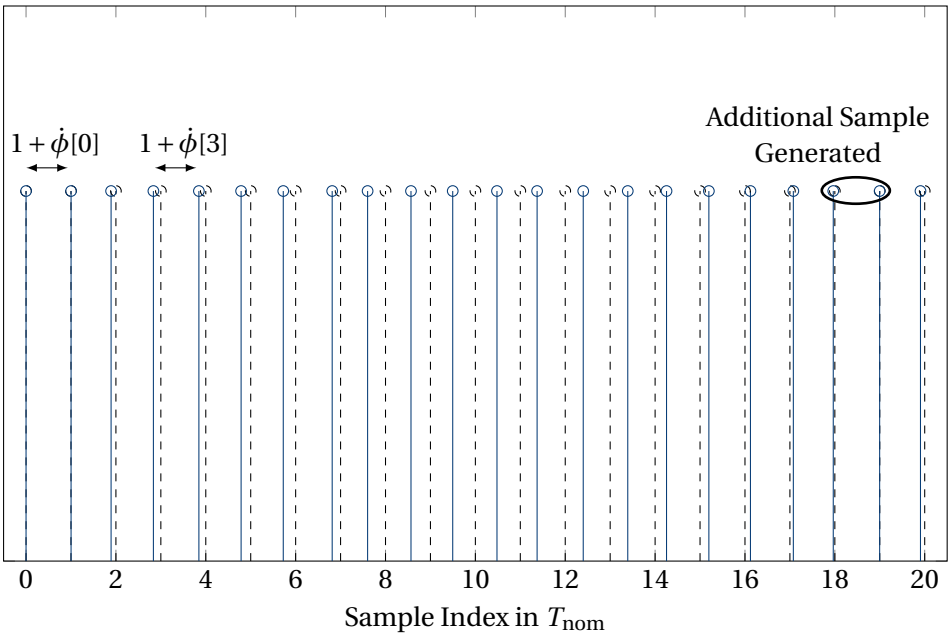


Figure 2.8: Qualitative example of the generation of the sampling instants from a time reference using: an ideal oscillator with T_{nom} (dashed), and a practical oscillator with a varying phase process (blue). The difference between the two contains a fixed mean offset and random variations, e.g., due to jitter, see (2.33). In this example, the practical reference is on average faster than the ideal nominal one, and the initial phase difference is zero, i.e., $\dot{\phi}_{\text{w}}[0] = 0$. The additional sample problem is further explained in appendix A.3.

carrier frequency, the period varies over time, i.e., has a fixed offset from the nominal value and a drift around it, which can be described by the model in (2.33), see the example realization in Figure 2.8. Continuing with the example this means $\sin\left(2\pi \frac{t}{T_{\text{Rx},m}(t)}\right)$, with $T_{\text{Rx},m}(t) = T_{\text{nom,Rx}} + \Delta T_{\text{Rx},m}(t)$, where $\Delta T_{\text{Rx},m}(t)$ is the continuous-time version of the sampling phase process offset of the mth front end (FE) describing the offset. The conversion to a discrete-time process is done by considering that only the difference with respect to the nominal value T_{nom} is important, refer to Figure 2.8. Thus, for the example, the relative period of two neighboring samples at discrete time k is given by $\frac{T_{\text{Rx},m}(kT_{\text{nom,Rx}})}{T_{\text{nom,Rx}}} = 1 + \dot{\phi}_{\text{Rx},m}[k]$. We can then model model the sampling phase processes w.r.t. their nominal period in a similar way as the carrier phase processes with $\dot{\phi}[k] = \dot{\phi}_{\text{w}}[k] + \dot{\phi}_{\text{n}}[k]$, which describes an offset from the nominal period or frequency, and a jitter process[15] of the corresponding DAC or ADC.

From the examples discussed above, it should be clear that for both cases, carrier and sampling frequency, the deviation from the nominal value is what is disturbing the system, and it can

[15] Jitter [51] in sampling devices has two components [56]. First, there is jitter due to the sampling circuitry, called aperture jitter, which can be assumed Gaussian and modeled by the white noise process with $\sigma^2_{\dot{\varphi}_{\text{n}}}$. Second, there is jitter due to the variation of the sampling frequency that determines the sampling instants, i.e., the variation due to phase noise, called clock jitter. The variance of the Wiener process $\sigma^2_{\dot{\varphi}_{\text{w}}}$ can be related to square root of the root-mean-square (RMS) cycle jitter. Both parameters are readily found in the data sheets of sampling devices and clock generators, respectively.

be described by the discrete phase processes $\phi[k]$ in rad/sample and $\dot{\phi}[k]$ in samples/sample, respectively. The two timing impairments are known as CFO, describing the differences in carrier frequencies, and SFO, describing the differences in sample timing. Furthermore, due to the one being defined for frequency and the other for time, it holds that the carrier frequency is higher than its nominal value, when $\mathrm{E}\left[\phi[k]\right] - \phi[0] > 0$, while the sampling frequency is higher than its nominal value, when $\mathrm{E}\left[\dot{\phi}[k]\right] - \dot{\phi}[0] < 0$, see also the two examples in the figures. Typical parameter ranges for both of them in mmWave systems can be found in Table 2.1. The units of the parameters characterizing the phase processes are to be understood per sampling interval, even though later on in the work this will sometimes be omitted for brevity, when noting the values, see also appendix A.1. For some general additional details on the topic, refer also to appendix A.

Table 2.1: Typical parameters found in the literature and data sheets of components for carrier and sampling frequency generation in mmWave systems, i.e., carriers in the tens of GHz range and bandwidths in the GHz range. Normalized values for the model of (2.33) are given per discrete sample with an assumed sample period of $T_{\mathrm{nom}} = 1\,\mathrm{ns}$, and an assumed carrier frequency of $f_{\mathrm{nom}} = 60\,\mathrm{GHz}$. For a description of how the conversion to the model parameters is done, see appendix A.

Parameter	Carrier Frequency	Sampling Frequency	Parameter
Accuracy in ppm	$\pm 1 \sim \pm 200$	$\pm 1 \sim \pm 200$	Accuracy in ppm
PN @ 1 MHz in dBc/Hz	$-125 \sim -85$	$10 \sim 400$	RMS Jitter in fs
far-out PN in dBc/Hz	$-140 \sim -120$	$50 \sim 500$	Apert. Jitt. in fs
$\mu_{\varphi_\mathrm{w}} / \frac{\mathrm{rad}}{\mathrm{ns}}$	$\pm 3.8 \cdot 10^{-4} \sim \pm 7.5 \cdot 10^{-2}$	$\pm 10^{-6} \sim \pm 2 \cdot 10^{-4}$	$\mu_{\dot{\varphi}_\mathrm{w}} / \frac{\mathrm{samples}}{\mathrm{ns}}$
$\sigma^2_{\varphi_\mathrm{w}} / \frac{\mathrm{rad}^2}{\mathrm{ns}^2}$	$1.25 \cdot 10^{-8} \sim 1.25 \cdot 10^{-4}$	$10^{-10} \sim 1.6 \cdot 10^{-7}$	$\sigma^2_{\dot{\varphi}_\mathrm{w}} / \frac{\mathrm{samples}^2}{\mathrm{ns}^2}$
$\sigma^2_{\varphi_\mathrm{n}} / \frac{\mathrm{rad}^2}{\mathrm{ns}^2}$	$10^{-5} \sim 10^{-3}$	$2.5 \cdot 10^{-9} \sim 2.5 \cdot 10^{-7}$	$\sigma^2_{\dot{\varphi}_\mathrm{n}} / \frac{\mathrm{samples}^2}{\mathrm{ns}^2}$
$\phi_\mathrm{w}[0]/\mathrm{rad}$	$\mathcal{U}(-\pi, \pi)$	$\mathcal{U}(-0.5, 0.5)$	$\dot{\phi}_\mathrm{w}[0]/\mathrm{samples}$

The final topic in this section will deal with deriving a high frequency oscillation from a low reference frequency, usable for both carrier and sampling frequency generation, see Figure 2.9. This is a particularly important technique and widely used, to have flexible frequency generation [47], and to synchronize multiple distributed systems. Application examples are the synchronization of clocks in measurement equipment using a 10 MHz reference [57], or using a 10 MHz reference for the synchronization of 100 front ends in a massive MIMO base station implementation [22]. In many of these cases, distributing the desired frequency directly is impractical due to significant power losses, or requires very demanding circuit design, e.g., splitting a 60 GHz sinusoid for upconversion to several widely-separated mixers. A better option is to have a shared reference that is multiplied by a factor, often using a PLL within a frequency synthesizer [47] for each device, in order to generate the actual desired frequency. The close-in PN of such a system is determined by the phase noise of the reference, the multiplication factor that is necessary to generate the desired frequency, and the parameters of the control loop [58]. In particular, the phase deviation is increased by the multiplication factor, such that there is a trade-off between simplicity of reference distribution, and phase noise

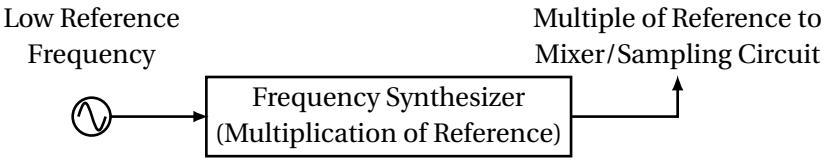


Figure 2.9: General model of a frequency synthesizer that generates a high frequency oscillation from a low frequency reference signal.

performance[16]. The far-out phase noise, on the other hand, is determined by the LO in each synthesizer and can be assumed independent for each device. To summarize, the accuracy of the generated frequency and the close-in phase noise is determined by the reference and the loop parameters, whereas the far-out PN is determined by the synthesizer implementation [47]. In general, the phase processes of such systems can be modeled by (2.33), but they experience certain correlations between the Wiener processes of the different devices, since they are determined by the shared reference. In this work, such system configurations will not be treated separately, but assumed to be equivalent to cases where the desired frequency is shared directly. Finally, consider the initial phases $\phi_{\mathrm{w}}[0]$ and $\dot{\phi}_{\mathrm{w}}[0]$ of the generated phase processes for different synthesizers with a common reference. Even though a shared reference is used, the output starting phases of the synthesizers cannot necessarily be assumed equal [57], as they depend on the distribution of the reference frequency, and the implementation of the frequency generation, i.e., where and how frequency dividers are used. Thus, for the rest of this work it will be assumed that the initial phases of the synthesizers are random and uniformly distributed over the complete phase range, e.g., $\phi_{\mathrm{w}}[0] \sim \mathcal{U}(-\pi,\pi)$, although some synthesizers do allow phase synchronization.

2.4.1 Oscillator Configurations for MIMO Setups

Depending on the system settings and requirements, a variety of synchronization configurations, i.e., distribution of sampling and carrier frequency, can be applicable for MIMO systems. There are four basic oscillator setups that will be considered, namely,

1. Shared between transmitter and receiver side, and all of their front ends,
2. Shared between all front ends on transmitter and receiver side, respectively,
3. Shared on one side, either Tx or Rx, and independent on the other side, either Rx or Tx,
4. Independent on both sides.

These setups are applicable separately to the carrier and sampling frequency distribution, such that different setups can be used for each of the two depending on system prerequisites. For

[16] The lower the reference, the easier it is to distribute between different devices [57]. However, the lower the reference, the higher the required multiplication factor and, thus, the amount of phase noise. Another benefit of higher reference frequencies is improved long-term phase stability.

the remainder of this work, we mostly consider cases where carrier and sampling frequency distribution has the same setup. Furthermore, subgroups composed of a couple of FEs could share clocks, which will not be considered, but could be viewed as one of the above setups with appropriate grouping. The impact of the different setups on estimation and equalization will become apparent later in this work. A system diagram of the different setups is given in Figure 2.10. Please note that, as described in the previous section, it is not necessary to directly share the sampling or carrier frequency for setups 1, 2, 3. Instead, a reference frequency can be shared that is then multiplied in a frequency synthesizer for each transceiver.

An example for each of the setups will be discussed next. Setup 1 occurs, when some additional resource or part of the spectrum, like in early amplitude modulation broadcasting, is allocated in order to transmit a reference clock or the Tx-shared frequency directly. It is then extracted in the receivers, and used in all front ends for downconversion and sampling. This setup generally requires a more complicated FE design and is spectrally less efficient, due to the spectrum part allocated to transmitting the reference. The benefit is that it should incur no frequency difference between transmitter and receiver. Setup 2 is, e.g., a common choice for MIMO systems where the antennas are closely co-located, as in single-user digital-beamforming systems [58, 59], or multi-user systems with multiple antennas per user and time-division or frequency-division multiple access. As in the previous setup, routing of the references to all transceivers needs to be handled on each side of the link, but no resources are spent for transmitting the frequency, making it more spectrally efficient than setup 1. Such a system incurs one frequency difference, due to the one independent oscillator on the transmitter and receiver side, respectively.

Setup 3 is, for example, relevant for the uplink of multi-user MIMO systems with spatial-division access. In such a system, the antennas and front ends on the basestation are co-located and can use a shared reference, although some effort has to be spend in order to assure synchronism [21]. Each user, on the other hand, has an independent LO. Thus, this type of system is influenced by as many frequency differences as there are independent oscillators on one side of the link. Finally, setup 4 represents, e.g., the case of a distributed MIMO system [18, 60], where multiple basestations cooperate in order to achieve a performance gain. Note that this setup is the most general and includes all of the previous ones, but also requires the most signal processing effort to achieve synchronization, as will become apparent later. It incurs all combinations of frequency differences that occur due to the independent oscillators on the transmitter and receiver side. The general model described in the next sections will, thus, be based on setup 4.

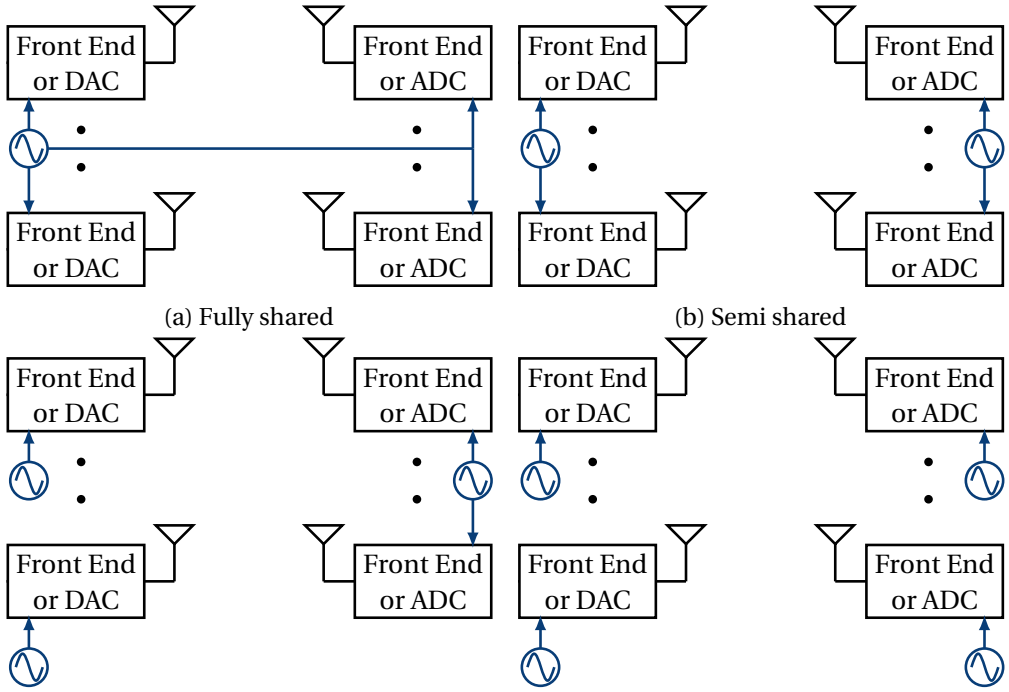


Figure 2.10: The four basic MIMO synchronization configurations, which are applicable separately to carrier and sampling frequency distribution. Setup 3 also includes the case, where sharing happens on the other side of the link (not shown here).

2.5 Discrete-Time Baseband MIMO Model with Realistic Oscillators

Restating the sampling equations (2.23), (2.24) to include the defined sampling phase processes, i.e., $\dot{\phi}_{\mathrm{Tx},n}[k]$ and $\dot{\phi}_{\mathrm{Rx},m}[k]$, yields

$$\bar{x}_n(t) = \sum_{k=-\infty}^{\infty} x_n[k] \cdot h_{\mathrm{Tx},n}\left(t - (k + \dot{\phi}_{\mathrm{Tx},n}[k]) \cdot T_{\mathrm{nom,Tx}}\right) \tag{2.34}$$

and

$$\begin{aligned} &\sum_{k=-\infty}^{\infty} y_m[k] \cdot \delta\left(t - (k + \dot{\phi}_{\mathrm{Rx},m}[k]) \cdot T_{\mathrm{nom,Rx}}\right) \\ &\qquad = \left(\bar{y}_m(t) * h_{\mathrm{Rx},m}(t)\right) \cdot \sum_{k=-\infty}^{\infty} \delta\left(t - (k + \dot{\phi}_{\mathrm{Rx},m}[k]) \cdot T_{\mathrm{nom,Rx}}\right), \text{ with} \end{aligned} \tag{2.35}$$

$$y_m[k] = \bar{y}_m\left((k + \dot{\phi}_{\mathrm{Rx},m}[k]) \cdot T_{\mathrm{nom,Rx}}\right) \star h_{\mathrm{Rx},m}\left((k + \dot{\phi}_{\mathrm{Rx},m}[k]) \cdot T_{\mathrm{nom,Rx}}\right). \tag{2.36}$$

Combining these equations and including the effects due to the channel, as was done for

(2.31), gives

$$\begin{aligned} y_m[k] = e^{-j\phi_{\mathrm{Rx},m}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right)} \cdot \Big(& \sum_{n=1}^{N} h_{mn}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \\ & \star \bar{x}_n\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \cdot e^{j\phi_{\mathrm{Tx},n}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right)} \\ & + n_m\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right)\Big) \star h_{\mathrm{Rx},m}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \end{aligned} \tag{2.37}$$

$$\begin{aligned} \approx e^{-j\phi_{\mathrm{Rx},m}(k\cdot T_{\mathrm{nom,Rx}})} \cdot \Big(& \sum_{n=1}^{N} h_{mn}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \\ & \star \bar{x}_n\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \cdot e^{j\phi_{\mathrm{Tx},n}(k\cdot T_{\mathrm{nom,Rx}})} \\ & + n_m(k\cdot T_{\mathrm{nom,Rx}})\Big) \star h_{\mathrm{Rx},m}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \end{aligned} \tag{2.38}$$

$$\begin{aligned} = h_{\mathrm{Rx},m}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \star & \sum_{n=1}^{N} e^{j\Delta\phi_{mn}[k]} \cdot h_{mn}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ & \star \sum_{l=-\infty}^{\infty} x_n[l] h_{\mathrm{Tx},n}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}} - (l+\dot{\phi}_{\mathrm{Tx},n}[l])\cdot T_{\mathrm{nom,Tx}}\right) \\ & + h_{\mathrm{Rx},m}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \star n_m[k] \end{aligned} \tag{2.39}$$

$$\begin{aligned} = h_{\mathrm{Rx},m}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \star & \sum_{n=1}^{N} e^{j\Delta\phi_{mn}[k]} \cdot h_{mn}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ & \star \sum_{l=-\infty}^{\infty} x_n[l] h_{\mathrm{Tx},n}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k] - (l+\dot{\phi}_{\mathrm{Tx},n}[l])\cdot \frac{T_{\mathrm{nom,Tx}}}{T_{\mathrm{nom,Rx}}}\right] \\ & + h_{\mathrm{Rx},m}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \star n_m[k]. \end{aligned} \tag{2.40}$$

The term $\Delta\phi_{mn}[k] = \phi_{\mathrm{Tx},n}[k] - \phi_{\mathrm{Rx},m}[k]$ denotes the difference between the transmitter and receiver carrier phase processes. An approximation has been made by assuming that $\phi_{\mathrm{Rx},m}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \approx \phi_{\mathrm{Rx},m}(k\cdot T_{\mathrm{nom,Rx}})$, and also $\phi_{\mathrm{Tx},n}\left((k+\dot{\phi}_{\mathrm{Rx},m}[k])\cdot T_{\mathrm{nom,Rx}}\right) \approx \phi_{\mathrm{Tx},n}(k\cdot T_{\mathrm{nom,Rx}})$. From a signal processing point of view this has no impact on further performance, since the errors that occur due to non-ideal sampling can be treated as part of the random variation[17] of the carrier phase processes $\phi_{\mathrm{Rx},m}[k]$ and $\phi_{\mathrm{Tx},m}[k]$. Furthermore, the noise process $n_m[k]$ is assumed to retain its white property, regardless if it is sampled at the nominal rate or the slightly impaired one, while the filtered noise process $h_{\mathrm{Rx},m}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \star n_m[k]$ will generally be colored.

Equation (2.40) is the general, sampling and carrier frequency offset impaired, discrete linear system model for MIMO links with realistic LOs. The goal is to get back to the transmitted samples (or symbols) $x_n[k]$, meaning that the impact of all of the other terms should be

[17] Consider a small example, where the carrier phase process is linear, i.e., linearly drifting, time-continuous, has $\mu_{\varphi_\mathrm{w}} = 10^{-2}\,\frac{\mathrm{rad}}{\mathrm{ns}}$, and is observed over a time interval of 20 ns. The total phase variation in that time interval is thus 0.2 rad. When sampled ideally with $T_{\mathrm{nom,Rx}} = 1\,\mathrm{ns}$ this gives of course 10^{-2} rad/sample. On the other hand, when sampling with $T_{\mathrm{nom,Rx}} + \Delta T_{\mathrm{Rx},m}(t) = 1.5\,\mathrm{ns}$, to exaggerate, yields a carrier phase variation of $1.5\cdot 10^{-2}$ rad/sample. Thus, the model is still valid, but the values of the carrier phase parameters change, when the sampling is not ideal.

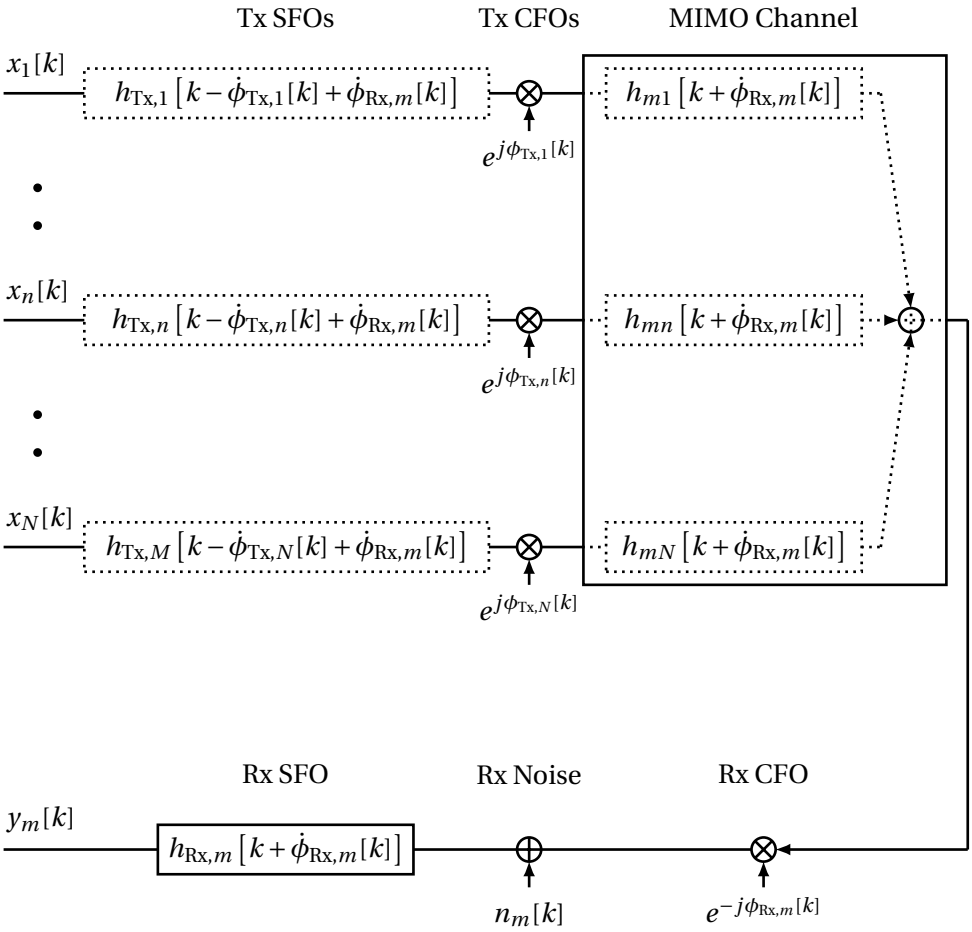


Figure 2.11: Discrete-time symbol-spaced baseband signal model of the mth received signal in a linear MIMO system with multiple timing impairments. The CFOs effects are expressed as parallel rotating complex exponentials with independent phase processes on the receiver and transmitter, respectively. The SFOs influences are conveyed through time-variant sampling on the Tx and Rx side, based on independent sampling phase processes. The parallel filters on the transmitter and receiver side contain the analog linear-filtering characteristics of each front end. They include, for example, interpolation, image rejection, and anti-aliasing filters.

reduced as much as possible. Each of the sampled received streams $y_m[k]$ is a superposition of multiple phase rotating and time-varying transmit streams $x_n[k]$ in correlated noise. Note that the physical wireless channel impulse responses $h_{mn}[k]$ are sometimes assumed to be linear and time invariant. However, due to the variations in the sampling and the variations of the transmit carriers, the observed responses vary in general slowly with time. A system block diagram of the equation is given in Figure 2.11. In the following, specific instances of (2.40) will be explored.

2.5.1 Synchronized Symbol-Spaced Model

Consider the case where the nominal Tx rate is equal to the symbol rate and the Rx rate, i.e., $1/T_{\text{nom,Tx}} = 1/T_s = 1/T_{\text{nom,Rx}}$. Additionally, $\dot{\phi}_{\text{Rx},m}[k] = \dot{\phi}_{\text{Tx},n}[k] = 0$ and $\phi_{\text{Rx},m}[k] = \phi_{\text{Tx},n}[k]$. Note that for the carrier phase processes it is sufficient to be equal to completely cancel their effect, whereas the sampling phase processes have to be equal and zero in order to always sample at the correct symbol time instant.

The model then simplifies to

$$y_m[k] = h_{\text{Rx},m}[k] \star \sum_{n=1}^{N} h_{mn}[k] \star \sum_{l=-\infty}^{\infty} h_{\text{Tx},n}[k-l]\, x_n[l] + h_{\text{Rx},m}[k] \star n_m[k]. \tag{2.41}$$

Assuming further that the transmitter and receiver characteristic is flat within the signal bandwidth, i.e., $h_{\text{Tx},n}[k] = h_{\text{Rx},m}[k] = \text{sinc}[\pi \cdot k]$, yields the common model[18]

$$y_m[k] = \sum_{n=1}^{N} \sum_{l=0}^{L_{mn}-1} h_{mn}[l]\, x_n[k-l] + n_m[k], \tag{2.42}$$

where the wireless channel impulse response is considered to be of finite length L_{mn}, with L being the maximum length across all responses. Continuing with the more common vector-matrix notation, and using TxRx-ST stacking, gives

$$y_m[k] = \underbrace{\begin{bmatrix} \mathbf{h}_m^{\text{T}}[0] & \mathbf{h}_m^{\text{T}}[1] & \cdots & \mathbf{h}_m^{\text{T}}[L-1] \end{bmatrix}}_{\mathbf{h}_{L,m}^{\text{T}}} \underbrace{\begin{bmatrix} \mathbf{x}[k] \\ \mathbf{x}[k-1] \\ \vdots \\ \mathbf{x}[k-L+1] \end{bmatrix}}_{\mathbf{x}_L[k]} + n_m[k], \tag{2.43}$$

where $\mathbf{h}_m[l] = \begin{bmatrix} h_{m1}[l] & h_{m2}[l] & \cdots & h_{mN}[l] \end{bmatrix}^{\text{T}}$ and $\mathbf{x}[k] = \begin{bmatrix} x_1[k] & x_2[k] & \cdots & x_N[k] \end{bmatrix}^{\text{T}}$. The signals from the different receive antennas can then be stacked, giving for flat transmit and receive characteristics

$$\mathbf{y}[k] = \mathbf{H}\mathbf{x}_L[k] + \mathbf{n}[k], \tag{2.44}$$

with the variables being $\mathbf{y}[k] = \begin{bmatrix} y_1[k] & y_2[k] & \cdots & y_M[k] \end{bmatrix}^{\text{T}}$, $\mathbf{H} = \begin{bmatrix} \mathbf{h}_{L,1} & \mathbf{h}_{L,2} & \cdots & \mathbf{h}_{L,M} \end{bmatrix}^{\text{T}}$, and $\mathbf{n}[k] = \begin{bmatrix} n_1[k] & n_2[k] & \cdots & n_M[k] \end{bmatrix}^{\text{T}}$, where $n_m[k] \sim \mathscr{CN}(0, \sigma_{n_m}^2)$. When the transmit and receive filters are not flat, the kth received sample from all antennas is given by the more general model as

$$\mathbf{y}[k] = \underbrace{\mathbf{H}_{\text{Rx}}\mathbf{H}_{L_{\text{Rx}}}\mathbf{H}_{\text{Tx},L_{\text{D}}}}_{\mathbf{H}_{\text{C}}} \mathbf{x}_{L_{\text{C}}}[k] + \mathbf{H}_{\text{Rx}}\mathbf{n}_{L_{\text{Rx}}}[k], \tag{2.45}$$

[18] In some works, it is assumed that $h_{mn}[k]$ contains the Tx and Rx characteristics. This will not be done here, as it obscures the effects of the timing impairments to some extent.

where it was also assumed that the receive and transmit filters $h_{\mathrm{Rx},m}[k]$ and $h_{\mathrm{Tx},n}[k]$ are of finite lengths $L_{\mathrm{Rx},m}$, $L_{\mathrm{Tx},n}$, with L_{Rx}, L_{Tx} being the maximum, respectively. Note that since there is a convolution of different discrete-time and finite-length filters for each received stream, the total spread of the signal in time can be calculated by combining their individual lengths with $L_{\mathrm{C}} = L_{\mathrm{Rx}} + L - 1 + L_{\mathrm{Tx}} - 1$. Additionally, define the maximum length of the spread excluding the transmitter part as $L_{\mathrm{D}} = L_{\mathrm{Rx}} + L - 1$.

Using these quantities, the vectors and matrices in (2.45) can be defined, with their dimensions below the name, as follows

$$
\begin{aligned}
\underset{(M\times ML_{\mathrm{Rx}})}{\mathbf{H}_{\mathrm{Rx}}} &= \begin{bmatrix} \mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[0]) & \mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[1]) & \cdots & \mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[L_{\mathrm{Rx}}-1]) \end{bmatrix}, \\
\underset{(ML_{\mathrm{Rx}}\times NL_{\mathrm{D}})}{\mathbf{H}_{L_{\mathrm{Rx}}}} &= \begin{bmatrix} \mathbf{H} & \mathbf{0}_{M\times N(L_{\mathrm{Rx}}-1)} & & \\ \mathbf{0}_{M\times N} & \mathbf{H} & \mathbf{0}_{M\times N(L_{\mathrm{Rx}}-2)} & \\ & & \ddots & \\ & & \mathbf{0}_{M\times N(L_{\mathrm{Rx}}-1)} & \mathbf{H} \end{bmatrix}, \\
\underset{(NL_{\mathrm{D}}\times NL_{\mathrm{C}})}{\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}} &= \begin{bmatrix} \mathbf{H}_{\mathrm{Tx}} & \mathbf{0}_{N\times N(L_{\mathrm{D}}-1)} & & \\ \mathbf{0}_{N\times N} & \mathbf{H}_{\mathrm{Tx}} & \mathbf{0}_{N\times N(L_{\mathrm{D}}-2)} & \\ & & \ddots & \\ & & \mathbf{0}_{N\times N(L_{\mathrm{D}}-1)} & \mathbf{H}_{\mathrm{Tx}} \end{bmatrix} \\
&\text{with } \underset{(N\times NL_{\mathrm{Tx}})}{\mathbf{H}_{\mathrm{Tx}}} = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_{\mathrm{Tx}}[0]) & \mathrm{diag}(\mathbf{h}_{\mathrm{Tx}}[1]) & \cdots & \mathrm{diag}(\mathbf{h}_{\mathrm{Tx}}[L_{\mathrm{Tx}}-1]) \end{bmatrix}, \\
\underset{(ML_{\mathrm{Rx}}\times 1)}{\mathbf{n}_{L_{\mathrm{Rx}}}[k]} &= \begin{bmatrix} \mathbf{n}[k] \\ \mathbf{n}[k-1] \\ \vdots \\ \mathbf{n}[k-L_{\mathrm{Rx}}+1] \end{bmatrix}, \underset{(NL_{\mathrm{C}}\times 1)}{\mathbf{x}_{L_{\mathrm{C}}}[k]} = \begin{bmatrix} \mathbf{x}[k] \\ \mathbf{x}[k-1] \\ \vdots \\ \mathbf{x}[k-L_{\mathrm{C}}+1] \end{bmatrix}.
\end{aligned}
$$

The matrices $\mathbf{H}_{\mathrm{Rx}}$ and $\mathbf{H}_{\mathrm{Tx}}$ are concatenated diagonal matrices of their filter taps $\mathbf{h}_{\mathrm{Rx}}[l_{\mathrm{Rx}}] = \begin{bmatrix} h_{\mathrm{Rx},1}[l_{\mathrm{Rx}}] & h_{\mathrm{Rx},2}[l_{\mathrm{Rx}}] & \cdots & h_{\mathrm{Rx},M}[l_{\mathrm{Rx}}] \end{bmatrix}^{\mathrm{T}}$ on the receiver side, and, on the transmitter side, $\mathbf{h}_{\mathrm{Tx}}[l_{\mathrm{Tx}}] = \begin{bmatrix} h_{\mathrm{Tx},1}[l_{\mathrm{Tx}}] & h_{\mathrm{Tx},2}[l_{\mathrm{Tx}}] & \cdots & h_{\mathrm{Tx},N}[l_{\mathrm{Tx}}] \end{bmatrix}^{\mathrm{T}}$. If there is additional coupling between different transmitter or receiver chains it could be included here, by removing the diagonal property. It could also be included directly in the channel matrix $\mathbf{H}$. The matrices $\mathbf{H}_{L_{\mathrm{Rx}}}$ and $\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}$ have a block banded and circulant structure. To get more familiar with the notation the small Example 1 will be considered next.

Example 1. *Consider a small MIMO system with $M = N = 2$ having mildly-selective transmit and receive filters of length $L_{Rx} = L_{Tx} = 2$ with $\mathbf{h}_{Rx}[0] = \mathbf{h}_{Tx}[0] = \begin{bmatrix} 1 & 1 \end{bmatrix}^{\mathrm{T}}$, $\mathbf{h}_{Rx}[1] = \begin{bmatrix} 0.05 & 0.1 \end{bmatrix}^{\mathrm{T}}$, and $\mathbf{h}_{Tx}[1] = \begin{bmatrix} -0.15 & 0 \end{bmatrix}^{\mathrm{T}}$. The wireless channel is also mildly selective with $L = 2$ and $\mathbf{h}_{L,1} =$*

$\begin{bmatrix}1 & -j & 0.1 & -0.1\end{bmatrix}^{\mathrm{T}}$, $\mathbf{h}_{L,2} = \begin{bmatrix}-j & 1 & 0.5 & 0\end{bmatrix}^{\mathrm{T}}$. *Using the notation, the matrices are given by*

$$\mathbf{H}_C = \mathbf{H}_{Rx}\mathbf{H}_{L_{Rx}}\mathbf{H}_{Tx,L_D} = \begin{bmatrix}1 & 0 & 0.05 & 0\\ 0 & 1 & 0 & 0.1\end{bmatrix} \cdot \begin{bmatrix}1 & -j & 0.1 & -0.1 & 0 & 0\\ -j & 1 & 0.5 & 0 & 0 & 0\\ 0 & 0 & 1 & -j & 0.1 & -0.1\\ 0 & 0 & -j & 1 & 0.5 & 0\end{bmatrix}$$
$$\cdot \begin{bmatrix}1 & 0 & -0.15 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & -0.15 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & -0.15 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\end{bmatrix}.$$

It is possible to combine $\mathbf{H}_{\mathrm{Rx}}$, $\mathbf{H}_{L_{\mathrm{Rx}}}$, and $\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}$ into $\mathbf{H}_{\mathrm{C}}$, and consider the complete system as one frequency-selective and time-invariant system with spread of L_{C}. However, there is some structure in the generation of that matrix which can be leveraged to reduce the computational complexity of equalization and synchronization, especially, once the timing impairments are considered.

2.5.2 Symbol-Spaced Model Containing Only CFOs

Similar to the previous section with $T_{\mathrm{nom,Tx}} = T_s = T_{\mathrm{nom,Rx}}$ and $\dot{\phi}_{\mathrm{Rx},m}[k] = \dot{\phi}_{\mathrm{Tx},n}[k] = 0$, we get from (2.40) the model containing only CFOs as

$$\begin{aligned} y_m[k] &= h_{\mathrm{Rx},m}[k] \star \sum_{n=1}^{N} e^{j\Delta\phi_{mn}[k]} \cdot h_{mn}[k] \star \sum_{l=-\infty}^{\infty} h_{\mathrm{Tx},n}[k-l]\, x_n[l] \\ &\quad + h_{\mathrm{Rx},m}[k] \star n_m[k] \end{aligned} \tag{2.46}$$

$$\begin{aligned} &= h_{\mathrm{Rx},m}[k] \star e^{-j\phi_{\mathrm{Rx},m}[k]} \cdot \sum_{n=1}^{N} h_{mn}[k] \cdot e^{j\phi_{\mathrm{Tx},n}[k]} \star \sum_{l=-\infty}^{\infty} h_{\mathrm{Tx},n}[k-l]\, x_n[l] \\ &\quad + h_{\mathrm{Rx},m}[k] \star n_m[k]. \end{aligned} \tag{2.47}$$

Using vector-matrix notation and assuming, again, that the Rx and Tx filters are ideally flat over the relevant signal bandwidth, i.e., they do not contribute to the characteristic of the complete transmission channel, including the shift due to the frequency offsets, gives

$$\mathbf{y}[k] = \mathbf{\Phi}_{\mathrm{Rx}}[k]\mathbf{H}\mathbf{\Phi}_{\mathrm{Tx},L}[k]\mathbf{x}_L[k] + \mathbf{n}[k] \tag{2.48}$$

$$= (\mathbf{\Delta\Phi}[k] \odot \mathbf{H})\,\mathbf{x}_L[k] + \mathbf{n}[k], \tag{2.49}$$

where $\mathbf{\Phi}_{\mathrm{Rx}}[k] = \mathrm{diag}(e^{-j\phi_{\mathrm{Rx},1}[k]}, e^{-j\phi_{\mathrm{Rx},2}[k]}, \ldots, e^{-j\phi_{\mathrm{Rx},M}[k]}) = \mathrm{diag}(\boldsymbol{\phi}_{\mathrm{Rx}}[k])$, $\mathbf{\Phi}_{\mathrm{Tx},L}[k] = \mathbf{I}_L \otimes \mathbf{\Phi}_{\mathrm{Tx}}[k]$ with $\mathbf{\Phi}_{\mathrm{Tx}}[k] = \mathrm{diag}(e^{j\phi_{\mathrm{Tx},1}[k]}, e^{j\phi_{\mathrm{Tx},2}[k]}, \ldots, e^{j\phi_{\mathrm{Tx},N}[k]}) = \mathrm{diag}(\boldsymbol{\phi}_{\mathrm{Tx}}[k])$. The complete carrier phase

rotation matrix is given by $\mathbf{\Delta\Phi}[k] = \boldsymbol{\phi}_{\mathrm{Rx}}[k](\mathbf{1}_{L\times 1} \otimes \boldsymbol{\phi}_{\mathrm{Tx}}[k])^{\mathrm{T}}$.

When the Tx and Rx filters are not flat over the relevant bandwidth, the more general model of the kth received sample from all antennas can be written as

$$\mathbf{y}[k] = \mathbf{H}_{\mathrm{Rx}}\mathbf{\Phi}_{\mathrm{Rx},L_{\mathrm{Rx}}}[k]\mathbf{H}_{L_{\mathrm{Rx}}}\mathbf{\Phi}_{\mathrm{Tx},L_{\mathrm{D}}}[k]\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}\mathbf{x}_{L_{\mathrm{C}}}[k] + \mathbf{H}_{\mathrm{Rx}}\mathbf{n}_{L_{\mathrm{Rx}}}[k], \tag{2.50}$$

using the previously defined matrices, and $\mathbf{\Phi}_{\mathrm{Rx},L_{\mathrm{Rx}}}[k]$, $\mathbf{\Phi}_{\mathrm{Tx},L_{\mathrm{D}}}[k]$ being defined as $\mathbf{\Phi}_{\mathrm{Tx},L}[k]$. Compared to (2.45), the impact of CFOs is thus that the transmission is time varying due to the carrier phase variations from the different antennas. In the following small Example 2, this should become more clear.

Example 2. *Consider a small MIMO system with $M = N = 2$, having ideal transmit and receive filters of length $L_{Rx} = L_{Tx} = 1$ with $\mathbf{h}_{Rx}[0] = \mathbf{h}_{Tx}[0] = \begin{bmatrix}1 & 1\end{bmatrix}^{\mathrm{T}}$. The wireless channel is also flat and orthogonal with $L = 1$ and $\mathbf{h}_{L,1} = \begin{bmatrix}1 & -j\end{bmatrix}^{\mathrm{T}}$, $\mathbf{h}_{L,2} = \begin{bmatrix}-j & 1\end{bmatrix}^{\mathrm{T}}$. Three different points in time will be considered, i.e., $k = \{0,1,2\}$, with the carrier phase process progressions $\phi_{Rx,1}[k] = \phi_{Rx,2}[k] = \{0, \frac{\pi}{2}, \frac{\pi}{2}\}$, $\phi_{Tx,1}[k] = \{0, 0, \frac{\pi}{2}\}$, and $\phi_{Tx,2}[k] = \{0, 0, -\frac{\pi}{2}\}$. Note that this corresponds to setup 3, where the same phase process is experienced at all Rx antennas and independent phase processes are experienced at all Tx antennas.*

$$\mathbf{\Phi}_{Rx,L_{Rx}}[0]\mathbf{H}_{L_{Rx}}\mathbf{\Phi}_{Tx,L_D}[0] = \begin{bmatrix} e^{-j\cdot 0} & 0 \\ 0 & e^{-j\cdot 0}\end{bmatrix}\begin{bmatrix}1 & -j \\ -j & 1\end{bmatrix}\begin{bmatrix} e^{j\cdot 0} & 0 \\ 0 & e^{j\cdot 0}\end{bmatrix} = \begin{bmatrix}1 & -j \\ -j & 1\end{bmatrix}$$
$$\mathbf{\Phi}_{Rx,L_{Rx}}[1]\mathbf{H}_{L_{Rx}}\mathbf{\Phi}_{Tx,L_D}[1] = \begin{bmatrix} e^{-j\cdot\frac{\pi}{2}} & 0 \\ 0 & e^{-j\cdot\frac{\pi}{2}}\end{bmatrix}\begin{bmatrix}1 & -j \\ -j & 1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} = e^{-j\cdot\frac{\pi}{2}}\begin{bmatrix}1 & -j \\ -j & 1\end{bmatrix}$$
$$\mathbf{\Phi}_{Rx,L_{Rx}}[2]\mathbf{H}_{L_{Rx}}\mathbf{\Phi}_{Tx,L_D}[2] = e^{-j\cdot\frac{\pi}{2}}\begin{bmatrix}1 & -j \\ -j & 1\end{bmatrix}\begin{bmatrix} e^{j\cdot\frac{\pi}{2}} & 0 \\ 0 & e^{-j\cdot\frac{\pi}{2}}\end{bmatrix} = \begin{bmatrix}1 & j \\ -j & -1\end{bmatrix}$$

2.5.3 Symbol-Spaced Model Containing Only SFOs

Consider $T_{\mathrm{nom,Tx}} = T_s = T_{\mathrm{nom,Rx}}$ and $\phi_{\mathrm{Rx},m}[k] = \phi_{\mathrm{Tx},n}[k]$. This yields

$$\begin{aligned} y_m[k] = {} & h_{\mathrm{Rx},m}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k]\right] \star \sum_{n=1}^{N} h_{mn}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ & \star \sum_{l=-\infty}^{\infty} x_n[l]\, h_{\mathrm{Tx},n}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k] - l - \dot{\phi}_{\mathrm{Tx},n}[l]\right] + h_{\mathrm{Rx},m}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k]\right] \star n_m[k]. \end{aligned} \tag{2.51}$$

Note that idealized Rx and Tx filters do not exist in this case because as the sampling phase processes progress, relevant filter coefficients, aside from a $\delta[k]$, will appear at the symbol positions generating ISI between consecutive symbols. In fact, these filters are infinitely long if the sample timing is not perfect. Fortunately, the higher coefficients have significantly less impact, i.e., their magnitude is very low, such that the infinite-length filters $h_{\mathrm{Rx},m}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k]\right]$

and $h_{\mathrm{Tx},n}\left[k+\dot{\phi}_{\mathrm{Rx},m}[k]-l-\dot{\phi}_{\mathrm{Tx},n}[l]\right]$ can be approximated by finite-length filters with lengths L_{Rx} and L_{Tx}. Example 3 shows the effect of a fixed sampling phase offset on the sampling of one analog receive filter.

Example 3. *Consider the analog Rx filter characteristic of the second antenna to be ideally flat with* $h_{Rx,2}(t) = \mathrm{sinc}\left(\pi \cdot \frac{t}{T_{nom,Rx}}\right) = \mathrm{sinc}\left(\pi \cdot \frac{t}{T_s}\right)$. *Assume its discrete-time version to be of finite length* $L_{Rx} = 7$ *with* $h_{Rx,2}[l=3] = 1$, $h_{Rx,2}[l \neq 3] = 0$, *for the ideal case of* $\dot{\phi}_{Rx,2}[k] = 0$. *Assume the case where* $\dot{\phi}_{Rx,2}[k]$ *is a constant for at least* L_{Rx} *symbols. Then, the filter coefficients are given by shifting the ideal filter to that constant, with non-integer values being obtained from interpolation using the analog filter response. Fig. 2.12 shows the filter coefficients for three different shift values, where the actual sampling clock is slower than the nominal one* ($\dot{\phi}_{Rx,m}[k] > 0$).

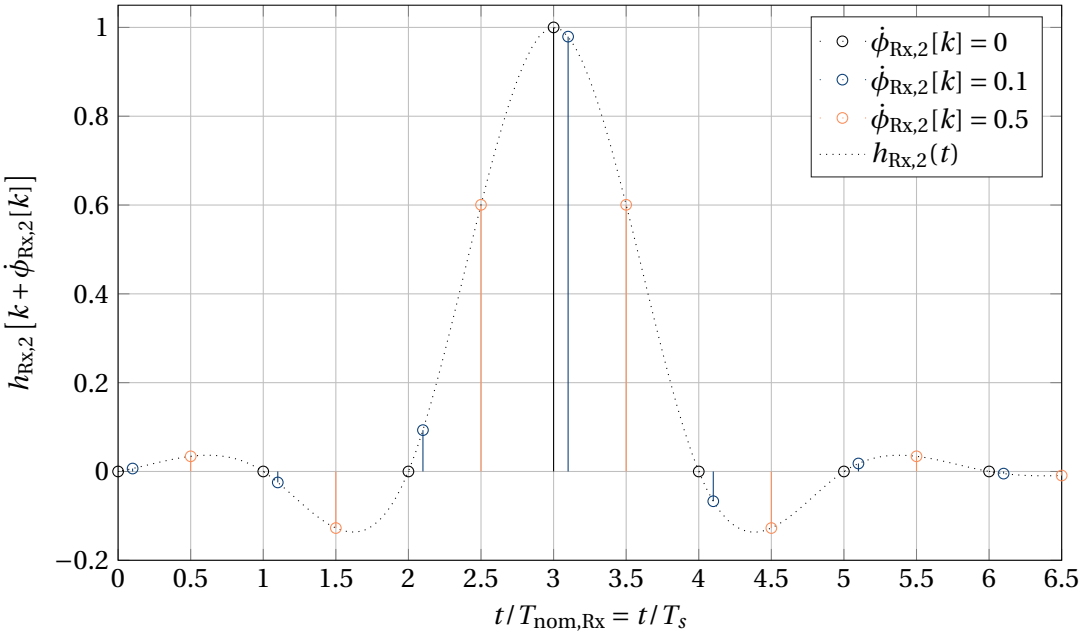


Figure 2.12: Example of the observed receive filter coefficients for different constant sampling phases, and with symbol rate sampling.

It is crucial to note that the observation of the wireless channel and transmit filters in each received stream also depends on the receiver sampling processes $\dot{\phi}_{\mathrm{Rx},m}[k]$. Thus, including the effects of multiple SFOs needs a slightly different notation. We can start with the kth sample for the mth receive antenna, given by

$$y_m[k] = \mathbf{h}^{\mathrm{T}}_{\mathrm{Rx},m}[k]\mathbf{H}_{L_{\mathrm{Rx}},m}[k]\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}},m}[k]\mathbf{x}_{L_{\mathrm{C}}}[k] + \mathbf{h}^{\mathrm{T}}_{\mathrm{Rx},m}[k]\mathbf{n}_{L_{\mathrm{Rx}},m}[k], \tag{2.52}$$

where all of the filters vary with time according to the sampling phase process of the mth

receiver. The following definitions are needed

$$\mathbf{h}_{\mathrm{Rx},m}[k] = \begin{bmatrix} h_{\mathrm{Rx},m}\left[0+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ h_{\mathrm{Rx},m}\left[1+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ \vdots \\ h_{\mathrm{Rx},m}\left[L_{\mathrm{Rx}}-1+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \end{bmatrix},$$

$$\mathbf{H}_{L_{\mathrm{Rx}},m}[k] = \begin{bmatrix} \mathbf{h}_{L,m}^{\mathrm{T}}[k] & \mathbf{0}_{1\times N(L_{\mathrm{Rx}}-1)} & & \\ \mathbf{0}_{1\times N} & \mathbf{h}_{L,m}^{\mathrm{T}}[k] & \mathbf{0}_{1\times N(L_{\mathrm{Rx}}-2)} & \\ & & \ddots & \\ & & & \mathbf{0}_{1\times N(L_{\mathrm{Rx}}-1)} \quad \mathbf{h}_{L,m}^{\mathrm{T}}[k] \end{bmatrix}$$

$$\text{with} \quad \mathbf{h}_{L,m}[k] = \begin{bmatrix} \mathbf{h}_m\left[0+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ \mathbf{h}_m\left[1+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ \vdots \\ \mathbf{h}_m\left[L-1+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \end{bmatrix},$$

and, additionally,

$$\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}},m}[k] = \begin{bmatrix} \mathbf{H}_{\mathrm{Tx},m}[k,0] & \mathbf{0}_{N\times N(L_{\mathrm{D}}-1)} & & \\ \mathbf{0}_{N\times N} & \mathbf{H}_{\mathrm{Tx},m}[k,1] & \mathbf{0}_{N\times N(L_{\mathrm{D}}-2)} & \\ & & \ddots & \\ & & & \mathbf{0}_{N\times N(L_{\mathrm{D}}-1)} \quad \mathbf{H}_{\mathrm{Tx},m}[k,L_{\mathrm{D}}-1] \end{bmatrix}$$

with

$$\mathbf{H}_{\mathrm{Tx},m}[k,l_{\mathrm{D}}] = \Big[\mathrm{diag}(\mathbf{h}_{\mathrm{Tx},m}[k,l_{\mathrm{D}},0]) \quad \mathrm{diag}(\mathbf{h}_{\mathrm{Tx},m}[k,l_{\mathrm{D}},1]) \quad \cdots \quad \mathrm{diag}(\mathbf{h}_{\mathrm{Tx},m}[k,l_{\mathrm{D}},L_{\mathrm{Tx}}-1])\Big],$$

$$\text{and} \quad \mathbf{h}_{\mathrm{Tx},m}[k,l_{\mathrm{D}},l] = \begin{bmatrix} h_{\mathrm{Tx},1}\left[l-\dot{\phi}_{\mathrm{Tx},1}[k-l-l_{\mathrm{D}}]+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ h_{\mathrm{Tx},2}\left[l-\dot{\phi}_{\mathrm{Tx},2}[k-l-l_{\mathrm{D}}]+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ \vdots \\ h_{\mathrm{Tx},N}\left[l-\dot{\phi}_{\mathrm{Tx},N}[k-l-l_{\mathrm{D}}]+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \end{bmatrix}, \quad \mathbf{n}_{L_{\mathrm{Rx}},m}[k] \begin{bmatrix} n_m[k] \\ n_m[k-1] \\ \vdots \\ n_m[k-L_{\mathrm{Rx}}+1] \end{bmatrix},$$

where the auxiliary variable l_{D} is needed in order to apply the appropriate value of the transmitter sampling phase process to the corresponding transmit symbol, and the lengths L_{C}, L_{D} are defined as before in section 2.5.1.

The general model for the kth received sample from all antennas, including the Tx and Rx filters, can then be specified by a more complex stacking, using the convention of TxRx-ST, with

$$\mathbf{y}[k] = \mathbf{H}_{\mathrm{Rx}}[k]\mathbf{H}_{L_{\mathrm{Rx}}}[k]\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}[k]\mathbf{x}_{L_{\mathrm{C}}}[k] + \mathbf{H}_{\mathrm{Rx}}[k]\mathbf{n}_{L_{\mathrm{Rx}}}[k], \tag{2.53}$$

where some of the matrices need to be defined differently compared to (2.45) due to the

sampling process, which will be done next. The received filters are defined by the matrix

$$\underset{(M\times ML_{\mathrm{Rx}})}{\mathbf{H}_{\mathrm{Rx}}[k]} = \begin{bmatrix}\mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[k,0]) & \mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[k,1]) & \cdots & \mathrm{diag}(\mathbf{h}_{\mathrm{Rx}}[k,L_{\mathrm{Rx}}-1])\end{bmatrix}$$

$$\text{with} \quad \mathbf{h}_{\mathrm{Rx}}[k,l] = \begin{bmatrix} h_{\mathrm{Rx},1}\left[l+\dot{\phi}_{\mathrm{Rx},1}[k]\right] \\ h_{\mathrm{Rx},2}\left[l+\dot{\phi}_{\mathrm{Rx},2}[k]\right] \\ \vdots \\ h_{\mathrm{Rx},M}\left[l+\dot{\phi}_{\mathrm{Rx},M}[k]\right]\end{bmatrix},$$

and are thus similar to the one used in (2.45) except for the variation with k. The wireless channel response can be written as

$$\underset{(ML_{\mathrm{Rx}}\times MNL_{\mathrm{D}})}{\mathbf{H}_{L_{\mathrm{Rx}}}[k]} = \begin{bmatrix} \mathbf{H}_L[k] & \mathbf{0}_{M\times MN(L_{\mathrm{Rx}}-1)} & & \\ \mathbf{0}_{M\times MN} & \mathbf{H}_L[k] & \mathbf{0}_{M\times MN(L_{\mathrm{Rx}}-2)} & \\ & & \ddots & \\ & & & \mathbf{0}_{M\times MN(L_{\mathrm{Rx}}-1)} \quad \mathbf{H}_L[k]\end{bmatrix}$$

$$\text{with} \quad \mathbf{H}_L[k] = \begin{bmatrix}\mathbf{H}[k,0] & \mathbf{H}[k,1] & \cdots & \mathbf{H}[k,L-1]\end{bmatrix},$$

$$\text{using} \quad \mathbf{H}[k,l] = \mathrm{blkdiag}\left(\mathbf{h}_1^{\mathrm{T}}[k,l],\mathbf{h}_2^{\mathrm{T}}[k,l],\ldots,\mathbf{h}_M^{\mathrm{T}}[k,l]\right)$$

$$\text{and} \quad \mathbf{h}_m[k,l] = \begin{bmatrix} h_{m1}\left[l+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ h_{m2}\left[l+\dot{\phi}_{\mathrm{Rx},m}[k]\right] \\ \vdots \\ h_{mN}\left[l+\dot{\phi}_{\mathrm{Rx},m}[k]\right]\end{bmatrix} = \mathbf{h}_m[l+\dot{\phi}_{\mathrm{Rx},m}[k]].$$

Finally, the transmit filters are given as

$$\underset{(MNL_{\mathrm{D}}\times NL_{\mathrm{C}})}{\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}[k]} = \begin{bmatrix} \mathbf{H}_{\mathrm{Tx}}[k,0] & \mathbf{0}_{MN\times N(L_{\mathrm{D}}-1)} & & \\ \mathbf{0}_{MN\times N} & \mathbf{H}_{\mathrm{Tx}}[k,1] & \mathbf{0}_{MN\times N(L_{\mathrm{D}}-2)} & \\ & & \ddots & \\ & & & \mathbf{0}_{MN\times N(L_{\mathrm{D}}-1)} \quad \mathbf{H}_{\mathrm{Tx}}[k,L_{\mathrm{D}}-1]\end{bmatrix}$$

$$\text{with} \quad \mathbf{H}_{\mathrm{Tx}}[k,l_{\mathrm{D}}] = \begin{bmatrix} \mathbf{H}_{\mathrm{Tx},1}[k,l_{\mathrm{D}}] \\ \mathbf{H}_{\mathrm{Tx},2}[k,l_{\mathrm{D}}] \\ \vdots \\ \mathbf{H}_{\mathrm{Tx},M}[k,l_{\mathrm{D}}]\end{bmatrix}.$$

Compared to (2.45), note in particular that the size of $\mathbf{H}_{L_{\mathrm{Rx}}}[k]$ and $\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}[k]$ is increased to $ML_{\mathrm{Rx}} \times MNL_{\mathrm{D}}$ and $MNL_{\mathrm{D}} \times NL_{\mathrm{C}}$, respectively, due to the independent sampling process for each receiver. It should also be noted that in a continuous transmission, as mentioned before and discussed in appendix A.3, the lengths of the filters change as time progresses, because eventually the same symbol from the transmitter will be sampled twice or it will be missed. This can be seen, for example, when considering that one of the receiver sampling

phase processes reaches a value of $|\dot{\phi}_{\mathrm{Rx},m}[k]| = 1$, while the others are smaller than one. In that case the corresponding receive filter will be shifted by one sample, also equal to a symbol here, forward or backward, which extends through to the transmitter side, i.e., one additional previous or next symbol is needed in order to compute $\mathbf{y}[k]$. This problem will not be treated further in this work, approaches and solutions can be found, for example, in [4].

2.5.4 Symbol-Spaced Model Containing CFOs and SFOs

Using (2.40) with $T_{\mathrm{nom,Tx}} = T_s = T_{\mathrm{nom,Rx}}$ and the definitions from the previous two sections, the symbol-spaced model containing both timing impairments, and the transmit and receive filter effects, can be specified in vector matrix notation, and is given by

$$\mathbf{y}[k] = \mathbf{H}_{\mathrm{Rx}}[k]\mathbf{\Phi}_{\mathrm{Rx},L_{\mathrm{Rx}}}[k]\mathbf{H}_{L_{\mathrm{Rx}}}[k]\mathbf{\Phi}_{\mathrm{Tx},L_{\mathrm{D}}}[k]\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}[k]\mathbf{x}_{L_{\mathrm{C}}}[k] + \mathbf{H}_{\mathrm{Rx}}[k]\mathbf{n}_{L_{\mathrm{Rx}}}[k], \tag{2.54}$$

with all the quantities being defined in the previous two sections. We can see that due to the timing impairments, the input-output relation is time varying through the different matrix terms, where the CFO terms influence mostly the phase and the SFO terms mostly influence the amplitude of the filters, see their respective definitions.

As in the synchronized case, it is possible to combine all effects into one frequency-selective and time-varying matrix with spread of L_{C}. However, the variation happens due to the MIMO mixing in a very complicated, albeit slow, way, such that estimating it accurately at every time step k poses a serious problem. Additionally, some of the processing steps can be parallelized by recognizing the structure of the system[19].

2.5.5 Intermission on Oversampling

So far, we have only considered the case, where the transmitter and receiver sample rates are equal and match the symbol rate, i.e., $1/T_{\mathrm{nom,Tx}} = 1/T_{\mathrm{nom,Rx}} = 1/T_s$. If the timing impairments are not compensated by analog means, such a sampling can only yield suboptimal digital synchronization performance [4, 61, 62]. The reason for this is the occurrence of aliasing due to the CFO and SFO, as will be explained next. The CFO shifts the received signals spectrum from ideal zero to the frequency offset, which creates aliasing when sampled at the symbol rate assuming no offset, violating the first Nyquist criterion. Likewise, when the transmitter sample rate is slightly different from the nominal one (this is true in the same way for the receiver sampling), the actual physical bandwidth is different, e.g., higher, compared to the nominal one. Sampling can then also create aliasing, violating the first Nyquist criterion.

Due to these limitations, most systems sample at a rate that is higher than the symbol rate on the receiver side , i.e., $1/T_{\mathrm{nom,Rx}} > 1/T_s$, in order to generate a sufficient statistic from

[19] Note for example that $M \cdot N$ different phase trajectories can be observed with setup 4 for both CFO and SFO, refer also to matrix $\Phi_{\Delta}[k]$ in equation (2.49). However, they are linear combinations of the $M+N$ actual phase processes from transmitter and receiver.

the continuous-time received signal. The oversampling factor $\bar{Q} = \frac{T_s}{T_{\text{nom,Rx}}} \geq 1$ is in practical systems often chosen in the range $1.1 < \bar{Q} < 2.5$, in order to achieve good performance, while not demanding excessive sampling speeds from the ADC. The actual required value depends on the level of CFO, SFO, and the shape of the spectrum of the received signal. In general, the oversampling factor should be chosen such that most of the signal power is captured within the sampling bandwidth, while simultaneously avoiding aliasing due to the timing impairments. Note that an increase in oversampling increases the required digital processing rate by the same amount, which can be a limiting factor especially for high-rate systems. As is common for discrete-time processing, we will assume that the signal has been interpolated such that an integer oversampling factor $Q = \lceil \bar{Q} \rceil$ can be used. In the following, the signal model for the ideally synchronized case, and for the case where both timing impairments are occurring, will be presented.

2.5.6 Synchronized Oversampled Model

For the oversampled case, all the filters and signals need to be specified on the oversampled grid, denoted through $\dot{(\cdot)}$. The general model for the kth sample from all receive antennas can be defined as

$$\dot{\mathbf{y}}[k] = \dot{\mathbf{H}}_{\text{Rx}} \dot{\mathbf{H}}_{\dot{L}_{\text{Rx}}} \dot{\mathbf{H}}_{\text{Tx},\dot{L}_{\text{D}}} \dot{\mathbf{x}}_{\dot{L}_{\text{C}}}[k] + \dot{\mathbf{H}}_{\text{Rx}} \dot{\mathbf{n}}_{\dot{L}_{\text{Rx}}}[k], \tag{2.55}$$

where the filters $\dot{\mathbf{H}}_{\text{Rx}}$, $\dot{\mathbf{H}}_{\dot{L}_{\text{Rx}}}$, and $\dot{\mathbf{H}}_{\text{Tx},\dot{L}_{\text{D}}}$, are essentially the same as in section 2.5.1, but contain the responses on the oversampled grid. Comparing between a symbol-spaced and an oversampled instance of the system, the filter lengths are increased, i.e., $\dot{L}_{\text{Rx}} = QL_{\text{Rx}}$, $\dot{L}_{\text{D}} = QL_{\text{D}}$, and $\dot{L}_{\text{C}} = QL_{\text{C}}$. The statistical properties of $\dot{\mathbf{n}}_{\dot{L}_{\text{Rx}}}[k]$ can still be assumed that of a white Gaussian noise process with equal power across all frequencies[20]. However, it should be visible that the actual noise that is added as a perturbation, i.e., the term $\dot{\mathbf{H}}_{\text{Rx}} \dot{\mathbf{n}}_{\dot{L}_{\text{Rx}}}[k]$ is not necessarily white, but correlated over time depending on the receive filters.

Finally, consider the relationship between the input symbols or samples, and their oversampled description $\dot{\mathbf{x}}_{\dot{L}_{\text{C}}}[k]$. For this, consider the kth sample at the nth antenna, where all of them are used to stack the previously mentioned vector, which can be defined as

$$\begin{aligned} \dot{x}_n[k] = &\begin{bmatrix} h_{\text{P},n}[1+\text{mod}(k,Q)] & h_{\text{P},n}[1+Q+\text{mod}(k,Q)] & \cdots & h_{\text{P},n}[1+(L_{\text{P}}-1)Q+\text{mod}(k,Q)] \end{bmatrix} \\ &\cdot \begin{bmatrix} x_n\left[\left\lfloor \frac{k}{Q} \right\rfloor\right] \\ x_n\left[\left\lfloor \frac{k}{Q} \right\rfloor - 1\right] \\ \vdots \\ x_n\left[\left\lfloor \frac{k}{Q} \right\rfloor - L_{\text{P}} + 1\right] \end{bmatrix}, \end{aligned} \tag{2.56}$$

[20] For fair comparison, it is important that the signal-to-noise ratio (SNR) inside the signal bandwidth is equal to the SNR of the symbol-spaced case.

where $h_{\mathrm{P},n}[l]$ is a filter of length $\dot{L}_{\mathrm{P}} = QL_{\mathrm{P}}$ on the receiver sampling grid that represents the digital pulse-shaping filter of the nth transmitter. The operators $\mathrm{mod}(\cdot)$ and $\lfloor\cdot\rfloor$ denote the modulo and flooring operation, respectively. Note that the transmit symbols, which influence the current sample, only change every Q samples. From a receiver perspective, the pulse-shaping filter can be combined with the analog filter $\dot{\mathbf{H}}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}$. However, the digital filter can, up to some limitations, be freely chosen in order to improve performance, whereas the analog filter is fixed for a given hardware setup.

