

Research-Internship

Winter Term 2016/2017

**Compressed Channel Sensing in THz
Communications**

**Institute for Digital Communications
(IDC)**

Viktoria Schram

supervised by Anna Moldovan and Prof. Dr. Gerstacker

June 22, 2017

Abstract

In this work, the problem of reliable wireless transmission over indoor Terahertz (THz) channels is studied. THz communication is a new emerging technology which offers significantly higher data rates compared to existing technologies, enables very-high-speed wireless communication for huge data file transfer and allows the development of novel applications with great advantages over already existing techniques. However, there are still difficulties to be overcome until a stable data transfer is possible. This work considers the estimation of the THz channel, whose characteristics are essential for reliable information transmission that suffers from high molecular absorption, high spreading loss, high reflection loss and thermal noise. Due to the specific conditions arising for THz wave propagation, the discrete-time channel impulse response (CIR) is characterized by only a few dominant taps and many (approximately-)zero taps. The channel is said to be approximately sparse. Conventional channel estimation techniques like least-squares (LS) estimation fail to fully exhibit this structure and face significant disadvantages. This work proposes the usage of compressive sensing (CS) channel estimation methods to take into account the sparsity of the THz channel. Two approaches are studied thoroughly, namely one method based on solving a convex program by using the *Dantzig selector* (DS) and one approach using a greedy pursuit method called *compressive sampling matching pursuit* (CoSaMP). Both methods are compared to the LS approach in terms of MSE (mean squared error), computational efficiency and numbers of observations needed. Results show that significant advantages over LS estimation are achievable in all categories.

Contents

1. Introduction	4
2. System Model	6
2.1. Characteristics of THz Channel	6
2.2. Design of Transmission Signal	8
3. Channel Estimation	9
3.1. Fundamentals of Compressed Sensing	10
3.2. Design of Sensing Matrix	11
3.3. Presentation of Reconstruction Algorithms	13
3.3.1. Least Squares Approach	14
3.3.2. Convex Program	14
3.3.3. Compressive Sampling Matching Pursuit Algorithm	17
3.4. Evaluation of Proposed Concept	19
3.4.1. Channel Estimation for the Entire Subchannel Length	20
3.4.2. Channel Estimation for the First 1000 Subchannel Taps	31
3.4.3. Performance Analysis of a Stochastic Sensing Matrix	37
4. Conclusion and Future Discussion	38
Appendix	40
A. THz Channel properties	40
B. THz Discrete Time CIR	42

1. Introduction

THz communication systems provide a possible solution to the problem of handling a steadily growing data traffic ([M⁺16b]) and a way to realize very-high-speed wireless communication networks. Furthermore, the narrow bandwidth, which is a fundamental limitation of current mobile networks, can be overcome by the increased THz operation frequency leading to higher channel capacity. The THz frequency range is defined from 0.1 THz to 3 THz with wavelengths between 3 mm and 30 μ m. The usage of THz communication can enhance existing applications which require high data rates like wireless short range communication between devices, ultra-high-speed data communication in small cells or secure wireless communication for military and defense applications. Moreover, this emerging communication technology can encourage the development of new applications like monitoring systems in the medical field or new paradigms like the Internet-of-Nano-Things. As THz radiation is non-ionizing, applications close to humans are possible without concern. [A⁺14, M⁺16b]

To apply THz communication successfully, some difficulties need to be overcome. Challenges occur for instance, when the THz channel is investigated closer [A⁺14]. It is characterized by a very high molecular absorption which leads to signal distortion and colored noise. Together with white thermal noise the generated colored noise defines the main noise source [M⁺16a]. Molecular absorption and spreading loss result in a very high and frequency selective path loss for line-of-sight (LOS) links. Moreover, high reflection loss depending on the characteristics of the reflecting surfaces influence THz wave propagation for non-line-of-sight (NLOS) links [M⁺16b].

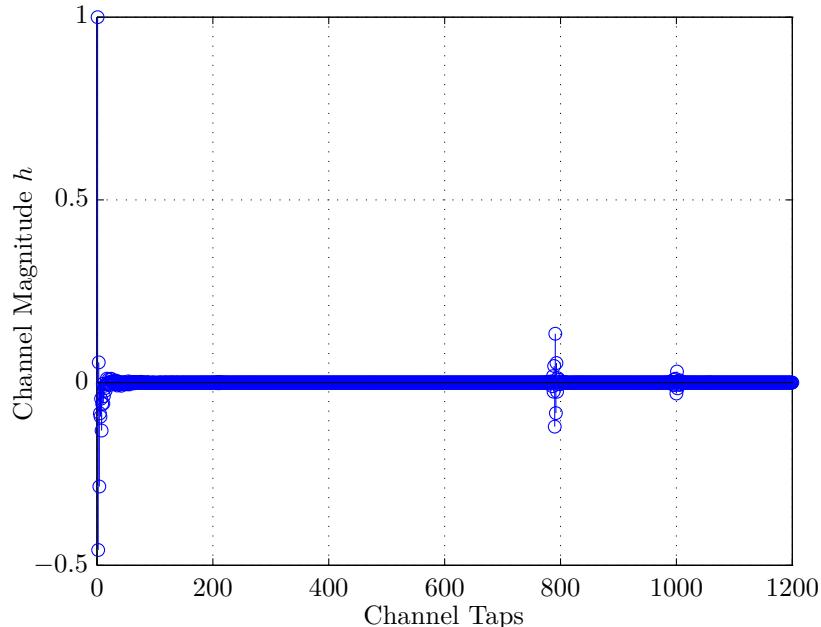


Figure 1: First 1200 taps of the first subband out of 32 subbands of the real part of the equivalent discrete-time CIR of the THz channel.

These difficult transmission characteristics limit THz wave propagation to a few meters ([M⁺16b]) and lead to an equivalent discrete-time channel impulse response (CIR) with special properties and peculiarities. The first channel taps of the first subchannel are shown in Figure 1, for the assumption that the whole channel is divided into 32 subchannels in total. Only a few distinct channel taps occur while predominantly (nearly-)zero taps characterize the CIR. Hence, the discrete-time channel is assumed to be approximately sparse.

The work of [M⁺16b] and [M⁺16a] prove that data rates in the order of Tbps can be achieved, while the authors of [M⁺16b] verify this for a distance of up to 1 m. [M⁺16a] proposes different bounds for the maximum distance, taking into account various transmission schemes and an efficient power allocation algorithm. These results encourage the implementation of THz communication systems, though perfect channel state information (CSI) is often assumed in the literature, which is not available in real systems. Therefore, a channel estimation technique has to be developed.

Usually, channel estimation is performed by using an LS approach. However, this method cannot fully exploit the characteristics of the THz channel and incorporate its sparsity [H⁺10]. Moreover, to obtain a reasonable result in terms of mean squared error (MSE), the number of used observations has to be higher or equal to the number of unknown channel taps [H⁺10]. Both aspects would lead to a high computational cost, as the considered THz channel exhibits a long CIR with few significant taps. Therefore, the scope of this work is to perform channel estimation based on the compressed sensing (CS) technique which exploits the property of sparsity and to analyze the achievable performance in terms of MSE, computational efficiency and necessary number of observations. It should also be noted that CS provides a possibility to reconstruct a signal from fewer measurements than the size of the signal [C⁺10].

CS has already been used for channel sensing before. There exists a great variety of publications about adaptation of channel sensing to the CS framework, which shows achievable bounds and gives recommendations about system design. In this context, the theories of [H⁺10, Baj09, Baj08, B⁺07, CT07, CR07, Can08] and [FR13] were used to establish the theoretical parts of this work. In this report, the main difference to the previously named literature is the application to a more difficult THz channel, which was to the best of our knowledge not considered before, and the presentation of simulation results. Beyond the above mentioned literature, publications showing applications of CS are available. In [B⁺10b] an application of compressed channel sensing to underwater acoustic communications is proposed. The concept is based on an OFDM (*orthogonal frequency division multiplexing*), i.e. multi-carrier, channel estimation approach. The authors provided an adaptation to the CS framework considering pilot and data symbol positions. However, in this work, only single-carrier transmission is considered. In [G⁺10] an approach for sparse wideband channel estimation was proposed. The authors used an improved sensing algorithm which combines low computational complexity and robustness for practical channel estimation. Their method is called CoSaMP. The authors managed to obtain a higher bandwidth efficiency and a lower computational complexity than preexisting concepts. In this report the CoSaMP algorithm is also applied in comparison with a convex program, to study the algorithm's behavior under the difficult THz channel conditions. The main differences to

the work of [G⁺10] are the following. The considered channel length is much larger than the one used in [G⁺10]. Moreover, due to the THz channel conditions, the channel suffers from much more noise influence and therefore the SNR is lower than the one assumed in [G⁺10].

The rest of this report is organized as follows. In Chapter 2, the used system model is described showing the composition of the used THz channel and the transmit signal structure. The process of channel estimation is going to be presented in Chapter 3. In this context, more information about compressed sensing is given in Subchapter 3.1. The used training sequence for channel sensing with requirements on the number of measurements and applied channel reconstruction algorithms with statements about their computational efficiency are shown in Subchapter 3.2 and 3.3. An evaluation of the proposed method is presented in Subchapter 3.4, containing results in terms of MSE. A conclusion and further considerations are given in Chapter 4.

2. System Model

In the process of developing a system model for the channel estimation task, two aspects have to be considered. Firstly, the channel has to be investigated closer and a channel model has to be established. Secondly, a suitable transmit signal waveform has to be designed. Both tasks are going to be discussed in Subchapter 2.1 and Subchapter 2.2.

