

Compressive Sensing based Dynamic Spectrum Access for Interference-Networks

Dissertation

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

vorgelegt dem Fachbereich 1 (Physik/Elektrotechnik)

der Universität Bremen

von

Dipl.-Ing. Dennis Wieruch

Tag des öffentlichen Kolloquiums:	25. Juni 2021
Gutachter der Dissertation:	Prof. Dr.-Ing. A. Dekorsy Prof. Dr.-Ing. A. Sezgin
Weitere Prüfer:	Prof. Dr. A. Förster Dr. rer. nat. P. Jung, Guest Prof.



Universität
Bremen

Berlin, August 2021

Vorwort

Diese Arbeit ist meinem Vater gewidmet.

Die vorliegende Dissertation entstand unter anderem während meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der Abteilung Drahtlose Kommunikation und Netze am Fraunhofer Heinrich-Hertz-Institut.

Mein besonderer Dank gilt Dr. rer. nat. Peter Jung, der mich nicht nur in das Forschungsfeld von Compressive Sensing einführte, sondern mich auch während meiner Forschungsarbeit immer unterstützte, und zum Abschluss als Prüfer fungierte. Weiterhin gilt mein besonderer Dank Prof. Dr.-Ing. Armin Dekorsy, der mir ermöglichte im Arbeitsbereich Nachrichtentechnik an der Universität Bremen als externer Doktorand zu promovieren, und mir besonders in den letzten Zügen meiner Arbeit durch detailliertes und konstruktives Feedback zu der jetzt vorliegenden Fassung verhalf.

Außerdem möchte ich auch Prof. Dr.-Ing. Aydin Sezgin von der Ruhr-Universität Bochum für sein Interesse an meiner Forschungsarbeit und die Übernahme des Zweitgutachtens danken. Prof. Dr. Anna Förster bin ich für ihre Tätigkeit als Prüfer verbunden.

Weiterhin möchte ich allen Weggefährten am Fraunhofer HHI sowie im Arbeitsbereich Nachrichtentechnik an der Universität Bremen für den konstruktiven und fachlichen Austausch sowie das freundschaftliche Arbeitsumfeld danken. Besonders hervorheben möchte ich dabei meinen langjährigen Kollegen und Freund Bernd Holfeld. Der Austausch und die zahlreichen fachlichen Diskussionen haben meine Arbeit sehr bereichert.

Abschließend möchte ich meiner gesamten Familie, insbesondere meinen beiden Kindern Liana und Nevian, meiner Frau Lisa und meiner Mutter, danken, welche mich immer unterstützt und an mich geglaubt haben. Ohne deren Unterstützung und Rückhalt wäre ich heute nicht da, wo ich jetzt bin.

Berlin, August 2021

Dennis Wieruch

Contents

Contents	VII
1 Introduction	1
1.1 Objectives	2
1.2 Thesis Structure and Contributions	3
1.3 Notation	6
2 Dynamic Spectrum Access	9
2.1 Overview and Contributions	9
2.2 Spectrum Sensing	10
2.3 Channel Models	13
2.4 Transmission Model	17
2.5 Binary Hypothesis Testing	20
2.5.1 Neyman-Pearson Lemma	23
2.5.2 Karlin-Rubin Theorem	24
2.5.3 Constant False Alarm Rate Property	26
2.6 Signal Detection	27
2.6.1 Simple Log-Likelihood Ratio Test	32
2.6.2 One-Sided Composite Log-Likelihood Ratio Test	36
2.7 Summary	41
3 Compressed Sensing	43
3.1 Overview and Contributions	43
3.2 Sparsity	44
3.3 Linear Systems	46
3.4 Recovery Criteria	50
3.4.1 Null Space Property	52
3.4.2 Mutual Coherence	54
3.4.3 Restricted Isometry Property	54
3.5 Estimation Algorithms	56

3.5.1	Reduced-Rank Least Squares	56
3.5.2	Basis Pursuit Denoising	61
3.5.3	Orthogonal Matching Pursuit	64
3.6	Partial Fourier Measurements	66
3.6.1	Pilot-Based Channel Estimation	69
3.6.2	Amplitude-Based Channel Estimation	72
3.7	Summary	80
4	Gray Space Detection	83
4.1	Overview	83
4.1.1	Main Contributions	83
4.1.2	Structure	85
4.2	System Model	86
4.3	Detection Procedure	90
4.3.1	Step 1: Energy Detection	93
4.3.2	Step 2: Amplitude-Based Channel Estimation	95
4.3.3	Step 3: Log-Likelihood Ratio Test	99
4.4	Simulation Setup	104
4.5	Performance Evaluation	105
4.5.1	Receiver Operating Characteristic	106
4.5.2	Impact of SNR	108
4.5.3	Impact of System Load	109
4.5.4	Impact of Blocksize	110
4.5.5	Impact of Threshold Estimation	111
4.6	Practical Channel	114
4.6.1	Measurement Campaign	114
4.6.2	Representative Channel Models	118
4.6.3	Impact of Channel Model	121
4.7	Summary	122
5	Allocation Map Retrieval	125
5.1	Overview	125
5.1.1	Main Contributions	125
5.1.2	Structure	126
5.2	System Model	127
5.2.1	Channel Transfer Function Measurements	132
5.2.2	Spectral Channel Power Measurements	133
5.2.3	General Measurements	134
5.3	Combinatorial Complexity	137
5.4	Problem Analysis	141
5.4.1	Unlabeled Sensing	143

5.4.2	Related Subproblems	147
5.4.3	Unlabeled Selection Sensing	148
5.4.4	Problem of Multiple Selections	149
5.5	Objective Function	150
5.5.1	Exchange of Arbitrary Element	151
5.5.2	Mixing of Sparse Signals	153
5.5.3	Joint Estimate	154
5.6	Simulative Analysis	155
5.6.1	BPDN	157
5.6.2	OMP	159
5.6.3	Noisy Measurements	161
5.6.4	Spectral Channel Power Measurements	163
5.7	Structured Allocation Map	165
5.8	Results	168
5.9	Summary	173
6	Summary	175
A	Probability Distributions	181
A.1	Normal Distribution	183
A.2	Complex Normal Distribution	185
A.3	Circularly-Symmetric Complex Normal Distribution	186
A.4	Chi Distribution	186
A.5	Rayleigh Distribution	187
A.6	Chi-Squared Distribution	188
A.7	Gamma Distribution	190
A.8	Non-Central Chi-Squared Distribution	194
A.9	Scaled Non-Central Chi-Squared Distribution	195
	Acronyms	201
	Notation	205
	List of Symbols	207
	Own Publication List	215
	Bibliography: Books	219
	Bibliography	231
	Index	233

Chapter 1

Introduction

Future wireless [information and communications technologies](#) shall address manifold scenarios and applications. Key drivers behind this diversification are especially the vertical industries such as automotive, manufacturing, and energy in areas such as Industrie 4.0, [autonomous driving](#), [smart grid](#), [Internet-of-Things](#), [augmented reality](#), and Tactile Internet. Due to the heavy diversification, the requirements on future wireless information and communications technologies are manifold, comprising different types of wireless communications such as [low-power wide area networks](#), [device-to-device](#) communications, low-latency communications, as well as ultra-reliable communications. Moreover, each application demands a different mix of requirements. Therefore, supporting various devices with diverse requirements on a single infrastructure is a highly anticipated requirement for future wireless information and communications technologies.

Current wireless access systems like LTE, [WLAN](#), and [ZigBee](#) either address only individual types of wireless communications or are not suitable for the special requirements of certain applications [[SCF+13](#); [HWW+16](#)]. In the course of the fifth generation of cellular mobile communications (5G) [[ABC+14](#); [3GP20h](#)] the manifold scenarios and applications shall be represented by a single communication infrastructure. Therefore, we need more flexibility in resource allocation to efficiently exploit the scarce wireless spectrum while supporting the quickly increasing number of applications. Furthermore, to enable efficient communication especially in device-to-device applications the signaling overhead has to be reduced.

In the recent two decades flexible radio communication systems have been studied within the areas of dynamic spectrum access, [cognitive radio](#) [[Hay05](#)], and [software-defined radio](#). There, the communication system

shall adapt itself on the environment by mere observation, and for example allocate the communication resource accordingly. Furthermore, the application of compressed sensing [Don06; CRT06] introduced diverse new methods to exploit sparse or compressible structure in general and in particular in wireless communications. More specifically, the wireless channel has compressible features [Mol05], which can be exploited by compressed sensing approaches.

1.1 Objectives

The objectives of this thesis are the introduction of dynamic spectrum access schemes, which are enabled by compressed sensing. Moreover, a measurement campaign is described, where practical channel models are obtained for performance evaluation of the dynamic spectrum access schemes.

In particular, we propose a gray space detection scheme for cognitive radio system, which detects temporary small fraction of unused resources within an already occupied primary user spectrum band. While energy detection is a uniformly most powerful test for mere noise statistics, the gray space detection method performs a uniformly most powerful one-sided composite hypothesis test exploiting noise and channel statistics. We will describe the gray space detection algorithm and give practical implementation considerations, such that gray space detection can also be applied if the channel statistics are only partially known. To evaluate the performance of gray space detection, we compare gray space detection with energy detection for different parameters and channel estimation methods.

Furthermore, we extend the gray space detection scheme considering multiple active transmitters simultaneously transmitting on orthogonal resources within a given spectrum band. There, we propose an objective function to solve the combinatorial problem of allocation map retrieval for frequency-division multiple access signals and show empirically that based on the objective function a receiver is able to identify the non-adjacent resources belonging to the same transmitter. Subsequently, we describe an efficient, but suboptimal breadth-first search algorithm, which solves the allocation map retrieval problem by traversing a decision tree and reconstructing the resources belonging to the same transmitter. Finally, we show the reconstruction capability of allocation map retrieval in the presence of artificial sparse channels as well as practical performance channels.

1.2 Thesis Structure and Contributions

The results of this thesis have been partially supported by the German funded BMBF (Federal Ministry of Education and Research) projects CoMoRa (16BU1200) and KoI (16KIS0195), the German funded BMWi (Federal Ministry for Economic Affairs and Energy) project IC4F (01MA17008) as well as the European Union funded Horizon 2020 project FANTASTIC-5G (ICT-671660). In the following we will describe the main contributions of this thesis on chapter basis.

- Chapter 2 will lay the foundation of the wireless communication theory of this thesis, including the sparse wireless channel model and the transmission model. A strong focus of Chapter 2 is the statistical hypothesis test for signal detection. There, we derive a uniformly most powerful test for signal detection with single channel observations. In particular, we state the simple hypothesis test exploiting the Neyman-Pearson lemma for mere energy detection, such that a desired detection or false alarm rate is achieved. Then, we extend statistical hypothesis test for applications, where the cumulative distribution function of the observed signal is known. Using the Karlin-Rubin theorem we obtain a one-sided composite hypothesis test, where a desired detection or false alarm rate can be selected to achieve a uniformly most powerful test. Furthermore for practical usage, we exploit the Glivenko–Cantelli theorem and derive approximations of the one-sided composite hypothesis test for empirical cumulative distribution functions. Some of the results on statistical hypothesis tests are published in [WPJ13; WJW+16b].
- Chapter 3 introduces the fundamentals of compressed sensing, including the classical channel estimation approaches of least squares and reduced-rank least squares as well as the compressed sensing based channel estimation algorithms of basis pursuit denoising and orthogonal matching pursuit. We consider two channel estimation schemes for partial (random) Fourier measurements: pilot-based channel estimation to retrieve the channel impulse response from channel transfer function measurements as well as amplitude-based channel estimation to retrieve the circular autocorrelation of the channel impulse response from spectral channel power measurements. Based on probability distributions we derive error bounds for signal recovery for both channel estimation schemes. In particular, for pilot-based channel estimation we state the error bound depending on an approximated Gamma distribution. Furthermore for amplitude-based channel estimation, we

derive an estimate for the expectation of the measurement error.

- Chapter 4 introduces our gray space detection scheme [WPJ13; WJW+16b] as well as describes our channel measurement campaign in an industrial environment [HWW+16; HWR+16; WHW16; DHC+19]. We develop a gray space detection scheme, which exploits small fractions of unused resources for interweave cognitive radio system while one primary user is actively transmitting within the spectrum band. The main feature of gray space detection is that it can cope with frequency-varying wireless channels. In particular, gray space detection distinguishes inactive resources from frequency fading while controlling the protection level of the primary user system. Our gray space detection scheme consists of three steps. In the first step, we observe the signal from the primary user and estimates a preliminary set of active primary user resources via energy detection. In the second step, the preliminary set is used to estimate the spectral channel power from the primary user via amplitude-based channel estimation introduced in the previous Chapter 3. Eventually in the third step, we perform the log-likelihood ratio tests introduced in Chapter 2 to classify all resources by marking them as gray space or active primary user resource. Since the uniformly most powerful test is only applicable with full statistical knowledge of the channel, we derive two practical thresholds to perform log-likelihood ratio tests on the available channel statistics. While the the individual energy detector just uses an individual threshold for each resource based on the instantaneous channel knowledge, the more sophisticated gray space test is based on a composite threshold obtained from the empirical cumulative distribution function of the channel distribution. There, we analyze the performance deviation of gray space detection via the individual energy detector as well as via the gray space test towards the uniformly most powerful test. Furthermore, we analyze the performance of the gray space detection scheme using the different compressed sensing estimators like orthogonal matching pursuit (OMP) and basis pursuit denoising (BPDN), as well as the classical reduced-rank least squares (RRLS) method. There, we evaluate the three amplitude-based channel estimation approaches with regard to SNR, primary user system load and allocated physical resource blocksize. To assess the performance of our gray space detection scheme also in practical scenarios, we conduct a measurement campaign to obtain channel statistics for representative manufacturing process in factory automation. In particular, we obtain short-range factory automation channel models, which are applied on the gray space detection scheme.

Gray space detection is eventually evaluated regarding our short-range factory automation channel models as well as the common ITU channel models. The results of Chapter 4 have been partially published in [WPJ13; WJW+16b; HWW+16; HWR+16; WHW16; DHC+19].

- Chapter 5 extends the framework discussed in Chapter 4 to multiple transmitters. While a receiver in Chapter 4 observes the power spectrum where a single transmitter is present with non-adjacent active resources, a receiver in Chapter 5 observes the superposition of multiple time signals containing non-adjacent, but disjoint active frequency resources. We elaborate this framework for two distinct cases: (i) channel transfer function measurements observing pilot transmissions, and (ii) spectral channel power measurements observing constant amplitude signals. Subsequently, we state the resulting joint estimation problem, where the receiver shall jointly recover all resources belonging to the same transmitter as well as the channel to each transmitter (channel impulse response for case (i) and circular autocorrelation of the CIR for case (ii)). In addition, we show that a part of the joint estimation problem is related to a problem called unlabeled sensing. To solve the joint estimation problem we derive an objective function based on sparsity of the transmission channel, such that the resources of the same transmitter can be identified. Furthermore, we propose an ℓ_1 -norm measure for basis pursuit denoising (BPDN) and a joint ℓ_0 -pseudonorm ℓ_2 -error measure for orthogonal matching pursuit (OMP). The feasibility of the objective function is shown via numerical simulations, where we verify the objective function by performing exhaustive search for small combinatorial problems. Based on these results, we developed an efficient but suboptimal breadth-first search algorithm, which traverses a decision tree and reconstructs the resources belonging to the same transmitter. The breadth-first search algorithm is evaluated via simulations based on artificial sparse channels as well as on practical performance channels. There, the applied practical performance channels also include the measured channels from Chapter 4. In addition, a transmission concept based on the discussed framework has been filed as an international patent application [WJH17]. The results of Chapter 5 have been partially published in [WJW+15; WJB+15; WJW+16a; WJH17].

1.3 Notation

In this Section, we will give a description of the conventions and the nomenclature utilized throughout the thesis. A summary of the notation and a list of symbols specifying the meaning of each used variable are also given at the end of the thesis.

The italic lower- and uppercase letters a, A will be used to denote scalars unless stated otherwise. In addition, lower- and uppercase bold letters \mathbf{a}, \mathbf{A} will be used to describe column vectors and matrices, respectively. Index sets are generally represented by blackboard bold uppercase letters \mathbb{A} . A subscript $(\cdot)_{\text{LS}}$ is used to distinguish different variables. Furthermore, a subscript $(\cdot)_n$ can also indicate dependency on a subscript variable.

The notation $a \in [n, m)$ can be rewritten to $n \leq a < m$, i.e., a can take all values between n and m excluding the value m . Real and imaginary part of variable a are denoted by $\Re\{a\}$ and $\Im\{a\}$, respectively. An element of matrix \mathbf{A} or matrix \mathbf{A}_a at row n and column m is denoted by $\mathbf{A}_{[n,m]}$ or $(\mathbf{A}_a)_{[n,m]}$, respectively. Thus, a column vector is described with $\mathbf{a} = [\mathbf{a}_{[0]} \ \mathbf{a}_{[1]} \ \dots \ \mathbf{a}_{[N-1]}]^\top$. Furthermore, multiple elements of a matrix or vector can be selected over sets. For example, $\mathbf{B} = \mathbf{A}_{[\mathbb{A}, \mathbb{B}]}$ is a submatrix of \mathbf{A} , containing the rows of \mathbf{A} marked by the index set \mathbb{A} and the columns of \mathbf{A} marked by the index set \mathbb{B} . In addition, $\mathbf{B} = \mathbf{A}_{[\cdot, \mathbb{B}]}$ contains all the rows of \mathbf{A} , but selects only the columns of \mathbf{A} marked by the index set \mathbb{B} .

Set of natural, real and complex numbers are denoted by \mathbb{N}, \mathbb{R} and \mathbb{C} , respectively. Moreover, the set with notation $\mathbb{R}_{\geq b} = \{a \in \mathbb{R} : a \geq b\}$ contains all real numbers equal or larger than b . A common special case is $\mathbb{R}_{\geq 0} = \{a \in \mathbb{R} : a \geq 0\}$, where the set contains all positive real numbers including zero. The dimension of matrices and vectors are generally described in combination with the domain. Thus, $\mathbf{A} \in \mathbb{C}^{N \times M}$ represents a complex matrix \mathbf{A} with N -rows and M -columns. Consequently, $\mathbf{a} \in \mathbb{R}^{N \times 1}$ or $\mathbf{a} \in \mathbb{R}^N$ represents a real column vector \mathbf{a} with N -rows. Furthermore, an $(N \times M)$ matrix contains N -rows and M -columns, but the domain of the matrix is not explicitly given.

A tilde above a variable (\tilde{a}) denotes the estimate of the variable (a). A modified or alternative meaning of a defined variable is expressed by $(\cdot)'$. In addition, the superscripts $(\cdot)^\top$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, the complex conjugate and the conjugate transpose (Hermitian transpose), respectively. The inverse and the pseudoinverse of matrix \mathbf{A} is given by \mathbf{A}^{-1} and \mathbf{A}^+ , respectively. The circumflex $\hat{(\cdot)}$ is used to describe the discrete Fourier transform of a variable. Thus, the discrete Fourier transform of a

vector is

$$\hat{\mathbf{a}}_{[m]} = \frac{1}{\sqrt{N}} \cdot \sum_{n=0}^{N-1} \mathbf{a}_{[n]} \cdot \exp\left(-i2\pi \frac{m \cdot n}{N}\right). \quad (1.3.1)$$

The symbols mod , \otimes and \cdot denote the operator of modulo, Kronecker product and matrix product, respectively. Furthermore, the absolute value or cardinality of a set is obtained by $|\cdot|$.

The ℓ_p -norm, quasinorm and "norm" are defined as

$$\|\mathbf{a}\|_p := \begin{cases} |\{n : \mathbf{a}_{[n]} \neq 0\}| & p = 0 & \text{"norm"} \\ (\sum_n |\mathbf{a}_{[n]}|^p)^{1/p} & p \in (0, 1) & \text{quasinorm} \\ (\sum_n |\mathbf{a}_{[n]}|^p)^{1/p} & p \in [1, \infty) & \text{norm} \\ \max_{\forall n} (|\mathbf{a}_{[n]}|) & p = \infty & \text{norm.} \end{cases} \quad (1.3.2)$$

For better readability at appropriate passages, we will refer to the ℓ_p -norm also for $p < 1$. In addition, we define the support of a vector \mathbf{a} by

$$\text{supp}(\mathbf{a}) := \{n : \mathbf{a}_{[n]} \neq 0\}. \quad (1.3.3)$$

Furthermore, the [indicator function](#) is defined as

$$\mathbf{1}_{\mathbb{A}}(a) := \begin{cases} 1 & a \in \mathbb{A} \\ 0 & \text{else.} \end{cases} \quad (1.3.4)$$

Chapter 2

Dynamic Spectrum Access

2.1 Overview and Contributions

This Chapter will lay the foundation of the wireless communication theory for the subsequent chapters. Therefore, we discuss the sparse wireless channel model and introduce our basic transmission model, which will be used throughout this thesis. Then we will introduce some basics on statistical hypothesis test and the notion of a uniformly most powerful test. Furthermore, we derive a uniformly most powerful statistical hypothesis test for signal detection with single sample observations and provide formulas to obtain desired detection or false alarm rates.

We will start by discussing relevant publications in the field of dynamic spectrum access and spectrum sensing in Section 2.2. Subsequently, we introduce the multipath channel model in Section 2.3 and state common performance channel models summarized in Table 2.1, which will be used for performance analysis throughout this thesis. In general, wireless channels have sparse features, which will be important for the use of the compressed sensing estimators provided in Section 3.5 as well as for our contributions in Chapters 4 and 5. Afterwards, we turn our attention to the general system model of this thesis in Section 2.4. In particular, we consider a cyclic-prefix based OFDM communication system similar to 3GPP LTE and 3GPP NR. Next, in Section 2.5 we discuss the statistical hypothesis test and the uniformly most powerful test, which is an optimal test to achieve a desired performance criterion. In particular, we discuss the uniformly most powerful likelihood ratio test established as simple hypothesis test (Neyman-Pearson lemma in Section 2.5.1) and as one-sided composite hypothesis test (Karlin-Rubin theorem in Section 2.5.2). Besides the uniformly most pow-

erful test, we introduce constant false alarm rate property in Section 2.5.3, which is desired in scenarios where the performance criterion of false alarm rate shall be achieved independent of the parameters within the family of distributions. To achieve uniformly most powerful test in our proposed gray space detection scheme in the later Chapter 4, we extend the statistical hypothesis test of our publications [WPJ13; WJW+16b] to a uniformly most powerful one-sided composite hypothesis test by exploiting the Karlin-Rubin theorem in Section 2.6. Furthermore, we derive the thresholds for the one-sided composite hypothesis test in Section 2.6.2 depending on the false alarm rate and on the detection rate in (2.6.36) and in (2.6.43), respectively. In particular, to enable statistical hypothesis test in practical settings, we exploit the Glivenko–Cantelli theorem to obtain the empirical CDF from i.i.d. observations. For sake of completeness, we also present the thresholds for the simple hypothesis test in Section 2.6.1 depending on the false alarm rate and on the detection rate in (2.6.23) and in (2.6.25), respectively. The chapter concludes with a short summary on dynamic spectrum access in Section 2.7.

Summary of own contributions within this Chapter:
[WPJ13; WJW+16b]

2.2 Spectrum Sensing

The licensed spectrum allocations carried out by regional, national and international regulation entities such as ITU experience a slow transformation to more flexibility. In the past and still currently, licenses are generally assigned in the magnitude of several years. However, in the past decades spectrum sharing schemes like licensed shared access have been evolved, enabling new spectrum license procedures [KHB+15; WWH+16]. Especially regarding 5G, new business models are likely to evolve, which demand for regional and local licenses. One model under increasing discussion is private networks [Qua17; PRI+18], which shall enable companies to deploy their own wireless networks in licensed spectrum. In particular, Germany is pushing forward private networks under the term “campus network” for usage in wide variety of industries [BMW20]. First trials on private networks have been conducted by network operators like the Deutsche Telekom [Deu18]. Nevertheless, the majority of the licensed spectrum is still fixed and investigations have shown its sparse utilization [VMB+10]. Thus, spectrum sensing will be an important technology component of future inter-network optimization for distinct private networks.

To increase the utilization of the licensed, but marginally exploited spectrum bands, tremendous research has been conducted in the field of cogni-

tive radio [Hay05] and dynamic spectrum access (DSA) [GJM+09; YA09]. Spectrum sensing has been one key technology to enable cognitive radio systems [Hay05]. The research was mainly driven by exploiting unused spectrum bands of the so called white spaces at certain time and location, which are an outcome of the underutilized spectrum. The classical spectrum sensing scheme of energy detection has been introduced by Urkowitz in 1967 [Urk67]. Classical energy detection indicates based on power threshold whether an observed spectrum band is occupied or empty. In particular, the proposed scheme of Urkowitz exploits the central limit theorem using the sum of i.i.d. random variables. A good introduction on classical energy detection can be found in [ATJ14]. Although the first paper was already published in the late 1960s, the major research on spectrum sensing has been done in the last two decades, leading to a large number of publications [YA09; SNW+13]. Spectrum sensing can be broadly divided into narrowband and wideband spectrum sensing. While the studies on spectrum sensing mainly focused on narrowband signals at the beginning of the last decade [YA09], wideband spectrum sensing became the major focus in recent years [SNW+13]. Besides the classical energy detection, other spectrum sensing approaches can be summarized in the group of cyclostationary as well as feature detection [YA09].

While the term white space is generally used for unoccupied spectrum bands (in wideband as well as narrowband spectrum sensing), the term gray space refers to a temporary unused small fraction of spectrum within an already occupied spectrum band. A similar terminology for gray space is introduced in [Sar12; Peh13], where they also present and discuss different scenarios regarding gray spaces. In particular, they list radar systems as well as (cellular) mobile communication systems as primary user system, which is not fully exploiting the time-frequency resources. However, Saruthirathanaworakun and Peha do not consider spectrum sensing, but spectrum sharing mechanism to exploit the gray spaces.

Spectrum sensing schemes on gray spaces would permit further exploitation of the scarce spectrum, which is not fully utilized by an existing wireless system. Even though spectrum sensing for narrowband signals has been studied intensively, the most of them address only white spaces. Furthermore, the most spectrum sensing application make extensive use of the central limit theorem and the Wilks theorem [Wil38], where the sensing decisions are based on averaging over multiple samples. However, in the case of observations with just single samples the central limit theorem or the Wilks theorem cannot be exploited. Therefore, we will study in Sections 2.5 and 2.6 spectrum sensing in the case where the central limit theorem and the Wilks theorem cannot be applied, which will be used later on in our

proposed gray space detection scheme in Chapter 4. For classical energy detection, some performance evaluations have already been conducted by Digiham et al. [DAS07] under fading channels, which are also applicable for gray spaces as well as single (complex) samples.

Even though wideband spectrum sensing is not the focus of this thesis, we discuss some relevant approaches like *edge detection* and *Haar wavelets* applied in wideband spectrum sensing. In the classical signal models that operate on Nyquist rate, wideband spectrum sensing is not feasible in practical systems due to the enormous hardware requirements. However, through the advances in compressed sensing [Don06; CRT06] (see Chapter 3) and sub-Nyquist sampling [VB01; WP11], the application of wideband spectrum sensing became practical. Sub-Nyquist sampling exploits the fact, that the reconstruction of multiband signals within a wideband spectrum is possible at Landau rate [Lan67a], which can be much smaller than the Nyquist rate. Tian and Giannakis [TG06; TG07a] proposed a wideband spectrum sensing scheme based on sub-Nyquist sampling. The sparse spectrum occupation is obtained through *edge detection*, which identifies the boundaries of piecewise smooth and almost flat power spectral density within a wideband spectrum. Recently, Beck et al. [BBD19] presented a practical approach for wideband spectrum sensing via edge detection. Further methods based on the compressed sensing framework have been proposed for spectrum sensing on wideband signals, e.g. multi-coset sampling and multirate wideband sampling [SNW+13]. As an alternative to the Fourier basis used in [TG07a], *Haar wavelets* were recently proposed as sparsity basis [EO15]. While the Fourier basis assumes sparse spectrum utilization, Haar wavelets are independent of the spectrum sparsity and directly exploit the edges between spectrum bands. Consequently, Haar wavelets lead to superior performance in strong utilized spectrum in comparison to [TG07a].

The techniques of *edge detection* as well as *Haar wavelets* are addressing similar problems as the ones discussed in our approaches in Chapters 4 and 5. However, a major drawback is that edge detection as well as Haar wavelets merely consider *constant functions* with some additive noise¹. but not general smooth functions like sine waves, Therefore, since propagation channels are in general not flat, these techniques are not applicable for narrowband spectrum sensing. In the subsequent Section 2.3 we discuss the properties and characteristics of propagation channels.

¹The noise within the power spectrum is in general following the Gamma distribution.

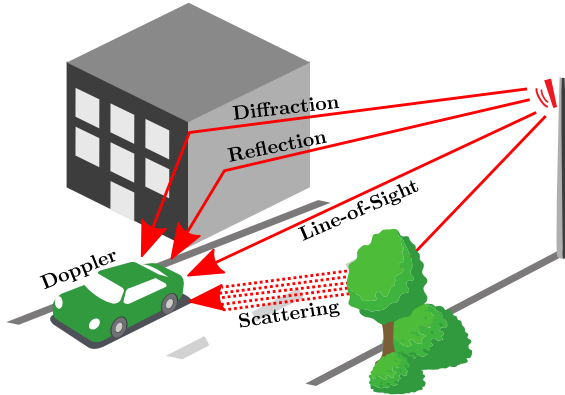


Figure 2.1: The multipath propagation channel model induced by obstacles of the environment and object velocity.

2.3 Channel Models

A signal emitted from a source via electromagnetic waves traverses through a wireless medium to its destination. The paths of the signal lead through an environment consisting of various obstacles, where each obstacle alters the signal and its direction. In particular, diffraction, reflection and scattering of obstacles as well as the velocity of obstacles, source and destination are the major effects on the signal. Consequently, multiple altered copies of the signal arrive at the destination over different ways at different times. These effects are described by channels following the multipath propagation model, or multipath channel model. A visual illustration of multipath propagation and the effects caused by the obstacle is shown in Figure 2.1.

In general, the paths can be distinguished into line-of-sight and non-line-of-sight. While the line-of-sight path represents the direct path of the signal from the transmitter to the receiver (see Figure 2.1), non-line-of-sight paths are traversing over obstacles. In general, if a dominant line-of-sight path is present, a propagation channel exhibits [Rice fading](#), which is described by the [Rice distribution](#), see Appendix A.9. In contrast to Rice fading with non-zero mean, [Rayleigh fading](#) with zero mean occurs in propagation channels, where mere non-line-of-sight paths are present or the line-of-sight path is not dominant. More on the practical distinction between line-of-sight and non-line-of-sight propagation channels will be discussed later in Section 4.6.2. Rayleigh fading is described by the [Rayleigh distribution](#), see Appendix A.5.

Due to the continuous altering nature of the environment, a multipath

channel is modeled by a linear time-variant system. In particular, the wide-sense stationary uncorrelated scatterers (WSSUS) model introduced by Bello in 1963 [Bel63] is assumed for multipath channels. There, the wide-sense stationary property describes that the first and second order statistics of the channel are time invariant. In addition, the uncorrelated scatterers property defines the statistical independence of the multipath components. Further explanation can also be found in [Mol11, Chapter 6.4].

A channel of a linear time-variant system is described by its continuous channel impulse response $h(t, \tau)$, which is a function of the absolute time t and delay τ of multipath components. In general, wireless channels are varying slowly in time. Thus, on appropriate small time scales a channel can be modeled as linear time-invariant (LTI), which is commonly also referred to as a block-fading channel. Within this block the channel impulse response (CIR) is fully described by $h(\tau)$, which is a function of the delay τ of multipath components only. The excess delay τ_{ED} describes the time span between the first and the last multipath component. However, in reality the last multipath component is not well defined. Therefore, the excess delay commonly describes the time span starting from the first multipath component in which a high percentage of the complete signal energy is present

$$\int_0^{\tau_{\text{ED}}} |h(\tau)|^2 d\tau \gg \int_{\tau_{\text{ED}}}^{\infty} |h(\tau)|^2 d\tau. \quad (2.3.1)$$

Thus, given a ratio ρ_{ed} between zero and one, the **excess delay** is calculated by

$$\tau_{\text{ED}} := \underset{\tau_{\text{ED}}}{\operatorname{argmin}} \left(\rho_{\text{ed}} < \frac{\int_0^{\tau_{\text{ED}}} |h(\tau)|^2 d\tau}{\int_0^{\infty} |h(\tau)|^2 d\tau} \right). \quad (2.3.2)$$

Since the actual CIR highly depends on the environment, different channel models are applied for realistic representation of the natural environment. For link level simulations performance channel models are commonly used, representing urban, rural and indoor environments with different velocities. There, the model is usually described by the tapped delay line model, where the average power and the corresponding delay is given for each multipath. From these parameters independent realizations of a block-fading channel can be generated, so that for each realization n , different CIR functions $h(n, \tau)$ can be derived. Usually, the power of the CIR realizations are

Model	Number of Multipaths	r.m.s. Delay Spread [ns]	Excess Delay [ns]
EPA	7	43	410
EVA	9	357	2510
IOA	6	37	310

Table 2.1: Summary of ITU and 3GPP channel model parameters.

normalized, such that the expectation of the accumulated power is one

$$\mathbb{E} \left[\sum_{\tau} |h(n, \tau)|^2 \right] = 1 \quad \forall n. \quad (2.3.3)$$

Performance channel models are mainly used for performance evaluation, comparing different link level simulations. Some of the most common performance channel models are

- ITU indoor office A (IOA) [ITU97, Chapter 1.2.2],
- 3GPP extended pedestrian A (EPA) [3GP20e, Chapter B.2.1] and
- 3GPP extended vehicular A (EVA) [3GP20e, Chapter B.2.1],

which will be used throughout this thesis. While EPA and EVA are used by 3GPP [3GP20e], IOA was defined in 1997 by ITU in [ITU97]. In particular, the 3GPP channel models EPA and EVA represent extended versions for 20 MHz bandwidth [SMF05] of the preliminary standardized ITU channel models pedestrian A and vehicular A with 5 MHz bandwidth [ITU97, Chapter 1.2.2]. EPA and EVA represent the extended channel models. While the 3GPP EPA channel model represents an outdoor scenario for pedestrian users, the 3GPP EVA represents an outdoor environment for vehicular devices. In addition, the ITU IOA models an indoor office scenario. The major difference between the three channel models lies within the number of multipath components, the distribution of its components and the excess delay. While IOA has the smallest excess delay, EPA suffers from a slightly larger excess delay. The largest excess delay is caused by EVA, which is more than a factor of six in relation to EPA. The parameters are summarized in Table 2.1. In addition, the power delay profile is visualized in Figure 2.2.

Since the channel transfer function is the Fourier transform of the CIR, the delay and values of the multipath components have direct affect on

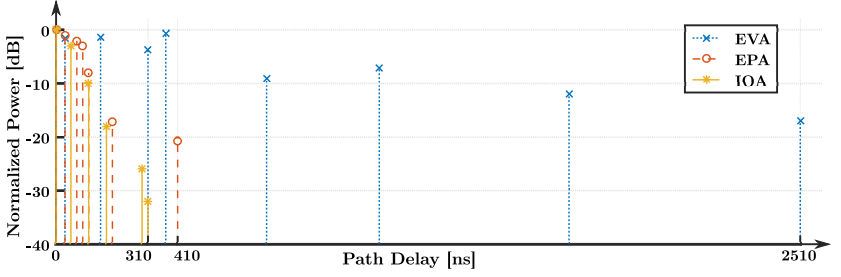


Figure 2.2: The power delay profile of the channel models ITU IOA [ITU97], 3GPP EPA [3GP20e] and 3GPP EVA [3GP20e].

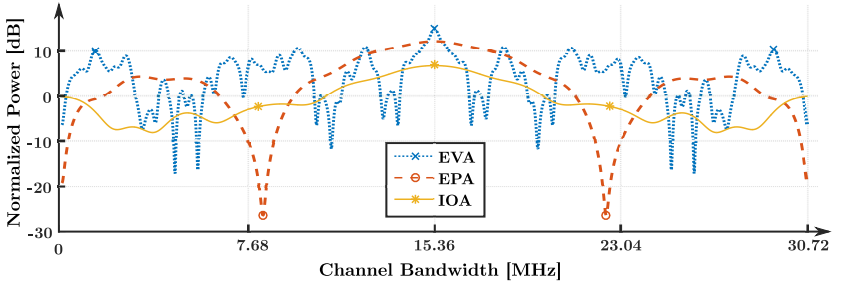


Figure 2.3: Examples for the spectral channel power of the models ITU IOA [ITU97], 3GPP extended pedestrian A (EPA) [3GP20e] and 3GPP extended vehicular A (EVA) [3GP20e].

it. While the vast EVA model exhibits several frequency fades, the shorter EPA and IOA models consist only of a few fading holes. An example for the different fading characteristics of the models is visualized in Figure 2.3, where the power of the channel transfer function is shown. Note, that the spectral channel power in Figure 2.3 is symmetric, because the examples are generated by setting the CIR merely to real values. In the following we will refer to the spectral channel power to describe the power of the channel transfer function. In particular, the spectral channel power describes the power spectral density of the CIR.

In real environments, the actual CIR as well as its transforms have to be estimated from measurements of the wireless channel, where the measurements are generally acquired through transmission of a-priori known signals. The estimation methods in general as well as the estimation of the wireless channel from measurements will be discussed later in Section 3.5 and Section 3.6, respectively. In addition, we describe the process to acquire the

CIR from a measurement campaign in Section 4.6.

Furthermore, a significant consequence of the multipath propagation model is that few multipath components are sufficient to describe the wireless channel in various environments. Therefore, sparsity (or compressibility as introduced in Section 3.2) is natural induced by multipath propagation. The effect is also validated in many measurement campaigns [Mol05]. Thus, a sparse channel model for the time domain representation of the channel can be used. Because this is an important aspect of this thesis, it will be further discussed in Chapter 3.

2.4 Transmission Model

Throughout this thesis, we consider a multi-carrier system like the cyclic-prefix (CP) based orthogonal frequency-division multiplexing (OFDM) communication system similar to 3GPP LTE [3GP20a] and NR [3GP20h]. There, one subcarrier describes the frequency modulated version of the complex frequency sample $\hat{\mathbf{u}}_{[f]}^2$ at OFDM frequency or subcarrier index $f \in \mathbb{F}$ within the contiguous band $\mathbb{F} := \{0, \dots, N_{\text{DFT}} - 1\}$, where N_{DFT} denotes the size of the discrete Fourier transform (DFT) matrix. Furthermore, we denote one resource as one complex frequency sample $\hat{\mathbf{u}}_{[f]} \in \mathbb{P}$ within the contiguous transmission band $\mathbb{P} \subseteq \mathbb{F}$, cf. Figure 2.4. Thus, subcarriers are only available as resources within the transmission band \mathbb{P} , whereby the remaining subcarriers are set to zero. Hence, *given the time sample indices $t \in \{0, \dots, N_{\text{DFT}} - 1\}$ the discrete samples of an **OFDM time domain symbol** are obtained by*

$$\mathbf{u}_{[t]} = \frac{1}{\sqrt{N_{\text{DFT}}}} \sum_{f \in \mathbb{P}} \exp\left(i2\pi \frac{ft}{N_{\text{DFT}}}\right) \hat{\mathbf{u}}_{[f]}, \quad (2.4.1)$$

which are generated via the inverse DFT from the complex frequency sample $\hat{\mathbf{u}} \in \mathbb{C}^{N_{\text{DFT}}}$. If the resource set \mathbb{P} is chosen with $|\mathbb{P}| < N_{\text{DFT}}$, then the transmission vector $\mathbf{u} \in \mathbb{C}^{N_{\text{DFT}}}$ is an oversampled version of the frequency vector $\hat{\mathbf{u}}_{[\mathbb{P}]} = \{\hat{\mathbf{u}}_{[f]} : f \in \mathbb{P}\}$. The transmission signal is considered solely in the discrete domain without transmit filters like the raised cosine filter. Therefore, we operate with time sample $\mathbf{u}_{[t]}$ with the time index t and frequency sample $\hat{\mathbf{u}}_{[f]}$ with the frequency index or subcarrier f .

For our analysis it is sufficient to operate on single OFDM symbols with CP. Therefore, we do not require an indexing of OFDM symbols and the column vectors \mathbf{u} as well as $\hat{\mathbf{u}}$ refer to this single OFDM symbol.

² $\hat{\mathbf{u}}_{[f]}$ denotes the element at index f of vector $\hat{\mathbf{u}}$

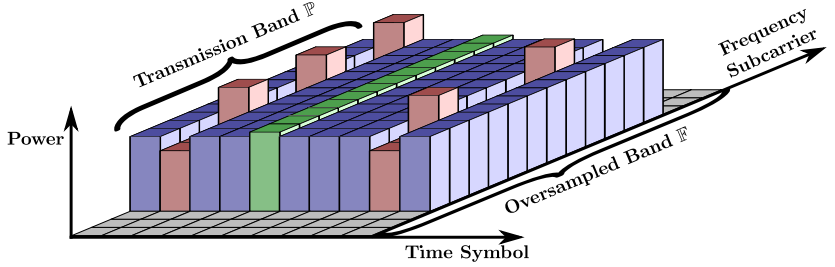


Figure 2.4: Example for transmit power allocation in OFDM based communication system.

Since our proposed approaches in Chapters 4 and 5 heavily rely on transmit signals with constant amplitudes in the frequency domain like frequency symbols with phase-shift keying, an important aspect is transmit power control and allocation in general OFDM based communication system. Power control and allocation are dynamically applied at the transmitter to increase power efficiency of the transmitter, to cope with the near-far problem and to harden sensitive control information. In general, these procedures are applied for the complete transmission band or for predefined sets of subcarriers. An example is visualized in Figure 2.4. In particular, 3GPP LTE describes simple power allocation for the base station [3GP20d, Chapter 5.2] and more advanced schemes of power control/allocation for the mobile device called user equipment [3GP20d, Chapter 5.1]. Moreover, 3GPP NR [3GP20f; 3GP20g] enhances power control/allocation further for 5G. However, power control/allocation is still applied only for predefined sets of subcarriers. Adjusting the power for each resource individually would require tremendous signal overhead, since receiver signal processing like demodulator and equalizer needs knowledge about the relative power level for correct operation. Consequently, the structure imprinted by the frequency samples $\hat{\mathbf{u}}$ is also kept after power control/allocation. Thus in general, phase-shift keying will be still transmitted with constant amplitude and (2.4.1) is valid for the transmitted signal.

To reduce signal overhead, transmission resources are aggregated to physical resource blocks (PRBs). In most OFDM communication modes, a PRB contains consecutive subcarriers over multiple OFDM symbols. However, in some modes (e.g. 3GPP single-carrier frequency-division multiple access (SC-FDMA)) the aggregated subcarriers of a PRB are distributed among the transmission band \mathbb{P} . In the default mode of 3GPP LTE (normal CP), one PRB aggregates 12 consecutive resources over 7 OFDM symbols [3GP20c], which also represents the smallest unit used for scheduling. Here, we will

refer to a PRB containing consecutive subcarriers unless it is stated otherwise. Since we just consider single OFDM symbols the aggregation of OFDM symbols over time is mostly irrelevant for this thesis. We will use the term blocksize N_B to denote the number of aggregated subcarriers within one PRB.

The transmission signal is altered by a block-fading channel and additive white Gaussian noise. Therefore, in the continuous domain, *the **received signal** w can be expressed as the convolution of the transmitted signal u with the continuous time-invariant channel impulse response h by*

$$w = h \circledast u + e, \quad (2.4.2)$$

where \circledast denotes the convolution operator. In general, we can expect the CIR h to be sparse, which is a realistic assumption as pointed out in [Mol05] and will be further discussed in Chapter 3. We assume that the receiver is frequency synchronized with the transmitter on a subcarrier basis, but asynchronous in time with an unknown time offset $\tau_{\text{offset}} \in \mathbb{R}$. To reflect the time inaccuracy at the receiver, the function h is the convolution of the wireless channel h_w and the Dirac delta function $\delta(\cdot)$ with the unknown time offset

$$h(\tau) := h_w(\tau - \tau_{\text{offset}}) = \delta(\tau - \tau_{\text{offset}}) \circledast h_w(\tau). \quad (2.4.3)$$

The maximum excess delay $\tau_{\text{ED,max}}$ of the CIR of the wireless channel h_w is assumed to be smaller than the cyclic-prefix length N_{CP} . Furthermore, we assume that unknown time offset is

$$\tau_{\text{offset}} \in [0, N_{\text{CP}} - \tau_{\text{ED,max}}). \quad (2.4.4)$$

Thus, the resulting CIR function h remains within the cyclic-prefix of length N_{CP} . Consequently, the asynchronicity for an OFDM system can be considered in the frequency domain by using the Fourier transform

$$\int_{-\infty}^{\infty} \exp\left(-i2\pi a \frac{b}{N_{\text{DFT}}}\right) \delta(b - \tau_{\text{offset}}) \, da = \exp\left(-i2\pi a \frac{\tau_{\text{offset}}}{N_{\text{DFT}}}\right). \quad (2.4.5)$$

Therefore, as long as the unknown time offset τ_{offset} stay within the bounds of (2.4.4) we can permit time asynchronicity of the receiver. Thus, the function h models both, the wireless channel h_w and the time asynchronicity by an unknown time offset τ_{offset} .

Now we turn back to the discrete domain. There, *the **received time samples** of (2.4.2) are expressed as the vector $\mathbf{w} \in \mathbb{C}^{N_{\text{DFT}}}$ calculated by the*

truncated discrete convolution

$$\mathbf{w}_{[t]} = \mathbf{n}_{[t]} + \sum_{\tau=0}^{N_{\text{CP}}} \mathbf{h}_{[\tau]} \cdot \mathbf{u}_{[(t-\tau) \bmod N_{\text{DFT}}]} \quad (2.4.6)$$

with $t \in \{0, \dots, N_{\text{DFT}} - 1\}$, where $\mathbf{h} \in \mathbb{C}^{N_{\text{CP}}+1}$ is the block-fading channel and $\mathbf{n} \in \mathbb{C}^{N_{\text{DFT}}}$ represents the additive white Gaussian noise vector. The modulo operation \bmod on the transmit vector \mathbf{u} from (2.4.1) extends the transmit vector by the CP. The additive white Gaussian noise vector \mathbf{n} is circularly-symmetric complex normal distributed $\mathbf{n}_{[t]} \sim \text{f}_{\mathbb{CN}}(\cdot; 0, \sigma_n^2)$. Exploiting the property of an OFDM system [Ste11, Page 131], the **received frequency samples** through an N_{DFT} -point DFT can be written in scalar notation as [Ste11, Equation 5.8]

$$\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]}, \quad f \in \mathbb{P}, \quad (2.4.7)$$

or as matrix notation

$$\hat{\mathbf{w}} = \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} + \hat{\mathbf{n}}, \quad (2.4.8)$$

where $\hat{\mathbf{w}}, \hat{\mathbf{h}}, \hat{\mathbf{u}}, \hat{\mathbf{n}} \in \mathbb{C}^{N_{\text{DFT}}}$. Consequently, the signal-to-noise ratio (SNR) in dB of a received signal is calculated by the ratio of expectations with

$$\text{SNR}_{\text{dB}} = 20 \cdot \log_{10} \left(\mathbb{E} \left[\left\| \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} \right\|_2 \right] \right) - 20 \cdot \log_{10} (\mathbb{E} [\|\hat{\mathbf{n}}\|_2]), \quad (2.4.9)$$

where $\hat{\mathbf{n}}_{[f]}$ is set to zero for $f \in \mathbb{F} \setminus \mathbb{P}$ so that noise is considered only at active resource indices. Furthermore, the instantaneous SNR is directly calculated from a received OFDM symbol and is given by

$$\text{SNR}_{\text{dB, Instant}} = 20 \cdot \log_{10} \left(\left\| \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} \right\|_2 \right) - 20 \cdot \log_{10} (\|\hat{\mathbf{n}}\|_2). \quad (2.4.10)$$

2.5 Binary Hypothesis Testing

A statistical hypothesis test is applied in decision theory to obtain a decision based on the probability distributions of the underlying process. Lehmann and Romano [LR05] provide a good overview on this topic and will be used as a reference in the following sections. Here, we will focus on probability density functions (PDFs) described in Appendix A, which restrict the probability distributions to **continuous functions**. Thus, its **cumulative distribution functions** are **absolute continuous**.

In this Section we will provide the definitions of binary hypothesis tests and in particular the log-likelihood ratio test. There, we consider two classes

of tests: (i) the simple hypothesis test and (ii) the composite hypothesis test. For these two test classes we will state the requirement to achieve a uniformly most powerful (UMP) test, which is highly desirable since a UMP test is optimal for the given PDFs. Furthermore, we will also introduce the property of constant false alarm rate, which may be easier achieved in scenario with unknown or partially known distributions.

A special case of statistical hypothesis tests is the binary hypothesis test, where a null hypothesis \mathcal{H}_0 is tested against an alternative hypothesis \mathcal{H}_1 . The random variable Y of each of these hypotheses follows a PDF $f_Y(\cdot; \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Theta$ is a parameter vector³ from a real-valued parameter set Θ . Furthermore, the parameter set $\Theta_0 \subset \Theta$ is defined as a proper subset of Θ . Consequently, $\Theta_1 = \Theta \setminus \Theta_0$ defines the complementary set of Θ_0 . Thus, given $|\Theta_{n \in \{0,1\}}| = 1$ the **binary hypotheses** are described by

$$\mathcal{H}_0 : Y \sim f_Y(\cdot; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta_0, \quad (2.5.1)$$

$$\mathcal{H}_1 : Y \sim f_Y(\cdot; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta_1. \quad (2.5.2)$$

However, in general one hypothesis does not consist of an exact PDF, but a family of PDFs. Therefore, we introduce **family of PDFs** described by

$$f_Y(\cdot; \boldsymbol{\theta}_{[1]} \in \Theta_{n,1}, \dots, \boldsymbol{\theta}_{[M]} \in \Theta_{n,M}) := \{f_Y(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta_n\} \quad (2.5.3)$$

with $n \in \{0,1\}$. There, $\boldsymbol{\theta}_{[m]}$ denotes a scalar parameter from the scalar parameter set $\Theta_{n,m}$ of hypothesis n . An example for a family of PDFs can be given by the normal distribution $f_{\mathcal{N},Y}(\cdot; \mu_Y > 1, \sigma_Y^2 \in \{1,2\})$, where the family consists of all normal distribution with mean parameter larger than one and variance parameter of one and two. For better readability, we will shorten the notation for family of PDFs at appropriate points by

$$f_{Y,\Theta_n}(\cdot) := \{f_Y(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta_n\}, n \in \{0,1\}. \quad (2.5.4)$$

Thus, we can describe the **binary hypotheses with families of PDFs** by

$$\mathcal{H}_0 : Y \sim f_{Y,\Theta_0}(\cdot), \quad (2.5.5)$$

$$\mathcal{H}_1 : Y \sim f_{Y,\Theta_1}(\cdot). \quad (2.5.6)$$

Since, the sets Θ_0 and Θ_1 are complementary, the two hypotheses \mathcal{H}_0 and \mathcal{H}_1 follow different PDFs. In general, only few parameters of the parameter vector $\boldsymbol{\theta}$ differ between the PDFs of the hypotheses. To differentiate between

³An example parameter vector is $\boldsymbol{\theta} = [0 \ 1]^\top$, which corresponds to $\mu = 0$ and $\sigma^2 = 1$ if the normal distribution $f_{\mathcal{N}}(\cdot; \boldsymbol{\theta})$ is selected.

Choice \ Event	\mathcal{H}_0 is true	\mathcal{H}_1 is true
accept \mathcal{H}_0	correct decision (detection) $P_d = P(\Lambda < \rho \mid \mathcal{H}_0)$	type 2 error (false alarm) $P_{fa} = P(\Lambda < \rho \mid \mathcal{H}_1)$
reject \mathcal{H}_0	type 1 error (missed detection) $1 - P_d = P(\Lambda \geq \rho \mid \mathcal{H}_0)$	correct decision (no detection) $1 - P_{fa} = P(\Lambda \geq \rho \mid \mathcal{H}_1)$

Table 2.2: Terminology of binary hypothesis testing, where Λ describes the log-likelihood ratio function as defined in (2.5.9) and (2.5.14).

parameters of direct interest and not of direct interest for the hypothesis test, we introduce the term **nuisance parameter**. Thus, parameters required, but not of immediate interest for the hypothesis test are called nuisance parameters [CB01, Page 378]. In the case of a hypothesis test of $f_{\mathcal{N},Y}(\cdot; \mu_Y = 0, \sigma_Y^2)$ against $f_{\mathcal{N},Y}(\cdot; \mu_Y = 1, \sigma_Y^2)$, the parameter σ_Y^2 is the nuisance parameter.

The decision between two hypotheses results in four cases, as depicted in Table 2.2. There, two cases represent the true state, where the underlying event is correctly accepted. The other two cases denote the incorrect state, namely type 1 and type 2 error, where the underlying event is falsely rejected. Note, that the definition of type 1 and type 2 error may differ in literature, since it depends on the considered scenario. *As performance metrics of the binary hypothesis test, the **detection probability** P_d (also called *detection rate*) and the **probability of false alarm** P_{fa} (also called *false alarm rate*) are commonly used*

$$P_d := P(\text{accept } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is true}), \quad (2.5.7)$$

$$P_{fa} := P(\text{accept } \mathcal{H}_0 \mid \mathcal{H}_1 \text{ is true}). \quad (2.5.8)$$

A widely used measure to describe the performance of statistical hypothesis tests is the receiver operating characteristic (ROC), where the false alarm rate versus the detection rate is plotted. More details on ROCs are found in [Tre01, Chapter 2.2.2].

In general, a hypothesis is classified as simple or composite [LR05, Chapter 3.2]. For a *simple hypothesis* the parameter vector of the underlying probability density function $f_Y(\cdot; \theta)$ is completely specified. In particular,

$\theta \in \Theta$ with the parameter set $\Theta = \{\theta\}$ consisting of one element only. Thus, the simple hypothesis has only one state and the exact probability density function is known. One example PDF of the simple hypothesis is the Gamma distribution of (A.7.19) with $f_{\Gamma}(\cdot; \sigma^2 = 1, 1)$.

In contrast to that, a *composite hypothesis* follows a family of probability density functions $f_{Y,\Theta}(\cdot)$. In particular, the parameter vector $\theta \in \Theta$, where the parameter set Θ consists of multiple elements $|\Theta| > 1$. Thus, a composite hypothesis consists of multiple states and therefore, the exact probability density function of a composite hypothesis is in general unknown.

Consequently, a *simple hypothesis test* decides between simple hypotheses only, while a *composite hypothesis test* performs a decision, where at least one hypothesis is a composite hypothesis.

A statistical hypothesis test shall maximize the detection probability P_d for a given probability of false alarm P_{fa} . The UMP test defines a statistical hypothesis test, which achieves the highest detection probability P_d among all possible tests for any given false alarm rate P_{fa} . However, UMP tests do not exist in general. Even for the class of binary hypothesis tests, UMP tests exist only for some distinct cases.

The theorems in the following sub-sections, the Neyman-Pearson lemma in Section 2.5.1 as well as the Karlin-Rubin theorem in Section 2.5.2, will provide the cases as well as the requirements to obtain a UMP test and are valid for probability density functions in general.

2.5.1 Neyman-Pearson Lemma

The Neyman-Pearson lemma described for example in [LR05, Theorem 3.2.1] and [CB01, Theorem 8.3.12] states that the uniformly most powerful binary simple hypotheses test is the likelihood ratio test. The Neyman-Pearson lemma was initially published by Jerzy Neyman and Egon Pearson in 1933 [NP33].

A likelihood ratio test is based on the likelihood ratio and accepts or rejects the null hypothesis depending on a threshold, which is chosen so that a given false alarm rate is met. The likelihood ratio is defined as a ratio of two likelihood functions or in our case two probability density functions. Usually, the natural logarithm of the likelihood ratio is used, namely the log-likelihood ratio. Consequently, a log-likelihood ratio test is based on the log-likelihood ratio.

For simple binary hypothesis test we have $|\Theta_n| = 1$ for $n \in \{0, 1\}$. Thus, the PDF of each hypothesis is exactly known. *Given some measurement y the **log-likelihood ratio function**⁴ of two probability density functions is*

⁴ Note that in the thesis the LLR is defined inverse to the general literature.

defined by

$$\Lambda(y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) := \ln \left(\frac{f_Y(y; \boldsymbol{\theta}_1)}{f_Y(y; \boldsymbol{\theta}_0)} \right), \boldsymbol{\theta}_n \in \Theta_n \quad (2.5.9)$$

with $n \in \{0, 1\}$. There, the PDF with parameter vector $\boldsymbol{\theta}_0$ represents the null hypothesis \mathcal{H}_0 , see (2.5.1). Consequently, $\boldsymbol{\theta}_1$ reflects the alternative hypothesis \mathcal{H}_1 , see (2.5.2).

The log-likelihood ratio test based on log-likelihood ratio function is equivalent to the likelihood ratio test, since maximizing the log-likelihood ratio is equivalent to maximizing the likelihood ratio. In particular, the natural logarithm $\ln(\cdot)$ is a **strictly increasing function**⁵ and the monotonic behavior of the likelihood ratio is preserved for the log-likelihood ratio (LLR). Thus, the log-likelihood ratio test of two simple hypotheses fulfills the requirements of the Neyman-Pearson lemma. Consequently, *the decision on accepting or rejecting the null hypothesis \mathcal{H}_0 given measurement y results from the **log-likelihood ratio test** by*

$$\Lambda(y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) < \rho : \text{accept } \mathcal{H}_0, \quad (2.5.10)$$

$$\Lambda(y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) \geq \rho : \text{reject } \mathcal{H}_0. \quad (2.5.11)$$

The threshold ρ is chosen, such that a given false alarm rate P_{fa} is achieved for drawing a random variable Y with

$$\rho = \underset{\rho'}{\operatorname{argmax}} \quad P(\Lambda(Y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) < \rho' \mid \mathcal{H}_1) \leq P_{\text{fa}}. \quad (2.5.12)$$

There, a threshold ρ is desired, such that $P(\Lambda(Y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) < \rho \mid \mathcal{H}_0) \stackrel{!}{=} P_{\text{fa}}$. Inversely, the threshold ρ for a given detection rate P_{d} can be obtain by

$$\rho = \underset{\rho'}{\operatorname{argmax}} \quad P(\Lambda(Y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) < \rho' \mid \mathcal{H}_0) \leq P_{\text{d}}. \quad (2.5.13)$$

2.5.2 Karlin-Rubin Theorem

While the Neyman-Pearson lemma of the last Section 2.5.1 is applicable on simple hypotheses only, the Karlin-Rubin theorem can also operate on composite hypotheses. Thus, the Karlin-Rubin theorem introduces uniformly most powerful tests for a broader range of cases.

In contrast to simple hypotheses, composite hypotheses consists of a family of probability density functions. Consequently, the log-likelihood ratio function depends on family of probability density functions and its calculation is generally not trivial. Therefore, the log-likelihood ratio function of

⁵The derivative of $\ln(a)$ $da = \frac{1}{a}$ is strictly positive for $a \in \mathbb{R}_{>0}$.

the composite binary hypothesis test is expressed as a ratio of two families of probability density functions given the parameter vectors $\boldsymbol{\theta}_n \in \Theta_n$ and the measurement y . At least one of the parameter sets Θ_n has more than one element, that means $|\Theta_1| \cdot |\Theta_2| > 1$. Thus, *the log-likelihood ratio function is defined as*

$$\Lambda(y, \Theta_0, \Theta_1) := \ln \left(\frac{f_{Y, \Theta_1}(y)}{f_{Y, \Theta_0}(y)} \right). \quad (2.5.14)$$

The Karlin-Rubin theorem by Samuel Karlin and Herman Rubin extended the Neyman-Pearson lemma in 1956 for one-sided (one-tailed) composite hypothesis tests based on monotone likelihood ratios [KR56]. In particular, the Karlin-Rubin theorem states that a one-sided binary composite hypothesis test is uniformly most powerful, if the test can be expressed as a monotone likelihood ratio function. Below we will introduce one-sided tests as well as define the monotone likelihood ratio property. More details on the Karlin-Rubin theorem are given in [CB01, Theorem 8.3.17] and [LR05, Theorem 3.4.1].

At first let us introduce the *one-sided* test [CB01, Page 391]. Given a family of PDFs $f_Y(\cdot; \theta \in \Theta)$ with a single parameter θ , then a one-sided hypothesis is present, if the hypothesis is directional. In particular, the parameter of a one-sided hypothesis is larger than a value, e.g. $\theta > c$, or smaller than a value, e.g. $\theta \leq c$. Consequently, we call a test one-sided, if the tested hypotheses are both one-sided. Note, that testing $f_Y(\cdot; \theta = 0)$ against $f_Y(\cdot; \theta > 0)$ with $\theta \in \mathbb{R}_{\geq 0}$ is also a one-sided test, which will be of use in the later Section 2.6. The one-sided test holds also for hypothesis test with PDF $f_Y(\cdot; \boldsymbol{\theta} \in \Theta)$, where all but one parameter of the parameter vector $\boldsymbol{\theta}$ are nuisance parameters and the parameter of immediate interest is directional.

Now we turn our attention to the *monotone likelihood ratio* property. The monotone likelihood ratio property is present, if

$$\begin{aligned} \exp(\Lambda(y_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1)) &\geq \exp(\Lambda(y_1, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1)), \\ \text{for every } y_0 > y_1, \forall \boldsymbol{\theta}_n \in \Theta_n, n \in \{0, 1\}. \end{aligned} \quad (2.5.15)$$

Since a logarithm $\log_a(f(\cdot))$ of a real positive function $f(\cdot) \in \mathbb{R}_{\geq 0}$ to a real positive base $a \in \mathbb{R}_{\geq 0}$ does not change the monotonic behavior of $f(\cdot)$, a monotone likelihood ratio function is also present, if

$$\begin{aligned} \Lambda(y_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) &\geq \Lambda(y_1, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1), \\ \text{for every } y_0 > y_1, \forall \boldsymbol{\theta}_n \in \Theta_n, n \in \{0, 1\}. \end{aligned} \quad (2.5.16)$$

Furthermore, *if the monotone likelihood ratio function is strictly monotone, then the **inverse log-likelihood ratio function** exists and is defined as*

$$\Lambda^{-1}(\rho, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) := \max (y \in \mathbb{R} : \rho \geq \Lambda(y, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1)). \quad (2.5.17)$$

Equation (2.5.17) will be of particular use to obtain the UMP test in the later Section 2.6.2.

Thus, *a **one-sided composite binary hypothesis test** based on a monotone log-likelihood ratio function fulfills the requirements of the Karlin-Rubin theorem and is defined as*

$$\Lambda(y, \Theta_0, \Theta_1) < \rho : \text{accept } \mathcal{H}_0, \quad (2.5.18)$$

$$\Lambda(y, \Theta_0, \Theta_1) \geq \rho : \text{reject } \mathcal{H}_0. \quad (2.5.19)$$

The threshold ρ holds for the given family of probability density functions and therefore depends on the parameter sets Θ_0 as well as Θ_1 . Note, that the threshold ρ is a scalar constant and has to hold for all parameter sets. The threshold ρ is chosen, such that a given false alarm rate P_{fa} is achieved with

$$\rho = \underset{\rho'}{\operatorname{argmax}} \quad P(\Lambda(Y, \Theta_0, \Theta_1) < \rho' \mid \mathcal{H}_1) \leq P_{\text{fa}}. \quad (2.5.20)$$

Thus, the threshold ρ has to be chosen, such that a given false alarm rate P_{fa} is achieved for the given family of probability density functions and is fixed for the complete one-sided composite binary hypothesis test. If the PDF of the parameters $\boldsymbol{\theta}_n \in \Theta_n$ for $n \in \{0, 1\}$ is known, the statistics can be used for calculation of the threshold. Inversely, the threshold ρ for a given detection rate P_d can be obtain by

$$\rho = \underset{\rho'}{\operatorname{argmax}} \quad P(\Lambda(Y, \Theta_0, \Theta_1) < \rho' \mid \mathcal{H}_0) \leq P_d. \quad (2.5.21)$$

However, in general the determination of the threshold to achieve the desired performance criterion like false alarm rate or detection rate is not trivial.

2.5.3 Constant False Alarm Rate Property

Besides a UMP test, another highly desirable property is constant false alarm rate, which is generally also easier to achieve. A statistical hypothesis test is called to have constant false alarm rate, if the false alarm rate P_{fa} is independent of the unknown parameters $\boldsymbol{\theta}_1 \in \Theta_1$ of the alternative hypothesis \mathcal{H}_1 [SD91; Láz12]. Thus, a simple hypothesis test achieves constant

false alarm rate because the parameters of the hypotheses are exactly given. In general, the constant false alarm rate property is achieved by adjusting the hypothesis test, so that the desired false alarm rate P_{fa} is obtained under the current parameter set of the scenario. For example, the expected signal power of the hypothesis test can be adapted according to the distance between transmitter and receiver to achieve constant false alarm rate.

Tests with constant false alarm rate are highly desirable in scenarios where the PDF are not or only partially known. Especially in these scenarios, the performance of the tests can be increased while keeping the desired false alarm rate by tracking the unknown parameters and adjust the parameters of the hypothesis test accordingly. Furthermore, it may be desirable that the false alarm rate is achieved independent of the parameters within a family of distributions. Taking the example above, the false alarm rate shall be kept independent of the distance between transmitter and receiver. In general, the constant false alarm rate property is also referred to a statistical hypothesis test, where the false alarm rate P_{fa} is approximately kept constant [BL98, Page 91].

A test with constant false alarm rate property may be easier to achieve than a uniformly most powerful (UMP) test. However, a constant false alarm rate achieved by tracking the unknown parameters is in general not a uniformly most powerful (UMP) test. Moreover, the constant false alarm rate property is generally contradictory to UMP tests, since the constant false alarm rate is achieved for each parameter set within the family of distributions. In contrast to that, a UMP test achieves a desired false alarm rate for the complete family of distributions altogether. Thus, while a UMP test can achieve different false alarm rate for each parameter set within the family of distributions, which "sums up" to the desired false alarm rate, tests with constant false alarm rate achieve the desired false alarm rate by setting each parameters set within the family of distributions to the desired false alarm rate. In summary, the constant false alarm rate property enables tests to achieve a desired false alarm rate where a UMP test is not desired or cannot be applied, in general with the cost of inferior performance in comparison to a UMP test.

2.6 Signal Detection

Binary hypothesis testing is used in signal detection theory to test the presence of a transmitted signal u against its absence in a noisy environment. Here, in contrast to the transmission model $w = h \cdot u + e$ presented in (2.4.7), we focus on a simplified model $w = u + e$, where u can be regarded solely as the transmitted signal or as a joint variable ($h \cdot u$). Later in this Section we

also consider u to follow a certain (but unspecified) distribution. There, the results on the distribution are particular applicable on a joint variable ($h \cdot u$) as used in the transmission model, cf. (2.4.2) and (2.4.7).

The general aim of this Section is to provide the hypothesis test for the signal detection procedure of the gray space detection scheme in Chapter 4. We show that the provided composite hypothesis test is a UMP test, which is therefore also valid for the provided simple hypothesis test. Furthermore, in extension to Section 2.5, we will elaborate how the threshold ρ is obtained to achieve a given performance criterion (false alarm rate or detection rate). There, like in Section 2.5, we will consider the two classes of hypothesis tests: (i) the simple hypothesis test in Section 2.6.1 and (ii) the composite hypothesis test Section 2.6.2. In summary, we will provide a UMP test including the calculation of the threshold ρ to achieve desired performance criteria.

To test the presence of a signal, we can describe the null hypothesis \mathcal{H}_0 and the alternative hypothesis \mathcal{H}_1 by

$$\mathcal{H}_0 : \text{mere noise (signal absent)}, \quad (2.6.1)$$

$$\mathcal{H}_1 : \text{signal} + \text{noise (signal present)}. \quad (2.6.2)$$

Consequently, a receiver is faced with a binary hypothesis testing problem and has to decide between the two events

$$w = \begin{cases} e & : \mathcal{H}_0 \\ u + e & : \mathcal{H}_1, \end{cases} \quad (2.6.3)$$

where $e \sim \text{f}_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_e^2)$ denotes the circularly-symmetric complex normal noise. The complex transmission signal $u \in \mathbb{C}$ follows an unknown probability density function (PDF). Note that in contrast to (2.4.2) we consider here simply u . However, u can also be modeled as a joint variable $\hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]}$ from (2.4.7).

The outcome of the decision at the receiver results in four cases, as depicted in Table 2.2. The detection probability P_d (correct decision that no signal is present) and the probability of false alarm P_{fa} (incorrect decision that no signal is present) describe the performance of the decision process at the receiver. Commonly the performance is visualized via receiver operating characteristic curves, as applied for example in [Her98].