2.5.7 Oversampled Model Containing CFOs and SFOs

For the CFO and SFO impaired oversampled model we get

$$\dot{\mathbf{y}}[k] = \dot{\mathbf{H}}_{\mathrm{Rx}}[k]\boldsymbol{\Phi}_{\mathrm{Rx},\dot{L}_{\mathrm{Rx}}}[k]\dot{\mathbf{H}}_{\dot{L}_{\mathrm{Rx}}}[k]\boldsymbol{\Phi}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}[k]\dot{\mathbf{H}}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}[k]\dot{\mathbf{x}}_{\dot{L}_{\mathrm{C}}}[k] + \dot{\mathbf{H}}_{\mathrm{Rx}}[k]\dot{\mathbf{n}}_{\dot{L}_{\mathrm{Rx}}}[k] \tag{2.57}$$

$$= \underbrace{\dot{\mathbf{H}}_{\mathrm{Rx}}[k]\boldsymbol{\Phi}_{\mathrm{Rx},\dot{L}_{\mathrm{Rx}}}[k]\dot{\mathbf{H}}_{\dot{L}_{\mathrm{Rx}}}[k]\boldsymbol{\Phi}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}[k]\dot{\mathbf{H}}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}[k]\mathbf{H}_{\mathrm{P}}[k]}_{\dot{\mathbf{H}}_{\mathrm{C}}[k]}\mathbf{x}_{L_{\mathrm{P}}}[k] + \dot{\mathbf{H}}_{\mathrm{Rx}}[k]\dot{\mathbf{n}}_{\dot{L}_{\mathrm{Rx}}}[k], \tag{2.58}$$

where the same definitions as above apply. Aside from the change in length for all filters, it just needs to be mentioned that the carrier and sampling phase processes $\phi_{\cdot}[k]$ and $\dot{\phi}_{\cdot}[k]$ are now defined on the receiver sampling grid.

2.6 Results from the Literature

In this section, some results for the synchronization of MIMO systems available in the literature, and their relation to the models specified above, will be laid out. A recent survey of synchronization research for SISO and MIMO systems can be found in [5]. Most of the works in the literature only deal with one of the two timing impairments, i.e., assume the other one compensated, and also assume block-wise frequency-flat channels. Furthermore, almost all of them assume that the phase processes characterizing the timing impairments are sufficiently well modeled, at least over the considered block, by linear phase ramps, or purely linear phase drifts, i.e., $\sigma^2_{\varphi_{\mathrm{w}}} = \sigma^2_{\varphi_{\mathrm{n}}} = 0$ and $\sigma^2_{\dot{\varphi}_{\mathrm{w}}} = \sigma^2_{\dot{\varphi}_{\mathrm{n}}} = 0$. Tables 2.2, 2.3, and 2.4 summarize the most relevant works. In case of a single CFO and/or SFO between transmitter and receiver, certain SISO approaches and results can be used. Specifically, one may use approaches, which are not influenced by the linear superposition of multiple signals such as data-aided algorithms based on correlation with orthogonal training sequences, or blind approaches where the utilized statistic is not affected by the superposition. More results about the synchronization of CFO and SFO in SISO systems can be found in [5], confer also the standard books [3, 4].

In the following, there will be additional details on some of the most important results.

2.6.1 Important Results for CFOs

Besson and Stoica [63] derive the Cramér-Rao bound (CRB) for flat-fading MIMO channels

Table 2.2: Summary of research on CFO synchronization for MIMO systems, under different oscillator setups and channel assumptions.

Authors	CFO	Channel	Contribution
Besson and Stoica [63]	multiple	flat fading	Derivation of CRB and ML estimator for CFOs and channel.
Yao and Ng [64]	multiple	flat fading	Correlation-based estimator and CRB for CFOs.
Qu et al. [65]	multiple	selective fading	Training sequences and subspace-based methods to lower estimation complexity.
Ghogho and Swami [66]	single	selective fading	Training sequences based on CRB, improving performance and lowering ML complexity.
Ahmed et al. [67]	multiple	selective fading	Approximate ML estimator and recursive MMSE equalizer.
Zhang et al. [68]	multiple	selective fading	Frequency-domain estimation and equalization, by approximating CFOs as constant phase shifts per sample block.
Mehrpouyan et al. [69]	multiple	flat fading	CRB, LS estimator, and tracking of phase processes with weighted LS and extended Kalman filter.
Cheng and Larsson [70]	multiple	flat fading	CRB and ML estimator for CFO, and achievable rate analysis for massive MIMO.

with multiple CFOs, i.e., setup 4. One important result from the work is that the CRB is block-diagonal, meaning that the estimation of the parameters of interest can be carried out for each receive antenna m independently. Furthermore, it is noted that the estimation of the channel and CFOs for each antenna depends on all channel parameters and frequency offsets associated with that antenna. On the other hand, when a LO is shared at the transmitter side, i.e., setup 3, the CRB for estimating the CFO is shown to become independent of the frequency offsets. This is intuitive since each received stream experiences only a single, albeit different for each receive antenna, frequency offset in this case. Since the exact expression of the general CRB is rather complicated w.r.t. the mentioned dependencies, the work also proposes using an asymptotic, or large-sample, CRB (asCRB). This result from [63] is given, in our notation, by

$$\text{asCRB}\left(h_{mn}[0]\right) = \frac{5\sigma_{n_m}^2}{2L_{\text{Tr}}\sigma_{x_{\text{Tr},n}}^2},\ \text{asCRB}\left(\Delta\mu_{\varphi_{\text{w}},mn}\right) = \frac{6\sigma_{n_m}^2}{L_{\text{Tr}}^3\sigma_{x_{\text{Tr},n}}^2|h_{mn}[0]|^2}, \tag{2.59}$$

where the frequency offset is $\Delta\mu_{\varphi_{\text{w}},mn} = \mu_{\varphi_{\text{w}},n} - \mu_{\varphi_{\text{w}},m}$ assuming, as mentioned above, $\sigma_{\varphi_{\text{w}}}^2 = \sigma_{\varphi_{\text{n}}}^2 = 0$. Additionally, the noise power is $\sigma_{n_m}^2 = \text{E}[|n_m[k]|^2]$ and $\sigma_{x_{\text{Tr},n}}^2 = \text{E}[|x_{\text{Tr},n}[k]|]^2$ is the

average power of the training signal from antenna n with total length L_{Tr}. The maximum-likelihood (ML) estimation of the parameters is proposed to be carried out in two steps. First, estimate the CFOs. Then, estimate the channels. CFO estimation requires an N-dimensional minimization of a likelihood function for each received antenna, which is computationally very complex. Significant simplifications are achieved by using uncorrelated training sequences from the different Tx antennas. In this case, each CFO can be estimated independently, since the contribution of the other ones are suppressed due to the uncorrelated nature of the transmitted signals. Then, either one-dimensional searches with periodogram techniques or correlation-based estimators can be used to get the offsets. For channel estimation, the obtained frequency offset estimates are used in a standard least-squares (LS) channel estimator. Numerical evaluations show that the asCRB is very close to the CRB, and that the simplified estimators achieve the CRB in the mid to high SNR regime.

In Ahmed et al. [67], the same simplified ML procedure as above is proposed, but for frequency-selective MIMO channels with multiple CFOs, i.e., setup 4. The work also derives a recursive minimum-mean-square-error (MMSE) equalizer, and shows that for the equalization of such a channel, the equalizer needs to be updated for every symbol due to the CFO. Fortunately, consecutive equalization matrices experience some structure, such that it is not necessary to update the complete matrix for every symbol. Numerical results show that the proposed approach performs almost as good as a system that does not incur timing impairments.

Mehrpouyan et al. [69] investigate the case of flat-fading MIMO channels with multiple CFOs, i.e., setup 4, under the assumption that the phase process is solely a Wiener process, i.e., $\Delta\mu_{\varphi_{\mathrm{w}},mn} = \sigma^2_{\varphi_{\mathrm{n}}} = 0$. One significant contribution is the proposal of tracking the phase during data transmission, using a weighted LS or an extended Kalman estimator. The estimation structure makes use of several training signals. First, there are long, orthogonal between different Tx, training sequences that are used for LS channel estimation to obtain the initial state of the channel, similar to the previous two works. Then, there are data symbols interspersed with pilot symbols, which are used for tracking in a decision-directed manner. The numerical results show that the approaches perform well, but incur a 3 dB to 6 dB performance degradation compared to a system with perfect channel estimation.

2.6.2 Important Results for SFOs

Wu et al. [62] derive an estimator for flat-fading MIMO channels that are oversampled and incur a single SFO, i.e., setup 2. As is common in the SFO literature, it is assumed that the sampling phase process is sufficiently well modeled by a constant for a block of symbols, see Example 3. The first discussion in the work is focused on the optimum sample selection algorithm, originally proposed in [61], which works by correlating the received oversampled signal containing the training signal with said training signal, and choosing the sampling phase as the sample that yields the highest correlation. In order to make the estimate independent of the carrier phase, and lower the impact of other channel influences, the magnitude squared of

the correlation is used for finding the sample phase. Additionally, since the same SFO is experienced at all Rx antennas, the result can be averaged across all training sequences and receive antennas, in order to improve the performance. Naturally, this algorithm cannot estimate sample phases that lie between two sample points and, thus, only achieves a good performance for high values of Q. A better performance is achieved by noting that the log-likelihood function, which is used to find the best sample, is smooth. Essentially, an interpolated version of the discrete log-likelihood function can be generated by employing a discrete Fourier transform (DFT), which can then be used to obtain a more accurate estimate of the sample phase. Similar approaches have also been proposed for SISO systems, e.g., in [71]. Numerical examples show a significant performance gain compared to the optimum sample selection algorithm, and suggest that the algorithm is viable for $Q \geq 4$. The results, furthermore, show that at high SNR, the performance is limited by aliasing of the sampled likelihood function. In other words, a higher Q is needed to improve the estimation performance further in such cases, rather than a higher SNR of the estimation procedure, i.e., a longer training sequence or more receive antennas.

Table 2.3: Summary of research on SFO synchronization for MIMO systems, under different oscillator setups and channel assumptions.

Authors	SFO	Channel	Contribution
Wu et al. [62]	single	flat fading	Correlation-based timing estimator working well with low Q, and analytical MSE expression for it.
Rajawat and Chaturvedi [72]	single	flat fading	Extension of [62], improving performance by exploiting knowledge about the Tx pulse shape.
Nasir et al. [73]	multiple	flat fading	CRB-based training sequence design and performance evaluation.

Nasir et al. [73] investigate optimal training sequences for flat-fading MIMO channels that are oversampled and incur multiple SFOs, i.e., setup 4, also modeled as deterministic constants for a block of symbols. The following conditions for optimal training sequences, in the sense of a hybrid CRB, are obtained: training sequences from every Tx exhibit a π radians phase shift every symbol, they are mutually orthogonal, and the training sequence from any Tx is orthogonal to $\pm T_s$-shifted training sequence from every other Tx. A MAP estimator for channel and SFO is derived, which works similar to the CFO estimators outlined in the previous section, by first estimating the SFO through some search, and then estimating the channel through LS. Simulation results show that a significant performance gain can be achieved when training sequences are used that meet the three conditions approximately, e.g., a certain set of Walsh-Hadamard codes, compared to ones that violate them. It is also seen that a relatively low estimation error is achieved with $Q = 2$.

Table 2.4: Summary of research on joint CFO and SFO synchronization for MIMO systems, under different oscillator setups and channel assumptions.

Authors	CFO	SFO	Channel	Contribution
Naguib et al. [61]	single	single	flat fading	Full MIMO system concept and performance evaluation, correlation-based estimators.
Kannan et al. [74]	multiple	multiple	flat fading	ML-based timing estimator and extended Kalman filter for channel tracking.
Komninakis et al. [75]	multiple	multiple	sel. fading	Kalman-filter-based channel estimation and tracking, as well as MMSE DFE equalization.

2.6.3 Important Results for Both Timing Impairments

As in the previous section, it is assumed for the following works that the SFO processes can be modeled as a constant for a block of samples or symbols.

In Naguib et al. [61], a full modem concept for flat-fading MIMO channels that are oversampled and incur a single CFO and a single SFO, i.e., setup 2 for each of them, is proposed. The frame structure is made up of a long training sequence at the start of each frame, and shorter pilot blocks between the data, both of which are assumed to be mutually orthogonal among different transmitters. SFO compensation is done by the optimum sample selection algorithm discussed in the previous section. CFO estimation is treated as part of the MMSE channel estimation, assuming that there is only a small residual offset of less than 0.1 ppm. In order to gain channel estimates including the CFO for the data part, an interpolation is used between the spread out pilot blocks. Numerical results show the viability of the concept, also including Doppler effects and mildly frequency-selective channels, with a training/pilot overhead of 20%.

Kannan et al. [74] explore flat-fading MIMO channels that are oversampled and incur multiple CFOs and multiple SFOs, i.e., setup 4 for each of them. Due to the complexity, it is proposed to separate the estimation and compensation of SFO and CFO. First, the sampling phase offset is estimated using a training sequence in an ML procedure, by performing a grid search over reasonable values. Compensation is done with a parallel filtering structure, whose weights are suggested to be precomputed for different timing offsets during system design, due to the immense complexity involved in determining them. The timing corrected output signal is then downsampled to symbol rate, and fed into the CFO and channel estimation and compensation part. Here, an extended Kalman filter, which is initialized with a known training signal, is used to deal with both of them at the same time. Decision directed updating of the filter is used during the data part in order to improve the performance. The simulation results show that the approach is feasible, leading to a performance degradation between 3 dB to 6 dB w.r.t. a system with perfect channel estimation and synchronization. They also suggest that sampling

offset errors are less severe compared to CFO/channel errors, and that catastrophic error propagation can occur in the decision directed part for certain SNR regions.

The paper of Komninakis et al. [75] investigates a general time-varying frequency-selective MIMO channel. The MIMO channel is modeled as an autoregressive process, and a Kalman filter that is able to track the time variations based on that model is designed. One drawback of this approach is that the timing impairments are not described explicitly, which can make the model more complex than necessary. A MMSE decision-feedback equalizer (DFE) is proposed for MIMO channel equalization, where due to the time gap between the Kalman filters channel estimates, an additional channel predictor is required. Simulation results show the viability of the approach, but also show that there are significant performance gaps compared to an ideally synchronized system for certain parameter ranges.

2.7 Summary & Main Results

In this chapter, the system fundamentals of frequency-selective MIMO systems with multiple timing impairments have been laid out. It was first shown, based on the general continuous-time system model, that the spatial multiplexing gain and frequency selectivity in LoS MIMO systems occurs due to spherical wave propagation, and depends on the geometric arrangement of the antennas. A general phase process model that can describe the timing impairments of independent, non-ideal oscillators, which are used for carrier and sampling frequency generation in the MIMO transceivers, has been introduced. It consists of a Wiener process with drift, and a white noise process. The possible oscillator configurations that can be used in MIMO systems have been described, and the discrete-time model for the general case with independent oscillators has been derived. It is seen that even for static pure-LoS transmission channels, the complete system including the timing impairments is time varying due to the CFOs, and both time varying and frequency selective due to the SFOs. It was also briefly discussed where the sample drop/add problem for continuous transmission comes from, which will, however, not be treated further in this work.

Finally, relevant results from the literature have been surveyed. It was seen that most of the results focus on one of the two impairments and assume the other one compensated or negligible. One important conclusion is that mutually uncorrelated training sequences from different transmitters simplify the estimation procedure for the channel and timing parameters significantly. Furthermore, if the channel is frequency selective, the training sequences should have a well-localized autocorrelation function to make channel estimation easier. Due to the number of parameters that need to be estimated, the procedure is generally very complex and most works suggest to perform the estimation consecutively. The most common estimation order seems to be: first SFO, then CFO, and finally channel. Correlation-based estimators are shown to yield good performance with reasonable complexity for CFO and channel. In the case of multiple SFOs, the few proposed methods rely on some form of grid search over the possible parameter range, making these approaches rather complex. It is

also seen that tracking of the channel, e.g., in a decision directed manner, can improve the performance during data transmission.

3 Parameter Estimation in MIMO Systems

This chapter deals with estimating the multiple channel and timing parameters that occur in MIMO systems and impair the transmission of data.

3.1 General MAP and ML Data Estimation

Ultimately, the receiver should be optimized in order to minimize the detection error of any sequence of transmitted symbols $\mathbf{x}_{L_\mathrm{T}}$, given the received sample sequence $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}$ [4]. In order to achieve this goal, the following derivations can be used for the MIMO case, similarly to the SISO case discussed in [4]. Generally, the sequence of symbols $\mathbf{x}_{L_\mathrm{T}}$ needs to be found that maximizes the a-posteriori probability, i.e.,

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \left[p\left(\mathbf{x}_{L_\mathrm{T}} | \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\right)\right] = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \left[\frac{p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)}{p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\right)} \right] \tag{3.1}$$

$$= \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right], \tag{3.2}$$

where $\hat{\mathbf{x}}_{L_\mathrm{T}}$ are the estimates of the transmitted symbols and $p(\cdot)$ is a PDF or a probability mass function (PMF). Assuming all possible transmitted symbol sequences $\mathbf{x}_{L_\mathrm{T}}$ to be equally probable, this MAP approach reduces to the ML approach that just depends on the conditional density $p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}\right)$, which describes the connection between input symbols and received samples. In order to generalize the following derivation, a generic form of the oversampled MIMO model, which includes all previously mentioned system setups and where all of the impairing effects are combined into a single time-varying matrix, will now be used.

The connection between L_T input symbols from N transmitters, denoted by the vector $\mathbf{x}_{L_\mathrm{T}}[k] = \begin{bmatrix}\mathbf{x}[k] & \mathbf{x}[k-1] & \cdots & \mathbf{x}[k-L_\mathrm{T}+1]\end{bmatrix}^\mathrm{T}$, and $\dot{L}_\mathrm{R}$ received samples from M receivers, denoted by $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k] = \begin{bmatrix}\dot{\mathbf{y}}[k] & \dot{\mathbf{y}}[k-1] & \cdots & \dot{\mathbf{y}}[k-\dot{L}_\mathrm{R}+1]\end{bmatrix}^\mathrm{T}$, at time k can then be written as

$$\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k] = \mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{L_\mathrm{T}}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k], \tag{3.3}$$

where $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ is a block-Toeplitz matrix containing all effects due to CFOs, SFOs, frequency selectivity due to the channel and Tx/Rx filters, as well as Tx pulse shaping. The structure of $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ will be revisited at a later stage. Furthermore, $\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k] = \left[\tilde{\mathbf{n}}[k] \quad \tilde{\mathbf{n}}[k-1] \quad \cdots \quad \tilde{\mathbf{n}}[k-\dot{L}_\mathrm{R}+1]\right]^\mathrm{T}$ with $\tilde{\mathbf{n}}[k] = \dot{\mathbf{H}}_{\mathrm{Rx}}[k]\dot{\mathbf{n}}_{\dot{L}_\mathrm{Rx}}[k]$, which is a complex Gaussian noise process that is shaped by the receive filters, see also section 2.5.7. The PDF of the noise is a multivariate complex[1] Gaussian, given by

$$p\left(\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}\right) = \frac{1}{\det\left(\pi\mathbf{C}_{\tilde{\mathbf{n}}}\right)} \exp\left(-\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}\right), \tag{3.4}$$

where the noise is assumed to have zero mean, i.e., $\mathrm{E}[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}] = \mathbf{0}_{M\dot{L}_\mathrm{R}\times 1}$ based on $\mathrm{E}[\dot{\mathbf{n}}_{\dot{L}_\mathrm{Rx}}[k]] = \mathbf{0}_{M\dot{L}_\mathrm{Rx}\times 1}$ with $\mathrm{E}[\dot{n}_m[k]] = 0$. To simplify the notation, the time dependence of the PDF has been omitted here. It will be seen in the next section that the noise correlation $\mathbf{C}_{\tilde{\mathbf{n}}} = \mathrm{E}\left[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}^\mathrm{H}[k]\right]$ varies with time due to the variation caused by the SFO. For a given trial value of $\mathbf{x}_{L_\mathrm{T}}$, the variable transformation $p\left(\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}\right) = p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}}\mathbf{x}_{L_\mathrm{T}}\right)$ yields the conditional density

$$p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}|\mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) = \frac{1}{\det\left(\pi\mathbf{C}_{\tilde{\mathbf{n}}}\right)} \exp\left(-\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}}\mathbf{x}_{L_\mathrm{T}}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}}\mathbf{x}_{L_\mathrm{T}}\right)\right). \tag{3.5}$$

Like for the correlation matrix $\mathbf{C}_{\tilde{\mathbf{n}}}$, the combined MIMO channel matrix $\mathbf{H}_{\tilde{\mathrm{C}}}$ is also time varying, due to the CFOs and SFOs, but the dependence is omitted in the notation for now. Note that compared to (3.2), this PDF is additionally dependent on the channel parameters $\mathbf{H}_{\tilde{\mathrm{C}}}$. With respect to optimal detection, those parameters should, thus, be seen as unwanted or a nuisance, and can be removed through marginalization, i.e.,

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \left[p\left(\mathbf{x}_{L_\mathrm{T}}\right) p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}|\mathbf{x}_{L_\mathrm{T}}\right)\right] \tag{3.6}$$

$$= \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \left[p\left(\mathbf{x}_{L_\mathrm{T}}\right) \int p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}|\mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \mathrm{d}\mathbf{H}_{\tilde{\mathrm{C}}}\right], \tag{3.7}$$

where $p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right)$ is the joint PDF of the MIMO channel parameters describing the potential a-priori knowledge about their statistical properties [4]. This optimal processing structure is very complex because the previous integration needs to be computed for all possible values of $\mathbf{H}_{\tilde{\mathrm{C}}}$ for each trial sequence $\mathbf{x}_{L_\mathrm{T}}$, especially considering that the parameters are time varying in general.

Another approach is to view the channel parameters as desired parameters and maximize the probability w.r.t. the symbols and the MIMO channel, i.e.,

$$\left[\hat{\mathbf{x}}_{L_\mathrm{T}}, \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\right] = \underset{\mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}|\mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right)\right], \tag{3.8}$$

where independence between data and channel is implied. It is seen that in this approach

[1] To be precise, this complex form of the distribution only contains all information about the signal if it is circular, i.e., real and imaginary part are uncorrelated and of equal variance [28, 76]. This property will be encountered again in chapter 5, which is concerned with practical implementation and where IQ imbalance is of importance.

there is no separate channel estimation or synchronization unit. Instead, joint estimation and detection needs to be carried out. In most cases this procedure cannot be implemented, due to the enormous parameter range of $\mathbf{H}_{\tilde{\mathrm{C}}}$. On the other hand, the possible values of $\mathbf{x}_{L_\mathrm{T}}$ are known, based on the type of modulation. Thus, a possible, but suboptimal strategy, because we may get stuck in local maxima, is to proceed as follows

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}(\mathbf{x}_{L_\mathrm{T}}) = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \; \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right] \tag{3.9}$$

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \; \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}(\mathbf{x}_{L_\mathrm{T}})\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right) \right], \tag{3.10}$$

where an estimate of the MIMO channel parameters $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}(\mathbf{x}_{L_\mathrm{T}})$ for each possible symbol sequence is first derived, which is then used in the second step as if it were the true value in order to compute the symbol estimates. Still, computing channel estimates for each possible symbol sequence is very complex, especially for time-varying MIMO systems. However, this approach leads towards separating the tasks of estimating the MIMO channel parameters $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$, and detecting (or estimating) the symbol sequence $\hat{\mathbf{x}}_{L_\mathrm{T}}$.

To be precise, the detection is performed with a channel parameter estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$, assuming that this estimate is close to the actual value of the channel and independent of the transmitted symbol sequence, i.e.,

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \; \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right) \right]. \tag{3.11}$$

Solving this problem will be treated further in the next chapter. For now, the focus is on obtaining the estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$. There are two approaches that are commonly taken to solve this problem. The first one, often referred to as data-aided (DA) estimation, works by doing the same thing for the channel estimation procedure that was done for the data detection above. Specifically,

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \; \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}} = \mathbf{x}_\mathrm{Tr}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right], \tag{3.12}$$

where $\mathbf{x}_\mathrm{Tr}$ is a training sequence[2] of known symbols that is sent separately from the data in order to identify the channel. If there is no prior information about the distribution of the channel parameters, which is often the case in practice, $p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right)$ can be omitted from the maximization. This implicitly means that all channel realizations are equally likely. A related approach is called decision-directed (DD) estimation, where decoded data symbols are used instead of a training sequence in (3.12). For this approach to work, it is necessary that the decoded data is correct with a high reliability. It is helpful for systems where the channel parameters vary with time, e.g., due to CFO and SFO, so that the estimate can be updated

[2] In general, the channel estimate depends on the training sequence, i.e., $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\left(\mathbf{x}_\mathrm{Tr}\right)$. With a suitable choice of training sequence, however, this dependence can be almost eliminated. It is also important that the training signal resembles the data signal that ought to be transmitted in certain aspects, e.g., in terms of occupied bandwidth.

without retransmission of a training sequence.

The second approach for estimating $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$ works similarly to (3.7), but instead of marginalizing out the channel, the data is removed. This is done as follows

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \sum_{\forall \mathbf{x}_{L_\mathrm{T}}} p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right) \right] \tag{3.13}$$

$$= \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{H}_{\tilde{\mathrm{C}}}\right) \right]. \tag{3.14}$$

These approaches are called non-data-aided (NDA) or blind.

3.2 DA Channel Estimation

The problem of estimating the channel, for which a solution is sought, is given for Gaussian noise by

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}} = \mathbf{x}_\mathrm{Tr}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right] \tag{3.15}$$

$$= \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[\frac{1}{\det\left(\pi \mathbf{C}_{\tilde{\mathbf{n}}}\right)} \exp\left(- \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr}\right) \right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right], \tag{3.16}$$

assuming that the training signal $\mathbf{x}_\mathrm{Tr}$ is known. This problem can be simplified by taking the logarithm of the PDF, i.e.,

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[\log\left(p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}} = \mathbf{x}_\mathrm{Tr}, \mathbf{H}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right) \right] \tag{3.17}$$

$$= \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \left[\log \frac{1}{\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}})} - \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{T}\right) + \log p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \right] \tag{3.18}$$

$$\begin{aligned} = \underset{\mathbf{H}_{\tilde{\mathrm{C}}}}{\arg\max} \Bigg[& \log \frac{1}{\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}})} - \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} + \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr} \\ & + \left(\mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \left(\mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{H}_{\tilde{\mathrm{C}}} \mathbf{x}_\mathrm{Tr} + \log p\left(\mathbf{H}_{\tilde{\mathrm{C}}}\right) \Bigg]. \end{aligned} \tag{3.19}$$

It is seen that in order to compute the above, knowledge about $\mathbf{C}_{\tilde{\mathbf{n}}}$ is required. This correlation matrix is given through

$$\mathbf{C}_{\tilde{\mathbf{n}}}[k] = \mathrm{E}\left[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k] \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}^\mathrm{H}[k] \right] \tag{3.20}$$

$$= \mathrm{E}\left[\left(\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \dot{\mathbf{n}}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}}[k] \right) \left(\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \dot{\mathbf{n}}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}}[k] \right)^\mathrm{H} \right] \tag{3.21}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \, \mathrm{E}\left[\dot{\mathbf{n}}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}}[k] \dot{\mathbf{n}}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}}^\mathrm{H}[k] \right] \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}^\mathrm{H}[k] \tag{3.22}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \cdot \dot{\mathbf{N}} \cdot \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}^\mathrm{H}[k], \tag{3.23}$$

where $\dot{\mathbf{N}} = \mathbf{I}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}} \otimes \mathrm{diag}(\sigma^2_{\dot{n}_1}, \sigma^2_{\dot{n}_2}, \ldots, \sigma^2_{\dot{n}_M})$ contains the noise powers of the different receivers, i.e., $\sigma^2_{\dot{n}_m} = \mathrm{E}[|\dot{n}_m[k]|^2]$, which are assumed to be constant over time and are often also assumed to be equal between different receivers. The length $\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}$ is defined by $\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}} = \dot{L}_{\mathrm{R}} + \dot{L}_{\mathrm{Rx}} - 1$, consisting of the length of the received sample sequences $\dot{L}_{\mathrm{R}}$ and the length of the receive filters $\dot{L}_{\mathrm{Rx}}$. Furthermore, the combined receive filter matrix $\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]$ is given by

$$\underset{(M\dot{L}_{\mathrm{R}} \times M\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}})}{\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]} = \begin{bmatrix} \dot{\mathbf{H}}_{\mathrm{Rx}}[k] & \mathbf{0}_{M\times M(\dot{L}_{\mathrm{R}}-1)} & & \\ \mathbf{0}_{M\times M} & \dot{\mathbf{H}}_{\mathrm{Rx}}[k] & \mathbf{0}_{M\times M(\dot{L}_{\mathrm{R}}-2)} & \\ & & \ddots & \\ & & \mathbf{0}_{M\times M(\dot{L}_{\mathrm{R}}-1)} & \dot{\mathbf{H}}_{\mathrm{Rx}}[k] \end{bmatrix}, \tag{3.24}$$

which is a block-Toeplitz matrix of the receive filters, see section 2.5.3 for their definition. Furthermore, the first two terms in (3.19) can in general not be omitted, because they are dependent on a part of the complete channel through $\dot{\mathbf{H}}_{\mathrm{Rx}}[k]$. Let us briefly revisit why the discrete-time noise correlation matrix $\mathbf{C}_{\tilde{\mathbf{n}}}[k]$ is time varying, when SFOs are present. In continuous time, the noise process is shaped by the linear analog filter $h_{\mathrm{Rx},m}(t)$ of each receiver front end. The statistics of this continuous-time colored-noise process are time invariant. However, when this process is sampled with a non-ideal sampling phase process, the statistics of the discrete-time colored-noise process change over time, due to the non-flat characteristic of the colored noise. The noise process can, thus, be equivalently modeled in the discrete-time domain with a time-invariant, in terms of its statistics, white Gaussian noise process $n_m[k]$, and a discrete-time low-pass filter $h_{\mathrm{Rx},m}\left[k + \dot{\phi}_{\mathrm{Rx},m}[k]\right]$, whose coefficients slightly change according to the sampling phase process.

The maximization of (3.19) requires a search over the continuous parameter space of $\mathbf{H}_{\tilde{\mathrm{C}}}$, which is practically infeasible for most systems. In order to continue with the ML derivation, the term $\mathbf{H}_{\tilde{\mathrm{C}}}\mathbf{x}_{\mathrm{Tr}}$ will be restructured to $\mathbf{X}_{\mathrm{Tr}}\dot{\mathbf{h}}_{\mathrm{C}}$, which consists of a matrix containing the training signals, and a vector containing all channel parameters. The estimation problem then becomes

$$\hat{\dot{\mathbf{h}}}_{\mathrm{C}} = \underset{\dot{\mathbf{h}}_{\mathrm{C}}}{\arg\max} \left[\log \frac{1}{\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}})} - \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}}\dot{\mathbf{h}}_{\mathrm{C}}\right)^{\mathrm{H}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}}\dot{\mathbf{h}}_{\mathrm{C}}\right) + \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right) \right] \tag{3.25}$$

$$\begin{aligned} = \underset{\dot{\mathbf{h}}_{\mathrm{C}}}{\arg\max} \Bigg[& \log \frac{1}{\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}})} - \dot{\mathbf{y}}^{\mathrm{H}}_{\dot{L}_{\mathrm{R}}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} + \dot{\mathbf{y}}^{\mathrm{H}}_{\dot{L}_{\mathrm{R}}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} \\ & + \dot{\mathbf{h}}^{\mathrm{H}}_{\mathrm{C}} \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \dot{\mathbf{h}}^{\mathrm{H}}_{\mathrm{C}} \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} + \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right) \Bigg]. \end{aligned} \tag{3.26}$$

$$\begin{aligned} = \underset{\dot{\mathbf{h}}_{\mathrm{C}}}{\arg\min} \Bigg[& \log\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}}) + \dot{\mathbf{y}}^{\mathrm{H}}_{\dot{L}_{\mathrm{R}}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \dot{\mathbf{y}}^{\mathrm{H}}_{\dot{L}_{\mathrm{R}}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} \\ & - \dot{\mathbf{h}}^{\mathrm{H}}_{\mathrm{C}} \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} + \dot{\mathbf{h}}^{\mathrm{H}}_{\mathrm{C}} \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr}} \mathbf{C}^{-1}_{\tilde{\mathbf{n}}} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} - \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right) \Bigg]. \end{aligned} \tag{3.27}$$

A way of minimizing this equation is to take the partial derivatives w.r.t. the desired channel

parameters, and set them equal to zero[3], i.e.,

$$\mathbf{0} = \frac{\partial}{\partial \dot{\mathbf{h}}_{\mathrm{C}}} \log\left(p\left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} | \mathbf{x}_{L_{\mathrm{T}}} = \mathbf{x}_{\mathrm{Tr}}, \dot{\mathbf{h}}_{\mathrm{C}}\right) p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right)\right) \tag{3.28}$$

$$\begin{aligned} &= \frac{\partial}{\partial \dot{\mathbf{h}}_{\mathrm{C}}} \big(\log\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}}) + \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} \\ &\qquad - \dot{\mathbf{h}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} + \dot{\mathbf{h}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} - \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right)\big) \end{aligned} \tag{3.29}$$

$$\approx -\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} + \dot{\mathbf{h}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} - \frac{\partial}{\partial \dot{\mathbf{h}}_{\mathrm{C}}} \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right), \tag{3.30}$$

where for (3.30) it was assumed that $\mathbf{C}_{\tilde{\mathbf{n}}}^{-1}$ is independent of $\dot{\mathbf{h}}_{\mathrm{C}}$. Although not exact, this approximation is well justified, when the receive filters have limited impact on the complete transmission characteristic, i.e., $h_{\mathrm{Rx},m}(t) \approx \mathrm{sinc}\left(\pi \frac{t}{T_{\mathrm{nom,Rx}}}\right)$.

It is now further assumed that there is no information about the prior distribution of the channel parameters $p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right)$, which leads to

$$\dot{\mathbf{h}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} = \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} \tag{3.31}$$

$$\mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}} = \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} \tag{3.32}$$

$$\implies \hat{\dot{\mathbf{h}}}_{\mathrm{C}}[k] = \left(\mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k] \mathbf{X}_{\mathrm{Tr}}\right)^{-1} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k], \tag{3.33}$$

with $(\mathbf{C}_{\tilde{\mathbf{n}}}^{-1})^{\mathrm{H}} = (\mathbf{C}_{\tilde{\mathbf{n}}}^{\mathrm{H}})^{-1} = \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}$ since $\mathbf{C}_{\tilde{\mathbf{n}}}$ is a Hermitian matrix, and accounting for the time dependence of the parameter estimates due to the timing impairments.

Similarly, an ML estimator for the noise correlation matrix can be obtained, by using the same method and assumptions, with

$$\mathbf{0} = \frac{\partial}{\partial \mathbf{C}_{\tilde{\mathbf{n}}}} \left(\log\det(\pi \mathbf{C}_{\tilde{\mathbf{n}}}) + \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}\right)^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}\right) - \log p\left(\dot{\mathbf{h}}_{\mathrm{C}}\right)\right) \tag{3.34}$$

$$\approx \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} - \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}\right) \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}} - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}\right)^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \tag{3.35}$$

$$\implies \hat{\mathbf{C}}_{\tilde{\mathbf{n}}}[k] = \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k] - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}[k]\right) \left(\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k] - \mathbf{X}_{\mathrm{Tr}} \dot{\mathbf{h}}_{\mathrm{C}}[k]\right)^{\mathrm{H}}, \tag{3.36}$$

using again the Hermitian property of $\mathbf{C}_{\tilde{\mathbf{n}}}$. It is seen that the estimators of (3.33) and (3.36) influence each other in general. In order to decouple them, we can consider a part of the training signals, where a zero signal is transmitted from all Tx simultaneously, i.e., $\mathbf{X}_{\mathrm{Tr}} = \mathbf{0}$ or $\mathbf{x}_{\mathrm{Tr}} = \mathbf{0}$. In that case, the estimator becomes $\hat{\mathbf{C}}_{\tilde{\mathbf{n}}}[k] = \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k] \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}}[k] = \tilde{\mathbf{n}}_{\dot{L}_{\mathrm{R}}}[k] \tilde{\mathbf{n}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}}[k]$, and averaging over several of these noise observations, i.e., $\hat{\mathbf{C}}_{\tilde{\mathbf{n}}}[k] = \frac{1}{L_{\mathrm{I}}} \sum_{l_{\mathrm{I}}} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k + l_{\mathrm{I}}] \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}^{\mathrm{H}}[k + l_{\mathrm{I}}]$, will give a good estimate. Averaging, and using the estimate for the estimation in (3.33) only makes sense if the correlation matrix is approximately time invariant over the corresponding time frame, i.e., $\mathbf{C}_{\tilde{\mathbf{n}}}[k] \approx \mathbf{C}_{\tilde{\mathbf{n}}}[k + l_{\mathrm{I}}]$ with $l_{\mathrm{I}} = \{1, \ldots, L_{\mathrm{I}} - 1\}$ being the time frame in samples. From here on

[3] Technically, the second derivative of the equation also needs to be checked.

forward, it will be assumed that the noise correlation matrix is known, or estimated accurately with the method just described.

3.2.1 Basics for Training Signals

DA channel estimation requires suitable training signals $\mathbf{x}_{\mathrm{Tr}}$ of length L_{Tr} to identify the MIMO channel parameters. Designing such signals is a wide area of research [77, 78, 79, 80], as it can lower estimation complexity and improve estimation performance, confer section 2.6. The following assumptions will be made about the training signals that will be used for parameter estimation:

1. Training signals $\mathbf{x}_{\mathrm{Tr}}$ are time multiplexed with data signals $\mathbf{x}[k]$
2. Each transmitter has a training signal $\mathbf{x}_{\mathrm{Tr},n}$ that is uncorrelated to all of the other transmitters training signals, i.e., $\sum_{l_{\mathrm{Tr}}} x_{\mathrm{Tr},n_1}[l_{\mathrm{Tr}}+l]x^*_{\mathrm{Tr},n_2}[l_{\mathrm{Tr}}] = 0$ or $\mathrm{E}\left[x_{\mathrm{Tr},n_1}[l_{\mathrm{Tr}}+l]x^*_{\mathrm{Tr},n_2}[l_{\mathrm{Tr}}]\right] = 0$ for $\forall l, \forall n_1 \neq n_2$
3. Training signals have an ideal (impulse-like) autocorrelation function, or are uncorrelated to time shifted versions of themselves on the symbol time grid, i.e., $\sum_{l_{\mathrm{Tr}}} x_{\mathrm{Tr},n}[l_{\mathrm{Tr}}+l]x^*_{\mathrm{Tr},n}[l_{\mathrm{Tr}}] = 0$ for $\forall l \neq 0, \forall n$, or $\mathrm{E}\left[x_{\mathrm{Tr},n}[l_{\mathrm{Tr}}+l]x^*_{\mathrm{Tr},n}[l_{\mathrm{Tr}}]\right] = \sigma^2_{x_{\mathrm{Tr},n}} \cdot \delta[l]$
4. Training signals are approximately uncorrelated to the data sections, i.e., $\sum_{l_{\mathrm{Tr}}} x_{n_1}[l_{\mathrm{Tr}}+k]x^*_{\mathrm{Tr},n_2}[l_{\mathrm{Tr}}] \approx 0$ for $\forall k, \forall n_1, n_2$, or $\mathrm{E}\left[x_{n_1}[l_{\mathrm{Tr}}+k]x^*_{\mathrm{Tr},n_2}[l_{\mathrm{Tr}}]\right] = 0$

A general treatment of signals with good correlation properties can be found in [81]. Some sequences that work good in practice and approximately fulfill those assumptions are: pseudo-noise- and m-sequences, as well as Zadoff-Chu sequences.

A final question that needs to be answered is how long training sequences need to be in order to identify the parameters. On the fundamental level, the channel has $M \cdot N \cdot \dot{L}_{\mathrm{C}}$ parameters, which change at every time step k due to the timing impairments. Clearly, in this case we need to continuously transmit training signals to keep track of the ever changing channel parameters. It is now assumed that the time variations are slow w.r.t. to the symbol duration, meaning that the channel is approximately time invariant over a certain time frame, i.e., $\dot{\mathbf{H}}_{\mathrm{C}}[k] \approx \dot{\mathbf{H}}_{\mathrm{C}}[k+l_{\mathrm{I}}]$ see appendix A.2. In such a frame, where the channel is approximately time invariant, at least $(N+1)\cdot\dot{L}_{\mathrm{C}}-1$ training symbols are required [82] from every Tx in order to uniquely identify the responses to all Rx. A simple scheme is obtained by time multiplexing a training signal $\mathbf{x}_{\mathrm{Tr,A}}$ with an ideal autocorrelation function, see the following Example 4.

Example 4. *Consider a small MIMO system with $M = N = 2$ and $\dot{L}_C = 2$, which is approximately time invariant with $\dot{\mathbf{H}}_C[k] = \dot{\mathbf{H}}_C[k+1] = \dot{\mathbf{H}}_C[k+2] = \dot{\mathbf{H}}_C[k+3]$, i.e., $L_I = 4$. Furthermore, assume the noise to be negligible. Accordingly, 5 training symbols are required. The training signals from the two antennas can be chosen as $\mathbf{x}_{Tr,1} = \begin{bmatrix}0 & \mathbf{x}^{\mathrm{T}}_{Tr,A} & 0 & 0\end{bmatrix}^{\mathrm{T}}$, $\mathbf{x}_{Tr,2} = \begin{bmatrix}0 & 0 & 0 & \mathbf{x}^{\mathrm{T}}_{Tr,A}\end{bmatrix}^{\mathrm{T}}$, where*

$\mathbf{x}_{Tr,A} = \begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathrm{T}}$ *is chosen in this example. Transmitting these training signals gives the following*

$$\begin{aligned} \dot{\mathbf{y}}[k] &= \dot{\mathbf{H}}_C[k]\mathbf{x}_{L_C}[k] = \dot{\mathbf{H}}_C[k]\mathbf{x}_{Tr}[0] \\ \dot{\mathbf{y}}[k+1] &= \dot{\mathbf{H}}_C[k+1]\mathbf{x}_{L_C}[k+1] = \dot{\mathbf{H}}_C[k+1]\mathbf{x}_{Tr}[1] \doteq \dot{\mathbf{H}}_C[k]\mathbf{x}_{Tr}[1] \\ &\vdots \end{aligned}$$

This can be written in a more compact form as

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{y}}[k] & \dot{\mathbf{y}}[k+1] & \dot{\mathbf{y}}[k+2] & \dot{\mathbf{y}}[k+3] \end{bmatrix} &= \dot{\mathbf{H}}_C[k] \begin{bmatrix} \mathbf{x}_{Tr}[0] & \mathbf{x}_{Tr}[1] & \mathbf{x}_{Tr}[2] & \mathbf{x}_{Tr}[3] \end{bmatrix} \\ &= \dot{\mathbf{H}}_C[k] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

which means that organizing the received samples as $\begin{bmatrix} \dot{\mathbf{y}}[k] & \dot{\mathbf{y}}[k+2] & \dot{\mathbf{y}}[k+1] & \dot{\mathbf{y}}[k+3] \end{bmatrix}$ *directly gives an estimate of the channel matrix for this training signal structure. Note that the first zeros in these particular training sequences, and the first* $\dot{L}_C - 1$ *symbols in general, are necessary to have known initial states of the channels.*

In practice, we would not choose such impulse-like training signals as in the example due to their bad envelope characteristics. However, any training sequence $\mathbf{x}_{\mathrm{Tr,A}}$ with ideal autocorrelation properties can be time multiplexed in this fashion to estimate the MIMO channel in a simple manner.

Some trade-offs and comments about training signals should be mentioned. In general, longer training sequences, i.e., allocating more power towards training, improves the quality of the estimates, but lowers the overall data transmission rate of the link. However, when the timing impairments influence[4] the correlation with the training signals, increasing the length of the training signals will not always improve estimation performance. On the other hand, the channel estimate obsolesces due to the timing impairments, which means that frequent reestimation is required given that there is no structure for the timing impairments. Finally, it was assumed that the training signals are sent from all transmitters simultaneously. For distributed systems and also for independent sample clock generation on the Tx side, this may not generally be the case. With training signals that are uncorrelated to the data, we can still estimate the MIMO channels, but observe the responses from different transmitters at different points in time. Whether or not they can be combined in a meaningful way depends on the time variation of the channels between the observation points.

[4] This is important when the time variation due to CFO and SFO is significant w.r.t. to the length of the training signal.

3.2.2 On the Need for DA Estimation in MIMO Systems

Blind methods have been mentioned as a means for parameter estimation at the beginning of this chapter. The problem with fully blind receiver structures is that, aside from the complexity that they often entail, they also suffer from a phase ambiguity, which needs to be somehow resolved. For MIMO systems this problem is further enhanced, as each received stream suffers from this ambiguity. Thus, as long as no additional coding across the streams is considered, some training signals are always required to resolve the ambiguities. In the following, the focus will, hence, be on transmission schemes using training signals. Nevertheless, the estimation of certain parameters in the system, such as CFOs and SFOs, can be carried out blind for some oscillator setups, as will be briefly discussed in chapter 4.

3.3 Estimators for Different MIMO Setups

In the following, MIMO system setups under different timing impairment conditions will be investigated, and the simplifications of the ML channel parameter estimators will be shown in these cases. Only oscillator setup 4, i.e., independent phase processes in each Tx/Rx chain for carrier and sampling phase generation, respectively, will be considered, as it encompasses all the other cases. For a discussion on the applicability of other known timing estimators to MIMO systems with different oscillator setups, refer to section 4.2 of the next chapter.

3.3.1 Synchronized Symbol-Spaced Model

When no timing impairments are present, the MIMO channel parameters do not depend on k in the case of a static link setup, e.g., a backhaul scenario. The combined MIMO channel matrix is given for symbol rate sampling ($Q = 1$) as

$$\mathbf{H}_{\tilde{\mathrm{C}}} = \begin{bmatrix} \mathbf{H}_{\mathrm{C}} & \mathbf{0}_{M\times N(L_{\mathrm{R}}-1)} & & \\ \mathbf{0}_{M\times N} & \mathbf{H}_{\mathrm{C}} & \mathbf{0}_{M\times N(L_{\mathrm{R}}-2)} & \\ & & \ddots & \\ & & \mathbf{0}_{M\times N(L_{\mathrm{R}}-1)} & \mathbf{H}_{\mathrm{C}} \end{bmatrix},$$

which is a block-Toeplitz matrix, depending only on the time-invariant transfer characteristic $\mathbf{H}_{\mathrm{C}} = \mathbf{H}_{\mathrm{Rx}}\mathbf{H}_{L_{\mathrm{Rx}}}\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}$. The vector of the desired channel parameters is given by

$$\mathbf{h}_{\mathrm{C}} = \mathrm{vec}\left(\mathbf{H}_{\mathrm{C}}^{\mathrm{T}}\right) = \begin{bmatrix} \mathbf{h}_{\mathrm{C},L_{\mathrm{C}},1}^{\mathrm{T}} & \mathbf{h}_{\mathrm{C},L_{\mathrm{C}},2}^{\mathrm{T}} & \cdots & \mathbf{h}_{\mathrm{C},L_{\mathrm{C}},M}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}},$$

and the training signal matrix follows from the convention in (3.3) as

$$\mathbf{X}_{\mathrm{Tr}} = \begin{bmatrix} \mathbf{I}_M \otimes \begin{bmatrix} \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{T}}-1] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{Tr}}-L_{\mathrm{C}}] \end{bmatrix} \\ \vdots \\ \mathbf{I}_M \otimes \begin{bmatrix} \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{C}}] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[1] \end{bmatrix} \\ \mathbf{I}_M \otimes \begin{bmatrix} \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{C}}-1] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[0] \end{bmatrix} \end{bmatrix},$$

with $\mathbf{x}_{\mathrm{Tr}}[l] = \begin{bmatrix} x_{\mathrm{Tr},1}[l] & x_{\mathrm{Tr},2}[l] & \cdots & x_{\mathrm{Tr},N}[l] \end{bmatrix}^{\mathrm{T}}$, and from the previous section $L_{\mathrm{Tr}} \geq (N+1) \cdot L_{\mathrm{C}} - 1$. As is pointed out in the literature and also seen here from the Kronecker structure, the estimation is decoupled for every receive stream m. This means that the channel estimation can be carried out at each receive stream in parallel[5] from the following

$$\underbrace{\begin{bmatrix} y_m[k] \\ \vdots \\ y_m[k-L_{\mathrm{R}}+1] \end{bmatrix}}_{\mathbf{y}_{L_{\mathrm{R}},m}[k]} = \underbrace{\begin{bmatrix} \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{Tr}}-1] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{Tr}}-L_{\mathrm{C}}] \\ \vdots & & \vdots \\ \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{C}}] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[1] \\ \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[L_{\mathrm{C}}-1] & \cdots & \mathbf{x}_{\mathrm{Tr}}^{\mathrm{T}}[0] \end{bmatrix}}_{\mathbf{X}_{\mathrm{Tr},M=1}} \mathbf{h}_{\mathrm{C},L_{\mathrm{C}},m} + \underbrace{\begin{bmatrix} \tilde{n}_m[k] \\ \vdots \\ \tilde{n}_m[k-L_{\mathrm{R}}+1] \end{bmatrix}}_{\tilde{\mathbf{n}}_{L_{\mathrm{R}},m}[k]}, \quad (3.37)$$

where $L_{\mathrm{R}} = L_{\mathrm{Tr}} - L_{\mathrm{C}} + 1$. With the estimate of the noise correlation matrix $\hat{\mathbf{C}}_{\tilde{\mathbf{n}}}$, this gives the estimator[6] for the complete system, or for each receive antenna as

$$\hat{\mathbf{h}}_{\mathrm{C}}[k] = \left(\mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \hat{\mathbf{C}}_{\tilde{\mathbf{n}}}^{-1} \mathbf{X}_{\mathrm{Tr}}\right)^{-1} \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \hat{\mathbf{C}}_{\tilde{\mathbf{n}}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{R}}}[k], \quad (3.38)$$

or

$$\hat{\mathbf{h}}_{\mathrm{C},L_{\mathrm{C}},m}[k] = \left(\mathbf{X}_{\mathrm{Tr},M=1}^{\mathrm{H}} \hat{\mathbf{C}}_{\tilde{\mathbf{n}}_m}^{-1} \mathbf{X}_{\mathrm{Tr},M=1}\right)^{-1} \mathbf{X}_{\mathrm{Tr},M=1}^{\mathrm{H}} \hat{\mathbf{C}}_{\tilde{\mathbf{n}}_m}^{-1} \cdot \mathbf{y}_{L_{\mathrm{R}},m}[k]. \quad (3.39)$$

In general, no further simplifications are possible on these estimators due to the structure of $\mathbf{C}_{\tilde{\mathbf{n}}}$.

When the receive filter characteristic is approximately ideal, i.e. $h_{\mathrm{Rx},m}(t) \approx \mathrm{sinc}\left(\pi \frac{t}{T_{\mathrm{nom,Rx}}}\right)$ meaning that the noise correlation matrix becomes diagonal, and using the assumptions about the cross- and autocorrelation properties of the training sequences from the previous section, i.e., $\mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr}} = \mathbf{I}_{MNL_{\mathrm{C}}}$ and $\mathbf{X}_{\mathrm{Tr},M=1}^{\mathrm{H}} \mathbf{X}_{\mathrm{Tr},M=1} = \mathbf{I}_{NL_{\mathrm{C}}}$, the estimators simplify further to

$$\hat{\mathbf{h}}_{\mathrm{C}}[k] = \mathbf{X}_{\mathrm{Tr}}^{\mathrm{H}} \cdot \mathbf{y}_{L_{\mathrm{R}}}[k], \quad (3.40)$$

[5] This is only possible if the noise correlation matrix $\mathbf{C}_{\tilde{\mathbf{n}}}$ can also be decoupled for each received stream, which holds true.

[6] The estimates of the channels in the right order are only obtained if the used received sample set $\mathbf{y}_{L_{\mathrm{R}}}[k]$ matches with the arrangement of the training signal matrix $\mathbf{X}_{\mathrm{Tr}}$. This requires some form of finding the training sequences in the received sample stream, which can, for example, be achieved with correlation and peak detection. Furthermore, the estimate varies with time as it depends on the current observation of the output $\mathbf{y}_{L_{\mathrm{R}}}[k]$, which varies with time due to the noise.

and

$$\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},m}[k] = \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr},M=1} \cdot \mathbf{y}_{L_\mathrm{R},m}[k]. \tag{3.41}$$

In other words, in this case the estimation for each mnth channel characteristic is done by using the mth received stream and the training signal from the nth antenna, and performing a matched filtering (or correlation) between the two. Note that this is an ML estimator only when the aforementioned conditions can be met.

3.3.2 Symbol-Spaced Model Containing Only CFOs

When only CFOs are considered, the MIMO channel varies according to the combined effects of the different carrier phase processes. The combined channel matrix is then, for the symbol rate case, given as

$$\mathbf{H}_{\tilde{\mathrm{C}}}[k] = \begin{bmatrix} \mathbf{H}_\mathrm{C}[k] & \mathbf{0}_{M\times N(L_\mathrm{R}-1)} & & \\ \mathbf{0}_{M\times N} & \mathbf{H}_\mathrm{C}[k] & \mathbf{0}_{M\times N(L_\mathrm{R}-2)} & \\ & & \ddots & \\ & & \mathbf{0}_{M\times N(L_\mathrm{R}-1)} & \mathbf{H}_\mathrm{C}[k] \end{bmatrix},$$

which is again a block-Toeplitz matrix, but depending on the time-variant transfer characteristic $\mathbf{H}_\mathrm{C}[k] = \mathbf{H}_{\mathrm{Rx}}\boldsymbol{\Phi}_{\mathrm{Rx},L_\mathrm{Rx}}[k]\mathbf{H}_{L_\mathrm{Rx}}\boldsymbol{\Phi}_{\mathrm{Tx},L_\mathrm{D}}[k]\mathbf{H}_{\mathrm{Tx},L_\mathrm{D}}$. The vector of the desired channel parameters can be given as

$$\mathbf{h}_\mathrm{C}[k] = \mathrm{vec}\left(\mathbf{H}^{\mathrm{T}}_\mathrm{C}[k]\right) = \left[\mathbf{h}^{\mathrm{T}}_{\mathrm{C},L_\mathrm{C},1}[k] \quad \mathbf{h}^{\mathrm{T}}_{\mathrm{C},L_\mathrm{C},2}[k] \quad \cdots \quad \mathbf{h}^{\mathrm{T}}_{\mathrm{C},L_\mathrm{C},M}[k]\right]^{\mathrm{T}}.$$

It is seen that the channel parameters change every time step k due to the carrier phase processes, which makes estimation difficult. Continue by assuming that the channel is approximately time invariant w.r.t. the length of the training sequence[7], i.e., $\mathbf{h}^{\mathrm{T}}_{\mathrm{C},L_\mathrm{C},1}[k] \approx \mathbf{h}^{\mathrm{T}}_{\mathrm{C},L_\mathrm{C},1}[k+l_{\mathrm{Tr}}]$, or $\boldsymbol{\phi}_{\mathrm{Tx}}[k] \approx \boldsymbol{\phi}_{\mathrm{Tx}}[k+l_{\mathrm{Tr}}]$ and $\boldsymbol{\phi}_{\mathrm{Rx}}[k] \approx \boldsymbol{\phi}_{\mathrm{Rx}}[k+l_{\mathrm{Tr}}]$ for $l_{\mathrm{Tr}} = \{1,\ldots,L_{\mathrm{Tr}}-1\}$. This leads to a very similar signal model as above, i.e.,

$$\mathbf{y}_{L_\mathrm{R},m}[k] = \mathbf{X}_{\mathrm{Tr},M=1}\mathbf{h}_{\mathrm{C},L_\mathrm{C},m}[k] + \tilde{\mathbf{n}}_{L_\mathrm{R},m}[k], \tag{3.42}$$

and, thus, the same estimators as above can be used, e.g.,

$$\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},m}[k] = \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr},M=1} \cdot \mathbf{y}_{L_\mathrm{R},m}[k], \tag{3.43}$$

but where the channel vector estimate now varies due to measurement noise as well as the CFOs.

When looking at the variation of the channel estimate for this model over time, investigated

[7] There are other results in the literature that do not require this assumption, but require the phase processes to be linearly increasing, or, in other words, linearly drifting [63, 65, 67].

more deeply in section 3.4.4, we can see that in the noise free case only the phases of the estimates, i.e., $\arg\left(\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},m}[k]\right)$, vary with time while the magnitudes are constant. A model for the l_Cth channel tap of the mnth antenna pair in this case is given by $h_{\mathrm{C},mn}[k,l_\mathrm{C}] = e^{j\cdot\Delta\phi_{mn}[k,l_\mathrm{C}]}\cdot h_{\mathrm{C},mn}[l_\mathrm{C}]$, where $\Delta\phi_{mn}[k,l_\mathrm{C}] = \phi_{\mathrm{Tx},n}[k-l_\mathrm{C}+1]-\phi_{\mathrm{Rx},m}[k]$. Thus, an estimate of the phase processes can be obtained, by observing the phase change of the estimates over time. In particular the phase change between two estimates, which corresponds to the carrier phase changes, can be given as

$$\mathbf{\Delta\hat{\phi}}_{L_\mathrm{C},m}[k] = \hat{\mathbf{h}}^{*}_{\mathrm{C},L_\mathrm{C},m}[k]\odot\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},m}[k+L_\mathrm{E}], \tag{3.44}$$

where L_E is the number of samples between the two estimates, and $\mathbf{\Delta\hat{\phi}}_{L_\mathrm{C},m}[k]$ contains the carrier phase differences between all transmitters and the mth receiver, weighted by the power of the corresponding tap. In general, the phase variation is the same for every tap of a given Tx/Rx antenna pair, which means that they can be combined in a sensible way depending on the quality of the estimate, e.g., by averaging the argument of different taps for an mn pair. Compared to the combined phase process matrix $\mathbf{\Delta\Phi}[k]$, which was defined in section 2.5.2, the estimate $\mathbf{\Delta\hat{\phi}}_{L_\mathrm{C},m}[k]$ contains the phase processes across L_E samples instead of one, is not normalized, and only corresponds to the mth row. When the phase processes are simple linearly-increasing (or linearly-drifting) functions for those L_E samples, i.e., $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} \approx 0$ for these samples, the mapping between the two becomes

$$\mathbf{\Delta\hat{\Phi}}[k] = \begin{bmatrix} e^{j\cdot\frac{1}{L_\mathrm{E}}\arg\left(\mathbf{\Delta\hat{\phi}}^\mathrm{T}_{L_\mathrm{C},1}[k]\right)} \\ e^{j\cdot\frac{1}{L_\mathrm{E}}\arg\left(\mathbf{\Delta\hat{\phi}}^\mathrm{T}_{L_\mathrm{C},2}[k]\right)} \\ \vdots \\ e^{j\cdot\frac{1}{L_\mathrm{E}}\arg\left(\mathbf{\Delta\hat{\phi}}^\mathrm{T}_{L_\mathrm{C},M}[k]\right)} \end{bmatrix} = \begin{bmatrix} e^{j\cdot\frac{1}{L_\mathrm{E}}\left(\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},1}[k+L_\mathrm{E}]\right)-\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},1}[k]\right)\right)} \\ e^{j\cdot\frac{1}{L_\mathrm{E}}\left(\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},2}[k+L_\mathrm{E}]\right)-\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},2}[k]\right)\right)} \\ \vdots \\ e^{j\cdot\frac{1}{L_\mathrm{E}}\left(\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},M}[k+L_\mathrm{E}]\right)-\arg\left(\hat{\mathbf{h}}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},M}[k]\right)\right)} \end{bmatrix}. \tag{3.45}$$

Depending on the shape of the phase processes, e.g., linear drift or more complicated function, the interval L_E can be increased and the mode of combining several estimates can be adapted. When the contribution[8] of the Wiener process is high, i.e., severe close-in phase noise, a low value for L_E is necessary and only the most recent estimates may be used for averaging. On the other hand, when the linear phase ramp dominates, averaging can be performed over a long time frame, or, equivalently less frequent estimation is required. However, even in such a case L_E cannot be arbitrarily large, as phase ambiguities need to be avoided. Considering the highest linear increase from section 2.4 μ_{φ_w}, which yields a combined phase process with twice that value as a slope, we need to measure the phase change before the drift due to that

[8] Note that this needs to hold for the oscillators on both the Tx and Rx side. In other words, combining two linearly drifting phase processes yields a combined linearly-drifting phase process, while a nonlinear process on one side leads to a combined nonlinear phase process.

increase reaches a value of 2π, or something smaller. This can be expressed by

$$\frac{2\pi}{a} > 2\mu_{\varphi_\mathrm{w}} \cdot L_\mathrm{E} \tag{3.46}$$

$$\implies L_\mathrm{E} < \frac{\pi}{a \cdot \mu_{\varphi_\mathrm{w}}} \tag{3.47}$$

$$\sim L_\mathrm{E} \leq \left\lceil \frac{\pi}{4\mu_{\varphi_\mathrm{w}}} \right\rceil, \tag{3.48}$$

where a is a factor that needs to be suitably chosen, e.g., $a = 4$ has been found to work well in practice, and $\lceil \cdot \rceil$ is the ceiling operation.

So far, the estimation of all $N \cdot M$, or $N \cdot M \cdot L_\mathrm{C}$, when they are different for each channel tap, phase processes that can be observed at the receiver has been considered without taking the structure of the problem into account. However, there are only $N+M$ phase processes present in the system, which are combined into the rank-one matrix $\mathbf{\Delta\Phi}[k]$ through a vector product. Taking this property into account will be briefly investigated in the next section, as it can improve estimation performance and allows for efficient equalization, as will be seen in the next chapter, especially for large system sizes.

Exploiting the Structure of the Combined Phase Processes

Consider that an estimate of the combined phase process matrix $\mathbf{\Delta\hat{\Phi}}[k]$ is available. It was described in section 2.5.2 that this matrix is generated by a multiplication of the transmit and receive phase process vectors, i.e., $\mathbf{\Delta\Phi}[k] = \boldsymbol{\phi}_\mathrm{Rx}[k](\mathbf{1}_{L_\mathrm{C}\times 1} \otimes \boldsymbol{\phi}_\mathrm{Tx}[k])^\mathrm{T}$. Thus, we need to find a method to get from the matrix estimate $\mathbf{\Delta\hat{\Phi}}[k]$ to the estimates of the individual processes $\hat{\boldsymbol{\phi}}_\mathrm{Tx}[k]$ and $\hat{\boldsymbol{\phi}}_\mathrm{Rx}[k]$. It is important to note that there is some scaling and ordering ambiguity, when considering just one instance of the estimate. For example, exchanging the phase process values between Tx and Rx yields the same matrix. The problem is further complicated by the fact that the matrix estimates are noisy, meaning that the observed matrix will not always be of rank one. Let us briefly elaborate on that point. Assume the estimates can be written as $\mathbf{\Delta\hat{\Phi}}[k] = \mathbf{\Delta\Phi}[k] + \mathbf{n}[k]\mathbf{1}_{1\times NL_\mathrm{C}}$, where $\mathbf{n}[k]$ is a vector of independent Gaussian noise entries. The last term thus presumes that the same noise process of a receiver[9] is seen for all phase difference estimates associated with it. Since the term $\mathbf{n}[k]\mathbf{1}_{1\times NL_\mathrm{C}}$ is a multiplication of two vectors, like $\mathbf{\Delta\Phi}[k]$, it is also a rank-one matrix. Then, we can use the subadditivity of ranks, which is $\mathrm{rank}\left(\mathbf{\Delta\Phi}[k] + \mathbf{n}[k]\mathbf{1}_{1\times NL_\mathrm{C}}\right) \leq \mathrm{rank}\left(\mathbf{\Delta\Phi}[k]\right) + \mathrm{rank}\left(\mathbf{n}[k]\mathbf{1}_{1\times NL_\mathrm{C}}\right)$, showing that the rank can be higher than one. A general iterative solution to the noisy rank-one decomposition can be found in [83].

Only a simple method, assuming that the noise is negligible, and aided by the fact that the magnitude of the matrix and vector entries should always equal to one, will be described here.

[9] Whether this is true or not is left open. However, the worst case assumption is that a noise matrix $\mathbf{N}[k]$ with completely independent entries is added to $\mathbf{\Delta\Phi}[k]$, in order to yield the estimate $\mathbf{\Delta\hat{\Phi}}[k]$. In this case, $\mathbf{N}[k]$ is full rank, to be specific $\mathrm{rank}\left(\mathbf{N}[k]\right) = \min\left(M, NL_\mathrm{C}\right)$. The rank can in this case, thus, also be higher than one.