2.1. Characteristics of THz Channel

Throughout this work, an *indoor* THz communication link is considered which leads to the channel model described in the following. Transmission over a THz channel is influenced by colored noise n_c , which is the result of molecular absorption taking place in the atmosphere. A continuous-time noise whitening filter $h_w(t)$ is used to whiten the received signal and transform n_c to i.i.d. (*independent identically distributed*) additive white Gaussian noise (AWGN). A simplified system model is depicted in Figure 2 (a), where $\{a_{TS,i}\}_{i=1}^k$ is the used training sequence

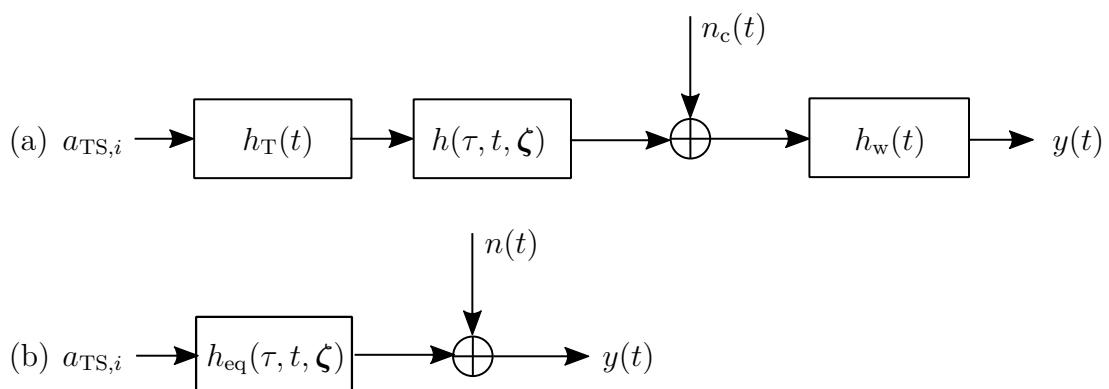


Figure 2: (a) THz channel model with colored noise n_c and whitening filter h_w .
(b) Equivalent THz channel model with i.i.d additive white Gaussian noise n .

to sense the channel, $y(t)$ denotes the continuous-time received signal, $h(\tau, t, \zeta)$ is the continuous-time CIR considering multi-path propagation characterized by delay variable τ and the scattering properties of the environment with vector ζ . According to [Ger16] this system model can be converted by introducing an equivalent continuous-time CIR $h_{\text{eq}}(\tau, t, \zeta)$, which contains the transmit filter $h_T(t)$, the channel and the continuous-time whitening filter. In Figure 2 (b) the equivalent system model, containing i.i.d. additive white Gaussian noise $n(t)$ with zero mean and variance σ^2 is shown. [M⁺16a, Ger16]

After sampling with parameters given in [M⁺16a] a discrete-time model can be obtained. However, the channel transfer function $H_{\text{eq}}(f, t, \zeta)$ is highly frequency-selective. Therefore, the total bandwidth BW is divided into N narrow, orthogonal subbands of equal width $\Delta f = \frac{BW}{N}$. For this work N is chosen as $N \in \{16, 32\}$ to ensure that each subchannel can be modeled as frequency weakly-selective. Moreover, the N -band approach is advantageous as it relaxes the requirements on current digital-to-analog converters (DAC) and analog-to-digital converters (ADC), which achieve sampling rates up to 100 Gbps. For the used system model in this work, constant scattering is assumed in each subband and N discrete-time CIRs can be obtained. This model is depicted in Figure 3. It shows the n^{th} subchannel denoted by $h_{\text{eq}_{\text{sub}_n}}[k]$ for $n \in \{1, 2, \dots, N\}$, the channel estimation part, the estimated subchannel $\hat{h}_{\text{eq}_{\text{sub}_n}}[k]$ and the final calculation of the MSE. For further considerations throughout this work, for means of simplicity $h_{\text{eq}_{\text{sub}_n}}[k]$ will in the following be referred to as $h[k]$. [M⁺16a]

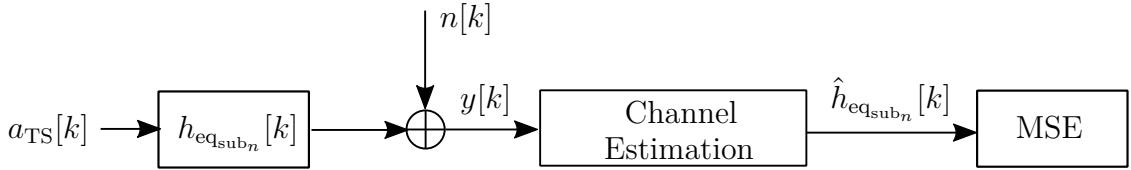


Figure 3: Whole discrete time system model, showing the channel estimation part and position of MSE calculation.

As was pointed out before, the discrete-time CIR has some peculiarities, which can also be observed in the CIR of each subband (see Figure 1). The channel can be characterized by few non-zero and mostly (nearly-)zero taps. Vectors with these properties can be characterized as follows.

1. A vector \mathbf{x} is said to be *S-sparse* if it has S non-zero taps, i.e. [FR13]

$$\|\mathbf{x}\|_0 := \text{card}(\{j : x_j \neq 0\}) \leq S.$$

2. A vector \mathbf{x} is said to be *compressible*, if it can be well approximated by sparse vectors [FR13]. A best \tilde{s} -term approximation $\mathbf{x}_{\tilde{s}}$ can be found by setting all but the \tilde{s} largest taps of \mathbf{x} to zero. It is quantified by the error of best \tilde{s} -term approximation given by [FR13]

$$\sigma_s(\mathbf{x})_p := \inf_{\|\mathbf{x}_{\tilde{s}}\|_0 \leq S} \|\mathbf{x} - \mathbf{x}_{\tilde{s}}\|_p,$$

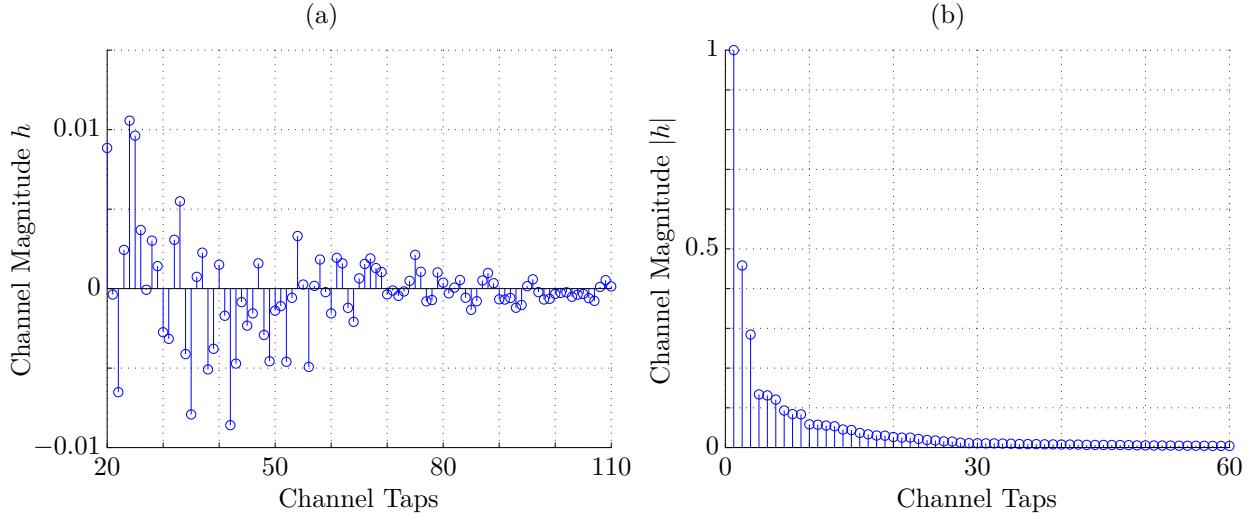


Figure 4: (a) Closer investigation of the real part discrete-time CIR, shown from tap 20 to tap 110. (b) Sorted CIR in descending order, shown for the first 60 channel taps of the real part discrete-time CIR

with $\|\cdot\|_p$ being any norm for $p > 0$ and $\mathbf{x}_{\tilde{s}}$ as an \tilde{s} -term approximation of \mathbf{x} . "A vector is *compressible* [or approximately sparse] if the error of its best \tilde{s} -term approximation decays quickly in \tilde{s} " [FR13].

3. A vector \mathbf{x} is said to be α -*compressible* with magnitude R if it obeys a specific decay structure which can be seen when the entries of \mathbf{x} are rearranged in decreasing order $|x_1|, |x_2|, \dots, |x_p|$ considering the j^{th} largest entry of the channel vector satisfies [CT07, NT09]

$$|x_j| \leq R \cdot j^{-\frac{1}{\alpha}}, \quad (1)$$

for some positive R and $\alpha \leq 1$. The smaller α , the better the \tilde{s} -term approximation, leading to an empirical error given by [NT09]

$$\|\mathbf{x} - \mathbf{x}_{\tilde{s}}\|_2 \leq C_\alpha \cdot R \cdot S^{0.5-1/\alpha},$$

for $C_\alpha = (2/p - 1)^{-1/2}$. [NT09]

The CIR \mathbf{h} considered in this work is not S -sparse. This fact is also depicted in Figure 4 (a) which illustrates the channel taps 20 - 110, showing that these channel taps are not exactly zero. Hence, the THz CIR is considered to be approximately sparse, i.e. it is said to be *compressible* and it can be well approximated by a sparse vector. According to [SB15] and [H⁺10], multi-path channels can be characterized by a specific decay structure satisfying Equation (1). This behavior is also satisfied by the THz channel, as can be seen in Figure 4 (b).