To distinguish between the two hypotheses \mathcal{H}_0 and \mathcal{H}_1 we exploit the underlying PDF. While for \mathcal{H}_0 the expectation of the received signal W is $E[W = w | \mathcal{H}_0] = 0$, for \mathcal{H}_1 we have $E[W = w | U = u, \mathcal{H}_1] = u$, where w and u are the complex realization of the complex random variables W and U ,

respectively. The complex mean parameter $\mu_w := \mathbb{E}[W = w]$ describes the expectation of the received signal W . Thus, the two hypotheses of (2.6.3) are described via the proper complex normal distribution (cf. Appendix A.2) by

$$\mathcal{H}_0 : W \sim \text{f}_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_e^2), \quad (2.6.4)$$

$$\mathcal{H}_{1,c} : W \sim \text{f}_{\mathbb{C}\mathcal{N}}(\cdot; \mu_w, \sigma_e^2), \mu_w = u \neq 0, \quad (2.6.5)$$

where σ_e^2 denotes the known noise variance. Since the mean parameter $\mu_w \neq 0$ of hypothesis $\mathcal{H}_{1,c}$ comprises a range of values and not a single value, hypothesis (2.6.5) is composite. Furthermore, since the complex normal distribution is defined on the whole complex plane $W \in \mathbb{C}$, (2.6.5) does not describe a one-sided hypothesis. However, to fulfill the Karlin-Rubin theorem and achieve a uniformly most powerful test, the composite binary hypothesis test has to be one-sided. Therefore, to obtain a one-sided composite binary hypothesis test, the absolute squared received symbol is used as the measurement $z = |w|^2$. Thus, *a receiver is faced with a **binary hypothesis testing problem** between the two events*

$$z = \begin{cases} |e|^2 & : \mathcal{H}_0 \\ |u + e|^2 & : \mathcal{H}_1. \end{cases} \quad (2.6.6)$$

Since the received signal w is a complex variable with statistically independent and identically distributed real and imaginary part, the absolute squared variable has two degrees of freedom and can be modeled as $Z = |W|^2$, cf. (A.9.13). Consequently, *the scaled non-central chi-squared distribution with two degrees of freedom* (cf. Appendix A.9) *describes the distribution of the measurement z with*

$$|W|^2 = Z \sim \text{f}_{\chi_{sn}^2}(\cdot; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2), \quad (2.6.7)$$

where $\mu_{\chi_{sn}^2} := \frac{1}{2}|\mu_w|^2 = \frac{1}{2}|u|^2$ denotes the non-central parameter of the scaled non-central chi-squared distribution, cf. (A.9.1), and σ_e^2 denotes the known noise variance. While the null hypothesis \mathcal{H}_0 is described by the exact value $\mu_{\chi_{sn}^2} = 0$, the alternative hypothesis $\mathcal{H}_{1,c}$ comprises the family of scaled non-central chi-squared distributions with two degrees of freedom and the unspecific non-central parameter of $\mu_{\chi_{sn}^2} \neq 0$. Thus, for the null hypothesis \mathcal{H}_0 in (2.6.6) the PDF of (2.6.7) can be simplified to the Gamma distribution (cf. Appendix A.7), since the non-central parameter $\mu_{\chi_{sn}^2}$ is exactly zero. Furthermore, since the non-central parameter is defined on the positive real line only, i.e. $\mu_{\chi_{sn}^2} \in \mathbb{R}_{\geq 0}$, the non-central parameter of $\mathcal{H}_{1,c}$ is

sufficiently described by $\mu_{\chi_{sn}^2} > 0$. Consequently, *given the events of (2.6.6) we obtain the two hypotheses, the **simple null hypothesis** \mathcal{H}_0 and the **composite alternative hypothesis** $\mathcal{H}_{1,c}$, by*

$$\mathcal{H}_0 : Z \sim f_{\Gamma}(\cdot; \sigma_e^2, 1), \quad (2.6.8)$$

$$\mathcal{H}_{1,c} : Z \sim f_{\chi_{sn}^2}(\cdot; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2), \mu_{\chi_{sn}^2} > 0. \quad (2.6.9)$$

Similar to (2.6.4) and (2.6.5), the null hypothesis \mathcal{H}_0 (2.6.8) is simple and the alternative hypothesis $\mathcal{H}_{1,c}$ (2.6.9) is composite. However, in contrast to $\mathcal{H}_{1,c}$ (2.6.5), the family of scaled non-central chi-squared distributions is described by the non-central parameter $\mu_{\chi_{sn}^2} > 0$. Thus, $\mathcal{H}_{1,c}$ represents a one-sided composite hypothesis. Furthermore, *we define the **simple alternative hypothesis** $\mathcal{H}_{1,s}$ with*

$$\mathcal{H}_{1,s} : Z \sim f_{\chi_{sn}^2}(\cdot; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2), \mu_{\chi_{sn}^2} = c, \quad (2.6.10)$$

*if the non-central parameter $\mu_{\chi_{sn}^2}$ is exactly described by a constant c . Therefore, testing the null hypothesis \mathcal{H}_0 (2.6.8) against the simple alternative hypothesis $\mathcal{H}_{1,s}$ (2.6.10) results in a simple binary hypothesis test. In contrast to that, a one-sided composite binary hypothesis test is present for testing the null hypothesis \mathcal{H}_0 (2.6.8) against the composite alternative hypothesis $\mathcal{H}_{1,c}$ (2.6.9). Since the PDF of the composite and the simple alternative hypothesis differ only in choice of the non-central parameter $\mu_{\chi_{sn}^2}$, the log-likelihood ratio function is the same. Thus, **the log-likelihood ratio function is calculated by applying the two probability density function from (2.6.8) and (2.6.9) into (2.5.14), resulting in***

$$\begin{aligned} \Lambda(z, \mu_{\chi_{sn}^2}) &= \Lambda(z, \mu_{\chi_{sn}^2, \mathcal{H}_0} = 0, \mu_{\chi_{sn}^2} = \mu_{\chi_{sn}^2, \mathcal{H}_1} > 0, \frac{1}{2}\sigma_e^2) \\ &= \ln \left(\frac{f_{\chi_{sn}^2}(z; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2)}{f_{\Gamma}(z; \sigma_e^2, 1)} \right). \end{aligned} \quad (2.6.11)$$

For ease of notation throughout this thesis, we will omit the nuisance parameter σ_e^2 , even though it is required to obtain the LLR. Furthermore, we will write $\Lambda(z, \mu_{\chi_{sn}^2})$ for log-likelihood ratios, where the null hypothesis \mathcal{H}_0 has $\mu_{\chi_{sn}^2, \mathcal{H}_0} = 0$. Consequently, $\mu_{\chi_{sn}^2}$ of $\Lambda(z, \mu_{\chi_{sn}^2})$ reflects the non-central parameter $\mu_{\chi_{sn}^2, \mathcal{H}_1}$ of the alternative hypothesis \mathcal{H}_1 . Thus, $\Lambda(z, \mu_{\chi_{sn}^2})$ is just written as a function of the measurement z and the parameter of interest $\mu_{\chi_{sn}^2}$ of the alternative hypothesis \mathcal{H}_1 .

Commonly, the Wilks theorem [Wil38] can be used to simplify the log-likelihood ratio function. Shortly summarized, the Wilks theorem showed that a log-likelihood ratio function constructed by two PDFs with a degree of

freedom approaching infinity will be asymptotically chi-squared distributed. Since the log-likelihood ratio function in (2.6.11) is only constructed by PDFs with two degrees of freedom, the Wilks theorem cannot be applied. However, to show that the log-likelihood ratio function is monotone we need further simplification of (2.6.11). Therefore, we substitute the PDFs by simpler PDFs with help of (A.7.19) and (A.9.14) and obtain

$$\begin{aligned}
 \Lambda(z, \mu_{\chi_{sn}^2}) &= \ln \left(\frac{f_{\chi_{nc}^2} \left(\frac{2z}{\sigma_e^2}; 2, \frac{2\mu_{\chi_{sn}^2}}{\sigma_e^2} \right)}{f_{\chi^2} \left(\frac{2z}{\sigma_e^2}; 2 \right)} \right) \\
 &= \ln \left(\frac{\exp \left(-\frac{\sigma_e \mu_{\chi_{sn}^2}}{\sqrt{8}} - \frac{z}{\sigma_e^2} \right) \sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z)^n}{2^{2n+1} \cdot \sigma_e^n \cdot (n!)^2}}{\frac{1}{2} \exp \left(-\frac{z}{\sigma_e^2} \right)}} \right) \\
 &= -\frac{\sigma_e \mu_{\chi_{sn}^2}}{\sqrt{8}} + \ln \left(\sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z)^n}{(4\sigma_e)^n \cdot (n!)^2} \right), \tag{2.6.12}
 \end{aligned}$$

with $\mu_{\chi_{sn}^2} = \frac{1}{2} |u|^2$. The inequality to prove the monotone behavior of (2.6.12) is given by

$$\Lambda(z_0, \mu_{\chi_{sn}^2}) > \Lambda(z_1, \mu_{\chi_{sn}^2}), \text{ for every } z_0 > z_1, \forall \mu_{\chi_{sn}^2} \in \mathbb{R}_{>0}, \tag{2.6.13}$$

with $z_0, z_1 \in \mathbb{R}_{\geq 0}$. In particular, applying (2.6.12) in (2.6.13) we obtain

$$\begin{aligned}
 \sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z_0)^n}{(4\sigma_e)^n \cdot (n!)^2} &> \sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z_1)^n}{(4\sigma_e)^n \cdot (n!)^2} \\
 \sum_{n=0}^{\infty} \left(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \right)^n \cdot \overbrace{\frac{(z_0)^n - (z_1)^n}{(4\sigma_e)^n \cdot (n!)^2}}^{\in \mathbb{R}_{>0}} &> 0. \tag{2.6.14}
 \end{aligned}$$

where for every $z_0 > z_1$ and $\forall \mu_{\chi_{sn}^2} > 0$ the left side of the inequality is always larger than the right side in each summation step. Thus, the LLR function (2.6.11) is a strictly increasing monotone function. Consequently, the LLR function (2.6.11) fulfills the monotone likelihood ratio property of (2.5.16) and the one-sided composite binary hypothesis test is UMP.

	Detection Rate	False Alarm Rate
\mathcal{H}_0 vs. $\mathcal{H}_{1,s}$	$z \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\geq}} -\sigma_e^2 \cdot \ln(1 - P_d)$ (2.6.25)	$z \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\geq}} F_{\chi_{sn}^2}^{-1}(P_{fa}; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2)$ (2.6.23)
\mathcal{H}_0 vs. $\mathcal{H}_{1,c}$	$\Lambda(z, \mu_{\chi_{sn}^2}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,c}}{\geq}} \rho_{c0,\Lambda}(P_d, N)$ (2.6.43)	$\Lambda(z, \mu_{\chi_{sn}^2}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,c}}{\geq}} \rho_{c1,\Lambda}(P_{fa}, N)$ (2.6.36)

Table 2.3: Overview of log-likelihood ratio test to achieve a certain *detection* or *false alarm* rate for *simple* or *composite* binary hypothesis test.

The one-sided composite ($\mu_{\chi_{sn}^2} > 0$) and the simple binary ($\mu_{\chi_{sn}^2} = c$) **log-likelihood ratio test** accepts the null hypothesis \mathcal{H}_0 if

$$\Lambda(z, \mu_{\chi_{sn}^2}) \underset{\mathcal{H}_0}{<} \rho. \quad (2.6.15)$$

There, we use the inequality symbol $\underset{\mathcal{H}_0}{<}$ to indicate the condition where the null hypothesis \mathcal{H}_0 is accepted. In general, the threshold ρ is selected, such that a desired performance criterion like false alarm rate P_{fa} or detection rate P_d is achieved. Note that the desired performance shall be achieved for the scalar threshold ρ applied for all $\mu_{\chi_{sn}^2} > 0$. Finding the threshold ρ for a desired performance criterion is not trivial.

In the following two subsections we will provide some insight on how to obtain this threshold ρ . At first we turn our attention to the simple binary hypothesis test in Section 2.6.1. Afterwards, given the performance criteria, false alarm rate or the detection rate, we calculate the threshold for the one-sided composite binary hypothesis tests in Section 2.6.2. Table 2.3 gives an overview of the different binary hypothesis tests derived in the following.

2.6.1 Simple Log-Likelihood Ratio Test

In case of the simple binary hypothesis test, the simple null hypothesis \mathcal{H}_0 of (2.6.8) is tested against the simple alternative hypothesis $\mathcal{H}_{1,s}$ of (2.6.10). Thus, \mathcal{H}_0 is accepted according to (2.5.10) if

$$\Lambda(z) \underset{\mathcal{H}_0}{<} \rho_{s,\Lambda}, \quad (2.6.16)$$

where $\Lambda(z) = \Lambda(z, \mu_{\chi_{sn}^2} = c)$ and $\rho_{s,\Lambda}$ denotes the threshold for the simple binary hypothesis test. There are two options to select the threshold $\rho_{s,\Lambda}$, (i) the threshold $\rho_{s,\Lambda} = \rho_{s1,\Lambda}(P_{fa})$ to achieve a desired *false alarm* rate or (ii) the threshold $\rho_{s,\Lambda} = \rho_{s0,\Lambda}(P_d)$ to achieve a desired *detection* rate. In the following, we will at first review (2.6.16) and simplify the simple binary hypothesis test. Afterwards, we will elaborate how to obtain the threshold $\rho_{s1,\Lambda}(P_{fa})$ and the threshold $\rho_{s0,\Lambda}(P_d)$.

Let us start by reviewing and simplifying (2.6.16) for the case of simple hypothesis test. Since we have a simple hypothesis test, we can replace the scalar threshold $\rho_{s,\Lambda}$ by the log-likelihood ratio function (2.6.12) with the scalar threshold $\rho_{s,z}$ regarding the measurement z . Thus, we have

$$\rho_{s,\Lambda} = \Lambda(\rho_{s,z}) = -\frac{\sigma_e \mu_{\chi_{sn}^2}}{\sqrt{8}} + \ln \left(\sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot \rho_{s,z})^n}{(4\sigma_e)^n \cdot (n!)^2} \right). \quad (2.6.17)$$

Applying the advertised threshold (2.6.17) and the LLR regarding the measurement z in (2.6.12) into the log-likelihood ratio test (2.6.16), we obtain

$$\ln \left(\sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z)^n}{(4\sigma_e)^n \cdot (n!)^2} \right) < \ln \left(\sum_{n=0}^{\infty} \frac{(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot \rho_{s,z})^n}{(4\sigma_e)^n \cdot (n!)^2} \right), \quad (2.6.18)$$

where the equal term is already canceled out. After taking the exponential on both sides and rearrange the inequality, we obtain

$$0 < \sum_{n=0}^{\infty} \frac{\overbrace{\left(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot \rho_{s,z} \right)^n}^{\in \mathbb{R}_{\geq 0}} - \overbrace{\left(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z \right)^n}^{\in \mathbb{R}_{\geq 0}}}{(4\sigma_e)^n \cdot (n!)^2}. \quad (2.6.19)$$

The two terms of the numerator at (2.6.19) are positive, because all variables of the numerator are positive. Since from $a \geq b$ it follows $a^n \geq b^n, \forall n$ with $a, b \in \mathbb{R}_{\geq 0}$, the inequality (2.6.19) is true if

$$\left(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot \rho_{s,z} \right)^n > \left(\sqrt{2} \cdot \mu_{\chi_{sn}^2} \cdot z \right)^n \quad \forall n. \quad (2.6.20)$$

Thus, the inequality (2.6.19) is true if $z < \rho_{s,z}$. Consequently, we simplified the log-likelihood ratio test of simple hypotheses in (2.6.16) to a test on the actual measurement z only. *The simple binary hypothesis test accepts the null hypothesis \mathcal{H}_0 if*

$$z <^{\mathcal{H}_0} \rho_{s,z}. \quad (2.6.21)$$

Since we just test on the measurement z , the threshold $\rho_{s,z}$ merely depends on the PDF of the measurement under the tested hypothesis. While the threshold $\rho_{s1,z}(P_{fa})$ to achieve certain false alarm rate P_{fa} is determined under the simple alternative hypothesis $\mathcal{H}_{1,s}$ (2.6.10), the threshold $\rho_{s0,z}(P_d)$ to achieve certain detection rate P_d is determined under the null hypothesis \mathcal{H}_0 (2.6.8). Urkowicz already presented this result on the simple hypothesis test in 1967 [Urk67].

Based on the simple binary hypothesis test derived in (2.6.21) we can now derive the threshold $\rho_{s,z}$ to achieve a certain performance criterion. We begin by deriving the threshold $\rho_{s1,z}(P_{fa})$ to achieve a desired *false alarm* rate P_{fa} and continue with the threshold $\rho_{s0,z}(P_d)$ to achieve a desired *detection* rate P_d .

False Alarm Rate Threshold

Since the simple binary hypothesis test is given by (2.6.21), the threshold $\rho_{s1,z}(P_{fa})$ to achieve a desired *false alarm* rate P_{fa} just depends on the distribution of the measurement z under the simple alternative hypothesis $\mathcal{H}_{1,s}$ (2.6.10). In particular, the measurement z follows $f_{\chi_{sn}^2}(\cdot)$ for $\mathcal{H}_{1,s}$. Thus, we obtain the **false alarm threshold** $\rho_{s1,z}(P_{fa})$ in conjunction with (2.5.12) by the scaled non-central chi-squared inverse cumulative distribution function (ICDF)

$$\begin{aligned} \rho_{s1,z}(P_{fa}) &= \underset{\rho'}{\operatorname{argmax}} (P_{fa} \geq P(z < \rho' \mid \mathcal{H}_{1,s})) \\ &= \underset{\rho' \in \mathbb{R}_{\geq 0}}{\operatorname{argmax}} \left(P_{fa} \geq F_{\chi_{sn}^2} \left(\rho'; 2, \mu_{\chi_{sn}^2}, \frac{1}{2} \sigma_e^2 \right) \right) \\ &= F_{\chi_{sn}^2}^{-1} \left(P_{fa}; 2, \mu_{\chi_{sn}^2}, \frac{1}{2} \sigma_e^2 \right). \end{aligned} \quad (2.6.22)$$

The scaled non-central chi-squared ICDF with two degrees of freedom is given in (A.9.15) of Appendix A.9. Furthermore, an approximations of the scaled non-central chi-squared ICDF is given in (A.9.10) of Appendix A.9, which can be used to obtain an approximated closed-form expression. Consequently, *given the false alarm rate we obtain the **simple binary hypothesis test**, testing the null hypothesis (2.6.8) against the simple alternative hypothesis (2.6.10) and accepting the null hypothesis \mathcal{H}_0 if*

$$z < F_{\chi_{sn}^2}^{-1} \left(P_{fa}; 2, \mu_{\chi_{sn}^2}, \frac{1}{2} \sigma_e^2 \right), \quad (2.6.23)$$

where $\mu_{\chi_{sn}^2} = \frac{1}{2}|u|^2 = c$. The simple hypothesis test (2.6.23) can be used to construct a test with constant false alarm rate property in the case of a composite alternative hypothesis (2.6.9) describing a family of distributions. There, the non-central parameter $\mu_{\chi_{sn}^2}$ is tracked, so that the simple hypothesis test (2.6.23) is conducted on measurement z given the non-central parameter $\mu_{\chi_{sn}^2}$. Consequently, the desired false alarm rate is always met. However, as discussed in Section 2.5.3, the constant false alarm rate property does not guarantee a UMP test. In general, there exists a better test by considering the false alarm rate performance for the complete family of distributions. How to obtain the composite threshold $\rho_{c1,\Lambda}(P_{fa})$ given a desired false alarm rate will be discussed later in Section 2.6.2.

Detection Rate Threshold

The threshold $\rho_{s0,z}(P_d)$ achieving a desired detection rate can be obtained similar to the calculation above for the false alarm rate threshold $\rho_{s1,z}(P_{fa})$. Since (2.6.21) depends on the measurement z only, we merely need to consider the null hypothesis \mathcal{H}_0 from (2.6.8) to describe simple binary hypothesis test given the detection rate P_d . In particular, Thus, *we obtain the detection threshold $\rho_{s0,z}(P_d)$ in conjunction with (2.5.13) by the Gamma inverse cumulative distribution function*

$$\begin{aligned} \rho_{s0,z}(P_d) &= \underset{\rho'}{\operatorname{argmax}} (P_d \geq P(z < \rho' \mid \mathcal{H}_0)) \\ &= \underset{\rho'}{\operatorname{argmax}} (P_d \geq F_{\Gamma}(\rho'; \sigma_e^2, 1)) \\ &= F_{\Gamma}^{-1}(P_d; \sigma_e^2, 1) \\ &= -\sigma_e^2 \ln(1 - P_d), \end{aligned} \tag{2.6.24}$$

where the Gamma ICDF is obtained from (A.7.21) in Appendix A.7. Apparently, the simple threshold $\rho_{s0,z}(P_d)$ for a desired detection rate is independent of the non-central parameter $\mu_{\chi_{sn}^2}$. Consequently, given the desired detection rate P_d we obtain the **simple binary hypothesis test**, testing the null hypothesis (2.6.8) against the simple alternative hypothesis (2.6.10). Thus, the null hypothesis \mathcal{H}_0 is accepted if

$$z \stackrel{\mathcal{H}_0}{<} -\sigma_e^2 \cdot \ln(1 - P_d). \tag{2.6.25}$$

Since (2.6.24) is independent of $\mu_{\chi_{sn}^2}$, the simple binary hypothesis test in (2.6.25) is also independent of the non-central parameter $\mu_{\chi_{sn}^2}$. Thus,

given a desired detection rate the non-central parameter $\mu_{\chi_{sn}^2}$ has not to be known to achieve a uniformly most powerful test between two simple hypotheses. Classical energy detection exploits this fact to detect signals under unknown non-central parameter $\mu_{\chi_{sn}^2}$ [Urk67].

However, in the case of composite hypothesis tests the simplification of (2.6.21) does not hold since we are confronted with a family of distributions. There, the knowledge of $\mu_{\chi_{sn}^2}$ is required to achieve a UMP test. The calculation of the composite threshold $\rho_{c1,\Lambda}$ to achieve such a UMP test will be conducted in the next Section 2.6.2.

2.6.2 One-Sided Composite Log-Likelihood Ratio Test

In contrast to the simple binary hypothesis tests, the one-sided composite binary hypothesis tests is performed on a family of distributions. Therefore, we have to consider alternative composite hypothesis (2.6.9) with the non-central parameter $\mu_{\chi_{sn}^2} > 0$. Thus, for each non-central parameter $\mu_{\chi_{sn}^2} > 0$ we can construct a log-likelihood ratio test fulfilling the requirements of the Neyman-Pearson lemma. There, for each non-central parameter a log-likelihood ratio test with an individual threshold $\rho_{s,z}$ is performed. However, a combination of these tests are not fulfilling requirements of the the Neyman-Pearson lemma for $\mu_{\chi_{sn}^2} > 0$. In other words, there exists a log-likelihood ratio test for a given performance criterion, which achieves at least as good performance (but generally better) as the combined log-likelihood ratio test, where each test fulfills the requirements of the Neyman-Pearson lemma.

In the case of a desired false alarm rate: There exists a log-likelihood ratio test for a given false alarm rate P_{fa} , which achieves at least as high detection rate P_d (but generally higher) as the combined log-likelihood ratio test, where each test fulfills the requirements of the Neyman-Pearson lemma. Thus, to achieve the uniformly most powerful (UMP) test, instead of selecting an individual threshold $\rho_{s1,z}$ for each $\mu_{\chi_{sn}^2}$, the false alarm rate has to be individually chosen for each non-central parameter $\mu_{\chi_{sn}^2}$, such that the overall desired false alarm rate P_{fa} is met. Consequently, a single composite threshold $\rho_{c1,z}(P_{fa})$ on z as in (2.6.22) for the simple binary hypothesis test valid for all $\mu_{\chi_{sn}^2} > 0$ cannot be found⁶. In other words, the individual false alarm rate for each non-central parameter has to be selected in such a way, that the overall desired false alarm rate P_{fa} is met by a single $\rho_{c1,\Lambda}(P_{fa})$ for all $\mu_{\chi_{sn}^2} > 0$. The above statements holds too, if a certain detection rate P_d is desired.

⁶In general, the composite threshold $\rho_{c1,z}(P_{fa})$ on z would be different for each $\mu_{\chi_{sn}^2} > 0$.

Since we cannot simplify the composite hypothesis test as done for the simple hypothesis test in (2.6.21), *the one-sided composite binary hypothesis test accepts the null hypothesis \mathcal{H}_0 according to (2.5.18) if*

$$\Lambda(z, \mu_{\chi_{sn}^2} > 0) \stackrel{\mathcal{H}_0}{<} \rho_{c,\Lambda}. \quad (2.6.26)$$

Before we turn our attention to the calculation of the composite threshold $\rho_{c,\Lambda}$, we will provide a more specific inverse log-likelihood ratio (LLR) than the one given in (2.5.17) for general PDFs. The inverse LLR is later used to calculate the composite threshold $\rho_{c,\Lambda}$. *Since the LLR function in (2.6.11) has strict monotone likelihood ratio property (2.6.14), the inverse LLR function exists and is defined as*

$$\begin{aligned} \Lambda^{-1}(\rho, \mu_{\chi_{sn}^2}) &:= \operatorname{argmax}_{z \in \mathbb{R}_{\geq 0}} (\rho \geq \Lambda(z, \mu_{\chi_{sn}^2})) \\ &= \operatorname{argmax}_{z \in \mathbb{R}_{\geq 0}} \left(\rho \geq \ln \left(\frac{f_{\chi_{sn}^2}(z; 2, \mu_{\chi_{sn}^2}, \frac{1}{2}\sigma_e^2)}{f_{\Gamma}(z; \sigma_e^2, 1)} \right) \right). \end{aligned} \quad (2.6.27)$$

The inverse LLR function selects the maximum measurement z amongst all those values whose LLR function is below or equal to a threshold ρ .

In the following sections, we will derive the composite threshold $\rho_{c1,\Lambda}$ (P_{fa}) for given false alarm rate P_{fa} and the composite threshold $\rho_{c0,\Lambda}$ (P_d) for given detection rate P_d .

False Alarm Rate Threshold

For ease of notation, we will use $\rho_{c1} = \rho_{c1,\Lambda}$ to denote the composite threshold to achieve a desired false alarm rate P_{fa} . Since the false alarm rate is given by the cumulative distribution function of the alternative hypothesis $\mathcal{H}_{1,c}$, in particular $F_{\chi_{sn}^2}(\cdot)$ given in (A.9.15)⁷, we can use the inverse LLR function (2.6.27) to obtain the false alarm rate P_{fa} for a given threshold ρ_{c1} . Therefore, *from (2.5.20) given a desired false alarm rate P_{fa} it follows the calculation of the composite false alarm threshold*

$$\rho_{c1}(P_{fa}) = \operatorname{argmax}_{\rho} (P_{fa} \geq P(\rho)), \quad (2.6.28)$$

$$\text{where } P(\rho) = \int F_{\chi_{sn}^2} \left(\Lambda^{-1}(\rho, a); 2, a, \frac{\sigma_e^2}{2} \right) \cdot f_X(a) \, da, \quad (2.6.29)$$

⁷Measurement z follows $f_{\chi_{sn}^2}(\cdot)$ for $\mathcal{H}_{1,c}$ (2.6.9).

where $f_X(a)$ is the PDF of the non-central parameter $\mu_{\chi_{sn}^2}$. The composite threshold $\rho_{c1}(\text{P}_{fa})$ in (2.6.28) is determined by selecting the maximum threshold ρ , such that the probability $\text{P}(\rho)$ is below or equal the desired false alarm rate P_{fa} . The probability $\text{P}(\rho)$ is obtained by the integral of the cumulative distribution function (CDF) of the composite alternative hypothesis $\mathcal{H}_{1,c}$ weighted with $f_X(a)$ denoting the PDF of the non-central parameter $\mu_{\chi_{sn}^2}$. Thus, given a desired false alarm rate P_{fa} and the inverse LLR function (2.6.27) the **one-sided composite binary hypothesis test** of (2.6.26) can be refined to

$$\begin{aligned} \Lambda(z, \mu_{\chi_{sn}^2} > 0) &\stackrel{\mathcal{H}_0}{<} \rho_{c1}(\text{P}_{fa}) \\ &= \underset{\rho}{\operatorname{argmax}} \left(\text{P}_{fa} \geq \int \text{F}_{\chi_{sn}^2} \left(\Lambda^{-1}(\rho, a); 2, a, \frac{\sigma_e^2}{2} \right) \cdot f_X(a) \, da \right). \end{aligned} \quad (2.6.30)$$

However in general, the PDF $f_X(a)$ of the non-central parameter $\mu_{\chi_{sn}^2}$ is hardly known. Therefore, in practical applications the composite threshold ρ_{c1} cannot be directly obtained from (2.6.28).

However, some observation of the non-central parameter distribution $f_X(a)$ may be available, enabling the estimation of the true CDF.

The **empirical CDF (ECDF)** is one approach to estimate the true CDF from a given sample set. In contrast to other estimates the ECDF is a non-parametric maximum likelihood estimate [Owe01, Chapter 2.1]. In other words, ECDF does not require knowledge of the underlying distribution family. The ECDF is obtained by independently drawing each measurement x_n of the random variable X with $\text{F}_X(a)$. Therefore, the **empirical CDF** is defined as

$$\check{\text{F}}_N(a) := \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{[-\infty, a]}(x_n), \quad (2.6.31)$$

where $\mathbf{1}_{[-\infty, a]}(x_n)$ denotes the indicator function. Following the **Glivenko–Cantelli theorem** [Can33; Gli33], the ECDF with N approaching infinity converges almost surely (a.s.) to its true CDF, thus

$$\lim_{N \rightarrow \infty} \left\| \check{\text{F}}_N(a) - \text{F}_X(a) \right\|_{\infty} = 0 \quad \text{a.s.} \quad (2.6.32)$$

In other words, for $(N \rightarrow \infty)$ i.i.d. observations of $\text{F}_X(a)$ the ECDF of the observations converges almost surely to its true $\text{F}_X(a)$. Consequently, the ECDF approximates the CDF with a step function. Originally, the

Glivenko–Cantelli theorem was formulated in Italian by Glivenko and Cantelli in 1933 [Can33; Gli33]. Further information can also be found in [LR05, Theorem 11.2.17].

Consequently, by applying the ECDF of (2.6.31) on (2.6.29), the **approximated composite false alarm threshold** ρ_{c1} is given by

$$\rho_{c1} (P_{fa}) \approx \rho_{c1} (P_{fa}, N) = \operatorname{argmax}_{\rho \in \mathbb{R}} (P_{fa} \geq P(\rho, N)), \quad (2.6.33)$$

$$\text{where } P(\rho, N) = \frac{1}{N} \sum_{n=1}^N F_{\chi_{sn}^2} \left(\Lambda^{-1}(\rho, x_n); 2, x_n, \frac{\sigma_e^2}{2} \right), \quad (2.6.34)$$

where each measurement x_n is independently drawn from the PDF of the non-central parameter with $X \sim f_X(a)$. Thus, (2.6.34) approximates (2.6.29) by calculating the arithmetic mean of the CDF over all measurements (x_1, \dots, x_N) . As stated by the Glivenko–Cantelli theorem, the approximation (2.6.33) converges to (2.6.29) for N approaching infinity

$$\rho_{c1} (P_{fa}) = \lim_{N \rightarrow \infty} \rho_{c1} (P_{fa}, N). \quad (2.6.35)$$

Thus, the larger N , the better the approximation of the composite threshold $\rho_{c1} (P_{fa})$. Finally, given an approximated composite threshold $\rho_{c1} (P_{fa}, N)$ based on a desired false alarm rate P_{fa} , (2.6.30) can be refined to a **one-sided composite binary hypothesis test** accepting the null hypothesis \mathcal{H}_0 if

$$\begin{aligned} \Lambda(z, \mu_{\chi_{sn}^2}) &\stackrel{\mathcal{H}_0}{<} \rho_{c1} (P_{fa}, N) \\ &= \operatorname{argmax}_{\rho \in \mathbb{R}} \left(P_{fa} \geq \frac{1}{N} \sum_{n=1}^N F_{\chi_{sn}^2} \left(\Lambda^{-1}(\rho, x_n); 2, x_n, \frac{\sigma_e^2}{2} \right) \right). \end{aligned} \quad (2.6.36)$$

Detection Rate Threshold

The composite threshold $\rho_{c0} = \rho_{c0, \Lambda}$ to achieve a desired detection rate P_d is derived similar to the false alarm threshold ρ_{c1} above. However, for calculation of the detection rate threshold ρ_{c0} we consider the null hypothesis \mathcal{H}_0 . Therefore, we can apply the inverse LLR function (2.6.27) on the CDF of the Gamma distribution $F_\Gamma(\cdot)$ given in (A.7.20)⁸. Thus, given a desired

⁸Measurement z follows $f_\Gamma(\cdot)$ for \mathcal{H}_0 (2.6.8).

detection rate P_d it follows from (2.5.21) the calculation of the **composite detection threshold**

$$\rho_{c0}(P_d) = \operatorname{argmax}_{\rho \in \mathbb{R}} (P_d \geq P(\rho)), \quad (2.6.37)$$

$$\text{where } P(\rho) = \int F_{\Gamma}(\Lambda^{-1}(\rho, a); \sigma_e^2, 1) \cdot f_X(a) \, da, \quad (2.6.38)$$

$$= 1 - \int \exp\left(-\frac{\Lambda^{-1}(\rho, a)}{\sigma_e^2}\right) \cdot f_X(a) \, da. \quad (2.6.39)$$

Consequently, given a desired detection rate P_d the **one-sided composite binary hypothesis test** of (2.6.26) can be refined to

$$\Lambda(z, \mu_{\chi_{sn}^2} > 0) \quad (2.6.40)$$

$$\stackrel{\mathcal{H}_0}{<} \operatorname{argmax}_{\rho} \left(P_d \geq 1 - \int \exp\left(-\frac{\Lambda^{-1}(\rho, a)}{\sigma_e^2}\right) \cdot f_X(a) \, da \right).$$

Similar to (2.6.34), the **approximation of the composite threshold** ρ_{c0} is obtained by applying the ECDF of (2.6.31) on (2.6.39). Thus, we have

$$\rho_{c0}(P_d) \approx \rho_{c0}(P_d, N) = \operatorname{argmax}_{\rho \in \mathbb{R}} (P_d \geq P(\rho, N)), \quad (2.6.41)$$

$$\text{where } P(\rho, N) = 1 - \frac{1}{N} \sum_{n=1}^N \exp\left(-\frac{\Lambda^{-1}(\rho, x_n)}{\sigma_e^2}\right), \quad (2.6.42)$$

where each measurement x_n is independently drawn from the PDF of the non-central parameter $\mu_{\chi_{sn}^2}$ with $X \sim f_X(a)$. The convergence result of (2.6.35) also holds for the approximated composite threshold $\rho_{c0}(P_d, N)$. Thus, given an approximated composite threshold $\rho_{c0}(P_d, N)$ based on a desired detection rate P_d , (2.6.40) can be refined to a **one-sided composite binary hypothesis test** accepting the null hypothesis \mathcal{H}_0 if

$$\Lambda(z, \mu_{\chi_{sn}^2}) \stackrel{\mathcal{H}_0}{<} \rho_{c0}(P_d, N) \quad (2.6.43)$$

$$= \operatorname{argmax}_{\rho \in \mathbb{R}} \left(P_d \geq 1 - \frac{1}{N} \sum_{n=1}^N \exp\left(-\frac{\Lambda^{-1}(\rho, x_n)}{\sigma_e^2}\right) \right).$$

The computation of the approximated composite threshold in (2.6.33) and (2.6.41) can be performed by **root-finding algorithms**. In particular,

a well-chosen initial threshold ρ is used to obtain an initial estimate of the performance criterion $P(\rho, N)$, reflecting the desired detection rate P_d or false alarm rate P_{fa} . Afterwards, the [secant method](#) can be applied to converge to the true performance criterion, since we are operating on strict monotone functions.

2.7 Summary

This Chapter introduced the concepts of wireless communications used in the subsequent chapters. Therefore, we discussed the channel model and provided the common performance channels ITU IOA as well as 3GPP EPA and EVA. Furthermore, we established the general system model based on cyclic-prefix OFDM similar to 3GPP LTE and NR. Finally, we derived the basics for the statistical hypothesis test used in Chapter 4, where we will heavily rely on the results for the one-sided composite hypothesis test in (2.6.36). In particular, (2.6.36) will be used to establish a log-likelihood ratio test for desired false alarm rate. The contributions on statistical hypothesis tests are extensions of the test discussed in [WPJ13; WJW+16b].

Chapter 3

Compressed Sensing

3.1 Overview and Contributions

In the recent decade compressed sensing (CS) has drawn immense attention, starting with the foundation laid by the joint work of Candès, Romberg and Tao [CRT06] as well as Donoho [Don06]. In particular, a breakthrough in sampling theory [CRT06] enabled the recovery of sparse signals with a sampling rate below the Nyquist rate. In this Chapter we will introduce the fundamentals of compressed sensing. In particular, we will consider the (partial) discrete Fourier transform as the basis for signal recovery and list possible estimation algorithms. Finally, we will give some error bounds for signal recovery, which are derived using probability theory.

We will start by introducing the notion of sparsity in Section 3.2, which comprises the strict form of sparsity (3.2.1) as well as the relaxed form of compressibility (3.2.5). Therefore, we also introduce the definition for the union of subspaces (3.2.2) as well as the ℓ_p -error of best s -term approximation (3.2.3). Afterwards, we provide a basic view on linear systems in Section 3.3, especially regarding the number of solutions of a linear system of equations. Based on some facts from probability theory, we limit the solution space on certain error bounds (3.3.8). With the foundation on linear systems, we introduce the framework of compressed sensing in Section 3.4. Besides the restricted isometry property (Section 3.4.3), which is a sufficient compressed sensing condition for signal reconstruction under certain error bounds, we also discuss the more tractable measure of mutual coherence (Section 3.4.2) and the sufficient and necessary condition for sparse recovery called the null space property (Section 3.4.1). In the subsequent Section 3.5 we introduce the basic estimation algorithm used within this

thesis. Reduced-rank least squares estimation (Section 3.5.1) is a classical ℓ_2 -norm approach used for general problems of linear systems without further structure like sparsity. To exploit sparsity within a given linear system, the subsequently introduced compressed sensing methods are used: basis pursuit denoising (Section 3.5.2), a convex optimization algorithm, as well as the greedy orthogonal matching pursuit approach (Section 3.5.3). Finally, the (probabilistic) partial Fourier model for compressed sensing is presented in Section 3.6. On the one hand we discuss linear partial Fourier measurements present in the channel transfer function (Section 3.6.1), and on the other hand we study absolute squared partial Fourier measurements present in the spectral channel power (Section 3.6.2). Depending on the measurements, we apply a corresponding channel estimation scheme: pilot-based channel estimation to retrieve the channel impulse response from channel transfer function measurements as well as amplitude-based channel estimation to retrieve the circular autocorrelation of the channel impulse response from spectral channel power measurements. Therefore, based on probability distributions we elaborate error bounds for the estimation algorithms in the previous Section 3.5. In particular, we derive the probability distribution of the measurement error in case of the channel transfer function (3.6.18) and an estimate for the expectation of the measurement error in case of the spectral channel power (3.6.42). The chapter concludes with a short summary on compressed sensing in Section 3.7.

3.2 Sparsity

In this Section, we will introduce the two definitions of sparsity, the strict concept of sparsity and the relaxed version of compressibility.

The straightforward definition of sparsity is as follows: *Let $\mathbf{x} \in \mathbb{C}^M$, then \mathbf{x} is called s -sparse if*

$$s \geq \|\mathbf{x}\|_0 := |\{n : \mathbf{x}_{[n]} \neq 0\}|. \quad (3.2.1)$$

In other words the vector \mathbf{x} describes a signal with at most s non-zero entries. An example for a 1-sparse vector is shown in Figure 3.1. Even though, the ℓ_0 -“norm” is neither norm nor quasinorm, it is frequently used in CS for counting the number of non-zero elements. Furthermore, *we define the set of all s -sparse vectors (union of s -dimensional subspaces) by*

$$\Sigma_s := \{\mathbf{x} : \|\mathbf{x}\|_0 \leq s\} \quad (3.2.2)$$

As an example, for all 2-sparse vectors in \mathbb{R}^3 , Figure 3.2 visualizes the union of 2-dimensional subspaces. Consequently, a 2-sparse vector can only be located on the plane spanned by each pair of axes.

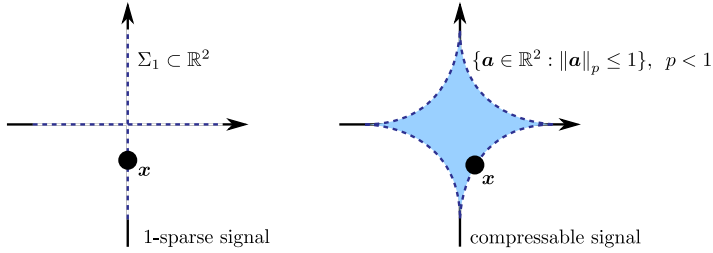


Figure 3.1: Notion of sparsity and compressibility in 2 dimensions.

However, the strict s -sparse definition (3.2.1) is not compatible with many practical settings, where a vector has only few significant entries, but many irrelevant entries as well. One example is the multipath channel described in Section 2.3. At first, to quantify the compressibility of a vector we can use the ℓ_p -error of best s -term approximation to the compressible vector $\mathbf{x} \in \mathbb{C}^M$. Therefore, we define at first the **ℓ_p -error of best s -term approximation** to a vector $\mathbf{x} \in \mathbb{C}^M$ by

$$\gamma_s(\mathbf{x})_p := \inf \left[\|\mathbf{x} - \mathbf{b}\|_p, \mathbf{b} \in \Sigma_s \subset \mathbb{C}^M \right]. \quad (3.2.3)$$

Consequently, the infimum is achieved by an s -sparse vector \mathbf{b} , whose non-zero entries match the entries of \mathbf{x} , which are the s absolute largest entries. The **ℓ_2 -error of best s -term approximation** is limited by [FR13, Theorem 2.5]

$$\gamma_s(\mathbf{x})_2 \leq \frac{1}{2\sqrt{s}} \cdot \|\mathbf{x}\|_1. \quad (3.2.4)$$

There exist different common notions to define a compressible vector. We will refer here to the notion via the ℓ_p -norm presented in [FR13, Chapter 2.1]. A vector \mathbf{x} is called **compressible** (or commonly also *sparse*), if for some small $0 < p < 1$ the vector \mathbf{x} is contained in the ℓ_p -ball

$$\mathbf{x} \in \mathbb{A}_p^M, \quad (3.2.5)$$

where the ℓ_p -ball is defined by

$$\mathbb{A}_p^M := \{\mathbf{a} \in \mathbb{C}^M : \|\mathbf{a}\|_p \leq 1\}. \quad (3.2.6)$$

An example for a compressible $\mathbf{x} \in \mathbb{R}^2$ contained in the ℓ_p -ball is shown in Figure 3.1. Informally, the smaller p the closer is the vector located to the union of s -dimensional subspaces Σ_s .

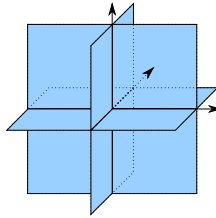


Figure 3.2: Union of 2-dimensional subspaces $\Sigma_2 \subset \mathbb{R}^3$ [EK12, Figure 1.5]: set of all 2-sparse vectors in 3 dimensions.

3.3 Linear Systems

Before introducing the framework of CS in the subsequent Section 3.4, we will briefly introduce different settings of linear systems as well as give some notions on solving them. In particular in noisy measurements we will state some probabilities, such that the solution exists within a certain bound.

Algorithms to estimate an unknown $\mathbf{x} \in \mathbb{C}^M$ from noisy measurement $\mathbf{y} \in \mathbb{C}^N$ through a **linear map** $\Phi \in \mathbb{C}^{N \times M}$ are widely used. **Signal processing** and in particular channel estimation [OA07; LTH+14] are applications, where solving linear equations is required, such that an unknown vector \mathbf{x} can be reconstructed from noisy samples. Here, we consider the complex-valued case. The real-valued case is derived similar.

We start with unaltered measurements, where the **system of linear equations** is written as

$$\mathbf{y} = \Phi \mathbf{x}. \quad (3.3.1)$$

The number of linear equations N are represented by the number of rows of measurement matrix $\Phi \in \mathbb{C}^{N \times M}$. Furthermore, the number of unknown parameters M is represented by the number of columns of measurement matrix Φ . Consequently, the relation between rows and columns of measurement matrix Φ indicates, whether the system of linear equations is **underdetermined** ($N < M$), **balanced** ($N = M$) or **overdetermined** ($N > M$) [Sel13]. However, the above classifications of underdetermined, balanced and overdetermined systems are neither necessary nor sufficient conditions to state the number of solutions of a system of linear equations.

To obtain the number of solutions of (3.3.1), the **Rouché–Capelli theorem** can be applied on matrix Φ given a measurement \mathbf{y} . Three results on the solution space of systems of linear equations are possible: infinite, one (and unique) or zero number of solutions. A system has no solution and is called inconsistent, if $\text{rank}(\Phi) < \text{rank}([\Phi \ \mathbf{y}])$, where $[\Phi \ \mathbf{y}]$ represents the **augmented matrix**. Consequently, a system has at least one solution and

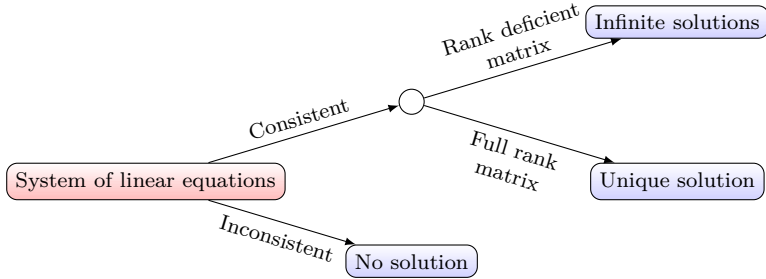


Figure 3.3: Solution space of a system of linear equations.

is called consistent, if $\text{rank}(\Phi) = \text{rank}([\Phi \ y])$. Note, since the row rank and the column rank of any matrix are equal [Mac95], the rank of a matrix describes both, the number of independent rows as well as the number of independent columns. A consistent system can be further distinguished in two matrix types. There, a unique solution of (3.3.1) is present if we have a full rank matrix, i.e. $\text{rank}(\Phi) = \min[N, M]$. Otherwise, we have an infinite number of solutions if the matrix is rank deficient with $\text{rank}(\Phi) < \min[N, M]$. A simple diagram on the possible solution spaces is depicted in Figure 3.3.

Even if the system of linear equations in (3.3.1) has zero or infinite solutions, relaxation or further constraints on the system can lead to a single (and in some cases even unique) solution [Sel13]. For example, solutions of an inconsistent system of linear equations can be found by minimizing the error of the system regarding a selected ℓ_p -norm

$$\tilde{\mathbf{x}} \in \underset{\mathbf{x}}{\text{argmin}} \|\mathbf{y} - \Phi \cdot \mathbf{x}\|_p. \quad (3.3.2)$$

For $p \geq 1$ the solutions of (3.3.2) can be found via a convex program, which can be solved with standard solvers and in particular for $p = 2$ in closed-form. In case of a consistent system of linear equations with a rank deficient matrix, a solution can be found by minimizing the unknown regarding a selected ℓ_p -norm under the constraint induced by the linear equation

$$\tilde{\mathbf{x}} \in \underset{\mathbf{x}}{\text{argmin}} \|\mathbf{x}\|_p, \text{ s.t. } \mathbf{y} = \Phi \mathbf{x}. \quad (3.3.3)$$

Common approaches are the classical least squares (LS) based on the ℓ_2 -norm and compressed sensing (CS) based on the ℓ_1 -norm, which is used as relaxation of the ℓ_0 -“norm”, see Section 3.4. Commonly, classical LS are applied on inconsistent systems, finding a least squares solution which

minimizes the energy of the error. Nevertheless, LS can also be applied on consistent systems, where the energy of the solution \mathbf{x} is minimized. Instead of the energy (ℓ_2 -norm) constraint, CS applies the ℓ_1 -norm to find solutions for systems of linear equations.

As stated at the beginning of this Chapter, measurements are generally error-prone due to noise. The disturbance can be taken into account by extending the system of linear equations in (3.3.1) by an adversarial error $\mathbf{e} \in \mathbb{C}^N$ representing the induced error on the system

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}. \quad (3.3.4)$$

Noisy measurements are present in communication systems, which is also reflected by our transmission model $\hat{\mathbf{w}} = \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} + \hat{\mathbf{n}}$ defined in (2.4.8) of Section 2.4. Consequently, given $\Phi \mathbf{x} = \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}}$, our transmission model can be represented in form of (3.3.4), where Φ represents the discrete Fourier transform matrix.

Since the actual realization of the adversarial error \mathbf{e} is modeled by a random variable, the exact value of \mathbf{e} is usually not known. Thus, solving (3.3.4) directly and retrieving the exact solution is impractical. Therefore, we seek recovery of an estimate $\tilde{\mathbf{x}}$ from noisy measurements. In the overdetermined case (with full rank), (3.3.4) will be generally an inconsistent system, because the addition of an adversarial error \mathbf{e} will cause $\text{rank}(\Phi) < \text{rank}([\Phi \ \mathbf{y}])$ almost surely. Thus, an estimate $\tilde{\mathbf{x}}$ of (3.3.4) can be found by applying (3.3.2). However, (3.3.2) does not consider any property of the adversarial error \mathbf{e} .

In general, some statistical properties of the adversarial error \mathbf{e} may be known, such as a [moment](#) or a [probability distribution](#) (family). Consequently, an estimate $\tilde{\mathbf{x}}$ of (3.3.4) can be obtained by considering the statistical properties of \mathbf{e} . For further discussion in this Section we will not consider a specific distribution of the adversarial error \mathbf{e} , but we simply assume an independent and identically distributed (i.i.d.) random variable $\mathbf{e}_{[n]}, n \in \{0, \dots, N-1\}$. However, a detailed discussion regarding the distribution of the adversarial error \mathbf{e} will follow in Section 3.6.

Assuming i.i.d. random variable $\mathbf{e}_{[n]}, n \in \{0, \dots, N-1\}$, we can use the ℓ_2 -norm of error \mathbf{e} , i.e. $\|\mathbf{e}\|_2^2$, to derive bounds on the estimation error. We know by (3.3.4) that for some $\rho > 0$ any \mathbf{x} lies in the space $\|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \leq \rho$ with some probability. So with use of (3.3.3) we can solve for solutions $\tilde{\mathbf{x}} \in \text{argmin}_{\mathbf{x}} \|\mathbf{x}\|_p$, s.t. $\|\mathbf{y} - \Phi \cdot \mathbf{x}\|_2^2 < \rho$, allowing for robust estimation. The term robustness is used to describe reconstruction schemes [FR13, Chapter 4.3], where the estimation error, e.g. $\|\mathbf{x} - \tilde{\mathbf{x}}\|_2$, can be controlled by $\rho > 0$.

To actually calculate the bounds on the estimation error, i.e. $\|\mathbf{e}\|_2^2$, we can exploit Kolmogorov's [strong law of large numbers \(LLN\)](#) [Kol33]. According to Kolmogorov's strong LLN, the scaled ℓ_2 -norm of the error vector $\mathbf{e} \in \mathbb{C}^N$ for N approaching infinity converges [almost surely \(a.s.\)](#) to the expected value of $\|\mathbf{e}\|_2^2$, i.e.

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \|\mathbf{e}\|_2^2 &= \mathbb{E} \left[\frac{1}{N} \|\mathbf{e}\|_2^2 \right] \forall n \text{ a.s.}, \\ \text{iff } \mathbb{E} \left[\frac{1}{N} \|\mathbf{e}\|_2^2 \right] &\text{ exists.} \end{aligned} \quad (3.3.5)$$

Furthermore, it follows from the Lindeberg–Lévy central limit theorem (CLT) [Lin22; Lévy26], that for finite variance and N approaching infinity the distribution of the properly translated and normalized ℓ_2 -norm [converges in distribution](#) to the standard normal distribution, i.e.

$$\lim_{N \rightarrow \infty} \frac{\|\mathbf{e}\|_2^2 - \mu}{\sigma^2} \sim f_{\mathcal{N}}(\cdot; 0, 1). \quad (3.3.6)$$

Note, that finite variance implies finite expected value. Since the [median](#) of the normal distribution is equal to its expectation, from (3.3.6) we can deduce

$$\lim_{N \rightarrow \infty} \text{median} \left(\frac{1}{N} \|\mathbf{e}\|_2^2 \right) = \mathbb{E} \left[\frac{1}{N} \|\mathbf{e}\|_2^2 \right]. \quad (3.3.7)$$

Consequently, given the noisy measurement (3.3.4) with i.i.d. random variables $\mathbf{e}_{[n]}$, we can

find \mathbf{x} ,

$$\text{s.t. } \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \leq \rho = \lambda \mathbb{E} \left[\|\mathbf{e}\|_2^2 \right] \approx \lambda \cdot \text{median} \left(\|\mathbf{e}\|_2^2 \right). \quad (3.3.8)$$

The constraint in (3.3.8) can be regarded as a relaxation of the constraint in (3.3.3) for noisy measurements. If a feasible solution exists, any $\tilde{\mathbf{x}}$ has to be in the solution space limited by ρ . Thus, the estimation error $\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2$ is bounded by $\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2 \leq \lambda \mathbb{E} \left[\|\mathbf{e}\|_2^2 \right]$. The tuning parameter $\lambda \in \mathbb{R}_{\geq 0}$ is applied to control the maximal permitted estimation error. For $\lambda = 1$ the bound can be approximated by $\text{median} \left(\|\mathbf{e}\|_2^2 \right)$, so that the solution \mathbf{x} lies within ρ with probability of 50%.

The ℓ_2 -error can be further simplified by exploiting the [fact](#), that the expectation of a sum of random variables is equal to the sum of expectations

of random variables. In addition, if \mathbf{e} consists of i.i.d. random variables, we have

$$\begin{aligned}\rho &= \lambda \mathbb{E}[\|\mathbf{e}\|_2^2] = \lambda \mathbb{E}\left[\sum_{n=0}^{N-1} |\mathbf{e}_{[n]}|^2\right] = \lambda \sum_{n=0}^{N-1} \mathbb{E}[|\mathbf{e}_{[n]}|^2] \\ &= \lambda N \cdot \mathbb{E}[|\mathbf{e}_{[n]}|^2].\end{aligned}\tag{3.3.9}$$

Since each $\mathbf{e}_{[n]}$ is obtained from i.i.d. random variables, we can refer to the expectation of any element of \mathbf{e} , e.g. the first element $\mathbf{e}_{[0]}$. Hence, the ℓ_2 -error of (3.3.8) is bounded by the expected value of $|\mathbf{e}_{[0]}|^2$ and some tuning parameter. As stated above, given $\lambda = 1$ we have $\rho \approx \text{median}\left(\|\mathbf{e}\|_2^2\right)$, so that the exact solution \mathbf{x} is found within ρ with probability of 50%. However, for $\lambda \neq 1$ the probability cannot be simply stated, since mere tuning of the expectation disregards the second central moment of $|\mathbf{e}_{[0]}|^2$. Therefore, including the variance into the calculation of the tuning parameter will reflect the distribution of the adversarial error. Since the error \mathbf{e} has property (3.3.6), a properly adjusted tuning parameter can be obtained by applying the desired probability on the inverse cumulative distribution function (also called the [quantile function](#)) of the normal distribution, see (A.1.8) in Appendix A.1.

3.4 Recovery Criteria

According to the [Nyquist-Shannon sampling theorem](#) [Lan67b; Sha48], bandlimited signals are firstly sampled at Nyquist rate and afterwards processed to extract the information within the signal. For the general class of bandlimited signals without extra information on the signal itself, sampling at Nyquist rate is the optimal choice to sample a signal, so that the information within the signal is fully preserved. However, signals containing further structure may have a more efficient way of sampling and processing. Sparse signals are one of such classes, which occur in diverse applications. Therefore, even before the term compressed sensing has been used, applications already exploited sparse properties of signals. Some well-known applications are the multimedia codecs MP3 and JPEG. The joint work of Candès, Romberg and Tao [CRT06] as well as Donoho [Don06] introduced ℓ_1 -norm minimization and showed that it works optimally for random matrices. For this they established the term [compressed sensing \(CS\)](#). Especially the breakthrough in [CRT06] regarding the required number of measurements with random Fourier samples for exact reconstruction of the signal

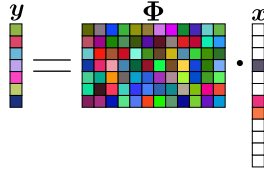


Figure 3.4: Visualization of the compressed sensing model.

with high probability drew a lot of attention, leading to further work on that subject. In the following we will give some basic statements for the field of CS. For more comprehensive reading we refer to the books of Foucart and Rauhut [FR13] as well as Eldar and Kutyniok [EK12].

In general, CS revolves around two aspects of general linear systems, a signal with sparsity property in a certain basis and a measurement matrix with random features. The CS model is formulated as a finite linear system (3.3.4) as discussed in the previous Section 3.3

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}. \quad (3.4.1)$$

In extension of (3.3.4), the vector $\mathbf{x} \in \mathbb{C}^M$ is sparse and $\Phi \in \mathbb{C}^{N \times M}$ is a (structured) random measurement matrix¹. A graphical visualization of CS is exemplary shown in Figure 3.4. Note, that in certain cases the measurement matrix is split into two matrices with $\Phi = \mathbf{A}\Phi' \in \mathbb{C}^{M \times M}$, such that we have

$$\mathbf{y} = \mathbf{A}\Phi' \mathbf{x} + \mathbf{e}. \quad (3.4.2)$$

There, \mathbf{A} and Φ' represent a (structured) random matrix and a fixed orthogonal matrix, respectively. This is normally the case, if the sparse property is not present in the canonical basis, but in some other basis. One example is image reconstruction, where the pixel visualization is usually not sparse, unlike its wavelet transform. While the model (3.4.1) will be of importance in Chapter 4, the model (3.4.2) is more or less present in Chapter 5. In the following, we will refer to (3.4.1), while (3.4.2) will be discussed in more detail in Chapter 5.

As also previously discussed in Section 3.3, \mathbf{x} shall be reconstructed from measurements \mathbf{y} . However here, we want to exploit the additional sparsity information for reconstruction. Thus, problem (3.3.8) can be rewritten

$$\min \|\mathbf{x}\|_0 \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \leq \rho. \quad (3.4.3)$$

¹ Φ can simply be an i.i.d. normal or Bernoulli random matrix, but also a structured random matrix like the partial random Fourier matrix, which contains a random set of rows of the DFT.

Summarizing (3.4.3), we search for the sparsest \mathbf{x} such that the inequality holds. If the solution is outside the bound, we state the problem to be infeasible. Unfortunately, problem (3.4.3) is **NP-hard** [Nat95], such that solving it is computationally infeasible. Moreover, changing the ℓ_0 -“norm” to an ℓ_p -quasinorm with $p < 1$, we still encounter an NP-hard problem [GJY11]. To overcome the computational infeasibility, CS uses the convex (and therefore tractable) ℓ_1 -norm as a proxy of the ℓ_0 -“norm”, which is a valid relaxation under conditions discussed later in this Section. Thus, we obtain the ℓ_1 -norm minimization problem

$$\min \|\mathbf{x}\|_1 \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \leq \rho, \quad (3.4.4)$$

which will also be addressed in Section 3.5.2. Regarding the ℓ_1 -norm relaxation, especially three notions are of importance to analyse reconstruction properties: the null space property, the mutual coherence and the restricted isometry property (RIP), which will be subsequently introduced.

3.4.1 Null Space Property

Starting with the noiseless case, we have (3.4.1) with $\mathbf{e} = \mathbf{0}$. Thus, we observe (3.3.3) with $p = 1$, which represents the basis pursuit approach. To introduce the **null space property (NSP)**, we firstly *define the right null space of a matrix $\Phi \in \mathbb{C}^{N \times M}$ by*

$$\text{NS}(\Phi) := \{\mathbf{a} \in \mathbb{C}^M : \Phi \mathbf{a} = \mathbf{0}\}. \quad (3.4.5)$$

In other words, the null space is the set of all vectors \mathbf{a} , which fulfills $\Phi \mathbf{a} = \mathbf{0}$. Consequently, we have the vector $\mathbf{a} = \mathbf{0}$ as a trivial element of the null space.

*Given a set $\mathbb{A} \subset \mathbb{B} := \{0, \dots, M-1\}$ and a complement set $\mathbb{A}^c = \mathbb{B} \setminus \mathbb{A}$, then a matrix $\Phi \in \mathbb{C}^{N \times M}$ fulfills the **null space property (NSP)** of order s , if [FR13, Definition 4.1]*

$$\|\mathbf{b}_{[\mathbb{A}]}\|_1 < \|\mathbf{b}_{[\mathbb{A}^c]}\|_1, \quad \forall \mathbf{b} \in \text{NS}(\Phi) \setminus \mathbf{0}, \quad (3.4.6)$$

for all $\mathbb{A} \subset \mathbb{B}$ with $|\mathbb{A}| \leq s$. Therefore, given a pair of distinct s -sparse vectors $\mathbf{x}, \mathbf{x}' \in \Sigma_s$, the difference between these vectors will result in a $2s$ -sparse vector ($\mathbf{b} = \mathbf{x} - \mathbf{x}'$). In the extreme case, \mathbf{b} will have exactly $2s$ non-zero entries, if all indices of the non-zero entries in \mathbf{x} and \mathbf{x}' differ, i.e. $\text{supp}(\mathbf{x}) \cap \text{supp}(\mathbf{x}') = \emptyset^2$. Informally speaking, the observation of \mathbf{x} and \mathbf{x}' via a measurement matrix Φ has to be different for all possible pairs \mathbf{x}, \mathbf{x}' . Concluding, if and only if a measurement matrix Φ fulfills the

² $\text{supp}(\mathbf{x}) = \{n : \mathbf{x}_{[n]} \neq 0\}$

NSP of order $2s$, then every s -sparse \mathbf{x} can be uniquely recovered from measurement \mathbf{y} via ℓ_1 -norm minimization (basis pursuit).

The NSP can also be extended to a stable form, such that the NSP is a sufficient and necessary condition for the reconstruction of compressible vectors, see [CDD08] and [FR13, Definition 4.11]. In general, stable recovery refers to the property, that small errors on the measurement will only induce small deviations to the solution. Here, stability refers to the recovery of compressible vectors while the recovery error is controlled by the distance to s -sparse vectors, e.g. the ℓ_2 -error of best s -term approximation.

The term null space property has been introduced by Cohen et al. [CDD08], even though it has been already implicitly used in several works before. The NSP is also described in more detail in [FR13, Chapter 4].

Switching to the case with noisy measurements, we observe the ℓ_1 -norm minimization problem in (3.4.4). There, the NSP as defined in (3.4.6) is too weak to guarantee robust recovery, since it does not account for measurement errors. To obtain reconstruction guarantees, the null space property can be strengthened to be robust in the presence of noise, then known as robust NSP. *Given a set $\mathbb{A} \subset \mathbb{B} := \{0, \dots, M-1\}$ and a complement set $\mathbb{A}^c = \mathbb{B} \setminus \mathbb{A}$, a matrix $\Phi \in \mathbb{C}^{N \times M}$ fulfills the **robust NSP** of order s , if [FR13, Definition 4.17]*

$$\|\mathbf{b}_{[\mathbb{A}]}\|_1 < \alpha \|\mathbf{b}_{[\mathbb{A}^c]}\|_1 + \beta \|\Phi \mathbf{b}\|_2, \quad \forall \mathbf{b} \in \mathbb{C}^M, \quad (3.4.7)$$

with constant parameter $0 < \alpha < 1$ as well as $\beta > 0$ for all $\mathbb{A} \subset \mathbb{B}$ with $|\mathbb{A}| \leq s$. Obviously, in contrast to the NSP (3.4.6), (3.4.7) permits vectors $(\mathbf{x} - \mathbf{x}') = \mathbf{b} \in \mathbb{C}^M \supset \text{NS}(\Phi)$ and includes the distance to $\text{NS}(\Phi)$ as a penalty term $\|\Phi \mathbf{b}\|_2$. The parameter β controls the influence of the penalty term, and therefore, controls the permitted measurement error on \mathbf{y} . Thus, the robust NSP (3.4.7) permits robust reconstruction, such that estimation error $\|\mathbf{x} - \tilde{\mathbf{x}}\|_2$ is controlled by the measurement error $\|\mathbf{y} - \Phi \mathbf{x}\|_2$. Furthermore, the robust NSP is necessary and sufficient condition for recovery of sparse signals with measurement errors (3.4.4).

However, in contrast to the subsequently introduced mutual coherence one drawback of the NSP is that it cannot be easily verified. In particular, Tillmann and Pfetsch [TP14] proved NP-hardness for the NSP. Note that, while in noiseless settings exact reconstruction under certain conditions is desired, stable and robust recovery is pursued in the noisy case, such that the estimated signal is close to the original one. In general, ℓ_1 -norm minimization (3.4.4) can be applied to just recover the exact support of an s -sparse vector, since for known $\text{supp}(\mathbf{x})$ standard LS methods can be used to stably and robustly recover \mathbf{x} .

3.4.2 Mutual Coherence

Mutual coherence is a measure of similarity between a pair of columns of a matrix. Therefore, *given a matrix $\Phi \in \mathbb{C}^{N \times M}$ with normalized columns, i.e. $\|\Phi_{[:,n]}\|_2 = 1 \ \forall n$, the **mutual coherence** introduced by Donoho and Elad [DE03] is defined by*

$$\text{MC}(\Phi) := \max_{1 \leq n < m \leq M} \left| (\Phi_{[:,n]})^H \cdot \Phi_{[:,m]} \right|. \quad (3.4.8)$$

The mutual coherence is related to the Welch bound [Wel74], such that the range is restricted by the Welch bound³ with $\text{MC}(\Phi) \in \left[\sqrt{\frac{M-N}{N(M-1)}}, 1 \right]$. In general it can be stated, that the smaller the mutual coherence the better the recovery performance via ℓ_1 -norm minimization (3.4.4). Therefore based on the mutual coherence concept, matrices achieving the Welch bound are most promising candidates for good reconstruction. For example, Welch bound achieving matrices can be generated by constructing DFT submatrices using difference sets [XZG05; XZG06]. However, even for wide matrices with more columns than rows $N \ll M$ close to the Welch bound the number of measurements required to solve problem (3.4.4) for all s -sparse signals scales quadratically in sparsity, see [FR13, Chapter 5]. Consequently, deterministically generated matrices based on low mutual coherence will just perform quadratic in scaling. Moreover, this result is valid for all currently known deterministically generated matrices.

3.4.3 Restricted Isometry Property

The notion of the **restricted isometry property (RIP)** introduced by Candès and Tao in 2005 [CT05] in conjunction with matrix random theory provides a tool to surpass the quadratic scaling of mutual coherence. Informally, the RIP can be regarded as a generalization of mutual coherence. Instead of just considering the similarity between all pairs of column vectors of a matrix, the RIP examines the joint similarity between all possible combinations of s columns of a matrix. Even though the RIP is generally used in conjunction with random matrices, the property can be applied on deterministic matrices as well.

*A matrix Φ has **restricted isometry property (RIP)** of order s , if there exists a δ_s with $0 \leq \delta_s < 1$ and*

$$(1 - \delta_s) \|\mathbf{b}\|_2^2 \leq \|\Phi \mathbf{b}\|_2^2 \leq (1 + \delta_s) \|\mathbf{b}\|_2^2, \quad \forall \mathbf{b} \in \Sigma_s. \quad (3.4.9)$$

³Note that the original bound described by Welch is actually more general.

Essentially, the RIP forces every subset of columns of Φ to be approximately **orthonormal**. The δ_s parameter controls how strict the orthonormality is enforced. Thus, the corner case of $\delta_s = 0$ just permits orthonormal column subsets of matrix Φ .

The RIP is connected to the previously discussed NSP. In particular, if a matrix fulfills RIP, then under certain conditions [EK12, Theorem 1.5] the matrix also fulfills NSP. Consequently, RIP represents a strictly stronger condition than NSP. Since NSP is a sufficient and necessary condition for successful recovery, RIP is a sufficient, but not a necessary condition.

In [EK12, Theorem 1.4] and [Dav10, Theorem 3.5] it has been shown, that if a matrix $\Phi \in \mathbb{C}^{N \times M}$ satisfies the **restricted isometry property (RIP)** of order $2s$ with **restricted isometry constant** $0 < \delta_{2s} \leq 1/2$, then

$$|\Lambda| \geq c \cdot s \cdot \ln(M/s), \text{ with } c \approx 0.28. \quad (3.4.10)$$

If Φ fulfills RIP with certain restricted isometry constant, then all submatrices $\Phi_{[\Lambda, \cdot]}$ will satisfy RIP with the corresponding restricted isometry constant as well. In particular, (3.4.10) gives a lower limit on the number of measurements, which are required to satisfy the RIP. For any number of measurements below that requirement a successful recovery cannot be guaranteed. Successful recovery refers to the exact solution for noiseless measurements, i.e. $\mathbf{e} = \mathbf{0}$ in (3.4.1). In the noisy setting, successful recovery refers to stable and robust reconstruction. Note, that the actual required restricted isometry constant depends on the estimation algorithms and has to be chosen appropriately, see Section 3.5. However, a value of $0 < \delta_{2s} \leq 1/2$ is a valid choice for many approaches.

In comparison to mutual coherence, proving the RIP of a matrix is NP-hard [TP14]. Therefore, deterministically generating a matrix with RIP is computationally intractable. However, [CT06] and subsequent work introduced probabilistic construction of matrices satisfying the RIP with high probability. In particular, the close entanglement of the RIP and the Johnson-Lindenstrauss lemma shown by Baraniuk et al. [BDD+08; KW11] enabled a better access to a probabilistic construction of RIP matrices. Some examples of how to construct matrices based on the Johnson-Lindenstrauss lemma can be found in [JW15]. Besides normal and Bernoulli random matrices, also the DFT submatrices obtained from random DFT rows are common candidates for probabilistic RIP matrices. A classic result [CT06; BDD+08] states that *for i.i.d. normal and Bernoulli random matrices $\Phi \in \mathbb{C}^{N \times M}$ the RIP is satisfied with high probability for*

$$N \geq \mathcal{O}(s \cdot \ln(M/s)). \quad (3.4.11)$$

Thus, (3.4.11) implies linear scaling of the number of measurements N and

sparsity s , but the size of the solution space M scales only logarithmically in N . Therefore, if $s \ll M$, we can choose a measurement matrix $\Phi \in \mathbb{C}^{N \times M}$ with $N < M$, so that reconstruction is still possible for an underdetermined linear system. In summary we can say that the linear scaling of probabilistic RIP matrices strongly outperforms deterministically generated matrices via mutual coherence with quadratic scaling.

3.5 Estimation Algorithms

In this Section, we will introduce the classical LS estimators, least squares (LS) and reduced-rank least squares (RRLS), as well as the common compressed sensing approaches, basis pursuit denoising (BPDN) and orthogonal matching pursuit (OMP). All discussed approaches estimate an unknown $\tilde{\mathbf{x}}$ from a linear system $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$ (cf. (3.3.4) and (3.4.1)) with *non-adaptive* measurements. Non-adaptivity refers to an approach, where the rows of the measurement matrix Φ used for reconstruction are chosen non-adaptively and independently of \mathbf{x} . Commonly, the rows of the measurement matrix Φ are constructed on deterministic basis or at random.

The initially considered LS estimators in Section 3.5.1 do not exploit sparsity. In contrast to that, the subsequently discussed CS methods, BPDN in Section 3.5.2 and OMP in Section 3.5.3, will consider sparsity of the unknown \mathbf{x} . Furthermore, for the CS methods we provide bounds for robust and stable reconstruction with respect to the previously established restricted isometry property (RIP).

3.5.1 Reduced-Rank Least Squares

The classical least squares (LS) estimator finds solutions of (3.3.8) by minimizing the ℓ_2 -norm of the adversarial error, which is frequently used in channel estimation [BES+95]. Thus, *given the linear system of (3.3.4) with full rank measurement matrix $\Phi \in \mathbb{C}^{N \times M}$ and $N \geq M$, a unique estimate of \mathbf{x} is found by the ℓ_2 -norm version of (3.3.2) with*

$$\begin{aligned} \tilde{\mathbf{x}} &= \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \Phi \cdot \mathbf{x}\|_2^2 \\ &= \left(\Phi^H \cdot \Phi \right)^{-1} \cdot \Phi^H \cdot \mathbf{y} \\ &= \Phi^+ \cdot \mathbf{y} = \mathbf{E}_{\text{LS}} \cdot \mathbf{y}, \end{aligned} \tag{3.5.1}$$

*which is referred to as the **LS approach**.* Hence, the LS estimator \mathbf{E}_{LS} is obtained by the pseudoinverse $\Phi^+ \in \mathbb{C}^{M \times N}$ of the measurement ma-

trix Φ . A full rank matrix is present almost surely for random matrices as used in (3.4.1). Note that the LS approach is also applicable for underdetermined systems ($N < M$) as well as rank deficient measurement matrices Φ . There, in contrast to (3.5.1), the solution will not be unique, i.e. $\tilde{\mathbf{x}} \in \arg\min_{\mathbf{x}} \|\mathbf{y} - \Phi \cdot \mathbf{x}\|_2^2$.

In noisy environments, the LS estimator may have poor estimation performance, since small eigenvalues of Φ can lead to strong noise amplification. One option to get a least squares estimator robust against noise is to consider an estimation matrix with reduced rank, which is obtained from the truncated singular value decomposition. Applying the singular value decomposition on the measurement matrix we obtain $\Phi = \mathbf{U}\Sigma\mathbf{V}^H$. The truncated version Σ_L is obtained by setting all non-dominant singular values of Σ to zero, i.e. $(\Sigma_L)_{[l,l]} = 0 \ \forall l \geq L$. Note that the singular values $\Sigma_{[l,l]}$ are of descending order, i.e. $\Sigma_{[l-1,l-1]} \geq \Sigma_{[l,l]} \ \forall l > 0$. Thus, the truncated version of LS estimator, namely the reduced-rank least squares (RRLS) estimator [BT90; EY36] of rank L , is defined as

$$\mathbf{E}_{\text{RRLS},L} = \mathbf{V} \cdot \Sigma_L^{-1} \cdot \mathbf{U}^H, \quad (3.5.2)$$

$$\text{where } (\Sigma_L^{-1})_{[m,n]} := \begin{cases} \frac{1}{\Sigma_{[n,m]}} & , \quad n = m, \ n < L \\ 0 & , \quad \text{else,} \end{cases}$$

where the number of dominant singular values of the RRLS estimator is $L = \text{rank}(\mathbf{E}_{\text{RRLS},L})$. The entries of the truncated diagonal $(M \times N)$ matrix Σ_L^{-1} are also given by $\text{diag}(\Sigma_L^{-1}) = \begin{bmatrix} \Sigma_{[0,0]}^{-1} & \dots & \Sigma_{[L-1,L-1]}^{-1} & 0 & \dots & 0 \end{bmatrix}^\top$. Note that since the singular values in Σ are of descending order, the non-zero diagonal elements of Σ_L^{-1} are of ascending order.