Given these assumptions, any row of $\mathbf{\Delta\Phi}[k]$ can be taken to be an estimate for the Tx processes, or any column of it to be an estimate for the Rx processes. Whichever row or column is chosen, the corresponding transmitter or receiver phase process acts as a reference for the rest of the system, meaning that the estimates only show the differences w.r.t. that phase process. The following operations[10] need to be carried out in order for the mth receiver to be the reference

$$\hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k] = \frac{1}{L_{\mathrm{C}}} \left(\mathbf{1}_{1\times L_{\mathrm{C}}} \otimes \mathbf{I}_N\right) \mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{T}}[k]\mathbf{e}_m \tag{3.49}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{Rx}}[k] = \frac{1}{NL_{\mathrm{C}}} \mathbf{\Delta}\hat{\mathbf{\Phi}}[k] \left(\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \hat{\boldsymbol{\phi}}^{*}_{\mathrm{Tx}}[k]\right), \tag{3.50}$$

where $\mathbf{e}_m$ is the mth basis vector, i.e., a zero vector with a one in the mth entry. For a transmitter to be the reference, the operations change to

$$\hat{\boldsymbol{\phi}}_{\mathrm{Rx}}[k] = \mathbf{\Delta}\hat{\mathbf{\Phi}}[k]\mathbf{e}_{n,l_{\mathrm{C}}} \tag{3.51}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k] = \frac{1}{ML_{\mathrm{C}}} \left(\mathbf{1}_{1\times L_{\mathrm{C}}} \otimes \mathbf{I}_N\right) \mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{T}}[k]\hat{\boldsymbol{\phi}}^{*}_{\mathrm{Rx}}[k]. \tag{3.52}$$

Without loss of generality, $\mathbf{e}_1$ can be chosen for both approaches, since it does not matter which stream is chosen as the reference[11]. Through substitution it can, furthermore, be seen that $\hat{\boldsymbol{\phi}}_{\mathrm{Rx}}[k](\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k])^{\mathrm{T}} = \mathbf{\Delta}\hat{\mathbf{\Phi}}[k]$. Finally, the effect when this approach is used for compensation will be investigated. The phase processes effects can be compensated by multiplying the received and equalized streams with the complex conjugate of the estimated phase process, see also the next chapter. If the multiplication of estimated value and actual value yields one, the effect is optimally compensated. Assume receiver one as the reference and $\mathbf{\Delta}\hat{\mathbf{\Phi}}[k] = \mathbf{\Delta\Phi}[k]$, which yields for the transmitter part

$$\mathbf{1}_{N\times 1} \stackrel{!}{=} \boldsymbol{\phi}_{\mathrm{Tx}}[k] \odot \hat{\boldsymbol{\phi}}^{*}_{\mathrm{Tx}}[k] \tag{3.53}$$

$$\stackrel{!}{=} \boldsymbol{\phi}_{\mathrm{Tx}}[k] \odot \frac{1}{L_{\mathrm{C}}} \left(\mathbf{1}_{1\times L_{\mathrm{C}}} \otimes \mathbf{I}_N\right) \mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{H}}[k]\mathbf{e}_1 \tag{3.54}$$

$$\stackrel{!}{=} \boldsymbol{\phi}_{\mathrm{Tx}}[k] \odot \frac{1}{L_{\mathrm{C}}} \left(\mathbf{1}_{1\times L_{\mathrm{C}}} \otimes \mathbf{I}_N\right) (\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \boldsymbol{\phi}^{*}_{\mathrm{Tx}}[k])\boldsymbol{\phi}^{\mathrm{H}}_{\mathrm{Rx}}[k]\mathbf{e}_1 \tag{3.55}$$

$$\stackrel{!}{=} \boldsymbol{\phi}_{\mathrm{Tx}}[k] \odot \boldsymbol{\phi}^{*}_{\mathrm{Tx}}[k]\phi^{*}_{\mathrm{Rx},1}[k] \tag{3.56}$$

$$\stackrel{!}{=} \mathbf{1}_{N\times 1} \cdot \phi^{*}_{\mathrm{Rx},1}[k]. \tag{3.57}$$

[10] If the phase processes are different for the taps of a Tx/Rx antenna pair, the Kronecker products can be removed and longer vectors can be used.

[11] If there are significant SNR differences between the Tx/Rx pairs, the reference should be chosen to be the one with the highest SNR. For example, if one transmitter has significantly higher transmit power than the other ones, it should be chosen as the reference.

Using this result, the receiver part is given by

$$\mathbf{1}_{M\times 1} \overset{!}{=} \boldsymbol{\phi}_{\mathrm{Rx}}[k] \odot \hat{\boldsymbol{\phi}}^{*}_{\mathrm{Rx}}[k] \tag{3.58}$$

$$\overset{!}{=} \boldsymbol{\phi}_{\mathrm{Rx}}[k] \odot \frac{1}{NL_{\mathrm{C}}} \boldsymbol{\Delta}\hat{\boldsymbol{\Phi}}^{*}[k] \left(\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k]\right) \tag{3.59}$$

$$\overset{!}{=} \boldsymbol{\phi}_{\mathrm{Rx}}[k] \odot \frac{1}{NL_{\mathrm{C}}} \boldsymbol{\phi}^{*}_{\mathrm{Rx}}[k] (\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \boldsymbol{\phi}^{*}_{\mathrm{Tx}}[k])^{\mathrm{T}} \left(\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k]\right) \tag{3.60}$$

$$\overset{!}{=} \boldsymbol{\phi}_{\mathrm{Rx}}[k] \odot \frac{1}{NL_{\mathrm{C}}} \boldsymbol{\phi}^{*}_{\mathrm{Rx}}[k] (\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \boldsymbol{\phi}^{*}_{\mathrm{Tx}}[k])^{\mathrm{T}} \left(\mathbf{1}_{L_{\mathrm{C}}\times 1} \otimes \boldsymbol{\phi}_{\mathrm{Tx}}[k]\phi_{\mathrm{Rx},1}[k]\right) \tag{3.61}$$

$$\overset{!}{=} \boldsymbol{\phi}_{\mathrm{Rx}}[k] \odot \boldsymbol{\phi}^{*}_{\mathrm{Rx}}[k]\phi_{\mathrm{Rx},1}[k] \tag{3.62}$$

$$\overset{!}{=} \mathbf{1}_{M\times 1} \cdot \phi_{\mathrm{Rx},1}[k]. \tag{3.63}$$

It is seen that the residual parts do not equal one. The term from the receiver side is seen to be the phase process of the chosen reference receiver, while the term from the transmitter side is seen to be the conjugate complex of that process. Since joint phase variations can be shifted throughout the processing chain, see the next chapter, the two terms actually cancel out and the estimates $\hat{\boldsymbol{\phi}}_{\mathrm{Tx}}[k]$ and $\hat{\boldsymbol{\phi}}_{\mathrm{Rx}}[k]$, derived here based on $\boldsymbol{\Delta}\hat{\boldsymbol{\Phi}}[k]$, can be used directly for equalization.

3.3.3 Symbol-Spaced Model Containing Only SFOs

It was already discussed that SFOs can only be compensated to a certain extent, when the received signal is not oversampled. The same also holds true for the estimation of the sampling phase processes because the aliasing, which creates ISI when the sampling is not aligned, is rarely known. Nevertheless, some simple methods can be derived, which yield satisfying results in certain circumstances. As for the CFOs, it will be assumed that the channel is approximately time invariant w.r.t. the length of the training sequence, i.e., constant sampling phase processes with $\dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k + l_{\mathrm{Tr}}]$ and $\dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k + l_{\mathrm{Tr}}]$ for $l_{\mathrm{Tr}} = \{1, \ldots, L_{\mathrm{Tr}} - 1\}$.

In this case, the channel taps are then time varying due to the differences in the sampling phase processes. The same signal model as in the CFO case above applies, and the same estimators can be used, for example,

$$\hat{\mathbf{h}}_{\mathrm{C},L_{\mathrm{C}},m}[k] = \mathbf{X}^{\mathrm{H}}_{\mathrm{Tr},M=1} \cdot \mathbf{y}_{L_{\mathrm{R}},m}[k]. \tag{3.64}$$

Looking at the variations of channel estimates over time, as was done in the CFO case, provides only little insight here, as the variations depend strongly on the aforementioned filters. However, since the variation is cyclical, an estimate can be generated by finding indices where channel estimates are approximately equal. This needs to be done for each Tx/Rx antenna pair, respectively. Assume that we have found an index $L_{\mathrm{E},mn}$ for an antenna pair where $\hat{\mathbf{h}}_{\mathrm{C},L_{\mathrm{C}},mn}[k] \approx \hat{\mathbf{h}}_{\mathrm{C},L_{\mathrm{C}},mn}[k + L_{\mathrm{E},mn}]$. It is then known that the difference sampling phase process between that pair, i.e., $\Delta\dot{\phi}_{mn}[k] = \dot{\phi}_{\mathrm{Rx},m}[k] - \dot{\phi}_{\mathrm{Tx},n}[k]$, generates a full symbol shift within

$L_{\mathrm{E},mn}$ samples. If the sampling phase processes are dominated by linearly-drifting processes, e.g., $\dot{\phi}_{\mathrm{Rx},m}[k] = \mu_{\dot{\varphi}_\mathrm{w},m} \cdot k$, the slope can be estimated[12] by

$$\Delta\mu_{\dot{\varphi}_\mathrm{w},mn} = \frac{1}{L_{\mathrm{E},mn}}, \tag{3.65}$$

with $\Delta\mu_{\dot{\varphi}_\mathrm{w},mn} = \mu_{\dot{\varphi}_\mathrm{w},m} - \mu_{\dot{\varphi}_\mathrm{w},n}$. When the phase processes have more complicated behavior, it becomes difficult to achieve meaningful estimation results with this method. Two other problems with this approach, which are left open, are finding a good similarity measure for comparing $\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},mn}[k]$ with $\hat{\mathbf{h}}_{\mathrm{C},L_\mathrm{C},mn}[k + L_{\mathrm{E},mn}]$, since the estimates are influenced by noise, and the problem that we need to continuously monitor the estimates in order to find $L_{\mathrm{E},mn}$. It should also be noted that the sampling phase processes are typically changing very slowly, such that estimates in the close vicinity of each other are very similar.

3.3.4 Estimation in the Oversampled Case

Continue with the description for the estimators based on the Q-oversampled received signal $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]$. Generally, all the estimators that have been derived for the symbol-spaced case can also be used when the received signal is oversampled. The time-varying channel taps associated with the mth receive antenna can thus be estimated with

$$\hat{\dot{\mathbf{h}}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k] = \dot{\mathbf{X}}^\mathrm{H}_{\mathrm{Tr},M=1} \cdot \dot{\mathbf{y}}_{\dot{L}_\mathrm{R},m}[k], \tag{3.66}$$

where the received signal and the estimated channel impulse responses are oversampled. Compared to the symbol-spaced case, the upsampled training matrix is given by

$$\dot{\mathbf{X}}_{\mathrm{Tr},M=1} = \begin{bmatrix} \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[\dot{L}_\mathrm{Tr}-1] & \cdots & \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[\dot{L}_\mathrm{Tr}-\dot{L}_\mathrm{C}] \\ \vdots & & \vdots \\ \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[\dot{L}_\mathrm{C}] & \cdots & \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[1] \\ \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[\dot{L}_\mathrm{C}-1] & \cdots & \dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr}[0] \end{bmatrix}, \tag{3.67}$$

where the upsampled training signal follows as

$$\dot{\mathbf{x}}^\mathrm{T}_\mathrm{Tr} = \begin{bmatrix} \mathbf{x}^\mathrm{T}_\mathrm{Tr}[0] & \mathbf{0}_{1\times N(Q-1)} & \mathbf{x}^\mathrm{T}_\mathrm{Tr}[1] & \mathbf{0}_{1\times N(Q-1)} & \cdots & \mathbf{x}^\mathrm{T}_\mathrm{Tr}[L_\mathrm{Tr}] & \mathbf{0}_{1\times N(Q-1)} \end{bmatrix}. \tag{3.68}$$

The full matrix $\dot{\mathbf{X}}_\mathrm{Tr}$ follows from a Kronecker product of the different rows of $\dot{\mathbf{X}}_{\mathrm{Tr},M=1}$, in the same manner as in section 3.3.1.

Except for the increased sizes of the vectors and matrices, the same estimators as in the symbol-spaced case can thus be applied for obtaining channel parameter values in the synchronized,

[12] When a matrix of the estimates for all antenna pairs is build, the same structure as in the previous section emerges, i.e., there are $M \cdot N$ estimates but only $M + N$ processes that need to be compensated. A separation of the processes, i.e., decomposing a rank-one matrix similar to the previous section, can again help for efficient equalization.

CFO only, and SFO only case. In the following, the SFO only case, and the case where both timing impairments are occurring simultaneously will be investigated, as symbol-spaced SFO estimation can only yield suboptimal performance.

3.3.5 Oversampled Model Containing Only SFOs

Consider an oversampled received signal, impaired by multiple SFOs. Using the assumption that the sampling phase processes are approximately constant during the length $\dot{L}_{\mathrm{Tr}}$ of the training sequences, i.e., $\dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k+\dot{l}_{\mathrm{Tr}}]$ and $\dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k+\dot{l}_{\mathrm{Tr}}]$ for $\dot{l}_{\mathrm{Tr}} = \{1,\ldots,\dot{L}_{\mathrm{Tr}}-1\}$, the same estimators as in the symbol-spaced case can be used, e.g.,

$$\hat{\dot{\mathbf{h}}}_{\mathrm{C},\dot{L}_{\mathrm{C}},m}[k] = \dot{\mathbf{X}}^{\mathrm{H}}_{\mathrm{Tr},M=1} \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}},m}[k]. \tag{3.69}$$

These channel estimates, associated with every mth receive antenna, vary again due to the difference in the sampling phase processes. Compared to the symbol-spaced case, the sampling criterion is now approximately fulfilled, which means that we have access to the full pulse shape of the received signal.

An approximate model[13] for the discrete channel taps of the mnth antenna pair in this case is given by $\dot{h}_{\mathrm{C},mn}[k,\dot{l}_{\mathrm{C}}] = \mathrm{sinc}\left[\pi \frac{\dot{l}_{\mathrm{C}}-\dot{L}_{\mathrm{C}}/2+\Delta\dot{\phi}_{mn}[k]}{Q}\right] \star \dot{h}_{\mathrm{C},mn}[\dot{l}_{\mathrm{C}}]$, where $\Delta\dot{\phi}_{mn}[k] = \dot{\phi}_{\mathrm{Rx},m}[k] - \dot{\phi}_{\mathrm{Tx},n}[k]$, and $\dot{\mathbf{h}}_{\mathrm{C},\dot{L}_{\mathrm{C}},m}[k] = \left[\dot{\mathbf{h}}_{\mathrm{C},m}[k,0] \quad \dot{\mathbf{h}}_{\mathrm{C},m}[k,1] \quad \cdots \quad \dot{\mathbf{h}}_{\mathrm{C},m}[k,\dot{L}_{\mathrm{C}}-1]\right]^{\mathrm{T}}$ with $\dot{\mathbf{h}}_{\mathrm{C},m}[k,\dot{l}_{\mathrm{C}}] = \left[\dot{h}_{\mathrm{C},m1}[k,\dot{l}_{\mathrm{C}}] \quad \dot{h}_{\mathrm{C},m2}[k,\dot{l}_{\mathrm{C}}] \quad \cdots \quad \dot{h}_{\mathrm{C},mN}[k,\dot{l}_{\mathrm{C}}]\right]^{\mathrm{T}}$. From the estimate for the channel taps of a Tx/Rx antenna pair $\hat{\dot{\mathbf{h}}}_{\mathrm{C},\dot{L}_{\mathrm{C}},mn}[k] = \left[\dot{h}_{\mathrm{C},mn}[k,0] \quad \dot{h}_{\mathrm{C},mn}[k,1] \quad \cdots \quad \dot{h}_{\mathrm{C},mn}[k,\dot{L}_{\mathrm{C}}-1]\right]^{\mathrm{T}}$, we can then estimate the corresponding difference in the sampling phase processes $\Delta\dot{\phi}_{mn}[k] = \dot{\phi}_{\mathrm{Rx},m}[k] - \dot{\phi}_{\mathrm{Tx},n}[k]$. Given the assumptions about the training signals[14] and white Gaussian noise, an ML solution for estimating the sampling phase differences is given by

$$\Delta\hat{\dot{\phi}}_{mn}[k] = \underset{\Delta\dot{\phi}_{mn}}{\arg\max} \left[\dot{\mathbf{y}}^{\mathrm{H}}_{\dot{L}_{\mathrm{R}},m}[k]\dot{\mathbf{X}}_{\mathrm{Tr},M=1,n}\left[\Delta\dot{\phi}_{mn}\right] \cdot \dot{\mathbf{X}}^{\mathrm{H}}_{\mathrm{Tr},M=1,n}\left[\Delta\dot{\phi}_{mn}\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}},m}[k]\right], \tag{3.70}$$

where $\dot{\mathbf{X}}_{\mathrm{Tr},M=1,n}[\Delta\dot{\phi}_{mn}]$ are trial matrices with entries $\dot{x}_{\mathrm{Tr},n}[\Delta\dot{\phi}_{mn},\dot{l}_{\mathrm{Tr}}] = \mathrm{sinc}\left[\pi \frac{\dot{l}_{\mathrm{Tr}}-\dot{L}_{\mathrm{C}}/2+\Delta\dot{\phi}_{mn}}{Q}\right] \star \dot{x}_{\mathrm{Tr},n}[\dot{l}_{\mathrm{Tr}}]$, which are used for finding the maximum of this equation for each mnth antenna pair. Since there is no prior knowledge about the values of $\Delta\dot{\phi}_{mn}$, the whole space of possible values has to be tested, with a finer grid of $\Delta\dot{\phi}_{mn}$ values leading to more accurate estimates. Even though the trial matrices $\dot{\mathbf{X}}_{\mathrm{Tr},M=1,n}[\Delta\dot{\phi}_{mn}]$ can be precomputed, and the maximization can be carried out in parallel for the $M \cdot N$ estimates given a desired grid spacing, this procedure seems too complex for MIMO systems.

[13] As the oversampling factor increases, the model becomes more accurate since the influence of the sinc function is reduced, and more samples of the actual pulse shape are obtained.

[14] For both CFOs and SFOs, the training signals are generally not completely uncorrelated due to the shift in frequency and sampling. This effect will be assumed negligible.

Simpler estimators can be obtained through some approximations. First, assume that the sinc function model for the SFOs impact on the channel taps is valid. Then, consider the DFT of these taps[15], given by

$$\mathscr{F}\left\{\dot{h}_{\mathrm{C},mn}[k,\dot{l}_\mathrm{C}]\right\} = \mathscr{F}\left\{\mathrm{sinc}\left[\pi\frac{\dot{l}_\mathrm{C}-\dot{L}_\mathrm{C}/2+\Delta\dot{\phi}_{mn}[k]}{Q}\right]\star\dot{h}_{\mathrm{C},mn}[\dot{l}_\mathrm{C}]\right\} \tag{3.71}$$

$$= \mathscr{F}\left\{\mathrm{sinc}\left[\pi\frac{\dot{l}_\mathrm{C}-\dot{L}_\mathrm{C}/2}{Q}\right]\star\delta\left[\dot{l}_\mathrm{C}+\Delta\dot{\phi}_{mn}[k]\right]\star\dot{h}_{\mathrm{C},mn}[\dot{l}_\mathrm{C}]\right\} \tag{3.72}$$

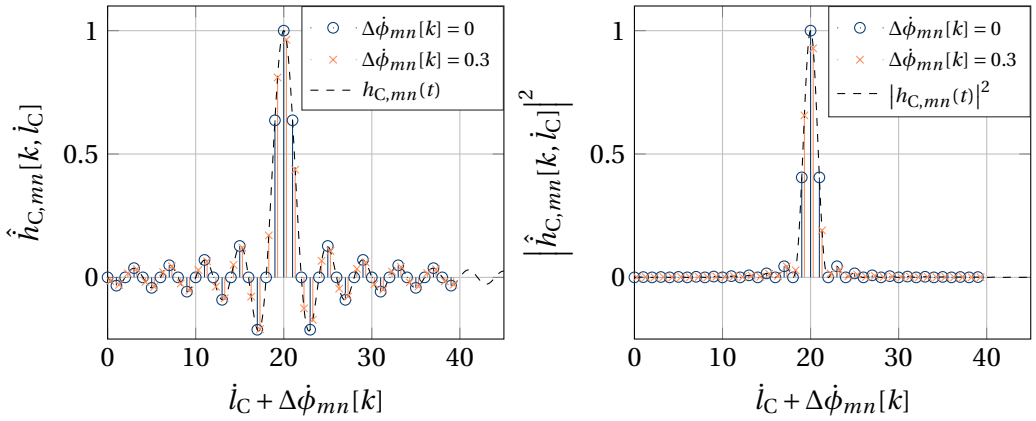
$$\approx \mathscr{F}\left\{\mathrm{sinc}\left[\pi\frac{\dot{l}_\mathrm{C}-\dot{L}_\mathrm{C}/2}{Q}\right]\right\}\cdot e^{j\cdot 2\pi\frac{\dot{l}_\mathrm{C}}{\dot{L}_\mathrm{C}}\cdot\Delta\dot{\phi}_{mn}[k]}\cdot\mathscr{F}\left\{\dot{h}_{\mathrm{C},mn}[\dot{l}_\mathrm{C}]\right\}, \tag{3.73}$$

where the approximation in the last line uses the circular shift property of the DFT. This approximation is exact, if the shift of the discrete channel taps due to the SFOs is circular, which is generally not the case. However, assuming that there are only a few taps with significant amplitude values, and that they are also somewhat localized to a certain sample region of the impulse response[16], the shift is approximately circular.

Assuming the validity of (3.73), it is seen that the only thing that changes in the DFT of the channel taps with k is a phase factor that depends on $\Delta\dot{\phi}_{mn}[k]$. Thus, one approach to estimating the sampling phase difference is to compare the phase change in the DFT of two time separated channel estimates, i.e., $\arg\left(\mathscr{F}\left\{\hat{\dot{h}}_{\mathrm{C},mn}[k+\dot{L}_\mathrm{E},\dot{l}_\mathrm{C}]\right\}\right)$ compared to $\arg\left(\mathscr{F}\left\{\hat{\dot{h}}_{\mathrm{C},mn}[k,\dot{l}_\mathrm{C}]\right\}\right)$ where $\dot{L}_\mathrm{E}$ is the number of samples between the two estimates, similar to the CFO approach in section 3.3.2. This approach will not be explicitly treated further. Instead, we continue with an approach that can obtain an estimate of the sampling phase difference from a single channel estimate. Looking at (3.73), it is seen that the left and right hand term prevent the direct estimation of $\Delta\dot{\phi}_{mn}[k]$. The first term is known and just depends on the oversampling factor Q, and the length of the channel impulse response $\dot{L}_\mathrm{C}$. The last term depends on the channel characteristics, and is generally not known beforehand. Using some reasonable assumptions, its impact can be made negligible for certain channel types. First, consider the channel to be flat, i.e., $\dot{h}_{\mathrm{C},mn}[\dot{l}_\mathrm{C}] = e^{-j\varphi_{mn}}\delta[\dot{l}_\mathrm{C}]$ or $\dot{\mathbf{h}}_{\mathrm{C},mn} = \left[e^{-j\varphi_{mn}} \quad \mathbf{0}_{1\times(\dot{L}_\mathrm{C}-1)}\right]^\mathrm{T}$, where φ_{mn} is a phase shift due to propagation. The DFT of such an impulse response is the just mentioned phase shift due to propagation multiplied by additional phase shifts depending on the delay of the pulse, due to the circular shift property of the DFT. Since the phase shift part due to the DFT is known from the position of the pulse, the only remaining unknown quantity is φ_{mn}. This unknown can be removed by taking the absolute value of the channel, i.e., $|\dot{h}_{\mathrm{C},mn}[\dot{l}_\mathrm{C}]| = \delta[\dot{l}_\mathrm{C}]$, before applying the DFT. Thus, for flat channels taking the absolute value of the channel estimate before applying the DFT, i.e., $\mathscr{F}\left\{\left|\hat{\dot{h}}_{\mathrm{C},mn}[k,\dot{l}_\mathrm{C}]\right|\right\}$, leads directly to a sampling phase estimate. Going back to more complex channel impulse responses containing multiple taps, there is no general solution to removing their influence on the phase characteristic of the corresponding DFT. This is because phase shifts appear in the DFT, depending on the position of the different taps and their significance is weighted by the amplitude of the corresponding tap.

[15] Standard finite-length DFT problems, such as spectral leakage, will be ignored here.

[16] How to strengthen this assumption will be discussed in the following, and was also used in, e.g., [62, 71].



(a) Standard estimate

(b) Magnitude squared of standard estimate

Figure 3.1: Noise free example of an oversampled impulse response estimate that can be used for sampling phase estimation. Parameters are $\dot{L}_C = 40$, $Q = 2$, and $\dot{h}_{C,mn}[\dot{l}_C] = \delta[\dot{l}_C]$: Left plot shows the standard estimate for two different sampling phase offsets; Right plot shows the magnitude squared of those estimates.

In the following, a solution for a type of channel impulse responses, which is often encountered in practice, will be described. Specifically, assume that there is a dominant tap in the channel impulse response and that there are several taps of significantly lower magnitude, e.g., 10 dB below the main tap [53]. The impact of the lower amplitude channel coefficients can be reduced by applying a monotonic function, which under weights lower values compared to higher values, to the channel estimate before computing the DFT, see also [62, 71]. Commonly used are power functions, and in particular squaring of the channel estimate, i.e., $\mathscr{F}\left\{\left|\hat{\dot{h}}_{C,mn}[k, \dot{l}_C]\right|^2\right\}$. It should be noted that the monotonic function changes the shape of the magnitude of the DFT output, but does not change the desired phase or time shift of the relevant signal components. This can also be visualized in the time domain, see the example in Figure 3.1. It is seen that the shape of the pulse is altered significantly, but the sample phase offset w.r.t. sample phase zero is the same. Even though this is a flat channel example, it can be seen here that lower amplitude coefficients are significantly attenuated by the squaring operation, thereby reducing their impact on the phase characteristic of the corresponding DFT. Another benefit of applying the monotonic function also becomes apparent. For channels where the taps far from the main tap carry significantly lower power, applying the monotonic function enhances the circular shift approximation because those taps become close to zero.

Finally, further complexity reduction, which can be achieved by using a smaller subset of the $\dot{L}_C$ available DFT values, will be considered. Looking at (3.73), it is seen that the sampling phase offset can be extracted from any of the $\dot{L}_C$ DFT values, given that the value has significant power. It is known that the $\dot{L}_C$-length DFT of a Q-fold oversampled signal has $\dot{L}_C/Q$ significant

values[17], which are split between the beginning and end of the DFT output sequence if the signal is in baseband. Since the signal is not exactly cyclically shifted, its DFT experiences an additional fixed phase shift after the middle of the sequence. Thus, the lowest $\lfloor \dot{L}_\mathrm{C}/(2Q) \rfloor$ DFT values are suggested to be used for the estimation, where the flooring operation is used in order to have integer DFT values. The lowest-complexity estimator, given that the phase impact of multi-path components is sufficiently attenuated by the squaring operation, is obtained by calculating a single dth DFT value, where $d = \left\{2, \ldots, \left\lfloor \frac{\dot{L}_\mathrm{C}}{2Q} \right\rfloor \right\}$, with

$$\begin{aligned} \Delta\hat{\dot{\phi}}_{mn}[k] = \frac{1}{2\pi} \cdot \frac{\dot{L}_\mathrm{C}}{d-1} \cdot \Bigg(& \arg\left(\mathscr{F}_d \left\{ \left| \hat{\dot{h}}_{\mathrm{C},mn}[k, \dot{\imath}_\mathrm{C}] \right|^2 \right\} \right) \\ & - \arg\left(\mathscr{F}_d \left\{ \left| \mathrm{sinc}\left[\pi \frac{\dot{\imath}_\mathrm{C} - \dot{L}_\mathrm{C}/2}{Q} \right] \right|^2 \right\} \right) \Bigg) \end{aligned} \tag{3.74}$$

$$\begin{aligned} = \frac{1}{2\pi} \cdot \frac{\dot{L}_\mathrm{C}}{d-1} \cdot \Bigg(& \arg\left(\sum_{\dot{\imath}_\mathrm{C}=0}^{\dot{L}_\mathrm{C}-1} \left| \hat{\dot{h}}_{\mathrm{C},mn}[k, \dot{\imath}_\mathrm{C}] \right|^2 \cdot e^{-j2\pi \cdot \frac{\dot{\imath}_\mathrm{C}}{\dot{L}_\mathrm{C}} \cdot (d-1)} \right) \\ & - \arg\left(\sum_{\dot{\imath}_\mathrm{C}=0}^{\dot{L}_\mathrm{C}-1} \left| \mathrm{sinc}\left[\pi \frac{\dot{\imath}_\mathrm{C} - \dot{L}_\mathrm{C}/2}{Q} \right] \right|^2 \cdot e^{-j2\pi \cdot \frac{\dot{\imath}_\mathrm{C}}{\dot{L}_\mathrm{C}} \cdot (d-1)} \right) \Bigg). \end{aligned} \tag{3.75}$$

It should be visible that the lower limit of d is chosen as two, since the first DFT value ($d = 1$) cannot give a useful estimate for the sampling phase difference. The second argument term in (3.75) is a fixed constant for a given $\dot{L}_\mathrm{C}$, d, Q, and, thus, does not increase complexity substantially. It may also be simplified for certain parameter combinations and alignments [62, 71].

Since the previous discussion was quite lengthy, the main points will now be briefly summarized here, also using vector-matrix notation. Given the estimate for the channel taps associated with the mth receive antenna $\hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]$, obtained with (3.69), check if there are significant multi-path components[18]. If that is the case, obtain a second channel estimate $\hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k + \dot{L}_\mathrm{E}]$, take the $\dot{L}_\mathrm{C}$-length DFT of each mnth sub-component, and compare the phases of the lower DFT values[19]. If no significant multi-path values are observed, generate

[17] The application of the monotonic function changes the amount of significant values. Power functions, for example, increase the amount of significant values towards the middle of the DFT sequence. However, their amplitude still drops towards the middle of the sequence, meaning that using the mentioned $\dot{L}_\mathrm{C}/Q$ values still captures the most significant values. Furthermore, while the sampling criterion is always fulfilled with $Q = 2$ for the normal channel estimate, it can be violated after applying the monotonic function. Thus, selecting the lower DFT values also limits the influence of aliasing on the estimation.

[18] Equivalently, we can also check the flatness of the channel in the frequency domain.

[19] Compared to the solution in (3.76), one replaces the second argument part with the first one, and inserts the second estimate $\hat{\mathbf{H}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k + \dot{L}_\mathrm{E}]$ into the first argument part. Furthermore, some normalization w.r.t. the separation $\dot{L}_\mathrm{E}$ has to be done, as this approach yields the sample shift over $\dot{L}_\mathrm{E}$ samples.

the mn ordered matrix $\hat{\mathbf{H}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k] = \left[\hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},m1}[k] \quad \cdots \quad \hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},mN}[k]\right]$, and use

$$\Delta\hat{\dot{\boldsymbol{\phi}}}_m^{\mathrm{T}}[k] = \frac{\dot{L}_\mathrm{C}}{2\pi}\cdot\mathbf{d}_D^{\mathrm{T}}\cdot\Bigg(\arg\Bigg(\underbrace{\left[\mathbf{0}_{D\times 1} \quad \mathbf{I}_D \quad \mathbf{0}_{D\times(\dot{L}_\mathrm{C}-D-1)}\right]\mathbf{D}_{\dot{L}_\mathrm{C}}}_{\tilde{\mathbf{D}}_{\dot{L}_\mathrm{C},D}}\left(\hat{\mathbf{H}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]\odot\hat{\mathbf{H}}^*_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]\right)\Bigg) - \arg\left(\tilde{\mathbf{D}}_{\dot{L}_\mathrm{C},D}\left(\dot{\mathbf{s}}_{\dot{L}_\mathrm{C},Q}\odot\dot{\mathbf{s}}^*_{\dot{L}_\mathrm{C},Q}\right)\right)\cdot\mathbf{1}_{1\times N}\Bigg), \tag{3.76}$$

where $\mathbf{D}_{\dot{L}_\mathrm{C}}$ is the DFT matrix of size $\dot{L}_\mathrm{C}$, D is the number of DFT values that are used for the estimation, e.g., $D = \left\lfloor\frac{\dot{L}_\mathrm{C}}{2Q}\right\rfloor - 1$ according to the previous suggestion, and $\tilde{\mathbf{D}}_{\dot{L}_\mathrm{C},D}$ is a matrix with D selected rows from the standard DFT matrix. Furthermore, the vector $\dot{\mathbf{s}}_{\dot{L}_\mathrm{C},Q} = \left[\mathrm{sinc}\left[\pi\frac{1-\dot{L}_\mathrm{C}/2}{Q}\right] \quad \mathrm{sinc}\left[\pi\frac{2-\dot{L}_\mathrm{C}/2}{Q}\right] \quad \cdots \quad \mathrm{sinc}\left[\pi\frac{\dot{L}_\mathrm{C}-\dot{L}_\mathrm{C}/2}{Q}\right]\right]^{\mathrm{T}}$ contains the reference pulse, which is fixed for a given $\dot{L}_\mathrm{C}$ and Q, and which is used for comparison of the DFT phases. Additionally, $\mathbf{d}_D = \frac{1}{D}\left[\frac{1}{1} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{D}\right]^{\mathrm{T}}$ is a vector that performs averaging of the D used DFT phase values, and removes the phase rotation that is inherent to each DFT value. In general, the solution in (3.76) averages D consecutive DFT values, starting from $d = 2$, for each of the N different sampling phase difference estimates. In order to use a different set of DFT values, the vector $\mathbf{d}_D$ and the matrix before $\mathbf{D}_{\dot{L}_\mathrm{C}}$ need to be adjusted accordingly. Note that for $D = 1$ and $N = 1$ (3.76) is equivalent to (3.75) with $d = 2$.

When the sampling phase difference estimation has been carried out for all receive antennas, as described in the previous sections, $M\cdot N$ estimates have been generated, similar to the CFO case. However, only $M+N$ sampling phase processes exist between transmitter and receiver. As in the CFO case, a complete phase difference matrix $\Delta\hat{\dot{\boldsymbol{\Phi}}}[k]$ can be defined with

$$\Delta\hat{\dot{\boldsymbol{\Phi}}}[k] = \begin{bmatrix}\Delta\hat{\dot{\boldsymbol{\phi}}}_1^{\mathrm{T}}[k]\\ \Delta\hat{\dot{\boldsymbol{\phi}}}_2^{\mathrm{T}}[k]\\ \vdots\\ \Delta\hat{\dot{\boldsymbol{\phi}}}_M^{\mathrm{T}}[k]\end{bmatrix}. \tag{3.77}$$

This matrix includes the estimate of the sampling phase differences between all Tx and Rx at time k. When more knowledge about the characteristics of the sampling phase processes is available, e.g., the linear term $\mu_{\dot{\phi}_\mathrm{w}}$ dominates, several of these estimates can be combined to further improve estimation performance. In the following, it will be briefly presented how the sampling phase difference matrix $\Delta\hat{\dot{\boldsymbol{\Phi}}}[k]$ can be separated into the transmitter and receiver contributions $\dot{\boldsymbol{\phi}}_\mathrm{Tx}[k]$ and $\dot{\boldsymbol{\phi}}_\mathrm{Rx}[k]$, as was also done for the carrier phase differences.

Exploiting the Structure of the Combined Phase Processes

Given an estimate of the combined sampling phase difference matrix, which can be defined as $\Delta\hat{\dot{\boldsymbol{\Phi}}}[k] = \dot{\boldsymbol{\phi}}_\mathrm{Rx}[k]\left(-\dot{\boldsymbol{\phi}}_\mathrm{Tx}^{\mathrm{T}}[k]\right)$, the estimates $\hat{\dot{\boldsymbol{\phi}}}_\mathrm{Tx}[k]$ and $\hat{\dot{\boldsymbol{\phi}}}_\mathrm{Rx}[k]$ should be obtained. Thus, as

in the CFO case discussed at the end of section 3.3.2, a rank-one matrix decomposition, for which a general solution is given in [83], needs to be found. A simpler solution to the problem, analogous to the CFO approach, assumes one Tx or Rx as a reference, and is given for the mth receiver as reference by

$$\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k] = -\boldsymbol{\Delta}\hat{\dot{\boldsymbol{\Phi}}}^{\mathrm{T}}[k]\mathbf{e}_m \tag{3.78}$$

$$\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k] = \frac{1}{N}\boldsymbol{\Delta}\hat{\dot{\boldsymbol{\Phi}}}[k]\left(-\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}^{\odot -1}[k]\right), \tag{3.79}$$

where $(\cdot)^{\odot -1}$ is the Hadamard (or entry-wise) inverse. For the nth Tx to be used as a reference, it can be defined as

$$\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k] = \boldsymbol{\Delta}\hat{\dot{\boldsymbol{\Phi}}}[k]\mathbf{e}_n \tag{3.80}$$

$$\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k] = -\frac{1}{M}\boldsymbol{\Delta}\hat{\dot{\boldsymbol{\Phi}}}^{\mathrm{T}}[k]\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Rx}}^{\odot -1}[k]. \tag{3.81}$$

3.3.6 Oversampled Model Containing CFOs and SFOs

Consider the most general case of an oversampled received signal $\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k]$ of a MIMO system experiencing both timing impairments, which are independent for each Tx and Rx, and a frequency-selective channel. Furthermore, as in the previous sections, assume that the carrier phase and sampling phase processes are approximately constant during the length of the training sequences, i.e., $\boldsymbol{\phi}_{\mathrm{Tx}}[k] \approx \boldsymbol{\phi}_{\mathrm{Tx}}[k+\dot{l}_{\mathrm{Tr}}]$ and $\boldsymbol{\phi}_{\mathrm{Rx}}[k] \approx \boldsymbol{\phi}_{\mathrm{Rx}}[k+\dot{l}_{\mathrm{Tr}}]$, and $\dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Tx}}[k+\dot{l}_{\mathrm{Tr}}]$ and $\dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k] \approx \dot{\boldsymbol{\phi}}_{\mathrm{Rx}}[k+\dot{l}_{\mathrm{Tr}}]$, for $\dot{l}_{\mathrm{Tr}} = \{1,\ldots,\dot{L}_{\mathrm{Tr}}-1\}$. Additionally, assume the frequency-selective behavior of the wireless channel to be time invariant over at least $\dot{L}_{\mathrm{E}}$ samples. The channels associated with the mth receive antenna, assuming $\dot{L}_{\mathrm{C}}$ significant taps, can be estimated with

$$\hat{\dot{\mathbf{h}}}_{\mathrm{C},\dot{L}_{\mathrm{C}},m}[k] = \dot{\mathbf{X}}_{\mathrm{Tr},M=1}^{\mathrm{H}} \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}},m}[k], \tag{3.82}$$

given ideal training sequences, see section 3.2.1. The channel estimates vary due to noise, as well as the timing impairments. Under the used assumptions, a model for the variation of the mnth channel over time can be given by

$$\dot{h}_{\mathrm{C},mn}[k,\dot{l}_{\mathrm{C}}] = e^{j\cdot\Delta\phi_{mn}[k]} \cdot \mathrm{sinc}\left[\pi\frac{\dot{l}_{\mathrm{C}} - \dot{L}_{\mathrm{C}}/2 + \Delta\dot{\phi}_{mn}[k]}{Q}\right] \star \dot{h}_{\mathrm{C},mn}[\dot{l}_{\mathrm{C}}], \tag{3.83}$$

where $\Delta\phi_{mn}[k] = \phi_{\mathrm{Tx},n}[k] - \phi_{\mathrm{Rx},m}[k]$ and $\Delta\dot{\phi}_{mn}[k] = \dot{\phi}_{\mathrm{Rx},m}[k] - \dot{\phi}_{\mathrm{Tx},n}[k]$ are the carrier and sampling phase difference processes of the mnth pair, respectively. Note that the shift of $\dot{L}_{\mathrm{C}}/2$ is used to make the finite-length pulse symmetric, and that functions other than $\mathrm{sinc}(\cdot)$ may be more beneficial, depending on the boundary conditions. From (3.83), it is seen that the carrier phase processes act as a phase shift on the channel coefficients. The sampling phase differences also generate phase shifts in general multi-path channels. These two observations lead to the following procedure for estimating the timing impairments in MIMO systems.

Given the channel estimates $\hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]$ and $\hat{\mathbf{h}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k+\dot{L}_\mathrm{E}]$ for each receive antenna, we should first estimate the sampling phase differences according to (3.76), when there is no significant multi-path, or using

$$\begin{aligned}\Delta\hat{\dot{\boldsymbol{\phi}}}_m^\mathrm{T}[k] = \frac{\dot{L}_\mathrm{C}}{2\pi}\cdot\frac{1}{\dot{L}_\mathrm{E}}\cdot\mathbf{d}_D^\mathrm{T}\cdot\Big(&\arg\Big(\tilde{\mathbf{D}}_{\dot{L}_\mathrm{C},D}\Big(\hat{\mathbf{H}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k+\dot{L}_\mathrm{E}]\odot\hat{\mathbf{H}}^*_{\mathrm{C},\dot{L}_\mathrm{C},m}[k+\dot{L}_\mathrm{E}]\Big)\Big)\\ &-\arg\Big(\tilde{\mathbf{D}}_{\dot{L}_\mathrm{C},D}\Big(\hat{\mathbf{H}}_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]\odot\hat{\mathbf{H}}^*_{\mathrm{C},\dot{L}_\mathrm{C},m}[k]\Big)\Big)\Big),\end{aligned} \tag{3.84}$$

when there is[20]. The approach in (3.84) assumes that the sampling phase difference process behaves linear between the two channel estimates separated by $\dot{L}_\mathrm{E}$ samples, i.e., $\Delta\dot{\boldsymbol{\Phi}}[k+\dot{L}_\mathrm{E}] = \Delta\dot{\boldsymbol{\Phi}}[k] + \dot{L}_\mathrm{E}\cdot\Delta\boldsymbol{M}_{\dot{\varphi}_\mathrm{w}}$ where $\Delta\boldsymbol{M}_{\dot{\varphi}_\mathrm{w}}$ has entries of $\Delta\mu_{\dot{\varphi}_\mathrm{w},mn} = \mu_{\dot{\varphi}_\mathrm{w},m} - \mu_{\dot{\varphi}_\mathrm{w},n}$. Sampling phase estimation should be done first, since taking the magnitude squared of the channel estimate, as proposed in the previous section, removes the influence of complex exponential phase shifts on the estimation process. Therefore, given that the carrier phase process is constant during the length of the training sequence, i.e., the model in (3.83) is valid, estimation of the sampling phase differences with (3.76) or (3.84) is not influenced by the CFOs.

With an estimate for the sampling phase differences, the initial channel estimates from (3.82) can be interpolated based on (3.83) in order to reduce the impact of the sampling phase differences on further processing. Based on those improved channel estimates[21], we can use (3.45) to estimate the carrier phase variations. Further improvement of estimation performance can be achieved by iterating between estimation of the parameters, and compensating them on $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]$ before using (3.82). Finally, once estimates for the $M\cdot N$ difference processes in the matrices $\Delta\hat{\dot{\boldsymbol{\Phi}}}[k]$ and $\Delta\hat{\dot{\tilde{\boldsymbol{\Phi}}}}[k]$ have been obtained, those matrices can be decomposed into their $M+N$ processes at the Tx and Rx side, as mentioned at the end of sections 3.3.2 and 3.3.5.

3.4 Adaptive Filtering Approach

In this section, a different approach to estimating and tracking the MIMO channel with timing impairments, based on adaptive filter theory, will be taken. It was seen in this chapter that the fundamental input/output problem is given by

$$\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k] = \mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k], \tag{3.85}$$

which means that in order to detect the data symbols, the time-varying channel matrix $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$, which changes for every block of input symbols, needs to be tracked and estimated. It is helpful to notice that two neighboring instances of the channel matrix, e.g., $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ and $\mathbf{H}_{\tilde{\mathrm{C}}}[k+1]$, are not arbitrarily different, but in fact very similar and, for example in the backhaul case, only

[20] Note that this approach just yields the change in the sampling phase between the two time instants. In order to get to zero sampling phase, it also needs to be compared to the phase of the DFT of the reference pulse, e.g., the sinc used in (3.76).

[21] Whenever the complex exponential phase contribution due to the CFOs is dominant, which is often the case in practice, we can also use the initial channel estimates of (3.82) directly.

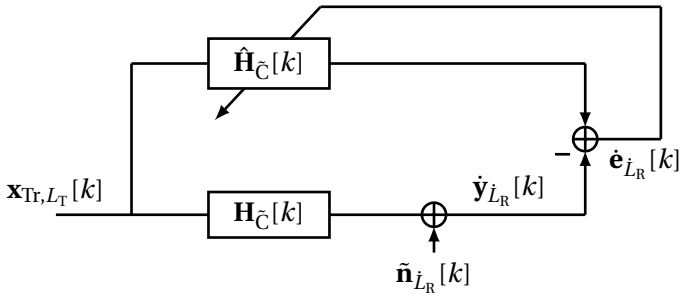


Figure 3.2: Basic working principle of adaptive algorithms for channel estimation, i.e., adapting the estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]$ based on the observed error signal $\dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]$, which contains the difference between the training signal $\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]$ transmitted through the real channel, and the training signal transmitted through the channel estimate. This setup is also known as the system identification configuration.

changing slowly due to the timing impairments. This means that once a good estimate of the channel at time k is available, only small adaptions are required to get a good estimate at time $k+1$.

The basic working principle of adaptive algorithms can be seen in Figure 3.2. A known training signal $\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]$[22] is sent both through the actual system and through an artificial system that uses the current estimate of the channel. The difference between the two outputs is the error signal $\dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]$, which is used to update the estimate for the next time step, i.e., $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k+1]$. If the actual system and the estimate are equal, the residual error signal is just the noise $\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]$. The problem addressed here is also known as system identification. Important considerations for these algorithms are:

- The initial setting of the artificial system, i.e., $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[0]$
- How to adapt the system based on the error signal $\dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]$, e.g., which characteristics of the error signal, such as correlations, are used and are models for the system changes available
- The rate of adaptation, i.e., how fast can the algorithm follow changes in the system and what is the residual error
- How robust is the algorithm w.r.t. unexpected changes in the system, e.g., a sudden large change in the system

There are different trade-offs between these properties, see the standard works [76, 84], which will also be outlined later on for the LoS MIMO systems with timing impairments.

[22] Note that L_T specifies the number of taps that are used per antenna to model the transmission channel in the adaptive filter. The required L_T, which is necessary in order to properly model the channel, is generally not known, but can be estimated for most channels.

The MAP/ML estimators were derived earlier in this chapter based on the optimal performance indicator for communications systems, i.e., the probability of symbol or estimation error. Adaptive algorithms are often derived based on different cost functions (performance indicators), meaning they are not necessarily optimal for all scenarios. The most popular cost function used for deriving adaptive algorithms is the mean squared error (MSE)[23] [76, 84], i.e., trying to minimize $\mathrm{E}[|\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k] - \mathbf{H}_{\tilde{\mathrm{C}}}[k]|^2]$, which is a quadratic function that has a unique minimum that varies in the case of timing impairments. In the following, the most basic adaptive algorithm will be outlined, and it will be shown how it can be adopted for time-varying MIMO systems. Before doing this, the second-order statistics of the MIMO system will be described, as they are fundamental in the derivation of adaptive algorithms and help in selecting the parameters for the algorithms.

3.4.1 Second-Order Statistics of the Signals in a MIMO System

It was already seen in section 3.2, that the noise has in general the following correlation matrix

$$\mathbf{C}_{\tilde{\mathbf{n}}}[k] = \mathrm{E}\left[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\tilde{\mathbf{n}}^\mathrm{H}_{\dot{L}_\mathrm{R}}[k]\right] \tag{3.86}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \cdot \dot{\mathbf{N}} \cdot \mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k], \tag{3.87}$$

in other words it varies with time due to the SFOs influencing the sampling of the receive filters. The autocorrelation matrix of the received signal is given as

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}[k] = \mathrm{E}\left[\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]\dot{\mathbf{y}}^\mathrm{H}_{\dot{L}_\mathrm{R}}[k]\right] \tag{3.88}$$

$$= \mathrm{E}\left[\left[\mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\right]\left[\mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\right]^\mathrm{H}\right] \tag{3.89}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}}}[k]\,\mathrm{E}\left[\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]\mathbf{x}^\mathrm{H}_{\mathrm{Tr},L_\mathrm{T}}[k]\right]\mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}}}[k] + \mathrm{E}\left[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\tilde{\mathbf{n}}^\mathrm{H}_{\dot{L}_\mathrm{R}}[k]\right] \tag{3.90}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}}}[k] \cdot \mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] \cdot \mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}}}[k] + \mathbf{C}_{\tilde{\mathbf{n}}}[k], \tag{3.91}$$

where it was assumed that the training or data signals are uncorrelated to the unfiltered noise processes, i.e., $\mathrm{E}[\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]\dot{\mathbf{n}}^\mathrm{H}_{\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}}[k]] = \mathbf{0}$. It is seen that this correlation matrix also generally varies with time[24], due to $\mathbf{C}_{\tilde{\mathbf{n}}}[k]$ as well as the impact of the CFO and SFO on the channel

[23] Note that for stationary white noise processes with equal power, i.e., $\mathbf{C}_{\tilde{\mathbf{n}}}[k]$ is diagonal with equal entries and constant over time, MSE and ML approach yield the same result, such that minimizing the MSE is also optimal in the ML sense.

[24] Without SFOs, but with CFOs, with ideal transmit filters, and with equal transmit and noise powers, the correlation matrix can become time invariant, i.e., $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}[k]$ does not depend on k, when, additionally, the MIMO wireless channel is optimal.

matrix $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$. The crosscorrelation matrix between output and input signal is given as

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \mathrm{E}\left[\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]\right] \tag{3.92}$$

$$= \mathrm{E}\left[\left[\mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\right]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]\right] \tag{3.93}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \mathbf{C}^{\mathrm{H}}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}[k], \tag{3.94}$$

using the same property of uncorrelatedness between data/training and noise. The autocorrelation matrix of the training or data signals is given by

$$\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \mathrm{E}\left[\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]\right] \tag{3.95}$$

$$= \mathbf{X} = \mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}, \tag{3.96}$$

which is, with the assumption of uncorrelated training signals across time and space, see section 3.2.1, and, similarly, the assumption of uncorrelated data signals across time and space, seen to be time invariant. In particular, $\mathbf{X} = \mathbf{I}_{L_\mathrm{T}} \otimes \mathrm{diag}(\sigma^2_{x_{\mathrm{Tr},1}}, \sigma^2_{x_{\mathrm{Tr},2}}, \ldots, \sigma^2_{x_{\mathrm{Tr},N}})$ contains the average transmit power of the different transmitters, i.e., $\sigma^2_{x_{\mathrm{Tr},n}} = \mathrm{E}[|x_{\mathrm{Tr},n}[k]|^2]$, which are assumed to be constant over time, and are often also assumed to be equal between different transmitters. When the data streams from the different transmit antennas are correlated as, for example, with precoding, one has to use the corresponding correlation matrix of these precoded signals. The analysis that will be carried out in the next few sections can be more complicated in such a case, since the eigenvalue structure of the correlation matrix may be more complicated. This will also be briefly discussed in the adaptive equalization case of chapter 4.

Finally, the correlation matrices of the transmitted and received signals, when the transmit power $\sigma^2_{x_{\mathrm{Tr},n}}$ is equal for every transmitter n, and when the noise power $\sigma^2_{\dot{n}_m}$ is equal for every receiver m, is given by

$$\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}} = \sigma^2_{x_{\mathrm{Tr},n}}\mathbf{I}_{NL_\mathrm{T}} \tag{3.97}$$

$$\mathbf{C}_{\tilde{\mathbf{n}}}[k] = \sigma^2_{\dot{n}_m} \cdot \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]\mathbf{H}^{\mathrm{H}}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k] \tag{3.98}$$

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \sigma^2_{x_{\mathrm{Tr},n}} \cdot \mathbf{H}_{\tilde{\mathrm{C}}}[k] \tag{3.99}$$

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}[k] = \sigma^2_{x_{\mathrm{Tr},n}} \cdot \mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{H}^{\mathrm{H}}_{\tilde{\mathrm{C}}}[k] + \sigma^2_{\dot{n}_m} \cdot \mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]\mathbf{H}^{\mathrm{H}}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]. \tag{3.100}$$

3.4.2 LMS

The LMS algorithm is one of the simplest, yet through that simplicity most powerful, adaptive algorithms, based on stochastic gradient descent. Its most basic form of adaptation can be

written as

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k+1] = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k] - \mu \cdot \left(\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k] - \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]\right)\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k] \tag{3.101}$$

$$= \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k] - \mu \cdot \dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k], \tag{3.102}$$

where μ is a step size that controls convergence and tracking speed, as well as steady-state noise performance. The algorithm works by adapting the channel estimate based on the momentary correlation between training signal $\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]$ and residual error signal $\dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]$. As described in the previous sections, if the estimate and the actual channel are equal, the error signal just consists of noise $\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]$, which is on average uncorrelated with the training signal $\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}[k]$, leading to the second term of (3.102) being zero on average and, thus, no adaptation of the estimate. If they are not equal, a correlation between the error signal and the training signal exists on average, which can be used as a gradient estimate towards which to move the estimate during the next step. The simplicity of the LMS algorithm comes from the fact that it replaces the true (averaged) correlation, between the error and training signals, with its momentary or instantaneous estimate, which makes computation of the gradient estimate very simple, but induces steady-state noise[25]. Furthermore, through this approximation of the gradient, it does not require knowledge about the exact statistics of the received signal [76, 84], which has the added benefit that it can automatically track variations in the statistics, e.g., occurring due to timing impairments as was seen in the previous section, while receiving the signal.

Some information about the statistics and time variation is still required in order to select μ properly. In the case of a time-varying MIMO system the algorithm has to solve two tasks:

1. Convergence, i.e., based on no prior information about the channel, generate a good channel estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]$.

2. Tracking, i.e., based on a converged estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]$, which is close to the true channel $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$, follow the relatively small system variations that occur over time, in other words, track the differences between $\mathbf{H}_{\tilde{\mathrm{C}}}[k], \mathbf{H}_{\tilde{\mathrm{C}}}[k+1], \mathbf{H}_{\tilde{\mathrm{C}}}[k+2], \ldots$

In the following, those two properties will be investigated more closely for the standard LMS algorithm. Other variants of the algorithm have been developed in the literature to improve convergence performance [28, 85, 86], reduce complexity, and make it more robust to input signal variations [76, 84].

[25] Even if the estimated and true channel are equal, the instantaneous estimate of the correlation between current noise realization and training signal, i.e., $\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]$, is not zero except for very high SNR cases. Only the average, i.e., $\mathrm{E}[\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]]$, tends towards zero. This residual adaptation in the algorithm even for perfect channel estimates, which occurs due to $\tilde{\mathbf{n}}_{\dot{L}_\mathrm{R}}[k]\mathbf{x}^{\mathrm{H}}_{\mathrm{Tr},L_\mathrm{T}}[k]$ not being zero for every realization, is the added steady-state noise or excess error. It is also immediately seen that in order to reduce this noise, μ should be chosen as small as possible.

3.4.3 Convergence Behavior of LMS

This property[26] is important, whenever there is a large change between the most recent and following observed realization of the channel. Two examples where this occurs are at the beginning of a transmission, where no previous estimate of the channel is available, i.e., the current realization needs to be learned from scratch, and when an event happens that changes the wireless transmission channel significantly between two observations, e.g., fast movement of transmitter or receiver, where the previous observation does not contain any useful information about the current realization. Thus, without any prior information, the initial value of the channel estimate for the LMS algorithm $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[0]$ is set to a zero matrix of appropriate size in this case.

Convergence in the time-invariant case, i.e., without timing impairments, is well covered in the literature and may be used if the channel is approximately time invariant over a certain time frame, i.e., $\mathbf{H}_{\tilde{\mathrm{C}}}[k] \approx \mathbf{H}_{\tilde{\mathrm{C}}}[k+l_{\mathrm{I}}]$ for $l_{\mathrm{I}} = \{1,\ldots,L_{\mathrm{I}}-1\}$ see appendix A.2. In this case, the standard solution for selecting μ that ensures convergence in the mean [28, 76, 84] is

$$0 < \mu < \frac{2}{\lambda_{\max}\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)} \tag{3.103}$$

$$\implies 0 < \mu < \frac{2}{\max_n \sigma^2_{x_{\mathrm{Tr},n}}}, \tag{3.104}$$

where $\lambda_{\max}(\cdot)$ is the largest eigenvalue of a matrix. The simplification in (3.104) can be made assuming that the training and data signals are uncorrelated across time and space, leading to a diagonal matrix, where all eigenvalues are equal to its diagonal values[27]. Furthermore, the misadjustment $\mathscr{M}$ [76, 84, 87], i.e., the ratio between the excess MSE due to steady-state noise

[26] It is also known as transient behavior.

[27] In practice, we may not exactly know the transmit power that is actually transmitted from the antennas and contained in the received signal. Note, however, that it can be estimated based on the received signal and the training signal, see the statistics section above.

and the MMSE of a filter derived from the normal equations, can be given as

$$\mathscr{M} = \frac{\lim_{k\to\infty} \mathrm{E}\left[\left|\dot{\mathbf{e}}_{\dot{L}_\mathrm{R}}[k]\right|^2\right] - \mathrm{E}\left[\left|\tilde{\dot{\mathbf{n}}}_{\dot{L}_\mathrm{R}}[k]\right|^2\right]}{\mathrm{E}\left[\left|\tilde{\dot{\mathbf{n}}}_{\dot{L}_\mathrm{R}}[k]\right|^2\right]} \tag{3.105}$$

$$= \mu \frac{\sum_{i=1}^{NL_\mathrm{T}} \frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}}{2-\mu\sum_{i=1}^{NL_\mathrm{T}} \frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}} = \mu L_\mathrm{T} \frac{\sum_{n=1}^{N} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}}{2-\mu L_\mathrm{T}\sum_{n=1}^{N} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \tag{3.106}$$

$$\approx \mu N L_\mathrm{T} \frac{\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}}{2-\mu N L_\mathrm{T}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \tag{3.107}$$

$$= \mu N L_\mathrm{T} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(NL_\mathrm{T}+2)}, \tag{3.108}$$

where the approximation holds, if the transmit power is equal for all Tx. It is a dimensionless quantity relative to one. In other words, the smaller the misadjustment is relative to unity, the closer the LMS algorithm performs to optimality in the MSE sense [84]. It can, thus, be substituted by $1-\epsilon$, and a further condition for μ from (3.108) can be derived. It is given by

$$\mathscr{M} = 1-\epsilon \tag{3.109}$$

$$\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}} = (1-\epsilon)\left(2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(NL_\mathrm{T}+2)\right) \tag{3.110}$$

$$\mu = \frac{1-\epsilon}{\sigma^2_{x_{\mathrm{Tr},n}}\cdot\left(NL_\mathrm{T}+1-\frac{\epsilon}{2}NL_\mathrm{T}-\epsilon\right)}. \tag{3.111}$$

A common goal is to have a misadjustment smaller than one [28], i.e., $0 \le \epsilon < 1$, which yields the following solution for μ from above

$$\mathscr{M} \le 1 \tag{3.112}$$

$$\mu_1 \le \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}}\cdot(NL_\mathrm{T}+1)}. \tag{3.113}$$

Another practical solution is to have a misadjustment of less than ten percent, i.e., $0.9 \le \epsilon < 1$. This gives

$$\mathscr{M} \le 0.1 \tag{3.114}$$

$$\mu \le \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}}\cdot\left(\frac{11}{2}NL_\mathrm{T}+1\right)}. \tag{3.115}$$

It is seen that achieving MSEs close to the MMSE, i.e., achieving negligible steady-state noise, generally requires a smaller μ than is needed just for convergence in the mean. Furthermore,

the higher the dimensions of the system[28], i.e., the number of transmit antennas N and multipath components L_T, the lower μ needs to be in order to reach the same error performance. It should also be noted that $\mathscr{M}$ is trending towards zero as $\mu \to 0$, as long as the system is time invariant. Generally, $\mathscr{M} > 1$ does not seem desirable. However, it will be seen later that such solutions can also be appropriate, when little information about the system is available.

Finally, since the channel is not actually time invariant, but changes even during the convergence phase due to CFO and SFO, it is of interest to determine the time it takes for the algorithm to move to the steady-state performance for a given μ. This time constant can be given approximately as

$$\tau_i = \frac{-1}{\ln\left(1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)\right)} \tag{3.116}$$

$$\tau_n = \frac{-1}{\ln\left(1-\mu\sigma^2_{x_{\mathrm{Tr},n}}\right)}, \tag{3.117}$$

which is to be understood in number of samples. Further, (3.113) can be used with equality and

$$\tau = \frac{-1}{\ln\left(1-\frac{1}{NL_\mathrm{T}+1}\right)} \tag{3.118}$$

$$\approx NL_\mathrm{T}, \tag{3.119}$$

can be derived, where the approximation holds if $NL_\mathrm{T} \gg 1$. It is important to note that there is a trade-off between the misadjustment $\mathscr{M}$ and the convergence time τ. The lower the desired $\mathscr{M}$, the higher the convergence time. Likewise, the larger the system dimensions, the longer the time the algorithm needs to converge to the same error performance. Nevertheless, due to the assumptions about the correlations of the training and data signals, convergence is in this case as fast as possible with the standard LMS algorithm for a selected μ.

3.4.4 Tracking Behavior of LMS

Tracking is important, when the statistics of the signals vary with time, as is the case for MIMO systems with timing impairments. In particular, from section 3.4.1 it is seen that the autocorrelation matrix of the received signal $\mathbf{C}_{\dot{\mathbf{y}}_{L_\mathrm{R}}}$, as well as the crosscorrelation matrix of the data and received signal $\mathbf{C}_{\dot{\mathbf{y}}_{L_\mathrm{R}}\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}$ vary with time. The following analysis is based on the fact that an estimate of the channel $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]$ is available, which is close to the true channel $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$. This can, for example, be achieved through the MAP/ML techniques derived in section 3.3, or through sufficiently quick convergence of the algorithm as described in the previous section. If a sufficiently accurate estimate of the channel is available at time k, the algorithm just needs

[28] Note, however, the absence of M in the equations, meaning that the updating and convergence for each received stream is decoupled and carried out in parallel, as was the case in the MAP/ML approach.

to be able to follow the variations over the next time steps. This can only be achieved if the channel changes are slower than the learning rate of the adaptive filter.

Tracking of CFOs

It is instructive to look at the two different timing impairments separately at first. Consider the model containing only CFOs, refer to section 2.5.2. The difference between two consecutive channel realizations is given as

$$\mathbf{H}_{\tilde{\mathrm{C}}}[k+1]-\mathbf{H}_{\tilde{\mathrm{C}}}[k]=\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k+1]\odot\mathbf{H}_{\tilde{\mathrm{C}}}-\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k]\odot\mathbf{H}_{\tilde{\mathrm{C}}} \tag{3.120}$$

$$=\left(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k+1]-\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k]\right)\odot\mathbf{H}_{\tilde{\mathrm{C}}} \tag{3.121}$$

$$=\left(e^{j\cdot(\arg(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k+1])-\arg(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k]))}-\mathbf{1}\right)\odot\mathbf{H}_{\tilde{\mathrm{C}}}[k] \tag{3.122}$$

$$\Longrightarrow\ \mathbf{H}_{\tilde{\mathrm{C}}}[k+1]=e^{j\cdot(\arg(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k+1])-\arg(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k]))}\odot\mathbf{H}_{\tilde{\mathrm{C}}}[k], \tag{3.123}$$

where it is seen that the changes in the channel matrices are phase shifts according to the difference in the combined phase processes from Tx and Rx from time k, where the channel is known, to time $k+1$. In order to derive some meaningful results, an assumption about the progression of the phase processes has to be made at this point. In particular, the phase processes are assumed to be sufficiently described by a linearly-increasing function[29], i.e., $\phi[k]=\mu_{\varphi_\mathrm{w}}\cdot k=2\pi\Delta f\,T_\mathrm{nom}\cdot k$ with $\sigma^2_{\varphi_\mathrm{w}}=0$ and $\sigma^2_{\varphi_\mathrm{n}}=0$, refer to section 2.4. Then, similar to [88] by assuming independent and identically distributed (i.i.d.) noise[30], the misadjustment[31] in the steady-state operation under a time-varying channel is given as

$$\mathscr{M}=\mu\frac{\sum_{i=1}^{NL_\mathrm{T}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}}{2-\mu\sum_{i=1}^{NL_\mathrm{T}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}}+\frac{1}{\mu}\frac{1}{2-\mu\sum_{i=1}^{NL_\mathrm{T}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}}\frac{\beta}{\sigma^2_{\dot{n}_m}} \tag{3.124}$$

$$\approx\mu NL_\mathrm{T}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_\mathrm{T})}+\frac{1}{\mu}\frac{1}{2-\mu NL_\mathrm{T}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}}\frac{\beta}{\sigma^2_{\dot{n}_m}}, \tag{3.125}$$

where the first term is equal to the time-invariant case from above, and the second term expresses the influence due to the time variation of the system. It is seen that w.r.t. misadjustment, or MSE in general, there exists a trade-off between the two terms. The first term

[29] It will be seen shortly that the maximum phase shift from sample to sample between any transmit and receive antenna is the crucial parameter for this derivation. This value is in general also influenced by $\sigma^2_{\varphi_\mathrm{w}}$ and $\sigma^2_{\varphi_\mathrm{n}}$, and the variation due to the Wiener process may be incorporated into that maximum phase shift, whenever it is significant.

[30] It is also assumed that the noise correlation matrix is time invariant. In actuality, however, it is cyclically varying due to the receiver sampling processes, see section 3.4.1.