2.2. Design of Transmission Signal

To ensure a reliable channel estimation, a training sequence is necessary to sense the channel, which has to be placed suitable in the transmit signal burst. Therefore, three burst designs are

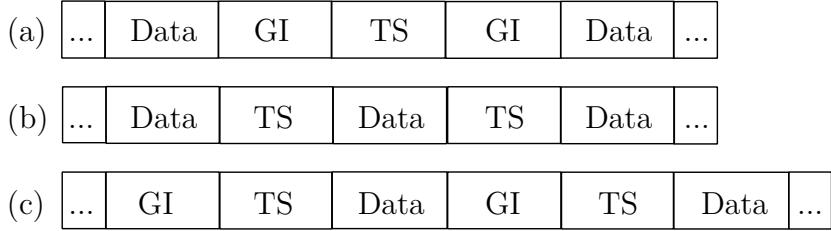


Figure 5: (a),(b),(c) Three different burst structures.

possible, which are depicted in Figure 5. Design (a) and (b) are proposed in [H⁺10]. The first concept suggests to incorporate a guard interval (GI) preceding and succeeding the training sequence, giving the possibility to use the whole training sequence for channel estimation [H⁺10]. Therefore, the length of the GI has to be longer than or equal to the subchannel length [Kam11]. However, a GI does not contain any useful information and thus, a degradation in efficiency is attained [Kam11]. Concept (b) does not consider any additional GI, leading to the fact, that not the whole training sequence can be used for channel estimation because of occurring ISI (*inter-symbol-interference*) [H⁺10]. The second concept also entails a degradation in efficiency as not all the transmitted training symbols can be used. To obtain a trade-off between both concepts, burst (c) in Figure 5 is proposed and used throughout this work, consisting of a preceding GI, the training sequence and succeeding data symbols.

3. Channel Estimation

A channel estimate $\tilde{\mathbf{h}}$ is obtained from the received signal vector $\mathbf{y} \in \mathbb{R}^{k+p-1}$, with dimension k depending on the length of the training sequence $\{a_{TS_i}\}_{i=1}^k$ and p being the channel length. As can be inferred from Figure 3 and according to [B⁺10a] the received signal vector \mathbf{y} is calculated by

$$y[k] = a_{TS}[k] * h[k] + n[k], \quad (2)$$

where $*$ is the discrete convolution operator, $\mathbf{n} \in \mathbb{R}^{n+p-1}$ denotes i.i.d. additive white Gaussian noise with zero mean and variance σ^2 and $\mathbf{h} \in \mathbb{R}^p$ being the discrete-time channel. Equation (2) can be equivalently rewritten to [B⁺10a]

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{n}, \quad (3)$$

with $\mathbf{A} \in \mathbb{R}^{(k+p-1) \times p}$ being a Toeplitz-structured convolution matrix, which is also called the *sensing matrix* as it is used to probe the channel. [B⁺10a]

Traditionally, a channel estimate is obtained by solving an LS problem. For this reason the number of measurements m has to be larger than or equal to the channel length p to get a reliable channel estimate, not taking into account any sparsity condition [H⁺10]. In contrast to LS, the main idea behind CS is the recovery of sparse signals from a relatively small number of

measurements m using a suitable reconstruction algorithm, i.e. the recovery of highly under-determined systems for $m \ll p$ [B⁺07]. This can be beneficial for THz subchannel estimation as they exhibit a highly sparse structure with just a few dominant taps (see Chapter 2.1).

When CS is applied in the case of reconstructing a signal which is contaminated by noise, there cannot be an *exact recovery* of the sparse channel, but the term used in this context is *reliable reconstruction* [Baj09]. Two main tasks arise while a CS solution is developed [FR13], namely

1. How should the *sensing matrix* \mathbf{A} be chosen?
2. What is an efficient and suitable reconstruction algorithm to reconstruct \mathbf{h} from \mathbf{y} ?

In the following, firstly, theoretical background behind CS is introduced in Subchapter 3.1. The design of the used *sensing matrix* and suitable reconstruction algorithms for signal recovery will be presented in subchapters 3.2 and 3.3. Finally, an evaluation of the used concept is shown in Subchapter 3.4.

3.1. Fundamentals of Compressed Sensing

The recovery of an S -sparse vector \mathbf{h} by using CS is said to be possible if the *sensing matrix* \mathbf{A} satisfies the *restricted isometry property* (RIP) ([H⁺10]), also called *restricted isometry hypothesis* ([CT07]). It is important to point out which condition has to be satisfied by the *sensing matrix*, to guarantee that S -sparse signals can be distinguished by their corresponding observations [NT09]. Afterwards, the matrix can be considered to sense the approximately sparse channels.

Definition 1 (RIP). *The sensing matrix $\mathbf{A} \in \mathbb{R}^{m \times p}$, with unit l_2 -normalized columns, satisfies the RIP of order $S \in \mathbb{N}$ with $\delta_S \in (0, 1)$, which is also termed $RIP(S, \delta_S)$, if*

$$(1 - \delta_S) \|\mathbf{h}\|_2^2 \leq \|\mathbf{A}\mathbf{h}\|_2^2 \leq (1 + \delta_S) \|\mathbf{h}\|_2^2 \quad (4)$$

holds for all $\mathbf{h} \in \mathbb{R}^p$ with no more than S non-zero taps. S denotes the sparsity of the channel and δ_S is the restricted isometry constant defining the smallest value δ_S which satisfies the inequalities in (4). [H⁺10, B⁺10a]

The RIP constraint must hold for all subsets $T \subset \{1, 2, \dots, p\}$, with \mathbf{A}_T being the $m \times |T|$ sub-matrices, with $|T| \leq S$, obtained by extracting the columns of \mathbf{A} corresponding to the indices in T [CT07]. It implies all singular values of all submatrices of \mathbf{A} , which are formed by every $|T| \leq S$ columns of \mathbf{A} , to be in the range $(\sqrt{1 - \delta_S}, \sqrt{1 + \delta_S})$ [H⁺10]. Moreover, another consequence is that disjoint sets of columns from \mathbf{A} are nearly orthogonal [NT09].

It is important that \mathbf{A} acts like a bijective structure-preserving map, i.e. S -sparse signals should be distinguished based on their samples [NT09]. Therefore, each $2S$ columns need to form a linearly independent set, i.e. need to be well conditioned and nonsingular. To achieve this, the authors of [CT07] proposed Equation (4) to be satisfied for δ_{2S} , i.e. $RIP(2S, \delta_{2S})$ has to hold. As a result, \mathbf{A} can be assumed to behave almost like an isometry for sparse vectors with S non-zero taps, i.e. it preserves the geometry in terms of the Euclidean metric of the set of all

S -sparse signals, if we assume $\mathbf{y} = \mathbf{A}\mathbf{h}$ to be a mapping from a higher dimensional space $\mathbf{h} \in \mathbb{R}^p$ to a lower dimensional space $\mathbf{y} \in \mathbb{R}^m$ with the minimum number of measurements satisfying at least $m \geq 2S$ [H⁺10, NT09].

3.2. Design of Sensing Matrix

To use the CS method successfully, a *sensing matrix* is needed which satisfies Equation (4) and the columns of \mathbf{A} are required to be linearly independent. No algorithm is known to date that can verify the RIP property for a given matrix in polynomial time. Nevertheless, there exist certain probabilistic constructions of matrices that are known to satisfy RIP with high probability and that justify the usage of CS [Baj09].

One kind of suitable matrices are i.i.d. matrices [Baj09]. In this case, the lower bound for the number of measurements m can be given by [H⁺10]

$$m \geq C_2 \cdot S \cdot \log p, \quad (5)$$

where C_2 is a universal constant depending on the used algorithm. This bound is sufficient to guarantee RIP for i.i.d. random sensing matrices.

Unfortunately, for a system identification task, a completely i.i.d. structured matrix cannot be applied. Instead, by using Equation (3), a Toeplitz-structured matrix is needed. To be useful in the context of CS, this matrix has to satisfy $\text{RIP}(2S, \delta_{2S} = 0.3)$, which is sufficient to guarantee reliable reconstruction. The Toeplitz-structure of a *sensing matrix* introduces statistical dependencies among observations, i.e. the rows of the matrix. However, in [H⁺10] and [Baj09] it was proven, that random Toeplitz matrices also satisfy the RIP condition and can be considered as CS-matrices for channel sensing. As the prove of this property is out of scope of this report, we refer the interested reader to [H⁺10] and [Baj09].

In the context of Toeplitz-structured matrices Equation (5) might be suboptimal, hence in [Baj09] the following number of measurements is proposed

$$m \geq C_2 \cdot S^2 \cdot \log p. \quad (6)$$

Throughout this work *stochastic* BPSK transmission is assumed for the transmission of training symbols, i.e. $\{a_i\}_{i=1}^k$ is given by a sequence of i.i.d. binary random variables, taking values +1 and -1 with probability 0.5 each. Thus, the elements of the training sequence are independent symmetric Bernoulli or Rademacher random variables [H⁺10]. The transmission signal structure, which is used in this work (see Figure 5 (c)) leads to a matrix \mathbf{A} with $\{a_{TS_i}\}_{i=1}^k$ being the sequence of training symbols, $\{x_i\}_{i=1}^{p-1}$ representing the succeeding sequence of the first $p-1$ data symbols and a preceding sequence of zeros ([H⁺10]), which is used as GI and is assumed to have length $p-1$. Therefore, $\mathbf{A} \in \mathbb{R}^{(k+p-1) \times p}$. However, the *full* Toeplitz matrix \mathbf{A}_{full} considered in the literature [Baj09] and [H⁺10] which was proved to satisfy $\text{RIP}(S, \delta_S = 0.3)$ was constructed differently. Both matrices \mathbf{A} and \mathbf{A}_{full} and their structures can be seen in the following.