The error induced by truncation to L singular values can be calculated by the **Frobenius norm**. In particular, regarding the RRLS estimator $\mathbf{E}_{\text{RRLS},L}$, the induced estimation error by truncation depends on the Frobenius norm between Φ and its truncation $\Phi_L = \mathbf{U}\Sigma_L\mathbf{V}^H$, which is obtained by

$$\|\Phi - \Phi_L\|_F^2 = \text{Tr}(\Sigma^2 - \Sigma_L^2) = \sum_{l>L} \Sigma_{[l,l]}^2, \quad (3.5.3)$$

where $\text{Tr}(\Sigma^2)$ denotes the matrix trace with $\sum_{\forall l} \Sigma_{[l,l]}^2$. Therefore, the error described by the Frobenius norm is the sum of the non-dominant singular values of Φ . The Eckart–Young theorem [EY36] proved, that given an error described by the Frobenius norm $\|\cdot\|_F$, the best truncation of Φ , i.e. the best rank L approximation to Φ , is obtained by the truncated singular value decomposition $\mathbf{U}\Sigma_L\mathbf{V}^H$. Consequently, regarding the Frobenius norm, Φ_L is the best rank L approximation of Φ .

However, given the noise variance of the noisy measurement (3.3.4) we still do not know how to choose the rank L of the RRLS estimator, so that the truncation error in conjunction with the noise amplification is minimized. Therefore, we consider the ℓ_2 -error between the RRLS estimate $\tilde{\mathbf{x}}_L$ of rank L and \mathbf{x} with the use of (3.3.4) and (3.5.2)

$$\begin{aligned} \|\mathbf{x} - \tilde{\mathbf{x}}_L\|_2^2 &= \|\mathbf{x} - \mathbf{E}_{\text{RRLS},L} \cdot \mathbf{y}\|_2^2 = \|\mathbf{x} - \mathbf{E}_{\text{RRLS},L} \cdot \Phi \cdot \mathbf{x} - \mathbf{E}_{\text{RRLS},L} \cdot \mathbf{e}\|_2^2 \\ &= \|\mathbf{x} - \mathbf{V} \Sigma_L^{-1} \mathbf{U}^H \mathbf{U} \Sigma \mathbf{V}^H \mathbf{x} - \mathbf{V} \Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2 \\ &= \|\mathbf{x} - \mathbf{V} \mathbf{I}'_L \mathbf{V}^H \mathbf{x} - \mathbf{V} \Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2. \end{aligned} \quad (3.5.4)$$

where with slight abuse of notation, the diagonal $(M \times M)$ matrix \mathbf{I}'_L contains ones on the first $L \leq M$ diagonal entries and zeros on the residual entries. Thus, \mathbf{I}'_L represents a truncated identity matrix. Multiplying the unitary matrix \mathbf{V}^H within the ℓ_2 -norm of (3.5.4) does not change the ℓ_2 -norm. So we have

$$\begin{aligned} \|\mathbf{x} - \tilde{\mathbf{x}}_L\|_2^2 &= \|\mathbf{V}^H \mathbf{x} - \mathbf{I}'_L \mathbf{V}^H \mathbf{x} - \Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2 \\ &= \|(\mathbf{I}_M - \mathbf{I}'_L) \mathbf{V}^H \mathbf{x} - \Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2. \end{aligned} \quad (3.5.5)$$

Applying the [Minkowski inequality](#) to (3.5.5), we split the ℓ_2 -norm in two parts. However, since matrix $(\mathbf{I}_M - \mathbf{I}'_L)$ and matrix Σ_L^{-1} are non-zero on disjoint rows, we have equality $\|\mathbf{a}_{[\mathbb{A}]}\|_2^2 + \|-\mathbf{b}_{[\mathbb{B} \setminus \mathbb{A}]}\|_2^2 = \|\mathbf{a}_{[\mathbb{A}]} - \mathbf{b}_{[\mathbb{B} \setminus \mathbb{A}]}\|_2^2$. Thus, we have

$$\|\mathbf{x} - \tilde{\mathbf{x}}_L\|_2^2 = \|(\mathbf{I}_M - \mathbf{I}'_L) \mathbf{V}^H \mathbf{x}\|_2^2 + \|\Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2. \quad (3.5.6)$$

Given the unitary matrix \mathbf{U}^H and the i.i.d. random variable $\mathbf{e}_{[n]}$, $n \in \{0, \dots, N-1\}$, the expectation of the noise term is written

$$\begin{aligned} \mathbb{E}[\|\Sigma_L^{-1} \mathbf{U}^H \mathbf{e}\|_2^2] &= \mathbb{E}[\text{Tr}(\Sigma_L^{-1} \mathbf{U}^H \mathbf{e} \cdot \mathbf{e}^H \mathbf{U} (\Sigma_L^{-1})^H)] \\ &= \text{Tr}(\Sigma_L^{-1} \mathbf{U}^H \mathbb{E}[\mathbf{e} \cdot \mathbf{e}^H] \mathbf{U} (\Sigma_L^{-1})^H) \\ &= \text{Tr}(\Sigma_L^{-1} \mathbf{U}^H \mathbb{E}[|e_{[0]}|^2] \mathbf{U} (\Sigma_L^{-1})^H) \\ &= \mathbb{E}[|e_{[0]}|^2] \cdot \text{Tr}(\Sigma_L^{-2}). \end{aligned} \quad (3.5.7)$$

Applying a unitary matrix $\mathbf{U}^H \mathbf{e}$, the statistical properties of \mathbf{e} are preserved [TV05, Chapter A.1]. Similar to (3.5.7), given the unitary matrix \mathbf{V}^H

and the i.i.d. random variable $\mathbf{x}_{[n]}$, $n \in \{0, \dots, N-1\}$, we have

$$\begin{aligned}
 \mathbb{E} \left[\left\| (\mathbf{I}_M - \mathbf{I}'_L) \mathbf{V}^H \mathbf{x} \right\|_2^2 \right] &= \text{Tr} \left((\mathbf{I}_M - \mathbf{I}'_L) \mathbf{V}^H \frac{\mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right]}{M} \mathbf{V} (\mathbf{I}_M - \mathbf{I}'_L)^H \right) \\
 &= \text{Tr} (\mathbf{I}_M - \mathbf{I}'_L) \cdot \frac{\mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right]}{M} \\
 &= \frac{M-L}{M} \cdot \mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right]. \tag{3.5.8}
 \end{aligned}$$

Thus, the expectation of the ℓ_2 -error between the RRLS estimate $\tilde{\mathbf{x}}_L$ of rank L and \mathbf{x} is given by

$$\mathbb{E} \left[\left\| \mathbf{x} - \tilde{\mathbf{x}}_L \right\|_2^2 \right] = \frac{M-L}{M} \cdot \mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right] + \lambda \mathbb{E} \left[|e_{[0]}|^2 \right] \cdot \text{Tr} (\mathbf{\Sigma}_L^{-2}), \tag{3.5.9}$$

where the tuning parameter $\lambda \in \mathbb{R}_{\geq 0}$ controls the noise impact. Eventually, the rank L of the RRLS estimator \mathbf{E}_{RRLS} is obtained by rearranging and minimizing (3.5.9)

$$\begin{aligned}
 L &= \underset{L'}{\text{argmin}} \left(\mathbb{E} \left[\left\| \mathbf{x} - \tilde{\mathbf{x}}_{L'} \right\|_2^2 \right] \right) \\
 &= \underset{L'}{\text{argmin}} \left(\text{Tr} (\mathbf{\Sigma}_{L'}^{-2}) + \frac{M-L'}{\lambda M} \cdot \frac{\mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right]}{\mathbb{E} \left[|e_{[0]}|^2 \right]} \right). \tag{3.5.10}
 \end{aligned}$$

We can further simplify (3.5.10) by retrieving the error induced by rank reduction

$$\begin{aligned}
\mathbb{E}[\|\mathbf{x} - \tilde{\mathbf{x}}_{L-1}\|_2^2] &= \text{Tr}(\boldsymbol{\Sigma}_{L-1}^{-2}) + \frac{M-L+1}{\lambda M} \cdot \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\mathbb{E}[|\mathbf{e}_{[0]}|^2]} \\
&= \text{Tr}(\boldsymbol{\Sigma}_L^{-2}) + \frac{M-L}{\lambda M} \cdot \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\mathbb{E}[|\mathbf{e}_{[0]}|^2]} \\
&\quad - \boldsymbol{\Sigma}_{[L,L]}^{-2} + \frac{1}{\lambda M} \cdot \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\mathbb{E}[|\mathbf{e}_{[0]}|^2]} \\
&= \mathbb{E}[\|\mathbf{x} - \tilde{\mathbf{x}}_L\|_2^2] - \boldsymbol{\Sigma}_{[L,L]}^{-2} + \frac{1}{\lambda M} \cdot \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\mathbb{E}[|\mathbf{e}_{[0]}|^2]}. \quad (3.5.11)
\end{aligned}$$

Thus, the error induced by reducing one rank is

$$\mathbb{E}[\|\mathbf{x} - \tilde{\mathbf{x}}_{L-1}\|_2^2] - \mathbb{E}[\|\mathbf{x} - \tilde{\mathbf{x}}_L\|_2^2] = -\boldsymbol{\Sigma}_{[L,L]}^{-2} + \frac{1}{\lambda M} \cdot \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\mathbb{E}[|\mathbf{e}_{[0]}|^2]}. \quad (3.5.12)$$

Since $\boldsymbol{\Sigma}^{-2}$ is of ascending order and the second term is constant, the truncation error in conjunction with the noise error is minimized for L such that

$$\boldsymbol{\Sigma}_{[L,L]} = \sqrt{\lambda \mathbb{E}[|\mathbf{e}_{[0]}|^2]} \cdot \sqrt{\frac{M}{\mathbb{E}[\|\mathbf{x}\|_2^2]}}. \quad (3.5.13)$$

In other words, if rank reduction leads to a larger truncation error than the error induced by noise amplification, then the overall estimation error will not benefit from further truncation. In particular, the truncation error would more than diminish the gain of reduced noise amplification.

Since achieving precise equality (3.5.13) is impractical, the optimal rank L of the RRLS estimator $\mathbf{E}_{\text{RRLS},L}$ is obtained by minimizing the difference between the left side and the right side of (3.5.13). Note that the singular

values $\Sigma_{[l,l]}$ are of descending order. Therefore, we have

$$L = \underset{l}{\operatorname{argmin}} \left| \Sigma_{[l,l]} - \sqrt{\lambda \mathbb{E}[|e_{[0]}|^2]} \cdot \sqrt{\frac{M}{\mathbb{E}[\|\mathbf{x}\|_2^2]}} \right| \quad (3.5.14)$$

$$\geq \left| \left\{ l : \Sigma_{[l,l]} \geq \rho = \sqrt{\lambda \mathbb{E}[|e_{[0]}|^2]} \cdot \sqrt{\frac{M}{\mathbb{E}[\|\mathbf{x}\|_2^2]}} \right\} \right|. \quad (3.5.15)$$

The approximation in (3.5.15) may return rank $L - 1$, which is smaller than the optimal rank obtained by (3.5.14). However, the error induced by the approximation will be generally not severe. Furthermore, standard methods for the reduced rank pseudoinverse operate with threshold ρ given in the form of (3.5.15). Therefore, *given the linear system of (3.3.4) and the singular value decomposition $\Phi = \mathbf{U}\Sigma\mathbf{V}^H$ the RRLS estimator used throughout this thesis is defined by*

$$\mathbf{E}_{\text{RRLS}} = \mathbf{V} \cdot \Sigma_L^{-1} \cdot \mathbf{U}^H, \quad (3.5.16)$$

$$\text{with } L = \left| \left\{ l : \Sigma_{[l,l]} \geq \rho = \sqrt{\lambda \mathbb{E}[|e_{[0]}|^2]} \cdot \sqrt{\frac{M}{\mathbb{E}[\|\mathbf{x}\|_2^2]}} \right\} \right|.$$

Thus, *given the RRLS estimator $\mathbf{E}_{\text{RRLS}} \in \mathbb{C}^{M \times N}$ and ($N \geq M$), an **RRLS solution** is found by*

$$\tilde{\mathbf{x}} = \mathbf{E}_{\text{RRLS}} \cdot \mathbf{y}. \quad (3.5.17)$$

Usually, we will consider $\mathbf{x} \in \mathbb{C}^M$ with $\|\mathbf{x}\|_2^2 = M$ (or at least to have $\mathbb{E}[\|\mathbf{x}\|_2^2] = M$), so that the term $\sqrt{M/\mathbb{E}[\|\mathbf{x}\|_2^2]}$ vanishes. Note that for $\mathbb{E}[|e_{[0]}|^2] = 0$, the RRLS estimator \mathbf{E}_{RRLS} reduces to the LS estimator \mathbf{E}_{LS} .

3.5.2 Basis Pursuit Denoising

In contrast to the LS approaches like RRLS, compressed sensing methods exploit $\operatorname{supp}(\mathbf{x})$ (support of the unknown $\mathbf{x} \in \mathbb{C}^M$). As described in Section 2.3 and [Mol05], typical wireless channels are sparse. Therefore, for s -sparse vectors with $s = |\operatorname{supp}(\mathbf{x})| \ll M$ the application of compressed sensing is

desirable. Consequently, given the measurement matrix $\Phi \in \mathbb{C}^{N \times M}$ and the measurement $\mathbf{y} \in \mathbb{C}^N$, the general linear problem stated in (3.3.8) as well as the reformulated ℓ_1 -norm minimization problem in (3.4.4) may also be uniquely solvable in the underdetermined case ($N < M$) if \mathbf{x} is s -sparse, cf. (3.4.11). The ℓ_1 -norm minimization approach is also often referred to as least absolute shrinkage and selection operator (**LASSO**) [Tib96] or basis pursuit denoising (**BPDN**) [CDS98; CT05], and is given by

$$\begin{aligned} \tilde{\mathbf{x}} &:= \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \\ \text{s.t.} \quad &\|\Phi \mathbf{x} - \mathbf{y}\|_2^2 \leq \rho = \lambda N \cdot \mathbb{E} \left[|e_{[0]}|^2 \right]. \end{aligned} \quad (3.5.18)$$

Hereinafter we will refer to (3.5.18) as the basis pursuit denoising (BPDN) approach. The BPDN approach can also be written as an ℓ_1 -norm constrained least squares version

$$\begin{aligned} \tilde{\mathbf{x}} &:= \underset{\mathbf{x}}{\operatorname{argmin}} \|\Phi \mathbf{x} - \mathbf{y}\|_2^2, \\ \text{s.t.} \quad &\|\mathbf{x}\|_1 \leq \rho'. \end{aligned} \quad (3.5.19)$$

Moreover, the unconstrained Lagrangian equivalent of the constrained versions (3.5.18) and (3.5.19) is

$$\tilde{\mathbf{x}} := \underset{\mathbf{x}}{\operatorname{argmin}} \left(\|\Phi \mathbf{x} - \mathbf{y}\|_2^2 + \lambda' \|\mathbf{x}\|_1 \right), \quad (3.5.20)$$

where λ' is a penalty operator for the ℓ_1 -norm term. In summary, each of the three approaches (3.5.18) to (3.5.20) searches for a solution depending on a certain parameter. In particular, the three versions solve a similar problem, but differ in the choice of parameter. More precisely [FR13, Proposition 3.2] states, that for each of the three versions a parameter exists, such that the same $\tilde{\mathbf{x}}$ is found, which uniquely minimizes the corresponding problem. However, the precise conversion from one parameter to another is not trivial.

As discussed in Section 3.4, the restricted isometry property (RIP)⁴ is a sufficient condition to give performance guarantees for solving (3.5.18). Informally, reconstruction via the BPDN approach (3.5.18) is stable and robust, if $\Phi \in \mathbb{C}^{N \times M}$ acts almost isometrically on every possible s -sparse vector. Thus, for a matrix Φ satisfying RIP of order s with a restricted isometry constant below a certain value, stable and robust recovery of the s -sparse \mathbf{x} can be guaranteed for the linear system $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$.

⁴RIP is a strictly stronger condition than NSP, where NSP is a sufficient and necessary condition solving (3.5.18).

Since the RIP has been introduced in 2005 [CT05], there has been a tremendous research effort to increase the sufficient restricted isometry constants for stable and robust recovery via BPDN. Andersson and Strömberg [AS14] obtained a sufficient restricted isometry constant $\delta_{2s} < 4/\sqrt{41} \approx 0.6246$ for the noisy case, which can be found also in [FR13, Theorem 6.12]. In more detail the theorem states, that having a matrix Φ satisfying the RIP of order $2s$ with $\delta_{2s} < 4/\sqrt{41}$, then for any $\mathbf{x} \in \mathbb{C}^M$ and $\mathbf{y} \in \mathbb{C}^N$ fulfilling $\|\Phi\mathbf{x} - \mathbf{y}\|_2^2 \leq \rho$ the s -sparse vector $\tilde{\mathbf{x}}$ of (3.5.18) approximates \mathbf{x} with errors restricted by

$$\begin{aligned} \|\mathbf{x} - \tilde{\mathbf{x}}\|_1 &\leq \alpha \cdot \gamma_s(\mathbf{x})_1 + \beta \cdot \sqrt{s \cdot \rho}, \\ \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 &\leq \frac{\alpha}{\sqrt{s}} \cdot \gamma_s(\mathbf{x})_1 + \beta \cdot \sqrt{\rho}, \end{aligned} \quad (3.5.21)$$

where $\alpha, \beta > 0$ are constant parameters only depending on δ_{2s} . Consequently, the error limit scales with the ℓ_1 -error of best s -term approximation $\gamma_s(\mathbf{x})_1$ and the magnitude the measurement errors $\sqrt{\rho}$. Obviously, a matrix Φ satisfying the RIP of order $2s$ is of most interest here, since it guarantees stable and robust reconstruction of all s -sparse vectors $\tilde{\mathbf{x}}$ from $\mathbf{y} = \Phi\mathbf{x} + \mathbf{e}$. The work of Cai and Zhang [CZ14] extended the restricted isometry constant for RIP of order $2s$ to a sharp bound of $\delta_{2s} < 1/\sqrt{2} \approx 0.7071$. Even more general, they provided a RIP bound of order as with $\delta_{as} < \sqrt{(1-a)/a}$, which is sharp for $a \geq 4/3$ [CZ14].

In addition, if the RIP statements above are fulfilled, then the BPDN approach is instance optimal [CT05; CDD08]. We call a reconstruction approach instance optimal⁵, if its approximation error has the same scaling in s (up to a constant c) as the best s -term approximation [FR13, Definition 11.1]

$$\|\mathbf{x} - \mathbf{x}'_s\|_1 \leq c \cdot \gamma_s(\mathbf{x})_1, \quad (3.5.22)$$

where \mathbf{x}'_s denotes the best s -term approximation. Thus, under the above RIP bounds, BPDN achieves in scaling as good performance as the best s -term approximation.

The convex BPDN problem can be solved via quadratic programming, where for example the interior point or the gradient projection method can be applied. In particular, to solve the BPDN problem we use the SPGL1-toolbox of Berg and Friedlander in [BF15], which implements a gradient projection method in polynomial time. The description and theory behind the SPGL1-toolbox can be obtained from [BF08].

⁵More precisely, ℓ_1 -instance optimal.

3.5.3 Orthogonal Matching Pursuit

Instead of solving (3.5.18) via a quadratic program, (3.5.18) can also be solved by greedy algorithms. One well-known greedy compressed sensing method is the orthogonal matching pursuit (OMP) algorithm [TG07b; EK12], which has reduced complexity in comparison to the basis pursuit denoising method. Tropp et al. [TG07b] noted that OMP originated from stagewise regression in the 1950s, and has been introduced independently in [MZ93; PRR+93]. Besides OMP, other greedy methods solving (3.5.18) exist. OMP can be counted to the family of greedy pursuit algorithm, where different variations and extensions exist. Furthermore, alternate methods like the thresholding approaches exist, where compressive sampling matching pursuit (COSAMP) [NT10] is a common approach. A list of greedy methods can be found in [EK12, Chapter 8.2]. Nevertheless, here we will focus on the conventional OMP algorithm as published in [EK12, Algorithm 8.3].

The conventional OMP is an iterative algorithm (see Algorithm 3.1), which solves a least squares problem at each iteration and successively reconstructs the unknown vector \mathbf{x} , cf. line 5 in Algorithm 3.1. The pseudoinverse $(\Phi_{[:,\mathbb{B}]})^+$ is constructed from a submatrix of Φ , which contains the columns selected by the index set \mathbb{B} . Starting with an empty index set \mathbb{B} , \mathbb{B} is extended at each iteration by the index with the largest projection of

Algorithm 3.1 Orthogonal matching pursuit (OMP) [EK12]

Input: $\mathbf{y} \in \mathbb{C}^N$, $\Phi \in \mathbb{C}^{N \times M}$, λ , $\mathbb{E}[\|\tilde{\mathbf{e}}_{[0]}\|^2]$, s

Initialize: $\mathbb{B} = \emptyset$, $\tilde{\mathbf{e}} = \mathbf{y}$, $\tilde{\mathbf{x}} = \mathbf{0}$, $s = \min[M, s]$

```

1: while  $\left(\|\tilde{\mathbf{e}}\|_2^2 \geq \rho = \lambda N \cdot \mathbb{E}[\|\tilde{\mathbf{e}}_{[0]}\|^2]\right)$  and  $(|\text{supp}(\tilde{\mathbf{x}})| < s)$  do
2:    $\mathbf{a} = \left| \Phi^H \cdot \tilde{\mathbf{e}} \right|$ 
3:    $m = \underset{m}{\operatorname{argmax}} \frac{\mathbf{a}_{[m]}}{\|\Phi_{[:,m]}\|_2}$ 
4:    $\mathbb{B} = \mathbb{B} \cup \{m\}$ 
5:    $\tilde{\mathbf{x}}_{[\mathbb{B}]} = (\Phi_{[:,\mathbb{B}]})^+ \cdot \mathbf{y}$ 
6:    $\tilde{\mathbf{e}} = \mathbf{y} - \Phi \cdot \tilde{\mathbf{x}}$ 
7: end while
8: return  $\tilde{\mathbf{x}}, \tilde{\mathbf{e}}$ 
```

the adversarial error, cf. line 2-4 in Algorithm 3.1. At the end of each iteration, the adversarial error is calculated, cf. line 6 in Algorithm 3.1. Note that at line 3 in Algorithm 3.1, we do not have to check for $m \notin \mathbb{B}$, because $\mathbf{a}_{[m]} = 0 \ \forall m \in \mathbb{B}$.

The OMP algorithm terminates if a stopping criterion is met. However, defining appropriate stopping criteria highly depend on the investigated problem. Here, two stopping criteria are considered, cf. line 1 in Algorithm 3.1. The first criterion terminates the algorithm, if the ℓ_2 -norm of the adversarial error is below a certain threshold, i.e.

$$\|\tilde{\mathbf{e}}\|_2^2 \geq \lambda N \cdot \mathbb{E} \left[|\tilde{\mathbf{e}}_{[0]}|^2 \right]. \quad (3.5.23)$$

Equivalent to (3.3.8), the threshold bounds the error of the estimate to be within the expectation of the adversarial error element with some scaling λN . The second criterion terminates the algorithm if a certain sparsity s of the estimate $\tilde{\mathbf{x}}$ is achieved, i.e.

$$|\text{supp}(\tilde{\mathbf{x}})| < s. \quad (3.5.24)$$

In this way, an s -sparse estimated solution can be enforced. If the sparsity s is not provided, then the algorithm runs until the first stopping criterion is met or all entries of $\tilde{\mathbf{x}}$ are set and, thus $\tilde{\mathbf{x}} \in \mathbb{C}^M$ has full support $|\text{supp}(\tilde{\mathbf{x}})| = M$. Note that as discussed earlier, compressed sensing (CS) methods are principally required to recover the support of an s -sparse vector, since problem (3.4.4) with known support can be simply solved by standard LS approaches. Similar to that, OMP iteratively assembles $\text{supp}(\mathbf{x})$, where the estimate $\tilde{\mathbf{x}}$ in each iteration step is primarily used to correctly update the support in the next iteration step.

In the recent decade, OMP has been intensively studied to obtain conditions for exact recovery for the noiseless case as well as for stable and robust reconstruction under noisy regime. Similar to the statement above, it is sufficient to prove exact support recovery for the noisy case to prove stable and robust reconstruction, since knowing the support of an s -sparse vector reduces the problem to a least squares problem, see (3.3.2) with $p = 2$. In 2017 Wen et al. [WZW+17] published sharp conditions for exact support recovery for noisy measurements. In particular, two theorems on recovery conditions are presented for OMP, where we pick the one that bounds the noise on the ℓ_2 -ball, which is the most relevant for us. [WZW+17, Theorem 1] states for the linear system in (3.4.1), that if $\|\mathbf{e}\|_2^2 \leq \rho$ and

$$\min_{n \in \mathbb{B}} |\mathbf{x}_{[n]}| > \frac{2 \cdot \sqrt{\rho}}{1 - \delta_{s+1} \cdot \sqrt{s+1}} \quad (3.5.25)$$

as well as Φ satisfy the RIP (3.4.9) with

$$\delta_{s+1} < \frac{1}{\sqrt{s+1}}, \quad (3.5.26)$$

then the OMP with stopping rule $\|\tilde{\mathbf{e}}\|_2^2 \leq \rho$ exactly reconstructs the $\mathbb{B} = \text{supp}(\mathbf{x})$ of any s -sparse vector \mathbf{x} in $|\mathbb{B}|$ iterations. Condition (3.5.25) enforces a magnitude on the entries of the s -sparse vector, such that the magnitude is exceeding the accumulated noise magnitude by a certain factor. Decreasing the restricted isometry constant δ_{s+1} and therefore enforcing orthonormality of Φ will also permit smaller magnitudes on the entries of \mathbf{x} . In addition, Wen et al. [WZW+17] also provided a necessary condition for the magnitudes of \mathbf{x} . Note that Liu et al. [LFL17] recently provided a weaker magnitude condition (3.5.25) for [WZW+17, Theorem 1], such that the magnitudes of \mathbf{x} can be smaller.

3.6 Partial Fourier Measurements

In the recent sections, we discussed the linear system $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$ and the requirements on the linear system for successful recovery. In particular, we discussed the requirements on the measurement matrix Φ , where the actual realization of Φ has been left fairly vague as (structured) random matrix.

In this Section, we will analyze partial Fourier measurements, which are reflected by Φ being a partial (random) Fourier matrix. Furthermore, we will discuss two types of partial Fourier measurements: (i) measurements of the channel transfer function and (ii) measurements of the spectral channel power (SCP). Depending on the type we will consider the following measurement model: (i) channel transfer function measurements $\hat{\mathbf{w}} = \Phi \cdot \mathbf{h} + \hat{\mathbf{n}}$ and (ii) SCP measurements $\hat{\mathbf{z}} = \Phi \cdot \mathbf{v} + \hat{\mathbf{e}}$. These two measurement models are obtained from our transmission model $\hat{\mathbf{w}} = \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} + \hat{\mathbf{n}}$ in (2.4.8) of Section 2.4. By exploiting the property of partial (random) Fourier matrices and sparsity, we show that we can perform channel estimation. In particular, we show that we can obtain (i) the channel impulse response (CIR) \mathbf{h} from the channel transfer function measurement $\hat{\mathbf{w}}$ with noise $\hat{\mathbf{n}}$ and (ii) the circular autocorrelation of the CIR (CAC) \mathbf{v} from the SCP measurement $\hat{\mathbf{z}} = |\hat{\mathbf{w}}|^2$ with error $\hat{\mathbf{e}}$. Therefore, we obtain a threshold, where (i) the exact solution of the CIR is found with certain probability and (ii) the solution of the CAC is found based on the expectation of the measurement error. Furthermore, we discuss the solution space of the two measurement models.

At first we start by defining the partial (random) Fourier matrix. An N_{DFT} -point DFT matrix is defined by

$$\mathbf{F}_{[f,t]} := \frac{1}{\sqrt{N_{\text{DFT}}}} \cdot \exp\left(-i2\pi \frac{ft}{N_{\text{DFT}}}\right), \quad (3.6.1)$$

where $f, t \in \mathbb{F} := \{0, \dots, N_{\text{DFT}} - 1\}$ as also defined in Section 2.4. Thus, we can define the measurement matrix $\Phi \in \mathbb{C}^{|\mathbb{S}_1| \times |\mathbb{A}|}$ by a N_{DFT} -point DFT submatrix with

$$\Phi := \mathbf{F}_{[\mathbb{S}_1, \mathbb{A}]}, \quad (3.6.2)$$

where $\mathbb{S}_1, \mathbb{A} \subseteq \mathbb{F}$. Successful recovery of \mathbf{x} from the measurement $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$, cf. (3.4.1), depends particularly on the dimension and the indices of the sets \mathbb{S}_1, \mathbb{A} . There, $\mathbb{C}^{|\mathbb{A}|}$ defines the solution space of $\mathbf{x} \in \mathbb{C}^{|\mathbb{A}|}$ and $\mathbb{C}^{|\mathbb{S}_1|}$ represents the measurement space of $\mathbf{y} \in \mathbb{C}^{|\mathbb{S}_1|}$. As discussed in Sections 3.4 and 3.5, while RRLS (3.5.17) considers problems of dimension $|\mathbb{S}_1| \geq |\mathbb{A}|$, BPDN (3.5.18) and OMP (Algorithm 3.1) consider problems of dimension $|\mathbb{S}_1| \geq |\text{supp}(\mathbf{x})| \ll |\mathbb{A}|$, cf. (3.4.11).

More specifically, let us consider our transmission model $\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]}$ in (2.4.7) with the transmission symbol $\hat{\mathbf{u}}_{[f]}$ and noise symbol $\hat{\mathbf{n}}_{[f]}$. There, $f \in \mathbb{S}_1 \subseteq \mathbb{P} \subseteq \mathbb{F}$ is limited on the transmission band \mathbb{P} , where \mathbb{S}_1 represents the allocated frequency resources of a single transmitter (user). To acquire information about the wireless channel, a subset of the transmission resources are used as pilots. For example, given \mathbb{S}_1 being the set of pilots of a single transmitter, the channel transfer function measurement $\hat{\mathbf{h}}_{[f]}$ can be deduced from measurements $\hat{\mathbf{w}}_{[f]}, f \in \mathbb{S}_1$. Consequently, given appropriate dimension and distribution of \mathbb{S}_1 and \mathbb{A} , the channel impulse response $\mathbf{h} \in \mathbb{C}^{N_{\text{CP}}+1}$ can be recovered from $\hat{\mathbf{h}}_{[f]}$ with the use of $\Phi \in \mathbb{C}^{|\mathbb{S}_1| \times |\mathbb{A}|}$. While hereinafter we discuss the general conditions on partial (random) Fourier matrices Φ regarding the sets \mathbb{S}_1 and \mathbb{A} , pilot-based channel estimation to retrieve the channel impulse response as well as amplitude-based channel estimation to retrieve the circular autocorrelation of the CIR will be the topic of the subsequent sections, Section 3.6.1 and Section 3.6.2.

Due to its high relevance, the recovery performance of partial Fourier matrices is an intensively studied topic. In general, the subject of partial Fourier matrices can be divided into the field of deterministically (see Section 3.4.2) and partially random (see Section 3.4.3) generated matrices. While deterministic DFT submatrices are generally obtained by decreasing the mutual coherence, the RIP is used to generate partial random Fourier matrices, where the rows are randomly selected from a Fourier matrix. Note that there also exists work addressing the intermediate subject

of deterministic design with only some partial probabilistic features, see for example [JKM18].

Commonly, **difference sets** are used to generate deterministic DFT submatrix with Welch bound achieving mutual coherence [XZG05; XZG06; TH17]. Xia et al. [XZG05; XZG06] originally showed the application of difference sets in construction of Welch bound achieving codebooks. Nevertheless, the general concept of difference sets has been already introduced by [Bru55]. Besides the drawback, that difference sets are just available for certain sizes of DFT matrices, the approach has the usual disadvantage of deterministic matrices, such that the number of measurements scale quadratically in sparsity, see Section 3.4.2.

In contrast to deterministic DFT submatrix via e.g. difference sets, the scaling between number of measurements and sparsity for partial random Fourier matrices is significantly better. Rudelson and Vershynin published already in 2008 [RV08], that a stable and robust reconstruction for partial random Fourier matrices can be achieved with high probability, if the number of measurements scales by $|\mathbb{S}_1| \geq \mathcal{O}(s \cdot \ln(|\mathbb{A}|)^4)$. The most recent extension on this result has been done by Haviv and Regev in [HR17], achieving a slightly enhanced scaling of

$$|\mathbb{S}_1| \geq \mathcal{O}(s \cdot \ln(|\mathbb{A}|)^3). \quad (3.6.3)$$

In addition, recently Bandeiry et al. [BLM17, Theorem 16] obtained also a lower bound on the scaling. Thus, the number of necessary measurements scales by

$$|\mathbb{S}_1| \geq \mathcal{O}(s \cdot \ln(|\mathbb{A}|)). \quad (3.6.4)$$

Therefore, given the results in (3.6.3) and (3.6.4), the implications of (3.4.11) are also valid for partial random Fourier matrices. Hence, given a partial random Fourier matrix $\Phi \in \mathbb{C}^{|\mathbb{S}_1| \times |\mathbb{A}|}$ with $(|\mathbb{S}_1| < |\mathbb{A}|)$, then a solution of the underdetermined linear system can be found via compressed sensing methods, if $s \ll |\mathbb{A}|$. Within this thesis, the considered size of DFT-matrices does not exceed $2^{11} = 2048$. Therefore, in the region of interest the sufficient scaling (3.6.3) is impractical, since already for small s the number of required measurements quickly explodes. However, the lower bound (3.6.4) shows, that the necessary condition for reconstruction is still accessible for the given DFT-matrix size.

In the following two sections we will concentrate on our transmission model (2.4.7) from Section 2.4

$$\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]}, \quad f \in \mathbb{S}_1, \quad (3.6.5)$$

where we restrict the transmission vector \mathbf{u} to a certain transmission alphabet. While we use pilot symbols in Section 3.6.1, we relax \mathbf{u} to constant amplitude symbols in Section 3.6.2.

3.6.1 Pilot-Based Channel Estimation

Classic channel estimation [OA07; LTH+14; JSW09] is concerned with the recovery of the wireless channel from measurements at the receiver. Here, we consider pilot-based channel estimation to reconstruct the channel impulse response (CIR) from a known frequency resource set \mathbb{S}_1 of complex pilots $\hat{\mathbf{u}}_{[\mathbb{S}_1]}$. Therefore, for simplification of notation and without loss of generality, the complex pilots are set to one, i.e.

$$\hat{\mathbf{u}}_{[f]} = 1, \quad f \in \mathbb{S}_1. \quad (3.6.6)$$

Consequently, the transmission model (3.6.5) simplifies to

$$\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{h}}_{[f]} + \hat{\mathbf{n}}_{[f]}, \quad f \in \mathbb{S}_1. \quad (3.6.7)$$

Since in CP based OFDM systems the support of the CIR has to fulfill (2.4.6)⁶

$$\text{supp}(\mathbf{h}) \subseteq \mathbb{N}_{\text{CIR}} \quad (3.6.8)$$

$$\text{with } \mathbb{N}_{\text{CIR}} := \{0, \dots, N_{\text{CP}}\}, \quad (3.6.9)$$

we restrict the N_{DFT} -point DFT submatrix (3.6.2) on the CP set $\mathbb{A} = \mathbb{N}_{\text{CIR}}$ ⁷. Thus, by applying the inverse discrete Fourier transform on the channel transfer function $\hat{\mathbf{h}}$ we obtain

$$\begin{aligned} \hat{\mathbf{h}}_{[f]} &= \frac{1}{\sqrt{N_{\text{DFT}}}} \sum_{t \in \mathbb{N}_{\text{CIR}}} \exp\left(-i2\pi \frac{ft}{N_{\text{DFT}}}\right) \cdot \mathbf{h}_{[t]} \\ &= \sum_{t \in \mathbb{N}_{\text{CIR}}} (\mathbf{F}_{[t,f]} \mathbf{h}_{[t]}) \end{aligned} \quad (3.6.10)$$

Therefore, (3.6.7) can be written as linear equations similar to (3.3.4) with

$$\hat{\mathbf{w}}_{[f]} = \sum_{t \in \mathbb{N}_{\text{CIR}}} (\mathbf{F}_{[t,f]} \mathbf{h}_{[t]}) + \hat{\mathbf{n}}_{[f]}, \quad f \in \mathbb{S}_1, \quad (3.6.11)$$

⁶In particular, the cyclic-prefix of systems like CP based OFDM is designed to be at least of size of the CIR, where the CIR contains the significant multipath entries, cf. (2.3.2).

⁷Note that $|\mathbb{N}_{\text{CIR}}| = N_{\text{CP}} + 1$.

where the noise variable $\hat{\mathbf{n}}$ refines the generic adversarial error \mathbf{e} . Furthermore, considering $\mathbf{\Phi} = \mathbf{F}_{[\mathbb{S}_1, \mathbb{N}_{\text{CIR}}]}$ and $\mathbf{h} \in \mathbb{C}^{|\mathbb{N}_{\text{CIR}}|}$, (3.6.11) can be written using matrix notation

$$\hat{\mathbf{w}}_{[\mathbb{S}_1]} = \mathbf{\Phi} \cdot \mathbf{h} + \hat{\mathbf{n}}_{[\mathbb{S}_1]}. \quad (3.6.12)$$

Thus, the estimation algorithms presented in Section 3.5 can be directly applied on (3.6.12) to estimate the CIR \mathbf{h} . However, since the noise on the channel transfer function measurements is assumed to be independent and identically circularly-symmetric complex normally distributed $\hat{\mathbf{n}}_{[f]} \sim \text{f}_{\mathbb{CN}}(\cdot; 0, \sigma_{\hat{\mathbf{n}}}^2)$, the bounds on the ℓ_2 -error in (3.3.8) can be more precisely stated.

Therefore, in the following we will provide different bounds on the ℓ_2 -error in closed-form, so that a solution of (3.3.8) is found with certain probability. We begin by simply bounding the ℓ_2 -error on the noise expectation. However, there a precise bound achieving a certain probability can be obtained for $\lambda = 1$, only. Then, to obtain bounds achieving arbitrary probability, we retrieve bounds on approximated probability distributions of $\hat{\mathbf{n}}$. There, we provide two approximation, where the latter gives more precise results especially for small number of measurements $|\mathbb{S}_1|$.

The ℓ_2 -norm of $\hat{\mathbf{n}}$ (also holds for the more general normal distribution with zero mean) results in the Gamma distribution (Appendix A.7)

$$\|\hat{\mathbf{n}}_{[\mathbb{S}_1]}\|_2^2 \sim \text{f}_{\Gamma}(\cdot; \sigma_{\hat{\mathbf{n}}}^2, |\mathbb{S}_1|). \quad (3.6.13)$$

Thus, we obtain the expected value of $\|\hat{\mathbf{n}}\|_2^2$ via the first raw [moment](#) of the Gamma distribution (A.7.11)

$$\mathbb{E}[\|\hat{\mathbf{n}}_{[\mathbb{S}_1]}\|_2^2] = |\mathbb{S}_1| \cdot \sigma_{\hat{\mathbf{n}}}^2. \quad (3.6.14)$$

Hence, the problem (3.3.8) can be refined, such that we simply solve

$$\begin{aligned} &\text{find } \mathbf{h}, \\ &\text{s.t. } \|\hat{\mathbf{w}}_{[\mathbb{S}_1]} - \mathbf{\Phi} \cdot \mathbf{h}\|_2^2 \leq \rho = \lambda |\mathbb{S}_1| \cdot \sigma_{\hat{\mathbf{n}}}^2. \end{aligned} \quad (3.6.15)$$

where the ℓ_2 -norm term is constrained by the expectation of the Gamma distribution $|\mathbb{S}_1| \cdot \sigma_{\hat{\mathbf{n}}}^2$ with some scaling $\lambda \in \mathbb{R}_{\geq 0}$. A solution of problem (3.6.15) can be found by RRLS estimation $\tilde{\mathbf{h}} = \mathbf{E}_{\text{RRLS}} \cdot \hat{\mathbf{w}}_{[\mathbb{S}_1]}$, cf. (3.5.17), where the RRLS estimator \mathbf{E}_{RRLS} (3.5.16) is constructed using

$\rho = \sqrt{\lambda} \cdot \sigma_{\hat{\mathbf{n}}} \cdot \sqrt{|\mathbb{N}_{\text{CIR}}| / \mathbb{E}[\|\mathbf{h}\|_2^2]}$. The threshold ρ is obtained by replacing the expectation $\mathbb{E}[\|\mathbf{e}_{[0]}\|^2]$ of (3.5.16) by (3.6.14) with $|\mathbb{S}_1| = 1$.

For $|\mathbb{S}_1| = 1$ the ICDF of the Gamma distribution can be expressed in closed-form (A.7.21). Thus, for $\lambda = 1$ the solution \mathbf{h} will be found within ρ with probability of approximately 63.2%. Further example values for $|\mathbb{S}_1| = 1$ are given in Table A.4. In addition, as stated for (3.3.8), for $|\mathbb{S}_1| \rightarrow \infty$ the expectation tends to the median. Thus, given $\lambda = 1$ the true channel impulse response \mathbf{h} is found within the bound with probability of around 50%. Consequently, given $\lambda = 1$ the solution \mathbf{h} exists within the solution space with probability between 63.2% and 50%.

Furthermore, for large $|\mathbb{S}_1|$ the Gamma distribution can be approximated by the standard normal distribution (A.7.15), such that the bound in (3.6.15) on mere expectation is replaced by the inverse cumulative distribution function

$$\begin{aligned}\rho &= F_{\Gamma}^{-1}(P; \sigma_{\mathbf{n}}^2, |\mathbb{S}_1|) \\ &\approx F_{\mathcal{N}}^{-1}(P; |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2, |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^4),\end{aligned}\quad (3.6.16)$$

where the mean ($|\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2$) and the variance ($|\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^4$) of the Gamma distribution is given by (A.7.11) and (A.7.13), respectively. With the use of (A.1.8), the ICDF of the normal distribution of (3.6.16) is replaced by

$$\begin{aligned}\rho &= |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2 + |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^4 \cdot \overbrace{\sqrt{2} \cdot \text{erf}^{-1}(2P - 1)}^{\lambda} \\ &= \lambda |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^4 + |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2,\end{aligned}\quad (3.6.17)$$

In this case, the probability that the solution exists within the bound (3.6.17) can be obtained for given λ from the CDF of the standard normal distribution (A.1.2). Some example values for λ can be found in Table A.1, where e.g. $\lambda = 2$ leads to probability of around 97.72%.

However, using directly the normal distribution as approximation will lead to high deviations for small $|\mathbb{S}_1|$. Another option is to apply a modified variable on the normal distribution, which results in more accurate probabilities than (3.6.17). One approximation is given by (A.7.9), which sets the bound dependent on an approximated Gamma distribution

$$\rho = |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2 \left(\frac{F_{\mathcal{N}}^{-1}(P)}{\sqrt{9|\mathbb{S}_1|}} + 1 - \frac{1}{9|\mathbb{S}_1|} \right)^3 \approx F_{\Gamma}^{-1}(P; \sigma_{\mathbf{n}}^2, |\mathbb{S}_1|). \quad (3.6.18)$$

The approximation of the noise distribution in (3.6.16) and (3.6.18) enables more precise adjustment of the threshold ρ on the solution space $\mathbb{C}^{|\mathbb{N}_{\text{CIR}}|}$. In particular, we can chose the probability that the solution exists within

the threshold. While RRLS (3.5.16) just restricts the ℓ_2 -error of the solution, BPDN (3.5.18) and OMP (Algorithm 3.1) additionally pose a sparsity constrained on the solution.

In the literature, pilot-based channel estimation to obtain the CIR via Fourier measurements utilizing compressed sensing has been intensively studied, see overview publications in [BWH+10; BHS+10]. As implied by (3.4.11), (3.6.3) and (3.6.4), we require $|\text{supp}(\mathbf{h})| \ll N_{\text{CP}} + 1$ to achieve successful reconstruction from an underdetermined system of linear equations ($|\mathbb{S}_1| < N_{\text{CP}} + 1$). As described in Section 2.3 and in [Mol05], typical wireless channels \mathbf{h} are compressible, see compressibility definition in (3.2.5) of Section 3.2. Hence, we require a small number of non-zero taps only, i.e.

$$|\text{supp}(\mathbf{h})| \ll N_{\text{CP}} + 1, \quad (3.6.19)$$

to have a good representation of the CIR. In other words, given the CIR vector \mathbf{h} with sorted entries $|\mathbf{h}_{[t]}| \geq |\mathbf{h}_{[t+1]}|$, the entries $|\mathbf{h}_{[t]}|$ will decay quickly, i.e. exponential. Due to the compressibility property of \mathbf{h} , CS methods are highly applicable to solve (3.6.15). In addition, the performance of Algorithm 3.1 (OMP) can be increased, if a-priori knowledge about $|\text{supp}(\mathbf{h})|$ is available. There, the stopping criterion based on sparsity is used to restrict the solution on the significant multipath components, such that an s -sparse CIR is enforced.

3.6.2 Amplitude-Based Channel Estimation

In the previous section, we discussed pilot-based channel estimation via channel transfer function measurements $\hat{\mathbf{w}} = \mathbf{\Phi} \cdot \mathbf{h} + \hat{\mathbf{n}}$ with the partial Fourier measurements matrix $\mathbf{\Phi}$, the CIR \mathbf{h} and noise $\hat{\mathbf{n}}$. In this section, we will focus on amplitude-based channel estimation⁸ via spectral channel power (SCP) measurements $\hat{\mathbf{z}} = \mathbf{\Phi} \cdot \mathbf{v} + \hat{\mathbf{e}}$, where \mathbf{v} and $\hat{\mathbf{e}}$ denote the circular autocorrelation of the CIR (CAC)⁹ and the noise term, respectively. The SCP measurements will be derived from our transmission model $\hat{\mathbf{z}} = |\hat{\mathbf{w}}|^2 = \left| \text{diag}(\hat{\mathbf{h}}) \hat{\mathbf{u}} + \hat{\mathbf{n}} \right|^2$ in (2.4.8) of Section 2.4. In particular, we will show that amplitude-based channel estimation can retrieve the CAC, if the amplitudes $|\hat{\mathbf{u}}|$ are known a-priori. This differs from pilot-based channel estimation discussed in the previous section, where the complex-valued pilots $\hat{\mathbf{u}}$ have to be known a-priori. Thus, we are able to obtain the SCP, where we exploit the linear Fourier transformation between CAC and SCP.

⁸We will use here also the term channel estimation, even though we do not estimate the channel directly, but its autocorrelation.

⁹The circular feature of the autocorrelation arises from the circular property of the discrete Fourier transform (DFT).

Furthermore, we use the expectation to obtain a bound on the ℓ_2 -error to solve (3.3.8) and discuss the solution space.

The advantage of amplitude-based channel estimation over pilot-based channel estimation is that instead of pilots only the amplitude of symbols have to be known, yielding in a further degree of freedom at the transmitter. However, while the pilot-based channel estimation in Section 3.6.1 is able to reconstruct the complex-valued channel impulse response (CIR), the here considered amplitude-based channel estimation with SCP measurements just reconstructs the CAC. In general, retrieving the original complex-valued CIR from the CAC is not feasible. However, we can obtain the SCP from the CAC via the linear Fourier transformation, which reveals information on the course of the wireless channel in the frequency domain. In particular, the channel gain over frequency can be obtained, which is of interest for e.g. signal quality measurements between transmitter and receiver.

At first we show how amplitude-based channel estimation estimates the CAC from SCP measurements. We begin by considering the power spectrum observed at a receiver, which is given by our transmission model (3.6.5) with absolute squared frequency measurements

$$\hat{\mathbf{z}}_{[f]} = |\hat{\mathbf{w}}_{[f]}|^2 = \left| \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]} \right|^2. \quad (3.6.20)$$

Just like in (3.6.6), the frequency resources $f \in \mathbb{S}_1$ are known a-priori. However, instead of pilots as in (3.6.6), we use transmit symbols with known amplitudes $|\hat{\mathbf{u}}|$, where $\hat{\mathbf{u}}_{[f]}$ is selected from a known transmit alphabet. Phase-shift keying symbols is an obvious choice for a transmit alphabet, since the amplitude is constant for all symbols. Therefore, only a real-valued scalar has to be known. Furthermore, for simplification of notation and without loss of generality, we normalize the amplitude of all transmission symbols to

$$|\hat{\mathbf{u}}_{[f]}| = 1, \quad f \in \mathbb{S}_1. \quad (3.6.21)$$

Thus, given (3.6.21) we can reformulate the power spectrum of (3.6.20) to

$$\begin{aligned}
 \hat{z}_{[f]} &= |\hat{w}_{[f]}|^2 = \left| \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]} \right|^2 \\
 &= \left| \hat{\mathbf{h}}_{[f]} + \hat{\mathbf{u}}_{[f]}^* \cdot \hat{\mathbf{n}}_{[f]} \right|^2 \\
 &= \left| \hat{\mathbf{h}}_{[f]} \right|^2 + \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]} + \left| \hat{\mathbf{n}}_{[f]} \right|^2 \\
 &= \left| \hat{\mathbf{h}}_{[f]} \right|^2 + \hat{\mathbf{e}}_{[f]} \\
 &= \hat{\mathbf{v}}_{[f]} + \hat{\mathbf{e}}_{[f]}, \quad f \in \mathbb{S}_1.
 \end{aligned} \tag{3.6.22}$$

Thus, we separated the spectral channel power (SCP) $\hat{\mathbf{v}}_{[f]} = \left| \hat{\mathbf{h}}_{[f]} \right|^2 \in \mathbb{R}_{\geq 0}$ and the error $\hat{\mathbf{e}}_{[f]} \in \mathbb{R}$ given by

$$\hat{\mathbf{e}}_{[f]} = \left| \hat{\mathbf{n}}_{[f]} \right|^2 + \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]}.$$
 \tag{3.6.23}

In particular, the SCP is just the absolute squared of the channel transfer function $\hat{\mathbf{h}}$. Hence, in general the SCP is missing the phase information of the channel transfer function. As stated in the previous Section 3.6.1, the relation between channel transfer function and channel impulse response is given by the DFT matrix (3.6.10)

$$\hat{\mathbf{h}} = \mathbf{F} \mathbf{h}.$$
 \tag{3.6.24}

In addition to (3.6.24) a connection between the SCP and the circular autocorrelation of the CIR can be established via the discrete Fourier transform using the [Wiener-Khintchine theorem](#) [Wie30; Khi34]

$$\underbrace{\left| \hat{\mathbf{h}}_{[f]} \right|^2}_{\hat{\mathbf{v}}_{[f]}} = \sum_t \left(\mathbf{F}_{[t,f]} \cdot \underbrace{\frac{1}{N_{\text{DFT}}} \cdot \sum_{n=0}^{N_{\text{DFT}}-1} \left(\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}]}]} \right)}_{\mathbf{v}_{[t]}} \right). \tag{3.6.25}$$

More general, the Wiener-Khintchine theorem states that the power spectral density is the Fourier transform of the autocorrelation function. Here, in addition, the circular property of the DFT enforces also the circular property on the autocorrelation function. Thus, due to the Wiener-Khintchine theorem, we obtain the power spectral density by the discrete Fourier transform of the circular autocorrelation function and vice versa.

In the following, we will describe $\hat{\mathbf{v}}$ as the spectral channel power (SCP) and \mathbf{v} as the circular autocorrelation of the CIR (CAC). As discussed, the SCP is connected to the CAC via the DFT matrix

$$\hat{\mathbf{v}} = \mathbf{F}\mathbf{v}. \quad (3.6.26)$$

In addition, the circular autocorrelation of the CIR in (3.6.25) is explicitly given by

$$\mathbf{v}[t] = \frac{1}{N_{\text{DFT}}} \cdot \sum_{n=0}^{N_{\text{DFT}}-1} \left(\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} \right). \quad (3.6.27)$$

Each element of the circular autocorrelation $\mathbf{v}[t]$ observes the linear relation of \mathbf{h} to its circular shifted version. If $\mathbf{v}[t] = 0$ for any t , then the sum for the circular shift t has to be zero. This occurs in two cases:

- (i) Even though $\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} \neq 0, n \in \text{supp}(\mathbf{h})$ for a fixed t , the summands can cancel each other out. For example, this happens for orthogonal sequences, which are particularly designed in that way.
- (ii) If $\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} = 0, \forall n \in \text{supp}(\mathbf{h})$ for a fixed circular shift t , which frequently occurs for sparse signals.

Using (3.6.25) we can rewrite (3.6.22) as linear equations similar to (3.3.4) with

$$\hat{\mathbf{z}}_{[f]} = \sum_t \left(\mathbf{F}_{[t,f]} \cdot \mathbf{v}[t] \right) + \hat{\mathbf{e}}_{[f]}, \quad f \in \mathbb{S}_1. \quad (3.6.28)$$

A visualization of the relation between CIR, channel transfer function, SCP and CAC is given in Figure 3.5. Furthermore, Figure 3.5 depicts the amplitude of an exemplary CIR and its corresponding CAC.

Note that given a channel impulse response with certain structure, phase information of the channel transfer function may be still present in the SCP. In particular this is the case for sparse CIRs, where algorithms exist to partially recover the phase information, e.g. phase retrieval [WBJ15].

We now turn our attention to the solution space, which is captured by the support of the circular autocorrelation of the CIR. In particular, we examine two cases: (i) the general solution space given by set \mathbb{N}_{CAC} , where no sparsity is assumed on the signal, and (ii) the solution space given by the support $\text{supp}(\mathbf{v})$ of the CAC with \mathbf{h} being sparse/compressible. At first we will focus on the first case. The latter case will be discussed later in this Section, see e.g. (3.6.45), where sparsity of the CAC is considered.

The set \mathbb{N}_{CAC} can be derived from the set \mathbb{N}_{CIR} (3.6.8). Since the circular autocorrelation is obtained by comparison of the CIR to a shifted version

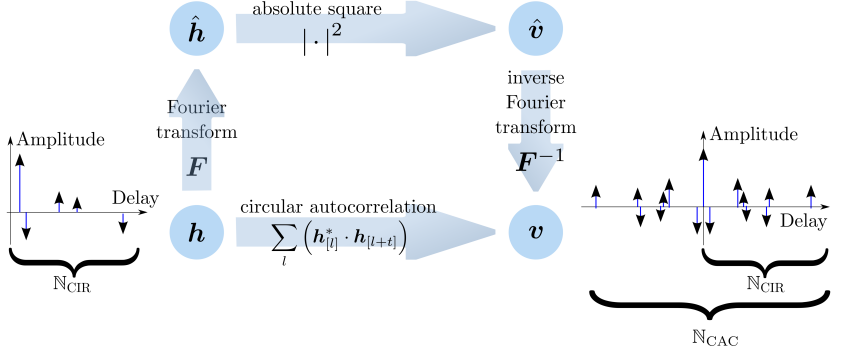


Figure 3.5: Relation between CIR, channel transfer function, SCP and CAC according to the Wiener-Khintchine theorem. The amplitude of an exemplary CIR and its corresponding CAC are visualized.

of itself, the region of the CAC is

$$\begin{aligned} \mathbb{N}_{\text{CAC}} &= \{0, \dots, N_{\text{CP}}\} \cup \{N_{\text{DFT}} - N_{\text{CP}}, \dots, N_{\text{DFT}} - 1\} \\ &= \mathbb{N}_{\text{CIR}} \cup \{N_{\text{DFT}} - N_{\text{CP}}, \dots, N_{\text{DFT}} - 1\}. \end{aligned} \quad (3.6.29)$$

Thus, if the support of the channel impulse response fulfills $\text{supp}(\mathbf{h}) \subseteq \mathbb{N}_{\text{CIR}}$, then $\text{supp}(\mathbf{v}) \subseteq \mathbb{N}_{\text{CAC}}$ has cardinality

$$|\mathbb{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1. \quad (3.6.30)$$

Hence, the general solution space captured by \mathbb{N}_{CIR} and \mathbb{N}_{CAC} is increased by approximately a factor of two. A rough example on the solution space and the actual support of \mathbf{h} and \mathbf{v} is depicted in Figure 3.5. Therefore, we can restrict the summation of (3.6.28) to the set, where the support of the CAC can occur

$$\hat{\mathbf{z}}_{[f]} = \sum_{t \in \mathbb{N}_{\text{CAC}}} (\mathbf{F}_{[t,f]} \cdot \mathbf{v}_{[t]}) + \hat{\mathbf{e}}_{[f]}, \quad f \in \mathbb{S}_1, \quad (3.6.31)$$

or with $\Phi = \mathbf{F}_{[\mathbb{S}_1, \mathbb{N}_{\text{CAC}}]}$ using matrix notation we write

$$\hat{\mathbf{z}}_{[\mathbb{S}_1]} = \Phi \cdot \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]} + \hat{\mathbf{e}}_{[\mathbb{S}_1]}. \quad (3.6.32)$$

Thus, the estimation algorithms presented in Section 3.5 can be directly applied on (3.6.32) to perform amplitude-based channel estimation. However, similar to Section 3.6.1 we want to have a more precise description of the

error $\hat{\mathbf{e}}_{[\mathbb{S}_1]}$, such that we can obtain a bound for the ℓ_2 -error in (3.3.8). In particular, to obtain a bound for (3.3.8)

$$\mathbb{E} \left[\|\hat{\mathbf{e}}_{[\mathbb{S}_1]}\|_2^2 \right] = \sum_{f \in \mathbb{S}_1} \mathbb{E} \left[|e_{[f]}|^2 \right] = |\mathbb{S}_1| \cdot \mathbb{E} \left[|\hat{\mathbf{e}}_{[0]}|^2 \right], \quad (3.6.33)$$

we have to find the second absolute raw [moment](#) of the error $\hat{\mathbf{e}}_{[f]}$ from (3.6.23), which is given by

$$|\hat{\mathbf{e}}_{[f]}| = \left| |\hat{\mathbf{n}}_{[f]}|^2 + \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]} \right|. \quad (3.6.34)$$

Therefore, we expand the absolute square of the error term (3.6.34) using the [triangle inequality](#) and the binomial formula

$$\begin{aligned} |\hat{\mathbf{e}}_{[f]}|^2 &= \left| |\hat{\mathbf{n}}_{[f]}|^2 + \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]} \right|^2 \\ &\leq |\hat{\mathbf{n}}_{[f]}|^4 + \left| \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]} \right|^2 \\ &= |\hat{\mathbf{n}}_{[f]}|^4 + 2 \left| \hat{\mathbf{h}}_{[f]} \hat{\mathbf{n}}_{[f]} \right|^2 + \left(\hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} \hat{\mathbf{n}}_{[f]}^* \right)^2 + \left(\hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{n}}_{[f]} \right)^2. \end{aligned} \quad (3.6.35)$$

As discussed in the previous Section 3.6.1, we have $\hat{\mathbf{n}}_{[f]} \sim \text{f}_{\mathbb{CN}}(\cdot; 0, \sigma_{\hat{\mathbf{n}}}^2)$. Thus, we exploit the first and the second raw moment of the Gamma distribution with 2 degrees of freedom (A.7.22) and obtain

$$\mathbb{E} \left[|\hat{\mathbf{n}}_{[f]}|^2 \right] = \sigma_{\hat{\mathbf{n}}}^2 \quad (3.6.36)$$

$$\mathbb{E} \left[|\hat{\mathbf{n}}_{[f]}|^4 \right] = 2\sigma_{\hat{\mathbf{n}}}^4. \quad (3.6.37)$$

Furthermore, we consider a propagation channel exhibiting [Rayleigh](#) or [Rice](#) fading, see Section 2.3. Thus, the random variable of $\left| \hat{\mathbf{h}}_{[f]} \right|^2$ follows the scaled non-central chi-squared distribution (see Appendix A.9). Therefore, to include the general case of Rice fading, we use the first raw moment of the scaled non-central chi-squared distribution (A.9.7)

$$\mathbb{E} \left[\left| \hat{\mathbf{h}}_{[f]} \right|^2 \right] = \sigma_{\hat{\mathbf{h}}}^2 + \left| \mathbb{E} \left[\hat{\mathbf{h}}_{[f]} \right] \right|^2. \quad (3.6.38)$$

Hence, given (3.6.36) to (3.6.38), the second absolute raw moment of the error is limited by

$$\mathbb{E}\left[|\hat{\mathbf{e}}_{[f]}|^2\right] \leq \mathbb{E}\left[|\hat{\mathbf{n}}_{[f]}|^4\right] + \mathbb{E}\left[2\left|\hat{\mathbf{h}}_{[f]}\hat{\mathbf{n}}_{[f]}\right|^2\right] \quad (3.6.39)$$

$$\begin{aligned} & + \underbrace{\mathbb{E}\left[\left(\hat{\mathbf{h}}_{[f]}\hat{\mathbf{u}}_{[f]}\hat{\mathbf{n}}_{[f]}^*\right)^2\right]}_{\mathbb{E}[\cdot]=0} + \underbrace{\mathbb{E}\left[\left(\hat{\mathbf{h}}_{[f]}^*\hat{\mathbf{u}}_{[f]}^*\hat{\mathbf{n}}_{[f]}\right)^2\right]}_{\mathbb{E}[\cdot]=0} \\ & = 2\sigma_{\mathbf{n}}^4 + 2\left(\sigma_{\mathbf{h}}^2 + \left|\mathbb{E}\left[\hat{\mathbf{h}}_{[f]}\right]\right|^2\right) \cdot \sigma_{\mathbf{n}}^2 \\ & = 2\sigma_{\mathbf{n}}^2 \left(\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2 + \left|\mathbb{E}\left[\hat{\mathbf{h}}_{[f]}\right]\right|^2\right). \end{aligned} \quad (3.6.40)$$

Now, we consider the case of a [Rayleigh](#) fading propagation channel only. Thus, the second absolute raw moment of $\hat{\mathbf{e}}$ are described by

$$\mathbb{E}\left[|\hat{\mathbf{e}}_{[f]}|^2\right] \leq 2\sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2). \quad (3.6.41)$$

Thus, (3.6.41) bounds the ℓ_2 -error of (3.3.8) with $2\sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2)$. Hence, the problem (3.3.8) based on linear equations (3.6.32) can be refined, such that we solve

$$\begin{aligned} & \text{find } \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}, \\ & \text{s.t. } \left\|\hat{\mathbf{z}}_{[\mathbb{S}_1]} - \Phi \cdot \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}\right\|_2^2 \leq 2\lambda |\mathbb{S}_1| \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2). \end{aligned} \quad (3.6.42)$$

where the second absolute raw moment of $\hat{\mathbf{e}}$ is applied with some scaling $\lambda \in \mathbb{R}_{\geq 0}$. The RRLS estimator \mathbf{E}_{RRLS} (3.5.16) can be constructed using

$$\rho = \sqrt{2\lambda \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2)} \cdot \sqrt{\frac{|\mathbb{N}_{\text{CAC}}|}{\mathbb{E}\left[\left\|\mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}\right\|_2^2\right]}}, \quad (3.6.43)$$

which is obtained by replacing the expectation $\mathbb{E}\left[|\mathbf{e}_{[0]}|^2\right]$ of (3.5.16) by (3.6.41). Consequently, we can estimate the CAC with RRLS (3.5.17).

Now, we return to the analysis of the solution space. As discussed in (3.6.29), if we just consider the dimension of the circular autocorrelation with $|\mathbb{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1$, then we have a scaling of almost a factor of two between the size of the CIR \mathbf{h} and the CAC \mathbf{v} . Now we turn our

attention to the dimension of the support $|\text{supp}(\mathbf{v})|$ of the circular autocorrelation of the CIR. For the channel impulse response we already justified $|\text{supp}(\mathbf{h})| \ll N_{\text{CP}} + 1$ in (3.6.19). Similar to (3.6.29) (and roughly visualized in Figure 3.5), the support (or number of relevant entries) of the CAC depends on the support of the channel impulse response (CIR) and the position of these entries within the CIR. Recalling the definition of the circular autocorrelation (3.6.27) calculating the sum just on the support of the CIR

$$\mathbf{v}_{[t]} = \frac{1}{N_{\text{DFT}}} \cdot \sum_{n \in \text{supp}(\mathbf{h})} \left(\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} \right), \quad (3.6.44)$$

we obtain by mere counting, that the cardinality of the support of the CAC is upper bounded by

$$|\text{supp}(\mathbf{v})| \leq |\text{supp}(\mathbf{h})|^2 - |\text{supp}(\mathbf{h})| + 1. \quad (3.6.45)$$

In more detail, given that the distribution of non-zero elements on the CIR follows the [power function](#)

$$\mathbf{h}_{[n]} = \begin{cases} a_n \neq 0, & n = 2^m - 1 \\ 0, & \text{else,} \end{cases} \quad (3.6.46)$$

where $m \in \{0, \dots, |\text{supp}(\mathbf{h})|\}$ with $|\text{supp}(\mathbf{h})| \leq \log_2(N_{\text{DFT}} + 1)$, then for each $n \neq m$ in (3.6.44) at most one overlap of two non-zero elements will occur. However, for $n = m$ all $|\text{supp}(\mathbf{h})|$ elements of \mathbf{h} will overlap, reducing the overall overlaps by $|\text{supp}(\mathbf{h})| - 1$. Thus, we obtain (3.6.45) as the upper bound.

The lower bound on $|\text{supp}(\mathbf{v})|$ is simply present, if \mathbf{h} has perfect circular autocorrelation properties, so that $\mathbf{v}_{[t]} = 0 \ \forall t \neq 0$. Thus, there we have the lower bound of $|\text{supp}(\mathbf{v})| = 1$ independent from $|\text{supp}(\mathbf{h})|$. However, if we are just interested in the number of overlaps of non-zero elements without considering the cases where the elements cancel each other out, then the lower bound will be higher. Therefore, modifying (3.6.44) to

$$\mathbf{v}'_{[t]} = \frac{1}{N_{\text{DFT}}} \cdot \sum_{n \in \text{supp}(\mathbf{h})} \left| \mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} \right| \quad (3.6.47)$$

we will have $\mathbf{v}'_{[t]} \neq 0$, if $\mathbf{h}_{[n]}^* \cdot \mathbf{h}_{[(n+t) \bmod N_{\text{DFT}}]} \neq 0$ for any n . There, the lower bound on the number of overlapping non-zero elements is achieved, if the non-zero elements of the CIR are equidistant. Then, the number of overlaps are equal to the number of non-zero elements of the CIR. Thus,

the lower bound on the number of overlapping non-zero elements is given by

$$|\text{supp}(\mathbf{v}')| = |\text{supp}(\mathbf{h})|. \quad (3.6.48)$$

Consequently, there are in general more non-zero entries within the CAC than the CIR (except for some designed cases like orthogonal sequences, where the summands of (3.6.44) cancel each other out). Nevertheless, the significant entries of \mathbf{v} still do not cover the complete set \mathbb{N}_{CAC} in many scenarios. Therefore, the sparse assumption on the CAC often holds

$$|\text{supp}(\mathbf{v})| \ll |\mathbb{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1. \quad (3.6.49)$$

Also note, that (3.6.45) is just an upper bound and the actual support of the circular autocorrelation can be much further below. Thus, having a sparse CAC, CS methods like BPDN (3.5.18) and OMP in Algorithm 3.1 can be applied. In addition similar to estimating the CIR, a-priori knowledge about $|\text{supp}(\mathbf{v})|$ can enforce an s -sparse CAC for OMP (Algorithm 3.1).

Note, that CS methods do not exploit the property of \mathbf{v} being a circular autocorrelation function. However, exploiting this fact is beyond the sparsity property and the compressed reception of bilinear combinations [JW15] has to be well understood. It can be seen, that the Fourier amplitude measurements and circular convolutions [WJP15] is a particular form of a phase retrieval problem, which can be approached by methods from low-rank matrix recovery.

In particular, we observe a simultaneous structured problem, where a circular autocorrelation function with sparse properties is present. Oymak et al. [OJF+15] and the refinement by Kliesch et al. [KSJ19] studied the problem of simultaneous structured problems and came to the conclusion, that there exists no convex algorithm jointly solving sparse and circular convolution problems with the current frameworks of CS and low-rank matrix recovery. In other words, convex algorithms can only solve a simultaneous structured problem for sparsity or circular autocorrelation separately.

3.7 Summary

Within this Chapter we introduced the basics of compressed sensing, building the foundation for the applications in the subsequent chapters. In particular, we provided recovery criteria for stable and robust reconstruction of sparse signals from (random) linear measurement with noise. Furthermore, we introduced the channel estimation algorithms: the classical reduced-rank

least squares method (3.5.17) as well as the compressed sensing based approaches of basis pursuit denoising (3.5.18) and orthogonal matching pursuit (Algorithm 3.1). Afterwards, we discussed (absolute squared) partial Fourier measurement and provided bounds from the literature, when successful recovery of sparse signals is possible. There, we considered two channel estimation scheme: pilot-based channel estimation (3.6.15) to retrieve the channel impulse response from channel transfer function measurements as well as amplitude-based channel estimation (3.6.42) to retrieve the circular autocorrelation of the channel impulse response from spectral channel power measurements. For these two schemes, we derived bounds on the ℓ_2 -error for noisy measurements. In particular, we stated the probability distribution of the measurement error (3.6.18), such that the channel impulse response can be estimated with a certain probability. Furthermore, we obtained an estimate for the expectation of the measurement error (3.6.42) to reconstruct the circular autocorrelation of the channel impulse response.

Chapter 4

Gray Space Detection

4.1 Overview

This Chapter introduces the *gray space detection* scheme, which identifies unused resources within active primary user spectrum. To enable detection of just single or few resources, composite binary hypothesis testing is combined with amplitude-based channel estimation based on compressed sensing. In particular, given measurements of the power spectrum, the spectral channel power from the primary user is reconstructed via amplitude-based channel estimation. Further on, each resource is classified as used or idle via the gray space test, which represents a composite log-likelihood ratio test utilizing empirical cumulative distribution function to achieve a desired false alarm rate. Finally, representative practical channel models from a measurement campaign are used to evaluate the gray space detection approach.

4.1.1 Main Contributions

The main contributions of this Chapter comprise the gray space detection scheme and the measurement campaign to retrieve representative practical channel models.

In more detail, we propose a gray space detection scheme [WPJ13; WJW+16b], which enables spectrum sensing in an already occupied, but partly exploited primary user spectrum band for interweave cognitive radio system [GJM+09]. In contrast to classical white space spectrum sensing for cognitive radio [Hay05], where generally completely unoccupied spectrum bands are exploited, gray space detection can exploit gray spaces, repre-

senting temporary small fractions of unused resources within an already occupied primary user spectrum band. While classical schemes like energy detection [Urk67; ATJ14] generally rely on the central limit theorem as well as the Wilks theorem [Wil38] to perform a simple hypothesis tests, spectrum sensing via gray space detection can also operate on composite hypotheses and does not require the central limit theorem and the Wilks theorem. Consequently, gray space detection can be beneficially applied on problems with random variables of low degree of freedom, like the here presented complex random variable with two degrees of freedom in form of the spectral channel power, and identify primary user activity on single resource basis. Since gray space detection exploits more realistic channel statistics via composite hypotheses, gray space detection achieves superior performance towards classical schemes like energy detection.

Furthermore, since gray space detection takes the varying channel into account, our proposed approach can control the false alarm probability, which represents the probability that a primary user occupied sample is falsely identified as a gray space. Therefore, the protection of the primary user system from interference of the cognitive radio can be controlled.

Our gray space detection approach consists of three steps: energy detection (step 1), amplitude-based channel estimation (step 2), and log-likelihood ratio test (step 3). Step 1 estimates a preliminary set of primary user resources, so that step 2 can estimate the spectral channel power on it. Eventually, step 3 estimates the final set of primary user resources, where the complement set represents the identified gray spaces.

In this Chapter we develop our gray space detection scheme and compare it with the conventional energy detector. The evaluation consists majorly of three parts: (i) comparison of amplitude-based channel estimation methods, (ii) comparison of our proposed thresholds for the log-likelihood ratio test, and (iii) comparison of different channel models including our measurement campaign.