[31] It has been noted [87] that for relatively fast time-varying systems, time averaged MSE or $\mathscr{M}$ may be inadequate measures for performance. Since we mostly deal with relatively slow time variations, time averaged MSE or $\mathscr{M}$ can be used as performance measures here.

represents the error due to the momentary gradient estimate, which decreases with smaller μ, while the second term expresses the error due to a lagging response of the algorithm to changes in the system, which decreases with smaller $1/\mu$. For the approximation in (3.125), the second term is expressed in the CFO case as

$$\beta = \mu\sigma^2_{x_{\mathrm{Tr},n}} \frac{\left|1 - e^{j\cdot\Delta\phi_{\max}}\right|^2}{\left|1 - e^{j\cdot\Delta\phi_{\max}} - \mu\sigma^2_{x_{\mathrm{Tr},n}}\right|^2} \left(2 - \mu\sigma^2_{x_{\mathrm{Tr},n}}\right) \tag{3.126}$$

$$\approx \frac{\Delta\phi^2_{\max}}{\mu\sigma^2_{x_{\mathrm{Tr},n}}} \left(2 - \mu\sigma^2_{x_{\mathrm{Tr},n}}\right), \tag{3.127}$$

where $\left|\Delta\phi_{\max}\right| = \max\left|\arg\left(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k+1]\right) - \arg\left(\mathbf{\Delta\Phi}_{\tilde{\mathrm{C}}}[k]\right)\right| = \max_{mn}\left|2\pi\Delta f_{mn} T_{\mathrm{Rx},m}\right| = \left|2\mu_{\varphi_{\mathrm{w}}}\right|$ is the maximum phase shift that occurs between any transmitter and receiver from sample to sample due to the CFOs, and the approximation can be made for $\Delta\phi_{\max} \ll 1$ and $\Delta\phi_{\max} \ll \mu\sigma^2_{x_{\mathrm{Tr},n}}$. Using (3.125) and (3.127), two approaches can be taken in order to find good values for μ. The first proceeds, as in the previous section, by defining that the misadjustment should be smaller than one, yielding the following equation

$$1 \geq \mathscr{M} \tag{3.128}$$

$$\begin{aligned} \Longrightarrow\ 0 \geq{}& (2NL_{\mathrm{T}} + 2)\cdot\mu^3 + \left(\frac{\Delta\phi^2_{\max}}{\sigma^2_{\dot{n}_m}} - \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}}\right)\cdot\mu^2 \\ &- \Delta\phi^2_{\max}\frac{3}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\cdot\mu + \Delta\phi^2_{\max}\frac{2}{\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}, \end{aligned} \tag{3.129}$$

where the full derivation can be found in appendix B. The second approach is to try to minimize $\mathscr{M}$ w.r.t. μ, i.e.,

$$0 = \frac{\partial\mathscr{M}}{\partial\mu} \tag{3.130}$$

$$\begin{aligned} \Longrightarrow\ 0 = {}& \left(\frac{2NL_{\mathrm{T}}}{\sigma^2_{x_{\mathrm{Tr},n}}} + \Delta\phi^2_{\max}\frac{2 + NL_{\mathrm{T}}}{\sigma^2_{\dot{n}_m}}\right)\cdot\mu^3 - \Delta\phi^2_{\max}\frac{6\,(2 + NL_{\mathrm{T}})}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\cdot\mu^2 \\ &+ \Delta\phi^2_{\max}\frac{6\,(3 + NL_{\mathrm{T}})}{\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\cdot\mu - \Delta\phi^2_{\max}\frac{8}{\sigma^6_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}, \end{aligned} \tag{3.131}$$

with the full derivation being available in appendix B.1. Thus, in both cases a cubic polynomial needs to be solved for a positive real root. The solutions of these polynomials can be obtained, but are for general parameters fairly complicated.

Considering very small values for μ, specifically $\mu \ll \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}} NL_{\mathrm{T}}}$ see appendix B, more functional solutions can be obtained. The two derivations simplify to

$$0 \geq NL_{\mathrm{T}}\cdot\mu^3 - \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}}\cdot\mu^2 + \Delta\phi^2_{\max}\frac{2}{\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}, \tag{3.132}$$

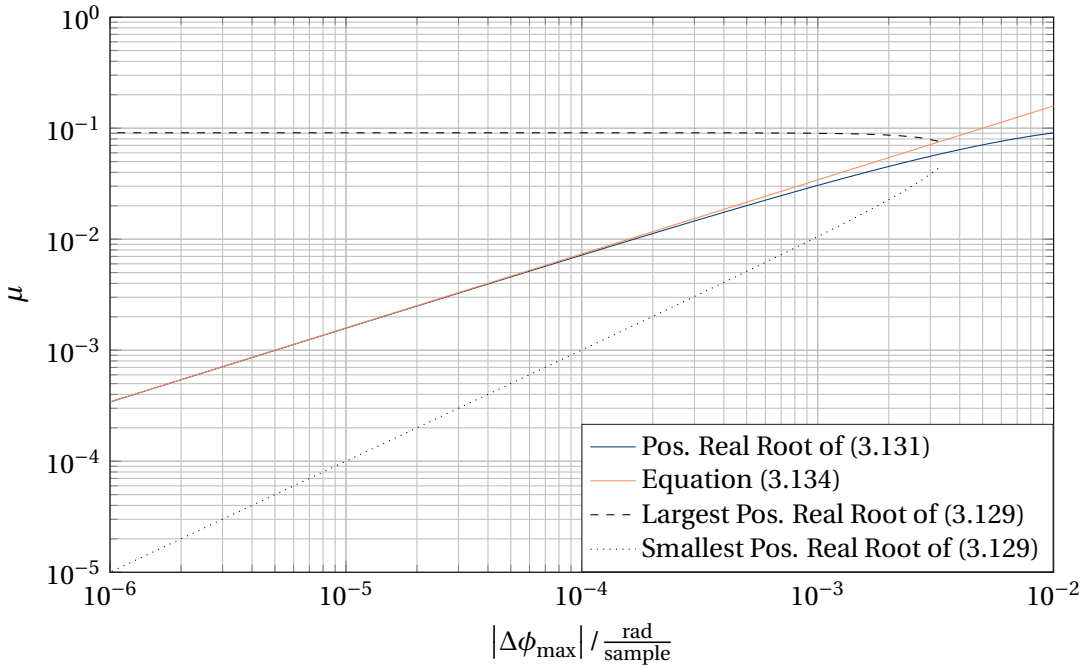


Figure 3.3: Comparison of the different solutions for selecting the step size μ of an LMS algorithm for channel estimation for MIMO systems with multiple CFOs, i.e., setup 4 where the maximum carrier frequency difference between any transmitter and receiver is $|\Delta\phi_{\max}|$. The parameter settings for this example are $\sigma^2_{x_{\text{Tr},n}} = 1$, $\sigma^2_{\dot{n}_m} = 0.01$, with a system dimension of $N \cdot L_{\text{T}} = 10$.

and

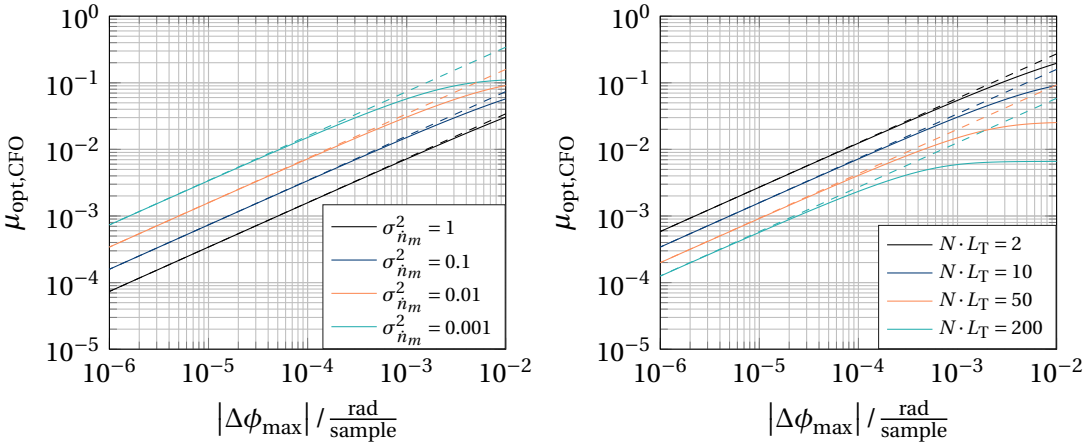
$$0 = NL_{\text{T}} \cdot \mu^3 - \Delta\phi^2_{\max} \frac{4}{\sigma^4_{x_{\text{Tr},n}} \sigma^2_{\dot{n}_m}}, \tag{3.133}$$

respectively. For the first approach using $\mathscr{M} \leq 1$, we still need to solve a cubic polynomial, even when using this approximation. For the second approach, a simple equation to choose μ optimally can directly be given, which is

$$\mu_{\text{opt,CFO}} = \left(4 \cdot \frac{\Delta\phi^2_{\max}}{\sigma^4_{x_{\text{Tr},n}} \sigma^2_{\dot{n}_m} \cdot NL_{\text{T}}} \right)^{\frac{1}{3}}. \tag{3.134}$$

Given that the difference in the carrier phase processes is very small, it can, furthermore, be seen from (3.129) and also (3.132) that a solution yielding $\mathscr{M} \leq 1$ can be obtained by selecting μ according to, e.g., (3.113), from the previous section.

In Figure 3.3, the different solutions for selecting μ in a MIMO system with multiple CFOs are compared. It is seen that for the chosen parameters, equation (3.134) approximates the true optimal value for μ, which is found by solving the cubic polynomial in (3.131), very closely if the CFOs are not too high, i.e., roughly if $|\Delta\phi_{\max}| < 10^{-3}$ in this case. The two roots of (3.129)



(a) $\sigma^2_{x_{\text{Ir},n}} = 1$, $N \cdot L_{\text{T}} = 10$

(b) $\sigma^2_{x_{\text{Ir},n}} = 1$, $\sigma^2_{\dot{n}_m} = 0.01$

Figure 3.4: Optimal step size μ for channel estimation for MIMO systems with multiple CFOs, i.e., setup 4 where the maximum carrier frequency difference between any transmitter and receiver is $|\Delta\phi_{\max}|$, under different parameter settings. The normal lines represent the positive real root of (3.131), while the dashed lines show the approximation given in (3.134).

give a range for μ where $\mathscr{M} \leq 1$. Furthermore, the larger root of (3.129) is almost equal to the solution for the time-invariant case, i.e., equation (3.113) with equality, and the actual optimum value for μ also seems to converge to that value. This last observation holds for a lot of practical parameter settings. Thus, when no information about the CFOs $\Delta\phi_{\max}$ or the noise powers $\sigma^2_{\dot{n}_m}$ is available, we may select μ according to (3.113) with equality. A similar observation was also made in [89]. An even better value for μ in such a scenario, particularly for large system sizes $N \cdot L_{\text{T}}$, can be obtained with

$$\mu_2 = \frac{4}{3} \cdot \frac{1}{\sigma^2_{x_{\text{Ir},n}} \cdot N L_{\text{T}}}, \tag{3.135}$$

see appendix B.1, which is very similar, but always greater than the solution in (3.113). It follows that this solution always has a misadjustment that is slightly larger than one, as will be seen later on.

Figure 3.4 shows the optimal values for μ according to the solutions from (3.131) and (3.134) (dashed) for different parameter settings. It is seen that the approximation for selecting μ in (3.134) holds well for small CFOs, low SNRs, and medium system sizes. Again, in cases where the approximation fails, we can get a value for μ in a simple manner according to (3.135). One important observation is that as the SNR increases, for a fixed maximum CFO, the optimum step size also increases. This is due to the fact that the higher the SNR is, the more impact the tracking/lag error has, which is reduced when μ is increased. The optimal value for μ finds the best balance between tracking error and noisiness of the gradient estimate.

From the optimal solution for μ in (3.134) and under the used approximations, we can derive

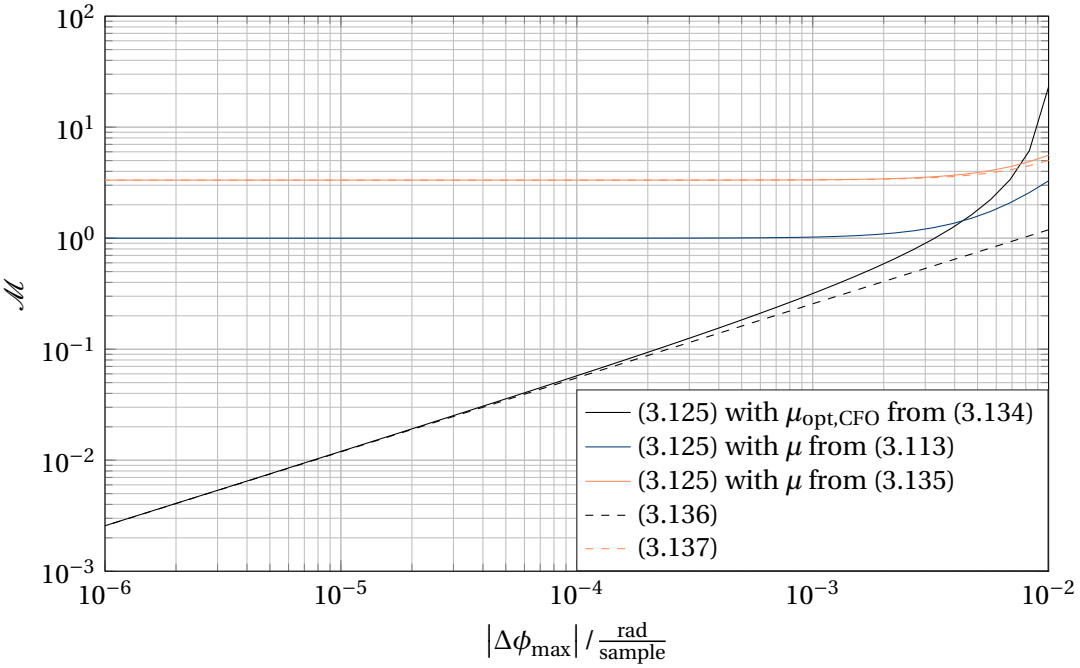


Figure 3.5: Comparison of the misadjustment of an LMS algorithm for channel estimation in MIMO systems with multiple CFOs, i.e., setup 4 where the maximum carrier frequency difference between any transmitter and receiver is $|\Delta\phi_{\max}|$, for different selections of the step size μ. Parameter settings for the shown example are $\sigma^2_{x_{\text{Tr},n}} = 1$, $\sigma^2_{\dot{n}_m} = 0.01$, with a system dimension of $N \cdot L_{\text{T}} = 10$.

the approximate misadjustment in this case, which is given by

$$\mathscr{M}_{\text{opt,CFO}} \approx \left(\frac{27}{16} \cdot \Delta\phi^2_{\max} \cdot \frac{\sigma^2_{x_{\text{Tr},n}}}{\sigma^2_{\dot{n}_m}} \cdot N^2 L^2_{\text{T}} \right)^{\frac{1}{3}}, \tag{3.136}$$

see also appendix B.1. Likewise, for the simple selection of the step size given in (3.135), the misadjustment can be approximated by

$$\mathscr{M} \approx 2 \cdot \frac{NL_{\text{T}}}{NL_{\text{T}} - 4} + \frac{27}{16} \cdot \Delta\phi^2_{\max} \cdot \frac{\sigma^2_{x_{\text{Tr},n}}}{\sigma^2_{\dot{n}_m}} \cdot N^2 L^2_{\text{T}}, \tag{3.137}$$

refer also appendix B. Figure 3.5 shows the misadjustment for different μ selection rules, based on the general $\mathscr{M}$ of (3.125) and the exact value[32] for the β term, i.e., (3.126). As before, the approximations for the misadjustment hold well for low CFO values. The selection solution for μ from the time-invariant case, i.e., equation (3.113) with equality, achieves $\mathscr{M} = 1$, while selecting μ according to (3.135) yields slightly higher misadjustment. Generally, it can be said that the misadjustment increases with the SNR, CFO, and system size, even when selecting

[32] Although it is not shown here, the difference between using (3.126) and (3.127) was checked for the parameter range shown here. It is significant, when roughly $|\phi_{\Delta,\max}| > 10^{-3}$ and $N \cdot L_{\text{T}} > 100$ occurs simultaneously.

μ optimally. Especially, the increase with the SNR seems somewhat counter intuitive, but it should be remembered that $\mathscr{M}$ is a relative quantity w.r.t. a solution from the normal equations. Thus, the MSE of the LMS algorithm may still decrease with the SNR, as will be seen in some of the results later on. Example 5 gives some practical, numerical values for a mmWave LoS MIMO system that can be can be obtained with the results from this section.

Example 5. *Consider a mmWave LoS MIMO system with $M = N = 2$ antennas on each side, $L_T = 5$ multi-path components, yielding $N \cdot L_T = 10$, and a sampling period of $T_{nom} = 1$ ns. It was seen in section 2.4, that a typical CFO value for a 60* GHz *oscillator is $|\Delta\phi_{\max}| = 10^{-3}$ in this case. Assume the transmit power to be $\sigma^2_{x_{Tr,n}} = 1$, and the noise power to be $\sigma^2_{\dot{n}_m} = 0.01$, yielding a SNR of 20* dB.

First, examine the case where we do not know the CFO level and noise power. Then, it is suggested to select μ according to (3.113) *or* (3.135). *With the former one, $\mu = 9.1 \cdot 10^{-2}$, leading to a convergence time from* (3.117) *of $\tau \approx 10$, which is to be understood in number of iterations, being equivalent to a time of 10* ns *here. From* (3.125), *the misadjustment for this μ is calculated as $\mathscr{M} \approx 1$. A step size selection according to* (3.135) *gives $\mu = 1.33 \cdot 10^{-1}$, $\tau \approx 7$, or 7* ns, *and $\mathscr{M} \approx 3.35$. It is thus seen that the former has significantly better error performance, at a slight increase in convergence time. A step size with less than ten percent misadjustment is given, for the time-invariant case, from* (3.115) *by $\mu = 1.8 \cdot 10^{-2}$, yielding $\tau \approx 55$, or 55* ns, *and $\mathscr{M} \approx 0.441$. It is seen that this solution does not achieve $\mathscr{M} = 0.1$, when CFOs are present.*

Secondly, consider the case, where we have information about the CFO level and noise power. We then select μ using (3.134). *This yields $\mu = 3.4 \cdot 10^{-2}$, $\tau \approx 29$, or 29* ns, *and $\mathscr{M} \approx 0.317$. This choice for μ has the best error performance and only requires a moderate time to converge.*

Tracking of SFOs

Consider the model containing only SFOs, refer to section 2.5.3. The difference between two consecutive channel realizations is then given by

$$\mathbf{H}_{\tilde{\mathrm{C}}}[k+1] - \mathbf{H}_{\tilde{\mathrm{C}}}[k] = \left(\mathbf{H}_{\tilde{\mathrm{C}}} + \boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k+1]\right) - \left(\mathbf{H}_{\tilde{\mathrm{C}}} + \boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k]\right) \tag{3.138}$$

$$= \boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k+1] - \boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k] \tag{3.139}$$

$$\implies \mathbf{H}_{\tilde{\mathrm{C}}}[k+1] = \mathbf{H}_{\tilde{\mathrm{C}}}[k] + \left(\boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k+1] - \boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k]\right). \tag{3.140}$$

The actual entries of $\boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k]$ cannot be directly determined in this case, as they are dependent on the actual wireless channel, as well as the transmit and receive filters that are used in the FEs, and vary according to the SFO processes. Two possible approaches can be used in order to model $\boldsymbol{\Delta}\dot{\boldsymbol{\Phi}}_{\tilde{\mathrm{C}}}[k]$. The first one is deterministic, similar to the CFO case, by using a constant that describes the change between two consecutive channel observations. The approach stems from the idea that the SFOs generate, in time-invariant channels, cyclical variations, which can be expressed in this manner. However, the values that need to be added depend on the already mentioned properties of the system, and are, hence, problem specific. For this reason,

this approach will not be further pursued here.

The other, more general, approach that has also been studied under several aspects in the literature [76, 84, 88], assumes the variations as a random process following a Markov model, specifically, an autoregressive model of order one. It can be described by

$$\mathbf{H}_{\tilde{\mathrm{C}}}[k+1] = \alpha \cdot \mathbf{H}_{\tilde{\mathrm{C}}}[k] + \boldsymbol{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}}, \tag{3.141}$$

where $0 \leq \alpha \leq 1$ is a parameter that controls the rate of change, usually very close to one, and $\boldsymbol{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}}$ are stationary random variables with some distribution and correlation properties. In order to derive values for the model parameters α and $\boldsymbol{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}}$, it needs to be determined how the complete channel matrix changes, as the sampling phases vary due to the SFO. This can be done by checking how the correlation properties of the different channel tap entries vary with the sampling phase processes, and matching them to the correlation properties of the autoregressive process. Given that the increments $\boldsymbol{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}}$ of each matrix entry follow a white noise series, one matrix entry of the model process has the following correlation function w.r.t. the lag l from the current observation

$$\mathrm{E}\left[h_{\tilde{\mathrm{C}},mn}[k, l_{\tilde{\mathrm{C}}}] h^*_{\tilde{\mathrm{C}},mn}[k-l, l_{\tilde{\mathrm{C}}}]\right] = \frac{\sigma^2_{\Delta\dot{\phi}}}{1-\alpha^2} \cdot \alpha^{|l|}. \tag{3.142}$$

It is seen that the correlation vanishes exponentially with the distance to the current channel entry depending on α, and that the first term is equal to the variance of the process. For the actual SFO behavior, the correlation also decreases first but would eventually increase again, giving rise to the cyclical nature of the process. Thus, when the autoregressive model is used to derive results, the estimation should always be able to track this first decrease in correlation, such that it always stays in a state where the approximation is valid. Since the variance of the channel entries should be fixed and similar to the amplitude of each channel tap entry, regardless of the level of SFO in the system, we use $\sigma^2_{\Delta\dot{\phi}} = \left(1-\alpha^2\right) \cdot \sigma^2_{h_{\tilde{\mathrm{C}},mn}[l_{\tilde{\mathrm{C}}}]}$. Qualitatively, the higher the SFOs in the system are, the lower α should be.

This autoregressive model has been used in the literature to model to model the Doppler spreading that is experienced in fading channels [4, 28, 88]. In this case, the correlation is assumed to follow a Bessel function, such that $\alpha = J_0\left(2\pi \cdot f_{\mathrm{D}} \cdot T_{\mathrm{nom,Rx}}\right)$, where $J_0(\cdot)$ is the Bessel function of the first kind and zeroth order, and f_{D} is the maximum Doppler frequency. In the SFO case, a similar argument w.r.t. spreading can be made. On the transmitter side, the signals bandwidth is increased or compressed according to the sampling phase process, compared to the nominal ideal process, by a factor of $\frac{1}{1+\dot{\phi}_{\mathrm{Tx},n}[k]}$. Since the exact shape of the sampling phase processes is not known for every system configuration, it will, as in the CFO case, be assumed that it is sufficient to investigate the fastest sampling phase drift, based on a linear model, i.e., $\left|\dot{\phi}_{\mathrm{max,Tx}}\right| = \max_n \left|\dot{\phi}_{\mathrm{Tx},n}[k+1] - \dot{\phi}_{\mathrm{Tx},n}[k]\right| = \left|\mu_{\dot{\phi}_{\mathrm{w}}}\right|$ assuming $\dot{\phi}[k] = \mu_{\dot{\phi}_{\mathrm{w}}} \cdot k$. As in the CFO case this means $\sigma^2_{\dot{\phi}_{\mathrm{w}}} = 0$ and $\sigma^2_{\dot{\phi}_{\mathrm{n}}} = 0$, refer also to section 2.4. The spread w.r.t. the desired signal bandwidth is then given by $\frac{|\dot{\phi}_{\mathrm{max,Tx}}|}{1-|\dot{\phi}_{\mathrm{max,Tx}}|}$, which yields a similar relation as in the Doppler

spreading case, with

$$\alpha = J_0\left(2\pi \cdot \frac{\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}{1-\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}\frac{1}{T_s}\cdot T_{\mathrm{nom,Rx}}\right) \tag{3.143}$$

$$= J_0\left(2\pi \cdot \frac{\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}{1-\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}\cdot\frac{1}{Q}\right). \tag{3.144}$$

This seems to suggest that the receiver sampling phase process has no influence on the correlation. However, the system is not sampling with the nominal rate, but with an impaired time reference, refer to section 2.4. Then, using the same assumptions and notation about the fastest sampling phase drift in the receiver, we can write

$$\alpha = J_0\left(2\pi \cdot \frac{\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}{1-\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|}\cdot\frac{1}{Q}\cdot\left(1+\left|\dot{\phi}_{\max,\mathrm{Rx}}\right|\right)\right), \tag{3.145}$$

where $\left|\dot{\phi}_{\max,\mathrm{Tx}}\right| < 1$ and $\left|\dot{\phi}_{\max,\mathrm{Rx}}\right| < 1$. Thus, as desired, the correlation decreases faster with higher differences in the sampling phase processes. Finally, in order to fully describe the autoregressive process, the correlation matrix of the increments $\mathbf{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}}$ needs to be determined. It was described previously that they follow a complex Gaussian distribution. The simplifying assumption[33] that the channel variations are uncorrelated between different receive antennas, and, also, between the different transmit antennas, as well as the different channel taps, will be used. Then, the correlation matrix of any row of the increments sufficiently describes the process, and is given by

$$\mathbf{C}_{\mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}} = \mathrm{E}\left[\mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}\mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}^{\mathrm{H}}\right] = \left(1-\alpha^2\right)\cdot\mathrm{diag}\left(\sigma^2_{\dot{h}_{\tilde{\mathrm{C}},m1}[1]},\ldots,\sigma^2_{\dot{h}_{\tilde{\mathrm{C}},mN}[L_{\tilde{\mathrm{C}}}]}\right), \tag{3.146}$$

with $\mathbf{\Delta\dot{\Phi}}_{\tilde{\mathrm{C}}} = \left[\mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},1} \quad \cdots \quad \mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},M}\right]^{\mathrm{T}}$, and using the variance normalization mentioned previously.

Given the two model parameters α and $\mathbf{C}_{\mathbf{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}$, which are used to express the time-varying behavior due to the SFOs, the goal is now to try and find good values for μ that can deal with these timing impairments. As in the previous section, the misadjustment can first be written as [84, 88]

$$\mathscr{M} = \mu\frac{\sum_{i=1}^{NL_{\mathrm{T}}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}}{2-\mu\sum_{i=1}^{NL_{\mathrm{T}}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}} + \frac{1}{\mu}\frac{1}{2-\mu\sum_{i=1}^{NL_{\mathrm{T}}}\frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}}[k]\right)}}\frac{\beta}{\sigma^2_{\dot{n}_m}} \tag{3.147}$$

$$\approx \mu NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_{\mathrm{T}})} + \frac{1}{\mu}\frac{1}{2-\mu NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}}\frac{\beta}{\sigma^2_{\dot{n}_m}}, \tag{3.148}$$

[33] For oversampled systems, there will at least be correlation on the channel tap level. However, this is a worst-case assumption that yields usable results for any case.

where the time-varying contribution is now given differently [88] with

$$\beta = \mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)\left(1+\frac{\left(1-\alpha^2\left(1-\mu\sigma^2_{x_{\mathrm{Tr},n}}\right)^2\right)}{\left(\alpha(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})-1\right)^2}\frac{(1-\alpha)^2}{1-\alpha^2}-\frac{2(\alpha-1)(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})}{\alpha(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})-1}\right) \tag{3.149}$$

$$\approx \left(1-\alpha^2\right)\left(1+\frac{\left(1-\alpha^2\left(1-\mu\sigma^2_{x_{\mathrm{Tr},n}}\right)^2\right)}{\left(\alpha(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})-1\right)^2}\frac{(1-\alpha)^2}{1-\alpha^2}-\frac{2(\alpha-1)(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})}{\alpha(1-\mu\sigma^2_{x_{\mathrm{Tr},n}})-1}\right), \tag{3.150}$$

where the second approximation holds for our definition of the correlation matrix $\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}$, and assuming unit sum power of all channel taps. The general solution for this β is complex, but can be computed for a given parameter set. In the following, some of the limiting cases for α, which lead to easier solutions, will be investigated. Consider first the case of $1-\alpha \approx 0$, i.e., small SFO, which yields

$$\beta = \mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right) \approx 1-\alpha^2 \tag{3.151}$$

$$\mathscr{M} = \mu N L_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_{\mathrm{T}})}+\frac{1}{\mu}\frac{1}{2-\mu NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}}\frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{\dot{n}_m}}. \tag{3.152}$$

Using the same approaches as for the CFO case we get

$$1 \geq \mathscr{M} \tag{3.153}$$

$$\implies 0 \geq -2\left(1+NL_{\mathrm{T}}\right)\cdot\mu^2+\left(\frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{\dot{n}_m}}+\frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}}\right)\cdot\mu-\frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \tag{3.154}$$

$$0 = \frac{\partial\mathscr{M}}{\partial\mu} \tag{3.155}$$

$$\begin{aligned}\implies 0 = &\left(2NL_{\mathrm{T}}-\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}}\left(2+NL_{\mathrm{T}}\right)\right)\cdot\mu^2+\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)\frac{2\left(2+NL_{\mathrm{T}}\right)}{\sigma^2_{\dot{n}_m}}\cdot\mu\\ &-2\frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}.\end{aligned} \tag{3.156}$$

The solution to these quadratic equations is obtainable, but rather complex. Assuming, as in the CFO case, that $\mu \ll \frac{2}{\sigma^2_{x_n}NL_{\mathrm{T}}}$ yields

$$\mathscr{M} \approx \mu N L_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2}+\frac{1}{\mu}\frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{2\sigma^2_{\dot{n}_m}}, \tag{3.157}$$

which simplifies the two approaches to solving the simpler quadratic equations

$$0 \geq NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} \cdot \mu^2 - \mu + \frac{\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\check{\mathrm{C}},m}}\right)}{2\sigma^2_{\dot{n}_m}}, \tag{3.158}$$

and

$$0 = NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} \cdot \mu^2 - \frac{\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\check{\mathrm{C}},m}}\right)}{2\sigma^2_{\dot{n}_m}}. \tag{3.159}$$

The second approach then leads, similar to the CFO case, to a usable expression, given by

$$\mu_{\mathrm{opt,SFO}} = \sqrt{\frac{\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\check{\mathrm{C}},m}}\right)}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m} \cdot NL_{\mathrm{T}}}} \tag{3.160}$$

$$\approx \sqrt{\frac{1-\alpha^2}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m} \cdot NL_{\mathrm{T}}}}, \tag{3.161}$$

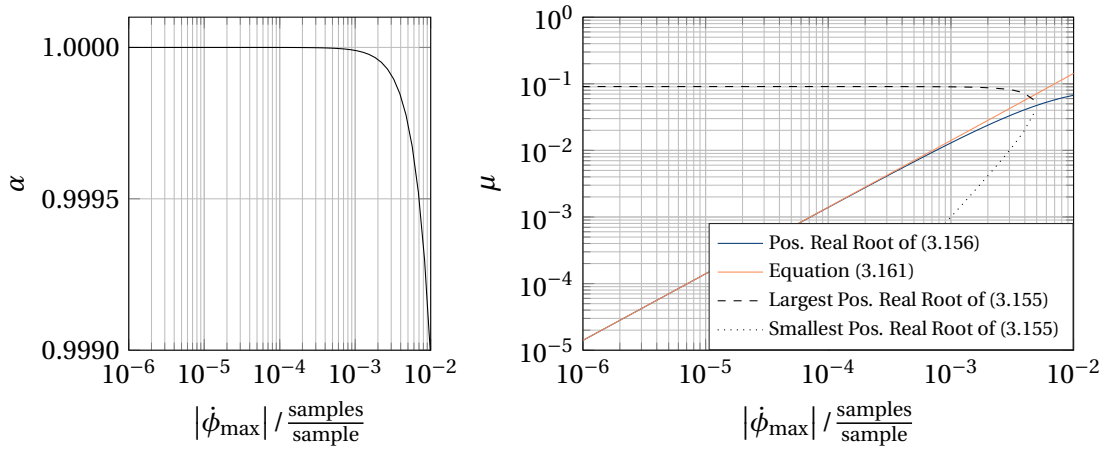
which has also been obtained in the literature. From this equation, it is seen that as expected, the lower α, i.e., the higher the SFOs $\left|\dot{\phi}_{\max}\right|$, the higher μ needs to be. Note that this expression only holds for small SFOs, i.e., α close to one or $\left|\dot{\phi}_{\max}\right| < 10^{-2}$, and small μ, as will be seen next. It is of particular importance, to check if $\mu_{\mathrm{opt,SFO}} \ll \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}} NL_{\mathrm{T}}}$ is true for the computed $\mu_{\mathrm{opt,SFO}}$ and the given parameters. From (3.157) and (3.161) we can, furthermore, calculate the misadjustment for this optimal choice of μ, which is given by

$$\mathscr{M}_{\mathrm{opt,SFO}} \approx \left(\left(1-\alpha^2\right) \cdot \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}} \cdot NL_{\mathrm{T}} \right)^{\frac{1}{2}}. \tag{3.162}$$

Figure 3.6 presents some example results for the correlation parameter α, as well as the selection of μ for the relevant SFO range[34], using the different solutions. The first plot shows the α values that correspond to the maximum SFOs $\left|\dot{\phi}_{\max}\right| = \left|\dot{\phi}_{\max,\mathrm{Tx}}\right| = \left|\dot{\phi}_{\max,\mathrm{Rx}}\right|$ using equation (3.145), where it is seen that for practical SFO values, α is indeed very close to one. The second figure has a similar shape as in the CFO case, where the selection of μ according to (3.161) is proper for low SFOs, i.e., $\left|\dot{\phi}_{\max,\mathrm{Rx}}\right| < 10^{-3}$. As in the CFO case it is seen that there is a fixed value for μ that generates a misadjustment that is comparable to one and independent of the SFO. It can be used when no information about the order of magnitude of $\left|\dot{\phi}_{\max,\mathrm{Tx}}\right|$ and $\left|\dot{\phi}_{\max,\mathrm{Rx}}\right|$ is available, and is given by

$$\mu = \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_{\mathrm{T}}}, \tag{3.163}$$

[34] Recall, for example, the $\mu_{\dot{\phi}_{\mathrm{w}}}$ values given for a 60 GHz system with a sample duration of 1 ns in section 2.4, which were between 10^{-6} and 10^{-3}.



(a) $Q = 1$

(b) $Q = 1$, $N \cdot L_{\mathrm{T}} = 10$, $\sigma^2_{x_{\mathrm{Tr},n}} = 1$, $\sigma^2_{\dot{n}_m} = 0.01$

Figure 3.6: Selection of the step size μ for channel estimation for MIMO systems with multiple SFOs under different parameter settings: Left plot shows the correlation parameter α w.r.t. practical SFO values $|\dot{\phi}_{\max}| = |\dot{\phi}_{\max,\mathrm{Tx}}| = |\dot{\phi}_{\max,\mathrm{Rx}}|$ according to equation (3.145); Right plot shows different selections of μ for the given parameters.

see appendix B.2.

Now, consider the case of $\alpha \approx 0$, i.e., very high SFOs of $|\dot{\phi}_{\max,\mathrm{Rx}}| > 10^{-1}$ that generate channels, which are basically uncorrelated from observation to observation. Similar to the CFO case, this approach can also be used when no information about the order of magnitude of $|\dot{\phi}_{\max,\mathrm{Tx}}|$ and $|\dot{\phi}_{\max,\mathrm{Rx}}|$ is available. This assumption leads to

$$\beta = \mu \mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)\sigma^2_{x_{\mathrm{Tr},n}} \approx \mu\left(1-\alpha^2\right)\sigma^2_{x_{\mathrm{Tr},n}} \tag{3.164}$$

$$\mathscr{M} = \mu N L_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_{\mathrm{T}})} + \frac{1}{2-\mu NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}}. \tag{3.165}$$

The general solutions, if they are obtainable, lead again to fairly complicated quadratic equations, yielding little insight. Using $\mu \ll \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}} NL_{\mathrm{T}}}$ with the approach of $1 \geq \mathscr{M}$, leads to the the following solution

$$\mu \leq \frac{2}{\sigma^2_{x_{\mathrm{Tr},n}} NL_{\mathrm{T}}} - \frac{\mathrm{Tr}\left(\mathbf{C}_{\boldsymbol{\Delta\dot{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{\dot{n}_m} NL_{\mathrm{T}}}. \tag{3.166}$$

It is seen that for low values of α, it becomes difficult to find good values for μ with the autoregressive model, as the observations are almost uncorrelated. In the most extreme case of $\alpha = 0$, the channel is completely uncorrelated from observation to observation, and thus the best estimate for the current channel $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ would be the instantaneous correlation between input and output, and neglecting the previous estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k-1]$. In such a case, the LMS

algorithm is no longer usable, as no tracking can be performed. However, such low α values do not seem relevant for the practical range of SFO values, see Figure 3.6, which is why this behavior is less critical. In Example 6, we give some numerical values for the selection of the step size in a mmWave LoS MIMO system with sampling phase variations, which can be obtained from the results in this section.

Example 6. *Assume a mmWave LoS MIMO system with $M = N = 2$ antennas on each side, $L_T = 5$ multi-path components, yielding $N \cdot L_T = 10$, a sampling period of $T_{nom} = 1\,\text{ns}$, and a symbol period of $T_s = 2\,\text{ns}$, yielding $Q = 2$. It was seen in section 2.4, that a typical SFO value for sampling rates in a 60 GHz transmission system is $|\Delta\dot{\phi}_{\max,Tx}| = |\Delta\dot{\phi}_{\max,Rx}| = |\Delta\dot{\phi}_{\max}| = 10^{-5}$. The transmit power is assumed to be $\sigma^2_{x_n} = 1$, and the noise power is assumed to be $\sigma^2_{\dot{n}_m} = 0.01$, yielding a SNR of 20 dB. When no information about the SFO level and SNR is available, the step size can be selected as mentioned in the first part of Example 5.*

Consider the case, where we have information about the SFO level and noise power. We then select the step size μ using (3.145) *and* (3.161)*. This gives $\mu \approx 7 \cdot 10^{-5}$, $\tau \approx 1.43 \cdot 10^4$, or 14.3 µs, and, from* (3.162)*, $\mathscr{M} \approx 7 \cdot 10^{-4}$. It is seen that both the suggested optimal step size μ and the misadjustment is lower than in the CFO case. For practical cases, where the variation due to the SFOs is often significantly smaller than the variation due to the CFOs, we can theorize that the performance of the adaptive filter is limited by the CFOs. This claim will be further investigated in the next section.*

Dealing with Both Timing Impairments

The total misadjustment, taking into account both CFO and SFO, can generally be computed [88], but takes on a rather complex form making it difficult to obtain solutions for μ. It is possible to use the sum of (3.127) and (3.151) for a combined β describing both effects, but also in such a case the expressions for μ are complicated. Thus, qualitative arguments for selecting μ under different scenarios will be made in the following.

Firstly, consider the case where no information about the CFOs and SFOs, which exist in the system, is available. The solution when only one of them exists, but no information about their order of magnitude is accessible, were given in (3.135) and (3.163), respectively. Accordingly, in this case it is suggested to use

$$\mu_2 = \frac{4}{3} \cdot \frac{1}{\sigma^2_{x_{\text{Tr},n}} \cdot N L_{\text{T}}}, \tag{3.167}$$

which is the larger of the two solutions, i.e., assuming the worst case scenario. Correspondingly, Figure 3.7 shows the misadjustment when μ is selected according to this equation, as well as the misadjustments for the cases of only CFOs or only SFOs, using their respective approximately optimal solutions from (3.134) and (3.161), thus, employing knowledge about their order of magnitude. It is seen that in cases where CFOs and SFOs are of a similar magnitude, the misadjustment due to CFOs is generally higher. Secondly, in cases where $|\phi_{\max}| \geq |\dot{\phi}_{\max}|$, i.e.,

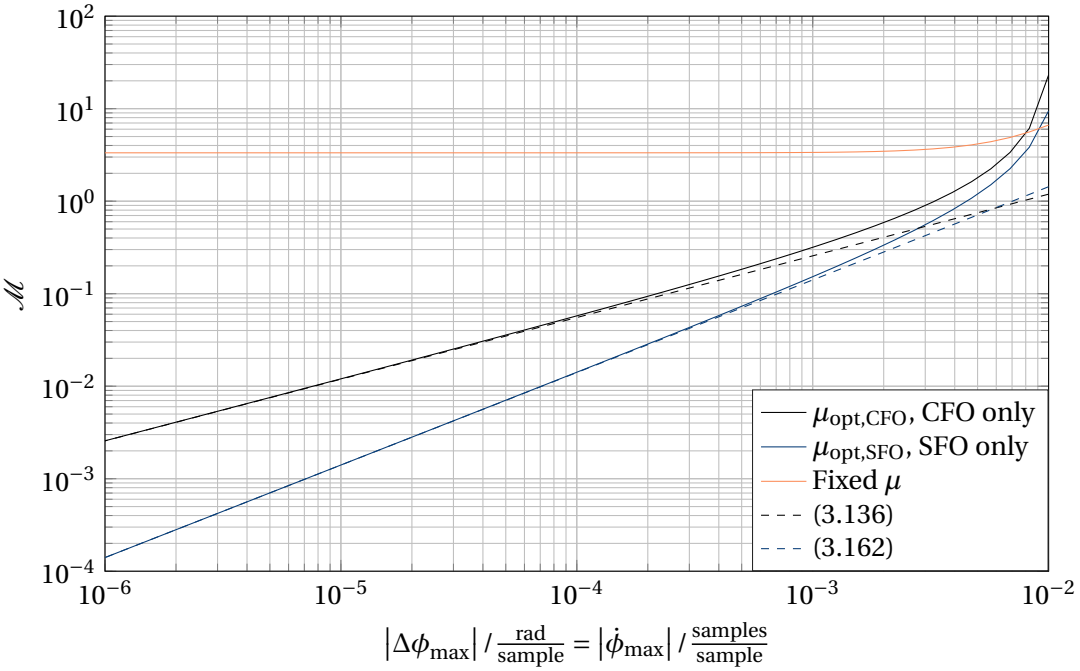


Figure 3.7: Comparison of the misadjustment of an LMS algorithm for channel estimation in MIMO systems with multiple CFOs, with a maximum value of $|\Delta\phi_{\max}|$, or SFOs, with a maximum value of $|\dot{\phi}_{\max}|$, for different selections of the step size μ. Parameter settings for the shown example are $\sigma^2_{x_{\text{Tr},n}} = 1$, $\sigma^2_{\dot{n}_m} = 0.01$, $N \cdot L_{\text{T}} = 10$, $Q = 1$, and $|\dot{\phi}_{\max}| = |\dot{\phi}_{\max,\text{Tx}}| = |\dot{\phi}_{\max,\text{Rx}}|$.

the SFOs are of similar or smaller magnitude compared to the CFOs, it is suggested to select μ according to the CFO case, given by

$$\mu_{\text{opt,CFO}} = \left(4 \cdot \frac{\Delta\phi^2_{\max}}{\sigma^4_{x_{\text{Tr},n}} \sigma^2_{\dot{n}_m} \cdot N L_{\text{T}}} \right)^{\frac{1}{3}}. \tag{3.168}$$

Thirdly, whenever the SFOs are significantly higher than the CFOs, i.e., $|\phi_{\max}| \ll |\dot{\phi}_{\max}|$, it is suggested to use the optimal solution of the SFO case, defined as

$$\mu_{\text{opt,SFO}} = \sqrt{\frac{1-\alpha^2}{\sigma^2_{x_{\text{Tr},n}} \sigma^2_{\dot{n}_m} \cdot N L_{\text{T}}}}. \tag{3.169}$$

Finally, whenever the approximations, that were used to derive these simplified optimal solutions do not hold, we may solve the polynomial function of the respective case.

3.4.5 Oversampling in LMS

So far, it was not specifically discussed whether the system is oversampled or not. Looking at the previous sections, it should be clear that it does not matter for the steady-state estimation

and tracking of the time-varying channel, except in terms of complexity, if the reference signal $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]$ and the channel to be estimated $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ are oversampled or not. It was seen that the performance of the adaptive filter only depends on the correlation matrix of the input signal $\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]$, the timing impairments, the SNR, and the step size μ. In particular, it was seen that the misadjustment of an adaptive filter, based on the LMS principle, depends in several ways on the term

$$\mathscr{M} \propto \sum_{i=1}^{NL_\mathrm{T}} \frac{\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)}{1-\mu\lambda_i\left(\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]\right)},$$

i.e., the eigenvalue spread of the input signals correlation matrix. Thus, it is interesting to investigate what the difference in $\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]$ is, when the input signal is symbol spaced or oversampled.

First, consider the case where the input signal to the channel and filter is not assumed oversampled[35]. Given the assumptions about the training signal from section 3.2.1 and assuming equal transmit power across all antennas, the correlation matrix is, as in section 3.4.1,

$$\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k] = \mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}} = \sigma^2_{x_{\mathrm{Tr},n}}\mathbf{I}_{NL_\mathrm{T}}. \tag{3.170}$$

The correlation matrix is, thus, time invariant and diagonal with equal entries, which are, for a diagonal matrix, also equal to the eigenvalues of that matrix. Thus, all of the results that were computed in the previous sections for selecting μ, and the corresponding misadjustment for MIMO systems with multiple timing impairments can be used.

Secondly, assume the case where the input signal to the channel and adaptive filter is already pulse shaped, i.e., $\dot{\mathbf{x}}_{\dot{L}_\mathrm{T}}[k]$, see also section 2.5.6. The main benefit of using the oversampled input signal is that the filter does not need to learn the transmit pulse shape, which is usually known, and thus the filter length can potentially be reduced. The correlation matrix for this case follows, using the same assumptions as above, as

$$\mathbf{C}_{\dot{\mathbf{x}}_{\mathrm{Tr},\dot{L}_\mathrm{T}}}[k] = \mathrm{E}\left[\dot{\mathbf{x}}_{\mathrm{Tr},\dot{L}_\mathrm{T}}[k]\dot{\mathbf{x}}^\mathrm{H}_{\mathrm{Tr},\dot{L}_\mathrm{T}}[k]\right] \tag{3.171}$$

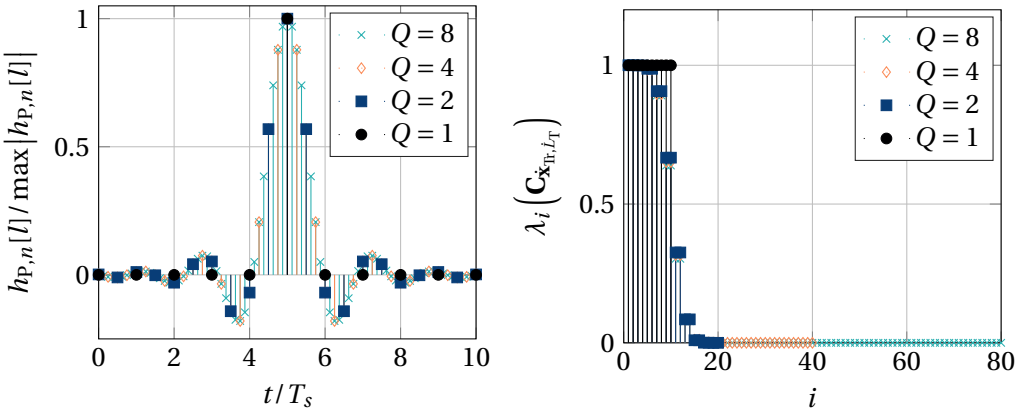
$$= \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{Q}\mathbf{H}_\mathrm{P}\mathbf{H}^\mathrm{H}_\mathrm{P}, \tag{3.172}$$

where $\mathbf{H}_\mathrm{P}$ is a block-Toeplitz matrix of the transmit pulse-shaping filters. It can be noted that they are still uncorrelated among different transmit streams, and for most practical filters also only have little correlation w.r.t. to symbol shifts, see Example 7.

Example 7. *Consider a set of two uncorrelated (across time and space) symbol streams, i.e.* $N = 2$*, that are each pulse shaped, except for the case of* $Q = 1$*, by a root-raised-cosine filter with roll-off* $\beta_T = 0.3$*, and with a length of 10 symbols, for different oversampling factors* Q*. An*

[35] The actual signal that is generated in baseband and then transmitted may still be oversampled (or pulse shaped), but from the adaptive filters perspective that filter is seen as part of the complete channel.

observation interval of $L_T = 5$ symbol periods is investigated, and $\sigma^2_{x_{Tr,n}} = 1$. Figure 3.8 shows the normalized pulse shape for these cases, as well as the corresponding eigenvalues of the correlation matrix in these cases. It is seen that regardless of the oversampling factor, there is only a finite number of significant eigenvalues, all of them being similar in magnitude. A lot of the insignificant eigenvalues are analytically zero, but will practically have some finite value. Nevertheless, their contribution to the overall steady-state performance is still negligible.



(a) (b)

Figure 3.8: Normalized pulse shape for different oversampling factors (left), and the corresponding eigenvalues of the correlation matrix (right), when using these pulse shapes on two uncorrelated symbol sequences across time and space.

From the example it is seen that the eigenvalues of the correlation matrix depend in the oversampled case on the pulse shape, specifically on the roll-off factor β_T, see also [90]. Looking at the relationship between $\mathscr{M}$ and the eigenvalues of the correlation matrix $\lambda_i\left(\mathbf{C}_{\mathbf{x}_{Tr,L_T}}[k]\right)$ it is seen that the largest eigenvalues have the highest contribution, and small ones, or ones that are zero, can be neglected. In the symbol-spaced case, there were $N \cdot L_T$ significant eigenvalues, corresponding to the power of each transmit stream. In the oversampled case, there are $N \cdot L_T \cdot (1 + \beta_T)$ significant eigenvalues, corresponding to the transmit power of each stream shaped by the pulse form. Thus, as $\beta_T \to 0$, the oversampled case becomes equal to the symbol-spaced case, and the results from the previous sections can be used. For high values of β_T, more significant eigenvalues appear, but their magnitude is reduced. It can be checked that the sum of eigenvalues from the beginning of this section, which determines $\mathscr{M}$, is approximately the same, irrespective of the roll-off β_T. This means that even when considering an oversampled input signal to the adaptive filter, all of the results that were derived in the previous sections can be employed. Note, however, that the convergence properties of the filter do depend on the complete eigenvalue structure. Typically, a large eigenvalue spread means that the adaptive filter converges slower, as the convergence speed is limited by the smallest eigenvalues [28, 76, 84]. To summarize, the convergence behavior of an adaptive filter is influenced by the eigenvalue spread caused by an oversampled input, while the steady-state error and tracking performance is not, at least for the channel estimation (or

system identification) case.

One further comment about the eigenvalues of oversampled inputs to an adaptive filter is important. The eigenvalues of a correlation matrix of several oversampled signals are determined by their spectral shape, i.e., their pulse shape, and the correlation between those several signals, as was seen in the example. This observation will become useful in the next chapter, where a version of the LMS algorithm is used for adaptive equalization, and the input signal is the oversampled received signal $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k]$, see [90] for the SISO case. In this case, the complete MIMO transmission channel has an influence on the eigenvalues [24, 28], see also section 3.4.1.

3.5 Simulation Results

In what follows, some exemplary results for the estimators that have been derived in this chapter will be given. The channel characteristics will be selected to fit mmWave LoS MIMO systems in terms of multi-path, e.g., delay spread in the nanosecond range and significantly reduced power of multi-path components see [12, 53, 91, 92, 93, 94], and timing impairments w.r.t. a symbol duration of $T_s = 1\,\mathrm{ns}$, see for example Table 2.1. In particular, the channel will be assumed[36] of length[37] $L_\mathrm{C} = L = 10$ with a Rician factor K_R of 10 dB. The LoS MIMO tap, i.e., $l_\mathrm{C} = 0$, will be considered orthogonal, i.e., designed according to (2.20) and following (2.12), while the other $L_\mathrm{C} - 1$ taps will be drawn from a complex Gaussian distribution, yielding[38]

$$h_{mn}[l_\mathrm{C} = 0] = \sqrt{\frac{10^{\frac{K_\mathrm{R}}{10}}}{1 + 10^{\frac{K_\mathrm{R}}{10}}}} \cdot e^{-j2\pi \frac{r_{mn,0}}{\lambda_{\mathrm{Tx},n}}} \tag{3.173}$$

$$h_{mn}[l_\mathrm{C} \neq 0] \sim \mathcal{CN}\left(0, \frac{1}{1 + 10^{\frac{K_\mathrm{R}}{10}}}\right), \tag{3.174}$$

while enforcing normalization of the channel power given by $\sum_{l_\mathrm{C}} |h_{mn}[l_\mathrm{C}]|^2 = 1$. Depending on the correlation properties of the training sequences across time and space, the channel characteristics can have an influence on the estimation performance, as will be seen in the results. When the training sequences are uncorrelated across space, which will usually be the case here, the orthogonality of the MIMO channel has no influence on the estimation.

[36] As briefly mentioned in the beginning of the adaptive filter section 3.4, the channel length is typically not known, but needs to be estimated. The efficiency of both estimation approaches can only be high if that estimate is of reasonable quality. In this work, we assume that such an estimate is available. In general, this is a model order selection problem [84, 95].

[37] With the mentioned symbol duration, this is comparable to a delay spread of 10 ns.

[38] In the simulations using these definitions, the wireless propagation channel is, thus, frequency flat and time invariant, when $K_\mathrm{R} \to \infty$. In other words, for this case it is a pure LoS channel where the channel entries are defined by the geometric arrangement through (2.12). For finite values of K_R, the propagation channel is frequency selective and time variant. For the simulations, the selectivity, i.e., influence of NLoS taps, is given by K_R, and the interval between changes in the NLoS taps, i.e., the time variation, was set to $10 \cdot L_\mathrm{Tr}$ symbols for the corresponding scenario. With the mentioned symbol duration and L_Tr values in the range of one hundred, this could be compared to a coherence time of 1 µs, or, similarly, a Doppler spread of approximately 420 kHz.

However, if information about the spatial characteristics of the channel, e.g., the antenna arrangements, is available, it may be exploited to improve estimation performance, as will be seen in the second part of section 3.5.1. The focus will be on symmetric antenna setups, i.e., $M = N$, as they yield the highest throughputs with the smallest array size in LoS MIMO systems [13]. As training sequences, the practically relevant pseudo-noise sequences, and Zadoff-Chu sequences [81], which approximately fulfill the criteria outlined in section 3.2.1, will be considered.

3.5.1 Channel and CFO Estimation in the CFO Impaired Case

Consider the case of a MIMO system that is affected by $M+N$ CFOs, i.e., oscillator setup 4, and is symbol spaced as described in section 2.5.2. As primary training, a Zadoff-Chu sequence of length L_C that is separated in time among the N different transmit antennas is considered. In order to obtain the ideal autocorrelation property of the sequence, it needs to be cyclically repeated at least once, and $L_\mathrm{C}-1$ zero symbols after the repeated training sequences are required for each antenna, to remove the interference between the different antennas due to the multi-path. This yields to a total training sequence length of $L_\mathrm{Tr} = N \cdot (L_\mathrm{C} \cdot 2 + L_\mathrm{C} - 1)$. Since the higher channel coefficients carry significantly lower power in mmWave systems, their influence is limited and we may reduce the length to $L_\mathrm{Tr} = N \cdot L_\mathrm{C} \cdot 2 + L_\mathrm{C} - 1$. Note that the average training power normalization $\mathrm{E}\left[|x_{\mathrm{Tr},n}[l_\mathrm{Tr}]|^2\right] = \sigma^2_{x_{\mathrm{Tr},n}} = 1$, and the channel normalization described above, lead to better estimation performance with an increase in M, N, and L_Tr, approximately proportional to the increase of that parameter. Since the CRBs are very complicated for this system setup, a benchmark scenario with ideal impulse-like training sequences across time and space of the same length and power is computed for comparison, which is thus a lower bound[39] on the achievable performance. Especially when the phase processes take on a more complicated shape, this bound helps to judge performance. For estimation of the channel (3.43) is employed, while for estimation of the CFO differences (3.45) with $L_\mathrm{E} = L_\mathrm{Tr}$ is used. To reduce complexity further, the latter only uses the LoS taps, i.e., $\hat{h}_{\mathrm{C},mn}[k,0]$, which carry most of the power, for the phase estimation.

First, assume purely linearly-drifting phase processes, i.e., $\sigma^2_{\varphi_\mathrm{w}} = 0$ and $\sigma^2_{\varphi_\mathrm{n}} = 0$, where the mean variations and the initial oscillator phases are drawn from $\mu_{\varphi_\mathrm{w},m/n} \sim \mathcal{U}(-\mu_{\varphi_\mathrm{w}}, \mu_{\varphi_\mathrm{w}})$ and $\phi_\mathrm{w}[0] \sim \mathcal{U}(-\pi,\pi)$. Thus, the maximum average CFO difference that can occur in the system is $\pm 2\mu_{\varphi_\mathrm{w}}$. The estimation performance for $M = N = 4$ and $L_\mathrm{C} = 10$ for channels and CFOs, using $L_\mathrm{Tr} = 89$, is given in Figure 3.9. Both the pure LoS case, i.e., $K_\mathrm{R} \to \infty$, and the standard case, i.e., $K_\mathrm{R} = 10\,\mathrm{dB}$, are shown. The MSE results show a performance gap of around 3 dB for both parameters in the LoS case, which is likely due to the required repetition of the Zadoff-Chu sequence, where only one of them can be used for reliable estimation. Although the MSE is generally low, some performance floors and performance gaps compared to the

[39] This approach essentially decouples the MIMO system into $M \cdot N$ SISO systems with impulse-like training of equivalent transmit power per transmit antenna. The results shown later on are then equivalent to the CRBs, for example [3, 4], if available for the considered system.

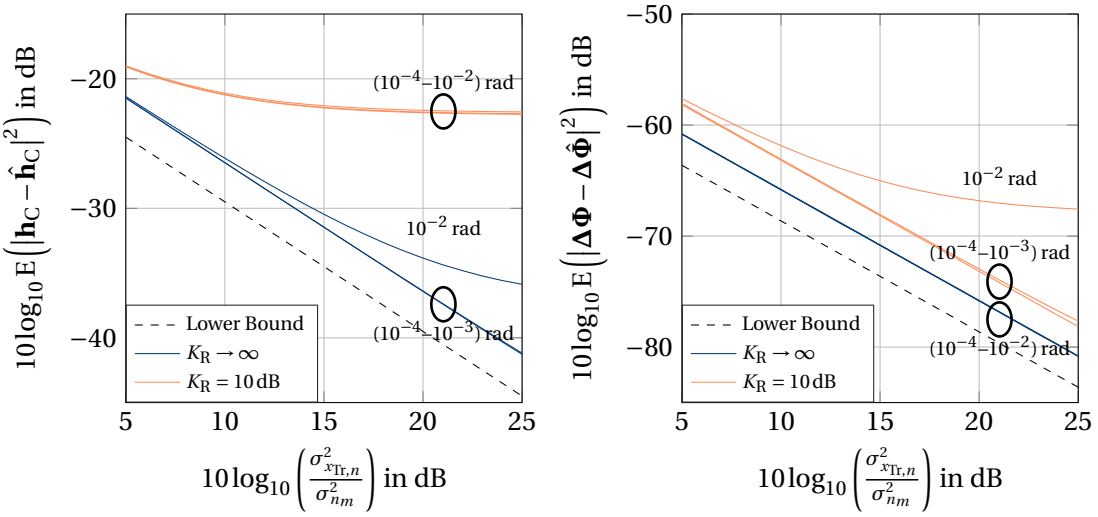


Figure 3.9: MSE of estimating the channels and CFOs with (3.43) and (3.45) for a symbol-spaced LoS MIMO system affected by independent linearly-drifting carrier phase processes, with $L_\mathrm{C}=10$ taps, $M=N=4$ antennas, and using Zadoff-Chu training of length $L_\mathrm{Tr}=L_\mathrm{E}=89$. Channel estimation error for different Rician factors K_R and normalized CFO values μ_{φ_w} (left). Phase difference estimation error for different K_R and μ_{φ_w} (right).

lower bound can be observed for the estimation of both parameters. For the estimation of the channel taps, it is seen that CFO values of $\mu_{\varphi_\mathrm{w}}=10^{-2}\,\mathrm{rad}$ are high enough that the approximation $\mathbf{h}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},1}[k]\approx\mathbf{h}^\mathrm{T}_{\mathrm{C},L_\mathrm{C},1}[k+l_\mathrm{Tr}]$, or $\boldsymbol{\phi}_\mathrm{Tx}[k]\approx\boldsymbol{\phi}_\mathrm{Tx}[k+l_\mathrm{Tr}]$ and $\boldsymbol{\phi}_\mathrm{Rx}[k]\approx\boldsymbol{\phi}_\mathrm{Rx}[k+l_\mathrm{Tr}]$ for $l_\mathrm{Tr}=\{1,\ldots,L_\mathrm{Tr}-1\}$, is not valid anymore and, thus, the correlation with the training sequence is influenced by the CFOs. The error floor in the case of channel estimation for $K_\mathrm{R}=10\,\mathrm{dB}$ is due to the ISI between the training sequences of the different antennas, i.e., the reduction of the training length from $L_\mathrm{Tr}=N\cdot(L_\mathrm{C}\cdot 2+L_\mathrm{C}-1)$ to $L_\mathrm{Tr}=N\cdot L_\mathrm{C}\cdot 2+L_\mathrm{C}-1$. The gap in the phase difference MSE between the two K_R cases is explained by the fact that only the LoS tap is used for the estimation, which carries different amounts of power for both cases, see the previous section. For CFO estimation, it should also be noted that the condition $L_\mathrm{E}\leq\left\lceil\frac{\pi}{4\mu_{\varphi_\mathrm{w}}}\right\rceil$ is not quite fulfilled for the highest CFO case, but since this is a conservative condition, the estimation still yields usable results. Generally, the MSE of all plotted scenarios is low enough to not significantly influence the performance of an equalizer. However, if more accurate estimations are required, simple solutions are: increase of the training length L_Tr, while removing the ISI between the antennas, and performing an iterative estimation, where carrier phase difference estimates are obtained first, which are then removed from the signal before carrying out the standard estimation procedures used here. For the pure LoS case, i.e., $K_\mathrm{R}\to\infty$, the results can also be compared to the asCRBs mentioned in section 2.6.1. It is seen that the asCRBs are higher than the lower bound used here, and slightly lower than the MSEs of the estimators in the LoS case.

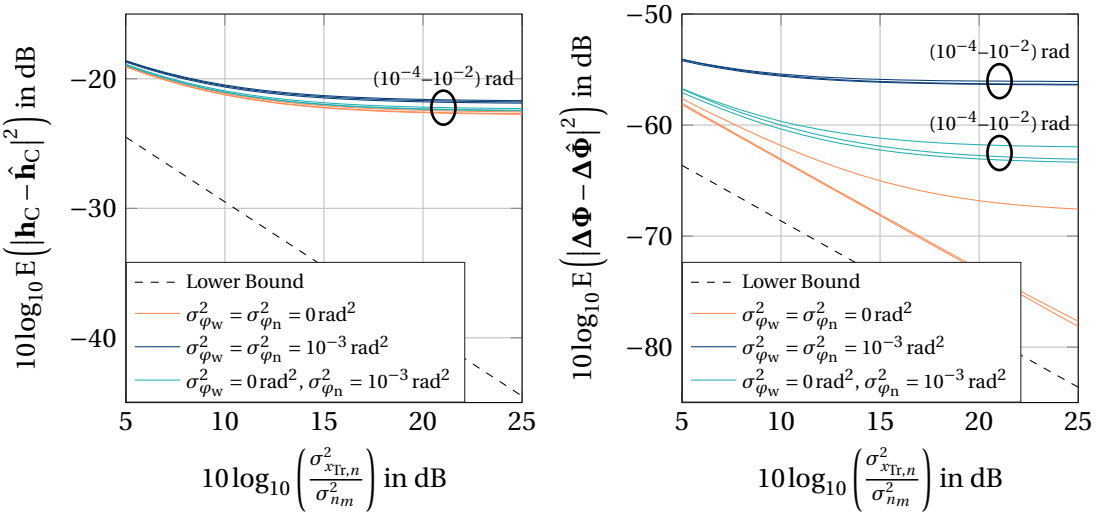
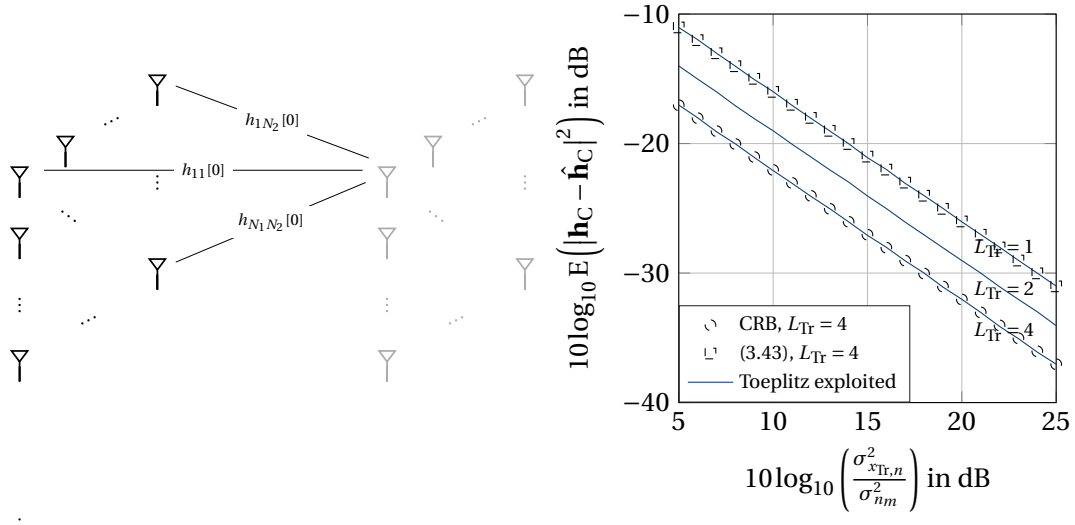


Figure 3.10: MSE of estimating the channels and CFOs with (3.43) and (3.45) for a symbol-spaced LoS MIMO system affected by independent practical carrier phase processes, i.e., $\sigma^2_{\varphi_\mathrm{w}}=\sigma^2_{\varphi_\mathrm{n}}\neq 0$, with $L_\mathrm{C}=10$ taps, Rician factor $K_\mathrm{R}=10\,\mathrm{dB}$, $M=N=4$ antennas, and using Zadoff-Chu training of length $L_\mathrm{Tr}=L_\mathrm{E}=89$. Channel estimation error for different normalized CFO values μ_{φ_w} (left). Phase difference estimation error for different μ_{φ_w} (right).