$$\mathbf{A} = \begin{pmatrix} a_{TS_1} & 0 & 0 & \cdots & \cdots & 0 \\ a_{TS_2} & a_{TS_1} & 0 & 0 & \cdots & 0 \\ a_{TS_3} & a_{TS_2} & a_{TS_1} & 0 & \cdots & 0 \\ a_{TS_4} & a_{TS_3} & a_{TS_2} & a_{TS_1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ a_{TS_p} & a_{TS_{p-1}} & \cdots & a_{TS_3} & a_{TS_2} & a_{TS_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{TS_k} & a_{TS_{k-1}} & \cdots & a_{TS_{p-1}} & a_{TS_{p-2}} & a_{TS_{p-3}} \\ x_1 & a_{TS_k} & a_{TS_{k-1}} & \cdots & a_{TS_{p-1}} & a_{TS_{p-2}} \\ x_2 & x_1 & a_{TS_k} & x_{TS_{k-1}} & \cdots & a_{TS_{p-1}} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ x_{p-2} & x_{p-3} & \ddots & x_1 & a_{TS_k} & a_{TS_{k-1}} \\ x_{p-1} & x_{p-2} & \ddots & \ddots & x_1 & a_{TS_k} \end{pmatrix},$$

$$\mathbf{A}_{\text{full}} = \begin{pmatrix} a_{TS_1} & 0 & 0 & \cdots & \cdots & 0 \\ a_{TS_2} & a_{TS_1} & 0 & 0 & \cdots & 0 \\ a_{TS_3} & a_{TS_2} & a_{TS_1} & 0 & \cdots & 0 \\ a_{TS_4} & a_{TS_3} & a_{TS_2} & a_{TS_1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ a_{TS_p} & a_{TS_{p-1}} & \cdots & a_{TS_3} & a_{TS_2} & a_{TS_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{TS_k} & a_{TS_{k-1}} & \cdots & a_{TS_{p-1}} & a_{TS_{p-2}} & a_{TS_{p-3}} \\ 0 & a_{TS_k} & a_{TS_{k-1}} & \cdots & a_{TS_{p-1}} & a_{TS_{p-2}} \\ 0 & 0 & a_{TS_k} & x_{TS_{k-1}} & \cdots & a_{TS_{p-1}} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & a_{TS_k} & a_{TS_{k-1}} \\ 0 & 0 & \ddots & \ddots & 0 & a_{TS_k} \end{pmatrix}.$$

$\mathbf{A}_{\text{full}} \in \text{RIP}(S, \delta_S)$ with probability exceeding $1 - \exp(-c_1 \cdot k/S^2)$ for any $\delta_S \in (0, 1)$ whenever $k \geq c_2 \cdot S^2 \log p$, with constants $c_1 < \delta_S^2$ and $c_2 = 12/(\delta_S^2 - 4c_1)$ which do only depend on δ_S [Baj09]. Moreover, in [B⁺10a], \mathbf{A}_{full} was said to satisfy $\text{RIP}(2S, \delta_{2S} = 0.3)$ with therefore slightly changed bounds and constants. Thus, the full Toeplitz-structured matrix $\mathbf{A}_{\text{full}} \in \text{RIP}(2S, \delta_{2S})$ with probability exceeding $1 - \exp(-\frac{c_1 \cdot k}{4S^2})$ for any $\delta_{2S} \in (0, 1/3)$ whenever $k \geq c_2 \cdot S^2 \log p$, with constants $c_1 \in (0, \delta_{2S}^2/4)$ and $c_2 = 48/(\delta_{2S}^2 - 4c_1)$ which do only depend on δ_S [B⁺10a, H⁺10].

As can be seen from matrix \mathbf{A} , the full matrix can not be used for channel estimation. The last $p - 1$ observations do also contain contributions from data, which is *unknown* to the

receiver [H⁺10]. Hence, the *sensing matrix* is $\mathbf{A}_{\text{TS}} \in \mathbb{R}^{k \times p}$

$$\mathbf{A}_{\text{TS}} = \begin{pmatrix} a_{\text{TS}_1} & 0 & 0 & \cdots & \cdots & 0 \\ a_{\text{TS}_2} & a_{\text{TS}_1} & 0 & 0 & \cdots & 0 \\ a_{\text{TS}_3} & a_{\text{TS}_2} & a_{\text{TS}_1} & 0 & \cdots & 0 \\ a_{\text{TS}_4} & a_{\text{TS}_3} & a_{\text{TS}_2} & a_{\text{TS}_1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ a_{\text{TS}_p} & a_{\text{TS}_{p-1}} & \cdots & a_{\text{TS}_3} & a_{\text{TS}_2} & a_{\text{TS}_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{\text{TS}_k} & a_{\text{TS}_{k-1}} & \cdots & a_{\text{TS}_{p-1}} & a_{\text{TS}_{p-2}} & a_{\text{TS}_{p-3}} \end{pmatrix}.$$

According to the structure of the *sensing matrix*, the number of measurements (observations) m taken, equals the training signal length k , i.e. $m = k$. In general it holds for CS theory that as long as the measurements satisfy Equation (4), their number needs only to commensurate with the complexity of the considered signal and as a consequence, preserve its information content [C⁺10].

If matrix \mathbf{A}_{full} satisfies $\text{RIP}(2S, \delta_{2S})$, there is no abuse in the RIP condition if only the first k observations are taken for channel estimation. Therefore \mathbf{A}_{TS} can be used for application of CS theory.

3.3. Presentation of Reconstruction Algorithms

As was said before, LS channel estimation is incapable of taking into account the sparsity of the considered THz channel. Therefore, the consequences of this condition are going to be shown in Subchapter 3.3.1.

A channel estimate $\hat{\mathbf{h}}$ for noiseless transmission, taking into account the sparsity of the channel \mathbf{h} , obtained from $\mathbf{y} = \mathbf{A}_{\text{TS}}\mathbf{h}$ can be given by [Baj09]

$$\begin{aligned} \hat{\mathbf{h}} &= \underset{\mathbf{h} \in \mathbb{R}^p}{\operatorname{argmin}} \quad \|\tilde{\mathbf{h}}\|_0 \\ \text{subject to} \quad \mathbf{A}_{\text{TS}}\tilde{\mathbf{h}} &= \mathbf{y}. \end{aligned}$$

Unfortunately, an l_0 -optimization problem is NP-hard. This problem can be overcome by using convex relaxation, leading to the following convex optimization problem also called *basis pursuit* [Baj09]

$$\begin{aligned} \hat{\mathbf{h}} &= \underset{\mathbf{h} \in \mathbb{R}^p}{\operatorname{argmin}} \quad \|\tilde{\mathbf{h}}\|_1 \\ \text{subject to} \quad \mathbf{A}_{\text{TS}}\tilde{\mathbf{h}} &= \mathbf{y}. \end{aligned}$$

The global optimum of this convex problem can be determined with limited computational complexity, since it can be recast as a linear program.

However, transmission signals in reality are contaminated with noise, i.e. Equation (3) holds. Therefore, in subchapters 3.3.2 and 3.3.3 two methods are explained, how noisy observations can

be used to obtain a channel estimate $\hat{\mathbf{h}}$. The first approach is a solution to a convex program, which is characterized by stability, high estimation accuracy but long computation time [G⁺¹⁰]. The second algorithm that is proposed is the CoSaMP approach, which combines the advantages of a convex program and a greedy algorithm, and is therefore additionally characterized by a low implementation effort. [G⁺¹⁰]

3.3.1. Least Squares Approach

For the sake of completeness, a short overview of LS channel estimation should be given.

An LS channel estimate can be obtained by [H⁺¹⁰]

$$\hat{\mathbf{h}}_{\text{LS}} = (\mathbf{A}_{\text{TS}}^H \mathbf{A}_{\text{TS}})^{-1} \mathbf{A}_{\text{TS}}^H \mathbf{y}. \quad (7)$$

If the training sequence $a_{\text{TS}_i} \in \{+1, -1\}$ for $i = 1, 2, \dots, k$ with probability 0.5 each, the MSE lower bound of an LS channel estimate for some $k \geq p$ can be given by [H⁺¹⁰]

$$\mathbb{E}[\|\hat{\mathbf{h}}_{\text{LS}} - \mathbf{h}\|_2^2] \geq \frac{p\sigma^2}{k}. \quad (8)$$

For comparison, an ideal *oracle* equipped LS channel estimation approach is analyzed closer. This *oracle* tells us the indices of the non-zero entries of \mathbf{h} . This ideal estimation strategy cannot be used in practice, but gives an idea about the gain, that could be achieved by LS estimation if knowledge about sparsity could be incorporated. If $T_* \subset \{1, 2, \dots, p\}$ is the set of indices of the non-zero entries of \mathbf{h} and the sparsity pattern of the S non-zero entries is known an optimal LS estimate can be obtained for some $k \geq S$ by [H⁺¹⁰]

$$\hat{\mathbf{h}}_{\text{LS}_*} = (\mathbf{A}_{\text{TS}_*}^H \mathbf{A}_{\text{TS}_*})^{-1} \mathbf{A}_{\text{TS}_*}^H \mathbf{y}, \quad (9)$$

with \mathbf{A}_{TS_*} being the *sensing matrix* with columns defined by the indices T_* . The MSE lower bound of this *oracle* based estimator is then given by [H⁺¹⁰]

$$\mathbb{E}[\|\hat{\mathbf{h}}_{\text{LS}_*} - \mathbf{h}\|_2^2] \geq \frac{S^2\sigma^2}{\text{trace}\{\mathbf{A}_{\text{TS}_*}^H \mathbf{A}_{\text{TS}_*}\}} = \frac{S\sigma^2}{k}. \quad (10)$$

Comparing Equation (8) and Equation (10), the disadvantage of LS channel estimation becomes clearly visible. If the sparsity assumption is not incorporated, the lower bound gets much higher, since $S \geq k$ [H⁺¹⁰].

3.3.2. Convex Program

To obtain a channel estimate $\hat{\mathbf{h}}$ for a more realistic and therefore noisy case, i.e. from Equation (3), various works like [H⁺¹⁰, B⁺⁰⁷, FR13] and [B^{+10a}] propose to use the *Dantzig selector* (DS).

The DS guarantees near optimal reconstruction based on the RIP characterization of the *sensing matrix* [Baj09]. Moreover, it is said to have the following advantages [B^{+10a}, CT07]

- It is one of few reconstruction algorithms in CS that are guaranteed to perform almost optimally in the presence of stochastic noise.
- It is computationally tractable, as it can be reformulated to a linear program.
- There are clean and most interpretable reconstruction error bounds for sparse and approximately sparse signals available.