First [WPJ13; WJW+16b], amplitude-based channel estimation in step 2 is performed via the different estimation methods introduced in Section 3.5. There, the primary focus lies on compressed sensing estimators like greedy orthogonal matching pursuit (OMP) [TG07b] and basis pursuit denoising (BPDN) [CT05], which are compared with the traditional reduced-rank least squares (RRLS) estimator. Simulation results regarding several aspects are provided: SNR, primary user system load, and allocated physical resource blocksize.

Second, since the uniformly most powerful hypothesis test is only applicable with full statistical knowledge of the channel, we derive a gray space test based on approximated composite threshold for the log-likelihood ra-

tio test in step 3. Furthermore, we use individual energy detection with an individual threshold for comparison. While the empirical cumulative distribution function is exploited to calculate the approximated composite threshold from partial local channel statistics, the individual threshold is estimated on individual sample basis and thus, applied for each decision independently. The gray space test is compared to individual energy detection and the UMP test. This second part is not yet covered by our official publications.

Third, a measurement campaign was conducted to obtain channel statistics for representative manufacturing process in factory automation [HWW+16; HWR+16; WHW16; DHC+19]. These short-range factory automation channel models are considered for our gray space detection scheme and are used for evaluation in conjunction with the ITU models.

Summary of own contributions within this Chapter:

[WPJ13; WJW+16b; HWW+16; HWR+16; WHW16; DHC+19]

4.1.2 Structure

The outline of this Chapter is as follows: The subsequent Section 4.2 introduces the system model of gray space detection and the problem description. It is followed by Section 4.3 describing the complete gray space detection algorithm, consisting of three steps: step 1 applies energy detection without requiring knowledge about the spectral channel power to estimate a preliminary base station resource allocation \tilde{S}_E (Section 4.3.1), step 2 exploits \tilde{S}_E to estimate the spectral channel power $\hat{\mathbf{v}}$ from the base station via amplitude-based channel estimation (Section 4.3.2), and step 3 applies the log-likelihood ratio test based on $\hat{\mathbf{v}}$ to estimate the final base station resource allocation \tilde{S}_G (Section 4.3.3). The link level simulation setup is described in Section 4.4. At the subsequent Section 4.5 we present our simulation results, including the receiver operating characteristic at different SNRs (Sections 4.5.1 and 4.5.2), the primary user system load (Section 4.5.3), multiple physical resource blocksizes (Section 4.5.4) and different thresholds for the hypothesis test (Section 4.5.5). The following Section 4.6 contains our results on different channel models including practical channels (Section 4.6.3). The representative practical channel models are obtained from our measurement campaign (Section 4.6.1) and described in Section 4.6.2. The chapter concludes with the summary on the gray space detection scheme in Section 4.7.

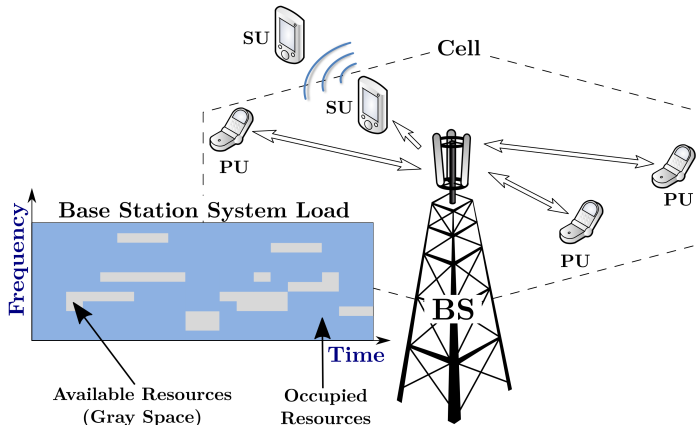


Figure 4.1: System setup with one base station (BS), several primary users (PUs) and secondary users (SUs) [WJW+16b].

4.2 System Model

In this Section we introduce our system model based on the transmission model in Section 2.4 for the application of gray space detection. Therefore, we will describe the scenario of an interweave cognitive radio system, where we apply the result of (3.6.22) on our transmission model (2.4.8) and formulate the binary hypothesis testing problem similarly to (2.6.6). Given the binary hypothesis testing problem at the end of this Section, we will elaborate the statistical hypothesis testing, cf. Sections 2.5 and 2.6, as well as amplitude-based channel estimation, cf. Section 3.6.2 in the subsequent Section 4.3.

We consider a scenario as depicted in Figure 4.1, where multiple primary users (PUs) are served in a single cell by a single-antenna downlink signal from a base station (BS). The BS communicates only in the transmission band $\mathbb{P} \subseteq \mathbb{F}$. A secondary user is located within the cell of the BS and wants to communicate with another secondary user (SU) within the band \mathbb{P} occupied by the BS for downlink transmission. Thus, we have an interweave cognitive radio system, where an SU intends to find unused time-frequency resources, so that the SU itself can exploit these unused resources for its own communication. Therefore, the SU observes the signal emitted by the BS and searches for gray spaces, where the term gray space refers to a temporary small fraction of unused resources within an already occupied spectrum band.

If the PU system operates at full system load, then the available transmis-

sion resources in the downlink are completely utilized by the BS. However, in practice, there are several reasons that a PU system operates at lower system load. Among others, poor channel conditions or only few active PUs in the cell can lead to low system load. Thus, only a subset of the transmission band $\mathbb{S}_1 \subseteq \mathbb{P}$ is used for downlink transmission, where \mathbb{S}_1 denotes the resource allocation of the BS. Therefore, we have two events

$$\begin{aligned} \hat{\mathbf{u}}_{[f]} &= 0, & f \in \mathbb{P} \setminus \mathbb{S}_1 & : \mathcal{H}_0 \\ \hat{\mathbf{u}}_{[f]} &\neq 0, & f \in \mathbb{S}_1 & : \mathcal{H}_1. \end{aligned} \quad (4.2.1)$$

Consequently, the BS operates with a system load below 100%, leaving resources at subcarrier $f \in \mathbb{P} \setminus \mathbb{S}_1$ vacant.

Furthermore, resources are scheduled on the basis of scheduling units, where a scheduling unit generally refers to a block of consecutive time-frequency resources. Thus, a PU either utilizes a complete scheduling unit or leaves it completely unexploited. Consequently, unallocated scheduling unit will leave distributed, but consecutive spectrum parts of the transmission band \mathbb{P} unoccupied, as depicted in the time-frequency diagram of Figure 4.1. In 3GPP LTE a scheduling unit is referred to as a transmission time intervals (TTIs), where one TTI consists of up to two PRBs [3GP20b, Chapter 5.2.2, 5.3.2], i.e. a duration of up to 1 ms. One PRB is regarded as the smallest scheduling unit, generally consisting of 7 consecutive OFDM symbols and $N_B = 12$ consecutive subcarriers [3GP20c], cf. Section 2.4. There, we introduce blocksize N_B to describe the number of adjacent subcarriers scheduled by the BS as a single scheduling unit for downlink transmission.

As illustrated in Figure 4.1, an SU in form of a cognitive radio is located in the PU cell. All devices, the SU and the PUs, receive the same downlink signal from the BS. Thus, the SU observes the BS signal through a sparse block-fading channel as described in (2.4.6). Since the resources at the BS are allocated according to (4.2.1), the received samples of the null hypothesis \mathcal{H}_0 (idle resources) contain only noise, i.e., $\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{n}}_{[f]}, f \in \mathbb{P} \setminus \mathbb{S}_1$. Considering our transmission model from (2.4.7), we observe two events similar to (2.6.3) with

$$\hat{\mathbf{w}}_{[f]} = \begin{cases} \hat{\mathbf{n}}_{[f]}, & f \in \mathbb{P} \setminus \mathbb{S}_1 : \mathcal{H}_0 \\ \hat{\mathbf{h}}_{[f]} \hat{\mathbf{u}}_{[f]} + \hat{\mathbf{n}}_{[f]}, & f \in \mathbb{S}_1 : \mathcal{H}_1. \end{cases} \quad (4.2.2)$$

Obviously, measurements of the channel transfer function $\hat{\mathbf{h}}_{[f]}$ from the BS are available in the case of the alternative hypothesis \mathcal{H}_1 in (4.2.2) only. Since the channel \mathbf{h} of (4.2.2) can also incorporate an unknown time offset τ_{offset} , cf. (2.4.3), \mathcal{H}_1 is also valid if the SU receives the BS signal with

certain inaccuracy in time. Note that the BS resource allocation \mathbb{S}_1 for downlink transmission is unknown to the SU. Therefore, the SU aims to estimate \mathbb{S}_1 .

Signal detection, cf. Sections 2.5 and 2.6, via statistical hypothesis tests can be applied to detect signals emitted from the BS, i.e. distinguish between the two events of (4.2.2). Each event represents a state of a resource, \mathcal{H}_0 being a gray space, i.e., $f \in \mathbb{P} \setminus \mathbb{S}_1$, and \mathcal{H}_1 being a resource occupied by the BS altered through a channel, i.e., $f \in \mathbb{S}_1$. To measure the performance of signal detection schemes, detection probability $P_d := P(\text{accept } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is true})$ defined in (2.5.7) or probability of false alarm $P_{fa} := P(\text{accept } \mathcal{H}_0 \mid \mathcal{H}_1 \text{ is true})$ defined in (2.5.8) are used. There, detection rate refers to the probability that a gray space is correctly identified and false alarm rate refers to the probability that a signal from the BS is not detected. While higher detection rate increases the number of gray spaces for SU exploitation, lower false alarm rate reduces the interference on the PU system. Therefore, low false alarm rate will protect the occupied BS resources from SU exploitation. In general, the operation of cognitive radio system in PU bands are more likely to be permitted, if certain protection of the PU can be guaranteed. For example, the operation point of 3GPP LTE is at 10% packet error rate [3GP20d], leaving the practical false alarm rate significantly below 10%. Hence, the practical false alarm rate should be significantly below 10%, so that a PU system is only marginally affected by an SU. Therefore, a hypothesis test controlling the false alarm rate is highly desirable.

Note that the final aim of an SU is the exploitation of the gray spaces. Thus, an SU has to identify and utilize the gray space within a single scheduling unit. One scheduling unit in e.g. 3GPP LTE [3GP20a] consist of up to two PRBs, i.e. 14 OFDM symbols. Consequently, identifying a gray space in the first OFDM symbols leaves 13 symbols for communication. However, the amount of available resources depends on how fast the decision process is executed.

Now, let us exploit the results obtained in Section 3.6.2. To enable amplitude-based channel estimation at an SU, the SU has to know the amplitudes of the PU symbols $|\hat{\mathbf{u}}_{[f]}|$, which are transmitted from the BS. One option to know the amplitudes at an SU is to restrict the transmission alphabet of the BS to complex symbols with constant amplitude. There, the SU knows the amplitude of the PU symbols up to a global amplitude. As discussed in Section 2.4, power allocation in the downlink is rather simple [3GP20d, Chapter 5.2], i.e. transmit power is constant for a complete set of resources. Consequently, the imprinted power structure of the restricted transmit alphabet is still present after transmit power allocation.

Thus, the gray space detection scheme is particularly easy to realize if we assume a transmit alphabet with complex symbols of constant amplitude, i.e. phase-shift keying with amplitude one. Hence, we have two events

$$|\hat{\mathbf{u}}_{[f]}| = \begin{cases} 0, & f \in \mathbb{P} \setminus \mathbb{S}_1 : \mathcal{H}_0 \\ 1, & f \in \mathbb{S}_1 : \mathcal{H}_1. \end{cases} \quad (4.2.3)$$

With increase of complexity, gray space detection may also be applicable for higher-order modulation scheme, i.e. quadrature amplitude modulation (QAM) with multiple distinct amplitudes.

Applying the transmit alphabet with constant amplitude symbols of (4.2.3) on the power spectrum of (4.2.2) and follow the derivation of (3.6.22), we observe two events of the power spectrum similar to (2.6.6). Thus, *we end up with a **binary hypothesis testing problem** between two power spectrum observations*

$$\hat{\mathbf{z}}_{[f]} = |\hat{\mathbf{w}}_{[f]}|^2 = \begin{cases} |\hat{\mathbf{n}}_{[f]}|^2, & f \in \mathbb{P} \setminus \mathbb{S}_1 : \mathcal{H}_0 \\ \hat{\mathbf{v}}_{[f]} + \hat{\mathbf{e}}_{[f]}, & f \in \mathbb{S}_1 : \mathcal{H}_1, \end{cases} \quad (4.2.4)$$

where the spectral channel power (SCP) is described by

$$\hat{\mathbf{v}}_{[f]} = |\hat{\mathbf{h}}_{[f]}|^2$$

and the error term is given by

$$\hat{\mathbf{e}}_{[f]} = \hat{\mathbf{u}}_{[f]} \hat{\mathbf{h}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{n}}_{[f]} + |\hat{\mathbf{n}}_{[f]}|^2.$$

Similar to (4.2.2), only in the case of the alternative hypothesis \mathcal{H}_1 in (4.2.4) observes the SCP of the transmission channel from the BS to an SU.

If the resource allocation \mathbb{S}_1 of the BS would be known to an SU, the power spectrum from the BS could be directly measured. Consequently, we could estimate the SCP from the BS via amplitude-based channel estimation as discussed in Section 3.6.2. However, here the SU is ignorant of the BS resource allocation \mathbb{S}_1 . Therefore, we have to estimate \mathbb{S}_1 from the observation at the SU. In the subsequent section we will recap energy detection as well as introduce our gray space detection approach. In contrast to energy detection, gray space detection operates on composite hypotheses and, thus, is applicable under frequency-varying wireless channels, where single spectrum samples are considered for signal detection.

4.3 Detection Procedure

In this Section, we will introduce our *gray space detection* approach. Given the binary hypothesis testing problem in (4.2.4) of the previous Section, we will exploit the main results on statistical hypothesis testing, cf. Sections 2.5 and 2.6, as well as on amplitude-based channel estimation, cf. Section 3.6.2, to derive the gray space detection procedure. Therefore, we will also recap the classical scheme of energy detection.

The final aim of an SU is to identify the BS resource allocation \mathbb{S}_1 , so that unused resources $f \in \mathbb{P} \setminus \mathbb{S}_1$ can be exploited for secondary user (SU) communication. Our proposed method of gray space detection shall identify available transmission resources within an active and used transmission band for the interweave cognitive radio system by estimating the utilized frequencies. Furthermore, besides the capability of distinguishing gray spaces from channel fades, our gray space detection approach shall also permit tuning of the primary user (PU) protection level, so that the interference of an SU on the PU system can be controlled.

Given the binary hypothesis testing problem in (4.2.4), an SU observes two cases of resource allocation embedded in the power spectrum. Hence, we can apply statistical hypothesis testing with the log-likelihood ratio test as given in (2.6.15) by

$$\Lambda(\hat{\mathbf{z}}_{[f]}, \hat{\mathbf{v}}_{[f]}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \rho. \quad (4.3.1)$$

There, the log-likelihood ratio Λ is written as a function of the observed power spectrum $\hat{\mathbf{z}}$ and the spectral channel power (SCP) $\hat{\mathbf{v}}$ of the alternative hypothesis \mathcal{H}_1 . The nuisance parameter $\hat{\mathbf{n}}$ as well as $\hat{\mathbf{v}} = 0$ of the null hypothesis \mathcal{H}_0 are omitted, cf. definition of LLR function in (2.6.11). The null hypothesis \mathcal{H}_0 is accepted, i.e. the spectrum sample $\hat{\mathbf{z}}_{[f]}$ is regarded as a gray space, if the LLR lies above a certain threshold ρ . Thus, for each spectrum sample we perform a test, whether we have $\mathcal{H}_1 : \hat{\mathbf{z}}_{[f]}, f \in \mathbb{S}_1$, i.e. $\hat{\mathbf{z}}_{[f]}$ is part of the BS resource allocation \mathbb{S}_1 , or $\mathcal{H}_0 : \hat{\mathbf{z}}_{[f]}, f \notin \mathbb{S}_1$. Hence, we eventually estimate the BS resource allocation \mathbb{S}_1 .

The test in (4.3.1) describes a more general log-likelihood ratio test, neither specifying the type of the alternative hypothesis \mathcal{H}_1 nor the type of the threshold ρ . Thus, given our system model in Section 4.2, we (i) have to determine whether the test involves simple hypotheses or composite hypotheses, and (ii) have to decide how to choose the threshold ρ . Both questions are related to the discussion in the previous Section about controlling the SU interference on the PU system, and therefore, guaranteeing a PU protection level.

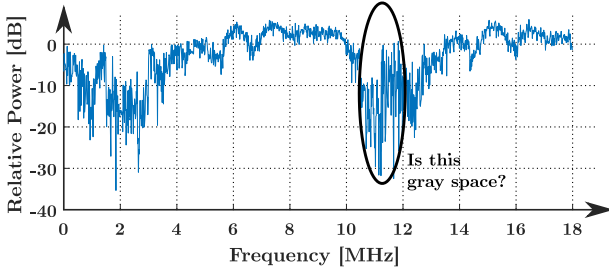


Figure 4.2: Observed power spectrum at a secondary user with primary user activity [WJW+16b]. Can the power spectrum tell us something about unused spectrum regions?

While simple hypotheses are used by classical schemes like energy detection and resemble constant SCP, i.e. $\hat{\mathbf{v}}_{[f]} = c$, a composite hypothesis describes varying SCP with $\hat{\mathbf{v}}_{[f]} > 0$, which reflects the actual channel characteristic captured by $\hat{\mathbf{v}}$ in our system model, cf. Section 2.3. In addition, the threshold can be chosen to achieve a desired performance criterion. In our case, to guarantee certain PU protection, the false alarm rate P_{fa} is a suitable performance metric. In the following, given our system model, we will discuss in more detail how we eventually obtain a log-likelihood ratio test for one-sided composite binary hypotheses achieving a desired false alarm rate P_{fa} .

An exemplary power spectrum is depicted in Figure 4.2, where we can already observe that the SCP is not a constant function. There, we ask if we can identify the gray spaces by power spectrum observation, even though the SCP is not a constant function. Each spectrum sample $\hat{z}_{[f]}$ with low power may indicate an unoccupied resource, i.e. gray space. However, channel fading has a quite similar appearance, so that samples with low power may also indicate a channel fade induced by the SCP.

In contrast to gray spaces, classical white spaces are present in the form of complete spectrum bands, where multiple independent observations of the spectrum band can be taken. Thus, if we consider the summation of multiple independent observations, then the summation can be modeled by a random variable with large degree of freedom. There, given the CLT [LR05, Chapter 11.2], the distribution is well approximated by the normal distribution, cf. (3.3.6). In consequence as showed by the Wilks theorem [Wil38], the log-likelihood ratio function is well approximated by the chi-squared distribution. Classical schemes like energy detection [Urk67; ATJ14] exploit this fact and identify white spaces based on simple hypotheses, where the average path gain obtained by the summation of multiple independent ob-

servations is approximated by a constant function. Simple hypotheses are generally a good approximation, if the summation of multiple independent observations is considered.

However, for single measurements (complex spectrum samples with two degrees of freedom) on a frequency-varying wireless channel, the constant property does apparently not hold for the SCP, see also the varying power spectrum in Figure 4.2. I.e., the CLT as well as the Wilks theorem are not valid and energy detection may misinterpret samples with low power as gray spaces, increasing the interference to the PU system. Thus, an SU is faced with the problem to distinguish gray spaces from channel fades, which cannot be solved by a mere simple hypothesis test on the power spectrum at the receiver. For successful identification of gray spaces we need to exploit the varying SCP from the BS, which is represented by the composite alternative hypothesis $\mathcal{H}_{1,c}$ with $\hat{\mathbf{v}}_{[f]} > 0$. Furthermore, as explained in Section 2.5.2, the composite alternative hypothesis $\mathcal{H}_{1,c}$ with $\hat{\mathbf{v}}_{[f]} > 0$ is one-sided, i.e. directional. Consequently, given $\hat{\mathbf{z}}, \sigma_n^2$ and varying SCP with $\hat{\mathbf{v}}_{[f]} > 0$, we can apply the log-likelihood ratio test for one-sided composite binary hypotheses as given in (2.6.26) by

$$\Lambda(\hat{\mathbf{z}}_{[f]}, \hat{\mathbf{v}}_{[f]} > 0) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,c}}{\geq}} \rho_{c,\Lambda}. \quad (4.3.2)$$

Thus, the null hypothesis \mathcal{H}_0 is accepted, if the LLR lies above a certain composite threshold $\rho_{c,\Lambda}$.

So now we turn our attention to the composite threshold $\rho_{c,\Lambda}$. We want to select a composite threshold $\rho_{c,\Lambda}$, such that an SU can guarantee certain protection to the PU system. In particular, to enable protection of occupied PU resources, the log-likelihood ratio test has to control the interference of an SU imposed on the PU system by regulating the false alarm rate P_{fa} . Therefore, to control the interference, we have to calculate the composite threshold $\rho_{c,\Lambda}(P_{fa})$ for a desired false alarm rate P_{fa} , cf. (2.6.28). Eventually, *we have the **log-likelihood ratio test for one-sided composite binary hypotheses achieving a desired false alarm rate P_{fa} as given in (2.6.30) by***

$$\Lambda(\hat{\mathbf{z}}_{[f]}, \hat{\mathbf{v}}_{[f]} > 0) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,c}}{\geq}} \rho_{c1,\Lambda}(P_{fa}). \quad (4.3.3)$$

The precise definition of the LLR function as well as the threshold will be part of Section 4.3.3, where we will adapt the definitions of (2.6.11) and (2.6.28) for our system model. Moreover, we will derive our gray space test for practical implementations, where we exploit the results on empirical cumulative distribution function and utilize the threshold given in (2.6.33).

Even though we now have the test (4.3.3) for gray space detection at hand, we miss several information to successfully apply this test. Initially, an SU just has information about its power spectrum $\hat{\mathbf{z}}$. In addition, the variance σ_n^2 of additive white Gaussian noise (AWGN) can be obtained from measurements at idle time-frequency regions like guard bands between two adjacent PU spectrum bands. However, to apply the test in (4.3.3) an SU requires (i) the SCP $\hat{\mathbf{v}}$ corresponding to the power spectrum measurement $\hat{\mathbf{z}}$ for the calculation of Λ , and (ii) the distribution of the SCP to obtain the composite threshold $\rho_{c,\Lambda}(\mathbf{P}_{\text{fa}})$. To obtain the missing information about the SCP, *we propose **gray space detection procedure** consisting of three steps.*

1. **Step:** We perform **energy detection** based on simple hypotheses to obtain a preliminary estimate of the BS resource allocation $\tilde{\mathbf{S}}_{\text{E}}$. There we exploit the fact that, given a desired detection rate, energy detection does not need information about the SCP, cf. (2.6.25). The set $\tilde{\mathbf{S}}_{\text{E}}$ shall be used to estimate the SCP via amplitude-based channel estimation in step 2.

Step 1 will be discussed in Section 4.3.1, where we will primarily recap energy detection.

2. **Step:** Given the set $\tilde{\mathbf{S}}_{\text{E}}$ obtained in step 1, we estimate the SCP $\tilde{\mathbf{v}}$ from the BS via **amplitude-based channel estimation**.

Step 2 will be discussed in Section 4.3.2, where we will exploit the results on amplitude-based channel estimation, cf. Section 3.6.2, and refer to the estimation algorithms in Section 3.5.

3. **Step:** Given $\tilde{\mathbf{v}}$ obtained in step 2, we perform the **log-likelihood ratio test** described in (4.3.3) to obtain the final estimate of the BS resource allocation $\hat{\mathbf{S}}_{\text{G}}$, which will be discussed in Section 4.3.3. In particular, the log-likelihood ratio test classifies the spectrum samples and mark them as gray space or occupied by the BS.

Step 3 will be discussed in Section 4.3.3, where we will also derive the LLR function $\Lambda(\hat{\mathbf{z}}_{[f]}, \hat{\mathbf{v}}_{[f]} > 0)$ as well as the composite threshold $\rho_{c1,\Lambda}(\mathbf{P}_{\text{fa}})$ required for the log-likelihood ratio test. Furthermore, we derive our gray space test, which exploits the empirical cumulative distribution function to operate on a practical composite threshold.

4.3.1 Step 1: Energy Detection

Step 1 of gray space detection effectively comprises the application of energy detection [Urk67; ATJ14]. In general, classical energy detection is used for

unknown transmit signals and multiple independent observation. However, here we deal with single observations of a frequency-varying wireless channel. Therefore, energy detection can be generally outperformed by other detection schemes. However, in step 1 we exploit the fact that energy detection in (2.6.25) does not require information about the spectral channel power (SCP) to estimate BS resource allocation $\tilde{\mathbf{S}}_E$.

Initially, an SU just has information about its power spectrum $\hat{\mathbf{z}}$ and the noise variance $\sigma_{\mathbf{n}}^2$, where $\sigma_{\mathbf{n}}^2$ can be obtained from measurements at idle spectrum parts like the guard band. However, the SU has no information about the non-central parameter, cf. (A.9.1), here in the form of $\hat{\mathbf{v}}$ denoting the SCP. Thus, to perform the log-likelihood ratio test in (4.3.3), we have to estimate $\hat{\mathbf{v}}$. Since only spectrum samples $\hat{\mathbf{z}}_{[f]}$ at $f \in \mathbb{S}_1$ carry information about $\hat{\mathbf{v}}$, cf. (4.2.4), we firstly have to estimate a preliminary BS resource allocation $\tilde{\mathbf{S}}_E$. This preliminary BS resource allocation is used to estimate the SCP from the BS, where the SCP will be used to refine the BS resource allocation $\tilde{\mathbf{S}}_G$ via the log-likelihood ratio test. Thus, based on $\hat{\mathbf{z}}$ and $\sigma_{\mathbf{n}}^2$ without $\hat{\mathbf{v}}$, in step 1 of gray space detection an SU has to estimate the resource allocation $\tilde{\mathbf{S}}_E$ of the BS, i.e. has to decide between the null hypothesis \mathcal{H}_0 and the alternative hypothesis \mathcal{H}_1 of (4.2.4).

To find a suitable test depending on $\hat{\mathbf{z}}$ and $\sigma_{\mathbf{n}}^2$ only, we exploit the theory on hypothesis testing as well as signal detection introduced and discussed in Sections 2.5 and 2.6. Reviewing Table 2.3, we observe that the simple hypotheses test to achieve a desired detection probability P_d (probability that a gray space is correctly detected) in Section 2.6.1 just depends on $\hat{\mathbf{z}}$ and $\sigma_{\mathbf{n}}^2$. Consequently, we apply the log-likelihood ratio test for simple hypotheses to achieve a desired P_d from (2.6.25)

$$\hat{\mathbf{z}}_{[f]} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\gtrless}} \rho_{s,\hat{\mathbf{z}}}(\mathbf{P}_d) = -\sigma_{\mathbf{n}}^2 \cdot \ln(1 - P_d). \quad (4.3.4)$$

In the following, we will also refer to this test as energy detection achieving a desired detection probability P_d . The simple hypotheses considered in (4.3.4) are given by

$$\mathcal{H}_0 : \hat{\mathbf{z}}_{[f]} \sim f_{\Gamma}(\hat{\mathbf{z}}_{[f]}; \sigma_{\mathbf{n}}^2, 1), \quad (4.3.5)$$

$$\mathcal{H}_{1,s} : \hat{\mathbf{z}}_{[f]} \sim f_{\chi_{sn}^2}(\hat{\mathbf{z}}_{[f]}; 2, \tilde{\mathbf{v}}_{[f]}, \frac{1}{2}\sigma_{\mathbf{n}}^2), \tilde{\mathbf{v}}_{[f]} = c, \quad (4.3.6)$$

where the simple null hypothesis \mathcal{H}_0 is obtained from (2.6.8) and the simple alternative hypothesis $\mathcal{H}_{1,s}$ from (2.6.10).

However, recap that energy detection in (4.3.4) has two disadvantages for our scenario, cf. Section 4.3. First, (4.3.4) just uses simple hypotheses,

which does not consider frequency-varying wireless channel. Second, the detection rate P_d cannot directly control the SU interference imposed on the PU system. Therefore, energy detection in (4.3.4) is applied to acquire a sufficient number of power spectrum measurements with $\hat{\mathbf{z}}_{[f]}$ at $f \in \mathbb{S}_1 : \mathcal{H}_{1,s}$, while keeping the number of noise measurements with $\hat{\mathbf{z}}_{[f]}$ at $f \in \mathbb{P} \setminus \mathbb{S}_1 : \mathcal{H}_0$ low. Given energy detection (4.3.4) as well as $\hat{\mathbf{z}}$ and $\sigma_{\mathbf{n}}^2$, an SU estimates the BS resource allocation required for amplitude-based channel estimation in step 2 by

$$\tilde{\mathbb{S}}_E = \{f \in \mathbb{P} : \hat{\mathbf{z}}_{[f]} \geq -\sigma_{\mathbf{n}}^2 \cdot \ln(1 - P_d)\}. \quad (4.3.7)$$

There, the detection rate P_d controls the probability, that the BS resource allocation estimate $\tilde{\mathbb{S}}_E$ contains spectrum samples of the null hypothesis \mathcal{H}_0 , i.e. samples containing just noise with $\hat{\mathbf{z}}_{[f]} = |\hat{\mathbf{n}}_{[f]}|^2$ at $f \in \mathbb{P} \setminus \mathbb{S}_1$. A high ratio of $\mathcal{H}_{1,s}$ samples to \mathcal{H}_0 samples shall stabilize the amplitude-based channel estimation in step 2.

In simulations we will use $\hat{\mathbf{z}}_{[f]} \geq 4\sigma_{\mathbf{n}}^2$ in step 1, which corresponds to detection rate of $P_d \approx 98.17\%$, cf. Table A.4 of Appendix A.7. Therefore, we will contain less than 2% of noise samples, leaving a ratio of approximately 53.6. However, keep in mind that this is just valid for simple hypotheses. Here, we use simple hypothesis testing on actual composite hypotheses. Therefore, especially in scenarios with low SNR, the frequency-varying wireless channel will cause lower detection rate than the configured P_d of energy detection.

In addition, we will use the approach of energy detection also for comparison with our proposed gray space detection scheme. There, we will apply **general energy detection** with

$$\hat{\mathbf{z}}_{[f]} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\geq}} \rho_{s,\hat{\mathbf{z}}}, \quad (4.3.8)$$

where the simple threshold $\rho_{s,\hat{\mathbf{z}}}$ is selected to obtain a performance evaluation of energy detection via the receiver operating characteristic [Tre01, Chapter 2.2.2].

4.3.2 Step 2: Amplitude-Based Channel Estimation

In step 2 of gray space detection, we exploit the results of amplitude-based channel estimation derived in Section 3.6.2. Given the estimated BS resource allocation $\tilde{\mathbb{S}}_E$ from energy detection in step 1 as well as power spectrum measurements $\hat{\mathbf{z}}_{[f]}$ at $f \in \tilde{\mathbb{S}}_E$, the amplitude-based channel estimation in step 2 estimates the spectral channel power (SCP) from the BS. The estimate of the SCP is required for step 3 to perform the log-likelihood ratio test given in (4.3.3).

Let us start with the alternative hypothesis \mathcal{H}_1 of (4.2.4), where the power spectrum is given by

$$\hat{\mathbf{z}}_{[f]} = \hat{\mathbf{v}}_{[f]} + \hat{\mathbf{e}}_{[f]}, \quad f \in \tilde{\mathbb{S}}_E \subseteq \mathbb{S}_1 \quad : \mathcal{H}_1, \quad (4.3.9)$$

$$\begin{aligned} \text{where } \hat{\mathbf{e}}_{[f]} &= \hat{\mathbf{u}}_{[f]} \hat{\mathbf{h}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{n}}_{[f]} + |\hat{\mathbf{n}}_{[f]}|^2, \\ \hat{\mathbf{z}}_f &= |\hat{\mathbf{w}}_{[f]}|^2, \\ \hat{\mathbf{v}}_f &= |\hat{\mathbf{h}}_{[f]}|^2. \end{aligned}$$

Consequently, we assume that the estimated set $\tilde{\mathbb{S}}_E$ just contains elements of the true BS resource allocation \mathbb{S}_1 . However, in practice $\tilde{\mathbb{S}}_E$ also contains some spectrum samples at $f \in \mathbb{P} \setminus \mathbb{S}_1$, which result in an additional error of the amplitude-based channel estimation. In the following, we will discuss amplitude-based channel estimation for our scenario.

amplitude-based channel estimation will estimate the SCP $\hat{\mathbf{v}}$ via its time domain representation, the circular autocorrelation of the channel impulse response (CIR) denoted by \mathbf{v} , cf. (3.6.26). Therefore, we apply the derivations of Section 3.6.2 with the DFT submatrix

$$\Phi = \mathbf{F}_{[\tilde{\mathbb{S}}_E, \mathbb{N}_{\text{CAC}}]}, \quad (4.3.10)$$

where the definition of the DFT matrix is given in (3.6.1), so that (4.3.9) is reformulated to

$$\hat{\mathbf{z}}_{[\tilde{\mathbb{S}}_E]} = \Phi \cdot \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]} + \hat{\mathbf{e}}_{[\tilde{\mathbb{S}}_E]}. \quad (4.3.11)$$

Consequently, we just consider the circular autocorrelation of the CIR (CAC) within the set \mathbb{N}_{CAC} defined in (3.6.29), which is restricted by the dimension of cyclic-prefix with $|\mathbb{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1$. To solve (4.3.11) for $\mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}$, we have to

$$\begin{aligned} &\text{find } \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}, \\ &\text{s.t. } \left\| \hat{\mathbf{z}}_{[\tilde{\mathbb{S}}_E]} - \Phi \cdot \mathbf{v}_{[\mathbb{N}_{\text{CAC}}]} \right\|_2^2 \leq 2\lambda \cdot |\tilde{\mathbb{S}}_E| \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2), \end{aligned} \quad (4.3.12)$$

which is obtain from (3.6.42). The estimation error is bounded by the second raw moment of $\hat{\mathbf{e}}_{[\tilde{\mathbb{S}}_E]}$ obtained from (3.6.41), where we require knowledge about $\sigma_{\mathbf{n}}^2$ and $\sigma_{\mathbf{h}}^2$. As noted above, the noise variance $\sigma_{\mathbf{n}}^2$ can be obtained from measurements at idle spectrum parts like the guard band. A straightforward approach to obtain also a coarse approximation of $\sigma_{\mathbf{h}}^2$ is to consider

the variance of the power spectrum at $f \in \mathbb{S}_1$, i.e.

$$\sigma_{\mathbf{h}}^2 \approx \left\| \hat{\mathbf{z}}_{[\tilde{\mathbf{S}}_{\mathbf{E}}]} \right\|_2 / |\tilde{\mathbf{S}}_{\mathbf{E}}|. \quad (4.3.13)$$

However, in our simulations we have set $\sigma_{\mathbf{h}}^2 = 1/N_{\text{DFT}}$, since $\mathbb{E}[\|\mathbf{h}\|_2] = 1$ for our considered channel models, cf. (2.3.3). For detailed derivation of the second raw moment we refer to (3.6.35) to (3.6.40) in Section 3.6.2.

The solution of (4.3.12) can be found by the different estimation algorithms presented in Section 3.5, of which we will state the main equations in the following.

The **reduced-rank least squares (RRLS) solution** (3.5.17) is obtained by

$$\tilde{\mathbf{v}}_{[\text{N}_{\text{CAC}}]} = \mathbf{E}_{\text{RRLS}} \cdot \hat{\mathbf{z}}_{[\tilde{\mathbf{S}}_{\mathbf{E}}]}, \quad (4.3.14)$$

where given the singular value decomposition $\Phi = \mathbf{U}\Sigma\mathbf{V}^H$, the **RRLS estimator** (3.5.16) is calculated by

$$\mathbf{E}_{\text{RRLS}} = \mathbf{V} \cdot \Sigma_L^{-1} \cdot \mathbf{U}^H, \quad (4.3.15)$$

$$\text{with } L = \left\{ l : \Sigma_{[l,l]} \geq \sqrt{2\lambda \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2)} \cdot \sqrt{\frac{|\text{N}_{\text{CAC}}|}{\mathbb{E}[\|\mathbf{v}_{[\text{N}_{\text{CAC}}]}\|_2^2]}} \right\}.$$

The bound on the estimation error is refined by (3.6.43). The tuning parameter λ is typically set to one, which we have done for our simulations as well. Furthermore, for our simulations we have $\mathbb{E}[\|\mathbf{v}_{[\text{N}_{\text{CAC}}]}\|_2^2] = |\text{N}_{\text{CAC}}|$, so

that the term $\sqrt{|\text{N}_{\text{CAC}}| / \mathbb{E}[\|\mathbf{v}_{[\text{N}_{\text{CAC}}]}\|_2^2]}$ vanishes.

Thus, RRLS has to estimate $|\text{N}_{\text{CAC}}|$ number of entries of \mathbf{v} . In contrast to that, the CS methods exploit the sparsity of $\mathbf{v}_{[\text{N}_{\text{CAC}}]}$ and therefore, just have to estimate $|\text{supp}(\mathbf{v})|$ number of entries of \mathbf{v} , where $|\text{supp}(\mathbf{v})| \ll |\text{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1$, cf. (3.6.49).

The **basis pursuit denoising (BPDN) solution** (3.5.18) [CT05] is obtained by

$$\begin{aligned} \tilde{\mathbf{v}}_{[\text{N}_{\text{CAC}}]} &:= \underset{\mathbf{v}_{[\text{N}_{\text{CAC}}]}}{\text{argmin}} \left\| \mathbf{v}_{[\text{N}_{\text{CAC}}]} \right\|_1, \\ \text{s.t.} \quad &\left\| \Phi \mathbf{v}_{[\text{N}_{\text{CAC}}]} - \hat{\mathbf{z}}_{[\tilde{\mathbf{S}}_{\mathbf{E}}]} \right\|_2^2 \leq 2\lambda \cdot |\tilde{\mathbf{S}}_{\mathbf{E}}| \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2), \end{aligned} \quad (4.3.16)$$

Algorithm 4.1 Stopping criterion extension for OMP Algorithm 3.1

Input: $\hat{\mathbf{z}} \in \mathbb{C}^{|\tilde{\mathbb{S}}_{\text{E}}|}$, $\Phi \in \mathbb{C}^{|\tilde{\mathbb{S}}_{\text{E}}| \times |\mathbb{N}_{\text{CAC}}|}$, λ , $\sigma_{\hat{\mathbf{n}}}^2$, $\sigma_{\hat{\mathbf{h}}}^2$, s

Initialize: $\mathbb{B} = \emptyset$, $\tilde{\mathbf{e}} = \hat{\mathbf{z}}$, $\tilde{\mathbf{v}} = \mathbf{0}$, $s = \min[|\mathbb{N}_{\text{CAC}}|, s]$

```

1:  $\rho = 2\lambda \cdot |\tilde{\mathbb{S}}_{\text{E}}| \cdot \sigma_{\hat{\mathbf{n}}}^2 (\sigma_{\hat{\mathbf{n}}}^2 + \sigma_{\hat{\mathbf{h}}}^2)$ 
2: while  $\left( \|\tilde{\mathbf{e}}\|_2^2 \geq \rho \right)$  and  $(\|\tilde{\mathbf{v}}\|_0 < s)$  and  $\left( \|\tilde{\mathbf{e}}' - \tilde{\mathbf{e}}\|_2^2 > \frac{\rho}{|\mathbb{N}_{\text{CAC}}|} \right)$  do
3:    $\tilde{\mathbf{e}}' = \tilde{\mathbf{e}}, \tilde{\mathbf{v}}' = \tilde{\mathbf{v}}$ 
4:    $\mathbf{a} = |\Phi^H \cdot \tilde{\mathbf{e}}|$ 
5:    $m = \operatorname{argmax}_m \frac{\mathbf{a}_{[m]}}{\|\Phi_{[:,m]}\|_2}$ 
6:    $\mathbb{B} = \mathbb{B} \cup \{m\}$ 
7:    $\tilde{\mathbf{v}}_{[\mathbb{B}]} = (\Phi_{[:,\mathbb{B}]})^+ \cdot \hat{\mathbf{z}}$ 
8:    $\tilde{\mathbf{e}} = \hat{\mathbf{z}} - \Phi \cdot \tilde{\mathbf{v}}$ 
9: end while
10: if  $\left( \|\tilde{\mathbf{e}}' - \tilde{\mathbf{e}}\|_2^2 < \rho \right)$  then
11:    $\tilde{\mathbf{e}} = \tilde{\mathbf{e}}', \tilde{\mathbf{v}} = \tilde{\mathbf{v}}'$ 
12: end if
13: return  $\tilde{\mathbf{v}}, \tilde{\mathbf{e}}$ 

```

where we apply the bound of (4.3.12).

In addition to the stopping criteria described for orthogonal matching pursuit (OMP) in Section 3.5.3 [EK12, Algorithm 8.3], cf. line 1 in Algorithm 3.1, we applied one further stopping criterion for gray space detection. The complete **OMP algorithm** including the third stopping criterion is given in Algorithm 4.1.

The third criterion at line 2 in Algorithm 4.1 terminates the algorithm based on the gradient of the adversarial error. Thus, the noise improvement captured by the difference of the residual in each iteration has to be larger than the normalized adversarial error

$$\|\tilde{\mathbf{e}}' - \tilde{\mathbf{e}}\|_2^2 > \frac{\rho}{|\mathbb{N}_{\text{CAC}}|} = 2\lambda \cdot \frac{|\tilde{\mathbb{S}}_{\text{E}}|}{|\mathbb{N}_{\text{CAC}}|} \cdot \sigma_{\hat{\mathbf{n}}}^2 (\sigma_{\hat{\mathbf{n}}}^2 + \sigma_{\hat{\mathbf{h}}}^2). \quad (4.3.17)$$

For completeness, the first stopping criterion of (3.5.23) is given by the ℓ_2 -norm of the residual with

$$\left\| \tilde{\mathbf{e}} \right\|_2^2 > \rho = 2\lambda \cdot \left| \tilde{\mathbf{S}}_{\mathbf{E}} \right| \cdot \sigma_{\mathbf{n}}^2 (\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{h}}^2). \quad (4.3.18)$$

The third stopping criterion (4.3.17) does only slightly differ to the first stopping criterion (4.3.18). In particular, while (4.3.18) is compared to the total residual of the s -sparse solution $\tilde{\mathbf{v}}_{[\mathbb{B}]}$ in each iteration, (4.3.17) is just compared to the residual change in the current iteration. Therefore we can say, if the induced error $\frac{\rho}{|\mathbb{N}_{\text{CAC}}|}$ is larger than the gain due to the change of the residual, then the previous solution $\tilde{\mathbf{v}}'$ will be a better estimate. Thus, OMP in Algorithm 4.1 will terminate, if one of the three conditions at line 2 are not fulfilled. In empirical evaluations we have seen that OMP with the first stopping criterion (4.3.18) and the third stopping criterion (4.3.17) achieves better performance than if we just use the first stopping criterion.

Therefore, Algorithm 3.1 is extended to Algorithm 4.1 by some lines to include the third stopping criterion. The additions are highlighted in **dark red color** and comprise lines 2, 3, 10 and 11 in Algorithm 4.1. The lines 3, 10 and 11 in Algorithm 4.1 are required to load and store the previous estimate for comparison. In more detail, (i) at line 2 we included (4.3.17) for the while-loop condition, (ii) at line 3 we store the previously calculated residual $\tilde{\mathbf{e}}'$ and $\text{CAC } \tilde{\mathbf{v}}' = \tilde{\mathbf{v}}$ for comparison and later use, and (iii) at line 11 we revert the residual $\tilde{\mathbf{e}}$ and the $\text{CAC } \tilde{\mathbf{v}} = \tilde{\mathbf{v}}$ to the previous solutions, if line 10 is true and the current solution does not fulfill (4.3.17).

In summary, we have three amplitude-based channel estimation methods: RRLS (4.3.14), BPDN (4.3.16) and OMP in Algorithm 4.1. The estimated SCP based on these methods will be applied to the log-likelihood ratio test in the following step 3. Therefore, we have three different SCP estimates to evaluate the performance gray space detection.

4.3.3 Step 3: Log-Likelihood Ratio Test

Eventually, we arrive at step 3 of gray space detection, where based on the log-likelihood ratio test given in (4.3.3), we classify the spectrum samples within the transmission band \mathbb{P} as \mathcal{H}_0 : gray space or \mathcal{H}_1 : occupied by the BS. In the previous step, we obtained an estimate for the SCP, which is required to LLR function of (4.3.3). However, besides a coarse estimate of $\sigma_{\mathbf{h}}^2$ in (4.3.13), we still lack of information about the distribution of the SCP, which is required to obtain the composite threshold $\rho_{\text{c},\Lambda}(\text{P}_{\text{fa}})$.

In this Section we will describe the last step of gray space detection comprising the specific definition of the log-likelihood ratio test. Therefore, at

first we will recap the considered hypotheses, the simple null hypothesis \mathcal{H}_0 and the composite alternative hypothesis $\mathcal{H}_{1,c}$. Then, given the results obtained in Section 2.6, we will state the LLR function. Furthermore, we will derive estimates for the distribution of the SCP, so that we can calculate the composite threshold. Given the LLR function and the composite threshold, we will derive an approximated composite threshold and the corresponding gray space test. In addition, we will provide a individual threshold for a simple log-likelihood ratio test, i.e. individual energy detection. Finally we will discuss our proposed gray space test based on empirical cumulative distribution function and its performance towards a uniformly most powerful test.

As described in our system model in Section 4.2, we have an SU observing the power spectrum, where the two cases, \mathcal{H}_0 and \mathcal{H}_1 , are present, cf. (4.2.4). However here, in contrast to energy detection in step 1, we do not just consider simple hypotheses, but composite hypotheses. In particular, while \mathcal{H}_0 still represents the simple null hypothesis, $\mathcal{H}_{1,c}$ describes the composite alternative hypothesis.

While the simple null hypothesis \mathcal{H}_0 denotes a gray space event (first case of (4.2.4)), the composite alternative hypothesis $\mathcal{H}_{1,c}$ denotes a primary user (PU) event (second case of (4.2.4)). The distributions of each hypotheses are obtained from (2.6.8) and (2.6.9) with

$$\mathcal{H}_0 : \hat{\mathbf{z}}_{[f]} \sim f_{\Gamma}(\hat{\mathbf{z}}_{[f]}; \sigma_{\hat{\mathbf{n}}}^2, 1), \quad (4.3.19)$$

$$\mathcal{H}_{1,c} : \hat{\mathbf{z}}_{[f]} \sim f_{\chi_{sn}^2}(\hat{\mathbf{z}}_{[f]}; 2, \tilde{\mathbf{v}}_{[f]}, \frac{1}{2}\sigma_{\hat{\mathbf{n}}}^2), \tilde{\mathbf{v}}_{[f]} > 0. \quad (4.3.20)$$

Based on these two PDFs, the log-likelihood ratio derived at (2.6.11) and (2.6.12) is given by

$$\Lambda(\hat{\mathbf{z}}_{[f]}, \tilde{\mathbf{v}}_{[f]}) = \ln \left(\frac{f_{\chi_{nc}^2} \left(\frac{2 \cdot \hat{\mathbf{z}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2}; 2, \frac{2 \cdot \tilde{\mathbf{v}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2} \right)}{f_{\chi^2} \left(\frac{2 \cdot \hat{\mathbf{z}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2}; 2 \right)} \right). \quad (4.3.21)$$

Thus, with the LLR function in (4.3.21) we obtained the left side of the log-likelihood ratio test in (4.3.3).

Now, let us turn our attention to the right side of our log-likelihood ratio test in (4.3.3), the composite threshold $\rho_{c,\Lambda}(\mathbf{P}_{fa})$ for a desired false alarm rate \mathbf{P}_{fa} . The desired false alarm probability \mathbf{P}_{fa} denotes the probability, that the resource occupied by the BS is not detected. Therefore, the false alarm probability \mathbf{P}_{fa} significantly affects the protection of the PU

system against the interference of an SU. To obtain the composite threshold for the log-likelihood ratio test, the distribution of the spectral channel power (SCP) and the noise variance σ_n^2 is required, cf. (2.6.28). Since the distribution of the channel is not known, only an approximated composite threshold $\rho_{c,\Lambda}(\mathbf{P}_{\text{fa}}, N)$ can be estimated. Thus, the **approximated composite false alarm threshold** from (2.6.33) is given by

$$\rho_{c,\Lambda}(\mathbf{P}_{\text{fa}}, N) = \max(\rho \in \mathbb{R} : \mathbf{P}_{\text{fa}} \geq \mathbf{P}(\rho, N)), \quad (4.3.22)$$

$$\text{where } \mathbf{P}(\rho, N) = \frac{1}{N} \sum_{n=1}^N \mathbf{F}_{\chi_{sn}^2, \hat{\mathbf{z}}} \left(\Lambda^{-1} \left(\rho, \tilde{\mathbf{v}}_{[f_n]} \right); 2, \tilde{\mathbf{v}}_{[f_n]}, \frac{\sigma_n^2}{2} \right), \quad (4.3.23)$$

where f_n is sampled uniformly at random, without replacement, from \mathbb{P} . The inverse log-likelihood ratio function is taken from (2.6.27) with

$$\Lambda^{-1} \left(\rho_{c,\Lambda}, \tilde{\mathbf{v}}_{[f]} \right) := \max \left(\hat{\mathbf{z}}_{[f]} \in \mathbb{R}_{\geq 0} : \rho_{c,\Lambda} \geq \Lambda \left(\hat{\mathbf{z}}_{[f]}, \tilde{\mathbf{v}}_{[f]} \right) \right), \quad (4.3.24)$$

where the log-likelihood ratio function of (4.3.21) is applied.

As described in our system model in Section 4.2, we just consider a single OFDM symbol for spectrum power measurements. Therefore, the SCP estimated in step represents just a single realization of the channel. Consequently, the samples $\tilde{\mathbf{v}}_{[f_n]}$ of the estimated SCP just reflect the distribution of the current channel realization, cf. Glivenko–Cantelli theorem [Can33; Gli33] in (2.6.32). The distribution of the complete time variant channel is not known at the current observation. Since, we approximated only the PDF of a single channel realization, (4.3.23) does not reflect the probability of the whole channel statistics and the probability $\mathbf{P}(\rho, N)$ will differ. Thus, our approximated composite false alarm threshold in (4.3.22) will change for each channel realization, even if we increase N .

The log-likelihood ratio test in (2.6.15) is a UMP test, if the threshold $\rho_{c,\Lambda}$ is constant independent of the measurement, here in the form of the power spectrum $\hat{\mathbf{z}}$. Since our approximated threshold changes for each channel realization, the log-likelihood ratio test constructed with this threshold will not be a UMP test. However, we expect that the more stable a threshold over different channel realization, the closer the performance of the log-likelihood ratio test to a UMP test. Thus, the approximation of threshold in (4.3.22) can be regarded as a stabilization of the threshold by achieving a desired false alarm rate \mathbf{P}_{fa} .

Consequently, we can refine now the log-likelihood ratio test of (4.3.3) to achieve a desired false alarm rate \mathbf{P}_{fa} . *The one-sided composite log-likelihood*

ratio test obtained from (2.6.36), also called the **gray space test**, is

$$\Lambda \left(\hat{\mathbf{z}}_{[f]}, \tilde{\mathbf{v}}_{[f]} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,c}}{\geq}} \rho_{c,\Lambda} (P_{fa}, N). \quad (4.3.25)$$

Thus, we estimate BS resource allocation via the gray space test by

$$\tilde{\mathbb{S}}_G = \{f \in \mathbb{P} : \Lambda \left(\hat{\mathbf{z}}_{[f]}, \tilde{\mathbf{v}}_{[f]} \right) \geq \rho_{c,\Lambda} (P_{fa}, N)\}, \quad (4.3.26)$$

where the set $\mathbb{P} \setminus \tilde{\mathbb{S}}_G$ contains the estimated gray spaces.

Instead of approximating the composite threshold $\rho_{c,\Lambda}(P_{fa})$ for one channel realization $\tilde{\mathbf{v}}$, we can also obtain an **individual threshold** for each spectrum sample $\tilde{\mathbf{v}}_{[f]}$ by (2.6.22)

$$\rho_{s,\hat{\mathbf{z}}} (P_{fa}) = F_{\hat{\mathbf{z}}, \chi_{sn}^2}^{-1} \left(P_{fa}; 2, \tilde{\mathbf{v}}_{[f]}, \frac{1}{2} \sigma_{\hat{\mathbf{n}}}^2 \right). \quad (4.3.27)$$

The individual threshold describes a simple hypothesis for each spectrum sample. Thus, we obtain an **individual energy detector**

$$\hat{\mathbf{z}}_{[f]} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\geq}} F_{\hat{\mathbf{z}}, \chi_{sn}^2}^{-1} \left(P_{fa}; 2, \tilde{\mathbf{v}}_{[f]}, \frac{1}{2} \sigma_{\hat{\mathbf{n}}}^2 \right), \quad (4.3.28)$$

which represents the simple log-likelihood ratio test (2.6.23) to achieve a desired false alarm rate P_{fa} for each spectrum sample $\hat{\mathbf{z}}_{[f]}$ independently. Since the threshold changes for each spectrum sample, the individual energy detector is also not a uniformly most powerful test. Even more, the variance of the individual threshold will be larger than the approximated composite threshold of the gray space test in (4.3.25), since the gray space test uses an averaged threshold. However, the computation of the individual threshold is less complex than the approximated composite threshold.

Similar to (4.3.8), given the LLR in (4.3.21) and an threshold $\rho_{c,\Lambda}$ with an arbitrary constant value, we can refine (4.3.3) as the **general composite log-likelihood ratio test** by

$$\ln \left(\frac{f_{\chi_{nc}^2} \left(\frac{2 \cdot \hat{\mathbf{z}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2}; 2, \frac{2 \cdot \tilde{\mathbf{v}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2} \right)}{f_{\chi^2} \left(\frac{2 \cdot \hat{\mathbf{z}}_{[f]}}{\sigma_{\hat{\mathbf{n}}}^2}; 2 \right)} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{1,s}}{\geq}} \rho_{c,\Lambda}. \quad (4.3.29)$$

Moreover, since the threshold $\rho_{c,\Lambda}$ is constant independent of the channel realization, (4.3.29) represents a UMP test. However as discussed in Sections 2.5 and 2.6, the test (4.3.29) is not practical, since the performance

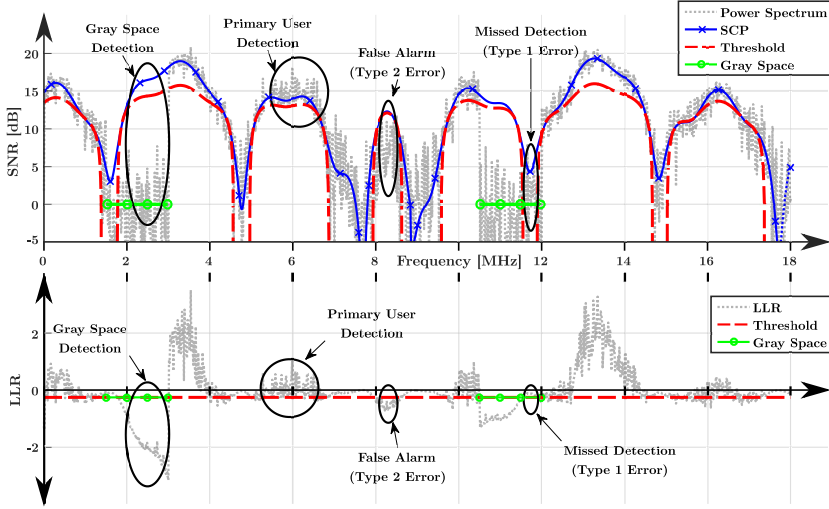


Figure 4.3: Illustration of the log-likelihood ratio test on the received power spectrum and a threshold based on perfect spectral channel power (SCP) knowledge for 10 dB SNR and false alarm rate of $P_{fa} = 0.2$.

in terms of detection rate and false alarm rate cannot be controlled. Nevertheless, the UMP test shows the performance deviation of our practical approaches.

Furthermore, we can use (4.3.29) to evaluate the performance of gray space detection regarding its different amplitude-based channel estimation methods provided in gray space detection step 2. Since the LLR is calculated on the basis of the SCP estimate, the estimated SCP of the amplitude-based channel estimation approaches will affect the performance of (4.3.29). There, the UMP test is achieved, if perfect SCP knowledge is applied on (4.3.29).

A visualization of the composite threshold $\rho_{c,\Lambda}(P_{fa})$ is illustrated in Figure 4.3, where $P_{fa} = 0.2$ and $\sigma_n^2 = 0.1$. The SNR domain and LLR domain are shown. While the threshold is constant in the LLR domain, the threshold differs for each spectrum sample in the SNR domain. The LLR or alternatively the power spectrum samples (gray dots) above the threshold (red dashed line) are detected as occupied BS resources. The remaining samples are considered as gray space.

The four cases of log-likelihood ratio test listed in Table 2.2 are also illustrated in Figure 4.3. While *type 2 errors* (false alarm) are provoked within the true PU regions, *type 1 errors* (missed detection) arouse within the true gray space regions (green). The primary cause for *type 1 errors* is frequency selective fading.

PU subcarriers $ \mathbb{P} $	1200
Cyclic-prefix	144
P_d for $\tilde{\mathbb{S}}_E$	$\approx 98.17\%$
λ	1
Channel Model	EPA
System Load	80%
Blocksize	12 subcarrier
SNR	10 dB
False Alarm Rate	10%
Algorithms	Energy detector Orthogonal matching pursuit (OMP) Reduced-rank least squares (RRLS) Basis pursuit denoising (BPDN)

Table 4.1: Default link level simulation parameters.

4.4 Simulation Setup

If not mentioned otherwise, the link level simulations are carried out with the default parameters in Table 4.1. Each simulation is performed for multiple independent channel realizations. We transmit a 20 MHz LTE downlink signal [3GP20a], where $|\mathbb{P}| = 1200$ subcarriers can be loaded. Thus, only 18 MHz of the complete transmission band \mathbb{P} are used by the PU for transmission. The cyclic-prefix is set to the 3GPP short cyclic-prefix value of 144 samples, representing duration of $4.7 \mu\text{s}$. Therefore, a CIR lying within 144 samples can be reconstructed. The duration of a CAC is approximately twice the duration of a CIR, in particular 287 samples. Thus, the estimators are configured, such that they are able to reconstruct a CAC with the maximum duration of 287 samples.

The transmission alphabet consists of complex samples with random phase and constant amplitude. In addition, the base station has a default system load of 80% and uses a PRB-size of 12, which reflects the number of consecutive subcarriers in one physical resource block (PRB). Consequently, a complete PRB of 12 subcarriers is either loaded or empty. A system load of 80% translates to 80 occupied PRBs with blocksize of 12 subcarriers. There-

fore, 20 PRBs are randomly selected and left blank. These *gray spaces* shall be detected by a secondary user (SU).

Three algorithms for amplitude-based channel estimation are considered at the SU, RRLS, BPDN and OMP. The SPGL1-toolbox of Berg and Friedlander [BF15; BF08] is used to solve the convex BPDN problem stated in (3.5.18). The BS resource allocation $\tilde{\mathbf{S}}_E$ is estimated in gray space detection step 1 with detection probability $P_d \approx 98.17\%$, so that approximately 98% of *gray spaces* are excluded from amplitude-based channel estimation in gray space detection step 2. The tuning parameter of each amplitude-based channel estimator is $\lambda = 1$. Furthermore, perfect knowledge about the SCP as well as the classical energy detector are considered in the simulations. The log-likelihood ratio test is performed with multiple thresholds, such that different pairs of false alarm rate P_{fa} and detection rate P_d are obtained. These pairs are used to generate the receiver operating characteristic.

4.5 Performance Evaluation

Our performance evaluations in this Section are divided in three parts, (i) comparison of amplitude-based channel estimation methods, (ii) comparison of our proposed thresholds for the log-likelihood ratio test, and (iii) comparison of different channel models including our measurement campaign. Each part is compared to classical energy detection as a baseline, cf. (4.3.8).

The first part in Sections 4.5.1 to 4.5.4 evaluates the three different estimation algorithms for amplitude-based channel estimation in step 2. There, we use (4.3.29) to perform the log-likelihood ratio test. We consider several aspects: receiver operating characteristic (ROC) as well as the impact of SNR, system load and physical resource blocksize. The curves of perfect channel are obtained by having full knowledge of the SCP.

The second part in Section 4.5.5 evaluates our gray space test of step 3. There, we compare our gray space test based on the approximated composite threshold (4.3.25) with individual energy detection (4.3.28) and the UMP test (4.3.29).

In the third part we evaluate the performance of the three estimation algorithms for amplitude-based channel estimation in step 2 under different channel models, including our the channel models derived from our measurement campaign. There, we apply again the log-likelihood ratio test of (4.3.29).

4.5.1 Receiver Operating Characteristic

The receiver operating characteristic is used to completely describe the performance of a statistical hypothesis test [Her98]. Therefore, the relation between gray space detection rate and false alarm rate is shown. The sub-figures of Figure 4.4 show the ROC for three different SNR values, a) at 0 dB, b) at 10 dB, c) at 15 dB and d) at 20 dB. The random process denotes the line, where the same results are obtained by just flipping a coin. Thus, the closer the curve of the statistical hypothesis test is to the random process, the worse is their performance. Moreover, for regions where an algorithm operates below the line of the random process, it is even better just flipping a coin for a decision.

At 0 dB SNR in Figure 4.4a) the approaches energy detection, OMP and BPDN perform similar. Still, there is a small gap at low false alarm rates to perfect channel knowledge. Due to the definition of RRLS in (3.5.16), the measurement matrix consists only of zeros and therefore, a vector with mere zeros is estimated. Thus, the ROC of RRLS is just a point at zero false alarm rate and zero detection rate. Consequently, at low SNR region of 0 dB the choice of the spectrum sensing approach does not matter. Further information by estimating the SCP does not get any advantage in the detection procedure. Nevertheless, at 0 dB SNR all algorithms have only a small gain in comparison to a random process.

At 10 dB SNR (Figure 4.4b) the performance differs significantly from the 0 dB SNR region. The performance of energy detection, OMP and BPDN increased strongly. Furthermore, OMP and BPDN achieves higher gray space detection rate than the energy detection, especially at the operation range of spectrum sensing below 10% false alarm rate. However, below 3% false alarm rate the performance of BPDN dramatically decreases due to some rare cases, where the SPGL1-toolbox does not converge. Nevertheless, OMP outperforms all algorithms at all false alarm rates. Thus, information of the SCP increases the performance of the detection procedure, because fading holes can be distinguished from *gray spaces*. However, in contrast to OMP and BPDN, gray space detection will suffer if the RRLS approach is selected for SCP estimation. In particular, at 10 dB a lot of measurements $\hat{\mathbf{z}}$ are excluded from the SCP estimation process, so that the $(|\mathbf{N}_{\text{CAC}}| \times \lceil \widetilde{\mathbf{S}}_1 \rceil)$ RRLS estimator \mathbf{E}_{RRLS} defined in (3.5.17) will be likely ill-posed. Thus, reconstructing the large autocorrelation vector of $|\mathbf{N}_{\text{CAC}}| = 287$ samples by RRLS estimation will likely fail. Finally, there is still a large gap especially at low false alarm rates between perfect channel knowledge and the applied algorithms.

Increasing the SNR by 5 dB to 15 dB (Figure 4.4c), the ROC curves

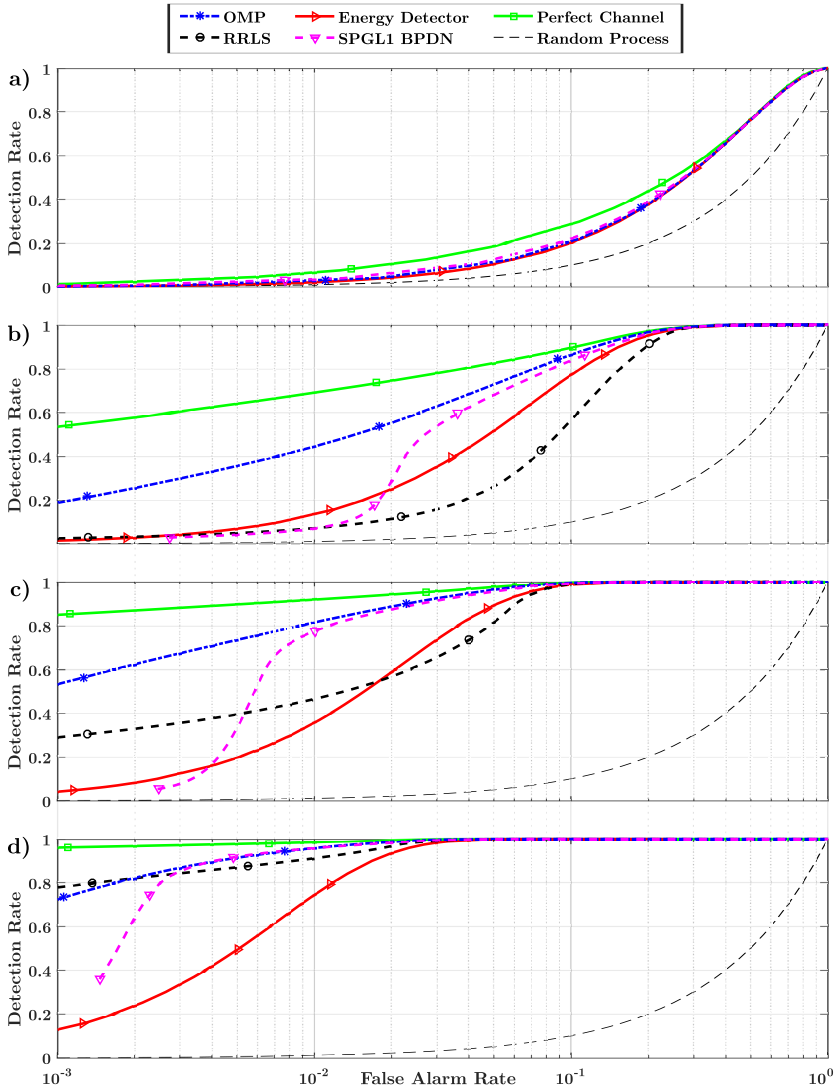


Figure 4.4: The receiver operating characteristic at a) 0 dB, b) 10 dB, c) 15 dB and d) 20 dB SNR for gray space detection vs. false alarm with orthogonal matching pursuit (OMP), basis pursuit denoising (BPDN) and reduced-rank least squares (RRLS).

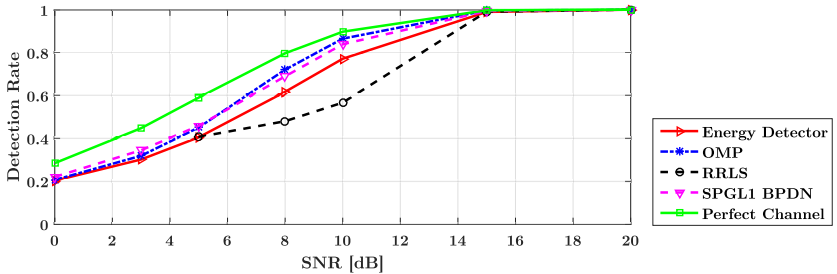


Figure 4.5: Gray space detection for fixed false alarm rate of 10% over SNR.

are shifted to the left, achieving higher performance. OMP still outperforms all algorithms. However, down to 1% false alarm probability BPDN achieves similar performance. Even RRLS achieves better results than energy detection at low false alarm rate. RRLS benefits from low receiver noise variance σ_n^2 and consequently from low noise threshold $\alpha \cdot \sigma_n^2$. Besides the frequency fading holes and the gray space regions, a large distributed number of spectrum samples are used for SCP estimation. Therefore, RRLS, OMP and BPDN are able to estimate a more accurate SCP. Nevertheless, there is still a gap between perfect channel knowledge and OMP as well as OMP and RRLS. Energy detection profits the least from the increased SNR, achieving almost the worst performance of all algorithms.

The trend continuous at the high SNR-regime, here illustrated with SNR of 20 dB (Figure 4.4c). While the ROC curves of the gray space detection algorithms becoming similar, the large gap to energy detection remains at low false alarm rate. Therefore, gray space detection is well suited for the high SNR-regime, where even RRLS is applicable for the selected physical resource blocksize of 12 subcarriers.

4.5.2 Impact of SNR

In general, a cognitive radio system is configured to accomplish a certain primary user (PU) protection. In particular, the target packet error rate of a PU system like LTE is 10% [3GP20d]. Therefore, spectrum sensing approaches have to achieve a false alarm rate below 10%. Otherwise, the PU system suffers from severe secondary user interference, leading to significant performance degradation of the PU. Moreover, to guarantee a smooth operation of the PU, the false alarm rate has to be even far below 10%. Nevertheless, we will focus here on a 10% false alarm rate, because for many systems this is the maximum tolerable error.

In Figure 4.5 we observe the gray space detection rate for a fixed false

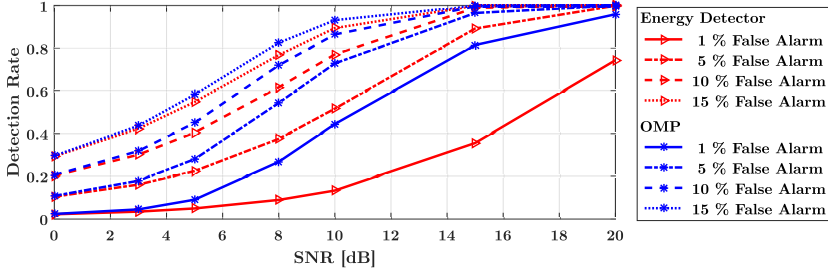


Figure 4.6: Gray space detection with orthogonal matching pursuit (OMP) for multiple fixed false alarm rates over SNR.

alarm rate of 10%. As mentioned before, the RRLS method is only able to achieve good detection results for high SNR regimes. At the other SNR regimes RRLS is simply outperformed by the other approaches. In contrast to RRLS, all other algorithms achieve better results. Gray space detection based on compressed sensing have up to 10% detection gain in comparison to energy detection. Furthermore, still a gap is visible between perfect channel knowledge and the compressed sensing approaches.

For better comparison, we picked OMP and energy detection in Figure 4.6 and give a more detailed view for multiple false alarm rates. To see the trend for higher false alarm rates, a false alarm of 15% is considered. Nevertheless, the focus lies on the lower false alarm rates towards 1%, since the secondary user (SU) interference on the PU system will be insignificant for the a considered primary user target packet error rate of 10%. The lower the false alarm rate, the higher is the gap between OMP and energy detection. In addition, for smaller false alarm rates the maximum gap between both approaches shift to higher SNR. E.g. a false alarm rate of 1% OMP is still able to perform fairly good detection results at 15 dB SNR. In contrast to that, energy detection suffers from significant performance loss. This result is already obvious in Figure 4.4c). Therefore, for practical requirements of PU protection, a gray space detector based on OMP leads to significant increase of the gray space detection rate.

4.5.3 Impact of System Load

Another view on gray space detection is taken in Figure 4.7. At 10 dB SNR and a false alarm rate of 10%, the detection rate is plotted over a PU system load from 10% to 90%. E.g. at 90% system load 10 out of 100 physical resource blocks consisting of 12 subcarriers are left empty.