Secondly, consider more complex phase processes using $\sigma^2_{\varphi_\mathrm{w}}=\sigma^2_{\varphi_\mathrm{n}}=10^{-3}\,\mathrm{rad}^2$, with all other parameters being the same as above. Figure 3.10 shows the MSE results for $K_\mathrm{R}=10\,\mathrm{dB}$, also including the lower bounds and the results for the linearly-drifting phase process from above. For channel estimation, a slight increase of the MSE is observed, which comes from the reduction of the maximum correlation gain due to the influence of the more randomly varying phases. The change in the estimation performance of the carrier phase differences is much more pronounced. It is seen that especially the Wiener contribution to the phase process, i.e., the phase variation due to close-in PN, significantly increases the error. The white noise contribution, i.e., far-out PN, does also create an error floor but has less impact as it can be averaged out to some extent. The same performance improvement suggestions as in the previous case can be used. However, depending on the impact of the Wiener process during the length of the training sequence, a performance limit may exist[40].



(a) Uniformly-spaced Tx & Rx antenna arrays

(b) MSE for channel estimation

Figure 3.11: LoS MIMO channel estimation exploiting the structure of the channel: (a) Antenna array setup generating the block-Toeplitz structure; (b) MSE for channel estimation without CFOs exploiting the block-Toeplitz structure. A Rician factor of $K_\mathrm{R}\to\infty$, $M=N=4$ antennas, and different training lengths L_Tr were used for the results.

Exploiting the Structure of Pure LoS MIMO in Frequency-Flat Channels

Consider the case of $L_\mathrm{C}=1$ and $K_\mathrm{R}\to\infty$ with ideal transmit and receive filters, i.e., $h_{\mathrm{Tx},n}(t)=h_{\mathrm{Rx},m}(t)=\mathrm{sinc}\left(\pi\frac{t}{T_\mathrm{nom,Rx}}\right)$. It was seen in section 2.2 that the LoS MIMO channel depends highly on the geometric setup of the antennas. If the setup of the antennas is known with a very high precision, specifically their relative distances, the channel can be calculated and no estimation is necessary. In practice, measuring (or manufacturing) those distances precisely is difficult and, thus, some form of channel estimation is always helpful. Nevertheless, it is still possible to exploit the knowledge about the geometric structure of the arrays. For all approximately optimal, symmetric ($M=N$), uniform array designs, the LoS channel matrix $\mathbf{H}[0]$ will have approximately a block-Toeplitz structure, i.e.,

$$\mathbf{H}[0]=\begin{bmatrix}\mathbf{H}_1[0] & \mathbf{H}_2[0] & \cdots & \mathbf{H}_{N_1}[0]\\ \mathbf{H}_2[0] & \mathbf{H}_1[0] & \cdots & \mathbf{H}_{N_1-1}[0]\\ \mathbf{H}_3[0] & \mathbf{H}_2[0] & \cdots & \mathbf{H}_{N_1-2}[0]\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{H}_{N_1}[0] & \mathbf{H}_{N_1-1}[0] & \cdots & \mathbf{H}_1[0]\end{bmatrix}, \tag{3.175}$$

[40] Consider the case, where the linear drift and the white noise part of the phase noise are negligible. This means that we only need to consider the Wiener contribution with variance $\sigma^2_{\varphi_\mathrm{w}}$. The variance of a Wiener process at time l_Tr is given by $l_\mathrm{Tr}\cdot\sigma^2_{\varphi_\mathrm{w}}$. We can see that the longer the training sequence, the more correlation gain is lost due to the Wiener process. The exact impact depends on the total length of the training sequence and the variance of the Wiener process.

where N_1 is the number of elements in the first array dimension and the sub-matrices of index n_1 have a Toeplitz structure with

$$\mathbf{H}_{n_1}[0] = \begin{bmatrix} h_{n_1 1}[0] & h_{n_1 2}[0] & \cdots & h_{n_1 N_2}[0] \\ h_{n_1 2}[0] & h_{n_1 1}[0] & \cdots & h_{n_1 N_2-1}[0] \\ h_{n_1 3}[0] & h_{n_1 2}[0] & \cdots & h_{n_1 N_2-2}[0] \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_1 N_2}[0] & h_{n_1 N_2-1}[0] & \cdots & h_{n_1 1}[0] \end{bmatrix}, \tag{3.176}$$

where N_2 is the number of elements in the second array dimension. Note that there are N_1 different sub-matrices with N_2 different entries. An example of a two-dimensional Tx/Rx antenna array setup with this structure is given in Figure 3.11a. It can, more generally, be said that for uniformly-spaced antenna arrangements, including non-symmetric cases, there will be a part of the channel matrix that has a block-Toeplitz structure and a part that depends on the specific chosen arrangement and alignment of the arrays.

From this structure, we can see that the first column or row of the channel matrix determines the complete matrix[41]. In other words, estimating the channel based on one received stream, e.g., using (3.43) with $m = 1$, is sufficient and the other ones can, for example, be used for averaging in order to improve the quality of the estimate. Figure 3.11b shows the estimation performance for an optimal LoS MIMO setup, i.e., designed according to (2.20), without CFOs and with $M = N = 4$, in terms of the MSE. The same normalizations as above are used, but the training length is reduced to $L_{\mathrm{Tr}} = N$. The CRB assuming this type of channel, see [96], as well as the standard solution from above not exploiting the structure are given as references. For the Toeplitz exploiting results, averaging was used over the appropriate entries [96], when $L_{\mathrm{Tr}} > 1$. It is seen that significant gains can be achieved by exploiting the structure of the channel. For different numbers of antennas, including the case of a single CFO, and also a discussion on the sensitivity of this approach w.r.t. to slightly imperfect channels, see [96]. In general, there is a trade-off regarding the potential gain of this strategy. The higher the number of antennas, the higher the gain of exploiting the structure. However, the higher the number of antennas, the more susceptible the setup becomes with respect to misalignments and small spacing offsets, as will be seen in the experimental results of chapter 5, which reduces the Toeplitz structure of the matrix.

3.5.2 Channel, CFO, and SFO Estimation in the General Case

Consider the general case of an oversampled received MIMO signal that is affected by both $M + N$ CFOs and $M + N$ SFOs. The same framework and lengths, in terms of number of symbols, as in section 3.5.1 are used. The oversampling factor is set as $Q = 2$, and on the transmit side a root-raised-cosine filter of 10 symbol durations with a roll-off $\beta_{\mathrm{T}} = 0.25$ is

[41] This only holds if oscillator setup 2 is used, while additionally assuming the same phase for the PLLs on the Tx/Rx side, respectively. Likewise, if the phase effects due to the oscillators have been sufficiently compensated, the structure is also observed.

used for pulse shaping. Note that the absolute lengths are increased approximately by the oversampling factor, i.e., $\dot{L}_{\mathrm{C/Tr/E}} \approx Q \cdot L_{\mathrm{C/Tr/E}}$, as mentioned previously. In order to make a fair comparison to the previous results, the parameters that define the carrier and sampling phase processes have to be slightly adopted. In particular, it is assumed that when the oversampling factor Q is increased, the phase variation per sample decreases[42]. Thus, the values that are mentioned in this section for μ_{φ_w}, $\mu_{\dot{\varphi}_\mathrm{w}}$, $\sigma^2_{\varphi_\mathrm{w}}$, and $\sigma^2_{\dot{\varphi}_\mathrm{w}}$ are to be understood per symbol. The sampling phase processes are defined by $\dot{\phi}_\mathrm{w}[0] \sim \mathscr{U}(-0.5, 0.5)$ and $\mu_{\dot{\varphi}_\mathrm{w},m/n} \sim \mathscr{U}(-\mu_{\dot{\varphi}_\mathrm{w}}, \mu_{\dot{\varphi}_\mathrm{w}})$, meaning that the maximum average sampling phase difference per symbol is $\pm 2\mu_{\dot{\varphi}_\mathrm{w}}$. When SFOs are present, the different antennas generate slightly different symbol rates, which means that after some time full sample, and eventually symbol, shifts will occur, as was discussed previously, see Figure 2.8 and appendix A.3. While the estimators that have been derived in this chapter are applicable to this case, as long as the shift is smaller than the window length $\dot{L}_\mathrm{C}$, the Zadoff-Chu sequence of the different transmit antennas will eventually overlap, creating interference in the correlation that is used for channel estimation. A simple remedy to this problem is to use different Zadoff-Chu sequences, having low crosscorrelation with each other [81], for different antennas. In order to use the same training signals as above, the results shown here, hence, only consider estimation of the relative difference on aligned signals, i.e., where full sample shifts have been removed, which are in the range of $-0.5 \leq \Delta\dot{\phi}_{mn}[k] \leq 0.5$.

For channel estimation (3.82) is used, while (3.45) with $\dot{L}_\mathrm{E} = \dot{L}_\mathrm{Tr}$ is used for the estimation of the carrier phase difference processes. As in the previous section, the latter one only uses the LoS taps in order to limit complexity. For estimation of the sampling phase differences, (3.76) is used in the case of almost frequency-flat channels, while (3.84) is used in other cases. Note that the second method only computes the relative variations of the phase processes over time, while the fixed phases offsets are also needed to obtain the correct symbol timings. In other words, this method yields the variations in the sampling over time, but not the initial starting sampling phases, whereas (3.76) computes the offset w.r.t. to the ideal time instants.

Figure 3.12 contains the MSEs for the estimation of channels and CFOs in a LoS MIMO system with $K_\mathrm{R} = 10\,\mathrm{dB}$. It contains results for both purely linearly-drifting processes, i.e., $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} = \sigma^2_{\dot{\varphi}_\mathrm{w}} = \sigma^2_{\dot{\varphi}_\mathrm{n}} = 0$, and more complex phase processes, i.e., $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} =\neq 0$ and $\sigma^2_{\dot{\varphi}_\mathrm{w}} = \sigma^2_{\dot{\varphi}_\mathrm{n}} \neq 0$. It is seen that the addition of the SFOs deteriorates the performance somewhat, compared to the results in Figure 3.10. The channel estimation MSE is increased by a couple of dBs, while the CFO MSE has almost the same behavior as in that previous case. It is also seen that the more complex phase processes slightly increase the MSE. As in the pure CFO case, the

[42] This is equivalent to assuming that fixed continuous-time phase processes exist, which do not change when the sample rate is changed. This also conforms with the conversion equations in appendix A.1, where a change in T_nom accordingly scales these values.

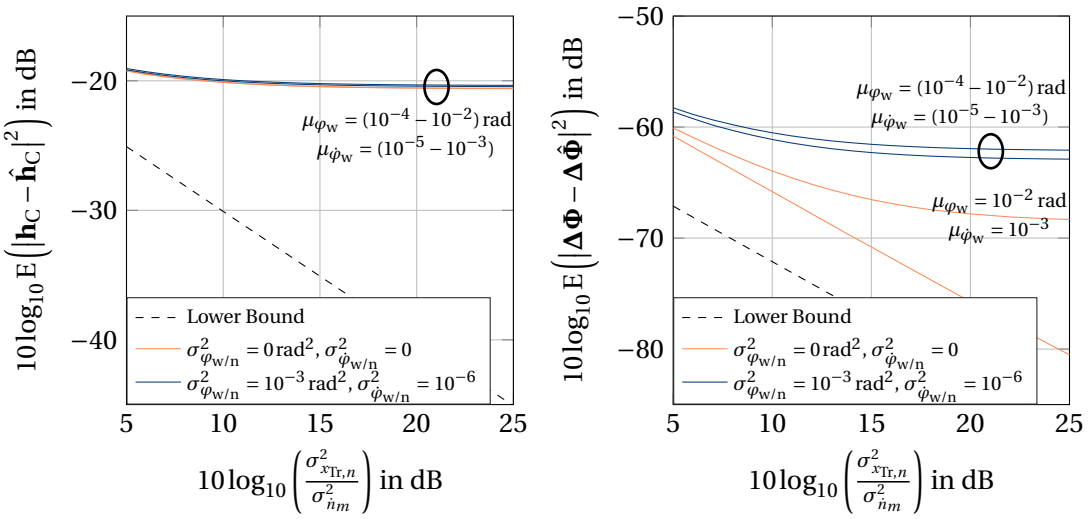
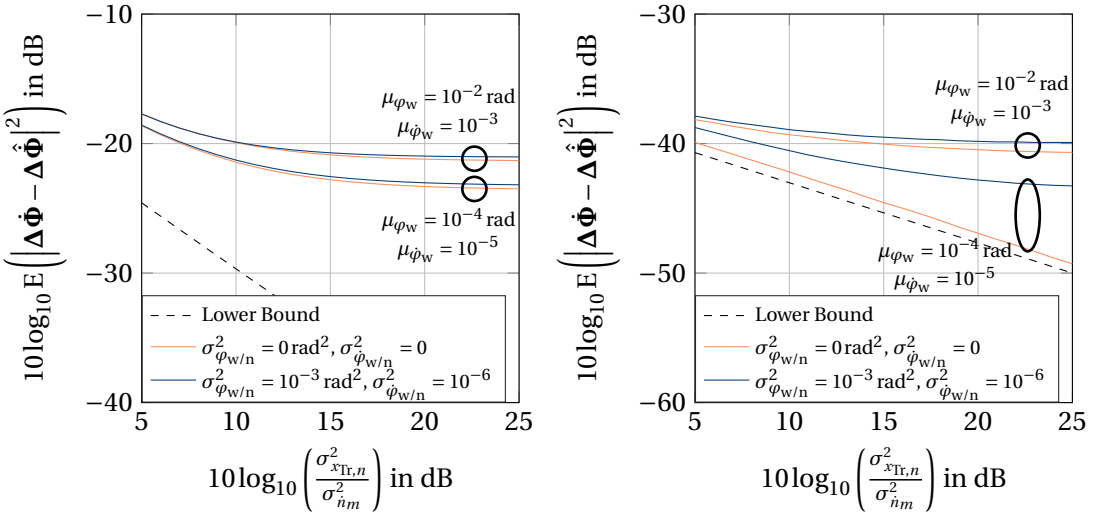


Figure 3.12: MSE of estimating the channels and CFOs, with (3.82) and (3.45), for a LoS MIMO system affected by independent carrier and sampling phase processes, with $\dot{L}_\mathrm{C} = 20$ taps, $M = N = 4$ antennas, Rician factor $K_\mathrm{R} = 10\,\mathrm{dB}$, using Zadoff-Chu training of length $\dot{L}_\mathrm{Tr} = 179$, and using an oversampling factor of $Q = 2$. Includes purely linearly-drifting, as well as more complicated phase processes. Channel estimation error for different normalized CFO values μ_{φ_w} and SFO values $\mu_{\dot{\varphi}_\mathrm{w}}$ (left). Carrier phase difference estimation error for different μ_{φ_w} and $\mu_{\dot{\varphi}_\mathrm{w}}$ (right).

main performance degradation[43] is due to the interference that is generated due to the timing impairments during correlation with the training sequences, as well as the interference from the training sequences of different antennas.

In Figure 3.13, MSE results for the estimation of the SFOs in the same system, but with both $K_\mathrm{R} \to \infty$ and $K_\mathrm{R} = 10\,\mathrm{dB}$, are shown. The observation window size for the sampling phase difference estimation is $\dot{L}_\mathrm{C} = 20$, and the DFT size is chosen as $D = \left\lfloor \frac{\dot{L}_\mathrm{C}}{2Q} \right\rfloor - 1 = 4$. For the case of $K_\mathrm{R} \to \infty$, the results were obtained with (3.76). This approach does not generate useful results for $K_\mathrm{R} = 10\,\mathrm{dB}$, since the squaring does not sufficiently suppress the multi-path contribution. Thus, for this scenario (3.84) is used. The results show that low MSEs can be achieved with both approaches for the respective cases. As for the estimation of the channel and the CFOs, the impact of the timing impairments on the correlation with the training sequence also increases the MSE of the SFO estimation, when higher CFO and SFO values or more complex phase

[43] How such estimation errors affect the complete transmission chain depends on how the estimates are used in the receiver. Such investigations may be found in the literature. A simple approach is to view the estimation errors as an additional noise term, which is scaled by the input samples. As such, it degrades performance markedly when it is in the vicinity of the SNR. For the results from Figure 3.12, this means that the channel estimation errors limit performance for SNRs higher than approximately 14 dB. This value is obtained by assuming that the channel estimation errors are insignificant, when the MSE is 6 dB above the SNR. Although this approach yields some intuition on the impact of estimation errors, it may not show the true influence of these errors, especially for MIMO systems.



(a) $K_R \to \infty$ (b) $K_R = 10\,\mathrm{dB}$

Figure 3.13: MSE of estimating the SFOs, with (3.76) and (3.84), for a LoS MIMO system affected by independent carrier and sampling phase processes, with $\dot{L}_C = 20$ taps, $M = N = 4$ antennas, different Rician factors K_R, using Zadoff-Chu training of length $\dot{L}_{Tr} = 179$, and using an oversampling factor of $Q = 2$. Includes purely linearly-drifting, as well as more complicated phase processes. Sampling phase difference estimation error using (3.76) for different normalized CFO values μ_{φ_w} and SFO values $\mu_{\dot{\varphi}_w}$ (left). Sampling phase difference estimation error using (3.84) for different μ_{φ_w} and $\mu_{\dot{\varphi}_w}$ (right).

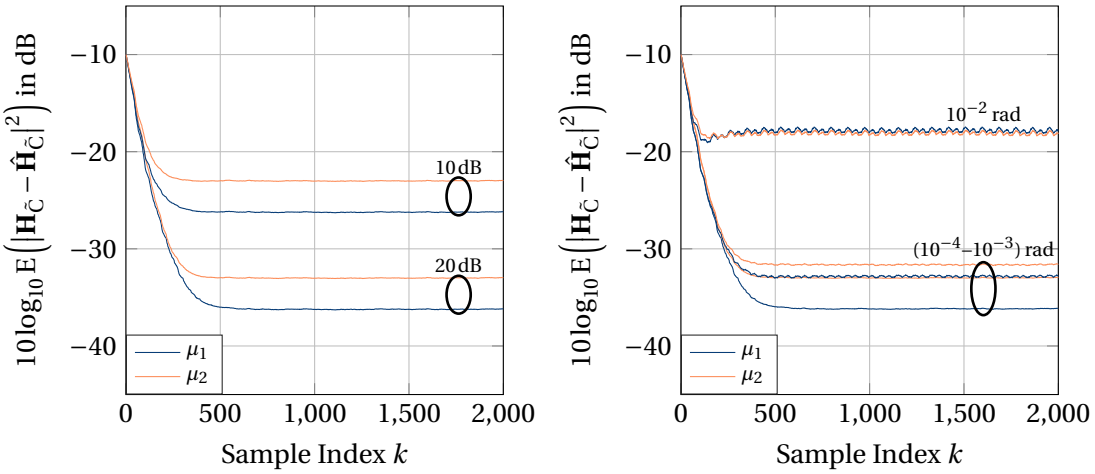
processes are considered. Generally, the method of (3.84), using two consecutive channel estimates for SFO estimation, can achieve lower MSEs and is closer to the bound, as it only estimates the difference in the sampling phase processes between two time instants, and also does not impose a certain channel characteristic.

3.5.3 Adaptive Filter: Convergence and Tracking in the CFO Impaired Case

In this section, simulations investigating the estimation performance of the adaptive filter will be provided. First, the focus is on a symbol-spaced, i.e., $Q = 1$, LoS MIMO system that is influenced only by multiple CFOs and a frequency-selective channel, i.e., $M = N = 4$, $K_R = 10\,\mathrm{dB}$, and $L_C = 10$ as in the previous sections. As training signal, a pseudo-noise sequence of length $L_T = L_C = 10$ per iteration is used. Compared to the previous sections, the timing impairments will not be specifically extracted[44], but only the MSE of the complete time-varying MIMO channel matrix will be computed. The adaptive filters are in the convergence case initialized with a zero matrix.

Figure 3.14 contains the convergence behavior, specifically the ensemble averages over many

[44] Given estimates of the time-varying MIMO matrix over time, we could extract the timing parameters in the same way as in the standard estimation approach above.



(a) $\mu_{\varphi_\mathrm{w}} = 0\,\mathrm{rad}$

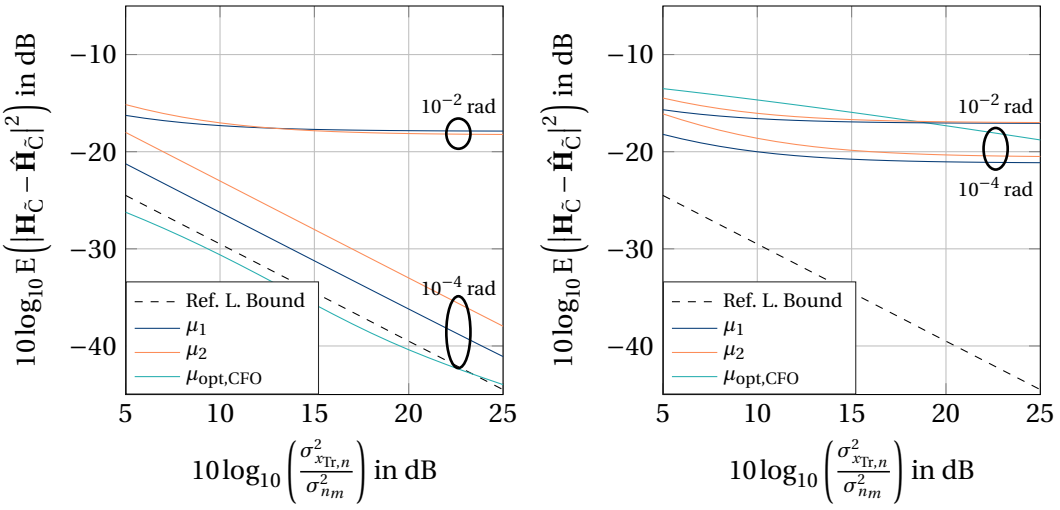
(b) $10\log_{10}\left(\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{n_m}}\right) = 20\,\mathrm{dB}$

Figure 3.14: Convergence MSE of channel estimation with an adaptive filter, with step sizes μ_1 and μ_2 from (3.113) and (3.135), for a LoS MIMO system affected by independent carrier phase processes, with $L_\mathrm{C} = 10$ taps, $M = N = 4$ antennas, a Rician factor of $K_\mathrm{R} = 10\,\mathrm{dB}$, using pseudo-noise training of length $L_\mathrm{T} = L_\mathrm{C} = 10$ per iteration, and using an oversampling factor of $Q = 1$. Behavior without CFOs for two different SNRs (left). Behavior for a SNR of 20 dB with different CFO levels μ_{φ_w} and $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} = 0\,\mathrm{rad}^2$ (right).

learning curves, for the system without and with CFOs for different SNRs, using

$$\mu_1 = \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} \cdot (NL_\mathrm{T}+1)}, \text{ and } \mu_2 = \frac{4}{3} \cdot \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} \cdot NL_\mathrm{T}}.$$

It is seen that the filters converge after roughly 400 iterations, compared to the approach from above where similar MSEs were achieved within a training length of $L_\mathrm{Tr} = 89$, and that the steady-state error for both μ is low, when no CFOs are present. In this case, it is also seen that the lower μ_1 always has a better performance, and that an increase in SNR equivalently improves the MSE. This shows that it is indeed the white noise that is added in the system, which is limiting the performance, and not the steady-state variation due the instantaneous gradient estimate, which was also the underlying assumption for deriving μ_1 in (3.113). Lower steady-state noise can be achieved by selecting μ even smaller than μ_1, which will on the other hand increase convergence time. In the CFO case, the behavior is different. While the slope of convergence is the same, the steady-state error is significantly influenced by the CFOs. For values of $\mu_{\varphi_\mathrm{w}} \geq 10^{-3}\,\mathrm{rad}$, the error term due to the lag of the adaptive filter, while following the variations due to the CFOs, is dominant. We can, furthermore, see slightly oscillatory behavior in the steady-state MSE, which comes from the fact that the changes are faster than the filters ability to track them. Specifically, this leads to time instances, where the cyclical variation of the channel catches up with the adaptive filter, yielding lower MSE, before it sort of overtakes the filter and the MSE increases again. In general, it is seen that the convergence



(a) $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} = 0\,\mathrm{rad}^2$ (b) $\sigma^2_{\varphi_\mathrm{w}} = \sigma^2_{\varphi_\mathrm{n}} = 10^{-3}\,\mathrm{rad}^2$

Figure 3.15: MSE of channel tracking with an adaptive filter, with step sizes μ_1, μ_2, and $\mu_{\mathrm{opt,CFO}}$ from (3.113), (3.135), and (3.134), for a LoS MIMO system affected by independent carrier phase processes, with $L_\mathrm{C} = 10$ taps, $M = N = 4$ antennas, a Rician factor of $K_\mathrm{R} = 10\,\mathrm{dB}$, using pseudo-noise training of length $L_\mathrm{T} = L_\mathrm{C} = 10$ per iteration, and using an oversampling factor of $Q = 1$. MSE with purely linearly-drifting CFO processes for different μ_{φ_w} (left). MSE with more complex CFO processes for different μ_{φ_w} (right).

behavior of the adaptive filter is, in this system, not significantly influenced by the CFOs, but the steady-state MSE is, which will be further investigated next.

The tracking MSE performance of the adaptive filter is given in Figure 3.15, including the lower bound that was also used in section 3.5.1. This bound is just meant to serve as a reference to prior results, as the adaptive filter has infinite memory, whereas the bound was derived for a finite length of $L_\mathrm{Tr} = 89$. Thus, for time-invariant channels and corresponding low step sizes, the performance can be better than the finite-sample reference lower bound. Aside from the choices μ_1 and μ_2 already discussed, the optimal selection for the step size μ in the CFO tracking case, derived in section 3.4.4 and given by

$$\mu_{\mathrm{opt,CFO}} = \left(4 \cdot \frac{\Delta\phi^2_{\max}}{\sigma^4_{x_{\mathrm{Tr},n}} \sigma^2_{\dot{n}_m} \cdot N L_\mathrm{T}}\right)^{\frac{1}{3}},$$

is also used. Note that in order to utilize $\mu_{\mathrm{opt,CFO}}$, knowledge about the SNR, and knowledge about the order of magnitude of the CFOs, i.e., $\Delta\phi_{\max} = 2\mu_{\varphi_\mathrm{w}}$ whereas the true values for the simulations are drawn according to $\mu_{\varphi_\mathrm{w},m/n} \sim \mathcal{U}(-\mu_{\varphi_\mathrm{w}}, \mu_{\varphi_\mathrm{w}})$, is assumed. Furthermore, it is assumed that the filter is in a converged state, either through a single estimate as in section 3.5.1 or through convergence with a suitable μ as above, and only the MSE during tracking is investigated. The first observation from Figure 3.15 is that for purely linearly-drifting carrier phase processes, using $\mu_{\mathrm{opt,CFO}}$ can yield significant performance improvements for

small CFOs. The lack of the curve with $\mu_{\text{opt,CFO}}$ for the higher CFO value is explained by the fact that roughly when $10\log_{10}\left(\frac{\sigma^2_{x_{\text{Tr},n}}}{\sigma^2_{n_m}}\right) > 15\,\text{dB}$ and $\mu_{\varphi_\text{w}} > 10^{-3}$ rad, the approximation $\mu \ll \frac{2}{\sigma^2_{x_{\text{Tr},n}} N L_\text{T}}$ that was used to derive $\mu_{\text{opt,CFO}}$, see appendix B, does not hold anymore and the resulting value for $\mu_{\text{opt,CFO}}$ leads to a divergence during tracking. Additionally, the crossing of the two curves for the high CFO value indicates the point where the lag error of the adaptive filter becomes dominant compared to the noise that is added by the system. The performance for more complex carrier phase processes is markedly different. The MSEs are higher, due to the more erratic phase processes. Importantly, the optimal solution from (3.134) does not give the best performance, even for small μ_{φ_w}. The white noise contribution of the phase process seems to have a significant impact, as the lowest μ, i.e., μ_1, achieves the best results. Generally, the MSE of the adaptive filter during tracking increases with the CFO, as the trade-off between gradient noise and lag error becomes worse.

3.5.4 Adaptive Filter: Tracking in the SFO Impaired Case

In this section, the tracking of the adaptive filter in the multiple SFO impaired case is investigated. The same parameters as above are assumed, with the exception of an oversampling factor of $Q = 2$ and a root-raised-cosine filter with $\beta_\text{T} = 0.25$. As training, an upsampled, but not pulse shaped, pseudo noise sequence of length $\dot{L}_\text{T} = \dot{L}_\text{C} = 20$ per iteration is used, see section 3.3.4. As a reference, the lower bound from section 3.5.2 is included additionally. The filter is assumed to be in a converged state, and the following choices for μ, derived in section 3.4.4 and given by

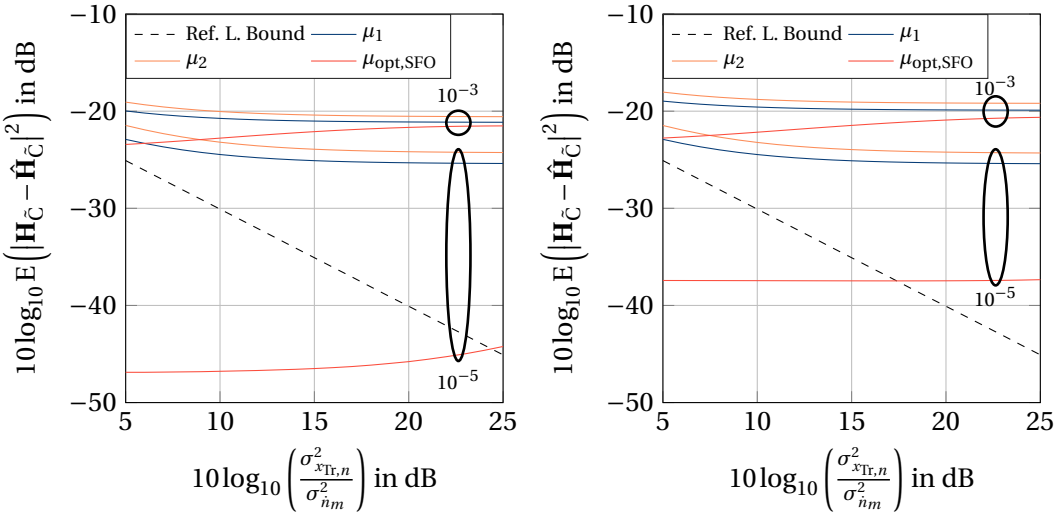
$$\mu_1 = \frac{1}{\sigma^2_{x_{\text{Tr},n}} \cdot (N\dot{L}_\text{T} + 1)},\ \mu_2 = \frac{4}{3} \cdot \frac{1}{\sigma^2_{x_{\text{Tr},n}} \cdot N\dot{L}_\text{T}},\ \text{and}\ \mu_{\text{opt,SFO}} = \sqrt{\frac{1-\alpha^2}{\sigma^2_{x_{\text{Tr},n}} \sigma^2_{\dot{n}_m} \cdot N\dot{L}_\text{T}}},$$

with

$$\alpha = J_0\left(2\pi \cdot \frac{\left|\dot{\phi}_{\text{max,Tx}}\right|}{1 - \left|\dot{\phi}_{\text{max,Tx}}\right|} \cdot \frac{1}{Q} \cdot \left(1 + \left|\dot{\phi}_{\text{max,Rx}}\right|\right)\right),$$

are used. As in the CFO case, $\mu_{\text{opt,SFO}}$ requires knowledge about the SNR and the level of the SFOs. Perfect knowledge about the SNR, and knowledge about the magnitude of the SFOs, i.e., $\left|\dot{\phi}_{\text{max,Tx}}\right| = \left|\dot{\phi}_{\text{max,Tx}}\right| = \mu_{\dot{\varphi}_\text{w}}$, whereas the true values for the simulations are drawn according to $\mu_{\dot{\varphi}_\text{w},m/n} \sim \mathcal{U}(-\mu_{\dot{\varphi}_\text{w}}, \mu_{\dot{\varphi}_\text{w}})$, is assumed.

Figure 3.16 contains the MSE results for the tracking of such LoS MIMO channels impaired by multiple SFOs with an adaptive filter based on the LMS principle. Again, note that the step sizes are very small, meaning that the LMS filter has a long memory, and can, therefore, be better than the finite-length lower bound, which is just given as a reference to the correlation-based results. It is seen that $\mu_{\text{opt,SFO}}$ has the best performance for all parameter settings and yields a significant performance gain, for both the linearly-drifting and more complex sampling phase processes, compared to the other solutions. Generally, all of the achieved



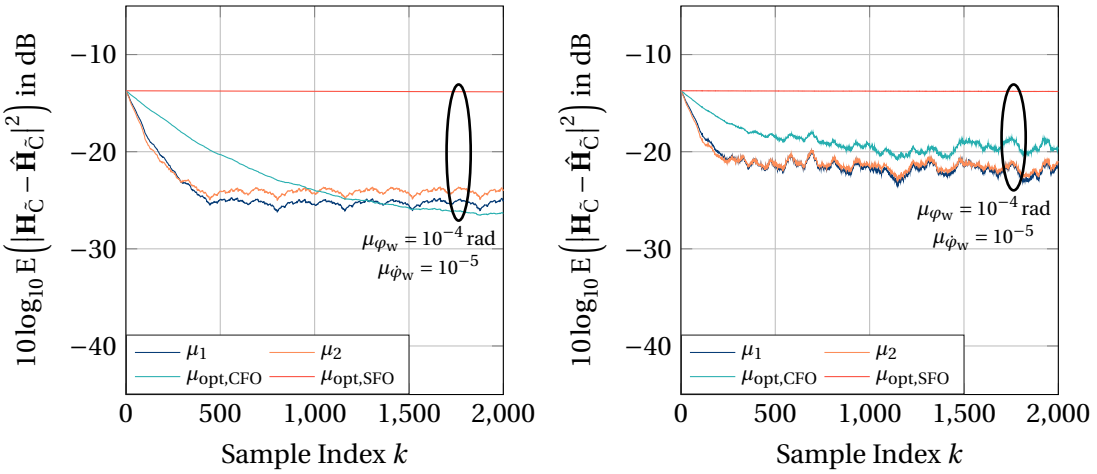
(a) $\sigma^2_{\dot{\varphi}_\text{w}} = \sigma^2_{\dot{\varphi}_\text{n}} = 0$ (b) $\sigma^2_{\dot{\varphi}_\text{w}} = \sigma^2_{\dot{\varphi}_\text{n}} = 10^{-6}$

Figure 3.16: MSE of channel tracking with an adaptive filter, with step sizes μ_1, μ_2, and $\mu_{\text{opt,SFO}}$ from (3.113), (3.135), and (3.161), for a LoS MIMO system affected by independent sampling phase processes, with $\dot{L}_\text{C} = 20$ taps, $M = N = 4$ antennas, a Rician factor of $K_\text{R} = 10\,\text{dB}$, using pseudo-noise training of length $\dot{L}_\text{T} = \dot{L}_\text{C} = 20$ per iteration, and using an oversampling factor of $Q = 2$. MSE with purely linearly-drifting SFO processes for different $\mu_{\dot{\varphi}_\text{w}}$ (left). MSE with more complex SFO processes for different $\mu_{\dot{\varphi}_\text{w}}$ (right).

MSEs are low, with the more complex phase processes increasing the MSE to some extent. The worsening in performance with increasing SNR, which is observed for some cases of $\mu_{\text{opt,SFO}}$, is probably explained by the accuracy of modeling the SFO variations as a Markov process, and the approximations that were used to derive $\mu_{\text{opt,SFO}}$ in section 3.4.4. The trade-off between noisiness of the gradient, and lag error due to tracking of this μ seems to be tilted towards the gradient. This means that for higher SNRs and SFO values, i.e., in regions where the SFO term dominates the behavior of (3.147), the MSE would increase. It is also seen that the MSE for the SFO case is lower than in the CFO case for the considered parameter range. This is in line with the discussion at the end of section 3.4.4, where it was observed that when CFOs and SFOs are of similar magnitude, the CFO contribution dominates the error performance. This will be further investigated in the next section.

3.5.5 Adaptive Filter: Convergence and Tracking in the General Case

Finally, the channel estimation performance of an adaptive LMS filter in a frequency-selective LoS MIMO system including both timing impairments is investigated. The same parameter setting as in the previous sections are used, i.e., $M = N = 4$, $L_\text{C} = 20$, $K_\text{R} = 10\,\text{dB}$, $Q = 2$, root-raised-cosine with $\beta_\text{T} = 0.25$, and pseudo-noise training of length $\dot{L}_\text{T} = \dot{L}_\text{C} = 20$ per iteration. Perfect knowledge about the SNR, and order of magnitude knowledge about the CFOs and SFOs is assumed, when utilizing $\mu_{\text{opt,CFO}}$ and $\mu_{\text{opt,SFO}}$, see also the sections above.



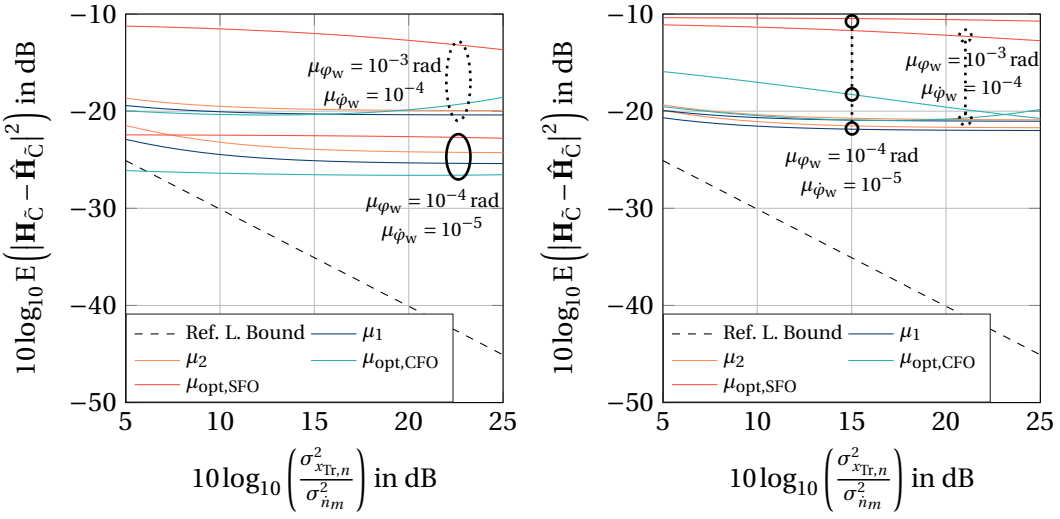
(a) $\sigma^2_{\varphi_w} = \sigma^2_{\varphi_n} = \sigma^2_{\dot{\varphi}_w} = \sigma^2_{\dot{\varphi}_n} = 0$

(b) $\sigma^2_{\varphi_w} = \sigma^2_{\varphi_n} = 10^{-3}\,\mathrm{rad}^2$, $\sigma^2_{\dot{\varphi}_w} = \sigma^2_{\dot{\varphi}_n} = 10^{-6}$

Figure 3.17: Convergence MSE of channel estimation with an adaptive filter, with step sizes $\mu_{\cdot}$ from (3.113), (3.135), (3.134) and (3.161), for a LoS MIMO system affected by independent carrier and sampling phase processes, with $L_C = 20$ taps, an SNR of 20 dB, $M = N = 4$ antennas, a Rician factor of $K_R = 10$ dB, using pseudo-noise training of length $L_T = L_C = 20$ per iteration, and using an oversampling factor of $Q = 2$. Behavior with purely linearly-drifting CFO and SFO processes for different μ_{φ_w} and $\mu_{\dot{\varphi}_w}$ (left). Convergence behavior with more complex CFO and SFO processes for different μ_{φ_w} and $\mu_{\dot{\varphi}_w}$ (right).

The convergence characteristics in terms of MSE are shown in Figure 3.17 for an SNR of 20 dB, where the filters were initialized with a zero matrix. The CFO and SFO mean values are chosen as $\mu_{\varphi_w} = 10^{-4}$ rad and $\mu_{\dot{\varphi}_w} = 10^{-5}$, respectively. It is seen that the filters converge in this case for all selections of μ, except for $\mu_{\mathrm{opt,SFO}}$. This is because, as will also be seen from the next results, the variation of the channel is dominated by the CFOs. The values of $\mu_{\mathrm{opt,SFO}}$ are, thus, generally too low to follow the variations due to the CFOs and ensure convergence. For linearly-drifting phase processes, convergence and the lowest MSE is obtained with $\mu_{\mathrm{opt,CFO}}$. However, for more complex phase processes, and for higher μ_{φ_w}, this does not hold. In particular, μ_1 and μ_2 ensure faster convergence and lower steady-state error for the more complex phase processes, and also ensure convergence for higher CFO values, compared to $\mu_{\mathrm{opt,CFO}}$. Generally, convergence requires a few hundred samples, is not significantly influenced by the complex phase processes, and is very similar to the CFO only case of section 3.5.3.

The MSE during tracking is shown in Figure 3.18, where the filter is assumed in a converged state when the tracking starts. As suspected in the previous section and at the end of section 3.4.4, the error performance is dominated by the CFOs. This is seen by the fact that the results for $\mu_{\mathrm{opt,SFO}}$ are the worst in all cases, since the resulting μ is too low. Using μ_1, μ_2, or $\mu_{\mathrm{opt,CFO}}$ leads to reasonably good performance with the two former leading to the most universal characteristics for the considered parameter range, including the more complex carrier and sampling phase process cases. As in the CFO only case, the adaptive filter diverges and does not track the channel properly with $\mu_{\mathrm{opt,CFO}}$ for values of $\mu_{\varphi_w} > 10^{-3}$ rad, since the



(a) $\sigma^2_{\varphi_w}=\sigma^2_{\varphi_n}=\sigma^2_{\dot{\varphi}_w}=\sigma^2_{\dot{\varphi}_n}=0$ (b) $\sigma^2_{\varphi_w}=\sigma^2_{\varphi_n}=10^{-3}\,\mathrm{rad}^2$, $\sigma^2_{\dot{\varphi}_w}=\sigma^2_{\dot{\varphi}_n}=10^{-6}$

Figure 3.18: MSE of channel tracking with an adaptive filter, with step sizes $\mu_\cdot$ from (3.113), (3.135), (3.134) and (3.161), for a LoS MIMO system affected by independent carrier and sampling phase processes, with $\dot{L}_C=20$ taps, $M=N=4$ antennas, a Rician factor of $K_R=10\,\mathrm{dB}$, using pseudo-noise training of length $\dot{L}_T=\dot{L}_C=20$ per iteration, and using an oversampling factor of $Q=2$. MSE with purely linearly-drifting CFO and SFO processes for different μ_{φ_w} and $\mu_{\dot{\varphi}_w}$ (left). MSE with more complex CFO and SFO processes for different μ_{φ_w} and $\mu_{\dot{\varphi}_w}$ (right).

approximations that were used to derive it do not hold anymore.

The main takeaway from these results is that μ_1 and μ_2 are good choices, when the MIMO channel needs to be initially estimated, i.e., during convergence, and also when a channel with more complex phase processes needs to be tracked. With linearly-drifting phase processes, $\mu_{\mathrm{opt,CFO}}$ may yield superior tracking performance, if the CFO values are not too high. A better MSE performance using $\mu_{\mathrm{opt,SFO}}$ is only obtained, when the SFOs have significantly higher deviations[45] than the CFOs.

3.6 Summary & Main Results

In this chapter, estimation of the channel parameters, based on training signals, for MIMO systems including multiple CFOs and SFOs was investigated. First, based on a MAP/ML approach, it was shown that matched filtering (or correlation) with each of the transmitted training sequences is an ML estimator for the channel impulse responses at time k, when the noise can be considered white, and when the training signals are uncorrelated across time and

[45] In practice, one often uses the same reference in a given transceiver to generate the carrier and the sampling frequency. The variations due to CFOs are then usually larger, since the carrier frequency is typically higher than the sampling frequency, meaning that the variations from the reference are magnified by the higher multiplication factor.

space. Extracting the timing impairments from these time-varying impulse response estimates was proposed to be done under the presumption that they are slow compared to the length of the training sequence. In particular, it was proposed to estimate them by using the difference of two, or more, time-separated channel impulse response estimates. Computationally, the proposed methods are fairly simple, as they only require a phase comparison, or a DFT and a phase comparison. Iterations between the different estimators could, additionally, be used to reduce the influence of one parameter on the others, and generate even better estimates. It was furthermore seen that these approaches can only yield the phase difference processes between transmitter and receiver. However, as will also be seen in the next chapter, it is beneficial to separate the Tx and Rx contribution, using the fact that $M \cdot N$ processes are estimated and only $M+N$ processes do exist. A simple way of doing this was proposed for CFO and SFO estimates, respectively. The simulation results show that the proposed estimation approaches yield low MSEs, especially for medium to low CFOs and SFOs. They also hint at the fact that more suitable training sequences, aside from the Zadoff-Chu sequences that have been used here, can lead to further performance improvement. Generally, the higher the timing impairments, the worse the estimation of a one-shot estimator, as the correlation is influenced by them. When more complex carrier and sampling phase processes are considered, the MSEs are still low, but exhibit a performance floor at high SNR values. This is because the estimators proposed here, and most of the ones available in the literature, are based on the assumption of linearly-drifting phase processes, which does not hold in all system configurations.

To deal with more general time-varying characteristics, this chapter also investigated adaptive filters as a means of estimating and tracking the MIMO channel with timing impairments. The focus is on the LMS algorithm, which is of low complexity and which automatically tracks changes in the channel with some lag. In general, the goal is to assure that the filter converges, and tracks the variations with the lowest MSE possible. For LMS, this depends on the step size or rate of adaptation μ, the timing impairments in the system, the correlation properties of the training signal, and the length of the channel impulse response as well as number of transmit antennas. Different choices for μ based on different optimization criteria, and different available information were derived. For example, when only CFOs are present, the optimal choice for the step size finds the best trade-off between lag error and steady-state noise. It thus increases with the level of the CFOs, as the lag error becomes more dominant. The general case containing both timing impairments is rather complex and only a qualitative discussion was presented. It was seen that for the parameter range considered in this work, the error due to the CFOs dominates, and, hence, its solutions may be used in a case were both CFOs and SFOs are present. A general solution for the step size, which works good for several cases but generates some excess steady-state noise, is given by

$$\mu = \frac{4}{3} \cdot \frac{1}{\sigma^2_{x_{\mathrm{Ir},n}} \cdot N L_{\mathrm{T}}}.$$

Since the optimal solutions were derived based on several assumptions, it is important to check whether the assumptions hold for a particular parameter setting. For example, the solutions

making use of knowledge about the order of magnitude of the CFOs or SFOs, are based on the assumption of linear phase progression, which does not always hold as mentioned previously. In these cases, where the phase processes take on a more complex shape, μ typically needs to be higher than what is predicted by these solutions. Simulation results show that convergence is fast for uncorrelated training signals with the selected step sizes, typically requiring a couple of hundred samples, which is equivalent to hundreds of nanoseconds in the mmWave case. Furthermore, convergence speed is not influenced by the timing impairments for the considered cases. Steady-state error, on the other hand, is influenced by them with higher values, or more complex phase processes, leading to a higher MSE during tracking. The results also reveal that the optimal solutions for the step size μ in the case of timing impairments yield a significant MSE improvement, given that the phase processes are linearly-drifting and relatively slow. Generally, the simulation results show that the μ mentioned above is a viable option in all of the considered cases.

Comparing the two approaches, it can be said that low MSEs can be achieved with both approaches with relatively low complexity and short training sequences of some hundred symbols[46]. The correlation approach can obtain lower MSEs, while the adaptive filtering approach is able to follow more general time-varying channel behavior in a more flexible way. It can also be used in a modified form directly as an equalizer, without the need for a separate estimation unit, as will be seen in the next chapter.

[46] For backhaul-like LoS channels, and linearly-drifting phase processes, one such training is sufficient for estimating the parameters, meaning that the efficiency loss due to training is negligible. For time-varying channels or more complicated phase processes, the training needs to be repeated. How often this is required, and how much it costs in terms of efficiency, depends on the time-varying characteristics of the channels and the phase processes. A treatment of this topic, w.r.t. the coherence time of a channel, can be found in [97].

4 Equalization and Synchronization in MIMO Systems

This part is concerned with describing receiver structures that can compensate the channel and timing impairments that have been described so far, using estimates from the techniques that have been laid out in the previous chapter.

4.1 MAP and ML Data Estimation

It was seen in the previous chapter, specifically section 3.1, that given a good channel estimate $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$, the task of detecting the unknown transmitted symbol sequence $\mathbf{x}_{L_\mathrm{T}}$ based on the received sample sequence $\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}$, can be formulated as follows

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \ \left[p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right] \tag{4.1}$$

$$= \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \ \left[\frac{1}{\det\left(\pi \mathbf{C}_{\tilde{\mathbf{n}}}\right)} \exp\left(- \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right)^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right)\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right]. \tag{4.2}$$

Simplifying by taking the logarithm yields

$$\hat{\mathbf{x}}_{L_\mathrm{T}} = \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \ \left[\log\left(p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right)\right] \tag{4.3}$$

$$= \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\max} \ \left[\log \frac{1}{\det\left(\pi \mathbf{C}_{\tilde{\mathbf{n}}}\right)} - \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right)^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right) + \log p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right] \tag{4.4}$$

$$= \underset{\mathbf{x}_{L_\mathrm{T}}}{\arg\min} \ \left[\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right)^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right) - \log p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right], \tag{4.5}$$

where the first term in (4.4) can be omitted because it does not depend on $\mathbf{x}_{L_\mathrm{T}}$[1], see (3.87). In order to solve (4.5), we need to find the $\mathbf{x}_{L_\mathrm{T}}$ that gives the minimum between the two terms, which can be done for time-invariant and finite-memory systems by using the Viterbi or BCJR algorithm [98, 99]. However, these approaches are already fairly complex for time-invariant frequency-selective SISO systems and time-invariant frequency-flat MIMO systems,

[1] It would also be removed when taking the derivative w.r.t. $\mathbf{x}_{L_\mathrm{T}}$, which is done next.

respectively. This is due to the fact that they scale exponentially with the cardinality of the transmit symbol alphabet and the dimensions of the system, which are in this case the numbers of antennas and the length of the channel memory.

Another way of minimizing the log-likelihood function in (4.5) is to take the partial derivatives w.r.t. the desired data symbols and set them equal to zero[2], i.e.,

$$\mathbf{0} = \frac{\partial}{\partial \mathbf{x}_{L_\mathrm{T}}} \log\left(p\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} | \mathbf{x}_{L_\mathrm{T}}, \mathbf{H}_{\tilde{\mathrm{C}}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}\right) p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right) \tag{4.6}$$

$$= \frac{\partial}{\partial \mathbf{x}_{L_\mathrm{T}}} \left(\left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right)^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}}\right) - \log p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right) \tag{4.7}$$

$$= \frac{\partial}{\partial \mathbf{x}_{L_\mathrm{T}}} \left(\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} - \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}} - \mathbf{x}_{L_\mathrm{T}}^\mathrm{H} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} + \mathbf{x}_{L_\mathrm{T}}^\mathrm{H} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}} - \log p\left(\mathbf{x}_{L_\mathrm{T}}\right)\right) \tag{4.8}$$

$$= -\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} + \mathbf{x}_{L_\mathrm{T}}^\mathrm{H} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} - \frac{\partial}{\partial \mathbf{x}_{L_\mathrm{T}}} \log p\left(\mathbf{x}_{L_\mathrm{T}}\right). \tag{4.9}$$

It is now assumed that the prior distribution of the data symbols $p\left(\mathbf{x}_{L_\mathrm{T}}\right)$ is not known or that they are independent and uniformly distributed (i.u.d.). Thus, the last term can be omitted yielding

$$\mathbf{x}_{L_\mathrm{T}}^\mathrm{H} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} = \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \tag{4.10}$$

$$\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}} \mathbf{x}_{L_\mathrm{T}} = \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}} \tag{4.11}$$

$$\Longrightarrow \hat{\mathbf{x}}_{L_\mathrm{T}}[k] = \left(\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H}[k] \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k] \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]\right)^{-1} \hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^\mathrm{H}[k] \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k] \cdot \dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}[k], \tag{4.12}$$

noting in the final step that the channel estimates $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}$ and the noise correlation $\mathbf{C}_{\tilde{\mathbf{n}}}$ vary in general with time k, and using the fact that $(\mathbf{C}_{\tilde{\mathbf{n}}}^{-1})^\mathrm{H} = (\mathbf{C}_{\tilde{\mathbf{n}}}^\mathrm{H})^{-1} = \mathbf{C}_{\tilde{\mathbf{n}}}^{-1}$ since $\mathbf{C}_{\tilde{\mathbf{n}}}$ is always a Hermitian matrix. It should be noted that for the general case considered so far, (4.12) can only be solved when $M\dot{L}_\mathrm{R} \geq N \cdot L_\mathrm{T}$. From the previous chapter it is known that $\dot{L}_\mathrm{R}$ received samples contain information about $L_\mathrm{T} = (\dot{L}_\mathrm{R} + \dot{L}_\mathrm{C})/Q - 1$ transmitted symbols. Thus, for most cases $M\dot{L}_\mathrm{R} \geq N \cdot L_\mathrm{T}$ is not fulfilled and we need to resort to the algorithms mentioned above, i.e., Viterbi or BCJR, in order to achieve ML performance.

We continue with an equalization approach that presumes[3] $M\dot{L}_\mathrm{R} \geq N \cdot L_\mathrm{T}$, is not ML for all cases, but is of low complexity. It can generally be said that the complete information about the kth transmitted symbol $\mathbf{x}[k]$ is contained in $\dot{L}_\mathrm{C}$ received samples. Thus, those samples are assumed to be sufficient for equalization, and the equalization equation for the

[2] To be certain that the obtained value is a minimum, the second derivative also needs to be checked. It should be visible in (4.9) that at least the first two terms are equal to zero, if the derivative of that function w.r.t. $\mathbf{x}_{L_\mathrm{T}}$ is taken again.

[3] This can be fulfilled when $M > N$.

kth transmitted symbols from all antennas may be written as

$$\hat{\mathbf{x}}[k] = \left(\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^{\mathrm{H}}[k]\mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k]\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k]\right)^{-1}\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}^{\mathrm{H}}[k]\mathbf{C}_{\tilde{\mathbf{n}}}^{-1}[k]\cdot\dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.13}$$

where $\hat{\mathbf{H}}_{\tilde{\mathrm{C}}}[k] = \left[\hat{\mathbf{H}}_{\mathrm{C}}^{\mathrm{T}}[k,0] \quad \hat{\mathbf{H}}_{\mathrm{C}}^{\mathrm{T}}[k,1] \quad \cdots \quad \hat{\mathbf{H}}_{\mathrm{C}}^{\mathrm{T}}[k,\dot{L}_{\mathrm{C}}-1]\right]^{\mathrm{T}}$, $\mathbf{C}_{\tilde{\mathbf{n}}}[k] = \mathrm{E}\left[\tilde{\mathbf{n}}_{\dot{L}_{\mathrm{C}}}[k]\tilde{\mathbf{n}}_{\dot{L}_{\mathrm{C}}}^{\mathrm{H}}[k]\right]$, and which is equivalent to (4.12) if $L_{\mathrm{C}} = 1$. Note that this equalization equation is not necessarily optimal in the ML sense, rather it is similar to the LS solution for the channel, and its performance depends strongly on the structure of the channel.

4.1.1 Separation of Timing Impairments

Based on (4.13), the goal is now to try to separate the different matrices contained in $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ and $\mathbf{C}_{\tilde{\mathbf{n}}}[k]$, in order to simplify the equalization. Consider the general notation with

$$\mathbf{H}_{\tilde{\mathrm{C}}}[k] = \mathbf{H}_{\mathrm{Rx}}[k]\cdot\mathbf{\Phi}_{\mathrm{Rx}}[k]\cdot\mathbf{H}[k]\cdot\mathbf{\Phi}_{\mathrm{Tx}}[k]\cdot\mathbf{H}_{\mathrm{Tx}}[k] \tag{4.14}$$

$$\mathbf{C}_{\tilde{\mathbf{n}}}[k] = \mathbf{H}_{\mathrm{Rx}}[k]\cdot\mathbf{N}\cdot\mathbf{H}_{\mathrm{Rx}}^{\mathrm{H}}[k], \tag{4.15}$$

omitting the lengths and oversampling notation for clarity, compare to section 2.5.7. Recall that the total length of the impulse response from the components is $L_{\mathrm{C}} = L_{\mathrm{Rx}} + L + L_{\mathrm{Tx}} - 2$. Suppose further that the transmit and receive filter influences are invertible with finite-length filters, i.e., the inverses $\mathbf{H}_{\mathrm{Tx}}^{-1}[k]$ and $\mathbf{H}_{\mathrm{Rx}}^{-1}[k]$ exist[4]. Then, substituting the above into (4.13) gives

$$\begin{aligned}\hat{\mathbf{x}}[k] = &\left(\mathbf{H}_{\mathrm{Tx}}^{\mathrm{H}}[k]\mathbf{\Phi}_{\mathrm{Tx}}^{\mathrm{H}}[k]\mathbf{H}^{\mathrm{H}}[k]\cdot\mathbf{N}^{-1}\cdot\mathbf{H}[k]\mathbf{\Phi}_{\mathrm{Tx}}[k]\mathbf{H}_{\mathrm{Tx}}[k]\right)^{-1}\\ &\cdot\mathbf{H}_{\mathrm{Tx}}^{\mathrm{H}}[k]\mathbf{\Phi}_{\mathrm{Tx}}^{\mathrm{H}}[k]\mathbf{H}^{\mathrm{H}}[k]\mathbf{\Phi}_{\mathrm{Rx}}^{\mathrm{H}}[k]\mathbf{N}^{-1}\mathbf{H}_{\mathrm{Rx}}^{-1}[k]\cdot\dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k]\end{aligned} \tag{4.16}$$

$$\approx \mathbf{H}_{\mathrm{Tx}}^{-1}[k]\cdot\mathbf{\Phi}_{\mathrm{Tx}}^{-1}[k]\cdot\underbrace{\left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{H}}}_{\substack{\text{Standard Zero-Forcing}\\ \text{MIMO Equalizer}}}\cdot\mathbf{\Phi}_{\mathrm{Rx}}^{\mathrm{H}}[k]\cdot\mathbf{H}_{\mathrm{Rx}}^{-1}[k]\cdot\dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.17}$$

using the fact that $\mathbf{\Phi}_{\cdot}[k]$ is a diagonal matrix of complex exponentials with the properties $\mathbf{\Phi}_{\cdot}^{-1}[k] = \mathbf{\Phi}_{\cdot}^{\mathrm{H}}[k] = \mathbf{\Phi}_{\cdot}^{*}[k]$ and $\mathbf{\Phi}_{\cdot}[k]\cdot\mathbf{\Phi}_{\cdot}^{\mathrm{H}}[k] = \mathbf{I}$. The second approximation can be made if the noise power is equal in all receivers and if the channel is time invariant, i.e., negligible movement during transmission, and if the variation due to receiver SFOs, refer to section 2.5.3, is sufficiently compensated by the first applied filter $\mathbf{H}_{\mathrm{Rx}}^{-1}[k]$.

From (4.17), a couple observations can be made. First, it is possible to separate the timing impairments from the standard MIMO equalization, i.e., the middle part of (4.17), by compensating the effects consecutively in the reverse order that they occur[5]. Furthermore, in order

[4] These filters inherently contain interpolation structures, see [4, 25], which are necessary to compensate the ISI due to the SFOs.

[5] In practice, refer to section 2.4, the phase change due to CFOs and SFOs is often very small from sample to sample, and can be assumed approximately constant over many samples. In such cases, it is possible to exchange SFO and CFO compensation before/after the standard MIMO equalizer.

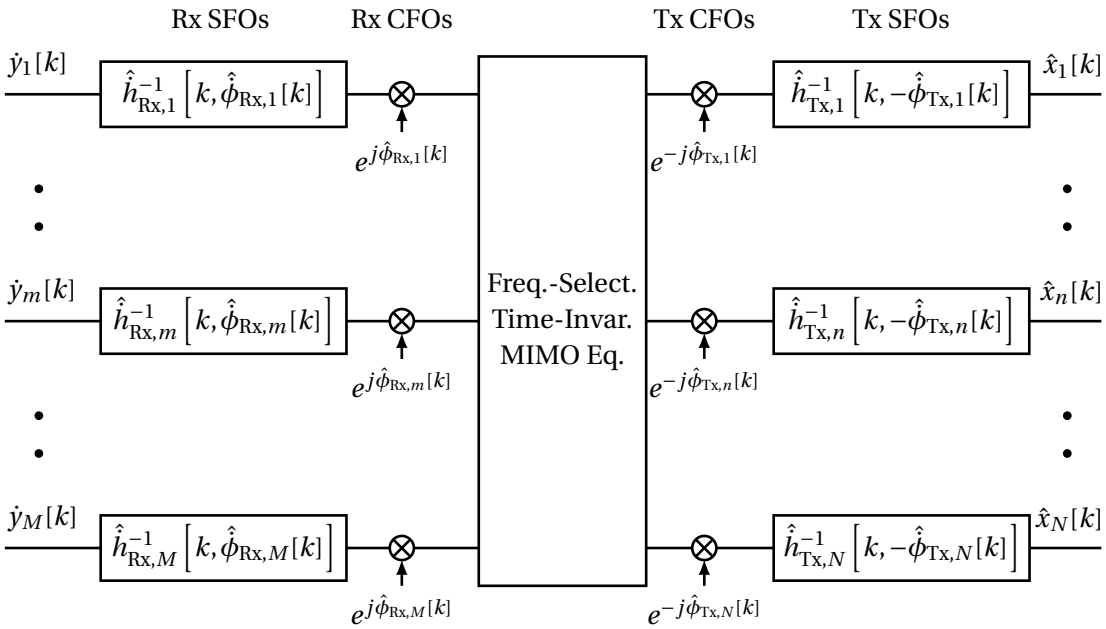


Figure 4.1: Basic equalization setup, compensating multiple timing impairments and a time-invariant frequency-selective MIMO channel.

to compensate the SFOs properly, the complete equalizer needs to work on an oversampled signal and down sampling to symbol rate is only possible after the final stage of the equalizer has compensated the Tx SFOs. Finally, in terms of complexity, it can be seen that all timing impairments can be compensated in parallel, refer to the equalizer diagram in Figure 4.1, meaning that it is not required to update the full equalization matrix at every time step k, as suggested in (4.13). Instead, it is sufficient to update the estimates and corresponding filters that are associated with the timing impairments. The structure can be further simplified, when joint processes are considered at the transmitter or receiver side, as will be seen in the next sections.

4.1.2 A More General Equalizer Structure

Before dealing with more specific instances of the equalizer, a more general equalization structure will be introduced that can improve the equalization performance. In the previous sections, suboptimal receivers that rely on $\dot{L}_{\mathrm{F}}$ received samples to estimate the transmitted symbols $\mathbf{x}[k]$ were discussed. It is known from the literature that the estimate of the symbols at time k can be improved by including knowledge about already processed symbols. Accordingly,

an equalizer can be written more generally [75, 100, 101] as

$$\hat{\mathbf{x}}[k] = \mathbf{W}_{\mathrm{F}}[k] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{F}}}[k] - \mathbf{W}_{\mathrm{B}}[k] \cdot \check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1] \tag{4.18}$$

$$= \underbrace{\begin{bmatrix} \mathbf{W}_{\mathrm{F}}[k] & \mathbf{W}_{\mathrm{B}}[k] \end{bmatrix}}_{\mathbf{W}_{\tilde{\mathrm{C}}}[k]} \begin{bmatrix} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{F}}}[k] \\ -\check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1] \end{bmatrix}, \tag{4.19}$$

where $\mathbf{W}_{\mathrm{F}}[k]$ and $\mathbf{W}_{\mathrm{B}}[k]$ are time-varying feedforward and feedback filters of length $\dot{L}_{\mathrm{F}}$ and L_{B}, respectively. The vector $\check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1]$ contains the L_{B} previous symbol decisions based on the previously estimated symbol vectors, where $\check{\mathbf{x}}[k] = \mathrm{dec}(\hat{\mathbf{x}}[k])$. These structures are known as DFEs. Compared to (4.12), it is seen that the feedforward part should be equivalent to the first N rows of that equalization matrix given that $\dot{L}_{\mathrm{F}} = \dot{L}_{\mathrm{R}}$. Selecting the lengths $\dot{L}_{\mathrm{F}}$ and L_{B} in the best way depends on the exact structure of the channel, and how reliable the decisions on the transmitted symbols are. An example, where such an equalizer can achieve almost optimal performance, is a strong LoS MIMO channel without timing impairments, by using $L_{\mathrm{F}} = 1$, corresponding to the inverse of the LoS tap, and $L_{\mathrm{B}} = L_{\mathrm{C}} - 1$, corresponding to subtraction of the ISI due to previously transmitted symbols, see [102].