The convex problem in form of a DS can be formulated as [H⁺10]

$$\begin{aligned} \hat{\mathbf{h}} = \operatorname{argmin}_{\tilde{\mathbf{h}} \in \mathbb{R}^p} & \|\tilde{\mathbf{h}}\|_1 \\ \text{subject to } & \|\mathbf{A}_{\text{TS}}^H(\mathbf{y} - \mathbf{A}_{\text{TS}}\tilde{\mathbf{h}})\|_\infty \leq \sigma\lambda\|\mathbf{A}_{\text{TS}}\|_{\infty,2}, \end{aligned} \quad (11)$$

with $\|\mathbf{A}_{\text{TS}}\|_{\infty,2}$ denoting the largest l_2 norm of the columns of \mathbf{A}_{TS} . In case of unit 2 normalized columns, this is equivalent to $\|\mathbf{A}_{\text{TS}}\|_{\infty,2} = 1$. If the DS is recast and solved as a linear program, the convex optimization problem can be described by [CT07]

$$\begin{aligned} \text{minimize } & \sum_i u_i \\ \text{subject to } & -\mathbf{u} \leq \tilde{\mathbf{h}} \leq \mathbf{u} \\ & -\mathbf{1}\sigma\lambda\|\mathbf{A}_{\text{TS}}\|_{\infty,2} \leq \mathbf{A}_{\text{TS}}^H(\mathbf{y} - \mathbf{A}_{\text{TS}}\tilde{\mathbf{h}}) \leq \mathbf{1}\sigma\lambda\|\mathbf{A}_{\text{TS}}\|_{\infty,2}. \end{aligned} \quad (12)$$

The parameter λ is chosen as [B⁺10a, H⁺10]

$$\lambda(p) = \sqrt{2(1+a)\log p}, \quad \text{for any } a \geq 0 \quad (13)$$

and if a real signal is assumed ([Baj09]), the following upper bound for the squared error of the DS estimate $\tilde{\mathbf{h}}$ for an S -sparse signal can be stated [Baj08]

$$\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2 \leq C_2 \cdot \log p \left(\frac{S\sigma^2}{k} \right), \quad (14)$$

where $C_2 = 8/(1 - \delta_S - \theta_{S,2S})$ and for very small values of $\delta_S + \theta_{S,2S}$, $C_2 \approx 8$ and $\theta_{S,2S}$ being the $S, 2S$ -restricted orthogonality constant defined as [CT07]

$$|\langle \mathbf{A}_{\text{TS}_T} \mathbf{h}, \mathbf{A}_{\text{TS}_{T'}} \mathbf{h}' \rangle| \leq \theta_{S,2S} \cdot \|\mathbf{h}\|_2 \|\mathbf{h}'\|_2,$$

with \mathbf{A}_{TS_T} and $\mathbf{A}_{\text{TS}_{T'}}$ being subsets $T \neq T'$ of matrix \mathbf{A}_{TS} (see Chapter 3.1). The squared error upper bound in Equation (14) is achieved with probability of success exceeding [Baj08, CT07]

$$1 - \left(\sqrt{\pi \log p} \cdot p^a \right)^{-1}. \quad (15)$$

Equivalent results can be achieved for complex-valued signals. [H⁺10] and [Baj09, B⁺10a] derived the following theorem for complex-valued signals with $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2)$ [Baj09].

Theorem 1. *Assume the sensing matrix \mathbf{A}_{TS} has unit 2-normalized columns and satisfies RIP ($2S, \delta_{2S} = 0.3$) [Baj09] for some $S \in \mathbb{N}$. The DS estimate achieves*

$$\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2 \leq C_4 \min_{1 \leq \tilde{s} \leq S} \left(\sigma\lambda\sqrt{\tilde{s}} + \frac{\|\mathbf{h} - \mathbf{h}_{\tilde{s}}\|_1}{\sqrt{\tilde{s}}} \right)^2, \quad (16)$$

with probability exceeding

$$1 - 2 \max \left\{ \left(\sqrt{\pi(1+a) \log p} \cdot p^a \right)^{-1}, \Pr(\mathbf{A}_{\text{TS}} \notin \text{RIP}(2S, \delta_{2S} = 0.3)) \right\}, \quad (17)$$

where $\mathbf{h}_{\tilde{s}}$ is the best \tilde{s} -term approximation of \mathbf{h} and $C_4 = 16/(1 - 3\delta_{2S})^2$.

The squared error in Equation (16) has the following two components. [Baj09]

- An *estimation error* (variance) occurs, when \tilde{s} unknown taps are determined from noisy observations.
- An *approximation error* (bias) is caused by estimating an unknown signal using \tilde{s} components only.

Theorem 2 can be adapted to a channel vector exhibiting an S -sparse structure or to a channel vector showing an approximately sparse behavior and a specific decay structure [H⁺10]. As the THz channel is approximately sparse and assumed to be α -compressible, only this case should be considered, for more information about S -sparse signals, the interested reader is referred to [H⁺10].

If an α -compressible signal is assumed, showing a specific decay structure according to Equation (1) and if the approximately sparse signal is additionally influenced by noise, the DS is only able to give a reliable estimate of taps which are significantly above the noise level, i.e. $|h_j| \geq \sigma$. Therefore, \mathbf{h} can be treated as having no more than $(\frac{R}{\sigma})^{(1/\alpha)}$ non-zero taps [SB15]. Moreover, if $S = |\{j : |h_j| > \sigma\}|$, then $\delta_{2S} + \theta_{S,2S} < 1$ has to hold, otherwise, there might be not enough observations to get a reliable estimate [CT07]. For $m \geq C_5 (\log p)^{\frac{2\beta-3}{2\beta-1}} (\sigma^2)^{\frac{-2}{2\beta-1}}$ and $\beta > 0.5$, the reconstruction error, i.e. Equation (16) can be reformulated to [H⁺10]

$$\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2 \leq C_0 (\log p)^{\frac{2\beta}{2\beta+1}} \cdot \left(\frac{\sigma^2}{k} \right)^{\frac{2\beta}{2\beta+1}} \quad (18)$$

and the *sensing matrix* has to satisfy $\text{RIP}(2S_*, 0.3)$ with $S_* \geq (\frac{R}{\sigma})^{(1/\alpha)}$ and $\delta_{2S_*} + \theta_{S_*,2S_*} < 1$ [SB15]. Equation (18) is the minimax rate if a decay structure is observed. For $S_* \leq (\frac{R}{\sigma})^{(1/\alpha)}$ the DS saturates because there is not enough data to recover the minimax rate, and a squared loss of only $\mathcal{O}(\log p)(R^2 S_*^{-2/\alpha+1} + S_* \cdot \sigma^2)$ can be achieved. In either case, the error is well controlled. [CT07, H⁺10]

In comparison to the LS channel estimation approach presented in Subchapter 3.3.1, the DS turns out to be computationally more efficient. The MSE lower bound $\frac{p\sigma^2}{k}$ for LS estimation given in Equation (8) is also valid under decay assumption. If the $\log p$ factor in Equation (18) is ignored, the estimation error behaves like $\mathcal{O}\left(\frac{\sigma^2}{k}\right)^{\frac{2\beta}{2\beta+1}}$. In comparison to the MSE lower bound, even in the case of approximately sparse channels, the DS channel estimate shows an improvement by a factor of (roughly) $n \cdot \left(\frac{\sigma^2}{k}\right)^{\frac{1}{2\beta+1}}$ over the LS MSE [H⁺10].

3.3.3. Compressive Sampling Matching Pursuit Algorithm

CoSaMP is based on a *greedy pursuit* method, i.e. "[it builds] up an approximation one step at a time by making locally optimal choices at each step." [NT09]

For a channel estimation task, the algorithm has the following advantages [NT09].

- CoSaMP is able to reconstruct the channel from minimal number of observations m .
- CoSaMP guarantees a robust signal recovery if the channel is contaminated with noise and if the signal is not exactly S -sparse.
- CoSaMP provides provably low computational cost and requires relatively small storage.

The CoSaMP procedure is described in Algorithm 1, for $\mathbf{A}_{\text{TS}}^\dagger = (\mathbf{A}_{\text{TS}}^H \mathbf{A}_{\text{TS}})^{-1} \mathbf{A}_{\text{TS}}^H$, $\mathbf{b}|_T$ being the signal restricted to the set of indices T , $\mathbf{b}|_{T^c}$ being the signal components which are not in the set T and \mathbf{A}_{TS_T} as the restriction of the *sensing matrix* to the column submatrix, with columns corresponding to the index set T . A required input to the algorithm is the sparsity S , which might be unknown a priori. There are two possibilities to overcome this problem. One method is to estimate S from Equation (6) to get $S \approx \sqrt{m/C_2 \log p}$. Another method might be to run the algorithm using various different S and select the best approximation $\hat{\mathbf{h}}_{\tilde{s}}$ by considering an empirical error metric e.g. given by [NT09]

$$\|\mathbf{A}_{\text{TS}} \hat{\mathbf{h}}_{\tilde{s}} - \mathbf{y}\|_2. \quad (19)$$

The algorithm can be described shortly in the following five steps [G⁺10, NT09, Stu11].