Since energy detection in (4.3.8) does not consider the SCP and only

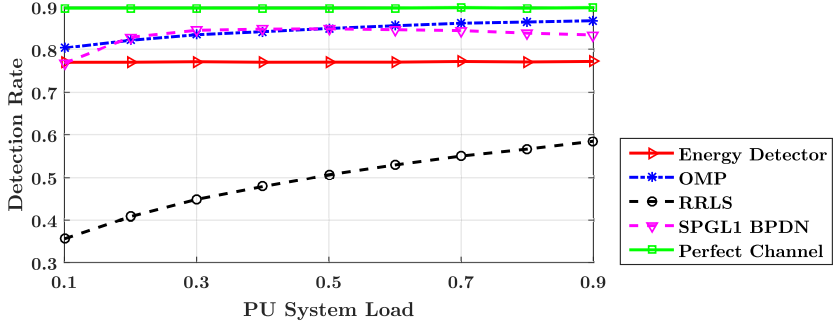


Figure 4.7: Gray space detection over primary user (PU) system load at fixed false alarm rate of 10% and 10 dB SNR.

depends on the receiver noise, the detection rate stays roughly constant at approximately 77%. In contrast to energy detection, RRLS, OMP and BPDN are influenced by the system load. The lower the system load, the less spectrum samples are available for amplitude-based channel estimation. However, with decreasing system load the detection performance of OMP is only slowly decreasing by a few percent. Even at 10% system load OMP achieves better results than energy detection. In contrast to OMP, BPDN stays coarsely constant down to 30% system load. Below 30% system load BPDN approaches quickly the performance of energy detection. In contrast to that, RRLS is strongly affected by system load. However, even at large system load RRLS performs significantly below energy detection. Consequently, OMP and BPDN are able to reconstruct the SCP even for low system load and therefore outperform both, the energy detection and the RRLS scheme.

4.5.4 Impact of Blocksize

Figure 4.8 takes a look on varying PRB-sizes. In 3GPP LTE the smallest scheduling unit in the frequency domain is one physical resource block (PRB) consisting of 12 subcarriers. Thus, we used 12 subcarriers as default blocksize for link level simulations. However, the effect of larger blocksize may be valuable for systems using larger PRBs. An increase in blocksize leads to an increase of the contiguous gray spaces. Thus, the estimate \hat{S}_E of the BS resource allocation in gray space detection step 1 will miss an increasing number of contiguous spectrum samples, so that the performance of the amplitude-based channel estimation methods in step 2 may suffer.

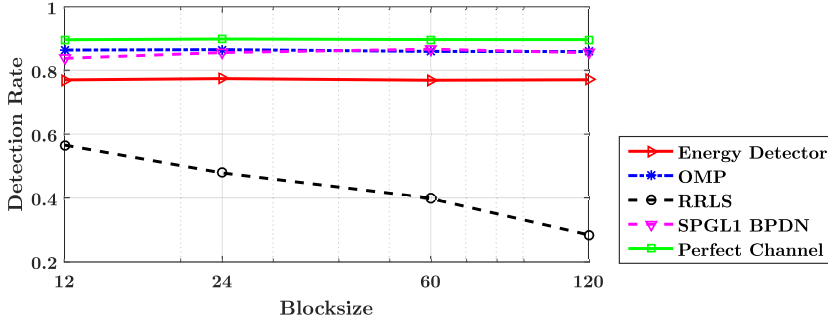


Figure 4.8: Impact of blocksize on gray space detection at fixed false alarm rate of 10% and 10 dB SNR.

Energy detection does not depend on the estimation process and therefore has a constant detection rate. However, also gray space detection based on the CS algorithms orthogonal matching pursuit (OMP) and basis pursuit denoising (BPDN) perform nearly constant. Thus, an increase of contiguous gray spaces hardly affects the decision accuracy of CS based gray space detection. Nevertheless, RRLS suffers dramatically from larger block sizes and goes even below 30% detection rate. Since RRLS estimates the CAC $\mathbf{v}_{[\mathbb{N}_{\text{CAC}}]}$ on the complete set \mathbb{N}_{CAC} , large contiguous gray spaces will severely degrade the estimation performance of RRLS.

4.5.5 Impact of Threshold Estimation

So far we selected an arbitrary constant hypothesis threshold for the log-likelihood ratio test, cf. (4.3.8) and (4.3.29), obtaining the uniformly most powerful test for a given gray space detection algorithm. However, if a specific false alarm probability is desired, the hypothesis threshold has to be calculated.

In Section 4.3.3, we derived our gray space test to obtain a desired false alarm rate. The gray space test (4.3.25) is based on composite log-likelihood ratio test and utilizes the approximated composite threshold $\rho_{c,\Lambda}(\mathbb{P}_{\text{fa}}, N = 10)$ from (4.3.22). The approximated composite threshold is obtained from $N = 10$ elements, which are sampled uniformly at random, without replacement, from \mathbb{P} for each channel realization. Thus, the approximated composite threshold is constant for one channel realization, but is recalculated for each channel realization. However, since we have $N = 10 < |\mathbb{P}|$, the threshold calculated in each channel realization is just an estimate of $\rho_{c,\Lambda}(\mathbb{P}_{\text{fa}}, |\mathbb{P}|)$. Therefore, $\widetilde{\mathbb{P}}_{\text{fa}} = \mathbb{P}(\rho_{c,\Lambda}(\mathbb{P}_{\text{fa}}, N), |\mathbb{P}|)$ in (4.3.23) is

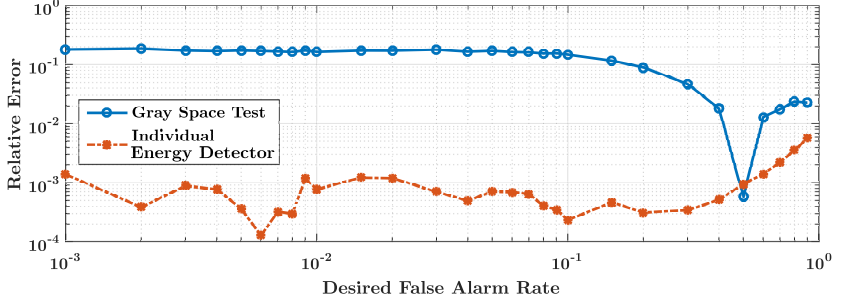


Figure 4.9: Relative error $\left| \widetilde{P}_{fa} - P_{fa} \right| / P_{fa}$ to achieve desired false alarm rate P_{fa} having perfect knowledge of the SCP with *individual energy detection* and *gray space test*.

also just an estimate of the desired false alarm rate P_{fa} .

For comparison we use the individual energy detection (4.3.28), where the individual simple threshold $\rho_{s,\hat{z}}(P_{fa})$ is calculated for each measurement $\hat{z}_{[f]}$ independently. Thus, the individual threshold changes for each channel realization. However, the individual threshold achieves the precise individual desired false alarm rate for each measurement. Therefore, independent of the actual channel statistics, the desired false alarm rate is precisely met.

Figure 4.9 shows the relative error $\left| \widetilde{P}_{fa} - P_{fa} \right| / P_{fa}$, where $\widetilde{P}_{fa} = F_{\chi^2_{2n}}(\rho; 2, \hat{\mathbf{v}}_{[f]}, \frac{1}{2}\sigma_n^2)$ is obtained by applying the threshold ρ on the distribution of \mathcal{H}_1 , cf. (4.3.6) and (4.3.20). The threshold of the corresponding approach is used, i.e. $\rho_{c,\Lambda}(P_{fa}, |\mathbb{P}|)$ for the gray space test and $\rho_{s,\hat{z}}(P_{fa})$ for the individual energy detection. To evaluate the impact of the gray space test only, perfect channel knowledge is assumed to calculate the corresponding thresholds. Selecting the threshold individually leads to low relative error of at least two order of magnitudes smaller than the desired false alarm rate P_{fa} , which results from inaccuracy in calculation of the probability density functions. However, the relative error of the gray space test stays mainly constant at approximately 0.2, resulting mainly from the estimation process. The notch at $P_{fa} = 0.5$ is due to a change of sign of $(\widetilde{P}_{fa} - P_{fa})$ from the lower to the higher false alarm rate. In summary, even though the individual threshold $\rho_{s,\hat{z}}(P_{fa})$ is more accurate, the approximated composite threshold $\rho_{c,\Lambda}(P_{fa})$ should be still feasible for practical implementations, where an relative error of approximately 0.2 will still leave \widetilde{P}_{fa} in the vicinity of P_{fa} .

In contrast to that, Figure 4.10 shows that the approximated compos-

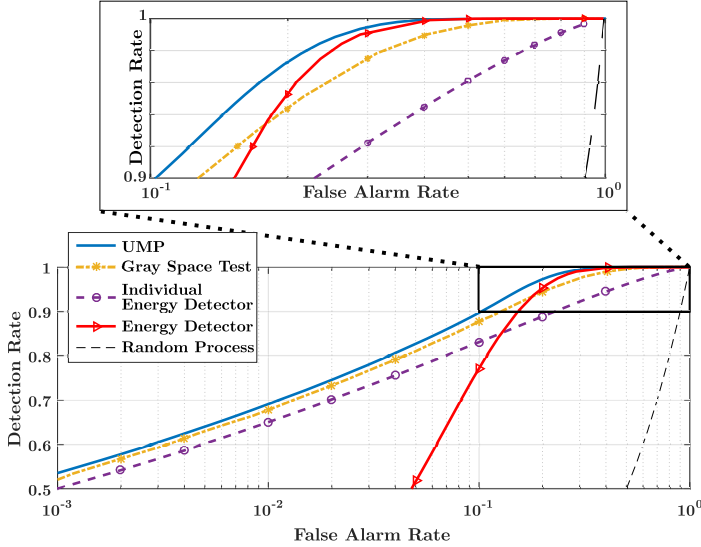


Figure 4.10: ROC of gray space test (4.3.25) in comparison with individual energy detection (4.3.28), uniformly most powerful (UMP) test (4.3.29), general energy detection (4.3.8) and random process.

ite threshold $\rho_{c,\Lambda}(P_{fa})$ of the gray space test achieves higher detection rate than $\rho_{s,\hat{z}}(P_{fa})$ of individual energy detection. Moreover, especially at low false alarm rates the approximated composite threshold is close to the performance of the uniformly most powerful test. Gray space test also outperforms classical energy detection in the false alarm region of interest below 10 %. In addition, individual energy detection also achieves better performance below 10 % towards classical energy detection, since classical energy detection assumes constant SCP. Therefore, at this region it is better to select $\rho_{s,\hat{z}}$ of energy detection individually than to fix the threshold to a constant. At large false alarm rates the impact of the frequency-varying wireless channel on classical energy detection decreases and the performance approaches quickly the detection rate of the uniformly most powerful test. However, as mentioned above, the operation point of gray space detection is below false alarm rate of 10%. Therefore, the approximated composite threshold of the Gray space test is the best choice for gray space detection, even if its accuracy regarding the desired false alarm rate is not as good as the individual threshold, see Figure 4.9.

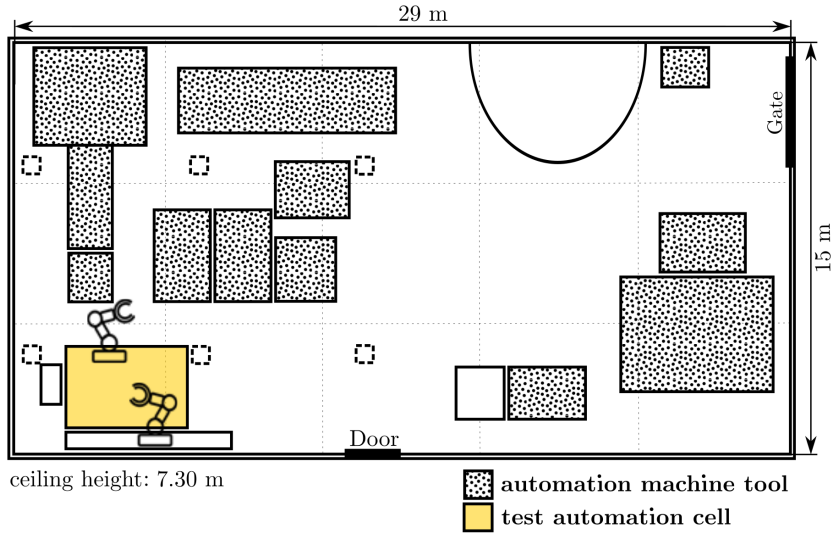


Figure 4.11: Map of the Smart Automation Lab at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University [HWR+16].

4.6 Practical Channel

The performance evaluations in Section 4.5 have been conducted with the 3GPP EPA performance channel model. In Section 2.3 we introduced also two other performance channel models, 3GPP EVA and ITU IOA. However, a performance evaluation on ground of practical wireless channels directly obtained from measurements gives a more realistic look on the detection scheme. Therefore, we use our measurement campaign [HWR+16; WHW16; DHC+19] to extract some representative practical performance channel models. The primary focus of the measurement campaign was the investigation of wear-and-tear free transmission in factory automation. Therefore, we picked the pick-and-place process as a representative manufacturing process.

4.6.1 Measurement Campaign

The measurement campaign took place in the Smart Automation Lab at the Laboratory for Machine Tools and Production Engineering at RWTH Aachen University, Germany, a highly sophisticated production environment. As shown in Figure 4.11 the dimension of the Smart Automation

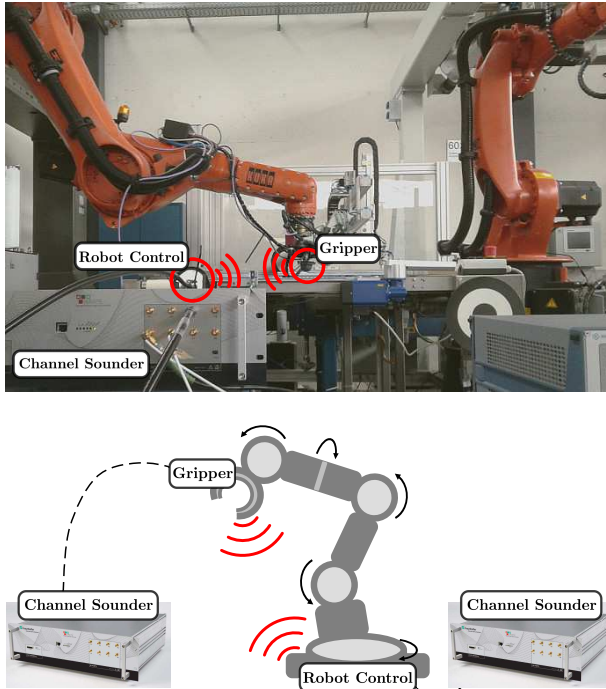


Figure 4.12: Measurement setup in the Smart Automation Lab at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University [HWR+16].

Lab is 29×15 m with a height of 7.3 m. In addition, the laboratory contains several automation machine tools, automated transportation systems and industrial robots as depicted in Figure 4.12. The building consists of a concrete floor and a metallic ceiling, where some open metallic joists are installed.

The channel measurements were conducted at the test automation cell in Figure 4.11 with a dimension of around 4×5 m. The automation machine tools nearby were actively used by workers during the measurements, thus representing a live production environment. Therefore, the wireless channel is time-variant due to the changing scattering and reflection characteristics. Nevertheless, [WHW16] shows that the impact on the short-range factory automation measurements is not significant.

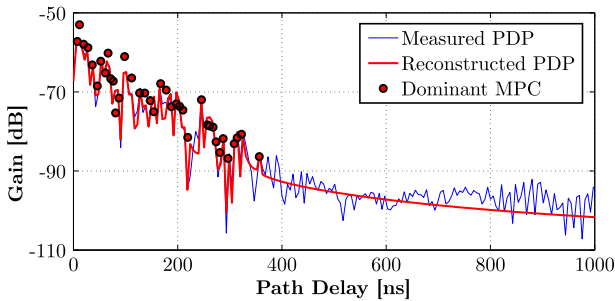
The two articulated industrial robots as depicted in Figure 4.12 were installed at the test automation cell, where the robot of interest was top-

mounted on an open metallic joist. The antennas were placed on the robot gripper and on the robot control, such that the wireless link bypasses the moving parts of the robot. Thus, a wear-and-tear free transmission is investigated for manufacturing processes, in particular a pick-and-place process. While the robot gripper is moving along programmed trajectories, the position of the robot control is fixed. The distance between both antennas was between 1.2 m and 2.7 m. The measurements were conducted at center frequencies 5.85 GHz and 2.25 GHz with 250 MHz bandwidth, because both frequency regions are targeted for the use of emerging use cases such as short range devices [CEP14]. However, only the 2.25 GHz center frequency is used to obtain the results below. During the measurement campaign no other interfering wireless system was active in the corresponding frequency ranges.

The measurement campaign was conducted with the High Performance Digital Radio Testbed (HIRATE) real-time channel sounder [KKP+13]. In particular, the HIRATE is a modular wireless radio testbed and consists of a digital signal processing platform, a wideband signal converter and reconfigurable radio frequency frontends. One HIRATE was used as the transmitter and one as the receiver. A clock cable connects both HIRATE units to enable frequency stability and time synchronization. Therefore, the frequency offset between both units is negligible. Furthermore, the absolute path delay between transmitter and receiver was recorded. Before the measurements, the channel sounder setup was calibrated to remove channel impairments induced by the measurement equipment itself. Therefore, a calibration setup of hard-wired transmitter and receiver antenna ports was used to obtain an ideal channel transfer function in the desired frequency regions. Thus, the Huber+Suhner antennas were considered as part of the wireless channel and were mostly removed in the measurement results. The used linear and vertical polarized Huber+Suhner multiband antennas are almost omni-directional with antenna gain of around 6 dBi at 2.25 GHz band and 8 dBi at 5.85 GHz band.

HIRATE was configured to support an effective signal bandwidth of $\Omega_{\text{BW}} = 250$ MHz. Thus, the corresponding sampling interval is 4 ns. The Frank-Zadoff-Chu sequence [Fra63] of length $n = 1024$ was implemented as sounding sequence, which was periodically repeated and correlated at the receiver. The time duration of the sequence was configured, such that $\tau_{\text{ED,max}} = n/\Omega_{\text{BW}} \ll \tau_{\text{coh}}$ is satisfied with τ_{coh} as coherence time of the channel. In particular, the maximum detectable channel excess delay was set to $\tau_{\text{ED,max}} = 4.096 \mu\text{s}$. One sounding sequence exploits the complete effective signal bandwidth of $\Omega_{\text{BW}} = 250$ MHz. In addition, the SNR of the measured channel impulse response is increased by 18 dB by averaging of 64

Scenario	Short-range factory automation (SFA)
Center frequency	2.25 GHz
Bandwidth	250 MHz
Sampling resolution	4 ns
Channel sounder	HIRATE
Antennas	Huber+Suhner model no. SWA 2459/360/7/20/V_2

Table 4.2: Measurement setup.**Figure 4.13:** Measured power delay profile (PDP) and reconstructed PDP based on the dominant multipath components (MPCs) of one measurement snapshot [HWR+16].

received sounding sequences. The resulting channel impulse response represents one measurement snapshot. An overview of the basic configuration parameters is given in Table 4.2.

The recorded complex channel impulse responses (CIRs) from the measurement snapshots are processed offline to obtain a set of dominant multipath components (MPCs) on a non-equidistant delay grid, as depicted in Figure 4.13. In this way, dominant MPCs for performance channel models can be retrieved. The dominant MPCs are obtained by a modified space-alternating generalized expectation-maximization (SAGE) algorithm [FTH+99], which is described in more detail in [KJT+15]. For better visualization, Figure 4.13 shows the real-valued power delay profile instead of the complex CIR. The modified SAGE algorithm iteratively determines the MPCs on a non-equidistant delay grid, such that the reconstructed CIR

calculated from the determined MPCs on the equidistant delay grid matches the measured CIR. Consequently, after determining all dominant MPCs, the measured and reconstructed CIR only differ by the noise floor.

4.6.2 Representative Channel Models

Among the reconstructed CIRs we extracted two representative short-range factory automation (SFA) channels with their corresponding MPC gains $|\mathbf{h}_{[n]}|^2$ and path delays $\tau_p(n)$. One representative SFA channel represents line-of-sight (LOS) characteristics and the other non-line-of-sight (NLOS) characteristics. To distinguish between LOS and NLOS we calculate the K-factor

$$\alpha_{KF} = \frac{(\max |\mathbf{h}|)^2}{\|\mathbf{h}\|_2^2 - (\max |\mathbf{h}|)^2}, \quad (4.6.1)$$

which is the ratio of the largest MPC gain to the power of the residual MPCs. A ratio above 0 dB indicates LOS and below 0 dB NLOS characteristics. Thus, we obtain two tapped delay line models for the performance channel models for short-range factory automation, one for LOS in Table 4.3 and one for NLOS in Table 4.4. For the performance channel models the path delay of the first path is set to zero and the following path delays are adjusted accordingly. In addition, the MPC gains are normalized, such that the largest MPC gain is 0 dB. We restrict the minimal MPC gains to -30 dB, having a range of up to 30 dB between the MPCs. The range of 30 dB is similar to ITU IOA, 3GPP EPA and 3GPP EVA performance channels discussed in Section 2.3.

The power delay profile (PDP) of the representative short-range factory automation (SFA) channel with LOS characteristics in Table 4.3 and with NLOS characteristics in Table 4.4 are visualized in Figure 4.14. At first glance, the LOS channel has fewer MPCs lower MPC gain than the NLOS channel. In addition, the excess delay of 188 ns for LOS is lower than for NLOS. The excess delay for given path delays is calculated by

$$\tau_{ED} = \tau_p(N) - \tau_p(1). \quad (4.6.2)$$

However, since the path delay of the first MPC is set to zero, the excess delay is equal to the path delay of the last MPC.

In addition, examples for the spectral channel power of the representative short-range factory automation (SFA) channel are visualized for 250 MHz bandwidth in Figure 4.15. Note, that the spectral channel power is symmetric, because the examples are generated by using channel impulse responses with real values only.

Path n	Path Delay $\tau_p(n)$ [ns]	MPC Gain $ h_{[n]} ^2$ [dB]	Path n	Path Delay $\tau_p(n)$ [ns]	MPC Gain $ h_{[n]} ^2$ [dB]
1	0	0	15	61.7	-27.6
2	1.7	-29.9	16	66.6	-20.3
3	4.5	-29.7	17	72.3	-18.2
4	4.8	-3.1	18	77.7	-21.6
5	8.5	-12.3	19	84.0	-23.1
6	12.0	-12.7	20	91.1	-26.5
7	15.8	-16.8	21	93.6	-22.9
8	19.3	-11.0	22	103.0	-26.9
9	27.7	-26.2	23	109.0	-26.1
10	31.8	-17.0	24	126.0	-28.9
11	35.9	-25.6	25	129.0	-24.4
12	42.5	-26.1	26	137.0	-29.4
13	51.8	-26.1	27	188.0	-29.6
14	56.8	-17.9			

Table 4.3: Tapped delay line model of short-range factory automation (SFA) with line-of-sight (LOS).

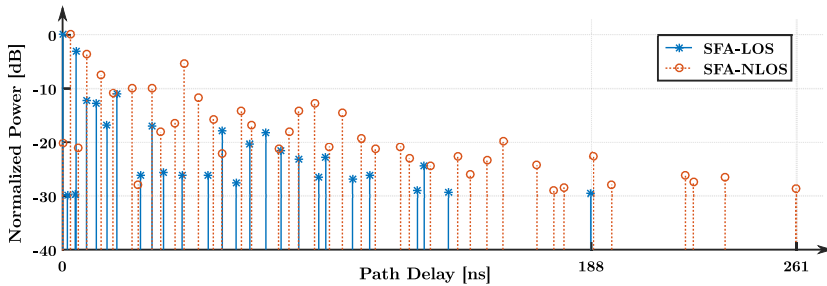


Figure 4.14: The power delay profile of channel models for representative short-range factory automation (SFA) with line-of-sight (LOS) and non-line-of-sight (NLOS) characteristics.

Path n	Path Delay $\tau_p(n)$ [ns]	MPC Gain $ \mathbf{h}_{[n]} ^2$ [dB]	Path n	Path Delay $\tau_p(n)$ [ns]	MPC Gain $ \mathbf{h}_{[n]} ^2$ [dB]
1	0	-20.1	22	94.9	-20.8
2	2.8	0	23	99.5	-14.6
3	5.6	-21.1	24	106.0	-19.3
4	8.6	-3.6	25	111.0	-21.2
5	13.7	-7.5	26	120.0	-20.8
6	17.9	-10.9	27	123.0	-23.0
7	24.7	-10.0	28	131.0	-24.4
8	26.9	-27.9	29	141.0	-22.7
9	31.8	-9.9	30	145.0	-26.0
10	34.9	-18.1	31	151.0	-23.3
11	40.0	-16.4	32	157.0	-19.9
12	43.2	-5.4	33	169.0	-24.3
13	48.3	-11.8	34	175.0	-28.9
14	53.6	-15.7	35	178.0	-28.4
15	56.6	-22.1	36	189.0	-22.7
16	63.6	-14.2	37	195.0	-27.9
17	67.2	-16.9	38	222.0	-26.1
18	76.8	-21.2	39	224.0	-27.4
19	80.5	-18.1	40	236.0	-26.6
20	83.9	-14.1	41	261.0	-28.6
21	89.6	-12.7			

Table 4.4: Tapped delay line model of short-range factory automation (SFA) with non-line-of-sight (NLOS).

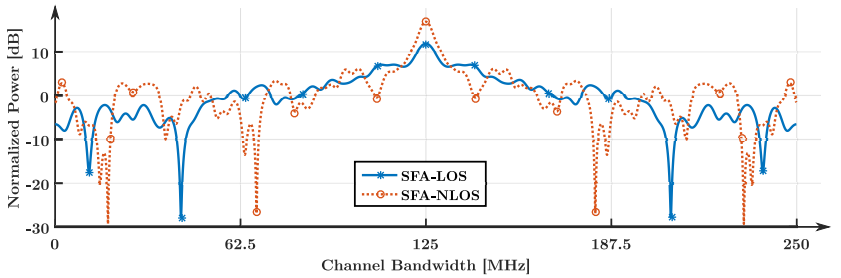


Figure 4.15: Examples for the spectral channel power of channel models for representative short-range factory automation (SFA) with line-of-sight (LOS) and non-line-of-sight (NLOS) characteristics.

Model	K-Factor α_{KF} [dB]	Number of Multipaths N	r.m.s. Delay Spread [ns] τ_{DS}	Excess Delay [ns] τ_{ED}
SFA-LOS	0.93	27	17	188
SFA-NLOS		41	33	261
ITU IOA		6	37	310
3GPP EPA		7	43	410
3GPP EVA		9	357	2510

Table 4.5: Parameters of short-range factory automation (SFA) as well as the ITU and 3GPP channel models.

To compare the representative SFA channel models with ITU and 3GPP, we calculate the root mean square (r.m.s.) delay spread

$$\tau_{DS} = \sqrt{\frac{\sum_{n=1}^N \left(|\mathbf{h}_{[n]}|^2 \cdot (\tau_p(n))^2 \right)}{\|\mathbf{h}\|_2^2} - \left(\frac{\sum_{n=1}^N \left(|\mathbf{h}_{[n]}|^2 \cdot \tau_p(n) \right)}{\|\mathbf{h}\|_2^2} \right)^2}. \quad (4.6.3)$$

An overview of the channel models of interest is given in Table 4.5. The representative SFA channel models are more comparable with the ITU IOA, since the r.m.s. delay spread and the excess delay are similar. However, the number of multipaths of SFA channels is significantly higher than the ITU and 3GPP channels.

4.6.3 Impact of Channel Model

To grasp the impact of different channel models, Figure 4.16 compares the algorithms under the selected channel models. The detection performance of all algorithms and even with perfect channel knowledge is decreasing over the channel models with larger excess delay and r.m.s. delay spread. Larger excess delay and r.m.s. delay spread lead to more frequency selective fading. Thus, more fading holes are present and decrease the gray space detection performance. Moreover, the spectral channel power estimation is influenced by the channel characteristics. Especially larger number of multipaths as well as larger excess delay decrease spectral channel power estimation accuracy of compressed sensing methods and RRLS estimator.

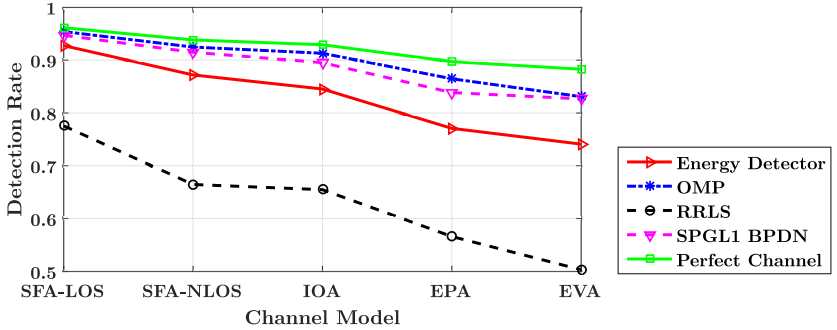


Figure 4.16: Behavior of gray space detection under different channel models at fixed false alarm rate of 10% and 10 dB SNR.

However, the impact on the compressed sensing methods is not as severe as for energy detection. The results indicate that gray space detection with OMP and BPDN are more robust against varying channel models.

4.7 Summary

In this Chapter, we proposed the *gray space detection* scheme for interweave cognitive radio system, which properly identifies empty resources within an active primary user band. Therefore, gray space detection copes with frequency fading holes and distinguishes them from inactive resources. In contrast to the classical energy detection, gray space detection can be configured to operate at desired false alarm rate. In other words, the primary user can be protected from the cognitive radio system with a desired probability. The gray space detection procedure comprises of three steps to acquire the unused resources: step 1 estimates a preliminary set of primary user resources via energy detection, step 2 estimates the spectral channel power via amplitude-based channel estimation on the given preliminary set of step 1, and step 3 estimates the final set of primary user resources via the log-likelihood ratio test, where the complement set represents the identified gray spaces. Different approaches for step 2, amplitude-based channel estimation, and step 3, log-likelihood ratio test, are applied.

In particular, amplitude-based channel estimation of gray space detection step 2 considers the conventional reduced-rank least squares (RRLS) method as well as the compressed sensing-based orthogonal matching pursuit (OMP) and basis pursuit denoising (BPDN) algorithm. The different gray space detection approaches were compared with energy detection un-

der several aspects: the receiver operating characteristic at different SNRs, the primary user system load, and physical resource blocksize. To summarize, gray space detection achieves in general better performance than energy detection. Moreover, gray space detection has the advantage that a desired primary user protection level can be achieved. For the different estimation algorithm of gray space detection, the greedy OMP approach of compressed sensing achieves the best performance results for a given false alarm rate. However, compressed sensing BPDN achieves mostly similar performance. The classical RRLS estimation is not well suited for gray space detection, since the performance suffers from the distribution of the primary user samples.

In addition, the gray space test is proposed for the log-likelihood ratio test of gray space detection step 3, where the utilized approximated composite threshold exploits the empirical cumulative distribution function of the wireless channel. We compared the gray space test with individual energy detection, where the threshold is individually applied for each measurement. There are generally two challenges to obtain a suitable threshold. First, to achieve a uniformly most powerful test, a threshold has to be constant under the considered channel model. Second, the threshold has to accurately resemble the desired performance criteria like false alarm rate. We showed that while the gray space test with the approximated composite threshold suffers slightly in accuracy in comparison to the individual energy detection, the false alarm rate performance of the gray space test is significantly closer to the uniformly most powerful test. The increased performance is achieved by stabilizing the approximated composite threshold to be less varying than the individual threshold. In conclusion, given that a relative accuracy of 10% is sufficient, the gray space test outperforms the simple log-likelihood ratio test.

Finally, we generated practical short-range factory automation channel models for a representative manufacturing process in factory automation. In addition to the ITU channel models, these short-range factory automation channel models were used to evaluate our gray space detection approach.

Chapter 5

Allocation Map Retrieval

5.1 Overview

This Chapter extends the system model of the previous Chapter 4 for multiple transmitters. Thus, in contrast to gray space detection of Chapter 4 we have more than two hypotheses. Furthermore, here we are interested in retrieving the distinct resource allocation map of each transmitter. We will propose an objective function to solve the combinatorial problem of allocation map retrieval for frequency-division multiple access signals and show that based on the objective function a receiver is able to identify the non-adjacent resources belonging to the same transmitter. Furthermore, we will propose a breadth-first search approach for decision trees, which solves the combinatorial problem of allocation map retrieval efficient but suboptimal.

5.1.1 Main Contributions

The main contribution of this Chapter comprises allocation map retrieval for frequency-division multiple access (FDMA) signals [WJW+15; WJB+15], where a receiver observes multiple transmitters exclusively allocating distributed frequency resources.

For this, we first extend the system model of Chapter 4 to multiple transmitters. In particular, in Chapter 4 we observed the power spectrum at the receiver consisting of a signal from a single transmitter and non-adjacent inactive resources. In this Chapter, we extend this system model to support multiple transmitters, where the superposition of these signals in form of an FDMA signal are observed by a receiver. Furthermore, we provide the system model for two distinct cases: (i) channel transfer function mea-

measurements observing pilots, and (ii) spectral channel power measurements observing constant amplitude signals. While in Chapter 4 we had only two hypotheses describing active and inactive resources, this Chapter is dealing with more than two hypotheses, each describing a different transmitter. In addition, in Chapter 4 we focused on the detection of gray spaces, where as a byproduct the resource allocation map of the single transmitter has been estimated. In contrast to Chapter 4, in this Chapter we are particularly interested in retrieving the distinct resource allocation map of each transmitter.

To understand our problem of allocation map retrieval, we put the problem into relation of the problem of unlabeled sensing studied by Unnikrishnan et al. [UHV15a]. There we show that a part of our allocation map retrieval problem, i.e. unlabeled selection sensing present in the single transmitter setting, is a subproblem of unlabeled sensing, and therefore, given sufficient i.i.d. measurements, a unique solution exists in the noiseless case with probability one. Furthermore, we derive an objective function based on sparsity for the allocation map retrieval problem. In particular, the objective function exploits the sparsity of the transmission channel to distinguish the signals arriving from different transmitters. To evaluate the feasibility of the objective function [WJW+16a], we provide combinatorial simulations for small problem sizes and show that the true resource allocation map generally corresponds to the global minimum of the objective function.

Finally, given the derived objective function, we develop a feasible but suboptimal breadth-first search algorithm for the allocation map retrieval problem, which traverses a decision tree and reconstructs the resources belonging to the same transmitter, i.e. retrieve the resource allocation maps. The breadth-first search algorithm is evaluated via simulations based on artificial sparse channels as well as on practical performance channels. There, the practical performance channels consist of the ITU performance channels as well as the SFA performance channels of our measurement campaign, see Section 4.6.2.

Based on the discussed framework of allocation map retrieval, a transmission concept has been filed as an international patent application [WJH17].

Summary of own contributions within this Chapter:

[WJW+15; WJB+15; WJW+16a; WJH17]

5.1.2 Structure

The outline of the chapter is as follows: We begin with the introduction of the system model for multiple transmitters in Section 5.2, where a receiver

observes the channel transfer functions in Section 5.2.1 and the spectral channel powers in Section 5.2.2 of the different transmitters. Furthermore, to ease evaluation of the two system models, we provide a unified system model in Section 5.2.3. Afterwards, to understand the complexity of the combinatorial problem of allocation map retrieval, we take a look on its problem size in Section 5.3. Then, in Section 5.4, we elaborate our allocation map retrieval problem in more detail. There, we introduce the problem of unlabeled sensing of Unnikrishnan et al. [UHV15a] (Section 5.4.1) and show that a part of our allocation map retrieval problem, i.e. unlabeled selection sensing, is a subproblem of unlabeled sensing. Furthermore, in Section 5.4.4 we discuss the extension of the unlabeled selection sensing problem to multiple transmitters, i.e. our original allocation map retrieval problem. Having a clearer view on the problem, we derive the objective function in Section 5.5 and give simulative verification in the subsequent Section 5.6. For transmission signals with some structures, like embedded physical resource blocks of certain blocksize, we develop a feasible but suboptimal algorithm based on orthogonal matching pursuit in Section 5.7, and provide simulations results in Section 5.8. The chapter concludes with the summary on allocation map retrieval in Section 5.9.

5.2 System Model

In this Section, we will extend our system model of Section 2.4 to multiple transmitters for the application of allocation map retrieval (AMR). For a single transmitter, the system model is similar to the system model introduced in Section 4.2. However, for multiple transmitters a frequency-division multiple access scenario is present, where each transmitter k allocates an exclusive resource allocation set (RAS) $\mathbb{S}_k \subseteq \mathbb{P}$ within the set \mathbb{P} of available resources in the transmission band, cf. Figure 5.1.

We will start by introducing the resource allocation framework in frequency-division multiple access (FDMA) scenarios, where we will define the set \mathbb{S} of RASs of all transmitters. Furthermore, we will introduce physical resource blocks (PRBs) to support resource allocation on PRB basis. Based on the RAS \mathbb{S}_k , we describe the FDMA signal $\hat{\mathbf{u}}_k$ generated by each transmitter $k \in \mathbb{K} := \{1, \dots, K\}$ and the superimposed signal $\hat{\mathbf{w}}$ at the receiver. At the end of this Section, we provide the system model for two distinct cases: (i) channel transfer function measurements observing pilots in Section 5.2.1, and (ii) spectral channel power measurements observing constant amplitude signals in Section 5.2.2. Furthermore, to ease evaluation of the two system models, we unify both measurement types in Section 5.2.3 to describe channel transfer function and SCP measurements in

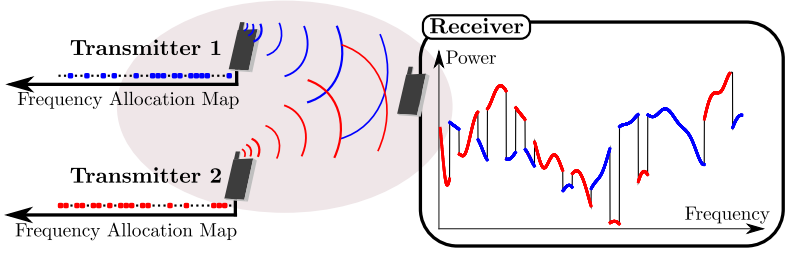


Figure 5.1: Transmitter 1 and 2 transmit on exclusive but fragmented frequency resources, while a receiver observes the superposition of both signals.

a single system model.

Wireless communication systems apply multiple access schemes to support multiple transmitters in a dedicated spectrum band, where the spectrum band is partitioned in a set \mathbb{P} of available resources. Here, in extension to our transmission model introduced in Section 2.4, multiple transmitters operate simultaneously within the same transmission band. However, each transmitter $k \in \mathbb{K} := \{1, \dots, K\}$ just allocates an exclusive RAS $\mathbb{S}_k \subseteq \mathbb{P}$ via a multiple access scheme, i.e.

$$\mathbb{S}_k \cap \mathbb{S}_{k'} = \emptyset, \quad \forall k, k' \text{ with } k \neq k'. \quad (5.2.1)$$

We focus on FDMA schemes like orthogonal FDMA (OFDMA) and deviates, which are e.g. applied for downlink, uplink and device-to-device (D2D) communication in 3GPP LTE [3GP20a] as well as NR [3GP20h]. Consequently, as depicted in Figure 5.1, the resource allocation within the transmission band is fragmented, such that each transmitter k utilizes a RAS $\mathbb{S}_k \subseteq \mathbb{P} \subseteq \mathbb{F} := \{0, \dots, N_{\text{DFT}} - 1\}$ distributed over the transmission band. Therefore, each transmitter k just allocates a subset of the available resources contained in the resource pool \mathbb{P} , where \mathbb{S}_k represents the RAS of transmitter k . The residual unallocated resources, not allocated to any transmitter, are given by

$$\mathbb{S}_0 := \mathbb{P} \setminus \left(\bigcup_{k \in \mathbb{K}} \mathbb{S}_k \right). \quad (5.2.2)$$

The set \mathbb{S}_0 can also be viewed as the RAS allocated by a single inactive transmitter. Therefore, we introduce the set

$$\mathbb{K}_0 := \{0\} \cup \mathbb{K} = \{0, \dots, K\}, \quad (5.2.3)$$

which extends \mathbb{K} by the set of unallocated resources. Thus, the union of the RASs $\mathbb{S}_k, k \in \mathbb{K}_0$ represent the set \mathbb{P} , i.e.

$$\bigcup_{k=0}^K \mathbb{S}_k = \mathbb{P}. \quad (5.2.4)$$

Consequently, the set of sets containing all RASs $\mathbb{S}_k, k \in \mathbb{K}_0$ are representing a [partition of the set](#) \mathbb{P} . Thus, we can describe the resource allocation of all transmitters by the set of RASs with

$$\mathbb{S} := \{\mathbb{S}_k\}_{k=0}^K = \{\mathbb{S}_0, \dots, \mathbb{S}_K\}, \quad (5.2.5)$$

where \mathbb{S} represents also a partition of the set \mathbb{P} ¹. Consequently, we have a bijective function from the resource set \mathbb{P} of the transmission band to the set \mathbb{S} of RASs, so that each subcarrier index $f \in \mathbb{P}$ is exclusively assigned to one RAS $\mathbb{S}_k, k \in \mathbb{K}_0$.

In typical communication systems, resources are not allocated on single subcarrier-basis, but as physical resource blocks (PRBs), see Section 2.4. Each PRB contains a predefined set $\mathbb{T}_m \subset \mathbb{P}$ of exclusive resources, where the PRB-size $N_B = |\mathbb{T}_m|, \forall m$ is fixed for all PRBs. Thus, the PRB-size N_B denotes the granularity, on which resources can be allocated by a transmitter. Hence, we introduce the set of all PRBs in the transmission band by²

$$\mathbb{T} := \{\mathbb{T}_0, \dots, \mathbb{T}_{N_{RB}-1}\}, \quad (5.2.6)$$

$$\text{where } \mathbb{T}_m \cap \mathbb{T}_{m'} = \emptyset, \quad \forall m, m' \text{ with } m \neq m',$$

where N_{RB} denotes the number of PRBs in \mathbb{P} and \mathbb{T} represents a set of sets. This definition also permits the use of PRBs containing resources distributed over the transmission band, of which we will take advantage in later sections. Now we can write the RAS \mathbb{S}_k as a function of \mathbb{T} with

$$\mathbb{S}_k := f(\mathbb{S}'_k, \mathbb{T}) = \bigcup_{m \in \mathbb{S}'_k} \mathbb{T}_m, \quad (5.2.7)$$

where $\mathbb{S}'_k \subset \{0, \dots, N_{RB} - 1\}$ represents the PRB allocation set of transmitter k .

¹A partition of a set can also be referred to as a Bell set, as discussed later in Section 5.3.

²In general, $|\mathbb{P}|$ is not an integer multiple of N_B , so that some subcarrier indices f are not assigned to any \mathbb{T}_m . However, if $|\mathbb{P}|$ is an integer multiple of N_B , then the set \mathbb{T} is a partition of the set \mathbb{P} and, moreover, a Stirling set of the second kind, as defined later in (5.3.3).

As an alternative to the set notation, we can also describe the RAS \mathbb{S}_k of transmitter k by a binary allocation matrix

$$\begin{aligned} \mathbf{S}_k &:= \text{diag}(\mathbf{s}_k), \\ \text{where } \mathbf{s}_k &:= \mathbf{1}_{\mathbb{S}_k}(f) \end{aligned} \quad (5.2.8)$$

defines the binary allocation vector $\mathbf{s}_k \in \{0, 1\}^{|\mathbb{P}|}$ via the [indicator](#) function. Moreover, the binary resource allocation map (RAM) \mathbf{S}_k can be concatenated to an overall RAM

$$\mathbf{S} := \begin{bmatrix} \mathbf{S}_0 & \dots & \mathbf{S}_K \end{bmatrix}, \quad (5.2.9)$$

where $\mathbf{S} \in \{0, 1\}^{|\mathbb{P}| \times (K+1) \cdot |\mathbb{P}|}$. Thus, the set \mathbb{S} of RASs containing the resource allocation information of all transmitters is represented by the single matrix \mathbf{S} .

Now we have defined the required variables, so that we can extend the transmit symbol calculation of (2.4.1) for multiple transmitters. Hence, *given the RAS \mathbb{S}_k and the time sample indices $t \in \{0, \dots, N_{\text{DFT}} - 1\}$, the discrete samples of an **OFDM time domain symbol at transmitter k** are obtained by*

$$(\mathbf{u}_k)_{[t]} = \frac{1}{\sqrt{N_{\text{DFT}}}} \sum_{f \in \mathbb{S}_k} \exp\left(i2\pi \frac{ft}{N_{\text{DFT}}}\right) \cdot (\hat{\mathbf{u}}_k)_{[f]}, \quad (5.2.10)$$

where each complex frequency sample $(\hat{\mathbf{u}}_k)_{[f]}$ is selected from a known transmit alphabet. Furthermore, each transmitter is located on a different spatial position. Thus, we can assume that the realizations of the transmission channels differ from each other with high probability

$$\forall_{k \neq m} \mathbf{h}_k \neq \mathbf{h}_m. \quad (5.2.11)$$

Consequently, the receiver in Figure 5.1 observes the superposition of multiple transmitted signals, each with a different channel \mathbf{h}_k . Besides the block-fading wireless channel, also the unknown time offset τ_{offset} of each transmitted signal is embedded in the channel \mathbf{h}_k , as described in (2.4.3). Therefore, time asynchronicity between the signals of the different transmitters is permitted as long as the maximal offset between the signals does not exceeds $(N_{\text{CP}} - \tau_{\text{ED, max}})$, i.e. the gap between CP length and the maximum excess delay $\tau_{\text{ED, max}}$ of the channel impulse response, see (2.4.4). Thus, we extend the block-fading transmission model of (2.4.6) and calculate the received time samples by

$$\mathbf{w}_{[t]} = \mathbf{n}_{[t]} + \sum_{k \in \mathbb{K}} \sum_{\tau=0}^{N_{\text{CP}}} (\mathbf{h}_k)_{[\tau]} \cdot (\mathbf{u}_k)_{[(t-\tau) \bmod N_{\text{DFT}}]} \quad (5.2.12)$$

with $t \in \{0, \dots, N_{\text{DFT}} - 1\}$, where $\mathbf{h}_k \in \mathbb{C}^{N_{\text{CP}}+1}$ is the block-fading channel of transmitter k and $\mathbf{n} \in \mathbb{C}^{N_{\text{DFT}}}$ represents the additive white Gaussian noise vector. The modulo operation (mod) on the transmit vector $\mathbf{u}_k \in \mathbb{C}^{N_{\text{DFT}}}$ from (5.2.10) extends the transmit vector by the CP.

The corresponding received frequency samples through an N_{DFT} -point DFT derived from (2.4.7) for multiple signals are calculated by

$$\hat{\mathbf{w}}_{[f]} = \hat{\mathbf{n}}_{[f]} + \sum_{k \in \mathbb{K}} \left(\hat{\mathbf{h}}_k \right)_{[f]} \cdot (\hat{\mathbf{u}}_k)_{[f]}. \quad (5.2.13)$$

Given the RAM \mathbf{S}_k , (5.2.13) can be written as matrix notation similar to (2.4.8) by

$$\hat{\mathbf{w}} = \hat{\mathbf{n}} + \sum_{k \in \mathbb{K}} \mathbf{S}_k \cdot \text{diag}(\hat{\mathbf{h}}_k) \cdot \hat{\mathbf{u}}_k. \quad (5.2.14)$$

Applying condition (5.2.1), permitting only exclusive RAMs \mathbf{S}_k on (5.2.14), the receiver observes merely the symbol $(\hat{\mathbf{u}}_k)_{[f]}$ of a single transmitter $k \in \mathbb{K}$ on subcarrier index $f \in \mathbb{S}_k$ altered by its channel and distorted by noise. All other transmitters are inactive on this subcarrier. In the case of an unallocated resource $f \in \mathbb{S}_0$, all transmitters remain inactive on the corresponding subcarrier, so that the resource contains only noise. Thus, given the actual RAS \mathbb{S}_k , the receiver observes

$$\hat{\mathbf{w}}_{[f]} = \begin{cases} \hat{\mathbf{n}}_{[f]}, & f \in \mathbb{S}_0 \\ \left(\hat{\mathbf{h}}_k \right)_{[f]} \cdot (\hat{\mathbf{u}}_k)_{[f]} + \hat{\mathbf{n}}_{[f]}, & f \in \mathbb{S}_k, k \in \mathbb{K}. \end{cases} \quad (5.2.15)$$

Thus, if the RASs \mathbb{S}_k are known at the receiver, then (5.2.15) results in the classical transmission model, where the unknown channel $\hat{\mathbf{h}}_k$ has to be estimated to achieve coherent wireless communication. However, here the receiver is ignorant about which resource belongs to which transmitter. Moreover, it is unknown to the receiver which set of non-adjacent resources belong to the same transmitter. Throughout this chapter we will mainly focus on the latter problem of allocation map retrieval, which will be discussed in more detail in Section 5.4. Solving the latter problem will simplify the former problem of identifying the corresponding transmitters, since having a known set of resources belonging to the same transmitter, a transmitter identification can be transmitted via these resources or may be retrieved via other techniques like wireless channel fingerprinting [YVS15].

In the following two sections, we consider two transmission alphabets: (i) a complex pilot alphabet to estimate CIRs as described in Section 3.6.1 and (ii) a complex alphabet with constant amplitude to estimate CACs as described in Section 3.6.2.

5.2.1 Channel Transfer Function Measurements

For the first transmit alphabet, we consider complex pilots. Thus, the pilot symbols $(\hat{\mathbf{u}}_k)_{[f]}$ are known at the receiver. For simplicity and without loss of generality the transmit alphabet is set to

$$(\hat{\mathbf{u}}_k)_{[f]} := \begin{cases} 0 & , \quad \text{else} \\ 1 & , \quad f \in \mathbb{S}_k. \end{cases} \quad (5.2.16)$$

Thus, if a pilot is transmitted on a subcarrier of transmitter k , then the corresponding transmit symbol is set to one, otherwise zero. Consequently, we directly observe one element of the channel transfer functions of a certain transmitter at the corresponding subcarrier. Since the resources are exclusively assigned to each transmitter, the received frequency domain samples in (5.2.15) using (3.6.10) and (5.2.16) simplify to

$$\hat{\mathbf{w}}_{[f]} = \begin{cases} \hat{\mathbf{n}}_{[f]} & , \quad f \in \mathbb{S}_0 \\ \mathbf{\Phi}_{[f,\cdot]} \cdot \mathbf{h}_k + \hat{\mathbf{n}}_{[f]} & , \quad f \in \mathbb{S}_k, k \in \mathbb{K}. \end{cases} \quad (5.2.17)$$

Thus, at each subcarrier index f the received frequency domain sample just contains a single channel transfer function measurement from one transmitter. Consequently, there exists no interference between different transmitters in the frequency domain.

The channel transfer function $\hat{\mathbf{h}}_k$ is obtained from the channel impulse response \mathbf{h}_k via the N_{DFT} -point discrete Fourier transform matrix $\mathbf{\Phi}$. To retrieve the CIR in CP based OFDM systems, the CIR has to be within the CP length N_{CP} , cf. (3.6.8). Thus, we can restrict $\mathbf{h}_k \in \mathbb{C}^{N_{\text{CP}}+1}$. Therefore, we have $\mathbf{\Phi} = \mathbf{F}_{[\mathbb{P}, \mathbb{N}_{\text{CIR}}]}$ being a N_{DFT} -point DFT submatrix as defined in (3.6.2), where $|\mathbb{N}_{\text{CIR}}| = N_{\text{CP}} + 1$, cf. (3.6.9). In matrix notation the channel transfer function is simply written

$$\hat{\mathbf{h}}_k = \mathbf{\Phi} \cdot \mathbf{h}_k \quad , \quad k \in \mathbb{K}. \quad (5.2.18)$$

An illustration of (5.2.17) without noise is given in Figure 5.2, where the received signal vector $\hat{\mathbf{w}} \in \mathbb{C}^{|\mathbb{P}|}$ emerges by selecting exclusive frequency components of each channel vector \mathbf{h}_k . In other words, there are no overlapping frequency components of different channel vectors. Furthermore, Figure 5.2 visualizes the discrete Fourier transform given in (5.2.18). Given $\mathbf{S}_k \in \{0, 1\}^{|\mathbb{P}| \times |\mathbb{P}|}$ and $\hat{\mathbf{n}} \in \mathbb{C}^{|\mathbb{P}|}$, the received frequency domain signal of (5.2.17) can be expressed in matrix notation as superposition of the system of linear equations (3.3.4) with

$$\hat{\mathbf{w}} = \hat{\mathbf{n}} + \sum_{k \in \mathbb{K}} \mathbf{S}_k \mathbf{\Phi} \mathbf{h}_k. \quad (5.2.19)$$

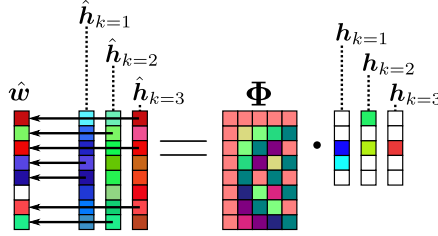


Figure 5.2: Each element of the received vector $\hat{\mathbf{w}}$ just contains the channel transfer function measurement $(\hat{\mathbf{h}}_k)_{[f]}$ of a single transmitter f . Thus, different transmitter signals do not interfere in the frequency domain.

5.2.2 Spectral Channel Power Measurements

The second transmit alphabet is a relaxed version of the first transmit alphabet, where constant amplitude symbols are permitted, yielding in a further degree of freedom at the transmitter. Thus, the phase of the transmit symbol can be arbitrarily chosen, cf. (3.6.21) of Section 3.6.2. Note that the amplitude of the same transmitter has to be constant for all frequency symbols in \mathbb{S}_k , but the amplitudes between different transmitters can vary. Furthermore, if the relative values of the amplitude symbols for each transmitter are known at the receiver, then the amplitudes can also differ within the same transmitter. However, to simplify the system model but without loss of generality, we require the transmission symbols to be of constant amplitude and scaled to

$$|(\hat{u}_k)_{[f]}| := \begin{cases} 0 & , \text{ else} \\ 1 & , f \in \mathbb{S}_k. \end{cases} \quad (5.2.20)$$

In contrast to the pilot transmission of the previous section, here the receiver observes the superposition of all transmitted signals in the power spectrum. Applying (3.6.22) and (3.6.25) on (5.2.20) we obtain the power spectrum measurements

$$\hat{z}_{[f]} = |\hat{\mathbf{w}}_{[f]}|^2 = \begin{cases} |\hat{\mathbf{n}}_{[f]}|^2 & , f \in \mathbb{S}_0 \\ \Phi_{[f], \cdot} \cdot \mathbf{v}_k + \hat{\mathbf{e}}_{[f]} & , f \in \mathbb{S}_k, k \in \mathbb{K}. \end{cases} \quad (5.2.21)$$

There, the adversarial error

$$\hat{\mathbf{e}}_{[f]} = |\hat{\mathbf{n}}_{[f]}|^2 + \hat{\mathbf{u}}_{[f]} \hat{\mathbf{h}}_{[f]} \hat{\mathbf{n}}_{[f]}^* + \hat{\mathbf{u}}_{[f]}^* \hat{\mathbf{h}}_{[f]}^* \hat{\mathbf{n}}_{[f]}$$

for $f \in \mathbb{S}_k$ is taken from (3.6.23), which is reduced to $\hat{\mathbf{e}}_{[f]} = |\hat{\mathbf{n}}_{[f]}|^2$ for frequencies $f \in \mathbb{S}_0$, where $|(\hat{\mathbf{u}}_k)_{[f]}| = 0$. Furthermore, as discussed in Section 5.2.1, the CIR in CP based OFDM systems is restricted to $\mathbf{h}_k \in \mathbb{C}^{N_{\text{CP}}+1}$. Thus, we have the circular autocorrelation of the CIR (CAC) $\mathbf{v}_k \in \mathbb{C}^{2N_{\text{CP}}+1}$ with $\text{supp}(\mathbf{v}) \subseteq \mathbb{N}_{\text{CAC}}$ and $|\mathbb{N}_{\text{CAC}}| = 2N_{\text{CP}} + 1$, cf. (3.6.29). Therefore, we have $\mathbf{\Phi} = \mathbf{F}_{[\mathbb{P}, \mathbb{N}_{\text{CAC}}]}$ being a N_{DFT} -point DFT submatrix as defined in (3.6.2). The DFT of the CAC is given by (3.6.26)

$$\hat{\mathbf{v}}_k = \mathbf{\Phi} \cdot \mathbf{v}_k, \quad k \in \mathbb{K}. \quad (5.2.22)$$

The observed power spectrum at the receiver at a certain subcarrier represents the noisy spectral channel power (SCP) of the wireless channel from the corresponding transmitter. Thus, Given $\mathbf{S}_k \in \{0, 1\}^{|\mathbb{P}| \times |\mathbb{P}|}$ and $\hat{\mathbf{e}} \in \mathbb{C}^{|\mathbb{P}|}$, the received frequency domain signal (5.2.21) can be expressed as superposition of the system of linear equations (3.3.4) with

$$\hat{\mathbf{z}} = \hat{\mathbf{e}} + \sum_{k \in \mathbb{K}} \mathbf{S}_k \mathbf{\Phi} \mathbf{v}_k. \quad (5.2.23)$$

5.2.3 General Measurements

In the previous two sections, Section 5.2.1 and Section 5.2.2, we provided the system model for channel transfer function and SCP measurements. To enable the evaluation for both transmission models at once, we will combine these two models to a more general system model.

We introduce a general system model based on the notation used for linear systems in Section 3.3 to represent both transmit alphabets, complex pilots (5.2.16) in Section 5.2.1 as well as constant amplitude symbols (5.2.20) in Section 5.2.2, and distinguish between them if required. Thus, given an error $\hat{\mathbf{e}} \in \mathbb{C}^{|\mathbb{P}|}$, measurement matrix $\mathbf{\Phi} \in \mathbb{C}^{|\mathbb{P}| \times M}$ and an unknown $\mathbf{x} \in \mathbb{C}^M$ representing $\mathbf{v} \in \mathbb{C}^{M=2N_{\text{CP}}+1}$ or $\mathbf{h} \in \mathbb{C}^{M=N_{\text{CP}}+1}$, the measurements of (5.2.17) and (5.2.21) are

$$\hat{\mathbf{y}}_{[f]} = \begin{cases} \hat{\mathbf{e}}_{[f]}, & f \in \mathbb{S}_0 \\ \mathbf{\Phi}_{[f, \cdot]} \cdot \mathbf{x}_k + \hat{\mathbf{e}}_{[f]}, & f \in \mathbb{S}_k, k \in \mathbb{K}. \end{cases} \quad (5.2.24)$$

Given $\mathbf{S}_k \in \{0, 1\}^{|\mathbb{P}| \times |\mathbb{P}|}$, in matrix notation of (5.2.19) and (5.2.23) we obtain

$$\hat{\mathbf{y}} = \hat{\mathbf{e}} + \sum_{k \in \mathbb{K}} \mathbf{S}_k \mathbf{\Phi} \mathbf{x}_k. \quad (5.2.25)$$

Thus, we have a received vector constructed as exemplary depicted in Figure 5.3. Since the resources are exclusively assigned, the received vector

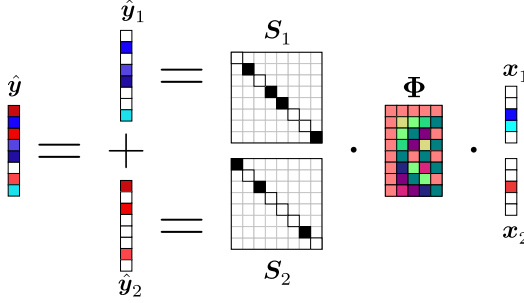


Figure 5.3: Example of transmission model in (5.2.25) with two transmitters, $k = 1$ and $k = 2$, with unknown exclusive RAMs S_k and unknown sparse vectors x_k . Allocation map retrieval shall retrieve the unknown RAMs S_k .

just contains a single transmitter signal at each element. Consequently, we have a received signal of overlapping transmitters in time domain, but not in frequency domain.

The receiver is ignorant about the resource allocation map (RAM) $S_k = \text{diag}(s_k)$, $s_k \in \{0, 1\}^{|\mathbb{P}|}$, cf. (5.2.8), and the sparse vector $x_k \in \mathbb{C}^M$ for all transmitters $k \in \mathbb{K}$. Note that the transmission model in (5.2.25) also permits the use of physical resource blocks (PRBs). In particular, while the set \mathbb{T} of PRBs is known at the receiver, the set S'_k of allocated PRBs of transmitter k is unknown. Therefore, we can obtain S_k from S'_k and \mathbb{T} via (5.2.7).

Given the overall RAM $S \in \{0, 1\}^{|\mathbb{P}| \times |\mathbb{P}|(|\mathbb{K}|+1)}$ defined in (5.2.9), the stacked vector $x \in \mathbb{C}^{M(|\mathbb{K}|+1)}$, and the block matrix $\Phi_B \in \mathbb{C}^{|\mathbb{P}|(|\mathbb{K}|+1) \times M(|\mathbb{K}|+1)}$, we can formulate the transmission model of (5.2.25) in mere matrix notation

$$\hat{y} = S\Phi_B x + \hat{e}, \quad (5.2.26)$$

$$\text{where } x := \begin{bmatrix} x_0^\top & \dots & x_{|\mathbb{K}|}^\top \end{bmatrix}^\top,$$

$$S := \begin{bmatrix} \text{diag}(s_0) & \dots & \text{diag}(s_{|\mathbb{K}|}) \end{bmatrix},$$

$$\Phi_B := I_{|\mathbb{K}|+1} \otimes \Phi.$$

An illustration of the model in (5.2.26) is given in Figure 5.4, where we observe $|\mathbb{K}| = 2$ transmitters and the set of unassigned resources. Note, that the allocation vector s_0 contains the unallocated resource, as defined

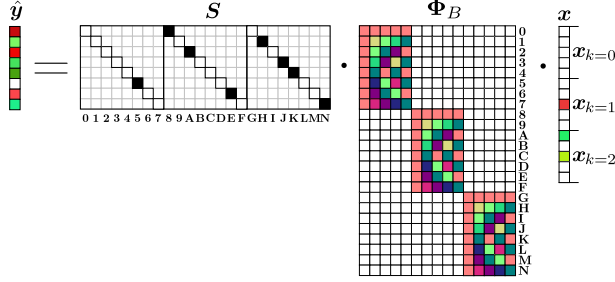


Figure 5.4: Example of transmission model in (5.2.26) with $|\mathbb{K}| = 2$ transmitters and the RAM S_0 containing the unassigned resources.

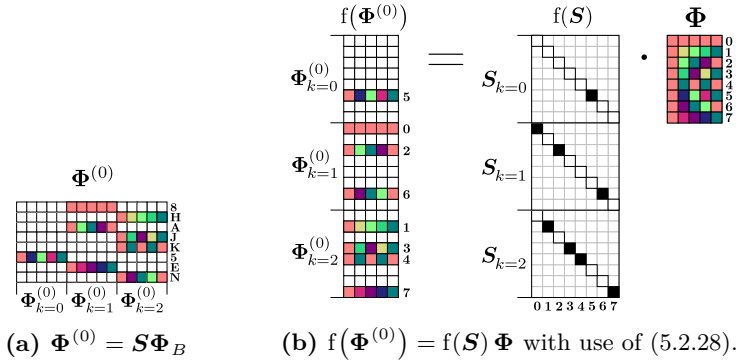


Figure 5.5: Illustration of linear transformation matrix $\Phi^{(0)}$ and $f(\Phi^{(0)})$ as well as its decomposition.

in (5.2.2). Thus, we can simply set the corresponding vector $\mathbf{x}_0 = \mathbf{0}$. Therefore, we have an overall RAM S with full rank. The unknown vector \mathbf{x} is just the stacked vector of all \mathbf{x}_k . And finally, the **Kronecker product** \otimes of the $(|\mathbb{K}| + 1) \times (|\mathbb{K}| + 1)^3$ identity matrix $I_{|\mathbb{K}|+1}$ with the measurement matrix $\Phi \in \mathbb{C}^{|\mathbb{P}| \times M}$ results in a block matrix Φ_B .

With slight abuse of notation, we define the actual transformation matrix by $\Phi^{(0)} = S \Phi_B$, see example in Figure 5.5a for $|\mathbb{K}| = 2$. Obviously, the size of the transformation matrix $\Phi^{(0)} \in \mathbb{C}^{|\mathbb{P}| \times M(|\mathbb{K}|+1)}$ is significantly smaller than its matrix factors S and Φ_B . Even though the block matrix Φ_B is constructed via the Kronecker product of $I_{|\mathbb{K}|+1}$ and Φ , it can become computational demanding due to its quickly growing size. To reduce the

³Matrix with $|\mathbb{K}| + 1$ rows and columns.

size of the block matrix Φ_B , we can reformulate (5.2.26) as

$$\hat{\mathbf{y}} = \mathbf{f}(\mathbf{f}(\mathbf{S}) \Phi) \mathbf{x} + \hat{\mathbf{e}}, \quad (5.2.27)$$

where $\mathbf{f}(\cdot)$ is a reshaping function

$$\mathbf{f}(\mathbf{A}) = \begin{bmatrix} (\mathbf{A}_0)^\top & \cdots & (\mathbf{A}_N)^\top \end{bmatrix}^\top. \quad (5.2.28)$$

Thus, we have $\mathbf{f}(\Phi^{(0)}) = \mathbf{f}(\mathbf{S}) \Phi$, see example in Figure 5.5b for $|\mathbb{K}| = 2$. Thereby, the complexity is similar to the transmission model in (5.2.25).

The receiver is now faced with the task of allocation map retrieval (AMR), where multiple hypotheses, each describing a different transmitter, have to be distinguished. Thus, since \mathbf{x}_k and \mathbf{S}_k are unknown for all transmitters k , AMR has to jointly estimate the stacked vector \mathbf{x} and the overall RAM \mathbf{S} . Before we elaborate our AMR problem in more detail in Section 5.4, we will take a brief excursion on combinatorial sets and the resulting problems in the subsequent Section 5.3, which will give some insights on the complexity of the AMR problem.

5.3 Combinatorial Complexity

Now we take a brief digression on some combinatorial sets and problems related to it. The focus of this Section will be on Bell sets and Stirling sets of the second kind, which are set of sets resulting from partitioning a set. These sets of sets are of particular importance in resource allocation, where exclusive resources are allocated by multiple transmitters from a single set of available resources, as given in (5.2.1) of Section 5.2. Therefore, these sets of sets are also of particular importance for AMR, so that we can assess the combinatorial problem size we are dealing with.

Since combinatorial problems evolve often in resource allocation, they are also relevant in the context of dynamic spectrum access, and in particular AMR. Many combinatorial problems are NP-hard [NW10] and thus, require exponential time to be solved. However, a variety of combinatorial problems can be solved suboptimal or in polynomial time considering additional information.

To grasp a feeling of the complexity of the considered problems in the subsequent Section 5.4, we give some numbers for different problem sizes. In particular, we want to assess the problem size of AMR, if the receiver has to traverse all possible resource allocations of (5.2.5), where a set \mathbb{B} of resource allocation sets of a set \mathbb{A} represents a partition of a set \mathbb{A} . Given a set \mathbb{A} containing e.g. single elements like indices/numbers, a set \mathbb{B} of sets

is the partition of a set \mathbb{A} consisting of a collection of subsets of \mathbb{A} . Thus, a [partition of a set](#) \mathbb{A} can be described by

$$\mathbb{B} := \{\mathbb{B}_0, \dots, \mathbb{B}_n, \dots, \mathbb{B}_{N-1}\}, \quad \mathbb{B}_n \neq \emptyset, \quad (5.3.1)$$

$$\text{where } \mathbb{B}_n \subseteq \mathbb{A}, \quad \bigcup_n \mathbb{B}_n = \mathbb{A}, \quad \mathbb{B}_n \cap \mathbb{B}_m = \emptyset, \quad \forall n \neq m.$$

The cardinality $|\mathbb{A}|$ of the set \mathbb{A} as well as the cardinality $|\mathbb{B}|$ of a set \mathbb{B} of sets denotes the number of elements of \mathbb{A} and the number of sets within \mathbb{B} , respectively. Our main interest lies in three different problem sizes, which are represented by (i) the [Bell number](#) N_{Bell} [Bru09, Chapter 8.2], (ii) the Stirling number of the second kind N_{Stirling} [Bru09, Chapter 8.2] and (iii) the number of all possible equal partition sets N_{eps} . The size of the three problems are related by $N_{\text{eps}} \leq N_{\text{Stirling}} \leq N_{\text{Bell}}$. In fact, the equal partition set captured by N_{eps} (iii) is a subset of the Stirling set of the second kind captured by N_{Stirling} (ii), which is again a subset of the partition of a set, also referred to as Bell set, captured by N_{Bell} (i). For comparison, we will also state the size of common combinatorial problems: the number of permutations of a set \mathbb{A} calculated by the factorial $|\mathbb{A}|!$ as well as the cardinality $|\mathfrak{P}(\mathbb{A})| = 2^{|\mathbb{A}|}$ of the power set $\mathfrak{P}(\mathbb{A})$, where e.g. the power set of $\mathfrak{P}(\{1, 2, 3\})$ is given by $\{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}, \{1, 3\}, \{2\}, \{2, 3\}, \{3\}\}$.

The Bell number N_{Bell} [Bru09, Chapter 8.2] captures the problem size of partitioning a set \mathbb{A} into indistinguishable, nonempty subsets as constructed in (5.3.1). Therefore, it describes the number of all possible sets of resource allocation sets \mathbb{S} in (5.2.5) for unknown number of transmitters $|\mathbb{K}|$. For more intuitive relation to the Bell number, we will also refer to a partition \mathbb{B} of a set \mathbb{A} as a Bell set \mathbb{B} of \mathbb{A} . For example, all bell sets of $\mathbb{A} = \{1, 2, 3\}$ are given by $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1\}, \{2, 3\}\}$, $\{\{2\}, \{1, 3\}\}$, $\{\{3\}, \{1, 2\}\}$, $\{\{1, 2, 3\}\}$. Note, that $|\mathbb{B}| \leq |\mathbb{A}|$ and therefore, the number of sets within the set of sets \mathbb{B} cannot exceed the number of elements within a set \mathbb{A} , since empty sets are not permitted in \mathbb{B} . The Bell number $N_{\text{Bell}}(M)$ [Bru09, Chapter 8.2] of \mathbb{A} with $|\mathbb{A}| = M$ is recursively calculated by

$$N_{\text{Bell}}(|\mathbb{A}|) = \sum_{n=0}^{|\mathbb{A}|-1} \binom{|\mathbb{A}|-1}{n} \cdot N_{\text{Bell}}(n) \quad (5.3.2)$$

$$= \sum_{n=0}^{|\mathbb{A}|-1} \frac{(|\mathbb{A}|-1)!}{n! \cdot (|\mathbb{A}|-1-n)!} \cdot N_{\text{Bell}}(n),$$

$$N_{\text{Bell}}(0) = 1.$$

The **Stirling number of the second kind**, also known as Stirling partition number N_{Stirling} [Bru09, Chapter 8.2], describes the number of possible partitions of a set \mathbb{A} containing only a fixed number of nonempty subsets. Thus, we introduce a Stirling set of the second kind \mathbb{B} of \mathbb{A} as a Bell set with fixed $|\mathbb{B}| = N$. A Stirling set of the second kind is described by

$$\begin{aligned} \mathbb{B} &:= \{\mathbb{B}_0, \dots, \mathbb{B}_n, \dots, \mathbb{B}_{N-1}\}, \quad \mathbb{B}_n \neq \emptyset, \quad |\mathbb{B}| = N \in \mathbb{N}, \\ \text{where } \mathbb{B}_n &\subseteq \mathbb{A}, \quad \bigcup_n \mathbb{B}_n = \mathbb{A}, \quad \mathbb{B}_n \cap \mathbb{B}_m = \emptyset, \quad \forall n \neq m, \end{aligned} \quad (5.3.3)$$

which extends the Bell set given in (5.3.1) by the additional constraint $|\mathbb{B}| = N \in \mathbb{N}$ to fix the cardinality of \mathbb{B} . For example, all Stirling sets of the second kind with $|\mathbb{B}| = 2$ of $\mathbb{A} = \{1, 2, 3\}$ are given by $\{\{1\}, \{2, 3\}\}$, $\{\{2\}, \{1, 3\}\}$, $\{\{3\}, \{1, 2\}\}$. Consequently, the set of all possible Stirling sets of the second kind in (5.3.3) is a subset of the set of all possible partitions of a set \mathbb{A} in (5.3.1). Given (5.3.3), the Stirling number of the second kind $N_{\text{Stirling}}(M, N)$ counts all possible Stirling sets of the second kind \mathbb{B} with $|\mathbb{B}| = N$ of \mathbb{A} with $|\mathbb{A}| = M$ and is calculated by

$$\begin{aligned} N_{\text{Stirling}}(|\mathbb{A}|, |\mathbb{B}|) &= \frac{1}{|\mathbb{B}|!} \cdot \sum_{n=0}^{|\mathbb{B}|} (-1)^{(|\mathbb{B}|-n)} \binom{|\mathbb{B}|}{n} \cdot n^{|\mathbb{A}|} \\ &= \begin{cases} |\mathbb{B}| \cdot N_{\text{Stirling}}(|\mathbb{A}| - 1, |\mathbb{B}|) \\ \quad + N_{\text{Stirling}}(|\mathbb{A}| - 1, |\mathbb{B}| - 1) \quad , \quad |\mathbb{A}| > |\mathbb{B}|, |\mathbb{B}| \neq 0 \\ 1 \quad , \quad |\mathbb{A}| = |\mathbb{B}| \\ 0 \quad , \quad \text{else.} \end{cases} \end{aligned} \quad (5.3.4)$$

Therefore, the Stirling number of the second kind describes the number of all possible sets of resource allocation sets \mathbb{S} in (5.2.5) for a given number of transmitters $|\mathbb{K}| = |\mathbb{B}|$. Furthermore, the summation of all Stirling numbers of the second kind results in the Bell number of set \mathbb{A}

$$N_{\text{Bell}}(|\mathbb{A}|) = \sum_{n=0}^{|\mathbb{A}|} N_{\text{Stirling}}(|\mathbb{A}|, n). \quad (5.3.5)$$

Therefore, a set captured by the Stirling number of the second kind is a subset of the set captured by the Bell number.