4.2 Equalizers for Different System Configurations

Equation (4.19) gives a general equalization equation for any time-varying MIMO system in colored noise, given i.u.d. data symbols. Computing this equation for every set of samples seems daunting, especially due to the fact that the equalization matrix $\mathbf{W}_{\tilde{\mathrm{C}}}[k]$ needs to be updated for every time step k. However, it was seen in the previous sections that for $L_{\mathrm{B}} = 0$ and equal noise variance in all receivers, the timing impairment compensation can be separated from the standard MIMO equalization. In the following, simplified equalizers for different MIMO configurations, experiencing the different timing impairments, will be presented. It was previously mentioned that in the presence of timing impairments, symbol-spaced signal processing can only achieve suboptimal results. Nevertheless, this type of signal model will be used to begin with in order to have a reference to results from the literature.

4.2.1 Synchronized Symbol-Spaced Equalizer

Consider the case where the timing impairments are negligible, e.g., they have been adequately compensated prior to equalization, and no oversampling is present ($Q = 1$), e.g., the signal has been decimated to symbol rate prior to equalization. This case is well covered in the literature [27, 28, 101, 102, 103, 104, 105]. Only the feedforward type ($L_{\mathrm{B}} = 0$) from equation (4.13) will be mentioned, given as

$$\hat{\mathbf{x}}[k] = \left(\hat{\mathbf{H}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \hat{\mathbf{H}}_{\mathrm{C}}\right)^{-1} \hat{\mathbf{H}}_{\mathrm{C}}^{\mathrm{H}} \mathbf{C}_{\tilde{\mathbf{n}}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.20}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}} \hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.21}$$

which is time invariant, and where the middle part of the second equation could be replaced by any MIMO equalizer as before. The filter contained in $\hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}$ and $\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}$ can be applied to each stream in parallel, see section 4.1.1, and the latter one is also commonly known as a whitening filter.

4.2.2 Symbol-Spaced Equalizer Containing Only CFOs

Assume the case where only CFOs exist and no oversampling is used. The general case with multiple CFOs has been considered to some extent in the literature [67, 68, 106]. From (4.17), we get in the general case

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Tx}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.22}$$

where, as above, the central part can be replaced by any MIMO equalizer. For the DFE of (4.19), we can accordingly write

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Tx}}^{-1}[k] \cdot \left(\mathbf{W}_{\mathrm{F}} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] - \mathbf{W}_{\mathrm{B}} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Tx}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Tx}} \cdot \check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1]\right), \tag{4.23}$$

where the feedforward and feedback filters $\mathbf{W}_{\mathrm{F}}$ and $\mathbf{W}_{\mathrm{B}}$ are time invariant.

So far, the most general case where independent oscillators are used for the generation of the carrier frequencies at Tx and Rx, i.e., oscillator setup 4, has been considered. In the following, equalizers for other oscillator configurations that can occur in a MIMO system, see section 2.4.1, will be presented. The focus will be on the feedforward type, the extension to DFE follows from the above.

Independent on Tx and Shared on Rx

Let us start with the case where the carrier phase process is the same for all receiving front ends, i.e., setup 3. Equation (4.22) simplifies to

$$\begin{aligned}
\hat{\mathbf{x}}[k] &\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Tx}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot e^{j\hat{\phi}_{\mathrm{Rx}}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] && (4.24)\\
&= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \mathbf{\Delta}\hat{\mathbf{\Phi}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] && (4.25)\\
&\approx \mathbf{\Delta}\hat{\mathbf{\Phi}}^{-1}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], && (4.26)
\end{aligned}$$

with $\mathbf{\Delta}\hat{\mathbf{\Phi}}^{-1}[k] = \hat{\mathbf{\Phi}}_{\mathrm{Tx}}^{-1}[k]e^{j\hat{\phi}_{\mathrm{Rx}}[k]} = \mathrm{diag}(e^{j(\hat{\phi}_{\mathrm{Rx}}[k]-\hat{\phi}_{\mathrm{Tx},1}[k])}, e^{j(\hat{\phi}_{\mathrm{Rx}}[k]-\hat{\phi}_{\mathrm{Tx},2}[k])}, \ldots, e^{j(\hat{\phi}_{\mathrm{Rx}}[k]-\hat{\phi}_{\mathrm{Tx},N}[k])})$ being a diagonal matrix that contains the compensation for the difference between the multiple transmit, and the one receive carrier phase processes. The approximations in the preceding equations can be made, given that the phase processes are approximately constant over the

length of the combined impulse response L_C[6].

Equation (4.26) shows that for a shared oscillator setup on the receiver side, we can employ under the given assumptions a standard time-invariant MIMO equalizer and compensate the effect due to the CFOs in parallel for each recovered transmitted stream after that equalizer. A straightforward way of doing this is to use decision-directed (DD) phase tracking in form of a digital PLL per recovered stream[7], which was seen to perform optimally for this type of setup in [55].

Shared on Tx and Independent on Rx

Consider the case where all transmitting front ends share the same carrier phase process, i.e., setup 3. It follows from (4.22) that

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot e^{-j\hat{\phi}_{\mathrm{Tx}}[k]} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.27}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.28}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{H}}[k] \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.29}$$

with $\mathbf{\Delta}\hat{\mathbf{\Phi}}^{\mathrm{H}}[k] = e^{-j\hat{\phi}_{\mathrm{Tx}}[k]}\hat{\mathbf{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k] = \mathbf{I}_{L_{\mathrm{F}}} \otimes \mathrm{diag}(e^{j(\hat{\phi}_{\mathrm{Rx},1}[k]-\hat{\phi}_{\mathrm{Tx}}[k])}, \ldots, e^{j(\hat{\phi}_{\mathrm{Rx},M}[k]-\hat{\phi}_{\mathrm{Tx}}[k])})$ being a diagonal matrix containing the difference between one transmit, and multiple receive carrier phase processes. The same reasoning for the approximations as in the previous section applies.

From (4.29), it is seen that in order to compensate the CFOs in systems with a shared oscillator on the transmitter side, synchronization prior to MIMO equalization is required. Due to the fact that the CFOs appear in this setup in parallel for each of the received streams we can, aside from the techniques mentioned in chapter 3, also use blind techniques [3, 4, 5, 28] per received stream in order to estimate the phase processes. One example is to use the well-known squarer and digital PLL, e.g, described in [107], on each receive stream separately.

Shared on Tx and Shared on Rx

Finally, let us look at the case where one carrier phase process is shared among all Txs, and where also one carrier phase process is shared among all Rxs, i.e., setup 2. The equalization

[6] It should be visible that the phase process matrices $\hat{\mathbf{\Phi}}_{\cdot}^{-1}[k]$ should be of larger dimensions than M or N due to the frequency selectivity of the system. However, given the small changes of the phase processes from sample to sample in practice, see section 2.4, the approximation is viable. Another way of looking at this physically is that the CFOs should be low enough such that their shifting of the spectrum does not change the impulse response of the filters significantly.

[7] It should be noted that standard SISO synchronization techniques, also the ones mentioned in the upcoming sections, do seldom account for in-band interference that may be present in the received signal. In general, the received MIMO signal contains several superimposed signals, i.e., in-band interference, prior to MIMO equalization, and residual in-band interference, depending on the channel and equalizer, after MIMO equalization. Nevertheless, this in-band interference can be approximately viewed as a higher modulation order signal, and noise, respectively, meaning that most techniques are still usable to some extent.

simplifies to

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot e^{-j\hat{\phi}_{\mathrm{Tx}}[k]} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot e^{j\hat{\phi}_{\mathrm{Rx}}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.30}$$

$$= e^{j\Delta\hat{\phi}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.31}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot e^{j\Delta\hat{\phi}[k]} \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.32}$$

where $e^{j\Delta\hat{\phi}[k]} = e^{j(\hat{\phi}_{\mathrm{Rx}}[k]-\hat{\phi}_{\mathrm{Tx}}[k])}$, and using the same approximations as in the previous two sections. Thus, it is seen that in the shared oscillator setup only a single phase process needs to be compensated before or after standard MIMO equalization. Both techniques mentioned in the previous two sections can be used in order to achieve this goal.

4.2.3 Symbol-Spaced Equalizer Containing Only SFOs

Assume the case where only SFOs exist and no oversampling is used. The general case with multiple SFOs has not been explicitly investigated in the literature[8]. From (4.17) we get in the general case

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}[k] \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.33}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right] \cdot \mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.34}$$

where, as mentioned previously, the central part can be replaced by any MIMO equalizer. The vectors $\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]$ and $\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]$ contain the estimated sampling phase processes of the N Tx and M Rx, refer also to section 2.5.3. The notation $\hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]$ and $\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]$, which will also be used in the next sections, needs some further explanation. These two filters are still of finite length, i.e., discrete, but they vary with time according to the sampling phase estimates $\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]$ and $\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]$. Hence, they correspond to the discrete-time versions of the filters, interpolated to the mentioned sampling phases. For the DFE of (4.19) we get

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right] \cdot \left(\mathbf{W}_{\mathrm{F}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right] \cdot \mathbf{y}_{L_{\mathrm{F}}}[k] - \mathbf{W}_{\mathrm{B}} \cdot \hat{\mathbf{H}}_{\mathrm{Tx}}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right] \cdot \check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1]\right), \tag{4.35}$$

where the feedforward and feedback filters $\mathbf{W}_{\mathrm{F}}$ and $\mathbf{W}_{\mathrm{B}}$ are, as in the CFO case, time invariant. We will continue with the different oscillator setups for sampling frequency generation, similar to the CFO section. Finally, as mentioned before, it is important to note that the effect of the SFOs can only be reduced, and not be fully compensated, in a symbol-spaced framework, as irrecoverable aliasing is generated in general in such a case.

[8] We could treat it as part of a general time-varying channel as in [75] and as will be done later on in this chapter. Furthermore, [74] treats a variant of the problem.

Independent on Tx and Shared on Rx

Consider the case, where the sampling frequencies of all receiving front ends is derived from the same phase process, i.e., setup 3. From (4.34), it follows that

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\phi}}_{\mathrm{Rx}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.36}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\left(\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\star\mathrm{sinc}\left[\pi\cdot\left(l-L_{\mathrm{Rx}}/2+\hat{\hat{\phi}}_{\mathrm{Rx}}[k]\right)\right]\right)\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.37}$$

$$\approx \left(\hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\star\mathrm{sinc}\left[\pi\cdot\left(l-L_{\mathrm{Rx}}/2+\hat{\hat{\phi}}_{\mathrm{Rx}}[k]\right)\right]\right)\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.38}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[\boldsymbol{\Delta}\hat{\hat{\boldsymbol{\phi}}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\cdot\mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.39}$$

with some abuse of notation, assuming that the channel is approximately constant over the sampling frequency change of the receivers, confer also section 2.5.3, and with the combined phase process $\boldsymbol{\Delta}\hat{\hat{\boldsymbol{\phi}}}[k] = \left[\hat{\hat{\phi}}_{\mathrm{Rx}}[k]-\hat{\hat{\phi}}_{\mathrm{Tx},1}[k] \quad \hat{\hat{\phi}}_{\mathrm{Rx}}[k]-\hat{\hat{\phi}}_{\mathrm{Tx},2}[k] \quad \cdots \quad \hat{\hat{\phi}}_{\mathrm{Rx}}[k]-\hat{\hat{\phi}}_{\mathrm{Tx},N}[k]\right]^{\mathrm{T}}$.

From (4.39), it is seen that the time-varying ISI due to the SFOs can, in such a setup under the given assumptions, be compensated by parallel time-varying filters after standard MIMO equalization. Thus, it is feasible to use standard interpolation filters [4, 25] in parallel, to compensate the combined SFO effect for each recovered transmitted stream. Additionally, given an oversampled signal, we can, aside from the methods mentioned in chapter 3, use standard SFO estimation techniques for SISO systems from the literature [3, 4, 5, 28, 71].

Shared on Tx and Independent on Rx

Assume setup 3, where all transmitters experience the same sampling phase process. Equation (4.34) yields

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\hat{\phi}}_{\mathrm{Tx}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.40}$$

$$= \left(\hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\star\mathrm{sinc}\left[\pi\cdot\left(l-L_{\mathrm{Tx}}/2-\hat{\hat{\phi}}_{\mathrm{Tx}}[k]\right)\right]\right)\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.41}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\left(\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\hat{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]\star\mathrm{sinc}\left[\pi\cdot\left(l-L_{\mathrm{Tx}}/2-\hat{\hat{\phi}}_{\mathrm{Tx}}[k]\right)\right]\right)\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.42}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\boldsymbol{\Delta}\hat{\hat{\boldsymbol{\phi}}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.43}$$

with $\boldsymbol{\Delta}\hat{\hat{\boldsymbol{\phi}}}[k] = \left[\hat{\hat{\phi}}_{\mathrm{Rx},1}[k]-\hat{\hat{\phi}}_{\mathrm{Tx}}[k] \quad \hat{\hat{\phi}}_{\mathrm{Rx},2}[k]-\hat{\hat{\phi}}_{\mathrm{Tx}}[k] \quad \cdots \quad \hat{\hat{\phi}}_{\mathrm{Rx},M}[k]-\hat{\hat{\phi}}_{\mathrm{Tx}}[k]\right]^{\mathrm{T}}$.

Compared to the case above, (4.43) shows that the SFOs can be compensated by parallel time-varying filters prior to standard MIMO equalization in this oscillator setup. Furthermore, in the oversampled case, we can use well-known blind SFO estimation techniques [3, 4, 5, 28, 107], e.g., squaring of the signal, in parallel for each received stream.

Shared on Tx and Shared on Rx

Finally, consider setup 2, where all transmitters experience the same sampling phase process, and all receivers experience the same sampling phase process, respectively. From (4.34)

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\dot{\phi}}_{\mathrm{Tx}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\dot{\phi}}_{\mathrm{Rx}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.44}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\Delta\hat{\dot{\phi}}[k]\right]\cdot\mathbf{y}_{L_{\mathrm{F}}}[k] \tag{4.45}$$

$$= \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[\Delta\hat{\dot{\phi}}[k]\right]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\cdot\mathbf{y}_{L_{\mathrm{F}}}[k], \tag{4.46}$$

where the combined phase process is $\Delta\hat{\dot{\phi}}[k] = \hat{\dot{\phi}}_{\mathrm{Rx}}[k] - \hat{\dot{\phi}}_{\mathrm{Tx}}[k]$, and using the same assumptions as above. Thus, all of the techniques mentioned above are feasible.

4.2.4 Oversampled Equalizer Containing CFOs and SFOs

Consider the case where both timing impairments, i.e., CFO and SFO, are present, and where the received signal is oversampled Q times. Only the cases, where the oscillator setup is the same for carrier and sampling frequency generation, because they are synonymous with independent transmitters or receivers, e.g., multiple users in a spatial division access scheme where sharing of a reference is cumbersome, will be presented here. Likewise, if a reference is already shared for carrier or sampling frequency generation, it may also be used for generation of the other one, confer the last part of section 2.4. Nevertheless, equalizers for mixed setups can be obtained in a similar way as the results from the previous and following sections.

Equalization for the general case with multiple CFOs and SFOs has not been explicitly treated in the literature[9]. From (4.17), we have

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\cdot\hat{\boldsymbol{\Phi}}_{\mathrm{Tx}}^{-1}[k]\cdot\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}}\cdot\hat{\boldsymbol{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k]\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]\cdot\dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.47}$$

again omitting the oversampled notation on the matrices. In comparison to the symbol-spaced case, the central part can be replaced by any fractionally-spaced MIMO equalizer. For the more general DFE of (4.19), we can write accordingly

$$\begin{aligned}\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\cdot\hat{\boldsymbol{\Phi}}_{\mathrm{Tx}}^{-1}[k]\cdot\Big(&\mathbf{W}_{\mathrm{F}}\cdot\hat{\boldsymbol{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k]\cdot\hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right]\cdot\dot{\mathbf{y}}_{\dot{L}_{\mathrm{F}}}[k]\\ &-\mathbf{W}_{\mathrm{B}}\cdot\hat{\boldsymbol{\Phi}}_{\mathrm{Tx}}[k]\cdot\hat{\mathbf{H}}_{\mathrm{Tx}}\left[-\hat{\dot{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right]\cdot\check{\mathbf{x}}_{L_{\mathrm{B}}}[k-1]\Big),\end{aligned} \tag{4.48}$$

where the feedforward and feedback filters $\mathbf{W}_{\mathrm{F}}$ and $\mathbf{W}_{\mathrm{B}}$ are time invariant, as in the symbol-spaced cases, but in contrast to these cases the filters now need to be fractionally spaced. From these equalization equations, it is seen that the complete equalization structure needs to be fractionally spaced in order to properly compensate the SFOs. In particular, the Tx SFOs create signals with slightly different bandwidths, meaning that whenever the received signal is

[9] With the exception of [74, 75], which treat a specific version of the problem.

downsampled to symbol rate prior to separation with the MIMO equalizer, it is very likely that some of the superimposed signals generate irreversible aliasing. In the following, different oscillator setups, based on (4.47), will be looked at, and the cases where part of the equalizer can be symbol spaced will be discussed.

Independent on Tx and Shared on Rx

Consider setup 3, where a phase process is shared for the generation of the carrier and sampling frequencies, respectively, on the receiver side. From (4.47) we get

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\tilde{\boldsymbol{\phi}}}_{\mathrm{Tx}}[k]\right] \cdot \hat{\boldsymbol{\Phi}}_{\mathrm{Tx}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot e^{j\hat{\phi}_{\mathrm{Rx}}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\tilde{\phi}}_{\mathrm{Rx}}[k]\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k] \tag{4.49}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[\Delta\hat{\tilde{\boldsymbol{\phi}}}[k]\right] \cdot \Delta\hat{\boldsymbol{\Phi}}^{-1}[k] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k] \tag{4.50}$$

$$\approx \Delta\hat{\boldsymbol{\Phi}}^{-1}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[\Delta\hat{\tilde{\boldsymbol{\phi}}}[k]\right] \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.51}$$

with $\Delta\hat{\boldsymbol{\Phi}}^{-1}[k]$ and $\Delta\hat{\tilde{\boldsymbol{\phi}}}[k]$ being defined as in the previous sections investigating this oscillator setup, and using the same assumptions as in those sections. It is seen that any fractionally-spaced time-invariant MIMO equalizer should be used first in order to recover the different transmitted streams. After separating them, the timing impairments can be compensated in parallel for each stream, similar to independent SISO systems, with the techniques mentioned previously. Downsampling to symbol rate is possible prior to CFO compensation.

Shared on Tx and Independent on Rx

Assume the case, where a phase process is shared for the generation of the carrier and sampling frequencies, respectively, on the transmitter side, i.e., setup 3. Starting from (4.47) yields

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\tilde{\phi}}_{\mathrm{Tx}}[k]\right] \cdot e^{-j\hat{\phi}_{\mathrm{Tx}}[k]} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\boldsymbol{\Phi}}_{\mathrm{Rx}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\tilde{\boldsymbol{\phi}}}_{\mathrm{Rx}}[k]\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k] \tag{4.52}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot \Delta\hat{\boldsymbol{\Phi}}^{\mathrm{H}}[k] \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\Delta\hat{\tilde{\boldsymbol{\phi}}}[k]\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.53}$$

where $\Delta\hat{\boldsymbol{\Phi}}^{\mathrm{H}}[k]$ and $\Delta\hat{\tilde{\boldsymbol{\phi}}}[k]$ are defined as in the previous sections investigating this oscillator setup, and employing the same assumptions as in those sections. For this setup, the timing impairments can thus be compensated prior to MIMO equalization in parallel for each of the received streams. Since the SFOs are compensated by this first set of parallel filters, the signal can be downsampled to symbol rate and the rest of the equalization can be carried out on a symbol-spaced level without loss of performance[10].

[10] This assumes perfect knowledge about the sampling phase processes. When the sampling phase processes are estimated, there will be estimation errors leading to irreversible aliasing if the rest of the processing is carried out on a symbol-spaced level.

Shared on Tx and shared on Rx

Finally, consider the case where a phase process is shared for the generation of the carrier and sampling frequencies on the transmitter and receiver side, respectively, i.e., setup 2. Equation (4.47) yields

$$\hat{\mathbf{x}}[k] \approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[-\hat{\tilde{\phi}}_{\mathrm{Tx}}[k]\right] \cdot e^{-j\hat{\phi}_{\mathrm{Tx}}[k]} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot e^{j\hat{\phi}_{\mathrm{Rx}}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\hat{\tilde{\phi}}_{\mathrm{Rx}}[k]\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k] \tag{4.54}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1}\left[\Delta\hat{\tilde{\phi}}[k]\right] \cdot e^{j\Delta\hat{\phi}[k]} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1} \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k] \tag{4.55}$$

$$\approx \hat{\mathbf{H}}_{\mathrm{Tx}}^{-1} \cdot \left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^{\mathrm{H}} \cdot e^{j\Delta\hat{\phi}[k]} \cdot \hat{\mathbf{H}}_{\mathrm{Rx}}^{-1}\left[\Delta\hat{\tilde{\phi}}[k]\right] \cdot \dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k], \tag{4.56}$$

with $\Delta\hat{\phi}[k]$ and $\Delta\hat{\tilde{\phi}}[k]$ being defined as in the previous sections investigating this oscillator setup, and using the same assumptions as in those sections. The signal can be downsampled to symbol rate after the SFO has been compensated, e.g., before CFO compensation and MIMO equalization.

4.3 Adaptive Equalization

So far, it was shown in this chapter that the timing impairment compensation can be separated from standard MIMO equalization, and that simplifications can be made depending on the oscillator setup. Another approach is to use adaptive filters in order to adjust the full time-varying MIMO equalizer, irrespective of the actual system configuration[11]. It was seen in the previous chapter in section 3.4 that adaptive filters are able to track general time-varying MIMO channels, which required knowledge about the transmitted symbols. In the current chapter, it was described that the equalizer generally needs to be time varying, i.e., needs to change for every set of received samples that correspond to a transmitted symbol, in order to remove the effects due to the timing impairments. It is possible to use an adaptive filter to replace the previous equalizer, which was fed by one-shot estimates of the parameters, and adopt it based on an error signal that is available in the receiver[12]. One of the simplest instances of this approach can be seen in Figure 4.2, where a hard decision[13] on the equalized symbols is used in order to derive the error signal [28], i.e., it is a decision-directed (DD) approach.

The equalization filter can take any of the forms that have been discussed so far in this chapter. The general equations for adaptive equalization using a feedforward and a feedback part based on the LMS algorithm can be written as follows. The estimates of the symbols are given

[11] Adaptive filters could also be modified to compensate the effects separately depending on the oscillator setup. This concept will not be pursued further in this work.

[12] Only the symbol updated versions of the adaptive algorithm will be presented, as they should yield the best tracking performance.

[13] Decoding could be included to increase the reliability of the decisions, but introduces additional delay.

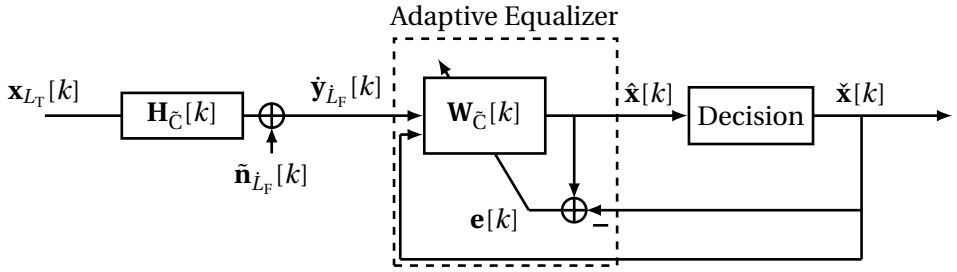


Figure 4.2: Basic working principle of adaptive equalization, i.e., adapting the equalizer $\mathbf{W}_{\tilde{\mathrm{C}}}[k]$ based on the observed error signal $\mathbf{e}[k]$, which contains the difference between the estimated signal $\hat{\mathbf{x}}[k]$ from the equalizer output, and the decided values $\check{\mathbf{x}}[k]$ of those estimated symbols.

from (4.19) as

$$\hat{\mathbf{x}}[k] = \mathbf{W}_{\tilde{\mathrm{C}}}[k] \underbrace{\begin{bmatrix} \dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}[k] \\ -\check{\mathbf{x}}_{L_\mathrm{B}}[k-1] \end{bmatrix}}_{\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}[k]}. \tag{4.57}$$

The update of the equalization matrix $\mathbf{W}_{\tilde{\mathrm{C}}}[k]$ is given as

$$\mathbf{W}_{\tilde{\mathrm{C}}}[k+1] = \mathbf{W}_{\tilde{\mathrm{C}}}[k] - \mu \cdot \mathbf{e}[k] \dot{\mathbf{y}}^{\mathrm{H}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}[k], \tag{4.58}$$

where, as in the estimation part of section 3.4, selecting the step size μ and the initialization of the equalization filter $\mathbf{W}_{\tilde{\mathrm{C}}}[0]$ is important in order to ensure convergence and good tracking performance. Especially initialization is important in order for the DD approach to work. The error signal is defined by

$$\mathbf{e}[k] = \hat{\mathbf{x}}[k] - \check{\mathbf{x}}[k], \tag{4.59}$$

with $\check{\mathbf{x}}[k]$ being the decision based on the estimate $\hat{\mathbf{x}}[k]$. The equations (4.57), (4.58), and (4.59) are the only things that need to be implemented for a DD time-varying adaptive MIMO equalizer, dealing with all channel impairments, based on the LMS algorithm.

4.3.1 Significant Statistics for Equalization Based on LMS

From the previous chapter exploring the channel estimation case, it is known that the convergence behavior and tracking performance of adaptive filters depends on the statistics of the filters input signal, in this case $\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}[k]$. The correlation matrix of this signal is defined by

$$\mathbf{C}_{\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}}[k] = \mathrm{E}\left[\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}[k] \dot{\mathbf{y}}^{\mathrm{H}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}[k]\right] \tag{4.60}$$

$$= \begin{bmatrix} \mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}}[k] & \mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}\check{\mathbf{x}}_{L_\mathrm{B}}}[k] \\ \mathbf{C}_{\check{\mathbf{x}}_{L_\mathrm{B}}\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}}[k] & \mathbf{C}_{\check{\mathbf{x}}_{L_\mathrm{B}}}[k] \end{bmatrix}. \tag{4.61}$$

Section 3.4.1 of the previous chapter showed that

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}}[k] = \mathrm{E}\left[\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}[k]\dot{\mathbf{y}}^\mathrm{H}_{\dot{L}_\mathrm{F}}[k]\right] \tag{4.62}$$

$$= \mathbf{H}_{\tilde{\mathrm{C}}}[k]\cdot\mathbf{C}_{\mathbf{x}_{L_\mathrm{T}}}[k]\cdot\mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}}}[k] + \mathbf{C}_{\tilde{\mathbf{n}}}[k] \tag{4.63}$$

$$\approx \sigma^2_{x_n}\cdot\mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}}}[k] + \sigma^2_{\dot{n}_m}\cdot\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]\mathbf{H}^\mathrm{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k], \tag{4.64}$$

Furthermore, assuming properly working equalization, see also [101], we can state

$$\mathbf{C}_{\check{\mathbf{x}}_{L_\mathrm{B}}}[k] = \mathrm{E}\left[-\check{\mathbf{x}}_{L_\mathrm{B}}[k-1]\left(-\check{\mathbf{x}}_{L_\mathrm{B}}[k-1]\right)^\mathrm{H}\right] = \mathbf{C}_{\mathbf{x}_{L_\mathrm{B}}}[k] = \mathbf{C}_{\mathbf{x}_{L_\mathrm{B}}} \tag{4.65}$$

$$= \mathrm{E}\left[\mathbf{x}_{L_\mathrm{B}}[k]\mathbf{x}^\mathrm{H}_{L_\mathrm{B}}[k]\right] \tag{4.66}$$

$$\approx \sigma^2_{x_n}\mathbf{I}_{NL_\mathrm{B}}, \tag{4.67}$$

meaning that the equalized output symbols are uncorrelated across time and space. Finally, using the same assumption, the cross terms are defined by

$$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}\check{\mathbf{x}}_{L_\mathrm{B}}}[k] = \mathrm{E}\left[\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}[k]\left(-\check{\mathbf{x}}_{L_\mathrm{B}}[k-1]\right)^\mathrm{H}\right] = -\mathrm{E}\left[\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}[k]\mathbf{x}^\mathrm{H}_{L_\mathrm{B}}[k-1]\right] \tag{4.68}$$

$$= -\mathrm{E}\left[\left[\mathbf{H}_{\tilde{\mathrm{C}}}[k]\mathbf{x}_{L_\mathrm{F}+L_\mathrm{C}-1}[k] + \tilde{\mathbf{n}}_{\dot{L}_\mathrm{F}}[k]\right]\mathbf{x}^\mathrm{H}_{L_\mathrm{B}}[k-1]\right] \tag{4.69}$$

$$= -\mathbf{H}_{\tilde{\mathrm{C}}}[k]\begin{bmatrix}\mathbf{0}_{N\times NL_\mathrm{B}}\\ \mathbf{C}_{\mathbf{x}_{L_\mathrm{B}}}[k]\\ \mathbf{0}_{N(L_\mathrm{F}+L_\mathrm{C}-L_\mathrm{B}-2)\times NL_\mathrm{B}}\end{bmatrix} = \mathbf{C}^\mathrm{H}_{\check{\mathbf{x}}_{L_\mathrm{B}}\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}}[k] \tag{4.70}$$

$$\approx -\sigma^2_{x_n}\mathbf{H}_{\tilde{\mathrm{C}}}[k]\begin{bmatrix}\mathbf{0}_{N\times NL_\mathrm{B}}\\ \mathbf{I}_{NL_\mathrm{B}}\\ \mathbf{0}_{N(L_\mathrm{F}+L_\mathrm{C}-L_\mathrm{B}-2)\times NL_\mathrm{B}}\end{bmatrix}. \tag{4.71}$$

In order to determine suitable values for the step size μ, the significant eigenvalues of the composite correlation matrix $\mathbf{C}_{\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_\mathrm{F,B}}}[k]$, and the time variations in the statistics, need to be investigated. In the following, the convergence and tracking characteristics of adaptive equalizers based on the LMS principle will be considered, and it is discussed how to select μ.

4.3.2 Convergence Behavior of LMS-Based Equalization

There are two approaches to deal with convergence for the equalization case. First, it can be assumed that a good channel estimate $\hat{\mathbf{H}}_\mathrm{C}[k]$ is available during initialization of the equalizer, e.g., by employing the techniques of chapter 3. From this estimate, the equalizer coefficients for the feedforward and feedback filters can be computed see, e.g., [100, 101, 102]. Assuming that there are no significant variations in the channel, or, if that is the case that the filter is reinitialized with a new channel estimate, we only need to worry about the tracking behavior, which will be described in the next section.

The other approach assumes that no initial channel knowledge is available, i.e., $\mathbf{W}_{\tilde{\mathrm{C}}}[0] = \mathbf{0}_{N\times(M\dot{L}_\mathrm{F}+NL_\mathrm{B})}$, such that the equalizer needs to be trained from scratch. In this case, train-

ing signals, as discussed in the estimation chapter, are required in order for the equalizer to converge to the proper solution. It is also assumed that the channel is approximately time invariant, see appendix A.2, during the convergence process. First, consider the simpler case of just using a feedforward equalizer, i.e., $L_\mathrm{B} = 0$, for which the correlation matrix reduces to $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{F}}}[k]$. In this case, the significant eigenvalues of the correlation matrix of a Q-fold oversampled MIMO signal need to be determined. It was discussed in section 3.4.5 that the eigenvalues of such a correlation matrix depend on the spectra of the received signals, as well as their correlations, see also [24] for the equalization case. Corresponding to section 3.4.5, and assuming linear channels, there will be approximately $M \cdot \frac{\dot{L}_\mathrm{F}}{Q} \cdot (1+\beta_\mathrm{T})$ significant eigenvalues[14], where β_T is the roll-off of the transmit pulse-shaping filter. If the MIMO channels are uncorrelated (or orthogonal) and flat, as in perfectly designed pure LoS MIMO systems, the eigenvalues are approximately equivalent, and the results of section 3.4.3 with the correlation matrix from (4.64) are directly applicable. In particular, the solutions

$$\mu = \frac{1}{\sigma^2_{\dot{y}_m} \cdot \left(M\frac{\dot{L}_\mathrm{F}}{Q}+1\right)} \text{ or } \mu = \frac{4}{3} \cdot \frac{1}{\sigma^2_{\dot{y}_m} \cdot M\frac{\dot{L}_\mathrm{F}}{Q}},$$

ensure convergence with low steady-state noise for small β_T[15], with $\sigma^2_{\dot{y}_m} = N\sigma^2_{x_n} + \sigma^2_{\dot{n}_m}$ being the average received power at the mth antenna. Going back to an equalizer containing both feedforward and feedback filters, the number of significant eigenvalues is approximately the same, i.e., $M \cdot \frac{\dot{L}_\mathrm{F}}{Q} \cdot (1+\beta_\mathrm{T})$. However, $N \cdot L_\mathrm{B}$ of the eigenvalues increase in value to approximately $(N+1)\sigma^2_{x_n} + \frac{N-1}{N}\sigma^2_{\dot{n}_m}$, compared to $N\sigma^2_{x_n} + \sigma^2_{\dot{n}_m}$ in the previous case. It is possible to reevaluate (3.106) using these eigenvalues, depending on L_B. Two solutions for μ, which have slightly slower convergence speed, but should ensure convergence and low steady-state noise in almost all scenarios, will be given. They are

$$\mu = \frac{1}{2\sigma^2_{\dot{y}_m} \cdot \left(M\frac{\dot{L}_\mathrm{F}}{Q}+1\right)} \text{ or } \mu = \frac{4}{3} \cdot \frac{1}{2\sigma^2_{\dot{y}_m} \cdot M\frac{\dot{L}_\mathrm{F}}{Q}},$$

meaning that μ is half as large as in the previous case. Note that the preceding discussion only holds when $M \cdot \dot{L}_\mathrm{F} > N \cdot L_\mathrm{B}$. When $N \cdot L_\mathrm{B} \geq M \cdot \dot{L}_\mathrm{F}$ as, for example, used for the equalizer proposed in [102], the term $M\frac{\dot{L}_\mathrm{F}}{Q}$ in the last two equations should be replaced with $NL_\mathrm{B}+1$.

When the MIMO channels are frequency selective and not orthogonal, μ generally needs to be smaller than these suggested values [24], which also increases the convergence time. Since the eigenvalues depend in such cases on the exact channel characteristics, see (4.64), no further recommendations for selecting μ will be given. It should be noted that different methods have been proposed in the literature in order to reduce the impact of correlated

[14] Note that the probability of a large eigenvalue spread is much higher for the equalization case, and they may influence the convergence speed of the adaptive equalizer significantly. This property will not be discussed here, where the focus is on the steady-state error performance. Furthermore, depending on the filter shape on the receiver side, there may exist more significant values for very low SNR environments. Such scenarios will also not be discussed here.

[15] For more conservative choices, we may substitute $M\frac{\dot{L}_\mathrm{F}}{Q}$ in the equations with $M\frac{\dot{L}_\mathrm{F}}{Q} \cdot \lceil 1+\beta_\mathrm{T} \rceil$ or $M\frac{\dot{L}_\mathrm{F}}{Q} \cdot 2$.

inputs on convergence, see for example the affine projection algorithm mentioned in [76]. Furthermore, for practical implementations one has to consider the impact of finite-resolution computations and other robustness-reducing factors, as briefly mentioned in section 3.4.2.

4.3.3 Tracking Behavior of LMS-Based Equalization

Assuming the equalizer coefficients have converged, i.e., the MIMO system is properly equalized such that most of the symbol decisions are correct, it needs to be investigated how the adaptive filter generates the changes in the equalizer coefficients, which are necessary due to the timing impairments. As discussed in the previous chapter, and also visible from (4.64) and (4.71), the correlation matrices $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}$ and $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\mathbf{x}_{L_\mathrm{T}}}$ generally vary with time because of the CFOs and SFOs. For the equalization case, it first needs to be checked whether the timing impairments do change the eigenvalues of the correlation matrix $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{\mathrm{C}},\dot{L}_\mathrm{F,B}}}[k]$ significantly.

For CFOs, it was seen in section 2.5.2 that their impact is row- and column-wise multiplication with a complex phase shift that increases on average over time. Thus, they act as diagonal matrices with unit-magnitude complex entries, which cannot change the correlation properties, in particular the eigenvalues, of the MIMO channel. This can also be seen from the feedforward equalizer in Figure 4.1, which shows that the algorithm needs to adopt the phases of the rows and columns of the equalizer according to the CFO processes, but where the inherent phase and amplitude structure, which determines the correlation across space and time, is constant[16]. For SFOs, the situation is slightly different, since sampling differences do influence the eigenvalues to some extent. In general, the correlation in the spatial domain, i.e., how correlated the $M \cdot N$ MIMO channels are, is only negligibly influenced by changes in the sampling phases. The temporal correlation on the other hand, is influenced by the sampling phase variations. However, assuming reasonably flat transmit and receive filters, this change in eigenvalues is small[17] for oversampled systems. Thus, it was qualitatively established that the timing impairments do not change the eigenvalue or correlation structure of the MIMO channel significantly in most cases. However, they do create small variations, which need to be tracked during equalization.

As in the estimation case, discussed in section 3.4.4, the CFO and SFO influences will be treated separately. It was discussed that CFOs correspond to small multiplicative phase variations, while SFOs can be modeled as small additive variations. The tracking error that is generated from each of them depends on the eigenvalues of the correlation matrix $\mathbf{C}_{\dot{\mathbf{y}}_{\dot{\mathrm{C}},\dot{L}_\mathrm{F,B}}}[k]$, the level of offset, and the step size μ, in the same manner as in (3.124), (3.126), and (3.149). Since the significant eigenvalues depend again on the exact channel characteristics, the flat and

[16] Another way of phrasing this is that the CFOs do not change the spatial correlation and the power spectra of the received signals, which are similar to the eigenvalues of the correlation matrix, see Example 7. As mentioned before, this only holds for small CFOs, where the channel does not change significantly due to the shift in frequency.

[17] This can again be visualized in the frequency domain, where the spectra change according to variations of the sampling phases, but most of the spectral characteristics (or significant eigenvalues) are constant. As for the CFO case, this only holds for small sampling phase differences, where the change in bandwidth due to the sampling does not significantly change the channel.

orthogonal MIMO channel case will be discussed first. Then, the optimal choice for μ in the CFO impaired case, similar to the previous chapter, is given by

$$\mu_{\text{opt,CFO}} \approx \left(4 \cdot \frac{\Delta\phi_{\max}^2}{\sigma_{\dot{y}_m}^4 \sigma_{\dot{n}_m}^2 \cdot M \frac{\dot{L}_\text{F}}{Q}} \right)^{\frac{1}{3}}, \tag{4.72}$$

where $\Delta\phi_{\max}$ is the highest normalized carrier frequency difference between all antennas. Likewise, the solution under SFO impairments is

$$\mu_{\text{opt,SFO}} \approx \sqrt{\frac{1-\alpha^2}{\sigma_{\dot{y}_m}^2 \sigma_{\dot{n}_m}^2 \cdot M \frac{\dot{L}_\text{F}}{Q}}}, \tag{4.73}$$

with α being a value very close to one that describes the amount of sampling phase difference, see the second part of section 3.4.4 and in particular (3.145). Note that these two are just an estimate of the true optimum, even in the frequency-flat and orthogonal MIMO channel case. This is because the roll-off β_T and the feedback filter generate unequal eigenvalues, for which finding the true optimum is more complicated, see section 3.4.4. When both impairments are present simultaneously, the solution becomes very complicated, see also the estimation chapter. In such a case, it is suggested to use the higher of the two μ_{opt} solutions, in order to be conservative such that the filter can always follow the faster of the two variations. When no information about the level of the timing impairments and of the noise is available, we can use the solutions from the previous section, assuming that CFOs and SFOs are reasonably small.

If the MIMO channel is not flat and orthogonal, μ also generally needs to be lower than these suggested values in this case. The reason is that the error due to steady-state noise, i.e., the first term of the misadjustment in (3.125), increases with the change in eigenvalues, while the second term is not affected by it for small μ. Thus, the error due to steady-state noise increases, while the error due to tracking lag stays the same. A smaller value for the step size should, hence, be closer to the real optimum.

4.4 Simulation Results

This section will give some example results for adaptive equalization of LoS MIMO channels with multiple timing impairments, under the same assumptions that were used in section 3.5. In particular, $M = N = 4$ Rx/Tx antennas, and a channel with a Rician factor of $K_\text{R} = 10\,\text{dB}$ and length $L_\text{C} = 10$, are considered. The non-line-of-sight (NLoS) taps are drawn from a complex random Gaussian distribution[18]. The equalization performance, specifically the separation of the spatial streams, is influenced by the orthogonality of the LoS MIMO channel. Hence, MIMO channels, which are not perfectly orthogonal, will also be considered to cover a broader range of scenarios. The reason to focus only on adaptive equalization in this section

[18] See section 3.5 for a more complete description of the simulation setup.

is that it offers a relatively low-complexity way to deal with all channel impairments in a simple configuration. Furthermore, MIMO equalization without timing impairments has been extensively treated in the literature, e.g., [102, 103, 105], and adding the timing impairment compensation to such an equalization can be done separately before and/or after it, see section 4.2. The performance of such a structure then just depends on the quality of the estimates of the channel and timing impairments [67]. As references, the same equalizers will be used with perfect channel knowledge, i.e., no channel estimation error, without timing impairments, and assuming no decision errors in the case of DFE. Modulation with 16-QAM symbols will be used exclusively for the simulations.

The focus will be on the two specific equalization structures that were mentioned previously in this chapter. The first one uses only a feedforward filter of length L_F, i.e., $L_\mathrm{B} = 0$. The other one is a decision feedback structure that has been shown to work well in LoS MIMO scenarios [102]. It has the settings $L_\mathrm{F} = 1$ and $L_\mathrm{B} = L_\mathrm{C} - 1$. The equalizers will be used directly, also during the convergence phase, meaning that a certain number of training symbols are required before unknown data can be transmitted. The convergence behavior of these structures will be investigated first in the next section. In order to select the step sizes μ for the adaptive equalizers, it was seen in the previous sections that knowledge about the received power, SNR, and sometimes the timing impairments is required. It is assumed that the former two quantities are known perfectly, while only the order of magnitude of the latter is available. It is assumed that $\sigma^2_{\dot{y}_m} \approx N \cdot \sigma^2_{\dot{x}_n}$.

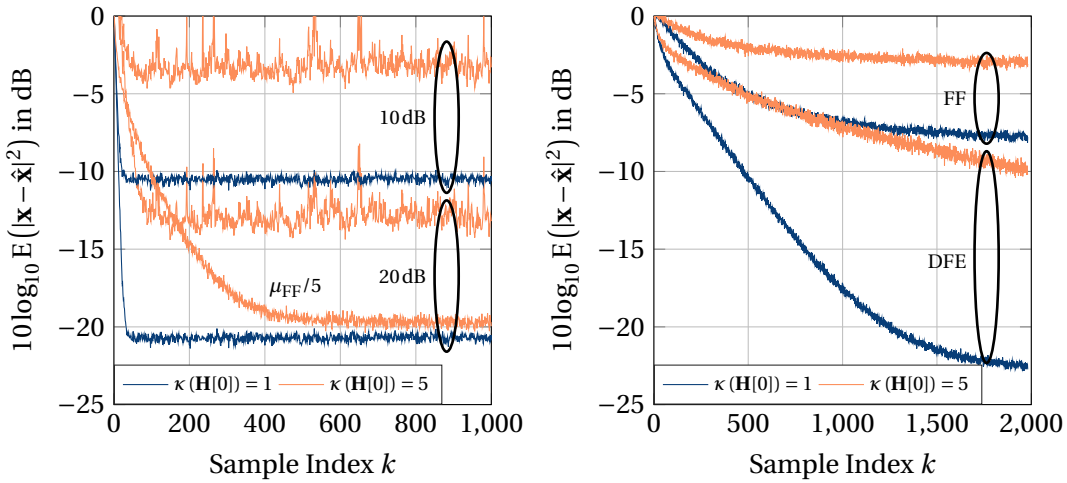
4.4.1 Adaptive Equalization in the CFO Impaired Case

First, consider the case where the LoS MIMO system is only impaired by CFOs and, thus, a symbol-spaced equalizer is sufficient, i.e., $Q = 1$. The convergence behavior of the adaptive equalizers will be investigated first. For it, the transmitted symbols are assumed to be known at the receiver, and suitable step sizes for the ideal LoS MIMO case can be given for the two equalizers as

$$\mu_\mathrm{FF} = \frac{4}{3} \cdot \frac{1}{\sigma^2_{\dot{y}_m} \cdot M L_\mathrm{F}}, \text{ and } \mu_\mathrm{DFE} = \frac{4}{3} \cdot \frac{1}{2\sigma^2_{\dot{y}_m} \cdot (N L_\mathrm{B} + 1)},$$

see the discussion in section 4.3.2. Figure 4.3 contains the MSE of the convergence process, averaged over 200 realizations, of the adaptive equalizers for some example parameters. LoS MIMO taps with different condition numbers[19] $\kappa(\mathbf{H}[k, l = 0])$ were used to cover a broader range of practical cases. The first plot shows the convergence properties for a pure, i.e., frequency flat with $L_\mathrm{C} = 1$, LoS MIMO channel in order to gauge the impact of the orthogonality of the channel. Both equalization structures are equivalent in this case, using $L_\mathrm{F} = 1$ and $L_\mathrm{B} = 0$, and μ_FF was chosen as the basic step size for all scenarios. It is seen that this selection for μ yields in the ideal channel case, i.e., $\kappa(\mathbf{H}[l = 0]) = 1$, fast convergence and steady-state

[19] The definition of condition number and some measured values can be found in section 5.2 of the next chapter.



(a) $L_{\mathrm{C}} = L_{\mathrm{F}} = 1$, $L_{\mathrm{B}} = 0$, $K_{\mathrm{R}} \to \infty$, $\mu_{\varphi_{\mathrm{w}}} = 10^{-3}$ rad

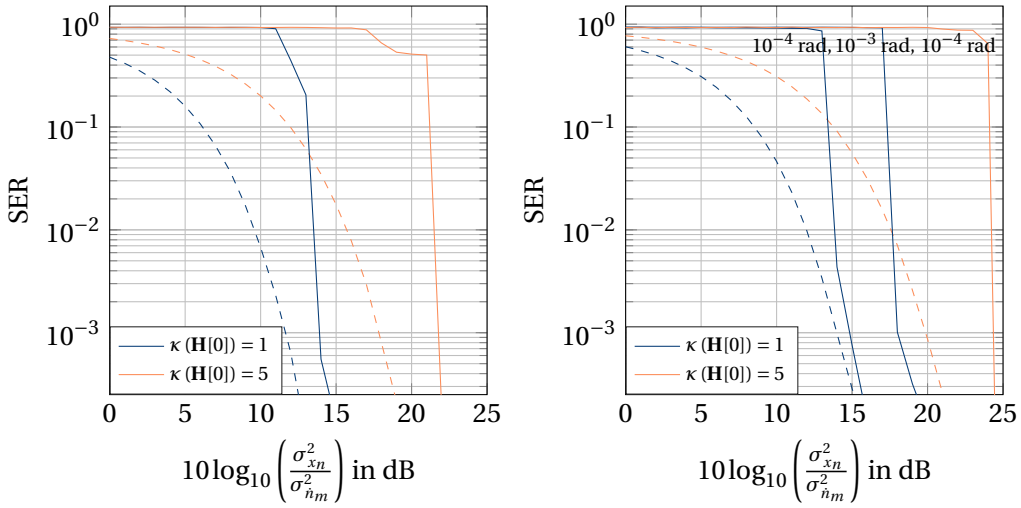
(b) $L_{\mathrm{C}} = 10$, $K_{\mathrm{R}} = 10\,\mathrm{dB}$, $\mu_{\varphi_{\mathrm{w}}} = 10^{-3}$ rad

Figure 4.3: Convergence MSE of adaptive channel equalization for a LoS MIMO system affected by independent purely linearly-drifting carrier phase processes, i.e., $\sigma^2_{\varphi_{\mathrm{w}}} = \sigma^2_{\varphi_{\mathrm{n}}} = 0$, with $M = N = 4$ antennas, assuming perfect knowledge about the transmitted symbols, and using an oversampling factor of $Q = 1$. Step sizes μ chosen according to the discussion in section 4.3.2. Behavior for a pure (frequency-flat) LoS MIMO channel for two different SNRs and two differently-conditioned channels (left). Behavior in a frequency-selective channel for a SNR of 20 dB with different equalization structures (right).

noise similar to the SNR. For channels with higher condition numbers, convergence is slower and the steady-state noise is increased, due to the non-orthogonal nature of the channel. As conjectured in section 4.3.2, μ needs to be smaller in such cases to achieve similar performance. In particular, for this scenario using $\mu_{\mathrm{FF}}/5$ showed comparable error behavior.

For the frequency-selective case, i.e., $L_{\mathrm{C}} = 10$ with $K_{\mathrm{R}} = 10\,\mathrm{dB}$, convergence is generally slower due to the increased length of the required equalization filters, as well as the additional correlation that occurs due to the ISI over time. The length of the adaptive feedforward equalizer is chosen as $L_{\mathrm{F}} = 3 \cdot L_{\mathrm{C}}$, and $\mu_{\mathrm{FF}}/\mu_{\mathrm{DFE}}$, as described above, are used for the respective equalizer. Since the FF equalizer did not converge for this solution[20], $\mu_{\mathrm{FF}}/2$ was actually used for the results shown in Figure 4.3. The results show that the DFE significantly outperforms the FF structure, both in terms of convergence speed, mainly due to the shorter total length, and in terms of steady-state error in such LoS channels. Note, however, that it is much more susceptible to symbol decision errors, which can become critical during data transmission. It is also important to mention that for shorter channels, i.e., lower L_{C}, the adaptive feedforward equalizer can achieve similar steady-state error performance as the feedback equalizer, but

[20] Recall that μ_{FF} is a viable solution if the input signals are uncorrelated across time and space, as discussed in section 4.3.2. Even if the LoS channel is orthogonal, the signals are at least correlated over time due to the frequency-selective channel. On the other hand, as also discussed in section 4.3.2, μ_{DFE} is a more conservative estimate. In fact, for the parameters used here using $2 \cdot \mu_{\mathrm{DFE}}$ yields faster convergence with approximately the same steady-state noise.



(a) $L_C = L_F = 1$, $L_B = 0$, $K_R \to \infty$, $\mu_{\varphi_w} = 10^{-3}$ rad

(b) $L_C = 10$, $K_R = 10$ dB

Figure 4.4: SER of adaptive DFE for a LoS MIMO system affected by independent purely linearly-drifting carrier phase processes, i.e., $\sigma^2_{\varphi_w} = \sigma^2_{\varphi_n} = 0$, with $M = N = 4$ antennas, using an oversampling factor of $Q = 1$, and μ_{DFE}. SER for a pure (frequency-flat) LoS MIMO channel for different SNRs and two differently-conditioned channels (left). Equalization performance in a frequency-selective channel for different CFO levels μ_{φ_w} and the two differently-conditioned channels (right). The dashed lines correspond to the SER of the equalizer in the same channel, but with perfect channel knowledge and without timing impairments.

always requires a longer filter and more convergence time compared to the DFE. Furthermore, as in the estimation case, the level of the CFOs does not impact the convergence properties significantly.

Figure 4.4 shows the symbol error rate (SER), i.e., the probability of a symbol error after adaptive equalization and hard decision, for the DFE in a frequency-flat and frequency-selective channel, with differently-conditioned channels, and CFO levels μ_{φ_w}. The focus is only on DFE, as the convergence results already showed that the MSEs achieved by the FF equalizer are not sufficient, even for high SNRs, to support a 16-QAM transmission in the frequency-selective case[21]. The step size is chosen as the μ_{DFE} described above. The most important feature of all results is the cliff-like behavior of the curves. This is explained by the fact that there is a certain threshold SNR for which the symbol decisions during the tracking phase are mostly correct, and where the adaptation of the equalizer relying on these decisions, thus, starts to work properly. In these regions, the performance gap compared to a system without timing impairments and having perfect channel knowledge is between 0.5 dB and 4 dB. Generally, higher CFOs lead to a larger gap, since the MSE of adaptive filters increase with the CFO, as was seen in chapter 3.

[21] Even without timing impairments, the feedforward equalizer never delivered SERs below 10^{-1} for the considered cases.

4.5 Summary & Main Results

In this chapter, the general derivation of MAP/ML data estimation was carried out, which yields a highly complex detection scheme that scales exponentially with the size of the modulation alphabet and the number of antennas. A simpler way to estimate the data is obtained by linearly filtering the received samples. Due to the memory, i.e., frequency selectivity, which is present in most practical channels, this approach is not ML. In the general case, it was seen that the equalizer needs to be updated at every time step and needs to be fractionally spaced, in order to compensate the SFOs properly. Based on a linear feedforward equalizer, it was shown that the timing impairment compensation can be separated from MIMO equalization, even in the case of multiple independent phase processes at the transmitters and receivers. It can be implemented with parallel time-varying filters and multiplications with complex exponentials before and after equalization on each of the streams. A more general equalization structure, allowing for feedback, e.g., to cancel interference across time from different streams, was then introduced. It uses symbol decisions from previous time steps, and a channel estimate, in order to remove their interference on the current symbols. For both equalization structures, the influence of different system configurations was investigated, showing that certain oscillator setups greatly simplify the equalization and synchronization process. In some cases, standard SISO solutions, which are readily available in the literature, can be applied directly and in parallel. For example, when the oscillator is shared on the receiver side, time-invariant MIMO equalization should be performed first, and the timing impairments can be compensated in parallel by, e.g., using DD PLLs for each stream.

Adaptive equalization, based on the symbol decisions of the estimated symbols, was introduced as an alternative, in order to directly update the equalizer and deal with the timing impairments, without the need for separate timing impairment estimation. It is closely related to the adaptive filtering that was used for channel estimation and tracking in the previous chapter. However, in contrast to it, its performance was shown to depend much more on the channel characteristics, while also depending on the equalization structure. Some guidelines on how to select the step size for adaptive equalization, during convergence and tracking in different scenarios, were also proposed. Generally, μ needs to increase with the timing impairments, while it needs to decrease with the increasing spatial and temporal correlation, which may be experienced due to the channel. Simulation results show the applicability of adaptive equalization in the CFO impaired case. It is seen that convergence is fast in a frequency-flat case, and takes significantly longer in the frequency-selective case. Likewise, a non orthogonal LoS MIMO tap increases convergence time. For the considered example scenarios it is, furthermore, seen that the feedforward equalizer does not sufficiently compensate the channel effects in frequency-selective scenarios to support 16-QAM transmission. SER results for the adaptive decision feedback equalizer show good performance for medium to low CFOs. The performance gap compared to a system without timing impairments and perfect channel knowledge is between 0.5 dB and 4 dB, depending on the scenario. Thus, DD adaptive equalization seems a viable low-complexity approach to compensate the interference

due to the channel, as well as the timing impairments, in one single structure.

5 Experimental Evaluation at 60 GHz

In this chapter, practical implementations of mmWave LoS MIMO systems, and practical tests of some of the methods and ideas that have been presented in the previous chapters, will be presented. The focus is on implementing an energy efficient LoS MIMO wireless backhaul link that is able to provide a transmission rate of several Gbit/s [8].

5.1 Hardware Description and Setup

The experimental system setup[1] consists of two subsystems at transmitter and receiver, respectively. The first one is the baseband signal generation/recording unit. The second one is the front end, including the antenna. A sketch of the system layout can be seen in Figure 5.1. Since the main focus of this chapter is a proof of concept of the developed ideas, rather than a real time implementation, all baseband processing is carried out offline. Thus, subsystem one consists of arbitrary waveform generators (AWGs) at the transmitters and real-time oscilloscopes (RTOs) at the receivers. An AWG can create an arbitrary waveform or signal within its specified parameter range, i.e., certain fixed nominal sampling rates $1/T_{\mathrm{nom,Tx}}$[2], analog bandwidth, and amplitude resolution. The RTO samples the output of the front end on the receiver side and saves it to memory, according to certain fixed nominal sampling rates $1/T_{\mathrm{nom,Rx}}$, analog bandwidth, and amplitude resolution. Another important parameter for MIMO systems are the number of inputs and outputs of each instrument, if the full system should be characterized at the same time. If multiple instruments have to be combined in order to have sufficient inputs or outputs, we have to take care of synchronization or will experience the effects mentioned in chapter 2. Finally, the memory of the instrument defines, in combination with the sampling rate, the maximum length of the generated signal or recording length of one acquisition. A

[1] The hardware and setup was provided and assembled as part of a cooperation between the Chair for Wireless Broadband Communication Systems at HU Berlin, the System Design Department at IHP Frankfurt (Oder), and the Chair of Communication Systems at UniBw Munich, within the DFG funded project "maximumMIMO", part of the SPP 1655 "Wireless Ultra High Data Rate Communication for Mobile Internet Access".

[2] Note that even though one sets a nominal value, these measurement instruments also sample according to some phase process characterized by their LO. Thus, the same effects that were described in earlier chapters can also be observed here, if the sampling devices are not synchronized, e.g., by using a shared reference.

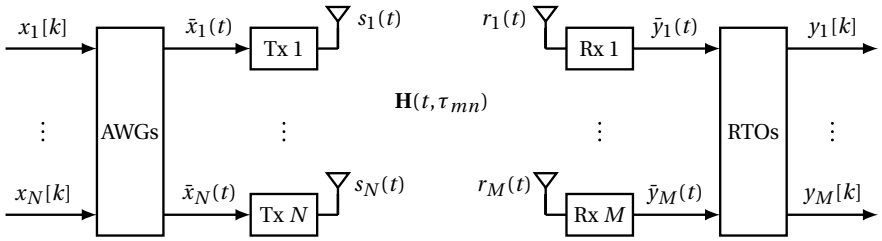


Figure 5.1: Basic system configuration used for the measurements, consisting of baseband signal generation and recording, using AWGs and RTOs, and Tx/Rx front ends with antennas.

summary of the different instruments that were used can be found in Table 5.1[3].

Table 5.1: Overview of measurement instruments with their characteristic parameters, which were used for baseband signal generation and recording in the different system setups.

Parameter	AWG 1	AWG 2	RTO 1	RTO 2
max. $1/T_{\text{nom,Tx/Rx}}$	12 GSa/s	65 GSa/s	20 GSa/s	40 GSa/s
max. Bandwidth	5 GHz	25 GHz	4 GHz	4 GHz
Resolution in ENOB	7.8	4.8	7.0	6.8
Memory	2 GSa	8 GSa	20 MSa	128 MSa
Out-/Inputs	2	8	4	4

Different types of transceiver front ends were available for the overall implementation of the system, all operating in the 60 GHz range with at least 1 GHz of bandwidth. Table 5.2 summarizes the most important properties of the front ends (FEs). Critical parameters for a mmWave MIMO backhaul system are the antenna gain, in order to focus the transmitted power towards the intended receiver, the amount of phase noise, so that the signal is not strongly impaired by it, and the possibility of supplying an external reference clock, for synchronizing the different transceivers. If synchronization is not possible, or intentionally not used, it was seen in the previous chapters that the system can still be synchronized based on the collected output samples.

5.2 Investigation of LoS MIMO Spatial Multiplexing

In this section, the spatial-multiplexing capabilities of LoS MIMO, which have been described in section 2.2, will be explored at 60 GHz. It was seen that the geometric arrangement of the antennas determines the spatial degrees of freedom. In order to measure this, multiple FE 1, see Table 5.2, were used in a one-dimensional MIMO configuration, with $M = N = 2$ and $M = N = 3$, and installed on an adjustable mount using different antenna spacings $d_{\text{Tx,1}} = d_{\text{Rx,1}} = d$, which are equidistant between antennas and equal on the Tx and Rx side, see Figure 5.2 for an example receiver setup. Furthermore, different link distances R were used.

[3] The values noted here are at the maximum sampling rate. The sampling rate can also be chosen lower, which improves, e.g., the resolution and the maximum acquisition time.

Table 5.2: Comparison of the different front ends with their characteristic parameters, which were used in the different system setups.

Parameter	FE 1	FE 2	FE 3
Frequency Range	57 GHz–64 GHz	57 GHz–64 GHz	57 GHz–64 GHz
RF Bandwidth	1.8 GHz	1.25 GHz	1.8 GHz
Tx Power, 1 dB Compr.	11 dBm	11.5 dBm	15 dBm
Tx/Rx Antenna Gain	7.5 dBi	23 dBi	23 dBi
Rx Noise Figure	7 dB	8 dB	8 dB
Phase Noise @1 MHz	-86 dBc/Hz	-104 dBc/Hz	-93 dBc/Hz
External Ref. Clock	yes	no	yes
Number Available	3 Tx, 3 Rx	2 Tx, 2 Rx	4 Tx, 4 Rx

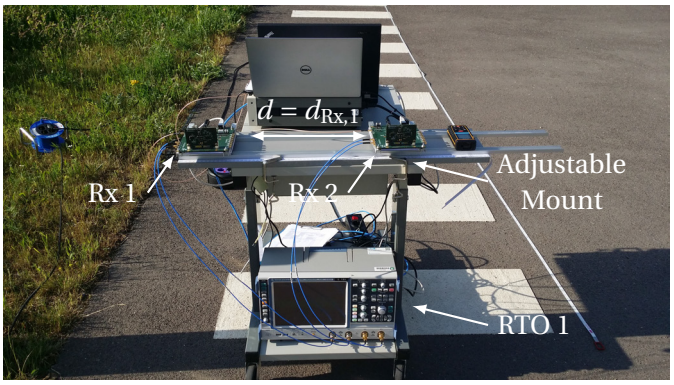


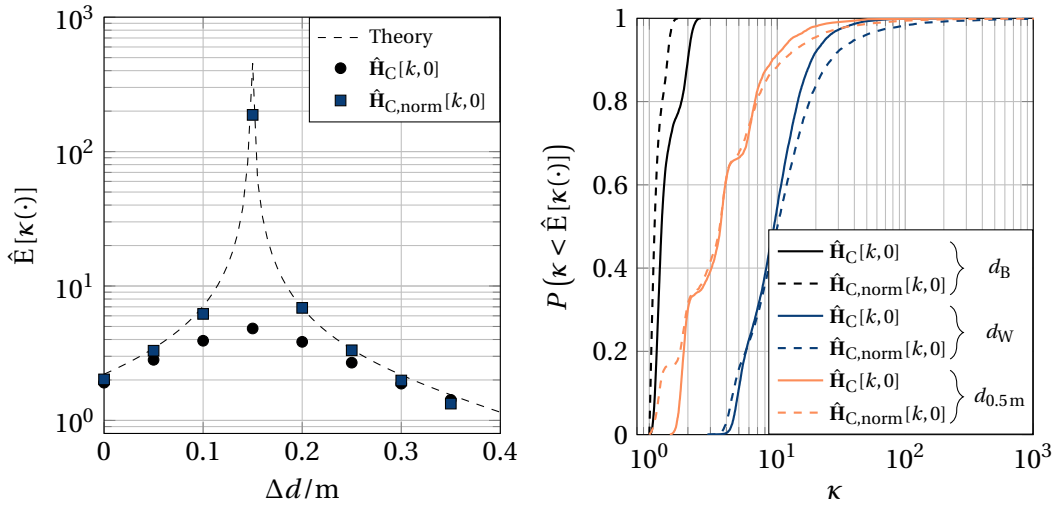
Figure 5.2: Receiver of an $M = N = 2$ LoS MIMO setup, showing two front end 1 modules, and oscilloscope 1 for recording the received baseband signals.

The estimation of the MIMO channel was performed using the correlation technique, with uncorrelated training sequences between different transmitters, as described in chapter 3, see also [108] for further details. Since FE 1 does not provide a lot of antenna gain, long training sequences were used for the larger distances, i.e., larger than 20 m, to yield enough correlation gain overcoming the path loss.