1. Firstly, dominant taps have to be identified. Therefore, in iteration i , a signal proxy \mathbf{p} is defined. A thresholding function Th_1 with threshold $\tau_{i,1} \geq 0$ is applied on each element of

Algorithm 1. CoSaMP Recovery Algorithm

```

Input : Sensing matrix  $\mathbf{A}_{\text{TS}}$ , a vector of noisy observations  $\mathbf{y}$  and sparsity  $S$ 
Output : An  $\tilde{s}$ -sparse approximation  $\hat{\mathbf{h}}_{\tilde{s}}$  of  $\mathbf{h}$ 
 $\hat{\mathbf{h}}_{\tilde{s}}^0 \leftarrow 0, \mathbf{r} \leftarrow \mathbf{y}, i \leftarrow 0$  /* Initialisation */
repeat
     $i \leftarrow i + 1$ 
     $\mathbf{p} \leftarrow \mathbf{A}_{\text{TS}}^H \mathbf{r}$  /* Form signal proxy */
     $W \leftarrow \text{supp}\{\text{Th}_1(|\mathbf{p}|; \tau_{i,1})\}$  /* Identify large components */
     $T \leftarrow W \cup \text{supp}\{\hat{\mathbf{h}}_{\tilde{s}}^{i-1}\}$  /* Merge supports */
     $\mathbf{b}|_T \leftarrow \mathbf{A}_{\text{TS}_T}^\dagger \mathbf{y}$  /* Signal estimation */
     $\mathbf{b}|_{T^c} \leftarrow 0$ 
     $\hat{\mathbf{h}}_{\tilde{s}}^i \leftarrow \text{supp}\{\text{Th}_2(|\mathbf{b}|; \tau_{i,2})\}$  /* Prune approximation */
     $\mathbf{r} \leftarrow \mathbf{y} - \mathbf{A}_{\text{TS}} \hat{\mathbf{h}}_{\tilde{s}}^i$  /* Update residual */
until some stopping criterion true

```

\mathbf{p} such that $2S$ maximum dominant taps are found. The positions of these dominant taps are denoted by index set W . It is merged with the index set of the dominant taps of $\hat{\mathbf{h}}_{\tilde{s}}^{i-1}$ also taken for the current approximation of $\hat{\mathbf{h}}_{\tilde{s}}^i$. The merged index set is denoted by T .

2. Secondly, CoSaMP calculates an LS channel estimate on the index positions T .
3. Non-dominant taps, i.e. taps, whose indices are not contained in the index set T are set to 0.
4. A new approximation $\hat{\mathbf{h}}_{\tilde{s}}^i$ is produced, by retaining the largest entries in the LS signal approximation. Therefore, a second thresholding function Th_2 is used and $\tau_{i,2} \geq 0$ is set such that a maximum of S elements of \mathbf{b} is retained.
5. Finally, the new residual is computed. This procedure is repeated until some stopping criterion holds. In [G⁺10] the proposed criterion is $\{i : i \geq 4S \mid \|\hat{\mathbf{h}}_{\tilde{s}}^i - \mathbf{h}\|_2^2 \leq 10^{-4}\}$. Other criteria can be found in [NT09].

If RIP($2S, \delta_{2S} \leq \sqrt{2} - 1$) is satisfied, the CoSaMP channel estimate achieves a squared error bound of [G⁺10]

$$\|\mathbf{h} - \hat{\mathbf{h}}_{\tilde{s}}\|_2 \leq C \cdot \max \left\{ \eta, 1/\sqrt{S} \|\mathbf{h} - \mathbf{h}_{2\tilde{s}}\|_1 + \|\mathbf{n}\|_2 \right\}, \quad (20)$$

with \mathbf{n} as the i.i.d. additive white Gaussian noise, an universal constant C , a given precision parameter η , $\mathbf{h}_{2\tilde{s}}$ being the best $2\tilde{s}$ -sparse approximation of \mathbf{h} and $\tilde{\mathbf{h}}_{\tilde{s}}$ being the output of the CoSaMP algorithm [G⁺10]. According to [NT09], CoSaMP performs best for sparse and *compressible* signals, producing a squared error bound for α -*compressible* signals of [G⁺10]

$$\|\mathbf{h} - \hat{\mathbf{h}}_{\tilde{s}}\|_2 \leq C \cdot \max \left\{ \eta, R \cdot S^{-2/\alpha+1} + \|\mathbf{n}\|_2 \right\}, \quad (21)$$

for $\eta = R \cdot S^{-2/\alpha+1}$. In comparison with the convex optimization algorithm proposed in Subchapter 3.3.2, CoSaMP has similar properties, however, it is computationally more efficient, which can be seen in Table 1. For both methods, the *sensing matrix* has to satisfy a RIP constraint to successfully recover S -sparse signals. Both concepts are able to reconstruct the channel estimate from a near optimal number of measurements, i.e. depending on the Toeplitz-structured *sensing matrix*, $m = \mathcal{O}(S^2 \log p)$. Both reconstruction algorithms are stable, even if the signal is *compressible* and contaminated by noise. The main difference between both methods, pointed out by [NT09], is the computational effort. For a convex optimization, LP(p, m) is the cost for solving a linear program with p variables and m constraints. If an interior point method is chosen, the complexity is $\mathcal{O}(m^2 p^{1.5})$. In contrast, for CoSaMP $\mathcal{O}(mp)$ is sufficient [NT09].

Table 1: Comparison of proposed signal recovery concepts

	CoSaMP	Convex Optimization
General Observations	RIP	RIP
Optimal # Observations	Yes	Yes
Stability	Yes	Yes
Running Time	$\mathcal{O}(mp)$	$\text{LP}(p, m)$

3.4. Evaluation of Proposed Concept

The error metric used throughout this work to evaluate the results is the MSE, which is defined as [G⁺¹⁰]

$$\text{MSE} = \frac{1}{R} \sum_{r=1}^R \|\mathbf{h} - \hat{\mathbf{h}}^r\|_2^2, \quad (22)$$

with $\tilde{\mathbf{h}}^r$ being the estimated channel after the r^{th} realization for in total R realizations.

The THz channel is complex having a real and an imaginary part. Therefore, channel estimation is first carried out for the real part and afterwards, for the imaginary part. Both components are divided into $N = \{16, 32\}$ subbands resulting in a subchannel length $p = \{3223, 1585\}$. The noise variance was taken from previous *indoor* THz channel modeling. All individual subchannel characteristics are collected in Appendix A. The channel estimate $\hat{\mathbf{h}}$ is obtained according to Figure 3. Throughout the simulations the channel estimation strategies presented in the previous chapters are going to be compared. Therefore CS, LS, CoSaMP and optLS performances are analyzed, with optLS being the *oracle* based LS approach. All reconstruction algorithms are first applied on the real part of the channel, afterwards they are taken to estimate the imaginary part. Each algorithm is performed using 1000 realizations for MSE calculation if not stated differently.

For application of the DS, the following parameters where assumed. The probability of success, according to Equation (15), is chosen as 0.999. This results in $a = 0.4684$ and $\lambda(p) = \{3.0656, 3.2098\}$ (see Equation (13)). The solution to the convex program is obtained by using *CVX*, which uses an interior point method. The DS as a stable convex optimization program was performed for training sequence length of $[1 : 200 : 2000]$ for $p = 1585$ and because of computational complexity, additionally, only a training sequence length of 3000 could be analyzed. Similar reasons hold for $p = 3223$. There, a training sequence length of $[1 : 200 : 2000]$ was used and additionally lengths of 2500, 3000, 3500, and 4000 were taken. The *sensing matrix* \mathbf{A}_{TS} is designed by using *stochastic* BPSK, i.e. the training sequence $\{a_i\}_{i=1}^k$ is given by a sequence of i.i.d. binary random variables, taking values +1 and -1 with probability 0.5 each. Because of computational complexity, the MSE is obtained for only 250 realizations. For each realization, the same *sensing matrix* is used.

When CoSaMP is applied, the sparsity S , which is needed as an input to the algorithm, is not known a priori. Therefore, like it was proposed in Chapter 1, different S where tried, by considering an empirical error metric given by Equation (19). The best estimate was finally chosen and used for MSE calculation. The stopping criterion given in [G⁺¹⁰] was also beneficial for the THz channel and therefore used in the simulations. The CoSaMP as a thresholding approach was performed for a training sequence length of $[200 : 200 : 2000]$ for $p = 1585$ and for a training sequence length of $[200 : 200 : 5000]$ for $p = 3223$. The algorithm appears to work unstable for a low number of observations, therefore $m = 200$ was chosen to begin with (see also [G⁺¹⁰]). The used code can be found on http://media.aau.dk/null_space_pursuits/2011/08/cosamp-and-cosaomp.html.

The *oracle* based LS, called optLS, is applied to the most dominant taps of the subchannels.

The subchannels exhibit different channel energies, therefore a thresholding method was applied to find the most dominant taps. The *oracle* based LS estimate is optimal and not achievable in practical systems. However, it can be seen as a benchmark performance.

LS estimation has a significant disadvantage for $m < p$ as there are no reliable channel estimates in this case. Therefore, LS was considered only for $m \geq p$.

First, the proposed methods are going to be used for the estimation of all taps within the whole subchannel length p . As the main energy is concentrated in the first channel taps (see Appendix B), afterwards, these methods should be applied for the first 1000 channel taps of each subchannel, in order to get either performance improvement in MSE or better insight into the influence of the channel energy on the MSE contained in the last taps. Furthermore, an analysis will be given, if a new random *sensing matrix* for each realization enhances the DS estimation compared to using the same random *sensing matrix* for each realization.

3.4.1. Channel Estimation for the Entire Subchannel Length

For the first channel estimation tasks using DS, LS, CoSaMP and optLS, the taps within the whole length p of each subchannel should be estimated. In the beginning, results for the first four subchannels out of 32 THz subchannels will be shown. Afterwards, the first four subchannels out of 16 THz subchannels will be analyzed. The first four subchannels were chosen, as they show the highest SNR (see Appendix A).

32 Subchannels

As was said before, the THz channel is complex. Therefore, firstly, the real part of the discrete-time CIR is estimated for $p = 1585$, leading to the results depicted in Figures 6 - 9.

All Figures 6 - 9 (a) show a poor performance of LS estimation in terms of MSE. The DS and the CoSaMP algorithm perform much better, even if there are more measurements m available than the channel length p . An advantage of both is, that they show reasonable results for $m < p$. With both concepts, the achieved MSE is close to the *oracle* based optLS. However, CoSaMP performs slightly better than the DS for subbands 2 and 3 (see Figures 7 (b) and 8 (b)) and for a higher number of measurements also for subbands 1 and 4 (see Figures 6 (b) and 9 (b)). A major advantage of CoSaMP is the computational efficiency and running time. The algorithm terminates much faster than *CVX*, which uses an interior point method.