Furthermore, we introduce an equal partition set as a partition of a set \mathbb{A} consisting only of a set \mathbb{B} of sets with equal cardinality. Thus, an equal

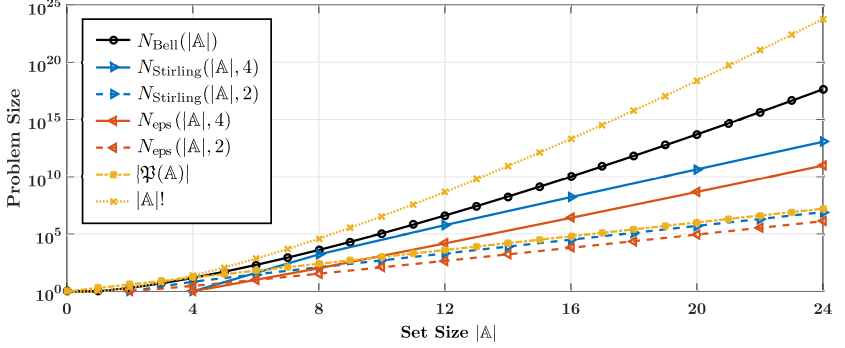


Figure 5.6: Increasing problem size over set size $|\mathbb{A}|$ for different combinatorial sets.

partition set is constructed as stated in (5.3.3), but additionally forces the subsets of \mathbb{B} to be of same cardinality

$$\mathbb{B} := \{\mathbb{B}_0, \dots, \mathbb{B}_n, \dots, \mathbb{B}_{N-1}\}, \quad \mathbb{B}_n \neq \emptyset, \quad (5.3.6)$$

$$\text{where } \mathbb{B}_n \subseteq \mathbb{A}, \quad |\mathbb{B}| = N \in \mathbb{N}, \quad |\mathbb{B}_n| = \frac{|\mathbb{A}|}{|\mathbb{B}|} \in \mathbb{N},$$

$$\bigcup_n \mathbb{B}_n = \mathbb{A}, \quad \mathbb{B}_n \cap \mathbb{B}_m = \emptyset, \quad \forall n \neq m.$$

For example, all equal partition sets with $|\mathbb{B}| = 2$ of $\mathbb{A} = \{1, 2, 3, 4\}$ are given by $\{\{1, 2\}, \{3, 4\}\}$, $\{\{1, 3\}, \{2, 4\}\}$, $\{\{1, 4\}, \{2, 3\}\}$. Thus, an equal partition set is only defined for sets \mathbb{A} and \mathbb{B} , where $|\mathbb{A}|$ is the integer multiple of $|\mathbb{B}|$. Consequently, the set of all possible equal partition sets in (5.3.6) is a subset of the set of all possible Stirling sets of the second kind in (5.3.3). The number $N_{\text{eps}}(M, N)$ of all possible equal partition sets \mathbb{B} with $|\mathbb{B}| = N$ of set \mathbb{A} with $|\mathbb{A}| = M$ is calculated by

$$\begin{aligned} N_{\text{eps}}(|\mathbb{A}|, |\mathbb{B}|) &= \frac{1}{|\mathbb{B}|!} \cdot \prod_{n=1}^{|\mathbb{B}|} \binom{\frac{n \cdot |\mathbb{A}|}{|\mathbb{B}|}}{\frac{|\mathbb{A}|}{|\mathbb{B}|}} \\ &= \frac{1}{|\mathbb{B}|!} \cdot \prod_{n=1}^{|\mathbb{B}|} \frac{\left(\frac{n \cdot |\mathbb{A}|}{|\mathbb{B}|}\right)!}{\left(\frac{|\mathbb{A}|}{|\mathbb{B}|}\right)! \cdot \left(\frac{(n-1) \cdot |\mathbb{A}|}{|\mathbb{B}|}\right)!}, \\ &\quad \text{where } \frac{|\mathbb{A}|}{|\mathbb{B}|} \in \mathbb{N}. \end{aligned} \quad (5.3.7)$$

$ \mathbb{A} $	4		8		16	
$ \mathbb{A} !$	24		40'320		$\approx 2 \cdot 10^{13}$	
$ \mathfrak{P}(\mathbb{A}) $	16		256		65'536	
$N_{\text{Bell}}(\mathbb{A})$	15		4'140		$\approx 10^{10}$	
$N_{\text{Stirling}}(\mathbb{A} , \mathbb{B})$	7	1	127	1'701	32'767	$\approx 2 \cdot 10^8$
$N_{\text{eps}}(\mathbb{A} , \mathbb{B})$	3	1	35	105	6'435	$\approx 3 \cdot 10^6$
$ \mathbb{B} $	2	4	2	4	2	4

Table 5.1: Problem size for selected cardinalities of $|\mathbb{A}|$ and $|\mathbb{B}|$.

For selected cardinalities of \mathbb{A} and \mathbb{B} (i) the Bell number N_{Bell} , (ii) the Stirling number of the second kind N_{Stirling} and (iii) the number of all possible equal partition sets N_{eps} are presented in Table 5.1. In addition, the cardinality of the power set $|\mathfrak{P}(\mathbb{A})|$ as well as the number of permutations $|\mathbb{A}|!$ are shown. Already at $|\mathbb{A}| = 16$ the Bell number indicates that it is impractical to solve a combinatorial problem on a Bell set. Furthermore, for $|\mathbb{A}| = 16$ and $|\mathbb{B}| = 4$ the Stirling number of the second kind as well as N_{eps} indicate an impractical combinatorial problem as well. However, for $|\mathbb{B}| = 2$ and for smaller sets \mathbb{A} a combinatorial problem may be still practical, at least for simulations. In addition to Table 5.1, we plotted the combinatorial problem size over set size $|\mathbb{A}|$ in Figure 5.6 for the considered combinatorial problems.

In summary, we have two relevant problem classes, represented by the Bell set and the Stirling set of the second kind. The Bell set addresses allocation problems, where the number of sets within the set of sets is unknown. The Bell set corresponds to problems, where the receiver does not know the number of active transmitters. If the number of active transmitters is known, then the allocation problem is described by the Stirling set of the second kind, where the number of sets within the set of sets is given.

5.4 Problem Analysis

In this Section we will elaborate our allocation map retrieval problem for our transmission model in (5.2.25) in more detail. First, we will formulate our allocation map retrieval problem, where the receiver is ignorant of the set \mathbb{S} of resource allocation sets (RASs) and the transmission channel \mathbf{h} .

Then, in Section 5.4.1 we will introduce the problem of unlabeled sensing formulated by Unnikrishnan et al. [UHV15a] and show that a part of our allocation map retrieval problem, i.e. unlabeled selection sensing with only a single transmitter, is a subproblem of the unlabeled sensing problem. In Section 5.4.2, we will discuss further related subproblems of unlabeled sensing and conclude in Section 5.4.3 with our subproblem of unlabeled selection sensing. Afterwards in Section 5.4.4, the extension of the unlabeled selection sensing problem to multiple transmitters is discussed, which represents our allocation map retrieval problem.

Given a fragmented spectrum band as depicted in Figure 5.1, where distinct RASs \mathbb{S}_k are associated to a vector \mathbf{x}_k of transmitter k , see our transmission model in (5.2.25), a spectrum observer or a receiver can carry out the following three different tasks:

1. identify unused fraction of spectrum,
2. identify the resource allocation map of each transmitter, or
3. perform channel estimation and obtain the transmitted data.

While the first task is addressed with classical spectrum sensing [YA09], the latter two tasks require more complex approaches. In particular, the first task wants to estimate \mathbb{S}_0 . Assuming frequency-selective channels, even the first task requires additional channel estimation, so that a fading hole is not mistaken as an unoccupied region as discussed in Chapter 4 as well as [WPJ13; WJW+16b]. Gray space detection in Chapter 4 jointly estimates \mathbb{S}_0 and $\mathbf{x}_k = \hat{\mathbf{z}}_k$ for $|\mathbb{K}| = 1$, cf. Section 4.3.

The second task is more demanding and, besides spectrum sensing capabilities, requires a further approach to distinguish the spectrum allocated by different transmitters. Edge-detection identifies the borders of an occupied spectrum band by exploiting edges within a flat channel spectrum [TG06; EO15]. Therefore, it goes a step beyond classical spectrum sensing, which is just indicating whether a selected spectrum band is occupied or not. However, edge-detection assumes a flat wireless channel, so that the SCP has constant amplitude. Furthermore, edge-detection only indicates the change of spectrum use like the switch between different transmitters. Consequently, edge-detection cannot identify non-adjacent fragmented spectrum belonging to the same transmitter. Therefore, other methods have to be applied to identify spectrum belonging to the same transmitter. In general, a joint estimation process has to be performed to recover the set \mathbb{S} of RASs in the second task, such that $\mathbb{S} := \{\mathbb{S}_0, \dots, \mathbb{S}_K\}$ and \mathbf{x}_k for all transmitters $k \in \mathbb{K}$ are jointly estimated.

The third task is partially addressed in the field of multi-user detection [ZG11; BSD13]. The activity and data of multiple devices are jointly detected. Most applications address sensor devices in the emerging market of Internet-of-Things. Nevertheless, current multi-user detection schemes are not applied in fragmented spectrum and therefore does not deal with task two. In general, to recover the transmitted data, joint estimation of $\mathbb{S} := \{\mathbb{S}_0, \dots, \mathbb{S}_K\}$ and \mathbf{x}_k for all transmitters $k \in \mathbb{K}$ like in task two has to be performed.

Consequently, we focus here on the underlying joint estimation problem invoked by task two and three, where we want to jointly recover the RAS \mathbb{S}_k and the sparse unknown \mathbf{x}_k of all transmitters $k \in \mathbb{K}$ in (5.2.25). The allocation map retrieval (AMR) problem can be written

$$\begin{aligned} & \text{find all } \mathbf{S}_k \in \{0, 1\}^{|\mathbb{P}| \times |\mathbb{P}|}, \mathbf{x}_k \in \Sigma_{s_k}, \\ & \text{s.t. } \left\| \hat{\mathbf{y}} - \sum_{k=1}^{|\mathbb{K}|} \mathbf{S}_k \Phi \mathbf{x}_k \right\|_2^2 \leq \rho, \sum_{k=0}^{|\mathbb{K}|} \mathbf{S}_k = \mathbf{I}, \end{aligned} \quad (5.4.1)$$

where Σ_{s_k} represents the union of s_k -dimensional subspaces as defined in (3.2.2).

In the next subsections we will take a closer look on the AMR problem. Therefore, we will at first consider (5.4.1) with $|\mathbb{K}| = 1$, which we will refer to as unlabeled selection sensing. There we will present the strongly related problem of unlabeled sensing by Unnikrishnan et al. [UHV15a] and elaborate the similarities between unlabeled selection sensing and unlabeled sensing. Furthermore, we will show that the findings on unlabeled sensing also hold for unlabeled selection sensing. Afterwards, in Section 5.4.4 we will consider (5.4.1) with $|\mathbb{K}| > 1$, i.e. our original AMR problem.

5.4.1 Unlabeled Sensing

Subsequently, we will introduce unlabeled sensing depicted in Figure 5.7 and its subproblems as well as related problems. In this Section, we slightly abuse the notation for $\hat{\mathbf{y}}$ and $\hat{\mathbf{e}}$ by describing not only the Fourier transform, but also other transforms depending on the used measurement matrix Φ . The transformation applied will be obvious from the context.

The aim of this Section is twofold: (i) we want to provide an overview of related problems to our unlabeled selection sensing problem representing the special case of our AMR problem with $|\mathbb{K}| = 1$, and (ii) want to show that the findings on unlabeled sensing in [UHV15a] also hold for unlabeled selection sensing.

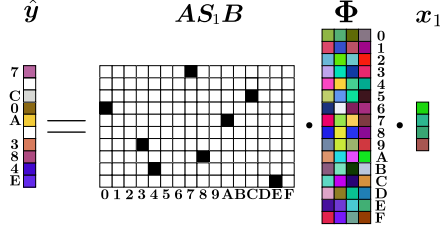


Figure 5.7: Unlabeled sensing with random linear measurements $\hat{\mathbf{y}}$, where the composite matrix $\mathbf{AS}_1\mathbf{B}$ and the vector \mathbf{x}_1 are unknown [UHV15a].

Before we go into more details on unlabeled sensing, we shortly detour to the related approach of [dictionary learning](#)

$$\hat{\mathbf{y}} = \Phi' \mathbf{x}_1 + \hat{\mathbf{e}}, \quad (5.4.2)$$

where an observation matrix $\Phi' \in \mathbb{C}^{N \times M}$ represents an unknown overcomplete dictionary, such that $N = \text{rank}(\Phi') < M$. Furthermore, $\mathbf{x}_1 \in \mathbb{C}^M$ is an unknown sparse vector and the measurement $\hat{\mathbf{y}}$ is disturbed by noise $\hat{\mathbf{e}}$. Dictionary learning now aims to find a fixed Φ' , so that as many measurements $\hat{\mathbf{y}}$ as possible have a sparse representation \mathbf{x}_1 . Various approaches like sparse Bayesian learning [WR04] can be applied to solve dictionary learning problems. However, note that in general solving the dictionary learning problem is NP-hard [Til15]. Furthermore, given just a single measurement, dictionary learning is not feasible in general.

In contrast to dictionary learning, unlabeled sensing is generally feasible for single measurements $\hat{\mathbf{y}}$, because unlabeled sensing just recovers a sparse linear transformation $\mathbf{AS}_1\mathbf{B}$ of a known matrix Φ as depicted in Figure 5.7. In addition, unlabeled sensing is generally also feasible for dense unknown vectors \mathbf{x}_1 . Thus, unlabeled sensing jointly recovers a partially known observation matrix $\Phi' = \mathbf{AS}_1\mathbf{B}\Phi$ and \mathbf{x}_1 from a measurement

$$\begin{aligned} \hat{\mathbf{y}} &= \Phi' \mathbf{x}_1 + \hat{\mathbf{e}} \\ &= \mathbf{AS}_1\mathbf{B}\Phi \mathbf{x}_1 + \hat{\mathbf{e}}, \end{aligned} \quad (5.4.3)$$

where the measurement matrix Φ is known, but not the sparse linear transformation $\mathbf{AS}_1\mathbf{B}$. Hence, the observation matrix $\Phi' \in \mathbb{C}^{N \times M}$ can be decomposed in an unknown composite matrix $(\mathbf{AS}_1\mathbf{B}) \in \{0, 1\}^{N \times L}$ and a known measurement matrix $\Phi \in \mathbb{C}^{L \times M}$. One example of unlabeled sensing without the adversarial error is visualized in Figure 5.7, where certain rows of the measurement matrix Φ are selected by the rectangular composite matrix $\mathbf{AS}_1\mathbf{B}$.

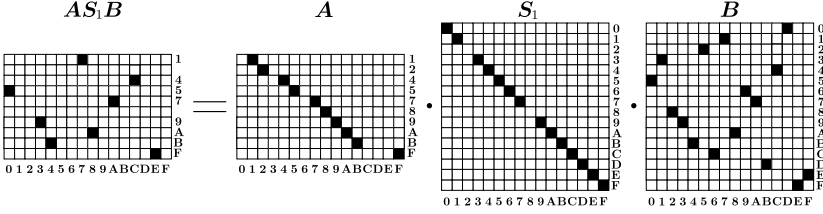


Figure 5.8: Construction of the composite matrix AS_1B from puncturing matrix A , selection matrix S_1 and permutation matrix B .

The unknown binary composite matrix AS_1B can be further decomposed into three components: a rectangular puncturing matrix $A \in \{0, 1\}^{N \times L}$, a square selection matrix $S_1 \in \{0, 1\}^{L \times L}$, and a permutation matrix $B \in \{0, 1\}^{L \times L}$. The composite matrix AS_1B as well as each matrix for itself have only binary elements, and therefore the elements are not representing a convex set. Thus, in general, there exists no efficient straightforward algorithm to recover the binary matrices. Note that Unnikrishnan et al. [UHV15a] defined unlabeled sensing only by $\hat{y} = AB\Phi x_1 + \hat{e}$, where Φ consists of random i.i.d. elements.

An illustration of the construction of the composite matrix AS_1B is given in Figure 5.8. While a puncturing matrix A can be constructed from the identity matrix I by omitting certain rows, the selection matrix has the form of our binary resource allocation map defined in (5.2.8), constructed by $S_1 = \text{diag}(s_1)$ with $s_1 \in \{0, 1\}^L$. Consequently, the puncturing matrix removes certain elements and the selection matrix sets certain elements to zero. While the number of non-zero entries of the puncturing matrix equals the number of rows of A , the number of rows of S_1 just gives an upper limit for the number of non-zero entries of the selection matrix. In contrast to both, A and S_1 , the binary permutation matrix B is only restricted by containing just a single non-zero entry in each row and column. Hence, just the order of the elements is lost, but not the contained information.

Unlabeled sensing comprises the problem of recovering an unknown vector x_1 and all three linear transformations as a single composite matrix AS_1B . There, A cuts out some rows of Φ and S_1 sets some rows of Φ to zero. Thus, not all rows of Φ are used to generate the measurement \hat{y} from the unknown x_1 . Consequently, each row and each column of the rectangular composite matrix AS_1B contain only zeros or just a single 1. For illustration of the effect of unlabeled sensing on measurement matrix Φ , Figure 5.9 visualizes puncturing, selection and permutation on a discrete Fourier transform matrix.

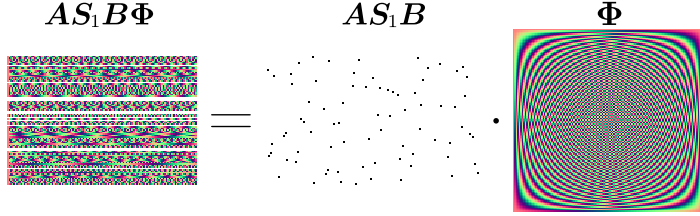


Figure 5.9: Unlabeled sensing with random Fourier measurements.

In [UHV15a; UHV15b] Unnikrishnan et al. analyzed the unlabeled sensing problem of recovering \mathbf{x}_1 from a measurement

$$\hat{\mathbf{y}}' = \mathbf{A} \cdot \mathbf{B} \cdot \Phi \cdot \mathbf{x}_1 + \hat{\mathbf{e}}, \quad (5.4.4)$$

where the composite matrix \mathbf{AB} is unknown. The model of Unnikrishnan et al. in (5.4.4) is similar to the formulated model in (5.4.3), but with $\mathbf{S}_1 = \mathbf{I}$. In principle, Unnikrishnan et al. [UHV15a, Theorem II.1] investigated the recovery of a vector \mathbf{x}_1 from random linear observation $\hat{\mathbf{y}}'$ and derived conditions, when successful recovery is possible for the noiseless case $\hat{\mathbf{e}} = \mathbf{0}$. In particular, given the real-valued case and a measurement matrix $\Phi \in \mathbb{R}^{L \times M}$ with i.i.d. random variables, then for measurement $\hat{\mathbf{y}}' \in \mathbb{R}^{N'}$ with $N' \geq 2M$ and probability one, a unique $\mathbf{x}_1 \in \mathbb{R}^M$ exists solving (5.4.4).

The statement of Unnikrishnan et al. for the noiseless case $\hat{\mathbf{e}} = \mathbf{0}$ in [UHV15a, Theorem II.1] can be extended to hold for (5.4.3) as well, which we will proof in the following. Given [UHV15a, Theorem II.1] with $\hat{\mathbf{y}}' = \mathbf{AB}\Phi\mathbf{x}_1$ and a predefined constant c , the probability of having a measurement with precisely $\hat{\mathbf{y}}'_{[f]} = c$ is zero. Thus, having a measurement with precisely $\hat{\mathbf{y}}'_{[f]} = 0$ is zero as well. Consequently, given now the noiseless model $\hat{\mathbf{y}} = \mathbf{AS}_1\mathbf{B}\Phi\mathbf{x}_1$ in (5.4.3), we have $\hat{\mathbf{y}}_{[f]} = 0$ if and only if the respective row of \mathbf{S}_1 is zero. Note, that rows of \mathbf{S}_1 are punctured by \mathbf{A} , whereby also rows with zeros may be removed. Therefore, given the joined matrix $\mathbf{A}' = \mathbf{AS}_1$, we can write

$$\hat{\mathbf{y}}'_{[\text{supp}(\hat{\mathbf{y}})]} = \mathbf{A}'_{[\text{supp}(\hat{\mathbf{y}}), \cdot]} \cdot \mathbf{B} \cdot \Phi \cdot \mathbf{x}_1, \quad (5.4.5)$$

with $\mathbf{A}'_{[\text{supp}(\hat{\mathbf{y}}), \cdot]}$ being a puncturing matrix, where the zero rows of \mathbf{A}' are removed. Thus, since (5.4.5) is equal to the formulation of Unnikrishnan et al. in (5.4.4), we proved recovery also for using a selection matrix in the noiseless setting. In particular, for measurement $\hat{\mathbf{y}} \in \mathbb{R}^N$ with $|\text{supp}(\hat{\mathbf{y}})| \geq 2M$ and probability one, a unique $\mathbf{x}_1 \in \mathbb{R}^M$ exists solving (5.4.3) with $\hat{\mathbf{e}} = \mathbf{0}$.

In addition, Unnikrishnan et al. [UHV15a, Proposition VI.1] analyzes the stability of the solution in the noisy case. However, the paper presents no

practical algorithm to solve the unlabeled sensing problem in (5.4.3).

5.4.2 Related Subproblems

As indicated by the decomposition of matrix $\mathbf{A}\mathbf{S}_1\mathbf{B}$ into three matrices in (5.4.3) and Figure 5.8, unlabeled sensing can be subdivided in three subproblems. We refer to these subproblems as:

1. unlabeled ordered sensing with \mathbf{A} ,
2. unlabeled selection sensing with \mathbf{S}_1 , and
3. unlabeled permutation sensing with \mathbf{B} .

In general, a structured matrix and a dense unknown vector \mathbf{x}_1 shall be jointly recovered in all three subproblems. Various publications are addressing one of these subproblems, but usually induce further structure on the unknown vector. Subsequently, we will discuss some recent results.

An application of unlabeled permutation sensing (subproblem 3) in a sparse setting has been already described in 2014 by Emiya et al. [EBD+14], where unknown sensor permutations and sparse \mathbf{x}_1 are jointly recovered

$$\hat{\mathbf{y}} = \mathbf{B}\Phi\mathbf{x}_1 + \hat{\mathbf{e}}. \quad (5.4.6)$$

The entries of the measurement matrix Φ are i.i.d. drawn from the normal distribution. Since we have a permutation problem, the inverse of the square permutation matrix \mathbf{B} exists, which is exploited by Emiya et al. [EBD+14]. Nevertheless, [Til15] proves that the pure sensor permutation problem is still computational intractable.

Besides a relaxed condition on \mathbf{x}_1 permitting dense vectors, the unlabeled permutation sensing problem of [PWC16] is the same as in [EBD+14]. In [PWC16], Pananjady et al. focused merely on the recovery of the permutation matrix from noisy measurements and provides SNR conditions for exact permutation recovery. Moreover, they analyze approximate permutation recovery, where the permutation matrix is recovered with some distortions. The distortion is measured by the [Hamming distance](#) between the approximation and the correct permutation. The approximate permutation recovery is linked to the SNR conditions by obtaining the probability of certain distortion. Publication [PWC17] extends [PWC16] for multivariate linear regression with shuffled data and noise.

Recently, various algorithms solving unlabeled permutation sensing were published. An algorithm based on lattice basis reduction has been proposed by Hsu et al. [HSS17] and Abid et al. [APZ17], presenting an approach based on least squares and [moments](#). Furthermore, Wang et al. [WZB+17]

analyzed unlabeled permutation sensing for binary quantized observations $\hat{\mathbf{y}}$ and proposed an algorithm to recover the sparse vector \mathbf{x}_1 . Most recently, Prasad published an approach to jointly recover the permutation matrix and a sparse vector [Pra18]. There, Prasad uses dictionary learning via sparse Bayesian learning [WR04].

To model jitter [Bal62], Haghighatshoar et al. study the unlabeled ordered sensing problem (subproblem 1) [HC18]

$$\hat{\mathbf{y}} = \mathbf{A}\Phi\mathbf{x}_1 + \hat{\mathbf{e}}. \quad (5.4.7)$$

Instead of a permutation matrix \mathbf{B} , unlabeled ordered sensing applies a puncturing matrix $\mathbf{A} \in \{0, 1\}^{N \times L}$. In general, a puncturing matrix is more structured than a permutation matrix, cf. Figure 5.8. However, in contrast to a permutation matrix the inverse of a puncturing matrix does not exist in general. Nevertheless, a puncturing matrix preserves the order of the measurement matrix Φ and has full rank with $\text{rank}(\mathbf{A}) = \min[N, L]$.

In [HC18], Haghighatshoar et al. present the duality between compressed sensing and unlabeled ordered sensing. Furthermore, they derive an alternating minimization algorithm to solve the problem of unlabeled ordered sensing. Although AMR (5.2.25) operates with a selection matrix \mathbf{S}_1 , the puncturing matrix \mathbf{A} of unlabeled ordered sensing and the selection matrix have some similarities. Both matrices, \mathbf{A} and \mathbf{S}_1 , preserve the order of the measurement matrix Φ . The difference is merely that the puncturing matrix \mathbf{A} removes certain rows of Φ , and the selection matrix \mathbf{S}_1 sets certain rows of Φ to zero.

5.4.3 Unlabeled Selection Sensing

Our AMR problem for $|\mathbb{K}| = 1$ represents unlabeled selection sensing (subproblem 2), but with \mathbf{x}_1 being sparse

$$\hat{\mathbf{y}} = \mathbf{S}_1\Phi\mathbf{x}_1 + \hat{\mathbf{e}}. \quad (5.4.8)$$

The problem of recovering the selection matrix \mathbf{S}_1 and sparse \mathbf{x}_1 for $|\mathbb{K}| = 1$ is already addressed by gray space detection in Chapter 4. Note, that in contrast to subproblems 1 and 3 with puncturing and permutation matrix, the selection matrix is rank deficient because some rows are zero. Nevertheless, the problem (5.4.8) can be reformulated, such that an unknown selection matrix \mathbf{S}'_1 has full rank. Therefore, we apply the Kronecker product \otimes on

the selection matrix and the measurement matrix, obtaining

$$\hat{\mathbf{y}} = \mathbf{S}'_1 \Phi' \mathbf{x}_1 + \hat{\mathbf{e}}, \quad (5.4.9)$$

$$\text{where } \mathbf{S}'_1 = \left(\mathbf{S}_1 \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} \right) + \left((\mathbf{I} - \mathbf{S}_1) \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} \right),$$

$$\text{and } \Phi' = \Phi \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}^\top.$$

Consequently, the modified unknown selection matrix $\mathbf{S}'_1 \in \mathbb{C}^{L \times 2L}$ is now a full rank rectangular matrix. The known measurement matrix Φ is replaced by a rank deficient but still known modified measurement matrix $\Phi' \in \mathbb{C}^{2L \times M}$. Therefore, the unlabeled selection sensing problem in (5.4.9) may be solved similar as in [HC18] via an alternating minimization algorithm.

In summary, our problem of interest represents unlabeled selection sensing, but with sparse \mathbf{x}_1 . Since unlabeled selection sensing is a subproblem of unlabeled sensing, joint recovery is also possible if no structure is induced on \mathbf{x}_1 . However, as we will discuss in Section 5.5, the sparse structure of \mathbf{x} is mandatory for the complete AMR problem with $|\mathbb{K}| > 1$. Furthermore, even for $|\mathbb{K}| = 1$, our AMR problem in (5.2.25) differs from the other two subproblems of unlabeled sensing in the presented publications. Briefly, while unlabeled selection sensing seems more tractable than unlabeled permutation sensing due to fewer combinatorial solutions, the inverse of a selection matrix does not exist. Furthermore, in contrast to both, permutation matrix as well as puncturing matrix, a selection matrix $\mathbf{S}_1 \in \{0, 1\}^{L \times L}$ is rank deficient, i.e. $\text{rank}(\mathbf{S}_1) < L$. The selection matrix can be converted to a full rank matrix, leaving the measurement matrix to be rank deficient. Therefore, the published algorithms are not fully compatible to our AMR problem with $|\mathbb{K}| = 1$.

5.4.4 Problem of Multiple Selections

Now we shift our focus from unlabeled selection sensing to the complete AMR model in (5.2.26). We encounter multiple sparse unlabeled selection sensing problems as depicted for the case of $|\mathbb{K}| = 2$ in Figure 5.3. In contrast to $|\mathbb{K}| = 1$, the support of $\mathbf{s}_k = \text{diag}(\mathbf{S}_k)$ for all $k \in \mathbb{K}$ cannot be obtained as in gray space detection of Chapter 4, since more than two hypotheses are present.

In general, to identify the correct resource allocation map \mathbf{S} , we are faced with two problems:

1. How many transmitters are active?

2. How are the resources distributed among these transmitters?

In other words, the former problem tells us, how many resource allocation sets (RASs) $\mathbb{S}_k, k \in \mathbb{K}$ are present and the latter problem tells us, whether a resource belongs to one or another transmitter. Both problems can be linked to one combinatorial problem discussed in Section 5.3. In particular, the first problem and the second problem are reflected by finding the correct set within all Bell sets of \mathbb{P} given in (5.3.1) and all Stirling sets of the second kind of \mathbb{P} given in (5.3.3), respectively. More precisely, while in the first problem the number of users $|\mathbb{K}|$ is unknown, the number of users $|\mathbb{K}| = c$ is fixed in the second problem. Therefore, while in the second problem, each resource has to be assigned to a fixed number $|\mathbb{K}|$ of RASs, in the first problem, the number of RASs is not fixed and therefore each resource can also be assigned to a new RAS \mathbb{S}_k . Consequently regarding these combinatorial sets, the second problem is a subproblem of the first problem. In summary, we have the following two problems:

1. problem: number of active transmitters is unknown and we find the correct set within all Bell sets of \mathbb{P} , as well as
2. problem: number of active transmitters is known and we find the correct set within all Stirling sets of the second kind of \mathbb{P} .

In the subsequent Section 5.5, we will derive the objective function, which shall identify the correct set among all possible sets.

5.5 Objective Function

In Section 5.4, we analyzed the structure of the AMR problem and stated two problem classes: (i) the set of all Bell sets of \mathbb{P} , and (ii) the set of all Stirling sets of the second kind of \mathbb{P} . To actually solve the AMR problem in (5.4.1), we first have to find a suitable objective function. Therefore, we will start by discussing (piecewise) continuous functions and the ℓ_1 -norm of its Fourier transforms. Based on this discussion, we turn to the discrete setting in Section 5.5.1 and analyze the ℓ_0 -pseudonorm of sparse vectors, where its corresponding discrete Fourier transforms is altered, i.e. arbitrary elements are exchange by continuous random variables. Furthermore, in Section 5.5.2 we analyze the impact on the ℓ_0 -pseudonorm, when the corresponding discrete Fourier transform of two sparse vectors are mixed. Based on these findings, we will propose the objective function for AMR in Section 5.5.3.

The aim of AMR is to identify all non-adjacent resources belonging to a single transmitter by just observing the received vector $\hat{\mathbf{y}}$ of (5.2.26). There-

fore, we follow the subsequent consideration. In the continuous setting, a *piecewise* differentiable function is *continuous*, if and only if the (inverse) Fourier transform is absolutely summable [Eng08]. In other words, the ℓ_1 -norm is bounded. Even though for discrete functions the continuity property cannot be applied, we can introduce the term *edge* to indicate a strong deviation from one sample to another. In image processing, *edge detection* is a common utility to detect strong deviations between two pixels. Furthermore, Engelberg [Eng08] applied edge detection via discrete Hilbert transform on discrete data sets. There, observing an edge in a Fourier domain signal for sample size approaching infinity, the ℓ_1 -norm of the time domain signal approaches infinity. Thus, for two transmitters as depicted in Figure 5.1, edges are generally present in the transition from one transmitter spectrum to another. Consequently, the ℓ_1 -norm might be a good measure to indicate, whether a measurement sample belongs to a specific transmitter, or not. In particular this is the case, if the time domain signals are sparse.

Subsequently, we will elaborate the above considerations for the discrete Fourier transform. Given an N -point DFT matrix Φ , an s -sparse vector \mathbf{x}_1 , and the corresponding DFT obtained by $\hat{\mathbf{x}}_1 = \Phi \mathbf{x}_1$, we will proof that exchanging an arbitrary index of $\hat{\mathbf{x}}_1$ by a random variable taken from an absolutely continuous probability distribution will increase the ℓ_0 -pseudonorm of the corresponding time domain vector, see (5.5.5). In addition, given the sparse vectors $\mathbf{x}_1, \mathbf{x}_2$, we will show that a mix up of vectors in the frequency domain will at least result at the same ℓ_0 -pseudonorm, see (5.5.8).

5.5.1 Exchange of Arbitrary Element

Given an N -point DFT matrix Φ and an s -sparse vector \mathbf{x}_1 , the corresponding DFT is obtained by $\hat{\mathbf{x}}_1 = \Phi \mathbf{x}_1$. However, only $\text{supp}(\mathbf{x}_1)$ of the s -sparse vector \mathbf{x}_1 is relevant for construction of $\hat{\mathbf{x}}_1$. Vice versa, the IDFT of $\hat{\mathbf{x}}_1$ will only have $(\mathbf{x}_1)_{[t]} \neq 0$ at indices $t \in \text{supp}(\mathbf{x}_1)$. All residual indices will cancel each other out since $(\mathbf{x}_1)_{[t]} = 0, \forall t \notin \text{supp}(\mathbf{x}_1)$.

Given two vectors $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{C}^N$, where entries of \mathbf{a}_2 or of its corresponding DFT $\hat{\mathbf{a}}_2$ are independent random variables taken from absolutely continuous probability distributions, then with probability one we have

$$(\mathbf{a}_1)_{[t]} \neq (\mathbf{a}_2)_{[t]} \quad \forall t \in \text{supp}(\mathbf{a}_2) \quad , \text{ and} \quad (5.5.1)$$

$$(\hat{\mathbf{a}}_1)_{[f]} \neq (\hat{\mathbf{a}}_2)_{[f]} \quad \forall f \in \text{supp}(\hat{\mathbf{a}}_2) . \quad (5.5.2)$$

Thus, exchanging an arbitrary index f of $\hat{\mathbf{x}}_1$ by a continuous random variable, then with probability one we have $(\hat{\mathbf{x}}_1)_{[f]} \neq (\hat{\mathbf{x}}'_1)_{[f]}$ at the exchanged

frequency index f . We denote the frequency vector with the exchanged element by $\hat{\mathbf{x}}'_1$. Thus, we obtain a 1-sparse vector $\hat{\mathbf{b}} = \hat{\mathbf{x}}'_1 - \hat{\mathbf{x}}_1$. Given the N -point DFT matrix Φ and a vector \mathbf{a} , the multiplicative uncertainty principle in [BLM17, Theorem 1] states that

$$\|\mathbf{a}\|_0 \geq \frac{N}{\|\Phi \mathbf{a}\|_0} \quad \forall \mathbf{a} \in \mathbb{C}^N \setminus \mathbf{0}. \quad (5.5.3)$$

Thus, given a frequency domain vector $\hat{\mathbf{b}} = \Phi \mathbf{b}$ with $\|\Phi \mathbf{b}\|_0 = 1$, we obtain the corresponding time domain vector \mathbf{b} with only non-zero entries, i.e. $\|\mathbf{b}\|_0 = N$.

Now, we want to obtain the lower limit of $\|\mathbf{x}'_1\|_0$ for $\mathbf{x}'_1 = \mathbf{x}_1 + \mathbf{b}$. Given two vectors $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{C}^N$, where the entries of $\hat{\mathbf{a}}_2$ are continuous independent random variables, then because of (5.5.1) the elements of $\mathbf{a}_1, \mathbf{a}_2$ are not canceling each other out. Thus, the lower limit is given by

$$\max(\|\mathbf{a}_1\|_0, \|\mathbf{a}_2\|_0) \leq \|\mathbf{a}_1 + \mathbf{a}_2\|_0. \quad (5.5.4)$$

There, the lower limit is achieved if the non-zero entries of both vectors are fully overlapping.

Consequently, given $\|\Phi \mathbf{b}\|_0 = 1$, we obtain $\|\mathbf{b}\|_0 = N$ from (5.5.3). Thus, given $\mathbf{x}'_1 = \mathbf{x}_1 + \mathbf{b}$ with $\mathbf{x}_1 \in \mathbb{C}^N$ being s -sparse and the one non-zero entry of $\hat{\mathbf{b}} \in \mathbb{C}^N$ being a continuous random variable, we obtain $\|\mathbf{x}'_1\|_0 = \|\mathbf{x}_1 + \mathbf{b}\|_0 \geq \max(\|\mathbf{x}_1\|_0, \|\mathbf{b}\|_0) = \|\mathbf{b}\|_0 = N$ from (5.5.4), i.e. $\|\mathbf{x}'_1\|_0 = N$. Consequently, given the s -sparse vector $\mathbf{x}_1 \in \mathbb{C}^N$ with $s < N$, exchanging an arbitrary index f of $\hat{\mathbf{x}}_1$ by a random variable taken from an absolutely continuous probability distribution, will result with probability one in

$$\|\mathbf{x}_1\|_0 < \|\mathbf{x}'_1\|_0. \quad (5.5.5)$$

Interestingly, the multiplicative uncertainty principle in (5.5.3) also offers insight on sparsity when more than one element are exchanged. Given a time vector $\mathbf{x}_1 \in \mathbb{C}^N$ and 1-sparse frequency vectors $\hat{\mathbf{b}}_m \in \mathbb{C}^N$, where each non-zero entry is a continuous independent random variable, then from (5.5.4) with probability one we have

$$\left\| \mathbf{x}_1 + \sum_{m=1}^M \hat{\mathbf{b}}_m \right\|_0 = N, \quad \forall M > 0. \quad (5.5.6)$$

Thus, the resulting time domain vector will be completely filled, i.e. the ℓ_0 -pseudonorm is N , independently of how many continuous independent random 1-sparse frequency vectors are added.

5.5.2 Mixing of Sparse Signals

Let us now consider a simple AMR setting, where a mix of two transmitters is received. Thus, given two sparse vectors $\mathbf{x}_1, \mathbf{x}_2$ with continuous independent random variables and two RAMs $\mathbf{S}_1, \mathbf{S}_2 = \mathbf{I} - \mathbf{S}_1$ obtained by $\text{diag}(\mathbf{S}_k) = \mathbf{1}_{\mathbb{S}_k}(f)$ from resource allocation sets \mathbb{S}_k , we have

$$\hat{\mathbf{y}} = \mathbf{S}_1 \hat{\mathbf{x}}_1 + \mathbf{S}_2 \hat{\mathbf{x}}_2. \quad (5.5.7)$$

Thus, if the correct \mathbf{S}_k are applied, then with sufficient measurements the receiver can reconstruct the sparse vectors \mathbf{x}_k via e.g. CS. Otherwise, if any other RAM is applied, the sparse vectors \mathbf{x}_k cannot be reconstructed. In the following, given the sparse vectors $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{S}_2 = \mathbf{I} - \mathbf{S}_1$, we want to proof that

$$\begin{aligned} \|\mathbf{x}_1\|_0 + \|\mathbf{x}_2\|_0 &\leq \|\Phi(\mathbf{I} - \mathbf{S}'_1) \hat{\mathbf{x}}_1 + \Phi \mathbf{S}'_1 \hat{\mathbf{x}}_2\|_0 \\ &\quad + \|\Phi(\mathbf{I} - \mathbf{S}'_2) \hat{\mathbf{x}}_2 + \Phi \mathbf{S}'_2 \hat{\mathbf{x}}_1\|_0 \end{aligned} \quad (5.5.8)$$

for any $\mathbb{S}'_1 \subset \mathbb{S}_2$ or $\mathbb{S}'_2 \subset \mathbb{S}_1$. Let us start by decomposing (5.5.7) to

$$\hat{\mathbf{y}} = \mathbf{S}_1 \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{x}}_2 + (\mathbf{S}_2 - \mathbf{S}'_1) \hat{\mathbf{x}}_2 \quad (5.5.9)$$

with $\mathbb{S}'_1 \subset \mathbb{S}_2$ and $\mathbf{S}_2 = \mathbf{I} - \mathbf{S}_1$, where $\widetilde{\mathbf{S}}_1 = \mathbf{S}_1 + \mathbf{S}'_1$ and $\widetilde{\mathbf{S}}_2 = \mathbf{S}_2 - \mathbf{S}'_1$ represent the false RAM of transmitters 1 and 2, respectively. Note that $\widetilde{\mathbf{S}}_2$ contains the correct resources of transmitter 2, but just misses the resources contained in \mathbf{S}'_1 . Assume that $\widetilde{\mathbf{S}}_2$ contains sufficient resources to reconstruct \mathbf{x}_2 , then we only have to deal with the mix up $\mathbf{S}_1 \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{x}}_2$ of (5.5.9). Taking a genie reconstruction algorithm, we obtain

$$\begin{aligned} \hat{\mathbf{x}}'_1 &= (\mathbf{I} - \mathbf{S}'_1) \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{x}}_2 \\ &= \hat{\mathbf{x}}_1 - \mathbf{S}'_1 \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{x}}_2 \\ &= \hat{\mathbf{x}}_1 + \mathbf{S}'_1 (\hat{\mathbf{x}}_2 - \hat{\mathbf{x}}_1) \\ &= \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{b}}, \end{aligned} \quad (5.5.10)$$

where $\hat{\mathbf{b}} = \hat{\mathbf{x}}_2 - \hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}'_1$ is reconstructed from $\hat{\mathbf{x}}_1$ except on the frequency positions contained in \mathbf{S}'_1 . Thus, since $\mathbf{x}_1, \mathbf{x}_2$ are constructed from continuous independent random variables, the ℓ_0 -pseudonorm of both vectors via (5.5.4) is

$$\begin{aligned} \|\mathbf{x}'_1\|_0 + \|\mathbf{x}_2\|_0 &= \|\mathbf{x}_1 + \Phi \mathbf{S}'_1 \hat{\mathbf{b}}\|_0 + \|\mathbf{x}_2\|_0 \\ &\geq \max\left(\|\mathbf{x}_1\|_0, \|\Phi \mathbf{S}'_1 \hat{\mathbf{b}}\|_0\right) + \|\mathbf{x}_2\|_0. \end{aligned} \quad (5.5.11)$$

Applying the multiplicative uncertainty principle of (5.5.3) on (5.5.11) and given an N -point DFT matrix Φ , we obtain

$$\|\mathbf{x}'_1\|_0 + \|\mathbf{x}_2\|_0 \geq \max\left(\|\mathbf{x}_1\|_0, \frac{N}{\|\mathbb{S}'_1\|_0}\right) + \|\mathbf{x}_2\|_0. \quad (5.5.12)$$

Thus, with decreasing number of entries in \mathbb{S}'_1 , the ℓ_0 -pseudonorm of (5.5.12) is increasing, except for $\mathbb{S}'_1 = \emptyset$. Consequently, (5.5.12) indicates a non-convex problem.

Decomposing (5.5.7) to

$$\hat{\mathbf{y}} = (\mathbf{S}_1 - \mathbf{S}'_2) \hat{\mathbf{x}}_1 + \mathbf{S}'_1 \hat{\mathbf{x}}_2 + \mathbf{S}'_2 \hat{\mathbf{x}}_1 + (\mathbf{S}_2 - \mathbf{S}'_1) \hat{\mathbf{x}}_2 \quad (5.5.13)$$

with $\mathbb{S}'_1 \subset \mathbb{S}_2$, $\mathbb{S}'_2 \subset \mathbb{S}_1$ and $\mathbf{S}_2 = \mathbf{I} - \mathbf{S}_1$, we obtain the ℓ_0 -pseudonorm similarly to (5.5.12) by

$$\|\mathbf{x}'_1\|_0 + \|\mathbf{x}'_2\|_0 \geq \max\left(\|\mathbf{x}_1\|_0, \frac{N}{\|\mathbb{S}'_1\|_0}\right) + \max\left(\|\mathbf{x}_2\|_0, \frac{N}{\|\mathbb{S}'_2\|_0}\right), \quad (5.5.14)$$

which is equal if $\|\mathbf{x}_1\|_0 \geq N / \|\mathbb{S}'_1\|_0$ and $\|\mathbf{x}_2\|_0 \geq N / \|\mathbb{S}'_2\|_0$. Thus, (5.5.14) proofs the inequality in (5.5.8), because we have

$$\|\mathbf{x}_1\|_0 + \|\mathbf{x}_2\|_0 \leq \|\mathbf{x}'_1\|_0 + \|\mathbf{x}'_2\|_0. \quad (5.5.15)$$

5.5.3 Joint Estimate

By taking the consideration on continuous functions as well as the ℓ_0 -pseudonorm in the discrete setting into account, the ℓ_1 -norm of an s -sparse signal seems to be a good objective to identify the correct RAM. Inspired by CS, we use the BPDN approach (3.5.18) and write the objective function to solve (5.4.1) with

$$\left(\tilde{\mathbf{x}}, \tilde{\mathbf{S}}\right) \in \underset{\mathbf{x}, \mathbf{S}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad (5.5.16)$$

$$\text{s.t. } \|\hat{\mathbf{y}} - \mathbf{S}(\mathbf{I}_{|\mathbb{K}|+1} \otimes \Phi) \cdot \mathbf{x}\|_2^2 \leq \lambda |\mathbb{P}| \cdot \mathbb{E}\left[\|\mathbf{e}_{[0]}\|^2\right], \quad (5.5.17)$$

$$\mathbf{1} = \sum_{k=0}^{|\mathbb{K}|} \mathbf{s}_k \in \{0, 1\}^{|\mathbb{P}|}, \quad (5.5.18)$$

$$\text{where } \mathbf{x} := \begin{bmatrix} \mathbf{x}_0^\top & \cdots & \mathbf{x}_{|\mathbb{K}|}^\top \end{bmatrix}^\top,$$

$$\mathbf{S} := \begin{bmatrix} \operatorname{diag}(\mathbf{s}_0) & \cdots & \operatorname{diag}(\mathbf{s}_{|\mathbb{K}|}) \end{bmatrix}.$$

The objective function returns a joint estimate of the resource allocation and the unknown. The main difference to original BPDN (3.5.18) lies in the second condition (5.5.18), where the exclusive RAM is enforced. The first condition (5.5.17) is still bounded by the ℓ_2 -error originating from the receiver noise, see (3.3.9), (3.6.15) and (3.6.42). Therefore, in certain settings, the tuning parameter λ can be obtained from a CDF to meet a certain probability. However, in our evaluation we set $\lambda = 1$ for simplicity.

Consequently, (5.5.16) finds the smallest ℓ_1 -norm of the stacked unknown vector \mathbf{x} for a given observation $\hat{\mathbf{y}}$ over all possible sets \mathbb{S} , which are given depending on the problem size by all Bell sets or all Stirling sets of the second kind. In order to solve (5.5.16), we have to jointly estimate $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{S}}$.

5.6 Simulative Analysis

To evaluate the qualification of the objective function proposed in (5.5.16), we perform brute-force simulations on the combinatorial allocation map retrieval (AMR) problem discussed in Section 5.3. To distinguish the different resource assignments in the brute-force search, we introduce a superscript on \mathbb{S} , so that candidate n of the set of resource allocation sets (RASs) is given by

$$\mathbb{S}^{(n)} := \left\{ \mathbb{S}_k^{(n)} \right\}_{k=1}^K. \quad (5.6.1)$$

While the correct resource assignment is present for $n = 0$, $n > 0$ denotes all other possible but incorrect sets of RASs. Similar notations are used for the set $\mathbb{S}'^{(n)}$ of PRB allocation sets, the overall resource allocation map (RAM) $\mathbf{S}^{(n)}$ and the overall PRB allocation map $\mathbf{S}'^{(n)}$. For simplicity and if appropriate, we will omit the superscript for $n = 0$. Hence, to solve (5.5.16), \mathbf{x} has to be estimated for each candidate $\mathbb{S}^{(n)}$.

In particular, we study two cases, the exhaustive search in all Stirling sets of the second kind with certain cardinality, as well as all Bell sets. Exhaustive search estimates \mathbf{x} for each $\mathbb{S}^{(n)}$ and evaluates if the correct $\mathbb{S}^{(0)}$ corresponds to the smallest ℓ_1 -norm of \mathbf{x} . To apply exhaustive search on feasible set sizes and still keep the reconstruction ability of the estimation algorithms presented in Section 3.5, we aggregate resources to PRBs as introduced in (5.2.7). In particular, we use a resource pool of $|\mathbb{P}| = 128$ resources with PRB-size $N_B = 16$ in the CIR case with $\mathbf{x} = \mathbf{h}$ and $|\mathbb{P}| = 256$ with $N_B = 32$ in the CAC case with $\mathbf{x} = \mathbf{v}$. In both transmission cases we have $N_{RB} = |\mathbb{P}| / N_B = 8$ PRBs within \mathbb{P} . These 8 PRBs are exclusively allocated by the transmitters, so that we observe $|\mathbb{K}| \leq 8$ active transmitters. Furthermore, we choose the CP length N_{CP} to be a multiple of 8 from 16

to $|\mathbb{P}|$. Note, that the CP length is equal to the solution space, where the sparse entries of the channel impulse responses are located. The simulation settings are summarized in Table 5.2.

In order to support good reconstruction performance via CS, the resources in each PRB are randomly distributed over the resource pool \mathbb{P} . Furthermore, exploiting $\mathbf{S} = \mathbf{f}(\mathbf{S}', \mathbb{T})$ obtained via (5.2.7) with the use of (5.2.8), we have a feasible number of Bell sets and Stirling sets of the second kind, enabling exhaustive search over all sets $\mathbb{S}'^{(n)}$ of PRB allocation sets to find the smallest ℓ_1 -norm of \mathbf{x} . In particular, the Bell number is $N_{\text{Bell}} = 4'140$ and the largest Stirling number of the second kind is $N_{\text{Stirling}} = 1'701$, cf. Table 5.1. Thus, the number of possible sets of PRBs allocation sets $\mathbb{S}'^{(n)}$ is restricted to $n = N_{\text{Bell}}$ or to $n \leq N_{\text{Stirling}}$ for the Bell sets or the Stirling sets of the second kind, respectively. Thus, we can write (5.2.26)

$$\hat{\mathbf{y}} = \mathbf{f}(\mathbf{S}', \mathbb{T}) \cdot (\mathbf{I}_{K+1} \otimes \Phi) \cdot \mathbf{x} + \hat{\mathbf{e}}, \quad (5.6.2)$$

where $K = 8$ and $\mathbf{f}(\mathbf{S}', \mathbb{T})$ is obtained via (5.2.7) with the use of (5.2.8). Note that generally some \mathbf{S}'_k of $k \in \mathbb{K}$ will be zero, whereby these transmitters have no resources allocated. Since we apply exhaustive search, we can split our objective function (5.5.16) in two parts. First, we simply estimate $\tilde{\mathbf{x}}$ for each candidate $\mathbf{S}'^{(n)}$ via BPDN (see Section 3.5.2)

$$\tilde{\mathbf{x}}^{(n)} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad (5.6.3)$$

$$\text{s.t.} \quad \left\| \hat{\mathbf{y}} - \mathbf{f}(\mathbf{S}'^{(n)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \Phi) \cdot \mathbf{x} \right\|_2^2 \leq \lambda \mathbb{E} \left[\|\hat{\mathbf{e}}\|_2^2 \right].$$

Then, we obtain a candidate by

$$\tilde{n} \in \underset{n}{\operatorname{argmin}} \left\| \tilde{\mathbf{x}}^{(n)} \right\|_1. \quad (5.6.4)$$

If $\tilde{n} = 0$, we identified the correct candidate. The Stirling sets of the second kind are generated for $|\mathbb{S}'^{(0)}|$, which is the cardinality of the correct set $\mathbb{S}'^{(0)}$ of PRB allocation sets. Thus, it represents the number of actual active transmitters.

Now, we will describe the simulation settings of the exhaustive search analysis. We select $\mathbb{S}'^{(0)}$ randomly from all Bell sets. In this way, we have a number of transmitters between 1 and 8 distributed according to the distribution of the Bell sets. Therefore, for each selected PRB each transmitter has N_{B} resources distributed within \mathbb{P} , enabling CS methods. In the CIR case, $\Phi \in \mathbb{C}^{N_{\text{DFT}} \times N_{\text{CP}}+1}$ is set to a submatrix of the DFT-matrix of size $N_{\text{DFT}} = |\mathbb{P}| = 128$, and in the CAC case we have $\Phi \in \mathbb{C}^{N_{\text{DFT}} \times 2N_{\text{CP}}+1}$

	$ \mathbb{P} = N_{\text{DFT}}$	N_{RB}	N_{B}	N_{CP}	s -sparse	$ \mathbb{K} $
CIR	128	8	16	16, 24, \dots, 128	1, 2, 4, 8, 16	≤ 8
CAC	256		32			

Table 5.2: Simulation settings for Figures 5.10 to 5.13.

with $N_{\text{DFT}} = 256$. Finally, the entries of each s -sparse unknown \mathbf{x}_k are randomly obtained from the circularly-symmetric complex normal distribution, whereas the location of the non-zero entries are uniformly randomly distributed within N_{CP} .

The evaluation of the brute-force AMR simulations is divided into four parts: First we will consider BPDN in Section 5.6.1 and evaluate its AMR performance based on the ℓ_1 -norm and channel transfer function measurements. Then, we will consider OMP in Section 5.6.2, where we introduce a joint ℓ_0 -pseudonorm ℓ_2 -error measure to improve AMR performance. Afterwards in Section 5.6.3, we evaluate the robustness of AMR via OMP with noisy measurements. Finally, instead of channel transfer function measurements, we will take a look on AMR with SCP measurements in Section 5.6.4.

5.6.1 BPDN

The results presented in Figure 5.10 for channel transfer function measurements are obtained via BPDN (SPGL1-toolbox [BF15]) for $N_{\text{DFT}} = 128$, $N_{\text{CP}} \in \{16, 24, \dots, 128\}$ and $s = 2^m$, $m \in \{0, \dots, 4\}$, see Table 5.2. A failure rate is shown for a given pair (N_{CP}, s) . In Figure 5.10, the performance of successfully recovering $\mathbb{S}^{(0)}$ is shown for BPDN (5.6.3) and ℓ_1 -norm (5.6.4). If reconstruction has been successful, then the ℓ_1 -norm of the correct solution with $\mathbb{S}^{(0)}$ is smaller than all other solutions $\mathbb{S}^{(n)}$, $n > 0$ within the given set of sets. While the black region indicates, that there is almost no recovery possible, the stained pattern marks the region, where all reconstructions were successful for up to 10^5 simulations. The brighter the region, the better is the reconstruction ability via exhaustive search.

In Figures 5.10a and 5.10b, the performance is given for recovering the correct set within all Bell sets and all Stirling sets of the second kind, respectively. Figure 5.10a shows, that taking the lowest ℓ_1 -norm of the BPDN solutions of all Bell sets does not retrieve the correct set in general. However, for large solution spaces (CP length) and small sparsity of 1-sparse and 2-sparse signals, the correct set is obtained in around 99% of the simulations. The picture changes completely in Figure 5.10b just considering all

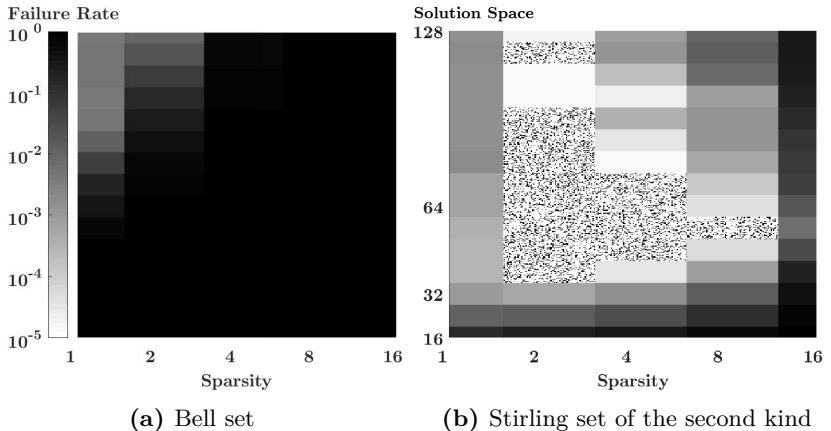


Figure 5.10: Failure rate of identifying the correct set via the ℓ_1 -norm of the BPDN solutions for the CIR over sparsity and solution space within 128 samples for SNR of infinity.

Stirling sets of the second kind (1'701 in the worst case). Even though the worst case set size is only about half the Bell number, the recovery ability is heavily increased. Especially in the center regions (stained pattern), exhaustive search over all Stirling sets of the second kind always obtained the correct set via the ℓ_1 -norm of BPDN for around 10^5 simulations. Therefore, a failure rate in this region below 10^{-5} is anticipated.

In the following, we will take a brief analysis of the results. At first we just consider the mere recovery ability of BPDN. In general, uniform random Fourier samples achieve the best recovery performance for partial Fourier measurements (Section 3.6). However, the higher s of s -sparse signals, the higher the requirements on the distribution of the Fourier samples. Consequently, the closer s is tending to the PRB-size of $N_B = 16$, the more likely that BPDN fails. Moreover, for s -sparse signals with $s \geq 16$ a single PRB is in general not sufficient for successful recovery via BPDN, see Section 3.4. Thus, each transmitter needs to allocate more PRBs to enable successful recovery for $s \geq 16$. Since in our simulations we made no restrictions on the number of PRBs per transmitter, it is highly probable that one transmitter will be left with a single PRB and cannot be recovered. Therefore, the recovery fail of this transmitter leads to complete failure in our simulations. This explains the increase of failure rate towards $s = 16$ in Figure 5.10b.

Further points are the two distinction problems discussed in Section 5.4.4. Since for Stirling sets of the second kind the number of transmitters is fixed,

only one of the two distinctions has to be made: finding the most probable transmitter for each PRB, such that all transmitters have at least one PRB. In contrast to the Stirling set of the second kind, the number of transmitters in the Bell set is unknown. Consequently, also the second distinction problem (activity detection) has to be solved within Bell sets: distinguish whether a set of PRBs belongs to a single transmitter or partitions of a set of PRBs belong to multiple distinct transmitters. Therefore, besides the larger size of all Bell sets, a further distinction problem has to be solved, explaining the recovery performance gap between Figure 5.10a and Figure 5.10b. In particular, BPDN (5.6.3) seems not well suited for activity detection. Taking a closer look on Figure 5.10a, we observe reconstruction performance for large solution spaces and high sparsity $s \leq 2$. Due to the high sparsity, a strong distinction between the resource sets of each transmitter may be given, enabling activity detection for up to 99% of the simulations. In summary, knowing the number of active transmitters in advance or acquiring them through activity detection can greatly reduce the problem size from the Bell number to the Stirling number of the second kind. Nevertheless, even though complete reconstruction is not possible, partial reconstruction may be still available.

5.6.2 OMP

Now, we turn our attention from the convex BPDN method to the greedy OMP approach, see Algorithm 3.1 in Section 3.5.3. Similar to the simulation via BPDN we want to know the brute-force performance of OMP via exhaustive search. We will use the same settings as for the BPDN simulations, see Table 5.2. However, in contrast to BPDN, the ℓ_1 -norm proposed in (5.6.4) for OMP will not suite as a measure to identify the correct overall RAM \mathbf{S} . In general, the BPDN approach using the SPGL1-toolbox will quickly converge, such that (5.6.3) is solved for an ℓ_2 -error very close to the constraint $\lambda \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]$. However, due to the design of OMP, the ℓ_2 -error strongly deviates for different solutions. Consequently, in contrast to BPDN, where the candidate is identified only via the ℓ_1 -norm of \mathbf{x} in (5.6.4), for OMP we have to include the ℓ_2 -error in the identification process as well. Therefore, instead of (5.6.4), we have to use an objective function which takes the ℓ_2 -error into account

$$\tilde{n} \in \underset{n}{\operatorname{argmin}} \left(\left\| \tilde{\mathbf{x}}^{(n)} \right\|_1 + \lambda' \left\| \hat{\mathbf{y}} - \mathbf{f}(\mathbf{S}'^{(n)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \Phi) \cdot \tilde{\mathbf{x}}^{(n)} \right\|_2^2 \right), \quad (5.6.5)$$

where the tuning parameter λ' has to be appropriately chosen. In addition, two further remarks should be taken into account. First, since we perform

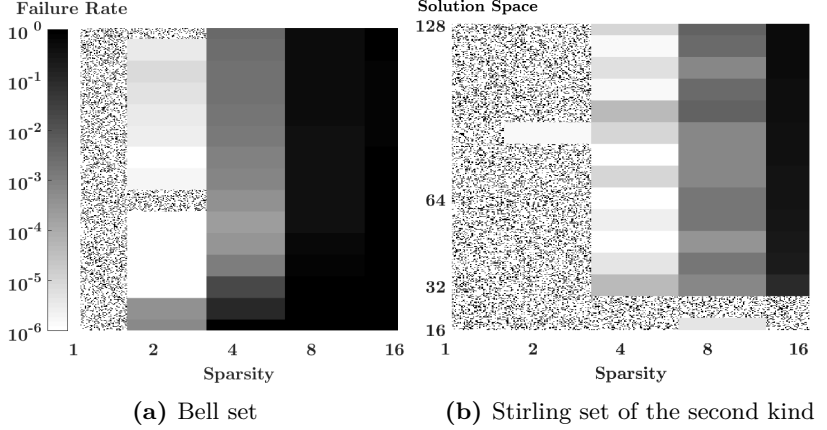


Figure 5.11: Failure rate of identifying the correct set via the joint ℓ_0 -pseudonorm ℓ_2 -error measure of the OMP solutions for the CIR over sparsity and solution space within 128 samples for SNR of infinity.

exhaustive search, we solve (5.6.3) via OMP given in Algorithm 3.1 for each candidate $\mathbf{S}'^{(n)}$ satisfying the ℓ_2 -error constraint, where the measurement matrix of OMP is given by $\mathbf{f}(\mathbf{S}'^{(n)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \Phi)$. Consequently, each solution $\mathbf{S}'^{(n)}$ will be paired with an $\tilde{\mathbf{x}}^{(n)}$, where the ℓ_2 -error is smaller or equal to the constraint $\lambda \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]$. Secondly, due to the iterative construction of $\tilde{\mathbf{x}}^{(n)}$, the number of iterations performed by the OMP algorithm directly specifies the number of non-zero entries of the candidate vector $\tilde{\mathbf{x}}^{(n)}$. Therefore, we first obtain $\tilde{\mathbf{x}}^{(n)}$ for each allocation $\mathbf{S}'^{(n)}$ via the OMP algorithm solving (5.6.3). Afterwards, we take the joint measure of ℓ_0 -pseudonorm and ℓ_2 -error to identify the candidate with

$$\tilde{n} \in \underset{n}{\operatorname{argmin}} \left(\|\tilde{\mathbf{x}}^{(n)}\|_0 + \frac{\|\hat{\mathbf{y}} - \mathbf{f}(\mathbf{S}'^{(n)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \Phi) \cdot \tilde{\mathbf{x}}^{(n)}\|_2^2}{\lambda \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]} \right). \quad (5.6.6)$$

Given $\lambda = 1$, the ℓ_2 -error is normalized by the constraint term, whereby merely a fractional part is added to the ℓ_0 -pseudonorm. Consequently, the ℓ_0 -pseudonorm represents the main classifier, whereas the ℓ_2 -error only effect the choice of the solution if multiple candidates $\mathbf{S}'^{(n)}$ have equal ℓ_0 -pseudonorm.

In Figure 5.11, we show the reconstruction performance of OMP using the joint ℓ_0 -pseudonorm ℓ_2 -error measure in (5.6.6). In comparison to the

BPDN simulations, we observe a completely different picture. Note, that the failure rate is one magnitude smaller than for Figure 5.10 reaching failure rates down to 10^{-6} . While BPDN with the ℓ_1 -norm measure proposed in (5.6.4) is in general not capable to reconstruct the correct resource allocation within all Bell sets, Figure 5.11a shows that OMP with (5.6.6) is able to reconstruct $\mathbf{S}'^{(0)}$ with high probability within almost all solution spaces for 1-sparse and 2-sparse signals. Only due to the smaller depicted magnitude compared to BPDN, the failure rate of 2-sparse signals is visible. Consequently, the performance results indicate that the OMP approach is suited to solve both distinction problems, activity detection and resource mapping. In addition, the OMP approach outperforms BPDN also for the Stirling sets of the second kind where Figure 5.11b shows almost no sign of failure for the complete region of 1-sparse and 2-sparse signals. Even for 4-sparse signals, the failure rate is reaching a level below the detection limit of the BPDN simulations. However, for 8-sparse signals and above, the performance of OMP degrades compared to the BPDN simulations in Figure 5.10b. The main reason for the worse performance is that OMP is approaching the recovery limit, set by the PRB-size $N_B = 16$, earlier than BPDN, since the RIP bound of OMP is larger than of BPDN, see Section 3.5. At small solution spaces up to $N_{CP} \leq 24$, OMP benefits from the blocksize $N_B = 16$, leading to reconstruction performance below the detection limit. In particular, for s -sparse signals with large s OMP simply reduces to the LS estimator, solving a mere balanced linear system.

Summarizing the obtained results, the OMP approach is better suited to obtain the correct resource allocation within both sets: the Bell sets and the Stirling sets of the second kind. Moreover, OMP can successfully solve both distinction problems, activity detection and resource mapping. In all, the superior performance of the OMP approach over BPDN may be explained by the different applied measures, the mere ℓ_0 -norm for BPDN and the joint ℓ_0 -pseudonorm ℓ_2 -error for OMP. However, the measure of OMP cannot be easily transferred to BPDN, and vice versa.

5.6.3 Noisy Measurements

To evaluate the robustness of the objective function proposed in (5.5.16) in noisy environments, we performed brute-force simulations for an SNR of 20 dB and the usual settings summarized in Table 5.2. We selected the OMP approach with measure proposed in (5.6.6), since it performed better than the BPDN approach with the mere ℓ_1 -norm. The failure rate is shown in Figure 5.12, where Figure 5.12a and Figure 5.12b present the performance for finding the correct set among all Bell sets and Stirling sets of the second

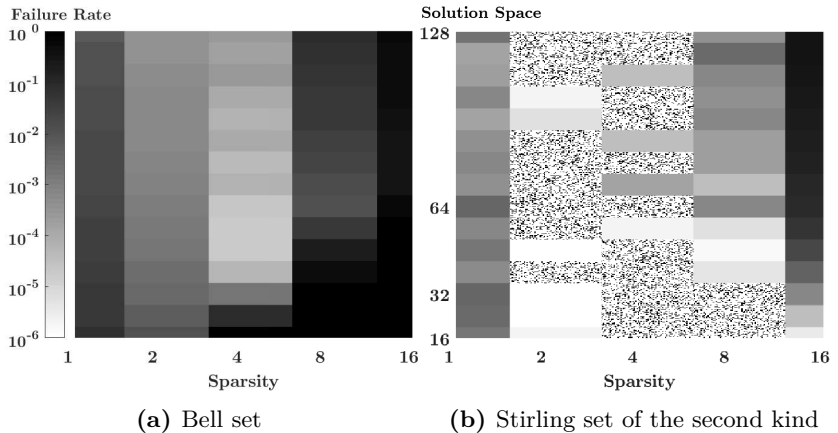


Figure 5.12: Failure rate of identifying the correct set via the joint ℓ_0 -pseudonorm ℓ_2 -error measure of the OMP solutions for the CIR over sparsity and solution space within 128 samples for SNR of 20 dB.

kind, respectively. Obviously and trivially, the performance decreases under noisy settings. However in general, recovery is still possible for an SNR of 20 dB. The best results are achieved with 4-sparse signals, where recovery among all Bell sets goes down to a failure rate of 10^{-4} and among all Stirling sets of the second kind is beyond the detection limit.

Comparing both plots Figure 5.12 with Figure 5.11, we see directly that the failure rate for 1-sparse signals increases significantly compared to other s -sparse signals. This increase majorly results from two effects. First, the s -sparse entries are randomly i.i.d. drawn from the circularly-symmetric complex normal distribution. Thus, the amplitude of each 1-sparse signal can differ strongly. At the receiver, AWGN is applied on the normalized superimposed frequency signal, such that an SNR of 20 dB is obtained. Consequently, while some 1-sparse signals experience only small distortion due to the large amplitude, others perceive strong distortion compared to the selected SNR. This is one reason why the failure rate increases more strongly for 1-sparse signals. Secondly, there is a chance that multiple 1-sparse signals coincide in the position of the non-zero entry, whereby the distinction between the 1-sparse signals is only given by the direct difference of the values on the complex plane. The smaller the difference between two 1-sparse signals, the larger the probability that these two signals are seen as one signal under noise distortion. Note, that due to the complex domain, 1-sparse

signals have two degrees of freedom. In particular, 1-sparse signals can be seen as 2-sparse signals, where the positions of the entries are fixed. Even though the complex plain offers more degrees of freedom, the chance that two 1-sparse signals are identified as one is still present. While the first effect can be seen for all selected solution spaces, the second effect decreases with increasing solution space. However, we did not evaluate how strong the second effect is influencing the performance.

Nevertheless, the influence of both effects change with increasing sparsity. The first one decreases with increasing sparsity, since multiple entries are randomly i.i.d. drawn from the circularly-symmetric complex normal distribution, leading to an averaging of each s -sparse signal. For the second effect, the chance of coinciding positions depends of the binomial distribution given s and the solution space, having the lowest probability at an $s = N_{\text{CP}}/2$. In contrast to that, the distance in the complex plain has to be small for all s entries between the two s -sparse signals, which is less probable for larger s .

Summarizing, from the simulation point of view the OMP approach is also robust against noise. In general, the OMP approach is able to recover the correct set among all Bell sets as well as all Stirling sets of the second kind. Especially, 4-sparse signals achieve good reconstruction performance.