For this section, the focus is on the estimate of the LoS MIMO channel coefficients, i.e., $\hat{\mathbf{H}}_{\mathrm{C}}[k, l_{\mathrm{C}} = 0] = \hat{\mathbf{H}}_{\mathrm{C}}[k, 0] = \left[\hat{\mathbf{h}}_{\mathrm{C},1}[k, 0] \quad \hat{\mathbf{h}}_{\mathrm{C},2}[k, 0] \quad \cdots \quad \hat{\mathbf{h}}_{\mathrm{C},M}[k, 0]\right]^{\mathrm{T}}$, and all other coefficients that may be present are neglected. To check the multiplexing capabilities of the matrix, the condition number is used, which is defined by

$$\kappa\left(\hat{\mathbf{H}}_{\mathrm{C}}[k, 0]\right) = \frac{\sigma_{\max}\left(\hat{\mathbf{H}}_{\mathrm{C}}[k, 0]\right)}{\sigma_{\min}\left(\hat{\mathbf{H}}_{\mathrm{C}}[k, 0]\right)}, \tag{5.1}$$

where $\sigma_{\max}(\cdot)$ and $\sigma_{\min}(\cdot)$ are the largest and smallest singular values of a matrix, respectively. The condition number is one when the matrix is orthogonal, i.e., there are $\min(M, N)$ spatial degrees of freedom, which corresponds to the best-case channel. It is infinity when the



(a) $M = N = 2$, $R \approx 2\,\mathrm{m}$ (b) $M = N = 2$, $R = \{10\,\mathrm{m}, \ldots, 60\,\mathrm{m}\}$

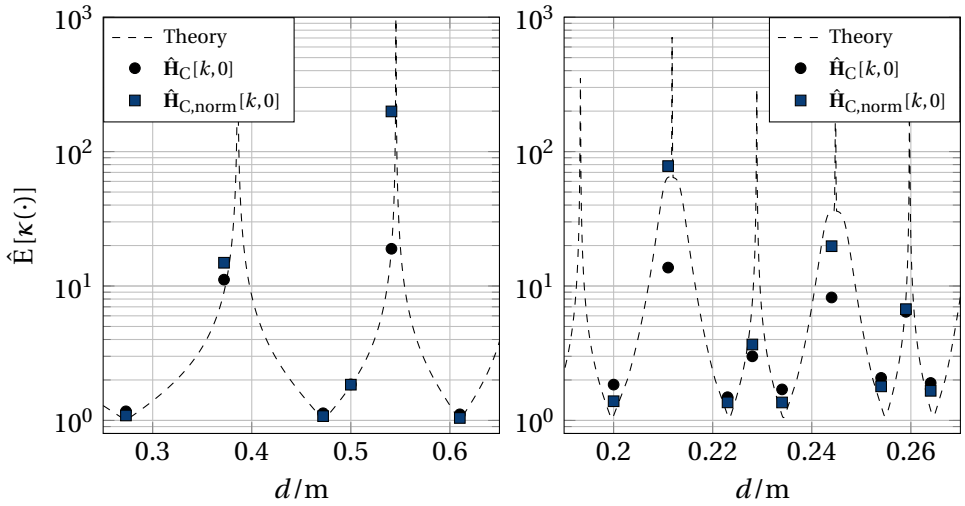
Figure 5.3: Condition number of the LoS MIMO channel tap at different link distances: (a) Changes when moving transmitter 2 by Δd, increasing antenna spacing $d_{\mathrm{Tx},1}$; (b) CDF of all distances, grouped into setups with best-case d_{B}, worst-case d_{W}, and fixed at 0.5 m $d_{0.5\,\mathrm{m}}$, antenna spacings.

columns of the matrix are fully dependent, i.e., there is one spatial degree of freedom and no spatial multiplexing is possible, which is the worst-case channel in terms of multiplexing. We are here mainly interested in the phase relations of $\hat{\mathbf{H}}_{\mathrm{C}}[k,0]$ and their dependence on the geometric arrangement of the antennas, see section 2.2. However, the general estimate also includes the magnitude values of the channel coefficients, which can be different and thereby influence the condition number. To remove that dependence the following normalized matrix is also used in the results

$$\left(\hat{\mathbf{H}}_{\mathrm{C,norm}}[k,0]\right)_{mn} = \frac{\hat{h}_{\mathrm{C},mn}[k,0]}{\left|\hat{h}_{\mathrm{C},mn}[k,0]\right|}. \tag{5.2}$$

A final note about the time dependence is in order. Although the channel estimate changes over time, even for a fixed setup and stable channel, due to the timing impairments, the impact of that effect on the condition number is rather small. To gain higher estimation accuracy, multiple condition number estimates of a certain setup were averaged w.r.t. time, using at least 400 realizations per spacing.

First, the assumed spherical wave propagation, and the corresponding geometric dependence of the LoS MIMO tap, were tested. For the results in Figure 5.3a, a fixed $M = N = 2$ setup with a link distance of $R = 1.993\,\mathrm{m}$ and an initial antenna spacing of $d_{\mathrm{Tx},1} = d_{\mathrm{Rx},1} = 0.180\,\mathrm{m}$ was used. Then, the spacing on the transmitter side was increased by Δd w.r.t. its initial position, i.e., $d_{\mathrm{Tx},1}^{\mathrm{new}} = d_{\mathrm{Tx},1} + \Delta d \neq d_{\mathrm{Rx},1}$, and the change in the condition number of the estimated LoS channel was observed. It can be seen that it follows the theory, when considering spherical



(a) $M = N = 2$, $R = 30\,\mathrm{m}$ (b) $M = N = 3$, $R = 3\,\mathrm{m}$

Figure 5.4: Measured condition number of the LoS MIMO channel tap at fixed distances, for different antenna spacings d, and for different numbers of antennas.

propagation of the waves, i.e., the channel to be given by (2.12). For Figure 5.3b, the cumulative distribution function (CDF) for the condition number of the LoS tap over a number of measured distances was computed. The system was deliberately set up, for each distance, with three different antenna spacing types. The first one is denoted by d_{B}, where, for example, equation (2.20) was used to derive the optimal[4] antenna spacing that yields the best-case channel for every distance. The term d_{W} denotes spacings, which theoretically yield $\kappa \to \infty$, in other words the worst-case channel for each distance. Finally, $d_{0.5\,\mathrm{m}}$ is a fixed spacing of 0.5 m at Tx and Rx that was used irrespective of the change of link distance R. The curves show that if the optimal spacings can be fulfilled, very low condition numbers ($\kappa < 2$) can be achieved, yielding good spatial-multiplexing gain. On the other hand, with the worst-case spacings, the condition number is one order of magnitude higher, while the fixed 0.5 m setup yields a κ between the two extremes.

Figure 5.4 gives two example measurements for an $M = N = 2$ and $M = N = 3$ setup at a fixed distance, where the antenna spacing types described above were used. The figures show that there are multiple optimal spacings, and that the measurements coincide with the theoretical prediction. There is greater variation in the $M = N = 3$ case, because the system becomes more sensitive to small errors in the setup, e.g., small antenna spacing differences and other misalignments, with an increasing number of antennas.

[4] As noted in section 2.2, there are multiple optimal spacings for each link distance.

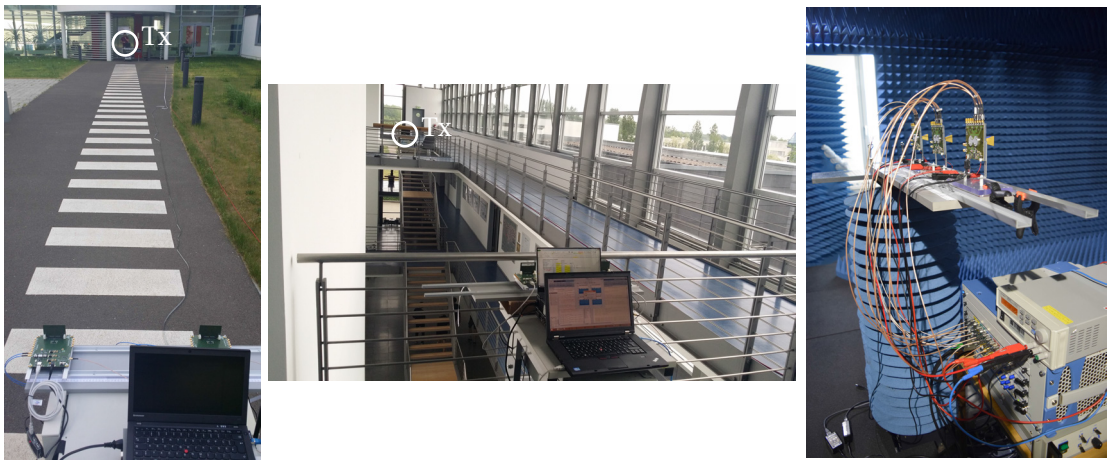


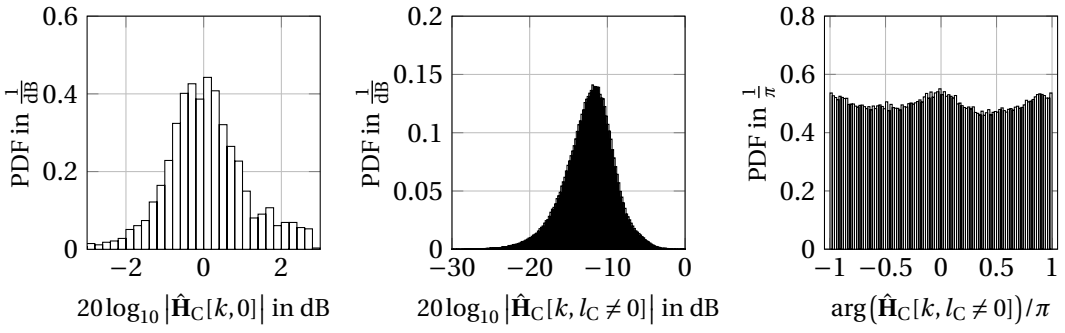
Figure 5.5: Three example measurement scenarios. Left: outdoor $M = N = 2$ scenario, FE 1, link distance between 10 m and 60 m. Center: indoor short-range backhaul $M = N = 2$ scenario, FE 1, distance of 15 m. Right: anechoic chamber $M = N = 3$ scenario, FE 2, distance of 5 m.

5.3 Statistical Properties of LoS MIMO for Backhaul-Like Scenarios

In the previous section, the spatial-multiplexing properties, in particular the required spherical wave propagation, of LoS MIMO in backhaul-like scenarios were established. In order to design a complete wireless communication system, other channel properties, like relative importance and amount of multi-path components, are also of interest. The following results were obtained using FEs 1 and 2, see Table 5.2, with $M = N = 2$ and $M = N = 3$ antennas, in 15 different setups, and with at least 40 snapshots per setup. Some example setups can be found in Figure 5.5. It includes spacings generating differently-conditioned channels, and link distances between 5 m and 60 m. All results are normalized w.r.t. the average LoS tap power of the corresponding setup, see [53] for further details.

Figure 5.6 contains the magnitude and phase distributions of the LoS and NLoS taps computed from the channel estimates of the different setups. It is seen that the NLoS taps are significantly reduced in power, with the average relative value w.r.t. the LoS taps being -11.2 dB, and that their phases are approximately uniformly distributed. The variations in the magnitude of the LoS components are due to the shape of the antenna patterns. For example, using FE 2 with its highly directive antennas, only the bore-sight channel entries have approximately the same gain. This effect is further enhanced by the fact that the systems were set up by hand, as was also observed in [109].

In Figure 5.7, the superposition of 75 consecutive channel impulse response estimates of an example recording of an $M = N = 2$ LoS MIMO setup with FE 2 in an anechoic chamber with a link distance of 5 m, and with a symbol duration of $T_s = 0.8$ ns, is shown. For the plot, an estimate is formed roughly every 0.8 µs, yielding a complete snapshot length of 60 µs (or $7.5 \cdot 10^4$ symbols). Aside from some small variations, the channel appears stable in terms of magnitude over this time frame, especially concerning the most significant taps. Since the



(a) LoS Magnitude (b) NLoS Magnitude (c) NLoS Phase

Figure 5.6: Measured magnitude and phase distributions of LoS and NLoS taps of 15 different measurement setups, each with 40 recordings, with differently-conditioned LoS MIMO taps, $M = N = 2$ and $M = N = 3$, FE 1 & FE 2, and frequency offsets removed. The power variation due to different distances has been removed by normalizing the power w.r.t. the average LoS power for each distance.

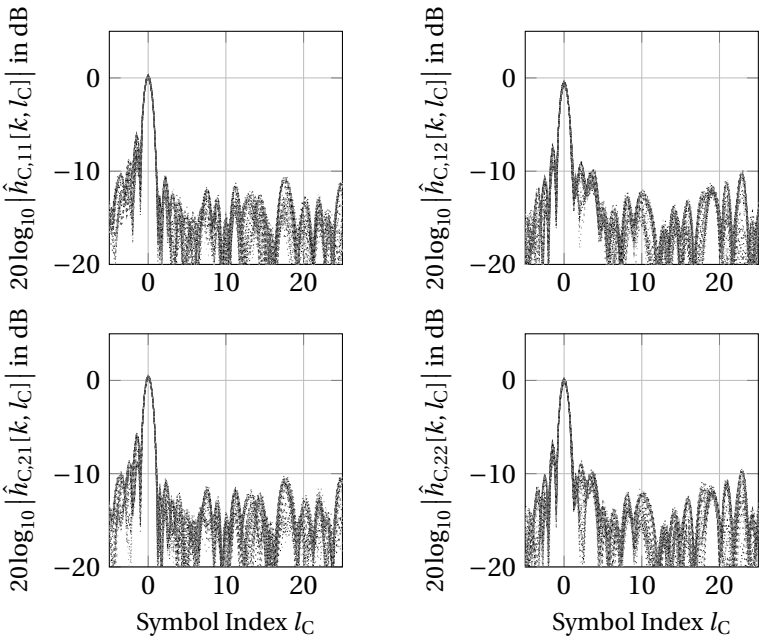
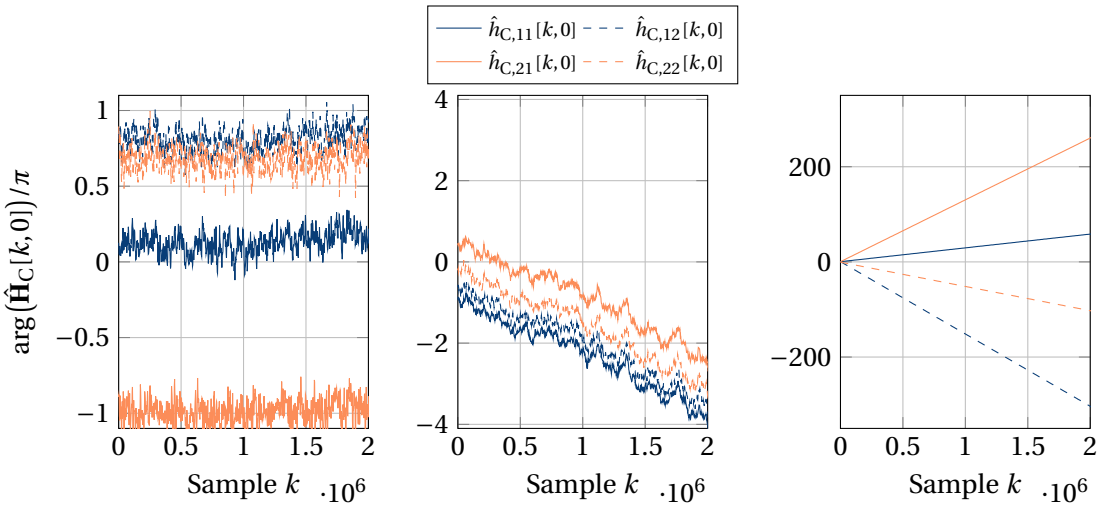


Figure 5.7: Superposition of 75 consecutive channel impulse response estimates of an example recording of an $M = N = 2$ LoS MIMO setup with FE 2 in an anechoic chamber. The sampling interval on the receiver side is $T_{\text{nom,Rx}} = 0.2\,\text{ns}$, the symbol duration is $T_s = 0.8\,\text{ns}$, and the link distance is 5 m. The power has been normalized w.r.t. the average LoS tap power across all antennas.

environment for this measurement was an anechoic chamber, almost no reflections ought to come from the environment. However, multiple taps, especially close to the LoS taps, are visible. It is conjectured that they come from the imperfections in the front ends, as was also mentioned in [7, 9, 109]. In fact, by inspecting the plots in the columns of Figure 5.7, it



(a) Oscillator Setup 1 (b) Oscillator Setup 2 (c) Oscillator Setup 4

Figure 5.8: Example measurement of the channel phases for three different MIMO oscillator setups for carrier frequency generation in an anechoic chamber, using FE 1, $M = N = 2$, and a sampling interval of $T_{\text{nom,Rx}} = 0.1\,\text{ns}$. Setups include three of the four possible configurations mentioned in section 2.4.1.

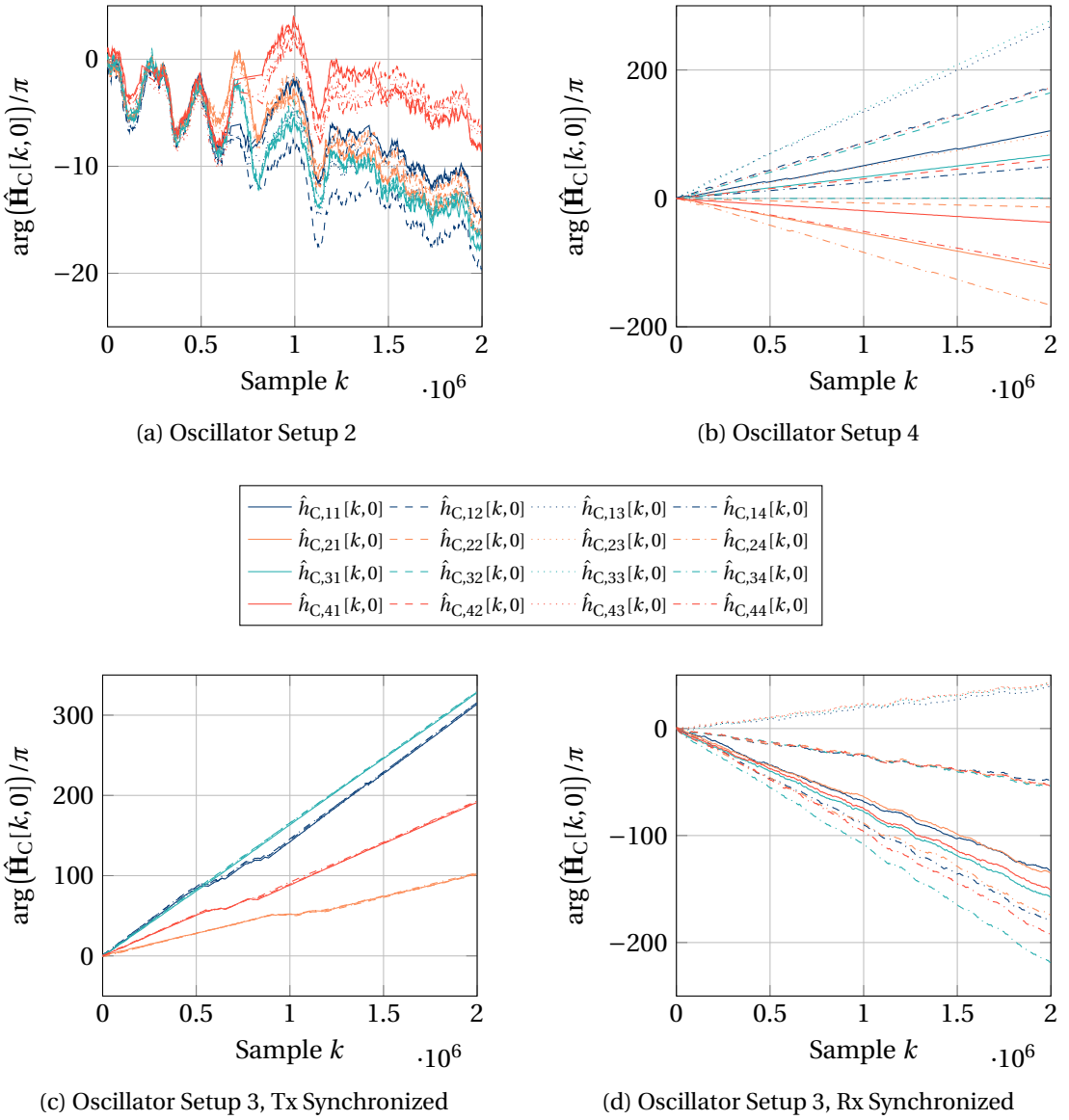
seems that most of the selectivity is coming from the transmitter side, as the column entries have similar significant components at similar positions. Measurements with FE 1 showed similar but less severe behavior. The influence can either be treated as part of the channel and equalized in the same manner, or could be reduced by some additional calibration scheme. Finally, the mean phase variation per sample, due to the CFOs, is computed as $\hat{\text{E}}[\Delta\hat{\phi}] \approx 4 \cdot 10^{-5}\,\text{rad}$, see [53] for details.

5.4 Investigation of Carrier Frequency Oscillator Setups

This section will show the behavior of the different oscillator setups, discussed in section 2.4.1, for carrier frequency generation. As in the previous two sections, the channel impulse response is estimated with a correlation-based estimator, before the phases of the LoS taps are extracted. The first setup uses FE 1, see Table 5.2, in an anechoic chamber, with a well-conditioned LoS MIMO channel, and a sampling interval of $T_{\text{nom,Rx}} = 0.1\,\text{ns}$. The total length of the recording shown in the plots is thus 0.2 ms. More information can be found in [53].

Figure 5.8 shows the extracted phases over time, normalized to π, for three different oscillator setups with $M = N = 2$. There is a fixed initial phase difference in all cases between the sub-channels, caused by the phase differences of the PLLs of the different FEs[5], as well as the phase shift due to the spherical wave propagation. It is seen that the carrier phase processes

[5]Note that although a reference clock is shared, each front end has an independent PLL, which generates the carrier frequency from the reference, and whose phase states are not equivalent.



(a) Oscillator Setup 2

(b) Oscillator Setup 4

(c) Oscillator Setup 3, Tx Synchronized

(d) Oscillator Setup 3, Rx Synchronized

Figure 5.9: Example measurement of the channel phases for three different MIMO oscillator setups for carrier frequency generation in a backhaul setup, using FE 3, $M = N = 4$, and a sampling interval of $T_{\text{nom,Rx}} = 0.2\,\text{ns}$. The bottom figures contain the cases where either all transmitters or all receivers share a reference oscillator.

are not simply linear increasing functions, but have a more complex shape, see also [52] and section 2.4. In particular for oscillator setup 2, we can observe that this behavior is caused by the reference oscillators and not by measurement noise, as all phase processes have a similar shape. For setup 1, no slope aside from the noisy variations is observed over the complete time frame, showing that there is no mean CFO, as expected for a shared reference between all Tx and Rx front ends. Setup 2 shows the same mean slope, corresponding to the difference of the two reference LOs, which were used at the transmitter and receiver, respectively. The

results for setup 4 show four different slopes, corresponding to the difference between the two independent carrier phase processes at the Tx and the two independent carrier phase processes at the Rx. Furthermore, the phase slope is significantly higher compared to the other two cases, as the internal LOs have orders of magnitude higher inaccuracies compared to the external ones. The highest CFO is also experienced in this case, with the average of the maximum being $\hat{\mathrm{E}}[|\Delta\hat{\phi}_{\max}|] \approx 5 \cdot 10^{-4}$ rad per sample, which is the carrier phase difference between the second Tx and first Rx.

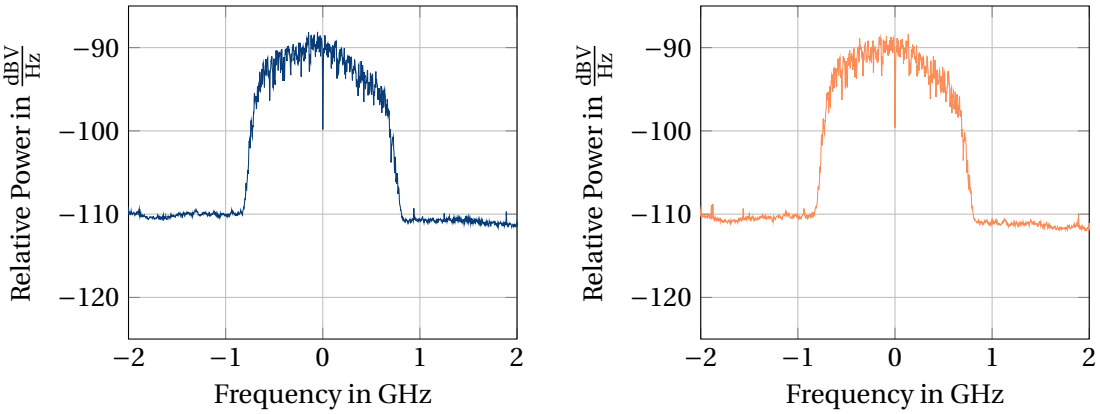
In Figure 5.9, additional example measurements of the time-varying channel phases for different oscillator setups, with $M = N = 4$ and FE 3, are shown. The LoS MIMO system was set up in a backhaul-like scenario, which will be explained in more detail in section 5.5.2. The sampling interval was chosen as $T_{\mathrm{nom,Rx}} = 0.2\,\mathrm{ns}$, yielding a total recording time of 0.4 ms. For oscillator setup 2 and 4, the same observations as above can be made. However, it can be seen that configuration 2 has a more erratic characteristic, also including some phase jumps. Similar behavior was also noticed in [53] for another setup. With the semi shared setups, there are groups of four phase processes that move together, depending on the side where the reference is shared. For example, when all Tx share the same reference for carrier frequency generation, all phase processes associated with one particular Rx antenna move together. The average phase variation per sample is computed as $\hat{\mathrm{E}}[|\Delta\hat{\phi}|] \approx 2.4 \cdot 10^{-4}$ rad, using all phase processes from oscillator setups 3 and 4. Generally, we can observe that the phase progressions are not simple linearly-increasing functions, but have a more complex shape. This means that DD tracking may be a more suitable solution for CFO estimation compared to one-shot estimation, especially when considering a relatively large number of consecutive symbols. The limits for tracking with an adaptive equalizer in a practical LoS MIMO channel can be inferred from Figure 3.4 and Figure 3.7. Tracking seems feasible in such a case up to average phase variations of $\mathrm{E}[|\Delta\phi|] \approx 3 \cdot 10^{-3}$ rad per sample.

5.5 Investigation of Adaptive Equalization Performance

In this section, the adaptive equalization performance for different LoS MIMO system setups at 60 GHz will be tested.

5.5.1 Setup with Two Antennas

First, consider an $M = N = 2$ antenna setup with FE 2, see Table 5.2, in an anechoic chamber, as seen in the rightmost photo of Figure 5.5, with a link distance of 5.5 m. Since FE 2 does not allow for external synchronization, oscillator setup 4 was used for carrier frequency generation. On the Tx side, QPSK symbols with a symbol duration of $T_s = 0.8\,\mathrm{ns}$, and shaped by a root-raised-cosine filter with $\beta_{\mathrm{T}} = 0.3$ and a finite-length of 10 symbols, generated by AWG 1, were transmitted. The receiver side used RTO 1 with $T_{\mathrm{nom,Rx}} = 0.1\,\mathrm{ns}$, leading an oversampling factor of $Q = 8$. Since the sample generation and recording happened in one instrument on the Tx and Rx side each, which was calibrated before the measurements, only a single SFO existed



(a) Receiver 1

(b) Receiver 2

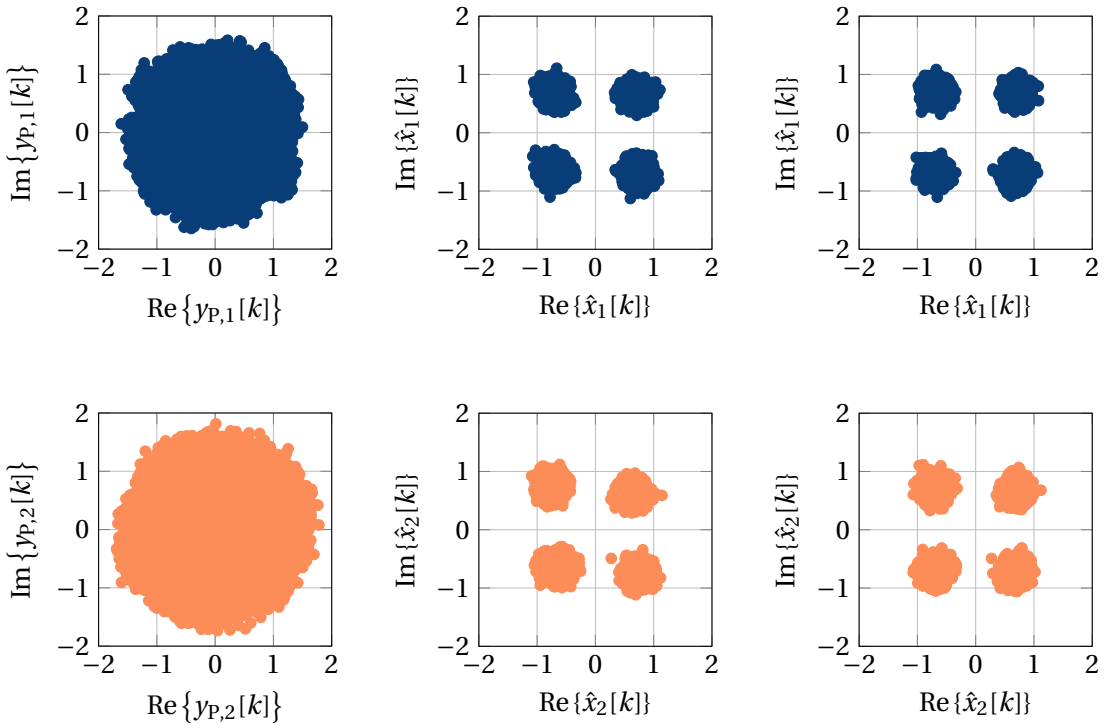
Figure 5.10: Example of the estimated power spectral densities of the received signals for an $M = N = 2$ LoS MIMO setup in an anechoic chamber. QPSK with a symbol duration of $T_s = 0.8\,\text{ns}$ and root-raised-cosine pulse shaping was transmitted. The average SNR estimate from these power spectral densities is 18.5 dB.

between their reference oscillators. The antenna separation was set to one of the optimal solutions. More information about the setup can also be found in [106].

Figure 5.10 shows the estimated power spectral densities of both receive antennas, containing the mixture of both transmitted streams according to the MIMO channel, for the given setup. The SNR at the receiver, defined using the estimated received in-band signal power and the estimated out-of-band noise power with $10\log_{10}\left(\frac{1}{M}\sum m \frac{\hat{\sigma}^2_{y_m}}{\hat{\sigma}^2_{\hat{n}_m}}\right)$, is calculated as roughly 18.5 dB, which serves as a bound[6] on achievable equalization performance. It can furthermore be seen that the signals cover a bandwidth of 1.25 GHz. The condition number of the LoS MIMO tap is estimated, using the aforementioned correlation techniques, as $\hat{\text{E}}[\kappa(\hat{\mathbf{H}}_\text{C}[k,0])] \approx 1.2$. The maximum average CFO per symbol, also estimated using correlation, is given by $\hat{\text{E}}[\Delta\hat{\phi}_\text{max}] \approx 1.5 \cdot 10^{-5}\,\text{rad}$, and is the difference between the oscillators of Tx 2 and Rx 2. For the equalization results shown next, the sampling frequency difference is estimated, through a correlation technique that is a special case to what is described in section 3.3.5, and compensated[7] prior to the further processing. Thus, the single SFO will be considered compensated, and the input signal to the adaptive equalizer is symbol spaced, such that the correct input indices have already been identified. Another processing step that is carried out prior to equalization is matched filtering with the root-raised-cosine filter. The equalizer is thus fed by the symbol-spaced preprocessed signal $\mathbf{y}_\text{P}[k]$, where matched filtering has been performed, and the SFO has been sufficiently compensated.

[6] Note that the amount of transmit power contained in each received stream is only approximately $\sigma^2_{x_n} = \frac{1}{N}\sigma^2_{y_m}$, given equal antenna gain in all directions. However, for the best case MIMO channel, the signals from the M receive antennas combine coherently, such that the power at the MIMO equalizer output is $\sigma^2_{\hat{x}_n} = M \cdot \sigma^2_{x_n}$. For $M = N$, this approximately means $\sigma^2_{\hat{x}_n} = \sigma^2_{y_m}$.

[7] See [106] for details.



(a) Equalizer Input (b) Equalizer Output (c) WL Equalizer Output

Figure 5.11: Constellation diagrams of 10^4 consecutive symbols of two streams, including the effects of multiple CFOs, before and after adaptive equalization, in an $M = N = 2$ LoS MIMO setup with a transmission rate of approximately 5 Gbit/s. A symbol-spaced feedforward LMS equalizer with $L_F = 10$, and its widely-linear version with $L_{F,L} = 10$ and $L_{F,WL} = 3$, were used. The step size is chosen as $\mu_{FF}/5$, and the estimated MSEs are -15.7 dB for the standard version, and -16.9 dB for the widely-linear version, respectively.

The constellation diagrams of 10^4 consecutive symbols per stream from this system setup before and after adaptive equalization, with a feedforward filter of length ten, i.e., $L_F = 10$ and $L_B = 0$, can be seen in Figure 5.11. The reason for using a feedforward equalizer is that it shows very good performance for this well-conditioned MIMO setup. The adaptive equalizers LoS taps are initialized with the pseudoinverse of the LoS tap $\hat{\mathbf{H}}_C[0,0]$, based on an estimate obtained through correlation at the start of transmission, while the other taps are initialized with zeros. After initialization, the equalizer works solely based on the symbol decisions. The step size is chosen as $\mu_{FF}/5$ for the reasons described in the previous chapter, but especially due to the low CFOs. It can be seen that the streams can be well separated and that no significant residual CFOs are visible. The average MSE of the two streams is computed for the standard equalizer as $10\log_{10}\hat{E}\left(|\mathbf{x}-\hat{\mathbf{x}}|^2\right) = -15.7\,\text{dB}$. It is thus shown that the adaptive feedforward equalizer can equalize a practical LoS MIMO channel, including multiple CFOs, yielding a gap compared to the estimated received SNR of less than 3 dB. The transmission rate of the system is approximately 5 Gbit/s, and could be doubled by using 16-QAM, which should be feasible according to the computed MSE. The equalization performance can be

further improved by optimizing the equalizer parameters, e.g., lower step size μ [106] since the CFOs are low, and by using a widely-linear (WL) equalization structure, as is shown in the rightmost constellation of Figure 5.11 and will be explained in the next section.

Widely-Linear Extension of the Adaptive Equalizer

During evaluation of the measurements, it was noticed that the received signals are also impaired by I/Q imbalance [56, 110], i.e., some part of the in-phase component of the signal leaks into the quadrature-phase and vice versa, leading to additional interference that should be compensated for. Several approaches can be found in the literature in order to compensate the effect [110, 111, 112], also including MIMO systems [113], where multiple different imbalances may be present. A very attractive approach is to employ the framework of WL signal processing [110, 114]. The basic premise is that the complete statistical information about a complex signal also requires knowledge about the interaction between real and imaginary part [76, 110, 114]. A simple way of doing this is to use the signal, as well as its conjugate complex for processing.

The equalization equation that was previously used, see section 4.3, can be adopted to include a WL part. Only the feedforward version will be presented here, the extension to a feedback version is straightforward. The estimates for the symbols are obtained through

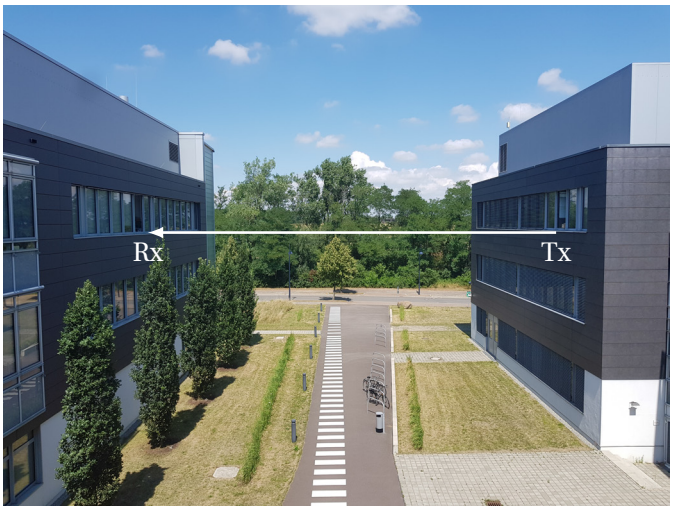
$$\hat{\mathbf{x}}[k] = \mathbf{W}_{\tilde{\mathrm{C}}}[k] \underbrace{\begin{bmatrix} \dot{\mathbf{y}}_{\dot{L}_{\mathrm{F,L}}}[k] \\ \dot{\mathbf{y}}^{*}_{\dot{L}_{\mathrm{F,WL}}}[k] \end{bmatrix}}_{\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_{\mathrm{F,WL}}}[k]}, \tag{5.3}$$

where the equalizer now has length $\dot{L}_{\mathrm{F}} = \dot{L}_{\mathrm{F,L}} + \dot{L}_{\mathrm{F,WL}}$, with $\dot{L}_{\mathrm{F,L}}$ being the length of the standard equalizer, and $\dot{L}_{\mathrm{F,WL}}$ being the length of the WL section. The equalizer is updated with

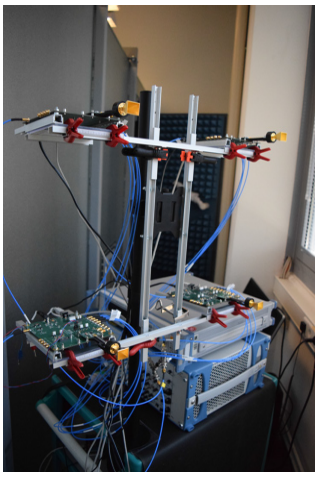
$$\mathbf{W}_{\tilde{\mathrm{C}}}[k+1] = \mathbf{W}_{\tilde{\mathrm{C}}}[k] - \mu \cdot \mathbf{e}[k] \dot{\mathbf{y}}^{\mathrm{H}}_{\tilde{\mathrm{C}},\dot{L}_{\mathrm{F,WL}}}[k]. \tag{5.4}$$

Compared to the standard solution, this means that the equalizer size is increased for the widely-linear version, which increases convergence time and requires lower μ. How to select $L_{\mathrm{F,WL}}$ depends on the type of I/Q imbalance, i.e., is it frequency selective or not. In most cases, small values for $L_{\mathrm{F,WL}}$ should give the most significant gains.

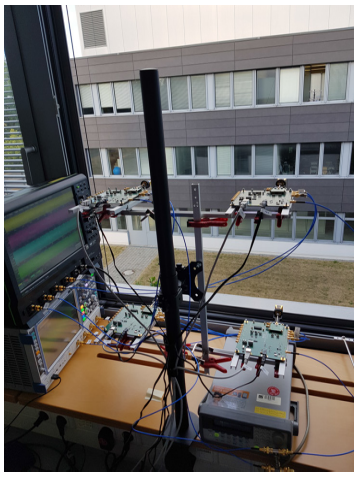
For the WL results shown in Figure 5.11, $L_{\mathrm{F,L}} = 10$ and $L_{\mathrm{F,WL}} = 3$ is used, with $\mu_{\mathrm{FF}}/5$ being adopted accordingly. Visually, no significant improvement seems to occur compared to the standard adaptive equalizer. However, computing the average MSE yields $10\log_{10}\hat{\mathrm{E}}\left(|\mathbf{x}-\hat{\mathbf{x}}|^2\right) \approx -16.9\,\mathrm{dB}$, leading to an improvement of 1.2 dB, and reducing the gap compared to the estimated received SNR further.



(a) Channel



(b) Transmitter

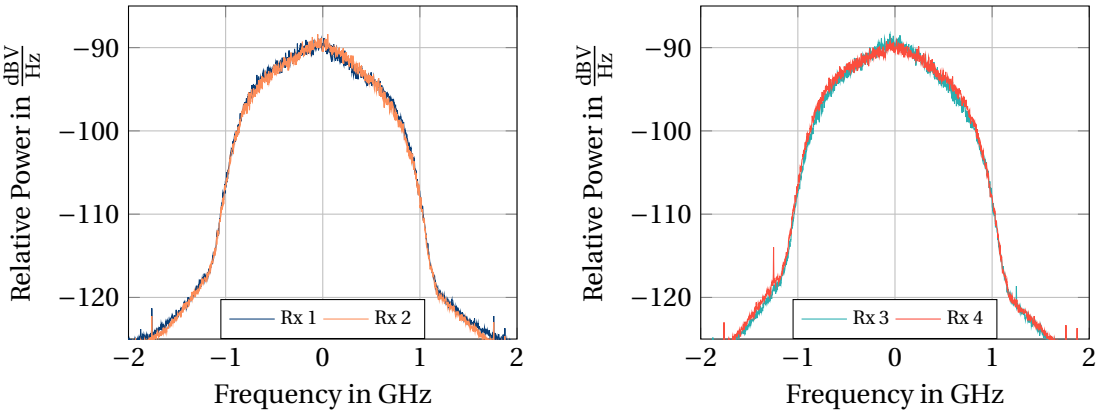


(c) Receiver

Figure 5.12: Backhaul LoS MIMO setup with $M = N = 4$ antennas, using FE 3. The distance between transmitter and receiver is 21.1 m.

5.5.2 Setup with Four Antennas

Consider the $M = N = 4$ antenna setup with FE 3, see Table 5.2, which is shown in Figure 5.12. The setup has a backhaul-like channel and transmits over a distance of 21.1 m. On the Tx side, AWG 2 was used to generate QPSK symbols with a duration of $T_s \approx 0.57\,\text{ns}$, and which are shaped by a root-raised-cosine filter with $\beta_\text{T} = 0.3$ and a finite-length of 20 symbols. In order to have enough sampling inputs on the receiver side, RTOs 1 and 2 were used jointly to sample the received signals. They were set to sample with an interval of $T_{\text{nom,Rx}} = 0.2\,\text{ns}$, yielding an oversampling factor of $\bar{Q} = 2.84$. Different oscillator setups for sampling and carrier frequency generation can be achieved with this setup, due to the available reference input for the carriers of FE 3, and the flexibility of the measurement equipment. Synchronized sampling between the two RTOs was ensured by sharing a reference clock, and aligning the triggers to



(a) Receiver 1 & 2

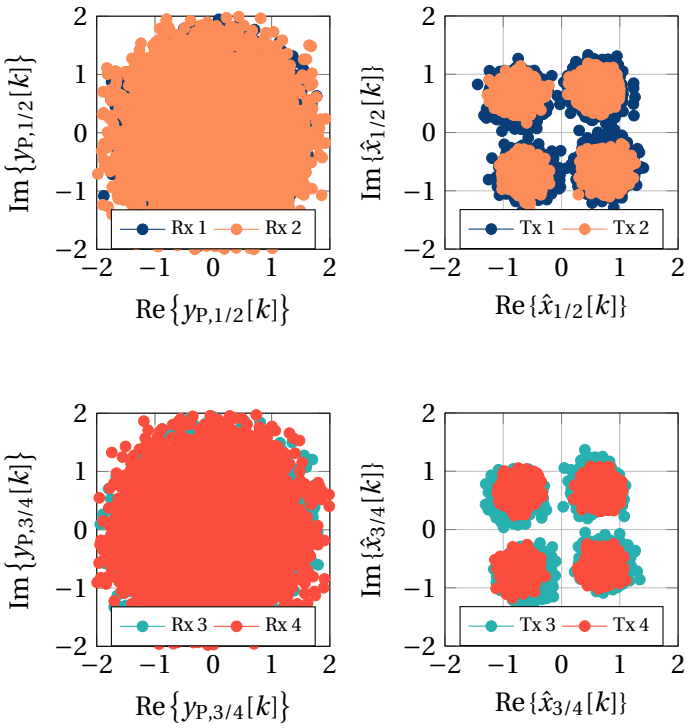
(b) Receiver 3 & 4

Figure 5.13: Example of the estimated power spectral densities of the received signals for an $M = N = 4$ LoS MIMO setup in a backhaul-like scenario. QPSK with a symbol duration of $T_s = 0.57\,\text{ns}$, and root-raised-cosine pulse shaping was transmitted. The estimated SNR from this setup is 24.3 dB.

within $\pm 0.08\,\text{ns}$. Figure 5.13 shows the received estimated power spectral densities of the four receivers. The signals roughly cover a bandwidth of 1.76 GHz. The estimated average SNR is 24.3 dB and the LoS MIMO channel has an estimated condition number of $\hat{\text{E}}[\kappa(\hat{\mathbf{H}}_{\text{C}}[k,0])] \approx 2.2$.

For the equalization, consider a case similar to the one with two antennas. In particular, assume the RTOs on the receiver side to be synchronized, such that only a single SFO exists, and that it has been sufficiently estimated and compensated. More broadly, the signals are pre-processed, prior to adaptive equalization, in the same manner as in the two-antenna case. For carrier frequency generation, oscillator setup 4, i.e., independent oscillators as in the previous section, was used. Then, a symbol-spaced adaptive equalizer may be sufficient for compensating the effects due to the MIMO channel and CFOs. Consider the same feedforward equalizer as above with $L_{\text{F}} = 10$, and using $\mu_{\text{FF}}/2$ as the step size. The adaptive equalizer is initialized directly with a training phase of $5 \cdot 10^3$ symbols[8], after which decision directed adaptation begins. The equalization results are shown in Figure 5.14. It is seen that the streams can be well separated, yielding an average MSE for the four streams of $10\log_{10}\hat{\text{E}}\left(|\mathbf{x}-\hat{\mathbf{x}}|^2\right) = -12.3\,\text{dB}$. The addition of a WL section to the equalizer does not lead to significant performance improvements, leading to the conclusion that these front ends experience less I/Q imbalance. Overall, there is still a significant performance gap compared to the estimated received SNR. The gap may be closed by considering fractionally-spaced equalization, which can deal better with residual SFO, or by using a DFE. Nevertheless, the used adaptive equalizer works error free with four parallel streams, dealing with multiple CFOs, and achieves a combined transmission rate of 14.1 Gbit/s.

[8] Although this seems like a lot compared to the 10^4 data symbols that are shown in the figures, the equalizer keeps working properly in DD mode over much more symbols.



(a) Equalizer Input (b) Equalizer Output

Figure 5.14: Constellation diagrams of 10^4 consecutive symbols of four streams, including the effects of multiple CFOs, before and after adaptive equalization, in an $M = N = 4$ LoS MIMO setup with a transmission rate of approximately 14.1 Gbit/s. A symbol-spaced feedforward adaptive LMS equalizer with $L_F = 10$ was used. The step size is chosen as $\mu_{FF}/2$, and the estimated MSE at the equalizer output is -12.3 dB.

5.6 Summary & Main Results

In this chapter, different topics regarding LoS MIMO with multiple timing impairments were investigated experimentally in the 60 GHz band. It was first shown that we can indeed observe the spherical wave propagation in practice, which leads to the possibility of spatial multiplexing in LoS conditions, as described in chapter 2. It was seen that when the optimal-spacing criterion can be fulfilled, low condition numbers are observed leading to highly orthogonal MIMO channels. This holds true for different system setups, in different conditions, and with different numbers of antennas, as long as a LoS component is present. In general, low condition numbers of $\kappa < 3$ are achievable, when care is taken during system set up. Some of the statistical variations occurring in these different LoS MIMO systems were investigated next. A main result is that on average the observed NLoS taps are approximately 11.2 dB reduced in power compared to the LoS taps at mmWave frequencies. Qualitative investigations of the measurements seem to agree with the literature [12, 92, 93, 94] in that the delay spread for LoS links in this frequency range is a few to some tens of nanoseconds. Furthermore, it was seen that the LoS taps experience some variations for each setup w.r.t. their mean value.

This is most likely due to the manual set up of the systems, where it is difficult for longer distances to properly align the antennas, which generates power variations due to the antenna patterns. It was also briefly discussed that frequency selectivity due to the front ends may play an important role for some of the wideband systems, as also mentioned in the literature.

Next, MIMO channel phase variations caused by CFOs in different oscillator setups were examined. As predicted in the previous chapters, when a reference oscillator is shared between the Tx and Rx front ends, respectively, a single carrier phase difference process is observed. When a reference is shared on only one side, groups of the carrier phase difference processes move together. With totally independent oscillators on both sides, we can observe $M \cdot N$ carrier phase difference processes, which are generated from all possible combinations of the $M+N$ offsets with respect to the nominal frequency, as described in section 3.3.2. Typical measured CFO values, meaning average phase increments per symbols, are in the range of $\hat{\mathrm{E}}[|\Delta\hat{\phi}|] \approx 10^{-5}$ rad to 10^{-3} rad, which is similar to the predicted values of chapter 2 for the given frequency range and bandwidth. Even though these average values are helpful for system design, it was also seen that the phase processes can have a more complex shape than simple linearly-increasing functions for many of the setups. Thus, the phase process model proposed in chapter 2, which on top of the linear drift also has a Wiener and a white noise component, and which was used for simulating the estimation performance in chapter 3, may give a more accurate prediction of practical performance than the commonly used purely linearly-drifting model. More generally, all of the observed MIMO channel variations indicate that decision-directed tracking can be a good option for these systems, potentially requiring less overhead compared to one-shot estimators, which may need more frequent estimations due to the variations from, e.g., the carrier phase processes.

The final part of this chapter was the application of adaptive equalizers, introduced in chapter 4, to two practical LoS MIMO setups. The first one consists of two transmit and receive antennas, an independent oscillator setup, and is able to provide a transmission rate of approximately 5 Gbit/s over a bandwidth of 1.25 GHz. The adaptive equalizer can recover both MIMO streams error free, while simultaneously dealing with the multiple CFOs. The performance gap compared to the receive SNR is less than 3 dB. Adding widely-linear processing to the equalizer, in order to deal with I/Q imbalance, was seen to provide a significant performance improvement. The second LoS MIMO setup uses four transmit and receive antennas, an independent oscillator setup, and provides a rate of 14.1 Gbit/s over a bandwidth of 1.76 GHz. In this case, the adaptive equalizer is also able to recover all four streams error free, but the gap compared to the receive SNR is significantly larger, leaving room for further improvement of the equalizer. In general, these practical equalization results show that adaptive equalization is a suitable approach for LoS MIMO systems with multiple timing impairments.

6 Epilogue

6.1 Summary and Conclusion

Synchronization is one of the key signal processing tasks in order to enable wireless communications. In most systems, we have to deal with a difference in carrier and a difference in sampling frequency, due to inherent hardware variations, while also being affected by the propagation through the channel. For MIMO systems, the number of these impairments affecting the transmission increases with the number of transmitters and receivers. Due to their number and their interdependence, low-complexity approaches are typically required in order to compensate them.

Depending on the ease of sharing reference frequencies between different Tx and Rx front ends, a wide variety of possible, or required, oscillator setups may be used for carrier and sampling frequency generation. Depending on the setup, a different number of frequency offsets exists in the system. The goal of this thesis has been to provide a unified approach to deal with these two timing impairments, namely CFOs and SFOs, in MIMO systems in the most general configuration, where multiple independent oscillators may exist in the system. The focus has been mostly on LoS MIMO links at mmWave frequencies in backhaul-like scenarios, i.e., stationary transmitters and receivers with one dominant transmission path and relatively large bandwidth in the GHz range. Nevertheless, many of the results can also be used for other MIMO scenarios. It was seen that the multiple timing impairments do not make spatial multiplexing, or MIMO transmission more broadly, infeasible. They require, however, that some of the processing steps are carried out in a certain order. It was discussed that multiple SFOs can only be properly compensated in the most general case, if a full fractionally-spaced equalizer is used, due to irreversible aliasing that would otherwise occur. When only multiple CFOs are present in the system, symbol-spaced processing is sufficient, given that the CFOs are relatively slow. Several low-complexity estimators have been proposed, most of them based on correlation, in order to estimate the MIMO channel and the timing impairments. Equalization and synchronization based on these estimates was shown to be possible with relatively low complexity, since the timing impairments can be compensated in parallel even

in the case of independent oscillators. However, in order to do that it is necessary to separate the Tx and Rx contributions from the estimated phase difference processes, which was shown to be possible. Furthermore, if an oscillator, or a reference, is shared on either side of the link, synchronization becomes significantly simpler, and several of the well-known techniques from the literature, such as digital PLLs, can be used.

The MIMO system with timing impairments was also shown to be equivalent to a large time-varying filter, through which the transmitted signals propagate. In order to estimate, track, and compensate this time-varying channel, adaptive filtering was proposed as an alternative strategy. The benefit of this approach is that we can use the decisions during data transmission to update and track the channel, which is superior when the carrier or sampling phase processes have a more complex shape than simple fixed linear phase increments over time, which are very often assumed in the literature. Suitable parameter selection for the adaptive filters was shown to depend on which of the two timing impairments dominates, and also requires that they are relatively slow compared to the symbol rate, which holds true for mmWave systems. In the equalization case, the performance characteristics of the adaptive equalization filter was shown to additionally depend on the MIMO channel. For a properly designed LoS MIMO system at mmWave frequencies, the channel is highly orthogonal and exhibits low frequency selectivity, yielding conditions where adaptive equalization filters can work very well.

Simulations and experimental results at 60 GHz verify the proposed processing approaches, and the characteristics of LoS MIMO at mmWave frequencies. In most cases, they are close to a performance lower bound, or close to the performance of a system without timing impairments. A data rate of 14.1 Gbit/s was achieved with a practical four-antenna LoS MIMO system, using an RF bandwidth of 1.76 GHz, four times spatial multiplexing, and adaptive equalization to deal with the CFOs. The spectral efficiency of the system follows as 8 bit/s/Hz. It was thus shown theoretically and experimentally that multiple timing impairments are not a hindrance for MIMO systems, neither in terms of performance loss nor in terms of complexity, as long as suitable processing is used. The proposed estimation and equalization approaches, as well as adaptive filtering, have been demonstrated to be two such processing options.

6.2 Future Work and Open Topics

There is a range of topics emerging from this work that requires further investigation. The largest topic, with several open problems demanding further examination, is systems experiencing multiple SFOs. In the literature, such systems have been scarcely treated, but they may come to prominence with rising interest in multi-user MIMO systems. This work has introduced their basic system descriptions, and proposed the most basic compensation schemes. However, several fundamental questions such as CRBs, or simplified compensation schemes to enable symbol-spaced equalization, which can lower computational complexity, remain open. Blind estimation approaches, which do not need training signals, have also been

popular to estimate timing impairments in SISO systems. It is not fully clear if and how they can be extended to the general MIMO case with multiple independent oscillators. Consider a typical squaring loop for timing impairment estimation [3, 4]. When it is used in MIMO systems, multiple maxima will be observed in the spectrum, due to the superposition of several transmitted signals of slightly different frequencies, which requires an additional matching of the CFO to the corresponding transmitter. Furthermore, since a nonlinear function is often applied to the signals with such an approach, several mixing products between the different frequencies will occur, making it even harder to identify the correct values. MIMO precoding with multiple timing impairments is another problem worth investigating. For example, in order to create parallel streams on the receiver side, without the need for additional MIMO equalization, the transmitter phase process differences need to be known. Typically, these differences can only be measured on the receiver side, and some delay occurs when feeding that information back to the transmitter. If the phase processes behave linear this is not a problem, but with more complex phase variations, as observed in this work, this becomes an issue. Most of this work was based on single-carrier transmission. There is a huge interest in MIMO systems using multi-carrier waveforms. How the multiple timing impairments are expressed in such a case is also not immediately clear.

For adaptive filtering, it is of importance to investigate the time of convergence, or initialization of the filter before DD tracking can begin, further. This is especially critical when these approaches are extended to more dynamic channel situations. In this work, we treated CFOs and SFOs separately for the LMS approach. It may be worthwhile, albeit complex, to consider both timing impairments simultaneously. This will give a better intuition for selecting the parameters, and the corresponding error performance of the adaptive filters, compared to the qualitative approach that has been taken in this work in cases, when both timing impairments are present. RLS or Kalman filters, although of significantly higher complexity, may also turn out suitable for such a system, as has been partly shown in the literature for the CFO case. A proper complexity comparison between the one-shot estimation and equalization approach, and the adaptive filtering approach is also of interest, with some results being available in [106]. Another open problem w.r.t. SFOs in adaptive filters should be mentioned. For a long enough transmission duration, the SFOs will eventually create sample shifts or removals. This means that the taps of the adaptive filter will wander, and eventually require a delay longer than the filter length. One simple solution to this problem is to track the most significant taps, and add parallel delays before the adaptive filters, when they reach certain thresholds.

From a practical standpoint, it is interesting to search for processing strategies that are of even lower complexity, compared to the proposed approaches of this work. For LoS MIMO at mmWave frequencies this work, and the literature [9, 109], has shown that some ISI may occur due to the hardware. If that is the case, we may use a simpler MIMO equalizer in conjunction with parallel, possibly adaptive, filters, as partly mentioned in this work and used in [106]. It was also discussed in this thesis that some automated antenna alignment strategy could potentially improve the link performance of LoS MIMO. Measurements with a multiple SFO setup are also of great interest, in order to verify the proposed approaches

and identify potential problems. Especially fractionally-spaced adaptive equalization in a practical scenario with such a MIMO setup is of practical importance, and no results from the literature are currently available. A deeper investigation of the hardware impairments, e.g., I/Q imbalance or phase process characteristics, of mmWave MIMO systems can, additionally, help to improve the system design, if typical values and their properties become known. Finally, the carrier and sampling frequency is often derived from the same reference frequency for a given transceiver in practice. This means that the same long-term variations could potentially be seen in the carrier and sampling phase processes. This property has not been considered so far.

A Additional Comments on Oscillators and Their Impact

This appendix presents some further discussions on the impact of oscillators on communications systems, with a focus on LoS MIMO. It is discussed, how data sheet parameters of oscillators, e.g., phase noise values, can be converted to values that can be used for the phase process model discussed in section 2.4.

In general, oscillators have a short-term, i.e., in the seconds range or faster, and a long-term, i.e., hours or longer, behavior. We focus here, and in the work, on the short-term characteristics of oscillators, which are often characterized for system components by their phase noise values. Long-term behavior is more commonly described by an Allan variance [46]. Both attributes can, of course, describe the complete behavior. This is the case when PN values below 1 Hz, or Allan variances below 1 s, are available. One can also convert between the two measures [46]. The short-term behavior is of critical importance here, as it immensely influences high-speed data transmission with symbol durations in the nanoseconds range. Furthermore, some of the long-term characteristics essentially correspond to shifts of the mean frequency, which are inherently compensated by the presented algorithms, given that they still lie within the estimation range. A general description of phase errors in oscillators, and the relation to phase noise and jitter can be found in [51].

A.1 Conversion of Oscillator Parameters to Model Values

The conversion from phase noise and jitter data to the model parameters, which was used in section 2.4 and throughout the work, will be described next. Recall that the model has three contributions, namely a mean drift, a Wiener noise part, and a white noise part. Note that in the following, we compute the values on the sample level, as this was also done in the main text. In general, it is of course more interesting to determine the impact of the model parameters in the signal bandwidth. This can be done by substituting the sample rate with the symbol rate in the following equations.

A.1.1 Carrier Phase Offset Process

The mean variation, or linear drift, of an oscillator is assumed to be described by its accuracy in ppm. Given this, the corresponding model parameter can be calculated with

$$\mu_{\varphi_\mathrm{w}} = 2\pi \cdot \Delta f \cdot T_\mathrm{nom} \tag{A.1}$$
$$= 2\pi \cdot \mathrm{Accuracy[ppm]} \cdot f_\mathrm{nom} \cdot T_\mathrm{nom}, \tag{A.2}$$

where T_nom is the nominal sampling interval at the receiver, Δf is the frequency offset w.r.t. the nominal carrier frequency f_nom, and μ_{φ_w} has the unit[1] of $\frac{\mathrm{rad}}{\mathrm{sample\,interval}}$. For the Wiener part, one can fit a slope to the drop-off region of the phase noise spectrum with

$$\sigma^2_{\varphi_\mathrm{w}} = \left(2\pi \cdot \Delta f_\mathrm{PN}\right)^p \cdot 10^{\frac{\mathrm{PN}(\Delta f_\mathrm{PN})\mathrm{[dBc/Hz]}}{10}} \cdot T_\mathrm{nom}^{p-1} \tag{A.3}$$
$$= \left(2\pi \cdot \Delta f_\mathrm{PN}\right)^2 \cdot 10^{\frac{\mathrm{PN}(\Delta f_\mathrm{PN})\mathrm{[dBc/Hz]}}{10}} \cdot T_\mathrm{nom}, \tag{A.4}$$

where p describes the slope, in the second line $p = 2$, i.e., 20 dB per decade drop, and $\mathrm{PN}(\Delta f_\mathrm{PN})$ is the phase noise power in dBc/Hz at offset frequency Δf_PN. The parameter $\sigma^2_{\varphi_\mathrm{w}}$ has the unit $\frac{\mathrm{rad}^2}{(\mathrm{sample\,interval})^2}$. In the literature, the Wiener part is also sometimes related to the 3 dB bandwidth of the PN spectrum [56]. Finally, the white noise contribution in the model from the data sheet PN values follows using

$$\sigma^2_{\varphi_\mathrm{n}} = \frac{10^{\frac{\text{far-out PN}\mathrm{[dBc/Hz]}}{10}}}{T_\mathrm{nom}} \tag{A.5}$$
$$= \frac{10^{\frac{\mathrm{PN}(\Delta f_\mathrm{PN}\to\infty)\mathrm{[dBc/Hz]}}{10}}}{T_\mathrm{nom}}, \tag{A.6}$$

which also has the unit $\frac{\mathrm{rad}^2}{(\mathrm{sample\,interval})^2}$. It is seen that the last contribution is the only one that increases with an increase in the sampling rate. It has been conjectured in the literature that this term can be performance limiting for wideband systems [50]. A more detailed discussion about the conversion between measured PN values and model parameters can be found in [49, 54].

A.1.2 Sampling Phase Offset Process

For the sampling phase process, the accuracy is also assumed to correspond to the mean variation, i.e.,

$$\mu_{\dot{\varphi}_\mathrm{w}} = \frac{\Delta T}{T_\mathrm{nom}} \tag{A.7}$$
$$= \mathrm{Accuracy[ppm]} \cdot \frac{T_\mathrm{nom}}{T_\mathrm{nom}}, \tag{A.8}$$

[1] In the main text we use $\frac{\mathrm{rad}}{\mathrm{sample}}$ and also just rad to shorten notation. The same applies to the squared quantities.

where ΔT is the sampling time interval change, and T_{nom} is the nominal receiver sampling interval. The unit[2] of $\mu_{\dot{\varphi}_\text{w}}$ is $\frac{\text{samples}}{\text{sample interval}}$. The Wiener part is computed from the RMS jitter value with

$$\sigma^2_{\dot{\varphi}_\text{w}} = \left(\frac{\text{RMS Jitter[s]}}{T_{\text{nom}}}\right)^2, \tag{A.9}$$

and has the unit $\frac{\text{samples}^2}{(\text{sample interval})^2}$. The white noise contribution is given from the aperture jitter with

$$\sigma^2_{\dot{\varphi}_\text{n}} = \left(\frac{\text{Aperture Jitter[s]}}{T_{\text{nom}}}\right)^2. \tag{A.10}$$

The model parameter $\sigma^2_{\dot{\varphi}_\text{n}}$ has also the unit $\frac{\text{samples}^2}{(\text{sample interval})^2}$.

A.2 Approximate Time Invariance

It was mentioned several times in the work that the channel is sometimes assumed to be approximately time invariant[3] over a frame of L_I samples, i.e., $\mathbf{H}_\text{C}[k] \approx \mathbf{H}_\text{C}[k+l_\text{I}]$ for $l_\text{I} = \{1,\ldots,L_\text{I}-1\}$. In this section, this statement will be quantified to some extent for the appropriate parameter range of the two timing impairments in mmWave LoS MIMO systems. It is sufficient to focus on each single entry $h_{\text{C},\cdot}[k]$ of the complete channel matrix $\mathbf{H}_\text{C}[k]$ individually. We can say that the channel is approximately time invariant if the variation within the L_I samples is below a certain threshold, i.e., $\text{E}[|h_{\text{C},\cdot}[k] - h_{\text{C},\cdot}[k+l_I]|^2] / \text{E}[|h_{\text{C},\cdot}[k]|^2] < a$. The parameter a controls how much variation is allowed. For example, in [4] it is proposed that CFO values of ten percent w.r.t. the symbol rate can be tolerated, when matched filtering should be carried out before CFO correction[4]. When making numerical evaluations, we thus also choose[5] $a = 0.1$.

The exact variation of $h_{\text{C},\cdot}[k]$ over time depends on the shape of the timing impairments, as well as the channel characteristic. For convenience, we only consider the linear drift parts here, and separate the two timing impairments. For the CFO case we can, without loss of generality, for the linear drift case state that $h_{\text{C},\cdot}[k] = 1$ and $h_{\text{C},\cdot}[k+l_\text{I}] = 1 \cdot \exp\left(j \cdot \mu_{\varphi_\text{w}} \cdot l_\text{I}\right)$. In the linear

[2] In the main text we use $\frac{\text{samples}}{\text{sample}}$ and also no unit to shorten notation. The same applies to the squared quantities.

[3] This concept is similar to the coherence time of a channel [4, 27, 28], which is usually given for fading channels excluding the timing impairments.

[4] This essentially means that the time variation due to the CFO is insignificant w.r.t. the length of the matched filter.

[5] This was also the threshold that was used to distinguish between frequency-flat and frequency-selective channels.

drift case, the maximum difference is reached when $l_\mathrm{I} = L_\mathrm{I} - 1$. We can then compute

$$a > \frac{\mathrm{E}\left[\left|h_{\mathrm{C},\cdot}[k] - h_{\mathrm{C},\cdot}[k+l_I]\right|^2\right]}{\mathrm{E}\left[\left|h_{\mathrm{C},\cdot}[k]\right|^2\right]} \tag{A.11}$$

$$a > \left|1 - \exp\left(j \cdot \mu_{\varphi_\mathrm{w}} \cdot (L_\mathrm{I} - 1)\right)\right|^2 \tag{A.12}$$

$$1 - \frac{a}{2} < \cos\left(\mu_{\varphi_\mathrm{w}} \cdot (L_\mathrm{I} - 1)\right) \tag{A.13}$$

$$\implies L_\mathrm{I} < \frac{\cos^{-1}\left(1 - \frac{a}{2}\right)}{\left|\mu_{\varphi_\mathrm{w}}\right|} + 1 \tag{A.14}$$

$$L_\mathrm{I} < \frac{\cos^{-1}(0.95)}{\left|\mu_{\varphi_\mathrm{w}}\right|} + 1, \tag{A.15}$$

where for the last line $a = 0.1$ as described above. We can now compute the number of samples L_I, over which the channel is approximately time invariant given the previously computed standard CFO values of Table 2.1. For the highest CFO value this yields $L_\mathrm{I} \approx 4$, and for the lowest one it yields $L_\mathrm{I} \approx 794$. We can see that for high CFO values, the channel changes relatively quickly.