Figures 6 (b) and 9 (b) show similar behavior of DS and CoSaMP in terms of MSE. For less observations m , the DS shows a better performance than CoSaMP. In [G⁺10], this property was explained as follows. The DS as a convex optimization algorithm can be transformed to a linear program to be solved stably. In contrast to this, CoSaMP is a ‘support set estimation’ which uses hard thresholding by selecting S dominant taps of the channel vector in each iteration (see also Chapter 3.3.3, Algorithm 1). If the training sequence is very short, the CoSaMP algorithm appears to be unstable on the range of hard thresholding. However, Figure 7 (b) and 8 (b) do not show this behavior, which indicates that the thresholding parameters for CoSaMP were more suitable for these subbands.

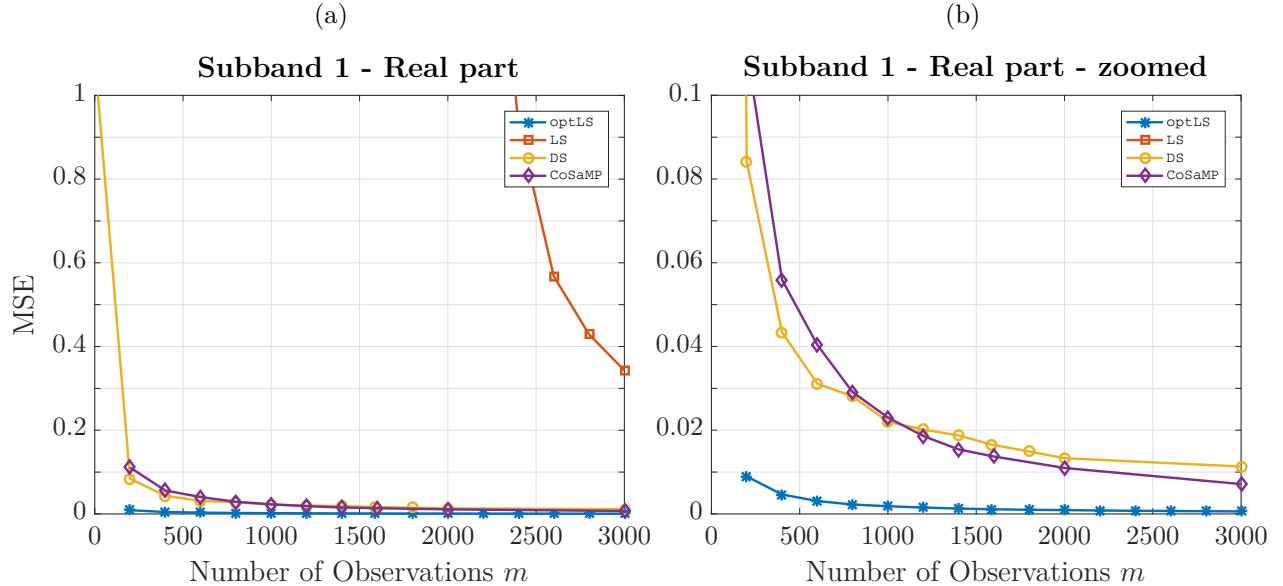


Figure 6: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of **32** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

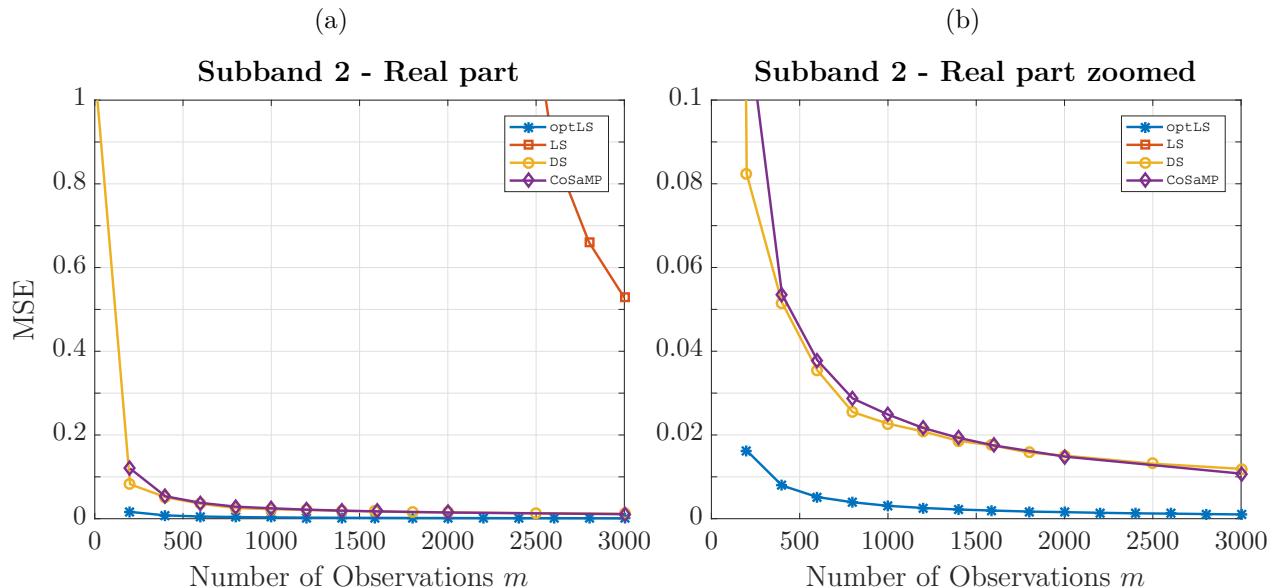


Figure 7: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of **32** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

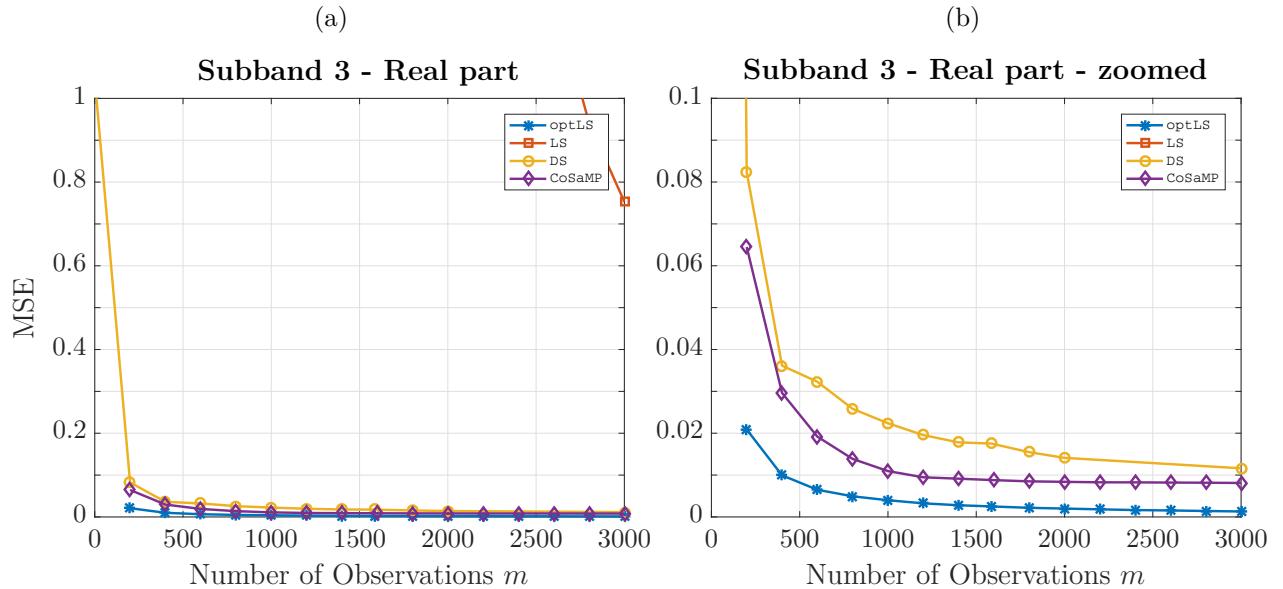


Figure 8: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of **32** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

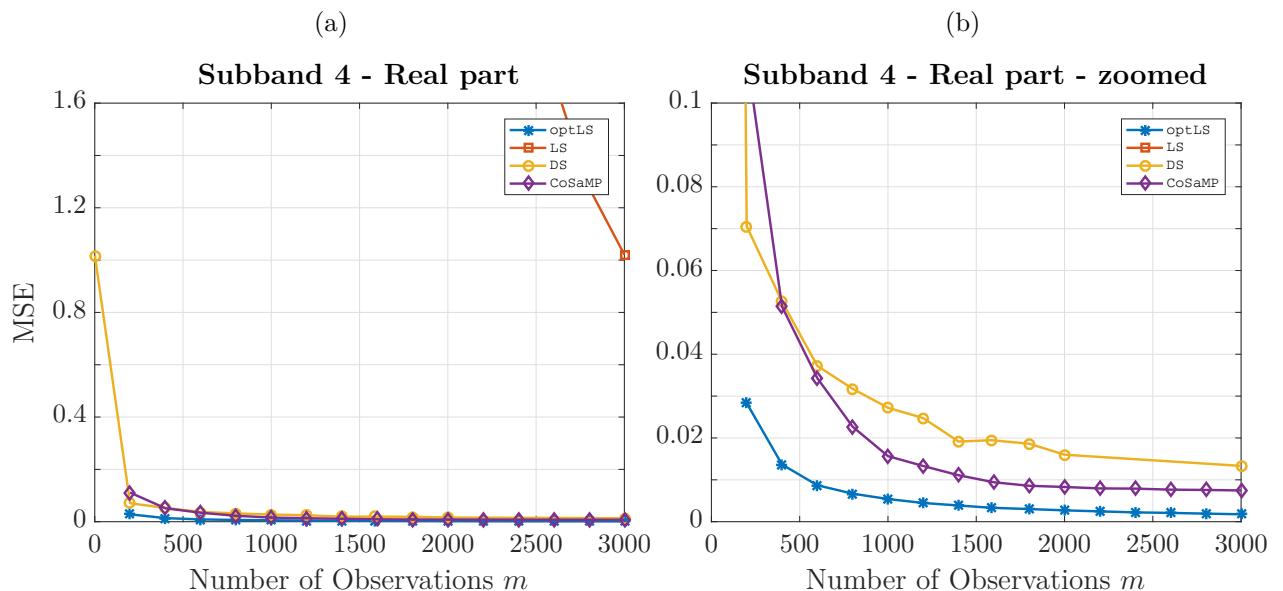


Figure 9: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of **32** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

Secondly, the imaginary part of the discrete-time CIR is estimated, leading to the following results depicted in Figures 10 - 13. Similar as for the channel estimation for the real part of the CIR, all Figures 10 - 13 (a) show the same poor performance in terms of MSE for LS channel estimation. DS and CoSaMP perform much better.