5.6.4 Spectral Channel Power Measurements

In the last simulations in Figure 5.13, we consider the identification performance via spectral channel power measurements. Consequently, the sparsity of the CAC $\mathbf{x} = \mathbf{v}$ is exploited to identify the correct set using the proposed objective function in (5.5.16). We consider just the Stirling sets of the second kind (Figure 5.13), but for noiseless setting (Figure 5.13a) and an SNR of 20 dB (Figure 5.13b). Note that the settings of the CAC simulations change compared to the CIR simulations, see Table 5.2. In particular, the resource pool and the PRB-size increase to $|\mathbb{P}| = 256$ and $N_{\text{B}} = 32$, respectively. This is due to the increased support region of the CAC compared to the CIR, see (3.6.29). Note that the sparsity is stated for the underlying CIR of the CAC. The precise sparsity of the CAC cannot be directly obtained, but is given by $|\text{supp}(\mathbf{v})| \leq \max(2N_{\text{CP}} + 1, s^2 - s + 1)$, see (3.6.45) and (3.6.49). Therefore, the results of the CAC in Figure 5.13 cannot be simply compared with the previous results. Nevertheless, we can obtain some performance indication for spectral channel power observations. Also note that the failure rate limits between Figure 5.13a and Figure 5.13b differ.

In the noiseless case, Figure 5.13a shows similar results than for the CIR in Figure 5.11b. In general, the similarity is plausible, since the s -sparse

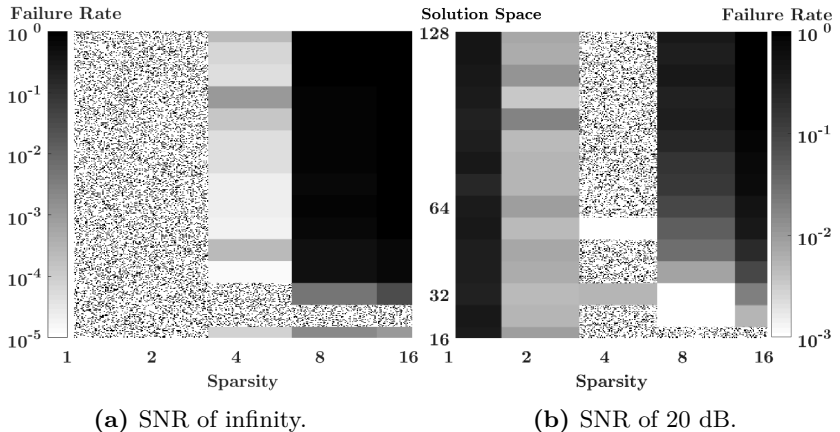


Figure 5.13: Failure rate of identifying the correct set in all Stirling sets of the second kind via the joint ℓ_0 -pseudonorm ℓ_2 -error measure of the OMP solutions for the CAC over sparsity and solution space within 256 samples.

entries of the underlying channel are restricted on the same solution space as for channel transfer function measurements. Therefore, the restriction of the solution space is also present for a CAC \mathbf{v} with larger sparsity. Consequently, in the noiseless setting good reconstruction performance among Stirling sets of the second kind can be achieved for underlying sparsity up to $s = 4$ and solution space up to 24.

In the noisy setting of 20 dB SNR in Figure 5.13b, the failure detection limit is already at 10^{-3} . However, similar results as for the CIR in Figure 5.12b are indicated. The best results are achieved for underlying sparsity of $s = 4$, which are beyond the detection limit of 10^{-3} . As discussed for Figure 5.12, an underlying 1-sparse signal is strongly distorted by noise, such that the recovery performance drops dramatically. Consequently, the correct set is unlikely to be retrieved from spectral channel power measurements, if the underlying sparsity is $s = 1$.

In summary, based on the exhaustive search simulations, the objective function proposed in (5.5.16) can be regarded as a good candidate to reconstruct the correct overall RAM among others from noisy channel transfer function measurements and spectral channel power measurements. However, the BPDN approach with the measure based on the mere ℓ_1 -norm proposed in (5.6.4) seems to be not suitable. Especially, BPDN hardly finds the correct set among all Bell sets. In contrast to BPDN, the OMP approach with

the joint measure of ℓ_0 -pseudonorm and ℓ_2 -error, see (5.6.6), achieves good reconstruction among Bell sets and Stirling sets of the second kind. Moreover, as observed in the brute-force simulations, the OMP method is also robust in noisy settings.

5.7 Structured Allocation Map

To obtain an efficient algorithm instead of the exhaustive search traversing through all possible candidates like in the previous Section 5.6, we have to significantly decrease the combinatorial complexity. Therefore, we propose an AMR algorithm that performs local decisions, which are generally sub-optimal on the global scale. In particular, we structure the combinatorial AMR problem as a [decision tree](#), which can be traversed via [breadth-first search](#). At each iteration, a local decision is made, assigning the most probable physical resource block (PRB) to a transmitter. The particularly proposed breadth-first search approach is given in Algorithm 5.1.

The input parameters of Algorithm 5.1 are almost identical to Algorithm 3.1, except that we also require the distinct PRB set \mathbb{T} . We initialize the estimated set $\tilde{\mathcal{S}}'$ of PRB allocation sets with one set containing one index of \mathbb{T} and initialize a temporary PRB set \mathbb{T}' with the remaining indices of \mathbb{T} . Then, we traverse through all indices of \mathbb{T}' (line 2) and all sets of $\tilde{\mathcal{S}}'$ including an empty set (line 3), constructing the resource allocation candidates $\mathcal{S}'^{(m,k)}$ for the current iteration (line 4-8). Furthermore, we calculate the estimate $\tilde{\mathbf{x}}^{(m,k)}$ for each resource allocation candidate via OMP with Algorithm 3.1 (line 9). After the construction of all candidates at the current level, we use the joint ℓ_0 -pseudonorm ℓ_2 -error measure from (5.6.6) in line 12 to obtain the most probable candidate in line 13. Furthermore, the index m' is removed from \mathbb{T}' in line 14. The iterative construction of $\tilde{\mathcal{S}}'$ is continued until the temporary PRB set \mathbb{T}' is empty (line 1). Finally, the last estimate $\tilde{\mathbf{x}}^{(m,k)}$ and the estimated set $\tilde{\mathcal{S}}'$ of PRB allocation sets are returned.

Since the order within sets are arbitrary, we cannot easily compare the estimate $\tilde{\mathcal{S}}'$ with the original set \mathcal{S}' of PRB allocation sets. Moreover, if $\tilde{\mathcal{S}}'$ is not correctly estimated, then we require a measure to assess the quality of our estimate. Informally, we need to find a sequence of sets, such that the sum of intersections between each $\tilde{\mathcal{S}}'_k$ and \mathcal{S}'_k is maximized. In other words, given a group of the permutations of the transmitter set \mathbb{K} with $\mathfrak{S}(\mathbb{K})^4$ and given a [permutation](#) $\Pi \in \mathfrak{S}(\mathbb{K})$, then we search the maximal number of intersections over all possible permutations in $\mathfrak{S}(\mathbb{K})$. Thus, we can assess

⁴Frequently also referred to as [symmetric group](#) on a set \mathbb{K} .

Algorithm 5.1 AMR for structured allocation pattern via OMP

Input: $\hat{\mathbf{y}}, \mathbb{T} = \{\mathbb{T}_0, \dots, \mathbb{T}_{N_{\text{RB}}-1}\}, \mathbf{\Phi} \in \mathbb{C}^{N \times M}, \lambda, \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]$

Initialize: $\tilde{\mathbb{S}}' = \{\{0\}\}, \mathbb{T}' = \{1, \dots, |\mathbb{T}| - 1\},$

```

1: while  $|\mathbb{T}'| > 0$  do
2:   for all  $m \in \mathbb{T}'$  do
3:     for  $k = 0$  to  $|\tilde{\mathbb{S}}'|$  do
4:       if  $k = |\tilde{\mathbb{S}}'|$  then
5:          $\mathbb{S}'^{(m,k)} = \left\{ \tilde{\mathbb{S}}'_0, \dots, \tilde{\mathbb{S}}'_{|\tilde{\mathbb{S}}'|-1}, m \right\}$ 
6:       else
7:          $\mathbb{S}'^{(m,k)} = \left\{ \tilde{\mathbb{S}}'_0, \dots, \tilde{\mathbb{S}}'_k \cup m, \dots, \tilde{\mathbb{S}}'_{|\tilde{\mathbb{S}}'|-1} \right\}$ 
8:       end if
9:       Via OMP (Algorithm 3.1) obtain
          
$$\tilde{\mathbf{x}}^{(m,k)} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1$$

          s.t.  $\|\hat{\mathbf{y}} - \mathbf{f}(\mathbf{S}'^{(m,k)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \mathbf{\Phi}) \cdot \mathbf{x}\|_2^2 \leq \lambda \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]$ 
10:      end for
11:    end for
12:    
$$\left( \tilde{m}, \tilde{k} \right) = \underset{m,k}{\operatorname{argmin}} \left( \left\| \tilde{\mathbf{x}}^{(m,k)} \right\|_0 \right. \\ \left. + \frac{\left\| \hat{\mathbf{y}} - \mathbf{f}(\mathbf{S}'^{(m,k)}, \mathbb{T}) \cdot (\mathbf{I} \otimes \mathbf{\Phi}) \cdot \tilde{\mathbf{x}}^{(m,k)} \right\|_2^2}{\lambda \mathbb{E}[\|\hat{\mathbf{e}}\|_2^2]} \right)$$

13:     $\tilde{\mathbb{S}}' = \mathbb{S}'^{(\tilde{m}, \tilde{k})}$ 
14:    Remove  $\tilde{m}$  from  $\mathbb{T}'$ 
15:  end while
16: return  $\tilde{\mathbb{S}}', \tilde{\mathbf{x}}^{(\tilde{m}, \tilde{k})}$ 

```

the quality of the estimated set $\tilde{\mathbb{S}}'$ of PRB allocation sets by the ratio of maximal number of intersections to the number of PRBs $N_{\text{RB}} = |\mathbb{T}|$ with

$$\alpha_{\text{RB}+\text{Inactive}} = \frac{1}{N_{\text{RB}}} \max_{\Pi \in \mathfrak{S}(\mathbb{K})} \sum_{k=0}^{|\mathbb{K}|} \left| \tilde{\mathbb{S}}'_{\Pi(k)} \cap \mathbb{S}'_k \right|. \quad (5.7.1)$$

Furthermore, (5.7.1) can be rewritten in matrix notation using the **inner product** between the estimates $\tilde{\mathbf{s}}'_{\Pi(k)}$ and the original allocation vectors \mathbf{s}'_k by

$$\alpha_{\text{RB}+\text{Inactive}} = \frac{1}{N_{\text{RB}}} \max_{\Pi \in \mathfrak{S}(\mathbb{K})} \sum_{k=0}^{|\mathbb{K}|} \left(\tilde{\mathbf{s}}'_{\Pi(k)} \right)^\top \cdot \mathbf{s}'_k. \quad (5.7.2)$$

Therefore, in order to obtain the success rate of our Algorithm 5.1, we just need to traverse through all possible permutations to find the best match between our estimate and the correct solution. However, as also discussed in Section 5.3, for large number of transmitters $|\mathbb{K}|$ the number of permutations quickly increases $(|\mathbb{K}| + 1!)$. To cope with large numbers of transmitters we apply the **Hungarian algorithm** [Kuh55], which had been originally proposed by Kuhn in 1955.

In addition, also the number of subsets can differ between $\tilde{\mathbb{S}}'$ and \mathbb{S}' . In particular, the estimated set $\tilde{\mathbb{S}}'$ of PRB allocation sets can consist of less or more transmitters $|\tilde{\mathbb{K}}'|$ than $|\mathbb{K}'|$. Furthermore, $\tilde{\mathbb{S}}'$ may also contain the set \mathbb{S}'_0 of inactive resources, but partitioned in multiple subsets. The evaluation in (5.7.2) interprets the inactive resources as a classical transmitter, which have to be correctly estimated. Thus, the **RB+Inactive ratio** is obtained via (5.7.2) and the Hungarian algorithm.

However, in many scenarios we are more interested in successful recovery of the active resources. In particular, the grouping of inactive resources is not of interest as long as the inactive resources are separately grouped from the active resources. Thus, we obtain the **RB ratio** with

$$\alpha_{\text{RB}} = \frac{1}{\sum_{k=1}^{|\mathbb{K}|} \|\mathbf{s}'_k\|_0} \cdot \max_{\Pi \in \mathfrak{S}(\mathbb{K})} \sum_{k=1}^{|\mathbb{K}|} \left(\tilde{\mathbf{s}}'_{\Pi(k)} \right)^\top \cdot \mathbf{s}'_k. \quad (5.7.3)$$

Note that the difference between (5.7.3) and (5.7.2) is also the start of the sum, which begins here with $k = 1$.

One further evaluation represents the complete recovery of all resources for a given transmitter. The **Tx ratio** is given by the number of transmitters completely recovered to the overall number of transmitters, which is

calculated by

$$\alpha_{Tx} = \frac{1}{|\mathbb{K}|} \max_{\Pi \in \mathfrak{S}(\mathbb{K})} \left[\frac{1}{\|\mathbf{s}'_k\|_0} \sum_{k=1}^{|\mathbb{K}|} \left(\tilde{\mathbf{s}}'_{\Pi(k)} \right)^\top \cdot \mathbf{s}'_k \right], \quad (5.7.4)$$

where $\lfloor \cdot \rfloor$ represents the flooring function, such that only complete recovered transmitters are counted.

Finally, to also evaluate the performance in comparison to mere *gambling*, we select two cases of resource assignment without estimation. In the first case, we just assign all resources to a single transmitter. Therefore, the ratio is obtained by the transmitter with largest number of resources divided by the overall number of PRBs

$$\alpha_{\text{Gambling01}} = \frac{1}{N_{\text{RB}}} \max_{k \in \mathbb{K}} [\mathbf{s}'_k]. \quad (5.7.5)$$

In the second case, we just assign one resource to each transmitter, resulting in a ratio simply calculated by $|\mathbb{K}|/N_{\text{RB}}$. To assess performance by the *gambling ratio*, we just take the best shot of both

$$\alpha_{\text{Gambling}} = \max \left[\frac{|\mathbb{K}|}{N_{\text{RB}}}, \frac{1}{N_{\text{RB}}} \max_{k \in \mathbb{K}} [\mathbf{s}'_k] \right]. \quad (5.7.6)$$

5.8 Results

We perform numerous simulations to evaluate our proposed Algorithm 5.1 for AMR. Therefore, we consider different SNRs to evaluate the robustness against noise in Figures 5.14 and 5.15. Furthermore, we apply different sparsity on the channel in Figure 5.16 and different physical resource blocksizes N_B on the transmission system in Figure 5.17. Finally, we also obtain the performance over commonly used performance channels in Figure 5.18, including the short-range factory automation performance channels obtained in our measurement campaign (see Section 4.6.1). The main settings of the simulations are summarized in Table 5.3.

As introduced in the previous Section 5.7, we consider three ratios to evaluate the performance and one ratio to compare the performance to mere *gambling* (5.7.6). The three performance ratios are *RB* (5.7.3), *Tx* (5.7.4), and *RB+Inactive* (5.7.2). In principle, we will focus on the recovery of active resources and the recovery of complete transmitters denoted as *RB* and *Tx*, respectively. Furthermore, we just plot one *gambling* curve per figure, which represents the best gambling for all settings within the corresponding figure.

	Model	SNR	N_{DFT}	$ \mathbb{P} $	N_{CP}	N_{B}	$ \mathbb{S}'_k $	$ \mathbb{K} $
CIR	8-sparse	10 dB	512	512	512	8	$\geq \frac{ \mathbb{P} }{8N_{\text{B}}}$	4
CAC					256	32		
CIR	EPA	0 dB	1024	600	72	24	$\geq \frac{ \mathbb{P} }{4N_{\text{B}}}$	
CAC								

Table 5.3: Simulation settings for Figures 5.14 to 5.18.

All the simulations are conducted with four transmitters. Furthermore, we consider two different channel models, the s -sparse channel model and the performance channel model. In addition, we also consider the two different frequency observations, channel transfer function measurements as introduced in Section 5.2.1 and spectral channel power measurements as introduced in Section 5.2.2. To directly observe the impact of sparse channels on the estimation performance, we consider the s -sparse channel model, where random i.i.d. variables are drawn from the circularly-symmetric complex normal distribution and uniformly distributed on s positions of the channel vector. The s -sparse channel model simulations are commonly performed for $s = 8$, SNR of 10 dB, and resource pool size of 512. The performance channel models shall assess the ability in practical scenarios. We mainly use the EPA model for performance evaluation at SNR of 0 dB. Furthermore, we base the transmission settings on a 10 MHz system commonly used in 3GPP LTE [3GP20a] and NR [3GP20h]. Therefore, the resource pool only consists of 600 resources with $N_{\text{DFT}} = 1024$ and CP length of 72. Finally, to enable reconstruction via compressed sensing methods and especially OMP, we enforced a minimum number of physical resource blocks $|\mathbb{S}'_k|$ allocated to each transmitter. The common settings for all simulations can be found in Table 5.3. If settings differ from Table 5.3, it will be separately stated in the text.

At first we turn our attention to the robustness of Algorithm 5.1 against noise for s -sparse channels in Figure 5.14. For observations of the channel transfer function we consider only 8-sparse channels with physical resource blocksize $N_{\text{B}} = 16$, but for observations of the spectral channel power we additionally present the plots for 4-sparse channels with blocksize $N_{\text{B}} = 32$. While we can distinguish the channel impulse response (CIR) of 4 transmitters with increasing SNR and by that retrieve the corresponding resource allocation, the distinction of 4 transmitters via their circular autocorrelation of the CIR (CAC) does not work for the same setting. This can be

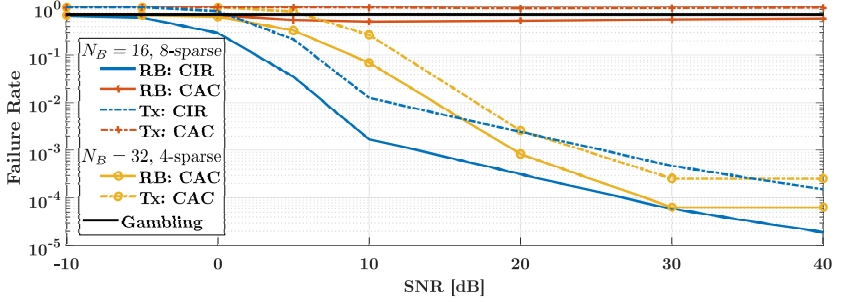


Figure 5.14: Success rate over SNR for s -sparse channels.

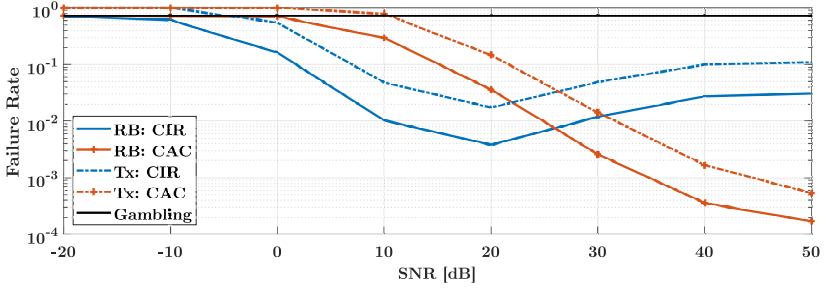


Figure 5.15: Success rate over SNR for 3GPP EPA performance channel.

particularly seen in comparison to mere gambling, where a similar success probability is achieved for identifying the resource allocation of the active resources. Obviously, the circular autocorrelation of 8-sparse channels cannot be reconstructed by mere blocksize $N_B = 16$, leading already to wrong decisions of Algorithm 5.1 within the first iterations. However, increasing the blocksize and sparsity enable also the estimation of the resource allocation for spectral channel power observations and therefore can distinguish the CAC of the 4 transmitters. Since it is more demanding to estimate a single transmitter correctly than just a certain number of overall resources, the Tx estimation performance is worse than the performance based on (5.7.3). At low SNR, the probability of Algorithm 5.1 identifying the correct physical resource blocks belonging to the same transmitter approaches the performance of mere gambling.

Changing to practical performance channels like 3GPP EPA in Figure 5.15, we observe lower CP length, but only compressible CIR. Channel transfer functions as well as spectral channel power observations are obtained from equal simulation settings. Consequently, AMR via the CIR

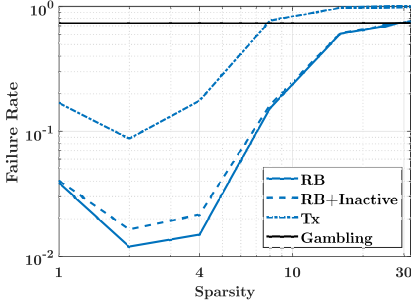


Figure 5.16: Success rate over s of s -sparse channels at an SNR of 10 dB.

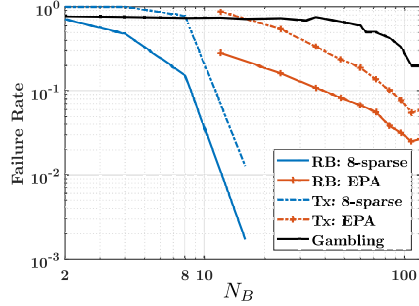


Figure 5.17: Success rate over the physical resource block-size N_B .

is already successful at a lower SNR than the estimation via the CAC, thus we already have some visible performance gain at 0 dB SNR. Strangely, however, starting with an SNR of 20 dB, the estimation performance via CIR deteriorates, so that an SNR of 20 dB achieves the best performance for a channel transfer function in the observed SNR region. This effect is not seen for estimation via the CAC, which curiously outperforms channel transfer function measurements for SNR of 30 dB and above. A reason for the deterioration may be the use of the joint ℓ_0 -pseudonorm ℓ_2 -error measure described in (5.6.6). In particular, for high SNR, the weighting of the ℓ_2 -error by $1/(\lambda E[\|\hat{e}\|_2^2])$ may be suboptimal. Another reason may be the selected tuning parameter of $\lambda = 1$ (5.5.16), which may be not appropriate for the CIR case. Nevertheless, the explanation of this behavior especially regarding the different effects on CIR and CAC remains an open issue and has to be studied further.

In Figure 5.16, we evaluate the reconstruction ability over different s -sparse channels for a physical resource blocksize $N_B = 8$ at an SNR of 10 dB. As discussed in Section 5.6 for Figure 5.12, the distinction of 1-sparse channels suffers under noisy measurements due to the low degree of freedom. Consequently, the resource allocation estimation achieves the best performance for 2-sparse CIRs. For sparsity level equal to the blocksize $N_B = 8$, the failure rate rapidly increases. In Figure 5.16, we exemplarily plotted also the recovery performance including the inactive resource. There we see, that in certain cases the inactive resource will not be aggregated to a single 'inactive' transmitter. Thus, the inactive resources are indicating multiple transmitters. Nevertheless, the deviation is not severe. Furthermore, as discussed in Section 5.7, in many scenarios we are more interested

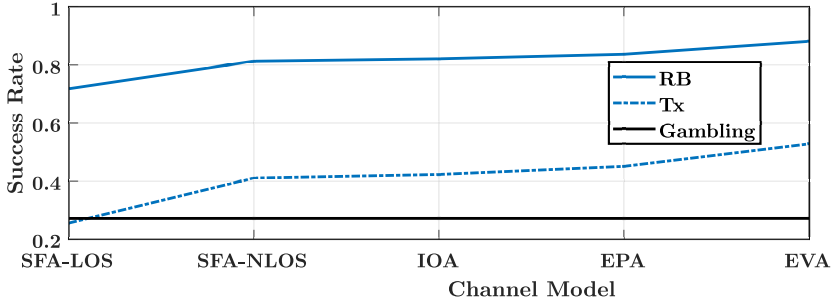


Figure 5.18: Success rate for the different performance channels (see Sections 2.3 and 4.6).

in the active resources. In such scenarios determining the resources of active transmitters suffice.

In addition to the sparsity point of view, we consider the impact of the physical resource blocksize N_B on 8-sparse CIRs as well as EPA performance channels in Figure 5.17. At first note, that obviously for increasing blocksize gambling is much easier, since guessing one physical resource block correctly results in a larger number of correct resources. Similar to Figure 5.16 for s -sparse channels, a blocksize of twice the number of s achieves good estimation performance. For a larger blocksize, the failure rate is below the detection limit. In contrast to that, the failure rate for distinguishing EPA CIRs decreases slowly but steadily for larger blocksizes. However, in comparison to the 8-sparse channel with 10 dB SNR, the EPA CIRs are distinguished at an SNR of 0 dB. Nevertheless, for both channel types an increasing blocksize significantly improves the identification of the resource allocation map.

Finally, in Figure 5.18, we perform simulations for the different performance channels from 3GPP and ITU (see Section 2.3) as well as the short-range factory automation (SFA) channels from our measurements (see Section 4.6.1). The performance channel models are sorted by the excess delay, see Table 4.5. Interestingly, the estimation performance increases for larger excess delay. In particular, Algorithm 5.1 performs visibly worse for an SFA channel with LOS characteristics than for an SFA channel with NLOS characteristics. The arguments are similar as stated before. For larger excess delay the sparse entries are more distinguishable, leading to more complex fingerprints of the CIRs. Thus, the CIRs can be easier distinguished from each other. Consequently, the estimation performance of Algorithm 5.1 can benefit from larger excess delay of compressible wireless channels.

5.9 Summary

In this Chapter, we introduced a framework to identify non-adjacent resources belonging to the same transmitter by frequency-division multiple access signal observation from multiple transmitters. In particular, we perform allocation map retrieval without prior knowledge of the actual overall resource allocation map. We analyzed two scenarios, (i) channel transfer function measurements observing pilots, and (ii) spectral channel power measurements observing constant amplitude signals. In general, allocation map retrieval is a combinatorial problem, which addresses activity detection as well as identifying all resources of the same transmitter. Therefore, we determined the combinatorial problem size of allocation map retrieval given by the Bell number in general and by the Stirling number of the second kind for a distinct number of transmitters. Furthermore, we elaborated the general problem of allocation map retrieval in more detail. We compared our problem with unlabeled sensing and pointed out similarities. Subsequently, we derived an objective function based on the ℓ_1 -norm to perform AMR. Furthermore, to enable the use of orthogonal matching pursuit (OMP), we proposed a joint ℓ_0 -pseudonorm ℓ_2 -error measure. To verify the objective function, we performed combinatorial simulations based on basis pursuit denoising (BPDN) and OMP, where the joint measure of OMP achieved superior performance over BPDN. Based on the objective function, we developed a feasible but suboptimal algorithm based on OMP to identify non-adjacent resources with certain structure. The performance is shown in simulations using s -sparse channels as well as practical performance channels from 3GPP, ITU, and our measurement campaign presented in Section 4.6.2. The performance results showed, that the proposed algorithm is able to identify the non-adjacent resources belonging to the same transmitter, if the channel representation can be reconstructed for the given physical resource block. Furthermore, we would like to emphasize that finding the correct resource allocation also yields the channel impulse response or the circular autocorrelation of the channel impulse response.

Chapter 6

Summary

In recent years the focus of wireless [information and communications technologies](#) laid in mobile broadband, which majorly addressed human-centric communication for smartphones and computers. In future, wireless communications will be strongly diversified. Besides human-centric communications also machine-to-machine communications with manifold requirements shall be supported. There, sporadic access as well as device-to-device communications play an important role. These new features demand for enhanced and new dynamic spectrum access schemes, which enable flexible use of spectrum resources as well as reduce signaling overhead in communications. In this thesis, we proposed two dynamic spectrum access schemes: gray space detection, which can be applied for cognitive radio system, and allocation map retrieval, which can be used for reduction of signaling overhead in device-to-device communications. In addition, we conducted a measurement campaign to measure practical wireless channels, which have been used to evaluate the dynamic spectrum access schemes on a practical basis.

In the first two chapters, Chapters 2 and 3, we introduced the foundations of wireless communications, dynamic spectrum access and compressed sensing. We introduced a multipath channel model based on the ITU and the 3GPP performance channels. In general, wireless channels exhibit sparse features and can be regarded as compressible signals, which is one key aspect exploited by our dynamic spectrum access schemes. As our basic wireless system model we considered a multi-carrier system like cyclic-prefix based OFDM communication system similar to the system used in 3GPP LTE and 3GPP NR. The considered channel estimation algorithms encompassed the reduced-rank least squares approach, a classical ℓ_2 -norm approach for general problems of linear systems without further structure like sparsity, and

compressed sensing methods, which exploit the sparsity of a signal. There, basis pursuit denoising represents a convex method for sparse signals, while orthogonal matching pursuit is just a greedy algorithm for sparse signals performing iterative least squares estimations.

One important aspect addressed by Chapter 2 is the statistical hypothesis tests. An optimal statistical hypothesis test is given by a uniformly most powerful test, which does not exist in general. Classically, the Neyman-Pearson lemma is exploited to obtain a uniformly most powerful test for simple hypotheses. However here, we elaborated a uniformly most powerful test based on the Karlin-Rubin theorem, which can be applied for signal detection where a one-sided composite hypothesis is present. In particular, we proposed a one-sided composite log-likelihood ratio test, where the false alarm rate (or alternatively the detection rate) can be controlled. Furthermore, we derived a practical approximated hypothesis test based on the empirical cumulative distribution function by exploiting the Glivenko–Cantelli theorem. Consequently, we were able to perform signal detection with the use of empirical channel statistics.

A further important aspect is addressed by Chapter 3, where bounds on the ℓ_2 -error for partial Fourier measurements are derived. There, we considered two observations of the received signal, measurements of the channel transfer function and the spectral channel power. While classic pilot-based channel estimation retrieves the channel impulse response from complex channel transfer function measurements, the circular autocorrelation of the channel impulse response is retrieved via amplitude-based channel estimation from real-valued spectral channel power measurements. Based on probability distributions we derived the error bounds for successful signal recovery for both observation types. In particular, for channel transfer function measurements we stated the error bound depending on an approximated Gamma distribution. In the case of spectral channel power measurements, we derived an estimate of the error bound depending on the expectation of the measurement error.

In Chapter 4 we proposed our gray space detection scheme for interweave cognitive radio system, which aims to find temporary small fractions of inactive resources within an already occupied primary user spectrum band. There, gray space detection copes with frequency fading holes and distinguishes them from inactive resources. In contrast to the classical energy detection, gray space detection can be configured to protect the primary user system from interference of a cognitive radio system with a desired probability. The gray space detection procedure comprises of three steps to acquire the unused resources: step 1 applies energy detection to estimate a preliminary primary user resource allocation, step 2 exploits the preliminary

allocation to estimate the spectral channel power from the primary user via amplitude-based channel estimation, and step 3 applies the log-likelihood ratio test based on the estimated spectral channel power to acquire the final primary user resource allocation. The complement set of the final allocation represents the desired set of gray spaces.

For amplitude-based channel estimation in step 2 three approaches have been applied: compressed sensing estimators like greedy orthogonal matching pursuit (OMP) and basis pursuit denoising (BPDN) as well as the traditional reduced-rank least squares (RRLS) estimator. Since primary user systems like 3GPP LTE target an error rate of 10%, the interesting range of false alarm rate lies below 10%. There, gray space detection based on the compressed sensing methods basis pursuit denoising and orthogonal matching pursuit achieved in general better performance than energy detection, especially in the interesting region below 10% false alarm rate. In addition, besides the impractical uniformly most powerful hypothesis test with complete knowledge of the cumulative distribution function of the primary user channel, we proposed practical thresholds for the log-likelihood ratio test in step 3 to achieve a desired false alarm rate. In particular, we derived the gray space test based on the empirical cumulative distribution function of the primary user channel. Furthermore, we provided the individual energy detector, which just utilizes the instantaneous spectral channel power, but no channel statistics. The simulations showed that the gray space test achieves a significant performance gain over individual energy detection. Moreover in comparison to the uniformly most powerful test, the gray space test only visibly suffers at large false alarm rates above 10% (low primary user protection).

To evaluate gray space detection also for practical wireless channels, we conducted a measurement campaign presented in Chapter 4, where wireless short-range factory automation channels of representative manufacturing process have been measured. The measured channels have been used to derive short-range factory automation channel models, such that the performance of dynamic spectrum access schemes can be evaluated and compared for practical channels.

Chapter 5 extended gray space detection of Chapter 4 to multiple transmitters, where the superposition of multiple signals in form of an FDMA signal are observed by a receiver. Thus, while in Chapter 4 only two hypotheses were present, describing active and inactive resources, Chapter 5 dealt with more than two hypotheses, each describing a different transmitter. Moreover, in contrast to gray space detection, in Chapter 5 we were particularly interested in retrieving the distinct resource allocation map of each transmitter. Therefore, we introduced the framework of allocation map

retrieval, which shall identify non-adjacent resources belonging to the same transmitter. While for fixed number of transmitters the combinatorial problem size is given by the Stirling number of the second kind, for unknown number of active transmitters the receiver also has to perform activity detection, and therefore, the problem size increases to the Bell number. We analyzed the allocation map retrieval problem for the two observation types introduced in Chapter 3, the channel transfer function measurements with embedded pilot tones, and the spectral channel power measurements with constant amplitude signals (like in Chapter 4). We showed that the considered allocation map retrieval problem for the case of a single transmitter is a subproblem of unlabeled sensing, where in general a permutation matrix and a unknown vector are jointly recovered, and therefore, given sufficient i.i.d. measurements, a unique solution exists for the single transmitter allocation map retrieval problem in the noiseless case with probability one.

For both observation types we proposed an objective function solving the combinatorial problem of allocation map retrieval by exploiting the sparse properties of the wireless channel. Informally, the objective function selects the pair of resource allocations and channel representations of the demixed receive signal with the smallest ℓ_1 -norm. To verify the objective function, we performed brute-force simulations based on BPDN and OMP, where the joint ℓ_0 -pseudonorm ℓ_2 -error measure of OMP achieved superior performance over BPDN.

In addition, we developed an efficient but suboptimal algorithm based on the OMP approach to perform AMR with the use of physical resource blocks. The simulation results showed, that the proposed algorithm is able to identify the non-adjacent resources belonging to the same transmitter, if the channel representation can be reconstructed for the given physical resource blocksize. Note that finding the correct resource allocation generally also yields in retrieving the channel impulse response or the spectral channel power depending on the used transmit alphabet.

Open Questions and Future Work

- In Chapter 3 we derived a bound on the ℓ_2 -error, such that the circular autocorrelation of the channel impulse response can be estimated from spectral channel power measurements. Unfortunately, the bound is not very precise since the underlying distribution is represented by the ℓ_4 -norm of complex normal random variable. The bound may be refined using [concentration inequalities](#) [CL06].
- Gray space detection is also faced with a joint estimation problem, where in the course of recovering the resource allocation of the pri-

mary user the spectral channel power is estimated as well. Currently, the spectral channel power is estimated based on a preliminary set of resources, where the preliminary set is obtained via energy detection. The turbo principle like in [HB03; IKW+09] may increase the detection performance, if the final estimate of the primary user resources allocation of step 3 is fed back to the spectral channel power estimator.

- Currently, the objective function for the allocation map retrieval problem in Chapter 5 is obtained mainly by heuristic considerations and evaluated via simulations. The mathematical proof, whether the ℓ_1 -norm can identify the correct allocation among all possible allocation, is still open.
- In Chapter 5 we investigated the allocation map retrieval problem for different physical resource block sizes. However, there are some further questions regarding the feasibility of the problem. First, let us assume that the channel representations are known in advance. Are there any restrictions on the physical resource block size, such that the reconstruction of the resource allocation is successful? Furthermore, let us assume that each transmitter selected sufficient resources to reconstruct the channel representation. Are there restrictions on the physical resource block size as well?
- In Chapter 5 an allocation map retrieval algorithm is proposed, which iteratively reconstructs the resource allocation of all transmitters by breadth-first search. The development of alternative algorithms based on belief propagation or alternating minimization may result in better estimation performance.

Appendix A

Probability Distributions

[Probability distributions](#) of a random variable X are used in the complete field of wireless communications. In particular the tuning of algorithms to obtain the correct solution requires good knowledge about the underlying distribution of variables (such as noise) and its derivatives. The two volumes of Johnson in 1994 [JKB94] and 1995 [JKB95] give detailed insights to the various probability distributions. In addition, the collection of Walck [Wal07] provides a quick and comprehensive compilation of several probability distributions. We will focus here on [continuous distributions](#), which can be described by probability density functions (PDFs).

A function $f_X(a)$ is called the PDF of a real-valued random variable X , if the following property is satisfied [Pap91, Chapter 2-2]

$$\int_{-\infty}^{\infty} f_X(a) \, da = 1. \quad (\text{A.0.1})$$

The cumulative distribution function (CDF) $F_X(a)$ describes the probability that $X < b$ and is obtained from the integral of the PDF

$$P(X < b) = F_X(b) = \int_{-\infty}^b f_X(a) \, da. \quad (\text{A.0.2})$$

Note, that $P(X < b) = P(X \leq b)$, since $P(X = b) = 0$ for continuous distributions. The probability that X lies within the interval of b and c is given by

$$P(b < X < c) = F_X(c) - F_X(b) = \int_b^c f_X(a) \, da. \quad (\text{A.0.3})$$

Since the CDF (A.0.2) is a monotone function, the inverse function of the CDF exists

$$F_X^{-1}(P) = \max(c \in \mathbb{R} : P_d \geq F_X(c)). \quad (\text{A.0.4})$$

The **inverse cumulative distribution function (ICDF)** describes the inverse of the CDF, such that for a given probability the threshold of the CDF is returned. The ICDF is also often referred to as percent-point function or quantile function.

Besides the PDF or CDF, probability distributions are further described by its properties. In general, central and raw **moments** of certain order are used to describe the course of the probability distributions. The most common properties are the first raw moment, mostly referred to as **mean** μ_X or expectation $E[X]$, and the second central moment, mostly referred to as **variance** σ_X^2 , which are defined as [Pap91, Chapter 5-3]

$$\begin{aligned} \mu_X &= E[X] &= \int_{-\infty}^{\infty} a \cdot f_X(a) \, da \\ &= \int_0^{\infty} (1 - F_X(b)) \, db - \int_{-\infty}^0 F_X(b) \, db, \end{aligned} \quad (\text{A.0.5})$$

$$\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (a - \mu_X)^2 \cdot f_X(a) \, da. \quad (\text{A.0.6})$$

Given a random variable X and a function $g(X)$ the expectation can be calculated by the integral of the function $g(a)$ and the PDF with $f_X(a)$ [Pap91, Chapter 5-3]

$$E[g(X)] = \int_{-\infty}^{\infty} g(a) \cdot f_X(a) \, da. \quad (\text{A.0.7})$$

Thus, following (A.0.7) the raw **moment** of order N is described by [Wal07, Chapter 2.2] and [Pap91, Chapter 5-3]

$$\mathbb{E}[X^N] = \int_{-\infty}^{\infty} a^N \cdot f_X(a) \, da \quad (\text{A.0.8})$$

For continuous distributions with support on the positive infinite interval $[0, \infty]$ (or alternatively $\mathbb{P}(X < 0) = 0$), the N -th raw moment is calculated with

$$\mathbb{E}[X^N] = \int_0^{\infty} a^N \cdot f_X(a) \, da \quad (\text{A.0.9})$$

$$= N \cdot \int_0^{\infty} b^{N-1} \cdot (1 - F_X(b)) \, db. \quad (\text{A.0.10})$$

A.1 Normal Distribution

One of the most commonly known distribution is the *normal distributed*, which is defined on the complete real line $(-\infty, \infty)$. In general, the normal distribution is present in diverse scenarios. However, in the domain of signal processing noise is often described via the normal distribution.

The *standard normal distribution* describes a normalized version of the normal distribution. The PDF and CDF of a *standard normal distributed* random variable $X \sim f_{\mathcal{N}}(\cdot; 0, 1)$ is given in (A.1.1) and (A.1.2), respectively [JKB94; Wal07]

$$f_{\mathcal{N},X}(a) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}a^2\right), \quad (\text{A.1.1})$$

$$\begin{aligned} F_{\mathcal{N},X}(b) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b \exp\left(-\frac{1}{2}a^2\right) \, da \\ &= \frac{1}{2} + \frac{1}{2} \cdot \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right). \end{aligned} \quad (\text{A.1.2})$$

The error function is defined by

$$\operatorname{erf}(c) = \frac{2}{\sqrt{\pi}} \int_0^c \exp(-a^2) \, da. \quad (\text{A.1.3})$$

$F_{\mathcal{N},X}(b)$	50%	84.13%	90%	95%	97.72%	99%	99.87%
b	0	1	1.28	1.64	2	2.33	3

Table A.1: Example values b for different $P(X < b)$.

Consequently, the ICDF of the standard normal distribution is defined by

$$F_{\mathcal{N},X}^{-1}(P) = \sqrt{2} \cdot \operatorname{erf}^{-1}(2P - 1). \quad (\text{A.1.4})$$

Some example probabilities of the standard normal distribution regarding b are given in Table A.1, which are obtained by (A.1.2) or by (A.1.4).

The *normal distributed* random variable $Y \sim f_{\mathcal{N}}(\cdot; \mu_Y, \sigma_Y^2)$ with mean μ_Y and variance σ_Y^2 represents the general version of the *standard normal distribution*. By substitution of $a' = \frac{a - \mu_Y}{\sigma_Y}$ in (A.1.1) the PDF of the normal distribution is derived by

$$\begin{aligned}
 f_{\mathcal{N},Y}(a; \mu_Y, \sigma_Y^2) \, da &= f_{\mathcal{N},X}(a') \, da' \\
 &= f_{\mathcal{N},X}\left(\frac{a - \mu_Y}{\sigma_Y}\right) \frac{da'}{da} \, da \\
 &= f_{\mathcal{N},X}\left(\frac{a - \mu_Y}{\sigma_Y}\right) \frac{d\left(\frac{a - \mu_Y}{\sigma_Y}\right)}{da} \, da \\
 &= f_{\mathcal{N},X}\left(\frac{a - \mu_Y}{\sigma_Y}\right) \cdot \frac{1}{\sigma_Y} \, da \\
 &= \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \left(\frac{a - \mu_Y}{\sigma_Y}\right)^2\right) \cdot \frac{1}{\sigma_Y} \, da. \quad (\text{A.1.5})
 \end{aligned}$$

Thus, the PDF of the normal distribution is

$$f_{\mathcal{N},Y}(a; \mu_Y, \sigma_Y^2) = \frac{1}{\sqrt{2\pi} \cdot \sigma_Y} \cdot \exp\left(-\frac{1}{2} \left(\frac{a - \mu_Y}{\sigma_Y}\right)^2\right). \quad (\text{A.1.6})$$

The corresponding CDF is obtained by substitution of $b = \frac{b' - \mu_Y}{\sigma_Y}$ in (A.1.2)

$$F_{\mathcal{N},Y}(b; \mu_Y, \sigma_Y^2) = F_{\mathcal{N},X}(b')$$

$$\begin{aligned}
&= F_{\mathcal{N},X} \left(\frac{b - \mu_Y}{\sigma_Y} \right) \\
&= \frac{1}{2} + \frac{1}{2} \cdot \operatorname{erf} \left(\frac{b - \mu_Y}{\sqrt{2} \cdot \sigma_Y} \right). \tag{A.1.7}
\end{aligned}$$

Consequently, the **ICDF of the normal distribution** (a generalization of (A.1.4)) is defined by

$$F_{\mathcal{N},Y}^{-1}(P; \mu_Y, \sigma_Y^2) = \mu_Y + \sqrt{2} \cdot \sigma_Y \cdot \operatorname{erf}^{-1}(2P - 1). \tag{A.1.8}$$

The example probabilities given in Table A.1 can also be obtained by appropriate normalization of b in (A.1.7) with $\frac{b - \mu_Y}{\sigma_Y}$.

A.2 Complex Normal Distribution

The **complex normal distribution** is obtained by addition of two normal distributed random variable $X = \Re\{X\} + i \cdot \Im\{X\}$. Therefore, a complex normal distribution can also be seen as a joint distribution of two real random variables. Both, the real and imaginary part of the random variable follow the normal distribution with

$$\begin{aligned}
\Re\{X\} &\sim f_{\mathcal{N}} \left(\cdot; \mu_{\Re\{X\}}, \sigma_{\Re\{X\}}^2 \right) \\
\Im\{X\} &\sim f_{\mathcal{N}} \left(\cdot; \mu_{\Im\{X\}}, \sigma_{\Im\{X\}}^2 \right).
\end{aligned}$$

The general complex normal distribution is fully described by three parameters: (i) the mean, (ii) the covariance matrix and (iii) the complementary covariance (or pseudo-covariance) matrix. Here we will consider the special case of proper complex normal distribution [SS14; NM93], where the complementary covariance matrix is zero. Furthermore, we assume a diagonal covariance matrix, where the values of the main diagonal are equal. Therefore, we have independently distributed $\Re\{X\}$ and $\Im\{X\}$ with identical variance. Thus, this proper complex normal distribution is fully described by two scalar parameters: (i) the mean $\mu_X \in \mathbb{C}$ and (ii) the variance $\sigma_X^2 \in \mathbb{R}_{\geq 0}$, which are calculated by

$$\mu_X = \mu_{\Re\{X\}} + i \cdot \mu_{\Im\{X\}}, \tag{A.2.1}$$

$$\sigma_X^2 = \sigma_{\Re\{X\}}^2 + \sigma_{\Im\{X\}}^2 = 2\sigma_{\Re\{X\}}^2 = 2\sigma_{\Im\{X\}}^2. \tag{A.2.2}$$

Note that $\sigma_{\Re\{X\}}^2 = \sigma_{\Im\{X\}}^2 = \frac{1}{2}\sigma_X^2$. Consequently, the proper complex normal distribution with identical variance is

$$X \sim f_{\mathbb{C}\mathcal{N}}(\cdot; \mu_X, \sigma_X^2) \tag{A.2.3}$$

and $\Re\{X\}$ and $\Im\{X\}$ are distributed by

$$\begin{aligned}\Re\{X\} &\sim f_{\mathcal{N}}\left(\cdot; \Re\{\mu_X\}, \frac{1}{2}\sigma_X^2\right), \\ \Im\{X\} &\sim f_{\mathcal{N}}\left(\cdot; \Im\{\mu_X\}, \frac{1}{2}\sigma_X^2\right).\end{aligned}\tag{A.2.4}$$

A.3 Circularly-Symmetric Complex Normal Distribution

The *circularly-symmetric complex normal distribution* is a special case of the proper complex normal distribution. There, a complex normal random variable $X = \Re\{X\} + i \cdot \Im\{X\}$ is circularly-symmetric complex normal distributed, if and only if $\Re\{X\}$ and $\Im\{X\}$ are statistically independent and identically distributed [Gal08] with

$$\Re\{X\}, \Im\{X\} \sim f_{\mathcal{N}}\left(\cdot; 0, \sigma_{\Re\{X\}}^2\right).$$

Thus, the variance σ_X^2 of a circularly-symmetric complex normal distributed variable

$$X \sim f_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_X^2)\tag{A.3.1}$$

is calculated in the same way as in (A.2.2) by [MC04]

$$\sigma_X^2 = \sigma_{\Re\{X\}}^2 + \sigma_{\Im\{X\}}^2 = 2\sigma_{\Re\{X\}}^2 = 2\sigma_{\Im\{X\}}^2,\tag{A.3.2}$$

where the distribution of real and imaginary part follows

$$\Re\{X\}, \Im\{X\} \sim f_{\mathcal{N}}\left(\cdot; 0, \frac{1}{2}\sigma_X^2\right).\tag{A.3.3}$$

E.g. the real and imaginary parts of a circularly-symmetric complex normal random variable $Y \sim f_{\mathbb{C}\mathcal{N}}(\cdot; 0, 1)$ are distributed with $\Re\{Y\}, \Im\{Y\} \sim f_{\mathcal{N}}(\cdot; 0, \frac{1}{2})$.

A.4 Chi Distribution

The *chi distribution* derives from standard normal distributed random values, such that $X = \sqrt{|X_1|^2 + \dots + |X_{N_f}|^2}$ with degree of freedom $N_f \in \mathbb{R}_{\geq 0}$, for further reference see [JKB94, Chapter 18] and [Wal07, Chapter 8.14]. While the random variable X is chi distributed $X \sim f_{\chi}(\cdot; N_f)$, the

$F_{\chi,X}(b; 1)$	50%	68.27%	90%	95%	95.45%	99%	99.73%
b	0.67	1	1.64	1.96	2	2.58	3

Table A.2: Example values b for different $P(-b < X < b)$, which represents the coverage probability of the standard normal distribution.

underlying random variables are $X_n \sim f_{\mathcal{N}}(\cdot; 0, 1)$. The PDF and CDF of the chi distribution on the supported semi-infinite interval $(0, \infty)$ are given in (A.4.1) and (A.4.2), respectively

$$f_{\chi,X}(a; N_f) = \frac{a^{N_f-1} \cdot \exp\left(-\frac{a^2}{2}\right)}{2^{\left(\frac{N_f}{2}-1\right)} \cdot \Gamma\left(\frac{N_f}{2}\right)} \quad (\text{A.4.1})$$

$$F_{\chi,X}(b; N_f) = \frac{\gamma\left(\frac{N_f}{2}, \frac{b^2}{2}\right)}{\Gamma\left(\frac{N_f}{2}\right)}, \quad (\text{A.4.2})$$

For degree of freedom $N_f = 1$ the CDF of the chi distribution simplifies to

$$F_{\chi,X}(b; 1) = \text{erf}\left(\frac{b}{\sqrt{2}}\right). \quad (\text{A.4.3})$$

From (A.4.3) the ICDF is obtained by

$$F_{\chi,X}^{-1}(P; 1) = \sqrt{2} \cdot \text{erf}^{-1}(P). \quad (\text{A.4.4})$$

In the context of standard normal distribution, the CDF of the chi distribution with $N_f = 1$ (A.4.3) states the **coverage probability** of standard normal distribution

$$\begin{aligned} F_{\chi,X}(b; 1) &= P(-b < X_1 < b) \\ &= F_{\mathcal{N},X_1}(b; 0, 1) - F_{\mathcal{N},X_1}(-b; 0, 1). \end{aligned} \quad (\text{A.4.5})$$

Some example probabilities of the chi distribution with $N_f = 1$ regarding b are given in Table A.2, which are obtained by (A.4.3) or by (A.4.4).

A.5 Rayleigh Distribution

The *Rayleigh distributed* is commonly used in communications for the propagation model called *Rayleigh fading*, see Section 2.3. In contrast to the previ-

$F_{\text{Rayleigh}, Y}\left(b; \frac{1}{\sqrt{2}}\sigma_X\right)$	63.21%	90%	95%	98.17%	99%
b	1	1.52	1.73	2	2.15

Table A.3: Example values b for different $P(Y < b \cdot \sigma_X)$.

ously characterized chi distribution in Appendix A.4, the Rayleigh describes only 2 degrees of freedom, but with arbitrary variance $\sigma_X^2 \in \mathbb{R}_{>0}$. If we have a circularly-symmetric complex normal random variable $X \sim f_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_X^2)$, then the derived absolute value is *Rayleigh distributed*

$$Y = \|X\|_2 = \sqrt{\Re\{X\}^2 + \Im\{X\}^2} \sim f_{\text{Rayleigh}}\left(\cdot; \frac{1}{\sqrt{2}}\sigma_X\right), \quad (\text{A.5.1})$$

where the mode $\frac{1}{\sqrt{2}}\sigma_X$ describes the Rayleigh distribution. The scaling of the mode results from the underlying circularly-symmetric complex normal distribution $f_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_X^2)$, see (A.3.3). Therefore, its PDF and CDF on the supported semi-infinite interval $(0, \infty)$ [JKB94; Wal07] is given by (A.5.2) and (A.5.3), respectively

$$f_{\text{Rayleigh}, Y}\left(a; \frac{1}{\sqrt{2}}\sigma_X\right) = \frac{2a}{\sigma_X^2} \cdot \exp\left(-\frac{a^2}{\sigma_X^2}\right) \quad (\text{A.5.2})$$

$$F_{\text{Rayleigh}, Y}\left(b; \frac{1}{\sqrt{2}}\sigma_X\right) = 1 - \exp\left(-\frac{b^2}{\sigma_X^2}\right), \quad (\text{A.5.3})$$

Setting $\sigma_X = \sqrt{2}$, the Rayleigh distribution equals the chi distribution with two degrees of freedom. From (A.5.3) the ICDF is obtained by

$$F_{\text{Rayleigh}, Y}^{-1}\left(P; \frac{1}{\sqrt{2}}\sigma_X\right) = \sigma_X \cdot \sqrt{-\ln(1 - P)}. \quad (\text{A.5.4})$$

Some example probabilities of the Rayleigh distribution regarding b are given in Table A.3, which are obtained for $\sigma_X^2 = 1$ by (A.5.3) or by (A.5.4). The same results are obtained by scaling b with σ_X^2 instead of setting $\sigma_X^2 = 1$.

A.6 Chi-Squared Distribution

The *chi-squared distribution* is supported on the semi-infinite interval $(0, \infty)$ and occurs frequently as the resulting distribution for squared ℓ_2 -norms of

standard normal random variables. The chi-squared distribution is strongly related to the chi distribution (see Appendix A.4), such that $\sqrt{X} \sim f_\chi(\cdot)$ for a chi-squared random variable X .

The *chi-squared distribution* with N_f degree of freedom as defined in [JKB94; Wal07] describes the distribution of a random variable $X = \sum_{n=1}^{N_f} (X_n)^2 \sim f_{\chi^2}(\cdot; N_f)$, where $X_n \sim f_{\mathcal{N}}(\cdot; 0, 1)$. The PDF and CDF of the chi-squared distribution on $(0, \infty)$ are given by

$$f_{\chi^2, X}(a; N_f) = \frac{\left(\frac{a}{2}\right)^{\left(\frac{N_f}{2}-1\right)} \cdot \exp\left(-\frac{a}{2}\right)}{2 \cdot \Gamma\left(\frac{N_f}{2}\right)} \quad (\text{A.6.1})$$

$$F_{\chi^2, X}(b; N_f) = \frac{\gamma\left(\frac{N_f}{2}, \frac{b}{2}\right)}{\Gamma\left(\frac{N_f}{2}\right)} \quad (\text{A.6.2})$$

Furthermore, for parameter $N_f/2 \in \mathbb{N}$ (present in sum of circularly-symmetric complex normal random variables) the CDF of the chi-squared distribution (A.6.2) in the interval $(0, \infty)$ can be simplified to [Wal07, Chapter 42.5.3]

$$F_{\chi^2, X}(b; N_f) = 1 - \exp\left(\frac{b}{2}\right) \sum_{n=0}^{\frac{N_f}{2}-1} \frac{b^n}{2^n \cdot n!}, \quad N_f/2 \in \mathbb{N}. \quad (\text{A.6.3})$$

The Gamma function $\Gamma(\cdot)$ and the lower incomplete Gamma function $\gamma(\cdot, \cdot)$ are defined by

$$\Gamma(b) = \int_0^\infty a^{b-1} \exp(-a) \, da, \quad (\text{A.6.4})$$

$$\gamma(b, c) = \int_0^c a^{b-1} \exp(-a) \, da. \quad (\text{A.6.5})$$

Since the argument of the Gamma function in (A.6.1) and (A.6.2) can only be a positive natural number divided by two, the calculation of the Gamma function (A.6.4) with $N \in \mathbb{N}$ can be simplified to a factorial expression

$$\Gamma(N+1) = N! , \quad (\text{A.6.6})$$

$$\Gamma\left(N + \frac{1}{2}\right) = \sqrt{\pi} \cdot \frac{(2N)!}{N! \cdot 4^N}. \quad (\text{A.6.7})$$

A.7 Gamma Distribution

The *Gamma distribution* is supported on the semi-infinite interval $(0, \infty)$ and occurs often as the resulting distribution for ℓ_2 -norms of (complex) normal random variables (e.g. the ℓ_2 -norm of noise). The chi-squared distribution of the previous Appendix A.6 is a special case of the Gamma distribution.

In the literature the *Gamma distribution* is described with different parameterizations. We focus here on the Gamma distribution with two-parameters as defined in [JKB94; Wal07]. A random variable X following the Gamma distribution is characterized by the shape parameter c and scale parameter d . Note, that [Wal07] uses the inverse scale parameter to describe the Gamma distribution. The PDF and CDF for $a, b > 0$ and $c, d > 0$ are given by

$$f_{\Gamma, X}(a; c, d) = \frac{a^{d-1} \exp(-a/c)}{c^d \Gamma(d)}, \quad (\text{A.7.1})$$

$$F_{\Gamma, X}(b; c, d) = \frac{\gamma(d, b/c)}{\Gamma(d)}. \quad (\text{A.7.2})$$

For the parameters $c = 2$ and $d = N_f/2$ the Gamma distribution represents the chi-squared distribution. However, the chi-squared distribution only supports the underlying *standard normal distribution* $f_{\mathcal{N}}(\cdot; 0, 1)$. Thus, an underlying *normal distributed* random variable $X_n \sim f_{\mathcal{N}}(\cdot; 0, \sigma_{X_1}^2)$ is only supported indirectly by a scaled chi-squared distributed random variable

$$X/\sigma_{X_1}^2 = \sum_{n=1}^{N_f} (X_n/\sigma_{X_1})^2 \sim f_{\chi^2}(\cdot; N_f).$$

In contrast to that, the Gamma distribution $X \sim f_{\Gamma}(\cdot; 2\sigma_{X_1}^2, N_f/2)$ directly supports the underlying normal distribution with variance $\sigma_{X_1}^2$ and zero mean. Considering circularly-symmetric complex normal random vari-

ables $Y_n \sim \text{f}_{\mathbb{CN}}(\cdot; 0, \sigma_{Y_1}^2)$, then

$$Y = \sum_{n=1}^{N_f/2} |Y_n|^2 \sim \text{f}_{\Gamma}(\cdot; \sigma_{Y_1}^2, N_f/2),$$

where $N_f/2$ is the number of complex normal random variables. For completeness, taking (A.3.2) with $\sigma_{Y_1}^2 = 2\sigma_{\Im\{Y_n\}}^2$, then we have $Y \sim \text{f}_{\Gamma}(\cdot; 2\sigma_{\Re\{Y_n\}}^2, N_f/2)$. Note, that if not stated otherwise we will only consider complex normal random variables. Thus, $N_f/2 \in \mathbb{N}$ is always fulfilled. Rewriting (A.7.1) and (A.7.2) for the complex normal random variables using the variance $\sigma_{Y_1}^2$ and $N_f/2$, we get

$$\text{f}_{\Gamma,Y}(a; \sigma_{Y_1}^2, N_f/2) = \exp\left(-\frac{a}{\sigma_{Y_1}^2}\right) \cdot \frac{a^{\frac{N_f}{2}-1}}{\sigma_{Y_1}^{N_f} \cdot \left(\frac{N_f}{2} - 1\right)!}, \quad (\text{A.7.3})$$

$$\begin{aligned} \text{F}_{\Gamma,Y}(b; \sigma_{Y_1}^2, N_f/2) &= \frac{\gamma\left(\frac{N_f}{2}, \frac{b}{\sigma_{Y_1}^2}\right)}{\Gamma\left(\frac{N_f}{2}\right)} \\ &= 1 - \exp(-b/\sigma_{Y_1}^2) \cdot \sum_{n=0}^{\frac{N_f}{2}-1} \frac{b^n}{\sigma_{Y_1}^{2n} n!}. \end{aligned} \quad (\text{A.7.4})$$

In relation to the chi-squared distribution, the Gamma distribution with regard to the variance $\sigma_{Y_1}^2$ can also be written

$$\text{f}_{\Gamma,Y}(a; \sigma_{Y_1}^2, N_f/2) = 2 \cdot \text{f}_{\chi^2,Y}\left(\frac{2a}{\sigma_{Y_1}^2}; N_f\right) / \sigma_{Y_1}^2, \quad (\text{A.7.5})$$

$$\text{F}_{\Gamma,Y}(b; \sigma_{Y_1}^2, N_f/2) = \text{F}_{\chi^2,Y}\left(\frac{2b}{\sigma_{Y_1}^2}; N_f\right). \quad (\text{A.7.6})$$

For large degree of freedom N_f the CDF of the Gamma distribution (A.7.4) can be approximated by the normal distribution

$$F_{\Gamma,Y}(b; \sigma_{Y_1}^2, N_f/2) \approx F_{\mathcal{N}}(b'),$$

$$\text{where } b' = \frac{2\sqrt{b}}{\sigma_{Y_1}} - \sqrt{2N_f - 1}, \quad (\text{A.7.7})$$

$$\text{or } b' = \sqrt{\frac{9N_f}{2}} \cdot \left(\sqrt[3]{\frac{2b}{N_f \sigma_{Y_1}^2}} + \frac{2}{9N_f} - 1 \right), \quad (\text{A.7.8})$$

which is derived similar to the approximation of the chi-squared distribution in [Wal07, Chapter 8.6]. While the simple approximation in (A.7.7) already achieves good results, (A.7.8) utilizing the cubic root is even more accurate.

Considering the above approximations, we can also derive an approximation for the ICDF of the Gamma distribution for large degree of freedom. Therefore, we plug the modified variable b' of (A.7.8) into the CDF of the standard normal distribution (A.1.2) and derive the inverse function. Thus, based on (A.7.8) we obtain an approximated ICDF of the Gamma distribution

$$F_{\Gamma,Y}^{-1}(P; \sigma_{Y_1}^2, N_f/2) \approx \frac{N_f}{2} \cdot \sigma_{Y_1}^2 \left(\sqrt{\frac{2}{9N_f}} F_{\mathcal{N}}^{-1}(P) + 1 - \frac{2}{9N_f} \right)^3. \quad (\text{A.7.9})$$

The raw [moment](#) of order N of the Gamma distribution is calculated by [Wal07, Chapter 17.3]

$$E[Y^N] = \sigma_{Y_1}^{2N} \cdot \frac{\Gamma\left(\frac{N_f}{2} + N\right)}{\Gamma\left(\frac{N_f}{2}\right)} = \sigma_{Y_1}^{2N} \cdot \frac{\left(\frac{N_f}{2} + N - 1\right)!}{\left(\frac{N_f}{2} - 1\right)!}. \quad (\text{A.7.10})$$

Note, that $\frac{N_f}{2} \in \mathbb{N}$. Equation (A.7.10) can be further simplified for the first raw moment. Thus, the mean of the Gamma distribution is given by

$$\mu_Y = E[Y] = N_f/2 \cdot \sigma_{Y_1}^2. \quad (\text{A.7.11})$$

Taking the approximation of (A.7.8) the probability at the point of the expectation (A.7.11) is simply written

$$F_{\Gamma,Y}(N_f/2 \cdot \sigma_{Y_1}^2; \sigma_{Y_1}^2, N_f/2) \approx F_{\mathcal{N}}\left(\sqrt{\frac{2}{9N_f}}\right). \quad (\text{A.7.12})$$

Furthermore, the variance of the Gamma distribution is obtained by

$$\begin{aligned}\sigma_Y^2 &= \mathbb{E}[(Y - \mathbb{E}[Y])^2] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &= \frac{N_f}{2} \left(\frac{N_f}{2} + 1 \right) \sigma_{Y_1}^4 - \left(\frac{N_f}{2} \right)^2 \sigma_{Y_1}^4 = \frac{N_f}{2} \cdot \sigma_{Y_1}^4.\end{aligned}\quad (\text{A.7.13})$$

Since the Gamma distribution tends to the normal distribution for large degree of freedom

$$\lim_{N_f \rightarrow \infty} \frac{2}{N_f} F_{\Gamma,Y}(b; \sigma_{Y_1}^2, N_f/2) = F_{\mathcal{N},Y}(b; \sigma_{Y_1}^2, \sigma_{Y_1}^4), \quad (\text{A.7.14})$$

we can approximate the Gamma distribution also by the standard normal distribution

$$\lim_{N_f \rightarrow \infty} \frac{F_{\Gamma,Y}(b; \sigma_{Y_1}^2, N_f/2) - N_f/2 \cdot \sigma_{Y_1}^2}{N_f/2 \cdot \sigma_{Y_1}^4} = F_{\mathcal{N},Y}(b). \quad (\text{A.7.15})$$

Thus, some approximated example probabilities of the normalized Gamma distribution in (A.7.15) for large degree of freedom N_f regarding b are given in Table A.1, which are obtained by (A.1.2).

A special case of the Gamma distribution is present for two random variables $X_1, X_2 \sim f_{\mathcal{N}}(\cdot; 0, \sigma_{X_1}^2)$, which can be rewritten as a circularly-symmetric complex normal random variable

$$Y_1 = \frac{X_1 + iX_2}{\sqrt{2}} \sim f_{\mathbb{C}\mathcal{N}}(\cdot; 0, \sigma_{Y_1}^2), \quad (\text{A.7.16})$$

which we have introduced in Appendix A.3. Thus, we have

$$\Re\{Y_1\}, \Im\{Y_1\} \sim f_{\mathcal{N}}\left(\cdot; 0, \frac{1}{2}\sigma_{Y_1}^2\right). \quad (\text{A.7.17})$$

Consequently, the *Gamma distribution* with $N_f = 2$ simply results from the ℓ_2 -norm with

$$Y = \|Y_1\|_2^2 \sim f_{\Gamma}(\cdot; \sigma_{Y_1}^2, 1). \quad (\text{A.7.18})$$

Note, that the special case with $N_f = 2$ represents the squared random variable of the Rayleigh distribution as discussed in Appendix A.5. The PDF and CDF from (A.7.3) and (A.7.4) simplifies to

$$f_{\Gamma,Y}(a; \sigma_{Y_1}^2, 1) = \frac{1}{\sigma_{Y_1}^2} \cdot \exp\left(-\frac{a}{\sigma_{Y_1}^2}\right), \quad (\text{A.7.19})$$

$F_{\Gamma,Y}(b; \sigma_{Y_1}^2, 1)$	50%	63.2%	90%	95%	98.2%	99%
b	$0.69\sigma_{Y_1}^2$	$\sigma_{Y_1}^2$	$2.3\sigma_{Y_1}^2$	$3\sigma_{Y_1}^2$	$4\sigma_{Y_1}^2$	$4.61\sigma_{Y_1}^2$

Table A.4: Example values b given $N_f = 2$ for different $P(Y < b)$.

$$F_{\Gamma,Y}(b; \sigma_{Y_1}^2, 1) = 1 - \exp\left(-\frac{b}{\sigma_{Y_1}^2}\right). \quad (\text{A.7.20})$$

Solving (A.7.20) for b results in

$$F_{\Gamma}^{-1}(P; \sigma_{Y_1}^2, 1) = -\sigma_{Y_1}^2 \cdot \ln(1 - P), \quad (\text{A.7.21})$$

representing the ICDF of the Gamma distribution with $N_f = 2$.

Some example probabilities of the Gamma distribution with $N_f = 2$ regarding b are given in Table A.4, which are obtained by (A.7.20) or by (A.7.21). The N -th order raw moment for two degrees of freedom simplifies from (A.7.10) to

$$E[Y^N] = N! \cdot \sigma_{Y_1}^{2N}. \quad (\text{A.7.22})$$

A.8 Non-Central Chi-Squared Distribution

The *non-central chi-squared distribution* $f_{\chi_{nc}^2}(\cdot)$ with N_f degrees of freedom and non-central parameter

$$\mu_{\chi_{nc}^2} = \sum_{n=1}^{N_f} (\mu_{X_n})^2 \quad (\text{A.8.1})$$

as defined in [Mui82] describes the distribution of a random variable

$$X = \sum_{n=1}^{N_f} (X_n)^2 \sim f_{\chi_{nc}^2}(\cdot; N_f, \mu_{\chi_{nc}^2}), \quad (\text{A.8.2})$$

where $X_n \sim f_{\mathcal{N}}(\cdot; \mu_{X_n}, 1)$. The PDF and CDF of the non-central chi-squared distribution on $(0, \infty)$ are given by

$$\begin{aligned}
f_{\chi_{nc}^2, X}(a; N_f, \mu_{\chi_{nc}^2}) &= \sum_{n=0}^{\infty} \left(\frac{\mu_{\chi_{nc}^2}}{2} \right)^n \cdot \frac{\exp\left(-\frac{\mu_{\chi_{nc}^2}}{2}\right)}{n!} \\
&\quad \cdot \underbrace{\left(\frac{a}{2} \right)^{\left(\frac{N_f}{2} + n - 1 \right)} \cdot \exp\left(-\frac{a}{2}\right)}_{f_{\chi^2, X}(a; N_f + 2n)} \\
&\quad \cdot \frac{2 \cdot \Gamma\left(\frac{N_f}{2} + n\right)}{f_{\chi^2, X}(a; N_f + 2n)} \\
&= \frac{a^{\frac{N_f}{2}-1} \exp\left(-\frac{1}{2}(\mu_{\chi_{nc}^2} + a)\right)}{2^{\frac{N_f}{2}} \sqrt{\pi}} \\
&\quad \cdot \sum_{n=0}^{\infty} \frac{(\mu_{\chi_{nc}^2} a)^n \Gamma\left(\frac{1}{2} + n\right)}{(2n)! \cdot \Gamma\left(\frac{N_f}{2} + n\right)}, \tag{A.8.3}
\end{aligned}$$

$$F_{\chi_{nc}^2, X}(b; N_f, \mu_{\chi_{nc}^2}) = \exp\left(-\frac{\mu_{\chi_{nc}^2}}{2}\right) \sum_{n=0}^{\infty} \frac{\left(\frac{\mu_{\chi_{nc}^2}}{2}\right)^n}{n!} \cdot \underbrace{\frac{\gamma\left(\frac{N_f}{2} + n, \frac{b}{2}\right)}{\Gamma\left(\frac{N_f}{2} + n\right)}}_{F_{\chi^2, X}(b; N_f + 2n)}, \tag{A.8.4}$$

where $f_{\chi^2, X}(a; N_f + 2n)$ (A.6.1) and $F_{\chi^2, X}(b; N_f + 2n)$ (A.6.2) represents the PDF and the CDF of the *chi-squared distribution* in Appendix A.6, respectively.

A.9 Scaled Non-Central Chi-Squared Distribution

Extending the non-central chi-squared distribution in the previous Appendix A.8, the *scaled non-central chi-squared distribution* operates on the normal distribution $X_n \sim f_{\mathcal{N}}(\cdot; \mu_{X_n}, \sigma_{X_n}^2)$ as defined in Appendix A.1. In particular, the *scaled non-central chi-squared distribution* describes the random variable

$$X = \sum_{n=1}^{N_f} |X_n|^2 \sim f_{\chi_{sn}^2}(\cdot; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) \tag{A.9.1}$$

with the non-central parameter

$$\mu_{\chi_{sn}^2} = \sum_{n=1}^{N_f} (\mu_{X_n})^2. \quad (\text{A.9.2})$$

By considering the scaling with appropriate substitution, the scaled non-central chi-squared distribution can be described via the non-central chi-squared distribution, see (A.8.3). Thus, for $a > 0$ we obtain the PDF

$$f_{\chi_{sn}^2, X}(a; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) = \frac{f_{\chi_{nc}^2, X}\left(\frac{a}{\sigma_{X_1}^2}; N_f, \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}\right)}{\sigma_{X_1}^2}. \quad (\text{A.9.3})$$

The corresponding CDF of the scaled non-central chi-squared distribution for $b > 0$ is also given via the non-central chi-squared distribution (see (A.8.4)) with

$$F_{\chi_{sn}^2, X}(b; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) = F_{\chi_{nc}^2, X}\left(\frac{b}{\sigma_{X_1}^2}; N_f, \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}\right). \quad (\text{A.9.4})$$

The full form of the PDF is written

$$\begin{aligned} f_{\chi_{sn}^2, X}(a; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) &= \sum_{n=0}^{\infty} \left(\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right)^n \cdot \frac{\exp\left(-\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right)}{2 \cdot n!} \\ &\quad \cdot f_{\Gamma, X}\left(\frac{a}{2}; \sigma_{X_1}^2, \frac{N_f}{2} + n\right) \\ &= \sum_{n=0}^{\infty} \left(\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right)^n \cdot \frac{\exp\left(-\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right)}{2 \cdot n!} \\ &\quad \cdot \frac{\left(\frac{a}{2\sigma_{X_1}^2}\right)^{\left(\frac{N_f}{2} + n - 1\right)} \cdot \exp\left(-\frac{a}{2\sigma_{X_1}^2}\right)}{\sigma_{X_1}^2 \cdot \Gamma\left(\frac{N_f}{2} + n\right)} \\ &= a^{\frac{N_f}{2}-1} \exp\left(-\frac{1}{2}\left(\sigma_{X_1} \cdot \mu_{\chi_{sn}^2} + \frac{a}{\sigma_{X_1}^2}\right)\right) \\ &\quad \cdot \sum_{n=0}^{\infty} \frac{\sigma_{X_1}^{-(n+N_f)} \cdot (\mu_{\chi_{sn}^2} \cdot a)^n}{2^{2n+\frac{N_f}{2}} \cdot n! \cdot \left(\frac{N_f}{2} + n - 1\right)!}, \quad (\text{A.9.5}) \end{aligned}$$

where $f_{\Gamma,X}\left(\frac{a}{2}; \sigma_{X_1}^2, \frac{N_f}{2} + n\right)$ represents the PDF from (A.7.1) of the Gamma distribution. The full form of the corresponding CDF is given by

$$\begin{aligned}
 F_{\chi_{sn}^2, X}(b; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) &= \exp\left(-\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right) \sum_{n=0}^{\infty} \frac{(\sigma_{X_1} \cdot \mu_{\chi_{sn}^2})^n}{2^n \cdot n!} \\
 &\quad \cdot \underbrace{\frac{\gamma\left(\frac{N_f}{2} + n, \frac{b}{2 \cdot \sigma_{X_1}^2}\right)}{\Gamma\left(\frac{N_f}{2} + n\right)}}_{F_{\Gamma, X}\left(\frac{b}{2}; \sigma_{X_1}^2, \frac{N_f}{2} + n\right)} \\
 &= \exp\left(-\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right) \sum_{n=0}^{\infty} \frac{\left(\frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{2}\right)^n}{n!} \quad (A.9.6) \\
 &\quad \cdot \left(1 - \exp\left(-\frac{b}{2 \cdot \sigma_{X_1}^2}\right) \sum_{m=0}^{\frac{N_f}{2} + n - 1} \frac{b^m}{2^m \cdot m! \cdot \sigma_{X_1}^{2m}}\right),
 \end{aligned}$$

where $F_{\Gamma, X}\left(\frac{b}{2}; \sigma_{X_1}^2, \frac{N_f}{2} + n\right)$ represents the CDF from (A.7.2) of the Gamma distribution in Appendix A.7.