For the SFO case, the situation is more complicated as the change of $h_{\mathrm{C},\cdot}[k]$ depends on the complete channel characteristic. We will make the simplifying assumption that the change is in the order of a if the sample shift is in the order of a. We can then state, with the assumption of only linear sampling phase offset drifts, where the maximum difference occurs at $l_\mathrm{I} = L_\mathrm{I} - 1$, that

$$a > \left|\mu_{\dot{\varphi}_\mathrm{w}}\right| \cdot (L_\mathrm{I} - 1) \tag{A.16}$$

$$\implies L_\mathrm{I} < \frac{a}{\left|\mu_{\dot{\varphi}_\mathrm{w}}\right|} + 1 \tag{A.17}$$

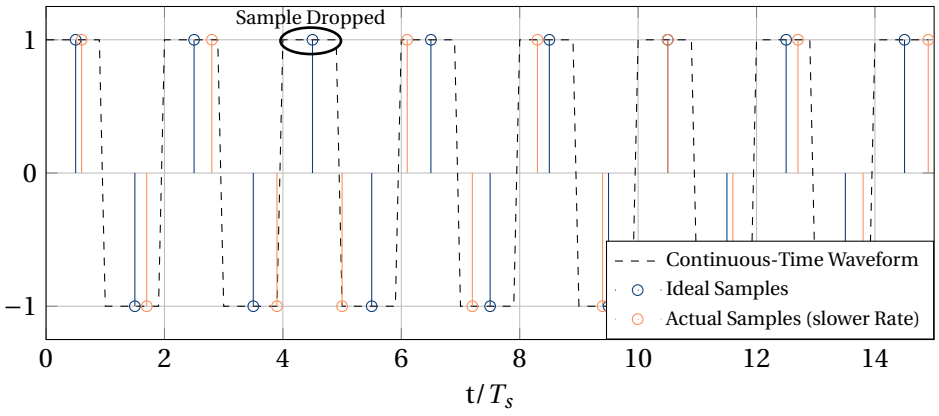
$$L_\mathrm{I} < \frac{0.1}{\left|\mu_{\dot{\varphi}_\mathrm{w}}\right|} + 1. \tag{A.18}$$

For the highest and lowest SFO value from Table 2.1, this gives $L_\mathrm{I} \approx 500$ and $L_\mathrm{I} \approx 10^5$, respectively. As mentioned for the adaptive filter channel tracking in section 3.4.4, it is also seen here that the main variation for typical mmWave components is due to the CFOs.

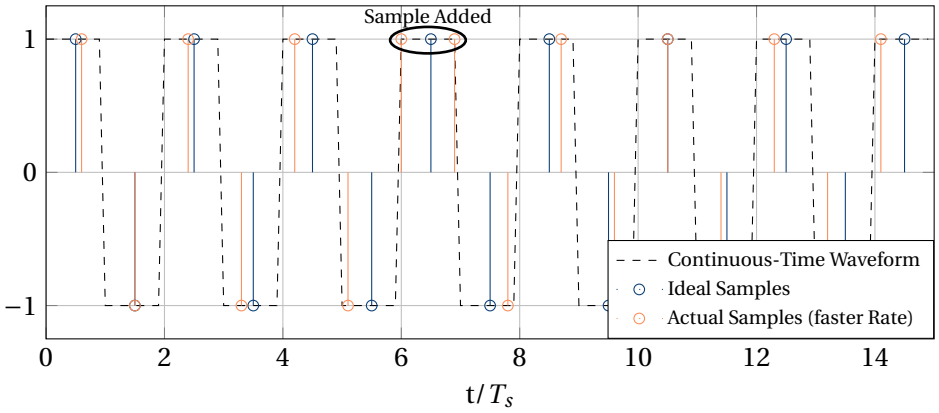
A.3 The Sample Drop and Add Phenomenon

In the work, we talked about the fact that the existence of SFOs will eventually lead to the sample add or drop phenomenon when continuous transmission is considered, see section 2.4 and Figure 2.8. We try to briefly explain, where this effect is coming from, and point to solutions for the problem.

Consider a prototype continuous-time waveform of alternating plus and minus one symbols,



(a) Sample Drop



(b) Sample Add

Figure A.1: Example prototype continuous-time waveform, which is sampled at the symbol rate, and at a slightly impaired symbol rate. In (a), the waveform is sampled with an actual rate that is slightly lower than the symbol rate, which leads to missing the symbol in period five. In (b), the actual sample rate is slightly higher than the nominal symbol rate, and the symbol in period seven is sampled twice.

which carry the information, and which are shaped by a rectangular pulse, as given in Figure A.1. If they are sampled at a fixed nominal rate, which is equivalent to the symbol rate $1/T_s$, one collects exactly one information symbol per symbol period[6], and gets the sequence of alternating plus and minus ones, see the ideal samples in Figure A.1. Then, assume that the actual sampling rate is smaller, i.e., the sample period is higher, than the nominal value. In this case, we sample the continuous-time waveform always a little bit later compared to the ideal case. Eventually, because the sample period is slightly larger than the symbol period, we will miss an information symbol, see the fifth symbol interval in Figure A.1a. Instead of the alternating sequence in the ideally sample case, we thus see a sequence, where two

[6]Although this example describes the symbol-spaced case, the problem occurs in the same way when oversampling is used. Imagine two samples per symbol. At some point only one sample per symbol period, or three samples per period will be generated, respectively. The benefit of oversampling is that we will always have at least one sample per symbol, whereas in the symbol-spaced case we can not recover the sample drop.

consecutive negative ones are seen in the sequence, and a positive one is missed.

In Figure A.1b, the actual sampling rate is slightly higher than the nominal one, i.e., the sampling period is lower. Hence, the continuous-time waveform is always sampled somewhat earlier compared to the ideal-sampling case. At some point, two samples will be generated within a symbol period, see interval seven in Figure A.1b. We see a sequence, where two consecutive positive ones are seen, instead of the alternating sequence. From the previous section, we can estimate when either of the two events is expected to occur. In the previous section, we said that ten percent sample shift is allowable, whereas now we are interested in the case when a full sample shift happens. With the values from before, we can compute that for high SFOs, such an event is expected approximately within each $5 \cdot 10^3$ samples, while for low SFOs it is expected approximately within 10^6 samples. This means that if we have blocks, which contain more than these numbers of samples, either of the two events is probably included in that block.

In a MIMO system, sample drops and adds occur at different points in time for different front end combinations. General remedies to this problem are the use of oversampling, the use of buffers, and mechanisms to detect when such an event has occurred. A more in-depth discussion for the SISO case can be found in [4], and references therein.

B Extended Derivations for LMS Adaptive Filter

This appendix includes more detailed derivations for the selection of the LMS step size μ, and the corresponding misadjustment $\mathscr{M}$, in the case when CFOs or SFOs are present, see section 3.4.4.

B.1 Tracking of CFOs

From section 3.4.4, with the given approximations for the phase term and equal transmit power for all Tx, the misadjustment is given by

$$\mathscr{M} \approx \mu \cdot NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_{\mathrm{T}})} + \frac{1}{\mu^2} \cdot \frac{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}. \tag{B.1}$$

Requesting a misadjustment smaller than one leads to

$$1 \geq \mathscr{M} \tag{B.2}$$

$$0 \geq \mu NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}(2+NL_{\mathrm{T}})} + \frac{1}{\mu^2} \frac{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu NL_{\mathrm{T}} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} - 1 \tag{B.3}$$

$$0 \geq 2\left(1+NL_{\mathrm{T}}\right)\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m} \cdot \mu^3 + \left(\Delta\phi^2_{\max}\sigma^2_{x_{\mathrm{Tr},n}} - 2\sigma^2_{\dot{n}_m}\right) \cdot \mu^2 - 3\Delta\phi^2_{\max} \cdot \mu + \Delta\phi^2_{\max}\frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}}, \tag{B.4}$$

where some approximations about the signs of different terms were used, and which is used with equality to get to the results. Typically, two roots will exist here, where for low CFO values one will be equal to the CFO free case.

Trying to minimize the misadjusment from above gives

$$0 = \frac{\partial \mathscr{M}}{\partial \mu} \tag{B.5}$$

$$\begin{aligned} 0 \approx & \left(2NL_{\mathrm{T}}\sigma^2_{\dot{n}_m} + \Delta\phi^2_{\max}\left(2+NL_{\mathrm{T}}\right)\sigma^2_{x_{\mathrm{Tr},n}}\right)\cdot\mu^3 - \Delta\phi^2_{\max}6\left(2+NL_{\mathrm{T}}\right)\cdot\mu^2 \\ & + \Delta\phi^2_{\max}\frac{6\left(3+NL_{\mathrm{T}}\right)}{\sigma^2_{x_n}}\cdot\mu - \Delta\phi^2_{\max}\frac{8}{\sigma^4_{x_{\mathrm{Tr},n}}}, \end{aligned} \tag{B.6}$$

which is almost exact, except when the parameters create a pole in the derivative, which happens when $2 = \mu\sigma^2_{x_{\mathrm{Tr},n}}\left(2+NL_{\mathrm{T}}\right)$. Furthermore, it can be shown that this equation has only one real solution.

Assuming small values of μ, particularly assuming $2 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}NL_{\mathrm{T}}$, which implies that $2 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}$, and for $NL_{\mathrm{T}} \geq 2$, i.e., any MIMO or frequency-selective case, that $1 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}$, the misadjustment simplifies to

$$\mathscr{M} \approx \mu\cdot NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} + \frac{1}{\mu^2}\cdot\frac{2}{2}\frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}. \tag{B.7}$$

Requesting again a misadjustment smaller than one leads to

$$1 \geq \mathscr{M} \tag{B.8}$$

$$0 \geq \mu NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} + \frac{1}{\mu^2}\frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} - 1 \tag{B.9}$$

$$0 \geq NL_{\mathrm{T}}\sigma^2_{x_{\mathrm{Tr},n}}\cdot\mu^3 - 2\cdot\mu^2 + \Delta\phi^2_{\max}\frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}. \tag{B.10}$$

With equality, it is still necessary to solve a cubic polynomial, but a simpler one.

Minimizing the simplified misadjustment equation leads to

$$0 = \frac{\partial \mathscr{M}}{\partial \mu} \tag{B.11}$$

$$0 = NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} + \frac{1}{\mu^3}\left(-2\right)\frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \tag{B.12}$$

$$= NL_{\mathrm{T}}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2}\cdot\mu^3 - \Delta\phi^2_{\max}\frac{2}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}, \tag{B.13}$$

from which an optimal solution for μ can be obtained. It is given by

$$\mu_{\mathrm{opt,CFO}} = \left(\frac{4\Delta\phi^2_{\max}}{NL_{\mathrm{T}}\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}}. \tag{B.14}$$

It was seen in section 3.4.4 that this solution does not hold for certain parameter settings, even for not so high CFO values when, for example, the system size is large. Assume now that the CFO term dominates the misadjustment, and start with the initial equation

$$\mathscr{M} \approx \frac{1}{\mu^2} \cdot \frac{2-\mu\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu N L_\mathrm{T} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \frac{\Delta\phi^2_\mathrm{max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \tag{B.15}$$

$$\approx \frac{1}{\mu^2} \cdot \frac{2}{2-\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}} \frac{\Delta\phi^2_\mathrm{max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}, \tag{B.16}$$

where it is still assumed that $1 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}$. We will only proceed with the derivative approach, in order to get a sort of upper bound on μ that may always be usable. The derivative leads to

$$0 = \frac{\partial\mathscr{M}}{\partial\mu} \tag{B.17}$$

$$\begin{aligned} 0 = {} & \frac{1}{\mu^3}(-2)\frac{2}{2-\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}} \frac{\Delta\phi^2_\mathrm{max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \\ & + \frac{1}{\mu^2} \frac{\Delta\phi^2_\mathrm{max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \frac{2}{\left(2-\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}\right)^2}(-1)\left(-N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}\right) \end{aligned} \tag{B.18}$$

$$= -4 + \mu \frac{2N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}}{2-\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}} \tag{B.19}$$

$$= -8 + 4\mu N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}} + \mu 2 N L_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}} \tag{B.20}$$

$$\implies \mu = \frac{4}{3}\frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} N L_\mathrm{T}}. \tag{B.21}$$

The misadjustment for the optimal choice of μ, employing the same approximations that were used to derive it, can be written as

$$\mathscr{M}_{\mathrm{opt,CFO}} \approx \left(\frac{4\Delta\phi^2_\mathrm{max}}{N L_\mathrm{T}\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}} \cdot N L_\mathrm{T}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2} + \left(\frac{4\Delta\phi^2_\mathrm{max}}{N L_\mathrm{T}\sigma^4_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}}\right)^{-\frac{2}{3}} \cdot \frac{\Delta\phi^2_\mathrm{max}}{\sigma^2_{x_n}\sigma^2_{\dot{n}_m}} \tag{B.22}$$

$$= \left(\frac{4}{8}\frac{\Delta\phi^2_\mathrm{max} N^2 L^2_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}} + \left(\frac{1}{16}\frac{N^2 L^2_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}\Delta\phi^2_\mathrm{max}}{\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}} \tag{B.23}$$

$$= \left(\frac{1}{2}\frac{\Delta\phi^2_\mathrm{max} N^2 L^2_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}} + \frac{1}{2}\left(\frac{1}{2}\frac{N^2 L^2_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}\Delta\phi^2_\mathrm{max}}{\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}} \tag{B.24}$$

$$= \frac{3}{2}\left(\frac{1}{2}\frac{\Delta\phi^2_\mathrm{max} N^2 L^2_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}}{\sigma^2_{\dot{n}_m}}\right)^{\frac{1}{3}}. \tag{B.25}$$

For the fixed choice of μ, which is independent of CFO and SNR from above, the misadjustment

can be approximated, using $1 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}$, as

$$\begin{aligned}\mathscr{M} \approx & \frac{4}{3}\frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}} \cdot NL_\mathrm{T} \frac{\sigma^2_{x_{\mathrm{Tr},n}}}{2-\left(\frac{4}{3}\frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}}\right)\cdot \sigma^2_{x_{\mathrm{Tr},n}}(2+NL_\mathrm{T})} \\ & + \left(\frac{4}{3}\frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}}\right)^{-2} \cdot \frac{2}{2-\left(\frac{4}{3}\frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}}\right)\cdot NL_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}} \frac{\Delta\phi^2_{\max}}{\sigma^2_{x_{\mathrm{Tr},n}}\sigma^2_{\dot{n}_m}} \end{aligned} \tag{B.26}$$

$$= \frac{2}{1-\frac{4}{NL_\mathrm{T}}} + \frac{2}{\frac{2}{3}}\frac{9}{16}\frac{\sigma^2_{x_{\mathrm{Tr},n}} N^2 L^2_\mathrm{T} \Delta\phi^2_{\max}}{\sigma^2_{\dot{n}_m}}. \tag{B.27}$$

B.2 Tracking of SFOs

Assuming that the variation due to the SFO dominates the misadjustment, it can be computed that

$$\beta = \mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}\right) \approx 1-\alpha^2 \tag{B.28}$$

$$\mathscr{M} \approx \frac{1}{\mu}\frac{1}{2-\mu NL_\mathrm{T}\frac{\sigma^2_{x_{\mathrm{Tr},n}}}{1-\mu\sigma^2_{x_{\mathrm{Tr},n}}}} \frac{\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{\dot{n}_m}} \tag{B.29}$$

$$\approx \frac{1}{\mu}\frac{1}{2-\mu NL_\mathrm{T}\sigma^2_{x_{\mathrm{Tr},n}}} \frac{\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}\right)}{\sigma^2_{\dot{n}_m}}, \tag{B.30}$$

where $1 \gg \mu\sigma^2_{x_{\mathrm{Tr},n}}$ was assumed. Minimizing the misadjustment leads to

$$0 = \frac{\partial\mathscr{M}}{\partial\mu} \tag{B.31}$$

$$0 = -\mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}\right)\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}\cdot\mu + \mathrm{Tr}\left(\mathbf{C}_{\mathbf{\Delta}\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}\right) \tag{B.32}$$

$$\Longrightarrow \mu = \frac{1}{\sigma^2_{x_{\mathrm{Tr},n}} NL_\mathrm{T}}. \tag{B.33}$$

Nomenclature

Basic Constants and Variables

$a_{mn,l}$ Path attenuation of the lth path between transmit antenna n and receive antenna m

β_{T} Roll-off of the discrete-time pulse-shaping filter

c Speed of light $\approx 3 \cdot 10^8 \, \frac{\mathrm{m}}{\mathrm{s}}$

d, $d_{\mathrm{Tx},1}$, $d_{\mathrm{Tx},2}$, $d_{\mathrm{Rx},1}$, $d_{\mathrm{Rx},2}$ Antenna spacing, antenna spacing on the transmit side in the 1st array direction, antenna spacing on the transmit side in the 2nd array direction, antenna spacing on the receive side in the 1st array direction, antenna spacing on the receive side in the 2nd array direction

f_{nom} Nominal (ideal) carrier frequency of the transmission system

$f_{\mathrm{Tx},n}$, $f_{\mathrm{Rx},m}$ Actual carrier frequency of the nth transmitter, actual carrier frequency of the mth receiver

K_{R} Rician K-factor

λ Wavelength of the nominal carrier frequency, i.e., $\lambda = \frac{c}{f_{\mathrm{nom}}}$

$\lambda_{\mathrm{Tx},n}$ Wavelength of the actual carrier frequency of the nth transmitter

M Number of receive antennas or front ends

N Number of transmit antennas or front ends

Q Integer oversampling factor

$\bar{Q}$ Fractional oversampling factor

R Transmission link range

T_s Symbol duration, equivalent to the inverse of the symbol rate

$\tau_{mn,l}$ Path delay of the lth path between transmit antenna n and receive antenna m

T_{nom}, $T_{\mathrm{nom,Tx}}$, $T_{\mathrm{nom,Rx}}$ Nominal (or ideal) sampling time interval, nominal sampling interval at the transmitter, nominal sampling interval at the receiver

$T_{\mathrm{Tx},n}$, $T_{\mathrm{Rx},m}$ Actual sampling interval of the nth transmitter, actual sampling interval of the mth receiver

Operators, Functions, and Distributions

$\mathscr{CN}(a,b)$ Complex Gaussian distribution with mean a and variance b, with uncorrelated real and imaginary part

$\star$ Discrete-time convolution

$\mathrm{E}[x]$ Expected value of random variable x

$\mathscr{F}\{x[l]\}$ Discrete Fourier transform of discrete-time signal $x[l]$

$\mathrm{Im}\{x\}$ Imaginary part of complex variable x

$J_0(x)$ Bessel function of the first kind and zeroth order evaluated at x

$\mathscr{N}(a,b)$ Gaussian distribution with mean a and variance b

$\mathrm{Re}\{x\}$ Real part of complex variable x

$\mathrm{sinc}(x)$ Sinc function of variable x, i.e., $\frac{\sin(x)}{x}$

$\mathscr{U}(a,b)$ Uniform distribution between a and b

Basic Vectors and Matrices, Operators and Functions

$\mathbf{0}_{m\times n}$ All-zeros vector or matrix of size $m \times n$

$\mathbf{1}_{m\times n}$ All-ones vector or matrix of size $m \times n$

$\mathrm{blkdiag}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$ Block diagonal matrix of a sequence of matrices $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$

x^*, $\mathbf{x}^*$, $\mathbf{X}^*$ Complex conjugate (entry-wise) of variable x, vector $\mathbf{x}$, or matrix $\mathbf{X}$

$\det(\mathbf{X})$ Determinant of matrix $\mathbf{X}$

$\mathrm{diag}(\mathbf{x})$ Diagonal matrix of vector $\mathbf{x}$

$\mathbf{e}_n$ nth basis vector

$\odot$ Hadamard (or entry-wise) product of two vectors, or matrices, of equal size

$\mathbf{x}^{\mathrm{H}}$, $\mathbf{X}^{\mathrm{H}}$ Conjugate transpose, or Hermitian transpose, of vector $\mathbf{x}$, or matrix $\mathbf{X}$

$\mathbf{I}_n$ Identity matrix of size $n \times n$

$\kappa(\mathbf{X})$ Condition number of matrix $\mathbf{X}$

$\otimes$ Kronecker product of two vectors, or matrices

$\lambda_i(\mathbf{X})$, $\lambda_{\max}(\mathbf{X})$ ith eigenvalue of matrix $\mathbf{X}$, maximum eigenvalue of matrix $\mathbf{X}$

$\mathrm{Tr}(\mathbf{X})$ Trace of matrix $\mathbf{X}$

$\mathbf{x}^{\mathrm{T}}$, $\mathbf{X}^{\mathrm{T}}$ Transpose of vector $\mathbf{x}$, or matrix $\mathbf{X}$

$\mathrm{vec}(\mathbf{X})$ Vectorization, i.e., column stacking, of matrix $\mathbf{X}$

Continuous-Time Signals, Filters, and Processes

$h_{\mathrm{BP},mn}(t)$, $h_{mn}(t)$ Continuous-time band-pass version of the wireless channel impulse response between transmit antenna n and receive antenna m, continuous-time baseband version of the wireless channel impulse response between transmit antenna n and receive antenna m

$h_{\mathrm{Tx},n}(t)$, $h_{\mathrm{Rx},m}(t)$ Continuous-time baseband linear impulse response characteristic of the nth transmitter front end, continuous-time baseband linear impulse response characteristic of the mth receiver front end

$n_m(t)$ Continuous-time band-pass/baseband noise process observed at the mth receiver

$r_m(t)$ Continuous-time band-pass version of the received signal of antenna m

$s_n(t)$ Continuous-time band-pass version of the transmitted signal of antenna n

$\bar{x}_n(t)$ Continuous-time baseband version of the transmitted signal of antenna n

$\bar{y}_m(t)$ Continuous-time baseband version of the received signal of antenna m

Discrete-Time Signals, Filters, and Processes

$h_{mn}[l]$ Discrete-time baseband version of the wireless channel impulse response between transmit antenna n and receive antenna m

$h_{\mathrm{C},mn}[l_{\mathrm{C}}]$, $h_{\mathrm{C},mn}[k,l_{\mathrm{C}}]$, $\dot{h}_{\mathrm{C},mn}[\dot{l}_{\mathrm{C}}]$, $\dot{h}_{\mathrm{C},mn}[k,\dot{l}_{\mathrm{C}}]$ Discrete-time baseband version of the composite, including linear transmit and receive filter characteristics, channel impulse response between transmit antenna n and receive antenna m, discrete-time baseband version of the composite channel impulse response between transmit antenna n and receive antenna m including the timing impairments, fractionally-spaced version of the former, fractionally-spaced version of the latter

$h_{\bar{\mathrm{C}},mn}[k,l_{\bar{\mathrm{C}}}]$ Extended discrete-time baseband version of the composite, including linear transmit and receive filter characteristics, channel impulse response between transmit antenna n and receive antenna m, including the timing impairments

$h_{\mathrm{P},n}[l_{\mathrm{P}}]$, $h_{\mathrm{P},m}[l_{\mathrm{P}}]$ Discrete-time finite-length pulse-shaping filter of the nth transmitter, discrete-time finite-length pulse-shaping filter of the mth receiver

$h_{\mathrm{Tx},n}[k]$, $h_{\mathrm{Rx},m}[k]$, $h_{\mathrm{Tx},n}[l_{\mathrm{Tx}}]$, $h_{\mathrm{Rx},m}[l_{\mathrm{Rx}}]$ Discrete-time baseband linear impulse response characteristic of the nth transmitter front end, discrete-time baseband linear impulse response characteristic of the mth receiver front end, discrete-time finite-length baseband linear impulse response characteristic of the nth transmitter front end, discrete-time finite-length baseband linear impulse response characteristic of the mth receiver front end

$\mu_{\varphi_{\mathrm{w}}}$, $\sigma^2_{\varphi_{\mathrm{w}}}$, $\sigma^2_{\varphi_{\mathrm{n}}}$ Mean linear drift of the carrier phase process, variance of the Wiener process contained in the carrier phase process, variance of the white noise process contained in the carrier phase process

$\mu_{\dot{\varphi}_{\mathrm{w}}}$, $\sigma^2_{\dot{\varphi}_{\mathrm{w}}}$, $\sigma^2_{\dot{\varphi}_{\mathrm{n}}}$ Mean linear drift of the sampling phase process, variance of the Wiener process contained in the sampling phase process, variance of the white noise process contained in the sampling phase process

$n_m[k]$, $\dot{n}_m[k]$ Symbol-spaced discrete-time baseband noise process sample observed at the mth receiver, fractionally-spaced discrete-time baseband noise process sample observed at the mth receiver

$\tilde{n}_m[k]$ Colored discrete-time baseband noise process sample observed at the mth receiver, including the effect of the linear impulse response characteristic of that receiver

$\phi_{\mathrm{Tx},n}[k]$, $\phi_{\mathrm{Rx},m}[k]$ Sample of the discrete-time carrier phase offset process of the nth transmitter with respect to its nominal carrier frequency f_{nom}, sample of the discrete-time carrier phase offset process of the mth receiver with respect to its nominal carrier frequency f_{nom}

$\Delta\phi_{mn}[k]$ Discrete-time baseband sample of the carrier phase difference process between transmitter n and receiver m, i.e., $\phi_{\mathrm{Tx},n}[k] - \phi_{\mathrm{Rx},n}[k]$

$\dot{\phi}_{\mathrm{Tx},n}[k]$, $\dot{\phi}_{\mathrm{Rx},m}[k]$ Sample of the discrete-time sampling phase offset process of the nth transmitter with respect to its nominal sampling interval $T_{\mathrm{nom,Tx}}$, sample of the discrete-time sampling phase offset process of the mth receiver with respect to its nominal sampling interval $T_{\mathrm{nom,Rx}}$

$\Delta\dot{\phi}_{mn}[k]$ Discrete-time baseband sample of the sampling phase difference process between transmitter n and receiver m, i.e., $\dot{\phi}_{\mathrm{Tx},n}[k] - \dot{\phi}_{\mathrm{Rx},n}[k]$

$\sigma^2_{n_m}$, $\sigma^2_{\dot{n}_m}$ Variance of the symbol-spaced noise process of the mth receiver, variance of the fractionally-spaced noise process of the mth receiver

$\sigma^2_{x_n}$ Variance (or average transmit power per symbol) of the signal of antenna n

$\sigma^2_{x_{\mathrm{Tr},n}}$ Variance (or average transmit power per symbol) of the training signal of antenna n

$\sigma^2_{y_m}$ Variance (or average received power per sample) of the received signal of antenna m

$x_n[k]$, $\dot{x}_n[k]$ Discrete-time baseband symbol of the transmitted signal of antenna n, discrete-time baseband sample of the transmitted signal of antenna n

$x_{\mathrm{Tr},n}[l_{\mathrm{Tr}}]$, $\dot{x}_{\mathrm{Tr},n}[\dot{l}_{\mathrm{Tr}}]$ Discrete-time baseband symbol of the transmitted finite-length training signal of antenna n, discrete-time baseband sample of the transmitted finite-length training signal of antenna n

$y_m[k]$, $\dot{y}_m[k]$ Symbol-spaced discrete-time baseband sample of the received signal of antenna m, fractionally-spaced discrete-time baseband sample of the received signal of antenna m

Discrete-Time Vector and Matrix Signals, Filters, and Processes

$\dot{\mathbf{e}}_{\dot{L}_{\mathrm{R}}}[k]$, $\mathbf{e}[k]$ Vector containing the current and previous $\dot{L}_{\mathrm{R}}-1$ error signal values from all antennas used for adaptive filtering, vector containing the current error signal value from all antennas used for adaptive equalization

$\mathbf{h}_m[l]$, $\mathbf{h}_{L,m}$ Vector containing the lth tap of the wireless channel impulse responses from all transmit antennas to receive antenna m, vector containing all L taps of the wireless channel impulse responses from all transmit antennas to receive antenna m

$\mathbf{H}$, $\mathbf{H}[k]$, $\mathbf{H}_{L_{\mathrm{Rx}}}$, $\mathbf{H}_{L_{\mathrm{Rx}}}[k]$, $\dot{\mathbf{H}}_{\dot{L}_{\mathrm{Rx}}}$, $\dot{\mathbf{H}}_{\dot{L}_{\mathrm{Rx}}}[k]$ Matrix containing all L taps of the time-invariant wireless channel impulse responses from all transmit antennas to all receive antenna, matrix containing all L taps of the time-varying wireless channel impulse responses from all transmit antennas to all receive antenna, extended matrix containing all symbol-spaced time-invariant wireless channel impulse responses, extended symbol-spaced time-varying matrix containing all wireless channel impulse responses, extended matrix containing the complete fractionally-spaced time-invariant wireless channel impulse response, extended matrix containing all fractionally-spaced time-varying wireless channel impulse responses

$\mathbf{h}_{\mathrm{C}}$, $\mathbf{h}_{\mathrm{C}}[k]$, $\dot{\mathbf{h}}_{\mathrm{C}}$, $\dot{\mathbf{h}}_{\mathrm{C}}[k]$ Vector containing all taps of the composite, including linear transmit and receive filter characteristics, symbol-spaced time-invariant channel impulse responses between all transmit and all receive antennas, vector containing all taps of the composite symbol-spaced time-varying channel impulse responses between all transmit and all receive antennas, vector containing all taps of the composite fractionally-spaced time-invariant channel impulse responses between all transmit and all receive antennas, vector containing all taps of the composite fractionally-spaced time-varying channel impulse responses between all transmit and all receive antennas

$\mathbf{H}_{\mathrm{C}}$, $\dot{\mathbf{H}}_{\mathrm{C}}[k]$ Matrix containing all taps of the composite, including linear transmit and receive filter characteristics, symbol-spaced time-invariant channel impulse responses between all transmit and all receive antennas, matrix containing all taps of the composite fractionally-spaced time-varying channel impulse responses between all transmit and all receive antennas

$\mathbf{H}_{\mathrm{C}}[k, l_{\mathrm{C}}]$ Matrix containing the l_{C}th tap of the composite, including linear transmit and receive filter characteristics, time-invariant channel impulse responses between all transmit and all receive antennas, in the thesis $l_{\mathrm{C}} = 0$ is often used to define the LoS tap

$\mathbf{H}_{\tilde{\mathrm{C}}}$, $\mathbf{H}_{\tilde{\mathrm{C}}}[k]$ Extended matrix containing all taps of the composite, including linear transmit and receive filter characteristics, time-invariant channel impulse responses between all transmit and all receive antennas, extended matrix containing all taps of the composite time-varying channel impulse responses between all transmit and all receive antennas

$\mathbf{H}_{\mathrm{P}}$ Matrix containing pulse-shaping filters

$\mathbf{H}_{\mathrm{Rx}}$, $\mathbf{H}_{\mathrm{Rx}}[k]$, $\dot{\mathbf{H}}_{\mathrm{Rx}}$, $\dot{\mathbf{H}}_{\mathrm{Rx}}[k]$ Matrix containing all taps of the time-invariant symbol-spaced linear receive filter characteristics from all receivers, matrix containing all taps of the time-varying symbol-spaced linear receive filter characteristics from all receivers, matrix containing all taps of the time-invariant fractionally-spaced linear receive filter characteristics from all receivers, matrix containing all taps of the time-varying fractionally-spaced linear receive filter characteristics from all receivers

$\mathbf{H}_{\tilde{\mathrm{C}},\mathrm{Rx}}[k]$ Extended matrix containing all taps of the time-varying linear receive filter characteristics from all receivers

$\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}$, $\mathbf{H}_{\mathrm{Tx},L_{\mathrm{D}}}[k]$, $\dot{\mathbf{H}}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}$, $\dot{\mathbf{H}}_{\mathrm{Tx},\dot{L}_{\mathrm{D}}}[k]$ Extended matrix containing all taps of the time-invariant symbol-spaced linear transmit filter characteristics from all transmitters, extended matrix containing all taps of the time-varying symbol-spaced linear transmit filter characteristics from all transmitters, extended matrix containing all taps of the time-invariant fractionally-spaced linear transmit filter characteristics from all transmitters, extended matrix containing all taps of the time-varying fractionally-spaced linear transmit filter characteristics from all transmitters

$\mathbf{n}[k]$, $\mathbf{n}_{L_{\mathrm{Rx}}}[k]$, $\dot{\mathbf{n}}_{\dot{L}_{\mathrm{Rx}}}[k]$ Vector containing the current symbol-spaced noise sample from all receive antennas, vector containing the current and $L_{\mathrm{Rx}} - 1$ previous symbol-spaced noise samples from all receive antennas, vector containing the current and $\dot{L}_{\mathrm{Rx}} - Q$ previous fractionally-spaced noise samples from all receive antennas

$\tilde{\mathbf{n}}[k]$, $\tilde{\mathbf{n}}_{\dot{L}_{\mathrm{R}}}[k]$ Vector containing the current colored noise sample from all receive antennas, vector containing the current and $\dot{L}_{\mathrm{R}} - Q$ previous fractionally-spaced colored noise samples from all receive antennas

$\boldsymbol{\phi}_{\mathrm{Tx}}[k]$, $\boldsymbol{\phi}_{\mathrm{Rx}}[k]$ Vector containing the current carrier phase offset process sample from all transmit antennas, vector containing the current carrier phase offset process sample from all receive antennas

$\boldsymbol{\Phi}_{\mathrm{Tx}}[k]$, $\boldsymbol{\Phi}_{\mathrm{Rx}}[k]$, $\boldsymbol{\Phi}_{\mathrm{Tx},L}[k]$, $\boldsymbol{\Phi}_{\mathrm{Tx},L_{\mathrm{D}}}[k]$, $\boldsymbol{\Phi}_{\mathrm{Rx},L_{\mathrm{Rx}}}[k]$ Diagonal matrix containing the current carrier phase offset process sample from all transmitters, diagonal matrix containing the current carrier phase offset process sample from all receivers, diagonal matrix containing the current and $L-1$ previous carrier phase offset process samples from all transmitters,

diagonal matrix containing the current and $L_{\mathrm{D}}-1$ previous carrier phase offset process samples from all transmitters, diagonal matrix containing the current and $L_{\mathrm{Rx}}-1$ previous carrier phase offset process samples from all receivers

$\mathbf{\Delta\Phi}[k]$, $\mathbf{\Delta\dot{\Phi}}[k]$ Matrix containing the carrier phase offset process difference samples between all transmit and all receive antennas, matrix containing the sampling phase offset process difference samples between all transmit and all receive antennas

$\mathbf{W}_{\tilde{\mathrm{C}}}[k]$ Matrix containing the time-varying equalizer coefficients for the complete system

$\mathbf{W}_{\mathrm{F}}$, $\mathbf{W}_{\mathrm{B}}$, $\mathbf{W}_{\mathrm{F}}[k]$, $\mathbf{W}_{\mathrm{B}}[k]$ Matrix containing the time-invariant symbol-spaced feedforward equalizer coefficients of the complete system, matrix containing the time-invariant symbol-spaced feedback equalizer coefficients of the complete system, matrix containing the time-varying fractionally-spaced feedforward equalizer coefficients of the complete system, matrix containing the time-varying fractionally-spaced feedback equalizer coefficients of the complete system

$\mathbf{x}[k]$, $\mathbf{x}_L[k]$, $\mathbf{x}_{L_{\mathrm{C}}}[k]$, $\mathbf{x}_{L_{\mathrm{P}}}[k]$, $\mathbf{x}_{L_{\mathrm{T}}}[k]$, $\dot{\mathbf{x}}_{\dot{L}_{\mathrm{C}}}[k]$ Vector containing the current transmitted symbol of all transmit antennas, vector containing the current and $L-1$ previous transmitted symbols of all transmit antennas, vector containing the current and $L_{\mathrm{C}}-1$ previous transmitted symbols of all transmit antennas, vector containing the current and $L_{\mathrm{P}}-1$ previous transmitted symbols of all transmit antennas, vector containing the current and $L_{\mathrm{T}}-1$ previous transmitted symbols of all transmit antennas, vector containing the current and $\dot{L}_{\mathrm{C}}-Q$ previous transmitted samples of all transmit antennas

$\mathbf{x}_{\mathrm{Tr}}[l_{\mathrm{Tr}}]$, $\mathbf{x}_{\mathrm{Tr}}$, $\dot{\mathbf{x}}_{\mathrm{Tr}}$ Vector containing the l_{Tr}th transmitted training symbol of all transmit antennas, vector containing all transmitted training symbols of all transmit antennas, vector containing all transmitted training samples of all transmit antennas

$\mathbf{X}_{\mathrm{Tr}}$, $\dot{\mathbf{X}}_{\mathrm{Tr}}$ Toeplitz matrix containing all transmitted training symbols of all transmit antennas, Toeplitz matrix containing all transmitted training samples of all transmit antennas

$\mathbf{x}_{\mathrm{Tr},L_{\mathrm{T}}}[k]$, $\dot{\mathbf{x}}_{\mathrm{Tr},\dot{L}_{\mathrm{T}}}[k]$ Vector containing the current and $L_{\mathrm{T}}-1$ previous transmitted training symbols of all transmit antennas, vector containing the current and $L_{\mathrm{T}}-1$ previous transmitted training samples of all transmit antennas

$\mathbf{y}[k]$, $\mathbf{y}_{L_{\mathrm{R}}}[k]$, $\mathbf{y}_{L_{\mathrm{F}}}[k]$ Vector containing the current received symbol of all receive antennas, vector containing the current and $L_{\mathrm{R}}-1$ previous received symbols of all receive antennas, vector containing the current and $L_{\mathrm{F}}-1$ previous received symbols of all receive antennas

$\dot{\mathbf{y}}[k]$, $\dot{\mathbf{y}}_{\dot{L}_{\mathrm{R}}}[k]$, $\dot{\mathbf{y}}_{\dot{L}_{\mathrm{C}}}[k]$, $\dot{\mathbf{y}}_{\dot{L}_{\mathrm{F}}}[k]$, $\dot{\mathbf{y}}_{\tilde{\mathrm{C}},\dot{L}_{\mathrm{F,B}}}[k]$ Vector containing the current received sample of all receive antennas, vector containing the current and $\dot{L}_{\mathrm{R}}-Q$ previous received samples of all receive antennas, vector containing the current and $\dot{L}_{\mathrm{C}}-Q$ previous received samples of all receive antennas, vector containing the current and $\dot{L}_{\mathrm{F}}-Q$ previous received samples of all receive antennas, vector containing the samples and symbols for decision-feedback equalization

Lengths of Discrete-Time Finite-Length Signals and Filters

L, $\dot{L}$ Length of the wireless channel impulse response in symbols, length of the wireless channel impulse response in samples

L_{C}, $\dot{L}_{\mathrm{C}}$ Length of the composite, including linear transmit and receive filter characteristics, channel impulse response in symbols, length of the composite channel impulse response in samples

$\dot{L}_{\tilde{\mathrm{C}},\mathrm{Rx}}$ Length of the extended receive filter channel impulse response in samples

L_{D}, $\dot{L}_{\mathrm{D}}$ Length of the combined channel impulse response of wireless channel and receive filter characteristic in symbols, length of the combined channel impulse response of wireless channel and receive filter characteristic in samples

L_{E}, $L_{\mathrm{E},mn}$, $\dot{L}_{\mathrm{E}}$ Number of symbols between two channel estimates, number of symbols after which the channel estimate between transmit antenna n and receive antenna m has been shifted by one symbol, number of samples between two channel estimates

L_{F}, $\dot{L}_{\mathrm{F}}$, L_{B} Length of the feedforward part of the equalizer in symbols per antenna, length of the feedforward part of the equalizer in samples per antenna, length of the feedback part of the equalizer in symbols per antenna

L_{I} Number of samples over which the channel is approximately time invariant

L_{P} Length of the pulse-shaping filter

L_{T}, $\dot{L}_{\mathrm{R}}$ Length of a block of transmit symbols per antenna, length of a block of receive samples per antenna

$L_{\mathrm{Tx},n}$, $L_{\mathrm{Rx},m}$, $\dot{L}_{\mathrm{Tx},n}$, $\dot{L}_{\mathrm{Rx},m}$ Length of the impulse response characteristic of the nth transmitter front end in symbols, length of the impulse response characteristic of the mth receiver front end in symbols, length of the impulse response characteristic of the nth transmitter front end in samples, length of the impulse response characteristic of the mth receiver front end in samples

L_{Tx}, L_{Rx}, $\dot{L}_{\mathrm{Tx}}$, $\dot{L}_{\mathrm{Rx}}$ Maximum length of the impulse response characteristic among all transmitter front ends in symbols, maximum length of the impulse response characteristic among all receiver front ends in symbols, maximum length of the impulse response characteristic among all transmitter front ends in samples, maximum length of the impulse response characteristic among all receiver front ends in samples

L_{Tr} Length of the training signal in symbols

Adaptive Filter Quantities

$\mathbf{C}_{\tilde{\mathbf{n}}}[k]$ Time-varying correlation matrix of the colored noise processes, between all antennas, of length $\dot{L}_{\mathrm{R}}$ per antenna

$\mathbf{C}_{\Delta\dot{\boldsymbol{\phi}}_{\tilde{\mathrm{C}},m}}$ Correlation matrix of the sampling phase variation model of the mth antenna

$\mathbf{C}_{\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]$ Time-varying correlation matrix of the training signals, between all antennas, of length L_T per antenna

$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}}[k]$ Time-varying correlation matrix of the received samples, between all antennas, of length $\dot{L}_\mathrm{R}$ per antenna

$\mathbf{C}_{\dot{\mathbf{y}}_{\dot{L}_\mathrm{R}}\mathbf{x}_{\mathrm{Tr},L_\mathrm{T}}}[k]$ Time-varying crosscorrelation matrix of the received samples, between all transmit antennas and receive antennas, of length $\dot{L}_\mathrm{R}$ per antenna with the transmitted training symbols of length L_T per antenna

$\mathscr{M}$, $\mathscr{M}_{\mathrm{opt,CFO}}$, $\mathscr{M}_{\mathrm{opt,SFO}}$ Misadjustment of the adaptive filter, misadjustment of the adaptive filter with optimal step size for the CFO-only case, misadjustment of the adaptive filter with optimal step size for the SFO-only case

μ, μ_1, μ_2, $\mu_{\mathrm{opt,CFO}}$, $\mu_{\mathrm{opt,SFO}}$ Step size of the adaptive filter, step size option one when no information is available, step size option two when no information is available, optimal step size in the CFO-only case, optimal step size in the SFO-only case

$\Delta\phi_{\max}$ Maximum carrier phase variation per sample

$\Delta\dot{\phi}_{\max}$, $\Delta\dot{\phi}_{\mathrm{max,Tx}}$, $\Delta\dot{\phi}_{\mathrm{max,Rx}}$ Maximum sampling phase variation per sample, maximum sampling phase variation per sample on the transmitter side, maximum sampling phase variation per sample on the receiver side

τ Convergence time of the adaptive filter

Abbreviations

ADC analog-to-digital converter
AWG arbitrary waveform generator

CDF cumulative distribution function
CFO carrier frequency offset
CRB Cramér-Rao bound

DA data-aided
DAC digital-to-analog converter
DD decision-directed
DFE decision-feedback equalizer
DFT discrete Fourier transform

ENOB effective number of bits

FE front end

i.i.d. independent and identically distributed
i.u.d. independent and uniformly distributed
ISI intersymbol interference

LMS least-mean-squares
LO local oscillator
LoS line-of-sight
LS least-squares

MAP maximum-a-posteriori
MIMO multiple-input and multiple-output
ML maximum-likelihood
MMSE minimum-mean-square-error
mmWave millimeter-wave
MSE mean squared error

NDA non-data-aided
NLoS non-line-of-sight

PDF probability density function
PLL phase-locked loop
PMF probability mass function
PN phase noise
ppm parts-per-million

RMS root-mean-square
RTO real-time oscilloscope
Rx receiver

SDoF spatial degrees of freedom
SER symbol error rate
SFO sampling frequency offset
SISO single-input and single-output
SNR signal-to-noise ratio

Tx transmitter
TxRx-ST transmitter-receiver space-time

w.r.t. with respect to
WL widely-linear

List of Figures

List of Tables

References

[1] S. H. Strogatz, *Sync: How Order Emerges from Chaos in the Universe, Nature, and Daily Life*. Hachette Books, 2015.

[2] J. F. Cliche and B. Shillue, "Precision Timing Control for Radioastronomy: Maintaining Femtosecond Synchronization in the Atacama Large Millimeter Array," *IEEE Control Syst. Mag.*, vol. 26, no. 1, pp. 19–26, 2006.

[3] U. Mengli and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.

[4] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. John Wiley & Sons, 1998.

[5] A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, and R. A. Kennedy, "Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years," *EURASIP J. Wirel. Commun. Netw.*, pp. 1–38, 2016.

[6] S. Chia, M. Gasparroni, and P. Brick, "The Next Challenge for Cellular Networks: Backhaul," *IEEE Microw. Mag.*, vol. 10, no. 5, pp. 54–66, 2009.

[7] J. Antes and I. Kallfass, "Performance Estimation for Broadband Multi-Gigabit Millimeter- and Sub-Millimeter-Wave Wireless Communication Links," *IEEE Trans. Microw. Theory Tech.*, vol. 63, no. 10, pp. 3288–3299, 2015.

[8] D. Cvetkovski, T. Hälsig, B. Lankl, and E. Grass, "Next Generation mm-Wave Wireless Backhaul Based on LOS MIMO Links," in *Proc. Ger. Microw. Conf.*, 2016, pp. 69–72.

[9] R. Wu, R. Minami, Y. Tsukui, S. Kawai, Y. Seo, S. Sato, K. Kimura, S. Kondo, T. Ueno, N. Fajri, S. Maki, N. Nagashima, Y. Takeuchi, T. Yamaguchi, A. Musa, K. K. Tokgoz, T. Siriburanon, B. Liu, Y. Wang, J. Pang, N. Li, M. Miyahara, K. Okada, and A. Matsuzawa, "64-QAM 60-GHz CMOS Transceivers for IEEE 802.11ad/ay," *IEEE J. Solid-State Circuits*, vol. 52, no. 11, pp. 2871–2891, 2017.

[10] R. G. Fellers, "Millimeter Waves and Their Applications," *Electr. Eng.*, vol. 75, no. 10, pp. 914–917, 1956.

[11] D. T. Emerson, "The Work of Jagadis Chandra Bose: 100 Years of Millimeter-Wave Research," *IEEE Trans. Microw. Theory Tech.*, vol. 45, no. 12, pp. 2267–2273, 1997.

[12] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson Education, 2014.

[13] X. Song, D. Cvetkovski, T. Hälsig, W. Rave, G. Fettweis, E. Grass, and B. Lankl, "High Throughput Line-of-Sight MIMO Systems for Next Generation Backhaul Applications," *Frequenz*, vol. 71, no. 9-10, pp. 389–398, 2017.

[14] P. F. Driessen and G. J. Foschini, "On the Capacity Formula for Multiple Input-Multiple Output Wireless Channels: A Geometric Interpretation," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, 1999.

[15] A. A. Hutter, F. Platbrood, and J. Ayadi, "Analysis of MIMO Capacity Gains for Indoor Propagation Channels with LOS Component," in *Proc. IEEE 13th Int. Symp. Pers. Indoor Mob. Radio Commun.*, 2002, pp. 1337–1341.

[16] S. Calabrò, B. Lankl, and G. Sebald, "Multiple Co-Polar Co-Channel Point-to-Point Radio Transmission," *Int. J. Electron. Commun.*, vol. 57, no. 5, pp. 1–7, 2003.

[17] C. Hofmann, K.-U. Storek, R. T. Schwarz, and A. Knopp, "Spatial MIMO over Satellite: A Proof of Concept," in *Proc. IEEE Int. Conf. Commun.*, 2016.

[18] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 4, pp. 1815–1831, 2014.

[19] Z. Wang, S. Zhou, and Z. Wang, "Underwater Distributed Antenna Systems: Design Opportunities and Challenges," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 179–185, 2018.

[20] C. Fulton, M. Yeary, D. Thompson, J. Lake, and A. Mitchell, "Digital Phased Arrays: Challenges and Opportunities," *Proc. IEEE*, vol. 104, no. 3, pp. 487–503, 2016.

[21] F.-L. Luo and C. Zhang, *Signal Processing for 5G: Algorithms and Implementations*. John Wiley & Sons, 2016.

[22] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The World's First Real-Time Testbed for Massive MIMO: Design, Implementation, and Validation," *IEEE Access*, vol. 5, pp. 9073–9088, 2017.

[23] B. Sadhu, X. Gu, and A. Valdes-Garcia, "The More (Antennas), the Merrier: A Survey of Silicon-Based mm-Wave Phased Arrays Using Multi-IC Scaling," *IEEE Microw. Mag.*, vol. 20, pp. 32–50, 2019.

[24] S. U. H. Qureshi, "Adaptive Equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, 1985.

[25] F. M. Gardner, "Interpolation in Digital Modems-Part I: Fundamentals," *IEEE Trans. Commun.*, vol. 41, no. 3, pp. 501–507, 1993.

[26] J. Treichler, I. Fijalkow, and C. Johnson, "Fractionally Spaced Equalizers," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 65–81, 1996.

[27] A. F. Molisch, *Wireless Communications*. John Wiley & Sons, 2011.

[28] J. G. Proakis and M. Salehi, *Digital Communications*. McGraw–Hill, 2008.

[29] T. Oguchi, "Electromagnetic Wave Propagation and Scattering in Rain and Other Hydrometeors," *Proc. IEEE*, vol. 71, no. 9, pp. 1029–1078, 1983.

[30] H. J. Liebe, T. Manabe, and G. A. Hufford, "Millimeter-Wave Attenuation and Delay Rates Due to Fog/Cloud Conditions," *IEEE Trans. Antennas Propag.*, vol. 37, no. 12, pp. 1617–1623, 1989.

[31] D. A. De Wolf and A. J. Zwiesler, "Rayleigh-Mie Approximation for Line-of-Sight Propagation Through Rain at 5-90 GHz," *IEEE Trans. Antennas Propag.*, vol. 44, no. 3, pp. 273–279, 1996.

[32] T. Hälsig and B. Lankl, "Array Size Reduction for High-Rank LOS MIMO ULAs," *IEEE Wirel. Commun. Lett.*, vol. 4, no. 6, pp. 649–652, 2015.

[33] F. Bøhagen, P. Orten, and G. E. Øien, "On Spherical vs. Plane Wave Modeling of Line-of-Sight MIMO Channels," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 841–849, 2009.

[34] E. Torkildson, U. Madhow, and M. Rodwell, "Indoor Millimeter Wave MIMO: Feasibility and Performance," *IEEE Trans. Wirel. Commun.*, vol. 10, no. 12, pp. 4150–4160, 2011.

[35] B. Mamandipoor, M. Sawaby, A. Arbabian, and U. Madhow, "Hardware-Constrained Signal Processing for mm-Wave LoS MIMO," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, 2015, pp. 1427–1431.

[36] F. Bøhagen, P. Orten, and G. E. Øien, "Construction and Capacity Analysis of High-Rank Line-of-Sight MIMO Channels," in *Proc. IEEE Wirel. Commun. Netw. Conf.*, 2005, pp. 432 – 437.

[37] P. Larsson, "Lattice Array Receiver and Sender for Spatially OrthoNormal MIMO Communication," in *Proc. IEEE 61st Veh. Technol. Conf.*, 2005, pp. 192–196.

[38] F. Bøhagen, P. Orten, and G. E. Øien, "Optimal Design of Uniform Planar Antenna Arrays for Strong Line-of-Sight MIMO channels," in *Proc. IEEE 7th Work. Signal Process. Adv. Wirel. Commun.*, 2006.

[39] E. Torkildson, C. Sheldon, U. Madhow, and M. Rodwell, "Nonuniform Array Design for Robust Millimeter-Wave MIMO Links," in *Proc. IEEE Glob. Telecommun. Conf.*, 2009.

[40] M. Matthaiou, A. M. Sayeed, and J. A. Nossek, "Maximizing LoS MIMO Capacity Using Reconfigurable Antenna Arrays," in *Proc. Int. ITG Work. Smart Antennas*, 2010, pp. 14–19.

[41] L. Zhou and Y. Ohashi, "Design of Non-uniform Antenna Arrays for Robust Millimeter-Wave LOS MIMO Communications," in *Proc. IEEE 24th Int. Symp. Pers. Indoor Mob. Radio Commun.*, 2013, pp. 1397–1401.

[42] X. Song and G. Fettweis, "On Spatial Multiplexing of Strong Line-of-Sight MIMO with 3D Antenna Arrangements," *IEEE Wirel. Commun. Lett.*, vol. 4, no. 4, pp. 393–396, 2015.

[43] M. Mishali and Y. C. Eldar, "Sub-Nyquist Sampling," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, 2011.

[44] N. A. Preyss, "Modulation, Coding, and Receiver Design for Gigabit mmWave Communication," Ph.D. dissertation, EPFL, 2016.

[45] A. Kipnis, Y. C. Eldar, and A. J. Goldsmith, "Analog-to-Digital Compression: A New Paradigm for Converting Signals to Bits," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 16–39, 2018.

[46] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, E. T. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. F. Vessot, and G. M. Winkler, "Characterization of Frequency Stability," *IEEE Trans. Instrum. Meas.*, vol. IM-20, no. 2, pp. 105–120, 1971.

[47] B. Razavi, *RF Microelectronics*. Pearson Education, 2011.

[48] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase Noise in Oscillators: A Unifying Theory and Numerical Methods for Characterization," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 47, no. 5, pp. 655–674, 2000.

[49] M. R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, "Calculation of the Performance of Communication Systems From Measured Oscillator Phase Noise," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 61, no. 5, pp. 1553–1565, 2014.

[50] J. Chen, D. Kuylenstierna, S. E. Gunnarsson, Z. S. He, T. Eriksson, T. Swahn, and H. Zirath, "Influence of White LO Noise on Wideband Communication," *IEEE Trans. Microw. Theory Tech.*, vol. 66, no. 7, pp. 3349–3359, 2018.

[51] I. Galton and C. Weltin-Wu, "Understanding Phase Error and Jitter: Definitions, Implications, Simulations, and Measurement," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 66, no. 1, pp. 1–19, 2019.

[52] D. Cvetkovski, E. Grass, T. Hälsig, and B. Lankl, "Hardware-in-the-Loop Demonstration of a 60GHz Line-of-Sight 2x2 MIMO Link," in *Proc. IEEE EUROCON*, 2017, pp. 631–636.

[53] T. Hälsig, D. Cvetkovski, E. Grass, and B. Lankl, "Statistical Properties and Variations of LOS MIMO Channels at Millimeter Wave Frequencies," in *Proc. 22nd Int. ITG Work. Smart Antennas*, 2018.

[54] M. R. Khanzadi, "Phase Noise in Communication Systems – Modeling, Compensation, and Performance Analysis," Ph.D. dissertation, Chalmers University of Technology, 2015.

[55] T. Hälsig and B. Lankl, "Performance Evaluation of LOS MIMO Systems under the Influence of Phase Noise," in *Proc. 19th Int. ITG Work. Smart Antennas*, 2015.

[56] G. Fettweis, M. Lohning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: A New Paradigm," in *Proc. IEEE 16th Int. Symp. Pers. Indoor Mob. Radio Commun.*, 2005, pp. 2347–2355.

[57] C. Tröster-Schmid and T. Bednorz, *Generating Multiple Phase Coherent Signals – Aligned in Phase and Time*, Rohde & Schwarz, 2016, Application Note.

[58] G. LaCaille, A. Puglielli, E. Alon, B. Nikolic, and A. Niknejad, "Optimizing the LO Distribution Architecture of mm-Wave Massive MIMO Receivers," University of California at Berkeley, Tech. Rep., 2019, arXiv.

[59] A. Puglielli, A. Townley, G. LaCaille, V. Milovanovic, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade, E. Alon, B. Nikolic, and A. M. Niknejad, "Design of Energy- and Cost-Efficient Massive MIMO Arrays," *Proc. IEEE*, vol. 104, no. 3, pp. 586–606, 2016.

[60] U. Madhow, D. R. Brown III, S. Dasgupta, and R. Mudumbai, "Distributed massive MIMO: algorithms, architectures and concept systems," in *Proc. Inf. Theory Appl. Work.*, 2014.

[61] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A Space-Time Coding Modem for High-Data-Rate Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1459–1477, 1998.

[62] Y.-C. Wu, S. C. Chan, and E. Serpedin, "Symbol-Timing Estimation in Space-Time Coding Systems Based on Orthogonal Training Sequences," *IEEE Trans. Wirel. Commun.*, vol. 4, no. 2, pp. 603–613, 2005.

[63] O. Besson and P. Stoica, "On Parameter Estimation of MIMO Flat-Fading Channels With Frequency Offsets," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 602–613, 2003.

[64] Y. Yao and T.-S. Ng, "Correlation-Based Frequency offset Estimation in MIMO system," in *Proc. IEEE 58th Veh. Technol. Conf.*, 2003, pp. 438–442.

[65] D. Qu, G. Zhu, and T. Jiang, "Training Sequence Design and Parameter Estimation of MIMO Channels with Carrier Frequency Offsets," *IEEE Trans. Wirel. Commun.*, vol. 5, no. 12, pp. 3662–3666, 2006.

[66] M. Ghogho and A. Swami, "Training Design for Multipath Channel and Frequency-Offset Estimation in MIMO Systems," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3957–3965, 2006.

[67] S. Ahmed, S. Lambotharan, A. Jakobsson, and J. A. Chambers, "MIMO frequency-selective channels with multiple frequency offsets: estimation and detection techniques," *IEE Proc. - Commun.*, vol. 152, no. 4, pp. 489–494, 2006.

[68] J. Zhang, Y. R. Zheng, C. Xiao, and K. B. Letaief, "Channel Equalization and Symbol Detection for Single-Carrier MIMO Systems in the Presence of Multiple Carrier Frequency Offsets," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 2021–2030, 2010.

[69] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannidis, and T. Svensson, "Joint Estimation of Channel and Oscillator Phase Noise in MIMO Systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, 2012.

[70] H. V. Cheng and E. G. Larsson, "Some Fundamental Limits on Frequency Synchronization in Massive MIMO," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2013, pp. 1213–1217.

[71] A. Gesell, J. B. Huber, B. Lankl, and G. Sebald, "Data-Aided Symbol Timing Estimation for Linear Modulation," *Int. J. Electron. Commun.*, vol. 56, no. 5, pp. 303–311, 2002.

[72] K. Rajawat and A. K. Chaturvedi, "A Low Complexity Symbol Timing Estimator for MIMO Systems Using Two Samples Per Symbol," *IEEE Commun. Lett.*, vol. 10, no. 7, pp. 525–527, 2006.

[73] A. A. Nasir, H. Mehrpouyan, S. Durrani, S. D. Blostein, R. A. Kennedy, and B. Ottersten, "Optimal Training Sequences for Joint Timing Synchronization and Channel Estimation in Distributed Communication Networks," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3002–3015, 2013.

[74] A. Kannan, T. P. Krauss, and M. D. Zoltowski, "Separation of Cochannel Signals Under Imperfect Timing and Carrier Synchronization," *IEEE Trans. Veh. Technol.*, vol. 50, no. 1, pp. 79–96, 2001.

[75] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-Input Multi-Output Fading Channel Tracking and Equalization Using Kalman Estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, 2002.

[76] A. H. Sayed, *Adaptive Filters*. John Wiley & Sons, 2011.

[77] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-Based Channel Estimation for Multiple-Antenna Broadband Transmissions," *IEEE Trans. Wirel. Commun.*, vol. 2, no. 2, pp. 384–391, 2003.

[78] L. Tong, B. M. Sadler, and M. Dong, "Pilot-Assisted Wireless Transmissions - General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12–25, 2004.

[79] X. Ma, L. Yang, and G. B. Giannakis, "Optimal Training for MIMO Frequency-Selective Fading Channels," *IEEE Trans. Wirel. Commun.*, vol. 4, no. 2, pp. 453–466, 2005.

[80] M. Coldrey and P. Bohlin, "Training-Based MIMO Systems-Part I: Performance Comparison," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5464–5476, 2007.

[81] H. D. Lüke, *Korrelationssignale*. Springer, 1992.

[82] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe, "Comparison of EM-Based Algorithms for MIMO Channel Estimation," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 216–226, 2007.

[83] S. Rangan and A. K. Fletcher, "Iterative Estimation of Constrained Rank-One Matrices in Noise," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 1246–1250.

[84] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 1996.

[85] R. H. Kwong and E. W. Johnston, "A Variable Step Size LMS Algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1633–1642, 1992.

[86] T. Aboulnasr and K. Mayyas, "A Robust Variable Step-Size LMS-Type Algorithm: Analysis and Simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 631–639, 1997.

[87] M. Hajivandi and W. A. Gardner, "Measures of Tracking Performance for the LMS Algorithm," *IEEE Trans. Acoust.*, vol. 38, no. 11, pp. 1953–1958, 1990.

[88] N. R. Yousef and A. H. Sayed, "Ability of Adaptive Filters to Track Carrier Offsets and Channel Nonstationarities," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1533–1544, 2002.

[89] M. Rupp, "LMS tracking behavior under periodically changing systems," in *Proc. 9th Eur. Signal Process. Conf.*, 1998.

[90] R. D. Gitlin and S. B. Weinstein, "Fractionally-Spaced Equalization: An Improved Digital Transversal Equalizer," *Bell Syst. Tech. J.*, vol. 60, no. 2, pp. 275–296, 1981.

[91] Z. Muhi-Eldeen, L. Ivrissimtzis, and M. Al-Nuaimi, "Modelling and measurements of millimetre wavelength propagation in urban environments," *IET Microwaves, Antennas Propag.*, vol. 4, no. 9, pp. 1300–1309, 2010.

[92] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, 2015.

[93] V. Raghavan, A. Partyka, L. Akhoondzadeh-Asl, M. A. Tassoudji, O. H. Koymen, and J. Sanelli, "Millimeter Wave Channel Measurements and Implications for PHY Layer Design," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6521–6533, 2017.

[94] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks-With a Focus on Propagation Models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, 2017.

[95] J. Ding, V. Tarokh, and Y. Yang, "Model Selection Techniques – An Overview," *IEEE Signal Process. Mag.*, vol. 35, pp. 16–34, 2018.

[96] T. Hälsig and B. Lankl, "Channel Parameter Estimation for LOS MIMO Systems," in *20th Int. ITG Work. Smart Antennas*, 2016, pp. 98–102.

[97] B. Hassibi and B. M. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[98] M. Tüchler and A. C. Singer, "Turbo Equalization: An Overview," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 920–952, 2011.

[99] F. Rusek, G. Colavolpe, and C. E. W. Sundberg, "40 Years with the Ungerboeck Model: A Look at Its Potentialities," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 156–161, 2015.

[100] N. Al-Dhahir and A. H. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921–2936, 2000.

[101] A. Lozano and C. Papadias, "Layered Space-Time Receivers for Frequency-Selective Wireless Channels," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 65–73, 2002.

[102] M. Chouayakh, A. Knopp, C. Ahokpossi, and B. Lankl, "Low Effort MIMO Detector for Frequency Selective Indoor Channels," in *Proc. 14th Eur. Wirel. Conf.*, 2008.

[103] H. Artés, D. Seethaler, and F. Hlawatsch, "Efficient Detection Algorithms for MIMO Channels: A Geometrical Approach to Approximate ML Detection," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2808–2820, 2003.

[104] E. Larsson, "MIMO Detection Methods: How They Work," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 91–95, 2009.

[105] N. Al-Dhahir, "FIR Channel-Shortening Equalizers for MIMO ISI Channels," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 213–218, 2011.

[106] X. Song, T. Halsig, D. Cvetkovski, W. Rave, B. Lankl, E. Grass, and G. Fettweis, "Design and Experimental Evaluation of Equalization Algorithms for Line-of-Sight Spatial Multiplexing at 60 GHz," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2570–2580, 2018.

[107] L. E. Franks, "Carrier and Bit Synchronization in Data Communication–A Tutorial Review," *IEEE Trans. Commun.*, vol. 28, no. 8, pp. 1107–1121, 1980.

[108] T. Hälsig, D. Cvetkovski, E. Grass, and B. Lankl, "Measurement Results for Millimeter Wave pure LOS MIMO Channels," in *Proc. IEEE Wirel. Commun. Netw. Conf.*, 2017.

[109] Z. Du, E. Ohlmer, K. Aronkytö, J. Putkonen, J. Kapanen, and D. Swist, "5G E-band Backhaul System Measurements in Urban Street-Level Scenarios," in *Proc. 47th Eur. Microw. Conf.*, 2017, pp. 372–375.

[110] L. Anttila, M. Valkama, and M. Renfors, "Circularity-Based I/Q Imbalance Compensation in Wideband Direct-Conversion Receivers," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2099–2113, 2008.

[111] A. Rezola, J. F. Sevillano, I. Gurutzeaga, D. Del Rio, R. Berenguer, and I. Velez, "Built-in-Self-Calibration for I/Q Imbalance in Wideband Millimeter-Wave Gigabit Transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 65, no. 11, pp. 4758–4769, 2017.

[112] H. Minn, Q. Zhan, N. Al-Dhahir, and H. Huang, "In-Phase and Quadrature Timing Mismatch Estimation and Compensation in Millimeter-Wave Communication Systems," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 7, pp. 4317–4331, 2017.

[113] C.-J. Hsu, R. Cheng, and W.-H. Sheen, "Joint Least Squares Estimation of Frequency, DC Offset, I-Q Imbalance, and Channel in MIMO Receivers," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2201–2213, 2009.

[114] B. Picinbono and P. Chevalier, "Widely Linear Estimation with Complex Data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, 1995.