The first raw **moment** (mean) of the scaled non-central chi-squared distribution is calculated by [Wal07, Chapter 31.3]

$$\mu_X = E[X] = N_f \cdot \sigma_{X_1}^2 + \mu_{\chi_{sn}^2}. \quad (A.9.7)$$

Approximations

To reduce the complexity, approximation of the scaled non-central chi-squared distribution can be applied. In particular, the *non-central chi-squared distribution* can be approximated by the chi-squared distribution for large b and N_f' [JKB95; Wal07]. Thus, we obtain an approximated scaled non-central chi-squared distribution by

$$F_{\chi_{sn}^2, X}(b; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) \approx F_{\chi^2, X}(b'; N'_f) \quad \text{with} \quad (\text{A.9.8})$$

$$b' = \frac{b}{\sigma_{X_1}^2} \cdot \frac{N_f + \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{N_f + 2 \cdot \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}},$$

$$N'_f = \frac{(N_f + \sigma_{X_1} \cdot \mu_{\chi_{sn}^2})^2}{N_f + \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}.$$

Further reduction in complexity can be achieved by using the *standard normal distribution* for large $(\sigma_{X_1} \cdot \mu_{\chi_{sn}^2})$, and for non-extreme values of $\frac{b}{\sigma_{X_1}^2}$ with [JKB95]

$$F_{\chi_{sn}^2, X}(b; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) = F_{\chi_{nc}^2, X}\left(\frac{b}{\sigma_{X_1}^2}; N_f, \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}\right),$$

$$\approx F_{\mathcal{N}, X}(b') \quad (\text{A.9.9})$$

$$\text{where } b' = \sqrt{\frac{2b}{\sigma_{X_1}^2 \cdot c}} - \sqrt{\frac{2d}{c} - 1},$$

$$c = 1 + \frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{d},$$

$$d = N_f + \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}$$

Moreover, for large $(\sigma_{X_1} \cdot \mu_{\chi_{sn}^2})$ and large $(P \cdot (1 - P))$ the ICDF of the approximation (A.9.9) is

$$F_{\chi_{sn}^2}^{-1}(P; N_f, \mu_{\chi_{sn}^2}, \sigma_{X_1}^2) = \sigma_{X_1}^2 F_{\chi_{nc}^2}^{-1}(P; N_f, \sigma_{X_1} \cdot \mu_{\chi_{sn}^2})$$

$$\approx \left(F_{\mathcal{N}}^{-1}(P) + \sqrt{\frac{2d}{c} - 1} \right)^2 \cdot \frac{\sigma_{X_1}^2 \cdot c}{2}, \quad (\text{A.9.10})$$

$$\text{where } c = 1 + \frac{\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}}{d}, \quad d = N_f + \sigma_{X_1} \cdot \mu_{\chi_{sn}^2}.$$

Note, that for values within invalid range the approximation is not a monotonic increasing function. More detailed, decreasing the probability P can increase the result of the ICDF. Empirically it can be said, that for $(P \cdot (1 - P)) > 10^{-7}$ and $(\sigma_{X_1} \cdot \mu_{\chi_{sn}^2}) > 50$ the approximation is within valid range and achieves good results.

Special Case

A special case of the scaled non-central chi-squared distribution is present for the proper complex normal distributed random variable with identical variance introduced in Appendix A.2. There, we rewrite the two random variables X_1, X_2 with $X_n \sim f_{\mathcal{N}}(\cdot; \mu_{X_n}, \sigma_{Y_1}^2)$ as a complex random variable

$$Y_1 = \frac{X_1 + iX_2}{\sqrt{2}} \sim f_{\mathbb{C}\mathcal{N}}(\cdot; \mu_{Y_1}, \sigma_{Y_1}^2) \quad (\text{A.9.11})$$

with $\mu_{Y_1} = (\mu_{X_1} + i\mu_{X_2})/\sqrt{2}$. Thus, we have

$$\begin{aligned} \Re\{Y_1\} &\sim f_{\mathcal{N}}\left(\cdot; \frac{1}{\sqrt{2}}\Re\{\mu_{Y_1}\}, \frac{1}{2}\sigma_{Y_1}^2\right), \\ \Im\{Y_1\} &\sim f_{\mathcal{N}}\left(\cdot; \frac{1}{\sqrt{2}}\Im\{\mu_{Y_1}\}, \frac{1}{2}\sigma_{Y_1}^2\right). \end{aligned} \quad (\text{A.9.12})$$

Consequently, the *scaled non-central chi-squared distribution* with $N_f = 2$ is described by

$$Y = \|Y_1\|_2^2 \sim f_{\chi_{sn}^2}\left(\cdot; 2, \frac{1}{2}|\mu_{Y_1}|^2, \frac{1}{2}\sigma_{Y_1}^2\right). \quad (\text{A.9.13})$$

The special case is frequently used in communication theory to describe e.g. the power of a signal. Furthermore, Y represents the squared random variable of the [Rice distribution](#) [MC04]. Among others, the Rice distribution is used in propagation of communication signals to model [Rice fading](#), see Section 2.3. The mean and the variance is obtained by (A.2.1) and (A.2.2), respectively. Consequently, the PDF and CDF for Y from (A.9.5) and (A.9.6) simplifies to

$$\begin{aligned} f_{\chi_{sn}^2, Y}\left(a; 2, \frac{1}{2}|\mu_{Y_1}|^2, \frac{1}{2}\sigma_{Y_1}^2\right) &= \exp\left(-\frac{1}{\sqrt{32}} \cdot \sigma_{Y_1} \cdot |\mu_{Y_1}|^2 - \frac{a}{\sigma_{Y_1}^2}\right) \\ &\cdot \sum_{n=0}^{\infty} \frac{\left(\sqrt{2} \cdot |\mu_{Y_1}|^2 \cdot a\right)^n}{2^{3n} \cdot \sigma_{Y_1}^{n+2} \cdot (n!)^2}, \end{aligned} \quad (\text{A.9.14})$$

$$F_{\chi^2_{sn}, Y} \left(b; 2, \frac{1}{2} |\mu_{Y_1}|^2, \frac{1}{2} \sigma_{Y_1}^2 \right) = \exp \left(-\frac{1}{\sqrt{32}} \cdot \sigma_{Y_1} \cdot |\mu_{Y_1}|^2 \right) \quad (\text{A.9.15})$$

$$\cdot \sum_{n=0}^{\infty} \frac{\left(\sqrt{2} \cdot \sigma_{Y_1} \cdot |\mu_{Y_1}|^2 \right)^n}{2^{3n} \cdot (n!)^2} \cdot \gamma \left(1 + n, \frac{b}{\sigma_{Y_1}^2} \right).$$

Acronyms

3GPP 3rd Generation Partnership Project.

5G the fifth generation of cellular mobile communications.

a.s. almost surely.

AMR allocation map retrieval.

AWGN additive white Gaussian noise.

BP basis pursuit.

BPDN basis pursuit denoising.

BS base station.

CAC circular autocorrelation of the CIR.

CDF cumulative distribution function.

cf. confer.

CIR channel impulse response.

CLT central limit theorem.

COSAMP compressive sampling matching pursuit.

CP cyclic-prefix.

CS compressed sensing.

D2D device-to-device.

DFT discrete Fourier transform.

DSA dynamic spectrum access.

e.g. for example.

ECDF empirical CDF.

EPA extended pedestrian A.

et al. and others.

EVA extended vehicular A.

FDMA frequency-division multiple access.

HIRATE High Performance Digital Radio Testbed.

i.e. that is.

i.i.d. independent and identically distributed.

ICDF inverse cumulative distribution function.

IDFT inverse discrete Fourier transform.

iff if and only if.

IOA indoor office A.

IoT Internet-of-Things.

ITU International Telecommunication Union.

LASSO least absolute shrinkage and selection operator.

LLN law of large numbers.

LLR log-likelihood ratio.

LOS line-of-sight.

LS least squares.

LTE Long-Term Evolution.

LTI linear time-invariant.

M2M machine-to-machine.

MLR monotone likelihood ratio.

MPC multipath component.

NLOS non-line-of-sight.

NP non-deterministic polynomial-time.

NR new radio.

NSP null space property.

OFDM orthogonal frequency-division multiplexing.

OFDMA orthogonal FDMA.

OMP orthogonal matching pursuit.

PDF probability density function.

- PDP** power delay profile.
- PRB** physical resource block.
- PSD** power spectral density.
- PU** primary user.
- QAM** quadrature amplitude modulation.
- r.m.s.** root mean square.
- RAM** resource allocation map.
- RAS** resource allocation set.
- RF** radio frequency.
- RIP** restricted isometry property.
- ROC** receiver operating characteristic.
- RRLS** reduced-rank least squares.
- s.t.** subject to.
- SAGE** space-alternating generalized expectation-maximization.
- SC-FDMA** single-carrier frequency-division multiple access.
- SCP** spectral channel power.
- SDR** software-defined radio.
- SFA** short-range factory automation.
- SNR** signal-to-noise ratio.
- SPGL1** spectral projected gradient for ℓ_1 -minimization.
- SU** secondary user.
- SVD** singular value decomposition.
- TTI** transmission time interval.
- UE** user equipment.
- UMP** uniformly most powerful.
- WLAN** wireless local area network.
- WSSUS** wide-sense stationary uncorrelated scatterers.

Notation

a, A	Italic lower- and uppercase letters denote scalars
\mathbb{A}	Blackboard bold uppercase letters denote index sets
\mathbf{a}	Lowercase bold letters denote column vectors
\mathbf{A}	Uppercase bold letters denote matrices
$(\cdot)_{\text{LS}}, (\cdot)_n$	Subscript distinguishes different variables or indicate dependency on a subscript variable
$a \in [n, m)$	$n \leq a < m$, i.e., a takes all values between n and m excluding value m
$\Re\{a\}, \Im\{a\}$	Real and imaginary part of variable a
$\mathbf{a}_{[n]}$	The n -th element of column vector \mathbf{a}
$\mathbf{A}_{[n, m]}$	The element of matrix \mathbf{A} at row n and column m
$\mathbf{A}_{[:, m]}$	Column m of matrix \mathbf{A}
$\mathbf{A}_{[n, :]}$	Row n of matrix \mathbf{A}
$\mathbf{A}_{[\mathbb{A}, \mathbb{B}]}$	Submatrix formed by selecting a set \mathbb{A} of rows and a set \mathbb{B} of columns from \mathbf{A}
$\mathbf{A} \in \mathbb{R}^{a \times b}$	Real matrix \mathbf{A} with a rows and b columns
$\mathbf{a} \in \mathbb{C}^a$	Complex vector \mathbf{A} with a rows
$(a \times b)$ matrix	Generic matrix with a rows and b columns
$(\cdot)^\top$	Transpose of vector or matrix
$(\cdot)^*$	Conjugate complex of vector or matrix
$(\cdot)^H$	Conjugate complex and transpose of vector or matrix
$\hat{(\cdot)}$	Fourier transform of variable

$(\cdot)'$	Modified version of variable
$\widetilde{(\cdot)}$	Estimate of a variable
\mathbf{A}^{-1}	Inverse of matrix \mathbf{A}
\mathbf{A}^+	Pseudoinverse of matrix \mathbf{A}
$a \bmod b$	Modulo operation
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of matrix \mathbf{A} and \mathbf{B}
$\mathbf{a}^\top \cdot \mathbf{b}$	Matrix product of row vector \mathbf{a}^\top and column vector \mathbf{b}
$\mathbf{A} \cdot \mathbf{B}$	Matrix product of matrix \mathbf{A} and \mathbf{B}
$ \cdot $	Absolute value or cardinality of a set
$\lfloor \cdot \rfloor$	Flooring function
$\ \mathbf{a}\ _p$	ℓ_p -norm, quasinorm or "norm" depending on value p
$\ \cdot\ _F$	Frobenius norm of a matrix

List of Symbols

Functions and Operators

$\mathfrak{S}(\mathbb{A})$	Group of the permutations of \mathbb{A} (symmetric group)
$\Gamma(\cdot)$	Gamma function
Π	Permutation
$\delta(\cdot)$	Dirac delta function
$\gamma_s(\boldsymbol{a})_p$	Error of best s -term approximation to a vector \boldsymbol{a}
$\gamma(\cdot, \cdot)$	Lower incomplete Gamma function
$\mathbb{E}[X]$	Expectation of X
$F_{(\cdot)}(\cdot)$	CDF of random variable or distribution
$\check{F}(\cdot)$	Empirical CDF
$F^{-1}(\cdot)$	ICDF (quantile function)
$\text{MC}(\boldsymbol{A})$	Mutual coherence of matrix
$\text{NS}(\boldsymbol{A})$	Null space of matrix
$\mathcal{O}(\cdot)$	Big O notation
$\mathfrak{P}(\mathbb{A})$	Power set of \mathbb{A}
$\text{P}(X)$	Probability of X
$\text{P}(Y \mid X)$	Conditional probability of Y given X
d	Differential operator
$\text{diag}(\boldsymbol{A})$	Diagonal elements of matrix \boldsymbol{A}
$\text{diag}(\boldsymbol{a})$	Diagonal matrix constructed from vector \boldsymbol{a}
$\text{erf}(\cdot)$	Error function

$\text{erf}^{-1}(\cdot)$. Inverse error function
$\exp(\cdot)$. Natural exponential function
$f(\cdot)$. Context specific function
$f_{(\cdot)}(\cdot)$. PDF of random variable or distribution
$f_{\Gamma}(\cdot; \sigma^2, N'_f)$. Gamma distribution (here PDF)
$f_{\mathcal{N}}(\cdot; \mu, \sigma^2)$. Real valued normal distribution (here PDF)
$f_{\mathbb{C}\mathcal{N}}(\cdot; \mu, \sigma^2)$. Complex normal distribution (here PDF)
$f_{\text{Rayleigh}}(\cdot; \sigma)$. Rayleigh distribution (here PDF)
$f_{\chi}(\cdot; N_f)$. Chi distribution (here PDF)
$f_{\chi^2}(\cdot; N_f)$. Chi-squared distribution (here PDF)
$f_{\chi^2_{nc}}(\cdot; N_f, \mu)$. Non-central chi-squared distribution
$f_{\chi^2_{sn}}(\cdot; N_f, \mu, \sigma^2)$. Scaled non-central chi-squared distribution
$g(\cdot)$. Context specific function
$\ln(\cdot)$. Natural logarithm
$\log_a(\cdot)$. Logarithm to base a
$\text{median}(\cdot)$. Median
$\text{rank}(\cdot)$. Matrix rank
$\text{supp}(\cdot)$. Vector support
$\text{Tr}(\cdot)$. Matrix trace
$\mathbf{1}_{\mathbb{A}}(\cdot)$. Indicator function of set \mathbb{A}

Matrices

Φ	. Measurement matrix
Σ	. Diagonal matrix
\mathbf{A}	. Context specific matrix variable
\mathbf{B}	. Context specific matrix variable
\mathbf{E}_{LS}	. LS estimator
\mathbf{E}_{RRLS}	. RRLS estimator
\mathbf{F}	. DFT matrix

\mathbf{I}_N	$N \times N$ identity matrix
\mathbf{S}	Overall resource allocation map, i.e. $\mathbf{S} = \begin{bmatrix} \mathbf{S}_0 & \dots & \mathbf{S}_K \end{bmatrix}$
\mathbf{S}_k	Resource allocation map of transmitter k
\mathbf{U}	Unitary matrix
\mathbf{V}	Unitary matrix

Scalars

Λ	Log-likelihood ratio
Ω_{BW}	Bandwidth
α	Context specific parameter
α_{KF}	K-factor
β	Context specific parameter
δ	Restricted isometry constant
λ	Tuning parameter
μ_X	Mean of X
$\mu_{\mathbf{a}}$	Mean of $\mathbf{a}_{[n]}$
ρ	Threshold
$\rho_{\text{c},a}$	Composite hypothesis threshold on a
ρ_{ed}	Threshold of excess delay
$\rho_{\text{s},a}$	Simple hypothesis threshold on a
σ_X	Standard deviation of X
$\sigma_{\mathbf{a}}$	Standard deviation of $\mathbf{a}_{[n]}$
σ_X^2	Variance of X
$\sigma_{\mathbf{a}}^2$	Variance of $\mathbf{a}_{[n]}$
τ	Delay
τ_{coh}	Coherence time
τ_{DS}	Delay spread
τ_{ED}	Excess delay
$\tau_{\text{ED,max}}$	Maximum excess delay

τ_{offset}	Time offset
τ_p	Path delay
A	Context specific scalar variable
\mathcal{H}_0	Null hypothesis
\mathcal{H}_1	Alternative hypothesis
$\mathcal{H}_{1,c}$	Composite alternative hypothesis
$\mathcal{H}_{1,s}$	Simple alternative hypothesis
K	Number of transmitters
L	Natural number
M	Natural number
N	Natural number
N_B	Blocksize
N_{Bell}	Bell number
N_{CP}	Cyclic-prefix length
N_{DFT}	DFT size
N_{eps}	Number of possible equal partition sets
N_f	Degree of freedom
N_{RB}	Number of resource blocks
N_{Stirling}	Stirling number of second kind
P_d	Detection probability
P_{fa}	False alarm probability
U	Random variable
W	Random variable
X	Random variable
Y	Random variable
Z	Random variable
θ	PDF parameter
a	Context specific scalar variable
b	Context specific scalar variable

c	Context specific scalar variable or constant
d	Context specific scalar variable
e	Adversarial error/noise symbol
f	Frequency/subcarrier index
h	Channel symbol
i	Imaginary number ($i = \sqrt{-1}$)
k	Transmitter index
l	Context specific index variable
m	Context specific index variable
n	Context specific index variable
p	Hyperparameter of ℓ_p -norm
s	Sparsity of vector, i.e. $\ \mathbf{a}\ _0 \leq s$
t	Time index
u	Transmit symbol
w	Received symbol
x	Unknown
y	Measurement
z	Absolute squared received symbol

Sets

Σ	Canonical subspaces
Θ	PDF parameter set
\mathbb{A}	Context specific set
\mathbb{B}	Context specific set
\mathbb{C}	Set of all complex numbers
\mathbb{F}	DFT index set
\mathbb{K}	Transmitter set
\mathbb{N}	Set of all natural numbers
\mathbb{N}_{CAC}	CAC index set
\mathbb{N}_{CIR}	Channel impulse response index set

\mathbb{P}	Orthogonal frequency resource set
\mathbb{R}	Set of all real numbers
\mathbb{S}	Set of frequency resource sets
\mathbb{S}_k	Frequency resource set of transmitter k
$\tilde{\mathbb{S}}_E$	Frequency resource set estimated by energy detection
$\tilde{\mathbb{S}}_G$	Frequency resource set estimated by gray space detection
\mathbb{T}	Set of PRBs
\mathbb{T}_n	Resource set of n -th PRB

Vectors

$\boldsymbol{\theta}$	PDF parameter vector
\boldsymbol{a}	Context specific column vector variable
\boldsymbol{b}	Context specific column vector variable
\boldsymbol{e}	Adversarial error/noise vector
\boldsymbol{h}	CIR vector
\boldsymbol{n}	Noise vector
\boldsymbol{s}	Resource allocation vector
\boldsymbol{u}	Transmit signal vector
\boldsymbol{v}	CAC vector
\boldsymbol{w}	Received signal vector
\boldsymbol{x}	Unknown variable
\boldsymbol{y}	Measurement or Observation vector
$\hat{\boldsymbol{z}}$	Power spectrum vector

Own Publication List

- M. Dungen, T. Hansen, R. Croonenbroeck, R. Kays, B. Holfeld, D. Wieruch, P. W. Berenguer, V. Jungnickel, D. Block, U. Meier, and H. Schulze, “Channel Measurement Campaigns for Wireless Industrial Automation”, *at - Automatisierungstechnik*, volume 67, no. 1, pp. 7–28, 2019.
- S. Eldessoki, B. Holfeld, and D. Wieruch, “Impact of Waveforms on Coexistence of Mixed Numerologies in 5G URLLC Networks”, in *Proc. of 21th International ITG Workshop on Smart Antennas (WSA)*, Berlin, Germany, VDE, 2017.
- D. Wieruch, P. Jung, and B. Holfeld, “Device and Method for Associating Resource Information with Channel Metric Information in Wireless Networks”, EN, no. WO/2019/068639, Patent, International Application, 2017.
- B. Holfeld, D. Wieruch, L. Raschkowski, T. Wirth, C. Pallasch, W. Herfs, and C. Breche, “Radio Channel Characterization at 5.85 GHz for Wireless M2M Communication of Industrial Robots”, in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2016.
- B. Holfeld, D. Wieruch, T. Wirth, L. Thiele, S. A. Ashraf, J. Huschke, I. Aktas, and J. Ansari, “Wireless Communication for Factory Automation: an Opportunity for LTE and 5G Systems”, *IEEE Communications Magazine*, volume 54, no. 6, pp. 36–43, 2016.
- J. Pilz, M. Mehlhose, T. Wirth, D. Wieruch, B. Holfeld, and T. Haustein, “A Tactile Internet Demonstration: 1ms Ultra Low Delay for Wireless Communications towards 5G”, in *Proc. of IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2016.
- D. Wieruch, B. Holfeld, and T. Wirth, “Wireless Factory Automation: Radio Channel Evolution in Repeated Manufacturing Processes”, in *Proc. of 20th International ITG Workshop on Smart Antennas (WSA)*, Munich, Germany, VDE, 2016.
- D. Wieruch, P. Jung, T. Wirth, and A. Dekorsy, “Identifying Non-Adjacent Multiuser Allocations by Joint l1-Minimization”, in *Proc. of Wireless Communications and Networking Conference (WCNC)*, IEEE, 2016.

- D. Wieruch, P. Jung, T. Wirth, T. Haustein, and A. Dekorsy, "Cognitive Radios Exploiting Gray Spaces via Compressed Sensing", *Frequenz*, volume 70, no. 7-8, pp. 289-300, 2016.
- T. Wirth, M. Mehlhose, J. Pilz, B. Holfeld, and D. Wieruch, "5G New Radio and Ultra Low Latency Applications: A PHY Implementation Perspective", in *Proc. of 50th Asilomar Conference on Signals, Systems and Computers (Asilomar)*, 2016.
- T. Wirth, D. Wieruch, B. Holfeld, M. Mehlhose, J. Pilz, T. Haustein, R. Halfmann, and K.-J. Friederichs, "Real-Time Demonstration of Optimized Spectrum Usage with LSA Carrier Aggregation", *Frequenz*, volume 70, no. 7-8, pp. 301-308, 2016.
- D. Wieruch, *Spectrum Sensing via Gray Space Detection by Exploiting Channel Sparsity*, Talk, Caltech (group of Babak Hassibi), Pasadena, CA, United States, 2015.
- D. Wieruch, P. Jung, C. Bockelmann, and A. Dekorsy, "Exploiting the Sparsity of Wireless Channels to Identify Non-Adjacent Spectrum Usage of Multiple Users", in *Poster Session of Matheon Conference on Compressed Sensing and its Applications*, TU Berlin, Berlin, Germany, 2015.
- D. Wieruch, P. Jung, T. Wirth, and A. Dekorsy, "Determining User Specific Spectrum Usage via Sparse Channel Characteristics", in *Proc. of 49th Asilomar Conference on Signals, Systems and Computers (Asilomar)*, IEEE, 2015.
- T. Wirth, M. Mehlhose, J. Pilz, R. Lindstedt, D. Wieruch, B. Holfeld, and T. Haustein, "An Advanced Hardware Platform to Verify 5G Wireless Communication Concepts", in *Proc. of IEEE 81st Vehicular Technology Conference (VTC-Spring)*, IEEE, 2015.
- T. Wirth, B. Holfeld, D. Wieruch, R. Halfmann, and K.-J. Friederichs, "System Level Performance of Cellular Networks Utilizing ASA/LSA Mechanisms", in *Proc. of 1st International Workshop on Cognitive Cellular Systems (CCS)*, IEEE, 2014.
- D. Wieruch, J. Pilz, and P. Jung, "Compressive Gray Space Detection for Interweaved Cognitive Radio Systems", in *Proc. of The Tenth International Symposium on Wireless Communication Systems Workshop (ISWCS Workshop)*, Ilmenau, Germany, VDE, 2013.
- D. Wieruch, T. Wirth, O. Braz, A. Dußmann, M. Mederle, and M. Müller, "Cognitive Repeaters for Flexible Mobile Data Traffic Offloading", in *Proc. of 8th International Conference on Cognitive Radio Oriented Wireless Networks (CrownCom)*, IEEE, 2013.
- D. Wieruch and V. Pohl, "A Cognitive Radio Architecture Based on Sub-Nyquist Sampling", in *Proc. of IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, IEEE, 2011.

-
- A. Ibing, D. Kühling, D. Wieruch, and H. Boche, “Software Defined Hybrid MMSE/QRD-M Turbo Receiver for LTE Advanced Uplink on a Cell Processor”, in *Proc. of IEEE International Conference on Communications Workshops (ICC)*, IEEE, 2009.
 - D. Wieruch, D. Kühling, and A. Ibing, “430MBit/s 64-State Soft-Decision Viterbi Decoder on a Cell CPU”, in *Mikroelektroniktagung*, Vienna, Austria, 2008.

Bibliography: Books

- [ATJ14] S. Atapattu, C. Tellambura, and H. Jiang, “Conventional Energy Detector”, in *Energy Detection for Spectrum Sensing in Cognitive Radio*, 2nd ed., New York, NY, USA, Springer, 2014, chap. 2, pp. 11–27 (cit. on pp. 11, 84, 91, 93).
- [BL98] D. K. Barton and S. A. Leonov, *Radar Technology Encyclopedia*, Artech House, 1998 (cit. on p. 27).
- [Bru09] R. A. Brualdi, *Introductory Combinatorics*, 5th ed., Pearson, 2009, p. 648 (cit. on pp. 138, 139).
- [CB01] G. Casella and R. Berger, *Statistical Inference*, Pacific Grove, CA, USA, Duxbury Resource Center, 2001, p. 660 (cit. on pp. 22, 23, 25).
- [EK12] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, New York, NY, USA, Cambridge University Press, 2012, p. 544 (cit. on pp. 46, 51, 55, 64, 98).
- [FR13] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Applied and Numerical Harmonic Analysis, Birkhäuser, 2013 (cit. on pp. 45, 48, 51–54, 62, 63).
- [HR17] I. Haviv and O. Regev, “The Restricted Isometry Property of Subsampled Fourier Matrices”, in *Lecture Notes in Mathematics*, Springer International Publishing, 2017, pp. 163–179 (cit. on p. 68).
- [Her98] A. Hero, “Signal Detection and Classification”, in *The Digital Signal Processing Handbook*, ed. by V. K. Madisetti and D. B. Williams, 1st ed., Boca Raton, Florida, USA, CRC Press LLC, 1998, chap. 13 (cit. on pp. 28, 106).
- [JKB94] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed., volume 1, Wiley Series in Probability and Mathematical Statistics, New York, NY, USA, Wiley, 1994 (cit. on pp. 181, 183, 186, 188–190).

- [JKB95] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed., volume 2, Wiley Series in Probability and Mathematical Statistics, New York, NY, USA, Wiley, 1995 (cit. on pp. 181, 197, 198).
- [JW15] P. Jung and P. Walk, “Sparse Model Uncertainties in Compressed Sensing with Application to Convolutions and Sporadic Communication”, in *Compressed Sensing and its Applications*, ed. by H. Boche, R. Calderbank, G. Kutyniok, and J. Vybiral, Springer, 2015, pp. 1–29 (cit. on pp. 55, 80).
- [Kol33] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und Ihrer Grenzgebiete, Julius Springer, Berlin, Heidelberg, 1933, p. 62 (cit. on p. 49).
- [LR05] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed., Springer Texts in Statistics, New York, NY, USA, Springer, 2005, p. 786 (cit. on pp. 20, 22, 23, 25, 39, 91).
- [Lév26] P. Lévy, *Calcul des probabilités*, Paris, Gauthier-Villars, 1926, p. 368 (cit. on p. 49).
- [MC04] S. L. Miller and D. Childers, *Probability and Random Processes: With Applications to Signal Processing and Communications*, Elsevier Science, 2004 (cit. on pp. 186, 199).
- [Mol11] A. F. Molisch, *Wireless Communications*, 2nd ed., John Wiley & Sons, 2011 (cit. on p. 14).
- [Mui82] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley Series in Probability and Statistics, New York, NY, USA, Wiley, 1982, p. 704 (cit. on p. 194).
- [NW10] F. Neumann and C. Witt, “Combinatorial Optimization and Computational Complexity”, in *Bioinspired Computation in Combinatorial Optimization: Algorithms and Their Computational Complexity*, Berlin, Heidelberg, Springer, 2010, chap. 2, pp. 9–19 (cit. on p. 137).
- [Owe01] A. B. Owen, *Empirical Likelihood*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, 2001 (cit. on p. 38).
- [Pap91] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed., McGraw-Hill Companies, 1991 (cit. on pp. 181, 182).
- [SD91] L. L. Scharf and C. Demeure, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley series in electrical and computer engineering, Boston, MA, USA, Addison-Wesley Publishing Company, 1991, p. 524 (cit. on p. 26).

- [SS14] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press, 2014, 330 pp. (cit. on p. 185).
- [Ste11] M. B. Stefania Sesia Issam Toufik, *LTE - The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed., Wiley, 2011, p. 792 (cit. on p. 20).
- [Tre01] H. L. van Trees, *Detection, Estimation, and Modulation Theory, Part 1: Detection, Estimation, and Linear Modulation Theory*, 1st ed., New York, NY, USA, Wiley, 2001, p. 716 (cit. on pp. 22, 95).
- [TV05] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005 (cit. on p. 58).
- [Wal07] C. Walck, *Hand-book on Statistical Distributions for experimentalists*, University of Stockholm Internal Report SUF-PFY/96-01, 2007 (cit. on pp. 181–183, 186, 188–190, 192, 197).

Bibliography

- [3GP20a] 3GPP, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer; General Description*, tech. rep. no. TS 36.201 V15.3.0, April 2020 (cit. on pp. 17, 88, 104, 128, 169).
- [3GP20b] 3GPP, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding*, tech. rep. no. TS 36.212 V15.9.0, April 2020 (cit. on p. 87).
- [3GP20c] 3GPP, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, tech. rep. no. TS 36.211 V15.9.0, April 2020 (cit. on pp. 18, 87).
- [3GP20d] 3GPP, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*, tech. rep. no. TS 36.213 V15.9.0, April 2020 (cit. on pp. 18, 88, 108).
- [3GP20e] 3GPP, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception*, tech. rep. no. TS 36.101 V15.10.0, April 2020 (cit. on pp. 15, 16).
- [3GP20f] 3GPP, *Technical Specification Group Radio Access Network; NR; Physical layer procedures for control*, tech. rep. no. TS 38.213 V15.9.0, April 2020 (cit. on p. 18).
- [3GP20g] 3GPP, *Technical Specification Group Radio Access Network; NR; Physical layer procedures for data*, tech. rep. no. TS 38.214 V15.9.0, April 2020 (cit. on p. 18).
- [3GP20h] 3GPP, *Technical Specification Group Radio Access Network; NR; Physical Layer; General Description*, tech. rep. no. TS 38.201 V16.0.0, January 2020 (cit. on pp. 1, 17, 128, 169).
- [APZ17] A. Abid, A. Poon, and J. Zou, “Linear Regression with Shuffled Labels”, in *ArXiv e-prints*, 2017 (cit. on p. 147).

- [AS14] J. Andersson and J.-O. Strömberg, “On the Theorem of Uniform Recovery of Random Sampling Matrices”, *IEEE Transactions on Information Theory*, volume 60, no. 3, pp. 1700–1710, March 2014 (cit. on p. 63).
- [ABC+14] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What Will 5G Be?”, *IEEE Journal on Selected Areas in Communications*, volume 32, no. 6, pp. 1065–1082, June 2014 (cit. on p. 1).
- [BHS+10] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, “Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels”, *Proceedings of the IEEE*, volume 98, no. 6, pp. 1058–1076, June 2010 (cit. on p. 72).
- [Bal62] A. Balakrishnan, “On the Problem of Time Jitter in Sampling”, *IRE Transactions on Information Theory*, volume 8, no. 3, pp. 226–236, April 1962 (cit. on p. 148).
- [BLM17] A. S. Bandeira, M. E. Lewis, and D. G. Mixon, “Discrete Uncertainty Principles and Sparse Signal Processing”, *Journal of Fourier Analysis and Applications*, volume 24, no. 4, pp. 935–956, June 2017 (cit. on pp. 68, 152).
- [BDD+08] R. G. Baraniuk, M. Davenport, R. DeVore, and M. B. Wakin, “A Simple Proof of the Restricted Isometry Property for Random Matrices”, *Constructive Approximation*, volume 28, no. 3, pp. 253–263, January 2008 (cit. on p. 55).
- [BT90] M. Barton and D. W. Tufts, “Reduced-Rank Least Squares Channel Estimation”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, volume 38, no. 8, pp. 1403–1410, August 1990 (cit. on p. 57).
- [BBD19] E. Beck, C. Bockelmann, and A. Dekorsy, “Compressed Edge Spectrum Sensing: Extensions and practical considerations”, *at - Automatisierungstechnik*, volume 67, no. 1, pp. 51–59, January 2019 (cit. on p. 12).
- [BES+95] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. Ola Borjesson, “On Channel Estimation in OFDM Systems”, in *Proc. of IEEE VTC*, Chicago, IL, USA, July 1995 (cit. on p. 56).
- [Bel63] P. A. Bello, “Characterization of Randomly Time-Variant Linear Channels”, *IEEE Transactions on Communications Systems*, volume 1, no. 4, pp. 360–393, December 1963 (cit. on p. 14).
- [BF08] E. van den Berg and M. P. Friedlander, “Probing the Pareto Frontier for Basis Pursuit Solutions”, *SIAM Journal on Scientific Computing*, volume 31, no. 2, pp. 890–912, 2008 (cit. on pp. 63, 105).

- [BF15] E. van den Berg and M. P. Friedlander, *SPGL1: A Solver for Large-Scale Sparse Reconstruction*, Version 1.9, April 2015, URL: <https://www.math.ucdavis.edu/~mpf/spgl1/> (visited on 09/14/2015) (cit. on pp. 63, 105, 157).
- [BWH+10] C. Berger, Z. Wang, J. Huang, and S. Zhou, “Application of Compressive Sensing to Sparse Channel Estimation”, *IEEE Communications Magazine*, volume 48, no. 11, pp. 164–174, November 2010 (cit. on p. 72).
- [BMW20] BMWi, *Leitfaden 5G-Campusnetze-Orientierungshilfe für kleine und mittelständische Unternehmen*, tech. rep., 2020 (cit. on p. 10).
- [BSD13] C. Bockelmann, H. F. Schepker, and A. Dekorsy, “Compressive Sensing Based Multi-User Detection for Machine-to-Machine Communication”, *Transactions on Emerging Telecommunications Technologies*, volume 24, no. 4, pp. 389–400, April 2013 (cit. on p. 143).
- [Bru55] R. H. Bruck, “Difference Sets in a Finite Group”, *Transactions of the American Mathematical Society*, volume 78, no. 2, pp. 464–481, March 1955 (cit. on p. 68).
- [CZ14] T. T. Cai and A. Zhang, “Sparse Representation of a Polytope and Recovery of Sparse Signals and Low-Rank Matrices”, *IEEE Transactions on Information Theory*, volume 60, no. 1, pp. 122–132, January 2014 (cit. on p. 63).
- [CRT06] E. J. Candès, J. K. Romberg, and T. Tao, “Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information”, *IEEE Transactions on Information Theory*, volume 52, no. 2, pp. 489–509, 2006 (cit. on pp. 2, 12, 43, 50).
- [CT05] E. J. Candès and T. Tao, “Decoding by Linear Programming”, *IEEE Transactions on Information Theory*, volume 51, no. 12, pp. 4203–4215, 2005 (cit. on pp. 54, 62, 63, 84, 97).
- [CT06] E. J. Candès and T. Tao, “Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?”, *IEEE Transactions on Information Theory*, volume 52, no. 12, pp. 5406–5425, December 2006 (cit. on p. 55).
- [Can33] F. P. Cantelli, “Sulla determinazione empirica delle leggi di probabilità. (On the empirical determination of a probability law)”, Italian, *Giornale dell’Istituto Italiano degli Attuari*, volume 4, pp. 421–424, 1933 (cit. on pp. 38, 39, 101).
- [CEP14] CEPT ECC Report 206, *Compatibility studies in the band 5725–5875 MHz between SRD equipment for wireless industrial applications and other systems*, tech. rep., CEPT ECC, January 2014 (cit. on p. 116).

- [CDS98] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic Decomposition by Basis Pursuit”, *SIAM Journal on Scientific Computing*, volume 20, no. 1, pp. 33–61, 1998 (cit. on p. 62).
- [CL06] F. Chung and L. Lu, “Concentration Inequalities and Martingale Inequalities: A Survey”, *Internet Mathematics*, volume 3, no. 1, pp. 79–127, January 2006 (cit. on p. 178).
- [CDD08] A. Cohen, W. Dahmen, and R. DeVore, “Compressed Sensing and Best k-Term Approximation”, *Journal of the American Mathematical Society*, volume 22, no. 1, pp. 211–231, July 2008 (cit. on pp. 53, 63).
- [Dav10] M. A. Davenport, “Random Observations on Random Observations: Sparse Signal Acquisition and Processing”, PhD thesis, Rice University, August 2010, p. 203 (cit. on p. 55).
- [Deu18] Deutsche Telekom, *Mehr Netz für Land und Leute*, November 2018, URL: <https://www.telekom.com/de/medien/medieninformationen/detail/mehr-netz-fuer-land-und-leute-545408> (visited on 02/18/2019) (cit. on p. 10).
- [DAS07] F. F. Digham, M.-S. Alouini, and M. K. Simon, “On the Energy Detection of Unknown Signals Over Fading Channels”, *IEEE Transactions on Communications*, volume 55, no. 1, pp. 21–24, 2007 (cit. on p. 12).
- [DE03] D. L. Donoho and M. Elad, “Optimally Sparse Representation in General (Nonorthogonal) Dictionaries via l1 Minimization”, *Proceedings of the National Academy of Sciences*, volume 100, no. 5, pp. 2197–2202, February 2003 (cit. on p. 54).
- [Don06] D. L. Donoho, “Compressed Sensing”, *IEEE Transactions on Information Theory*, volume 52, no. 4, pp. 1289–1306, April 2006 (cit. on pp. 2, 12, 43, 50).
- [DHC+19] M. Düngen, T. Hansen, R. Croonenbroeck, R. Kays, B. Holfeld, D. Wieruch, P. W. Berenguer, V. Jungnickel, D. Block, U. Meier, and H. Schulze, “Channel Measurement Campaigns for Wireless Industrial Automation”, *at - Automatisierungstechnik*, volume 67, no. 1, pp. 7–28, 2019 (cit. on pp. 4, 5, 85, 114).
- [EY36] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank”, *Psychometrika*, volume 1, no. 3, pp. 211–218, September 1936 (cit. on p. 57).
- [EO15] H. E. Egilmez and A. Ortega, “Wavelet-Based Compressed Spectrum Sensing for Cognitive Radio Wireless Networks”, in *Proc. of IEEE ICASSP*, April 2015 (cit. on pp. 12, 142).

- [EBD+14] V. Emiya, A. Bonnefoy, L. Daudet, and R. Gribonval, “Compressed Sensing with Unknown Sensor Permutation”, in *Proc. of IEEE ICASSP*, 2014 (cit. on p. 147).
- [Eng08] S. Engelberg, “Edge Detection Using Fourier Coefficients”, *The American Mathematical Monthly*, volume 115, no. 6, pp. 499–513, 2008 (cit. on p. 151).
- [FTH+99] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen, “Channel Parameter Estimation in Mobile Radio Environments Using the SAGE Algorithm”, *IEEE Journal on Selected Areas in Communications*, volume 17, no. 3, pp. 434–450, March 1999 (cit. on p. 117).
- [Fra63] R. Frank, “Polyphase codes with good nonperiodic correlation properties”, *IEEE Transactions on Information Theory*, volume 9, no. 1, pp. 43–45, January 1963 (cit. on p. 116).
- [Gal08] R. G. Gallager, “Circularly-Symmetric Gaussian Random Vectors”, January 2008 (cit. on p. 186).
- [GJY11] D. Ge, X. Jiang, and Y. Ye, “A Note on the Complexity of L_p Minimization”, *Mathematical Programming*, volume 129, no. 2, pp. 285–299, June 2011 (cit. on p. 52).
- [Gli33] V. I. Glivenko, “Sulla determinazione empirica delle leggi di probabilità. (On the empirical determination of a probability law)”, Italian, *Giornale dell’Istituto Italiano degli Attuari*, volume 4, pp. 92–99, 1933 (cit. on pp. 38, 39, 101).
- [GJM+09] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, “Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective”, *Proceedings of the IEEE*, volume 97, no. 5, pp. 894–914, 2009 (cit. on pp. 11, 83).
- [HC18] S. Haghighatshoar and G. Caire, “Signal Recovery from Unlabeled Samples”, *IEEE Transactions on Signal Processing*, volume 66, pp. 1242–1257, March 2018 (cit. on pp. 148, 149).
- [Hay05] S. Haykin, “Cognitive Radio: Brain-Empowered Wireless Communications”, *IEEE Journal on Selected Areas in Communications*, volume 23, no. 2, pp. 201–220, February 2005 (cit. on pp. 1, 11, 83).
- [HB03] B. M. Hochwald and S. ten Brink, “Achieving Near-Capacity on a Multiple-Antenna Channel”, *IEEE Transactions on Communications*, volume 51, no. 3, pp. 389–399, March 2003 (cit. on p. 179).
- [HWR+16] B. Holfeld, D. Wieruch, L. Raschkowski, T. Wirth, C. Pallasch, W. Herfs, and C. Breche, “Radio Channel Characterization at 5.85 GHz for Wireless M2M Communication of Industrial Robots”, in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2016 (cit. on pp. 4, 5, 85, 114, 115, 117).

- [HWW+16] B. Holfeld, D. Wieruch, T. Wirth, L. Thiele, S. A. Ashraf, J. Huschke, I. Aktas, and J. Ansari, “Wireless Communication for Factory Automation: an Opportunity for LTE and 5G Systems”, *IEEE Communications Magazine*, volume 54, no. 6, pp. 36–43, 2016 (cit. on pp. 1, 4, 5, 85).
- [HSS17] D. J. Hsu, K. Shi, and X. Sun, “Linear Regression without Correspondence”, *ArXiv e-prints*, May 2017 (cit. on p. 147).
- [IKW+09] A. Ibing, D. Kühling, D. Wieruch, and H. Boche, “Software Defined Hybrid MMSE/QRD-M Turbo Receiver for LTE Advanced Uplink on a Cell Processor”, in *Proc. of IEEE International Conference on Communications Workshops (ICC)*, IEEE, 2009 (cit. on p. 179).
- [ITU97] ITU-R M.1225 Recommendation, *Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*, tech. rep., ITU-R, February 1997 (cit. on pp. 15, 16).
- [JSW09] P. Jung, W. Schüle, and G. Wunder, “Robust Path Detection for the LTE Downlink based on Compressed Sensing”, in *Proc. of International OFDM-Workshop*, Hamburg, Germany, 2009 (cit. on p. 69).
- [JKM18] P. Jung, R. Kueng, and D. G. Mixon, “Derandomizing compressed sensing with combinatorial design”, pp. 1–12, December 2018 (cit. on p. 68).
- [KR56] S. Karlin and H. Rubin, “The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio”, *The Annals of Mathematical Statistics*, volume 27, no. 2, pp. 272–299, 1956 (cit. on p. 25).
- [KKP+13] W. Keusgen, A. Kortke, M. Peter, and R. Weiler, “A Highly Flexible Digital Radio Testbed and 60 GHz Application Examples”, in *European Microwave Conference (EuMC)*, October 2013 (cit. on p. 116).
- [Khi34] A. Khintchine, “Korrelationstheorie der stationären stochastischen Prozesse”, *Mathematische Annalen*, volume 109, no. 1, pp. 604–615, December 1934 (cit. on p. 74).
- [KSJ19] M. Kliesch, S. J. Szarek, and P. Jung, “Simultaneous Structures in Convex Signal Recovery—Revisiting the Convex Combination of Norms”, *Frontiers in Applied Mathematics and Statistics*, volume 5, May 2019 (cit. on p. 80).
- [KHB+15] A. Kliks, O. Holland, A. Basaure, and M. Matinmikko, “Spectrum and License Flexibility for 5G Networks”, *IEEE Communications Magazine*, volume 53, no. 7, pp. 42–49, July 2015 (cit. on p. 10).

- [KW11] F. Krahmer and R. Ward, “New and Improved Johnson-Lindenstrauss Embeddings via the Restricted Isometry Property”, *SIAM Journal on Mathematical Analysis*, volume 43, no. 3, pp. 1269–1281, January 2011 (cit. on p. 55).
- [Kuh55] H. W. Kuhn, “The Hungarian Method for the Assignment Problem”, *Naval Research Logistics Quarterly*, volume 2, no. 1–2, pp. 83–97, March 1955 (cit. on p. 167).
- [KJT+15] M. Kurras, S. Jaeckel, L. Thiele, and V. Braun, “CSI Compression and Feedback for Network MIMO”, in *Proc. of IEEE VTC’15 Spring*, Glasgow, May 2015 (cit. on p. 117).
- [Lan67a] H. J. Landau, “Necessary Density Conditions for Sampling and Interpolation of Certain Entire Functions”, *Acta Mathematica*, volume 117, no. 1, pp. 37–52, July 1967 (cit. on p. 12).
- [Lan67b] H. J. Landau, “Sampling, Data Transmission, and the Nyquist Rate”, *Proceedings of the IEEE*, volume 55, no. 10, pp. 1701–1706, 1967 (cit. on p. 50).
- [Láz12] J. A. Lázaro, “GNSS Array-based Acquisition: Theory and Implementation”, PhD thesis, Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), July 2012 (cit. on p. 26).
- [Lin22] J. W. Lindeberg, “Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung”, *Mathematische Zeitschrift*, volume 15, no. 1, pp. 211–225, December 1922 (cit. on p. 49).
- [LFL17] C. Liu, Y. Fang, and J. Liu, “Some New Results About Sufficient Conditions for Exact Support Recovery of Sparse Signals via Orthogonal Matching Pursuit”, *IEEE Transactions on Signal Processing*, volume 65, no. 17, pp. 4511–4524, September 2017 (cit. on p. 66).
- [LTH+14] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and Y. G. Li, “Channel Estimation for OFDM”, *IEEE Communications Surveys & Tutorials*, volume 16, no. 4, pp. 1891–1908, November 2014 (cit. on pp. 46, 69).
- [Mac95] G. Mackiw, “A Note on the Equality of the Column and Row Rank of a Matrix”, *Mathematics Magazine*, volume 68, no. 4, pp. 285–286, October 1995 (cit. on p. 47).
- [MZ93] S. G. Mallat and Z. Zhang, “Matching Pursuits with Time-Frequency Dictionaries”, *IEEE Transactions on Signal Processing*, volume 41, no. 12, pp. 3397–3415, December 1993 (cit. on p. 64).
- [Mol05] A. F. Molisch, “Ultrawideband Propagation Channels-Theory, Measurement, and Modeling”, *IEEE Transactions on Vehicular Technology*, volume 54, no. 5, pp. 1528–1545, 2005 (cit. on pp. 2, 17, 19, 61, 72).

- [Nat95] B. K. Natarajan, “Sparse Approximate Solutions to Linear Systems”, *SIAM Journal on Computing*, volume 24, no. 2, pp. 227–234, April 1995 (cit. on p. 52).
- [NT10] D. Needell and J. A. Tropp, “CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples”, *Communications of the ACM*, volume 53, no. 12, pp. 93–100, December 2010 (cit. on p. 64).
- [NM93] F. D. Neeser and J. L. Massey, “Proper Complex Random Processes with Applications to Information Theory”, *IEEE Transactions on Information Theory*, volume 39, no. 4, pp. 1293–1302, July 1993 (cit. on p. 185).
- [NP33] J. Neyman and E. S. Pearson, “On the Problem of the Most Efficient Tests of Statistical Hypotheses”, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, volume 231, pp. 289–337, 1933 (cit. on p. 23).
- [OJF+15] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi, “Simultaneously Structured Models With Application to Sparse and Low-Rank Matrices”, *IEEE Transactions on Information Theory*, volume 61, no. 5, pp. 2886–2908, May 2015 (cit. on p. 80).
- [OA07] M. Ozdemir and H. Arslan, “Channel Estimation for Wireless OFDM Systems”, *IEEE Communications Surveys & Tutorials*, volume 9, no. 2, pp. 18–48, July 2007 (cit. on pp. 46, 69).
- [PWC16] A. Pananjady, M. J. Wainwright, and T. A. Courtade, “Linear Regression with an Unknown Permutation: Statistical and Computational Limits”, *ArXiv e-prints*, August 2016 (cit. on p. 147).
- [PWC17] A. Pananjady, M. J. Wainwright, and T. A. Courtade, “Denoising Linear Models with Permuted Data”, *ArXiv e-prints*, April 2017 (cit. on p. 147).
- [PRR+93] Y. C. Pati, R. Rezaifar, Y. C. P. R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition”, in *Proc. of Asilomar*, 1993 (cit. on p. 64).
- [PRI+18] C. Patsioura, J. Robinson, P. Iacopino, and T. Hatt, *Radar*, December 2018, URL: <https://www.gsmaintelligence.com/research/2018/12/radar-december-2018/714/> (visited on 04/15/2019) (cit. on p. 10).
- [Peh13] J. M. Peha, “Spectrum Sharing in the Gray Space”, *Telecommunications Policy*, volume 37, no. 2-3, pp. 167–177, 2013 (cit. on p. 11).
- [Pra18] R. Prasad, “Unlabelled Sensing: A Sparse Bayesian Learning Approach”, in *ArXiv e-prints*, 2018 (cit. on p. 148).

- [Qua17] Qualcomm, *Private LTE Networks*, July 2017, URL: <https://www.qualcomm.com/media/documents/files/private-lte-networks.pdf> (visited on 02/18/2019) (cit. on p. 10).
- [RV08] M. Rudelson and R. Vershynin, “On Sparse Reconstruction from Fourier and Gaussian Measurements”, *Communications on Pure and Applied Mathematics*, volume 61, no. 8, pp. 1025–1045, 2008 (cit. on p. 68).
- [Sar12] R. Saruthirathanaworakun, “Gray-Space Spectrum Sharing with Cellular Systems and Radars, and Policy Implications”, PhD thesis, Carnegie Mellon University, December 2012, p. 164 (cit. on p. 11).
- [Sel13] I. Selesnick, “Least Squares with Examples in Signal Processing”, Brooklyn, New York, USA, 2013 (cit. on pp. 46, 47).
- [Sha48] C. E. Shannon, “A Mathematical Theory of Communication”, *The Bell System Technical Journal*, volume 27, pp. 379–423, July 1948 (cit. on p. 50).
- [SMF05] T. B. Sorensen, P. E. Mogensen, and F. Frederiksen, “Extension of the ITU Channel Models for Wideband (OFDM) Systems”, in *Proc. of IEEE VTC-Fall*, 2005 (cit. on p. 15).
- [SCF+13] P. Stenumgaard, J. Chilo, P. Ferrer-Coll, and P. Angskog, “Challenges and Conditions for Wireless Machine-to-Machine Communications in Industrial Environments”, *IEEE Communications Magazine*, volume 51, no. 6, pp. 187–192, 2013 (cit. on p. 1).
- [SNW+13] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, “Wideband Spectrum Sensing for Cognitive Radio Networks: A Survey”, *IEEE Wireless Communications*, volume 20, no. 2, pp. 74–81, April 2013 (cit. on pp. 11, 12).
- [TH17] M. Thill and B. Hassibi, “Low-Coherence Frames from Group Fourier Matrices”, *IEEE Transactions on Information Theory*, volume 63, no. 6, pp. 3386–3404, June 2017 (cit. on p. 68).
- [TG06] Z. Tian and G. B. Giannakis, “A Wavelet Approach to Wideband Spectrum Sensing for Cognitive Radios”, in *Proc. of IEEE Crown-Com*, June 2006 (cit. on pp. 12, 142).
- [TG07a] Z. Tian and G. B. Giannakis, “Compressed Sensing for Wideband Cognitive Radios”, in *Proc. of ICASSP*, 2007 (cit. on p. 12).
- [Tib96] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso”, *Journal of the Royal Statistical Society, Series B (Methodological)*, volume 58, no. 1, pp. 267–288, 1996 (cit. on p. 62).
- [Til15] A. M. Tillmann, “On the Computational Intractability of Exact and Approximate Dictionary Learning”, *IEEE Signal Processing Letters*, volume 22, no. 1, pp. 45–49, 2015 (cit. on pp. 144, 147).

- [TP14] A. M. Tillmann and M. E. Pfetsch, “The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing”, *IEEE Transactions on Information Theory*, volume 60, no. 2, pp. 1248–1259, February 2014 (cit. on pp. 53, 55).
- [TG07b] J. A. Tropp and A. C. Gilbert, “Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit”, *IEEE Transactions on Information Theory*, volume 53, no. 12, pp. 4655–4666, 2007 (cit. on pp. 64, 84).
- [UHV15a] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, “Unlabeled Sensing with Random Linear Measurements”, *ArXiv e-prints*, November 2015 (cit. on pp. 126, 127, 142–146).
- [UHV15b] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, “Unlabeled Sensing: Solving a Linear System with Unordered Measurements”, in *Proc. of IEEE Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015 (cit. on p. 146).
- [Urk67] H. Urkowitz, “Energy Detection of Unknown Deterministic Signals”, *Proceedings of the IEEE*, volume 55, no. 4, pp. 523–531, April 1967 (cit. on pp. 11, 34, 36, 84, 91, 93).
- [VMB+10] V. Valenta, R. Marsalek, G. Baudoin, M. Villegas, M. Suarez, and F. Robert, “Survey on Spectrum Utilization in Europe: Measurements, Analyses and Observations”, in *Proc. of IEEE CrownCom*, 2010 (cit. on p. 10).
- [VB01] R. Venkataramani and Y. Bresler, “Optimal Sub-Nyquist Nonuniform Sampling and Reconstruction for Multiband Signals”, *IEEE Transactions on Signal Processing*, volume 49, no. 10, pp. 2301–2313, October 2001 (cit. on p. 12).
- [WBJ15] P. Walk, H. Becker, and P. Jung, “OFDM Channel Estimation via Phase Retrieval”, in *Proc. of Asilomar*, November 2015 (cit. on p. 75).
- [WJP15] P. Walk, P. Jung, and G. E. Pfander, “On the Stability of Sparse Convolutions”, *Applied and Computational Harmonic Analysis*, September 2015 (cit. on p. 80).
- [WZB+17] G. Wang, J. Zhu, R. S. Blum, P. Willett, S. Marano, V. Matta, and P. Braca, “Maximum Likelihood Signal Amplitude Estimation Based on Permuted Blocks of Differently Binary Quantized Observations of a Signal in Noise”, in *ArXiv e-prints*, 2017 (cit. on p. 147).
- [Wel74] L. Welch, “Lower Bounds on the Maximum Cross Correlation of Signals (Corresp.)”, *IEEE Transactions on Information Theory*, volume 20, no. 3, pp. 397–399, May 1974 (cit. on p. 54).

- [WZW+17] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo, “A Sharp Condition for Exact Support Recovery With Orthogonal Matching Pursuit”, *IEEE Transactions on Signal Processing*, volume 65, no. 6, pp. 1370–1382, March 2017 (cit. on pp. 65, 66).
- [Wie30] N. Wiener, “Generalized Harmonic Analysis”, *Acta Mathematica*, volume 55, no. 0, pp. 117–258, 1930 (cit. on p. 74).
- [WHW16] D. Wieruch, B. Holfeld, and T. Wirth, “Wireless Factory Automation: Radio Channel Evolution in Repeated Manufacturing Processes”, in *Proc. of 20th International ITG Workshop on Smart Antennas (WSA)*, Munich, Germany, VDE, 2016 (cit. on pp. 4, 5, 85, 114, 115).
- [WJB+15] D. Wieruch, P. Jung, C. Bockelmann, and A. Dekorsy, “Exploiting the Sparsity of Wireless Channels to Identify Non-Adjacent Spectrum Usage of Multiple Users”, in *Poster Session of Math-eon Conference on Compressed Sensing and its Applications*, TU Berlin, Berlin, Germany, 2015 (cit. on pp. 5, 125, 126).
- [WJH17] D. Wieruch, P. Jung, and B. Holfeld, “Device and Method for Associating Resource Information with Channel Metric Information in Wireless Networks”, EN, no. WO/2019/068639, Patent, International Application, 2017 (cit. on pp. 5, 126).
- [WJW+15] D. Wieruch, P. Jung, T. Wirth, and A. Dekorsy, “Determining User Specific Spectrum Usage via Sparse Channel Characteristics”, in *Proc. of 49th Asilomar Conference on Signals, Systems and Computers (Asilomar)*, IEEE, 2015 (cit. on pp. 5, 125, 126).
- [WJW+16a] D. Wieruch, P. Jung, T. Wirth, and A. Dekorsy, “Identifying Non-Adjacent Multiuser Allocations by Joint ℓ_1 -Minimization”, in *Proc. of Wireless Communications and Networking Conference (WCNC)*, IEEE, 2016 (cit. on pp. 5, 126).
- [WJW+16b] D. Wieruch, P. Jung, T. Wirth, T. Haustein, and A. Dekorsy, “Cognitive Radios Exploiting Gray Spaces via Compressed Sensing”, *Frequenz*, volume 70, no. 7-8, pp. 289–300, 2016 (cit. on pp. 3–5, 10, 41, 83–86, 91, 142).
- [WPJ13] D. Wieruch, J. Pilz, and P. Jung, “Compressive Gray Space Detection for Interweaved Cognitive Radio Systems”, in *Proc. of The Tenth International Symposium on Wireless Communication Systems Workshop (ISWCS Workshop)*, Ilmenau, Germany, VDE, 2013 (cit. on pp. 3–5, 10, 41, 83–85, 142).
- [WP11] D. Wieruch and V. Pohl, “A Cognitive Radio Architecture Based on Sub-Nyquist Sampling”, in *Proc. of IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, IEEE, 2011 (cit. on p. 12).

- [Wil38] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *The Annals of Mathematical Statistics*, volume 9, no. 1, pp. 60–62, 1938 (cit. on pp. 11, 30, 84, 91).
- [WR04] D. P. Wipf and B. D. Rao, "Sparse Bayesian Learning for Basis Selection", *IEEE Transactions on Signal Processing*, volume 52, no. 8, pp. 2153–2164, 2004 (cit. on pp. 144, 148).
- [WWH+16] T. Wirth, D. Wieruch, B. Holfeld, M. Mehlhose, J. Pilz, T. Haustein, R. Halfmann, and K.-J. Friederichs, "Real-Time Demonstration of Optimized Spectrum Usage with LSA Carrier Aggregation", *Frequenz*, volume 70, no. 7-8, pp. 301–308, 2016 (cit. on p. 10).
- [XZG05] P. Xia, S. Zhou, and G. B. Giannakis, "Achieving the Welch Bound with Difference Sets", *IEEE Transactions on Information Theory*, volume 51, no. 5, pp. 1900–1907, May 2005 (cit. on pp. 54, 68).
- [XZG06] P. Xia, S. Zhou, and G. B. Giannakis, "Correction to "Achieving the Welch Bound With Difference Sets"", *IEEE Transactions on Information Theory*, volume 52, no. 7, pp. 3359–3359, July 2006 (cit. on pp. 54, 68).
- [YVS15] P. L. Yu, G. Verma, and B. M. Sadler, "Wireless physical layer authentication via fingerprint embedding", *IEEE Communications Magazine*, volume 53, no. 6, pp. 48–53, June 2015 (cit. on p. 131).
- [YA09] T. Yucek and H. Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications", *IEEE Communications Surveys & Tutorials*, volume 11, no. 1, pp. 116–130, 2009 (cit. on pp. 11, 142).
- [ZG11] H. Zhu and G. B. Giannakis, "Exploiting Sparse User Activity in Multiuser Detection", *IEEE Transactions on Communications*, volume 59, no. 2, pp. 454–465, 2011 (cit. on p. 143).

Index

- allocation map retrieval, AMR** 2, 125–127, 131, 135, 137, 138, 142–144, 148–151, 153, 155, 157, 165, 166, 169, 171, 173, 178, 179
- basis pursuit denoising, BPDN** 3–5, 44, 56, 62–64, 67, 72, 80, 81, 84, 97, 99, 104, 105, 107, 108, 110, 111, 123, 124, 155, 157–162, 165, 174, 176–178
- basis pursuit, BP** 52, 53
- Bell number** 138–141, 156, 158, 159, 173, 178, 210
 - set** 129, 137–139, 141, 142, 150, 155–159, 161–163, 165
- best s -term approximation** 43, 45, 53, 63, 207
- central limit theorem, CLT** 11, 12, 49, 84, 91, 92
 - Lindeberg–Lévy** 49
- channel estimation** 2, 3, 44, 46, 57, 66, 69, 72, 80, 81, 142, 175
 - amplitude-based** 3, 4, 44, 67, 72, 73, 76, 81, 83–86, 88–90, 93, 95, 96, 99, 103, 104, 106, 110, 111, 123, 176, 177
 - pilot-based** 3, 44, 67, 69, 72, 73, 81, 176
- channel impulse response, CIR** 3, 5, 14–17, 19, 20, 44, 66, 67, 69–76, 78–81, 96, 104, 118, 121, 130, 132, 134, 156–158, 161, 162, 164, 165, 169–174, 176, 178
- channel transfer function** 3, 5, 16, 44, 66, 67, 69, 70, 72, 74–76, 81, 88, 117, 126–128, 132–134, 157, 158, 164, 165, 169–171, 173, 176, 178
- circular autocorrelation of the CIR, CAC** 5, 66, 67, 72–76, 78–80, 96, 99, 104, 111, 132, 134, 156, 157, 164, 165, 169–171
- constant false alarm rate** 10, 21, 27, 28, 35, 36
- cumulative distribution function, CDF** 3, 21, 35, 36, 38–41, 71, 101, 112, 155, 177, 181–185, 187–200, 207
 - empirical CDF, ECDF** 3, 4, 10, 39–41, 83, 85, 93, 100, 124, 176, 177, 207
- delay spread** 122
- dictionary learning** 144, 148
- difference set** 54, 68
- discrete Fourier transform, DFT** 6, 17, 18, 20, 43, 48, 51, 54, 56, 67–69, 72, 74, 75, 96, 131–134, 146, 151, 152, 154, 157
- distribution** 3, 10, 13–15, 20–23, 27–31, 35–37, 39, 40, 44, 48–51, 56, 70, 71, 77, 81, 91, 93, 100, 101, 112, 147, 157, 163, 169, 176, 181–199
 - chi** 186–189, 208
 - chi-squared** 31, 32, 91, 100, 103, 189–192, 195, 197, 198, 208

- circularly-symmetric complex normal** 20, 29, 70, 157, 163, 169, 186, 188, 189, 191, 193
- complex normal** 20, 29, 70, 77, 178, 185, 186, 188, 190, 191, 193, 199, 208
- Gamma** 3, 13, 23, 30, 31, 36, 38, 40, 41, 70, 71, 77, 94, 100, 176, 190–194, 196, 197, 208
- non-central chi-squared** 32, 100, 103, 194–198, 208
- normal** 20–22, 29, 49–51, 56, 70, 71, 77, 91, 147, 183–195, 198, 199, 208
- Rayleigh** 14, 77, 78, 188, 193, 208
- Rice** 14, 77, 199
- scaled non-central chi-squared** 30–41, 77, 94, 100–102, 112, 195–200, 208
- Eckart–Young theorem** 58
- energy detection** 2–4, 11, 12, 36, 84, 85, 89–96, 100, 102, 104–106, 108–114, 123, 124, 176, 177
- error function** 183, 185, 187, 207
 - inverse** 71, 184, 185, 187, 208
- excess delay** 14–16, 19, 121, 122, 130, 172, 173, 209
- Glivenko–Cantelli theorem** 3, 10, 39, 40, 101, 176
- gray space** 4, 11, 12, 84, 86, 88, 90–94, 100, 102, 104, 106–112, 123, 126, 177
- gray space detection** 2, 4, 5, 10, 12, 28, 83–86, 89, 90, 93–95, 98–100, 103, 104, 108–112, 114, 122–125, 142, 149, 150, 175–178
- gray space test** 4, 83, 85, 93, 100, 102, 106, 112–114, 124, 177
- Hungarian algorithm** 167, 168
- hypothesis** 3, 9, 10, 21–42, 83, 85–92, 94–96, 100, 102, 103, 106, 112, 176, 177, 210
- composite** 2–4, 9, 10, 21, 23, 25–33, 35–42, 83, 84, 89–93, 95, 100–103, 112, 113, 176, 209
- one-sided** 2, 3, 9, 10, 25–27, 29–33, 37, 39–42, 91, 92, 102, 176
- simple** 3, 9, 10, 21, 23–25, 27, 28, 30–37, 84, 90–95, 100, 102, 112, 113, 176, 209
- indicator function** 7, 39, 130, 153, 208
- instance optimal** 63
- inverse discrete Fourier transform, IDFT** 152
- inverse cumulative distribution function, ICDF** 33, 35, 36, 50, 71, 102, 182, 184, 185, 187, 188, 192, 194, 198, 207
- Johnson–Lindenstrauss lemma** 56
- Karlin–Rubin theorem** 3, 9, 10, 23, 25, 26, 29, 176
- law of large numbers, LLN** 49
- least squares, LS** 3, 48, 54, 56, 57, 61, 62, 64, 65, 148, 162, 176
- likelihood ratio test** 9, 24
- log-likelihood ratio test** 4, 21, 24, 26, 32–41, 83–85, 90–96, 99–104, 106, 112, 123, 124, 177
- log-likelihood ratio, LLR** 23–26, 31, 32, 34, 38–40, 42, 90–93, 100–104, 176
- monotone likelihood ratio, MLR** 25, 26, 32, 38
- matrix**
 - Frobenius norm** 58, 206
 - full rank** 47, 48, 57, 136, 148, 149
 - rank** 47, 48, 57–61, 144, 148, 149, 208
 - rank deficient** 47, 57
 - singular value decomposition, SVD** 57, 58, 61, 97
 - trace** 58–60, 208

- median 49, 50, 208
- multiplicative uncertainty principle 152–154
- mutual coherence 43, 52, 54–56, 68
- Neyman-Pearson lemma 3, 9, 23–25, 37, 176
- NP-hard 52, 54, 56, 138, 144
- nuisance parameter 22, 26, 31, 90
- null space property, **NSP** 43, 52–55, 62
- orthogonal matching pursuit, **OMP** 3–5, 44, 56, 64–67, 72, 80, 81, 84, 98, 99, 104, 105, 107–111, 123, 124, 127, 157, 160–167, 170, 173, 174, 176–178
- partition of a set 129, 138–140, 159, *see* Bell set
- power delay profile, **PDP** 16, 118, 119, 121
- power spectral density, **PSD** 12, 16, 74
- PRB allocation map** 156
- PRB allocation set** 130, 135, 156, 165, 167
- probability 22, 23, 25–27, 29, 33–42, 71, 88, 91–95, 100–106, 112, 113, 181–185, 187, 188, 192–194, 198, 207, 210
 - conditional probability** 22, 23, 25, 27, 35, 36, 88, 207
 - detection** 3, 10, 22, 23, 25, 27, 29, 33–38, 40–42, 88, 94, 95, 104–106, 108–114, 182, 210
 - false alarm** 3, 10, 22, 23, 25–27, 29, 33–40, 42, 88, 91–93, 100–103, 105–114, 123, 124, 210
- probability density function, **PDF** 20–27, 29–32, 34, 35, 38–41, 49, 70, 77, 94, 100, 101, 103, 113, 181–191, 193–197, 199, 208
- receiver operating characteristic, **ROC** 22, 29, 85, 95, 106–109, 114, 123
- reduced-rank least squares, **RRLS** 3, 4, 44, 56–61, 67, 70–72, 78, 81, 84, 97, 99, 104, 105, 107–111, 122–124, 175, 177
- resource allocation map, **RAM** 125, 126, 130, 131, 135–137, 142, 145, 150, 153, 155, 156, 160, 165, 173, 177
- resource allocation set, **RAS** 127–131, 142, 143, 150, 153, 155, 156
- restricted isometry property, **RIP** 43, 52, 55, 56, 62, 63, 66, 68, 161
 - restricted isometry constant** 55, 63, 66
- root mean square, **r.m.s.** 16, 122
- Rouché–Capelli theorem 47
- set of sets 127, 129, 130, 137–139, 141, 142, 155–158, 165
- sparse 3, 9, 10, 12, 17, 19, 43–46, 53–56, 61–66, 68, 72, 75, 80, 87, 98, 99, 135, 143, 144, 147–153, 155, 157–159, 161–165, 169–172, 174, 207, 211, 213
 - compressible** 17, 44–46, 72
- spectral channel power, **SCP** 3–5, 16, 17, 44, 66, 72–76, 81, 83–85, 89–94, 96, 99–101, 103, 104, 106, 108, 110, 111, 113, 121–123, 126–128, 134, 143, 157, 164, 165, 169–171, 173, 176–179
- spectral projected gradient for ℓ_1 -minimization, **SPGL1** 63, 104, 108, 158, 160
- Stirling number of the second kind 138–141, 156, 159, 173, 178, 210
- Stirling set of the second kind** 129, 137–142, 150, 155–159,

161–165

uniformly most powerful, UMP 2–

4, 9, 10, 21, 23–29, 32, 36,

37, 85, 100–103, 106, 112–114,

124, 176, 177

union of subspace 43–46, 143

unlabeled sensing 5, 126, 127, 142–

147, 149, 173, 178

ordered 147, 148

permutation 147–149

selection 126, 127, 142–144, 147,

149, 150

vector

quasinorm 7, 44, 52

Welch bound 54, 68

Wiener-Khintchine theorem 74, 76

Wilks theorem 12, 31, 84, 91, 92