In Figures 10 - 13 (b), again, an unstable behavior for CoSaMP can be observed for low number of observations m . For these subbands, i.e. 1 and 4, CoSaMP even gets worse for a high number of observations m . The algorithm seems to be unsuitable for the estimation of the imaginary part of this subchannel. In contrast to this, the DS shows a stable behavior, getting better the more observations are available.

In this context, it should also be noted that the channel energy for the imaginary parts of the THz subchannels is very low (see Appendix A and B), which makes performing a thresholding algorithm like CoSaMP difficult and probably unsuitable for this kind of channel estimation.

Moreover, the imaginary parts of the subchannels show very low SNR. Therefore, the subchannels were difficult to estimate reliably because the noise level is very high. A compressive sensing approach, which tends to estimate zero and non-zero taps would therefore mainly observe zero taps as there are no significantly dominant taps available for estimation. Thus, the MSE of a channel estimate with many zero taps and a channel with low channel energy appears to be low. This indicates that the MSE might not be the right error metric for this case and it also explains the very low MSE in comparison to the estimation of the real part of the discrete-time CIR, which has a much better SNR in comparison to the imaginary part.

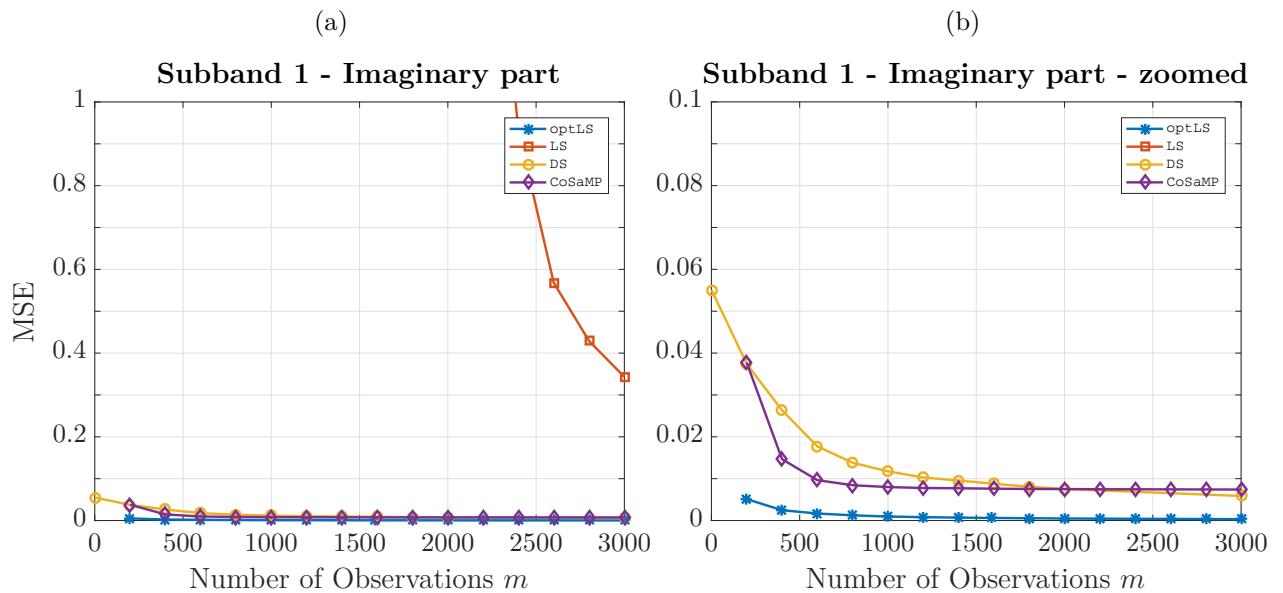


Figure 10: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of 32 subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

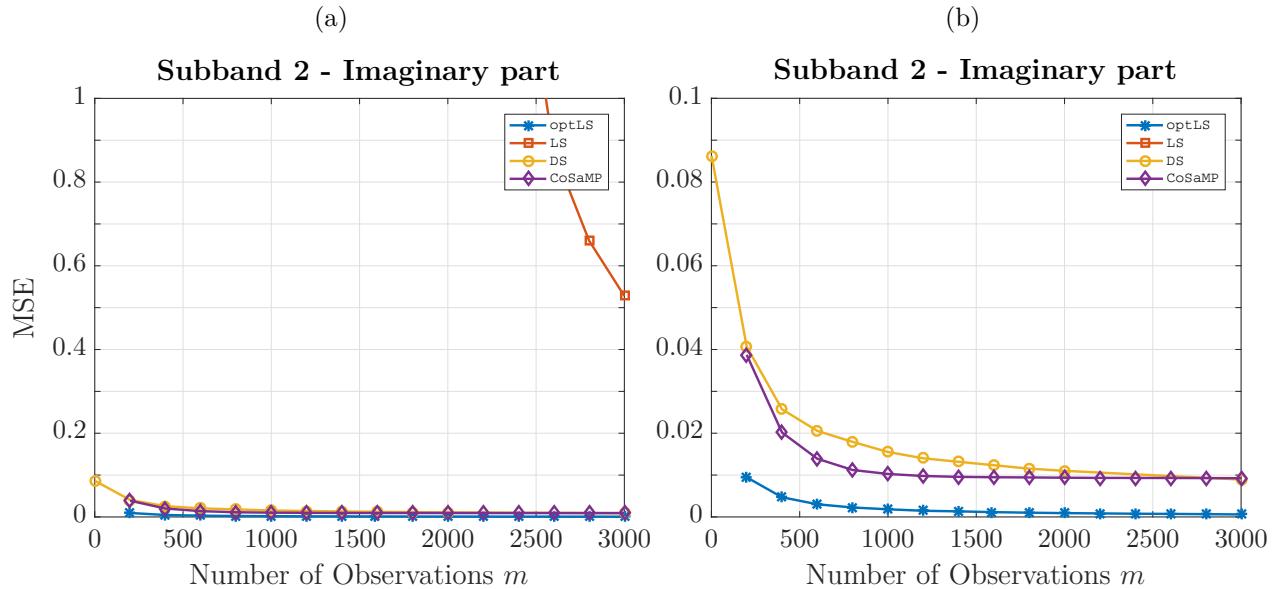


Figure 11: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of 32 subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

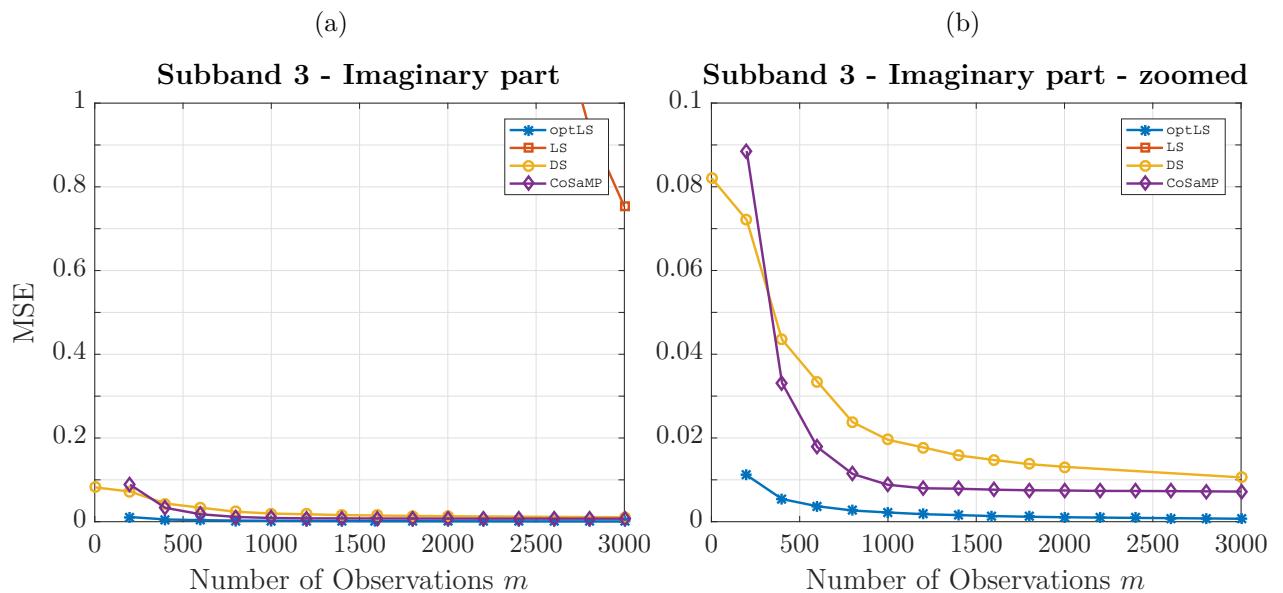


Figure 12: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of 32 subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

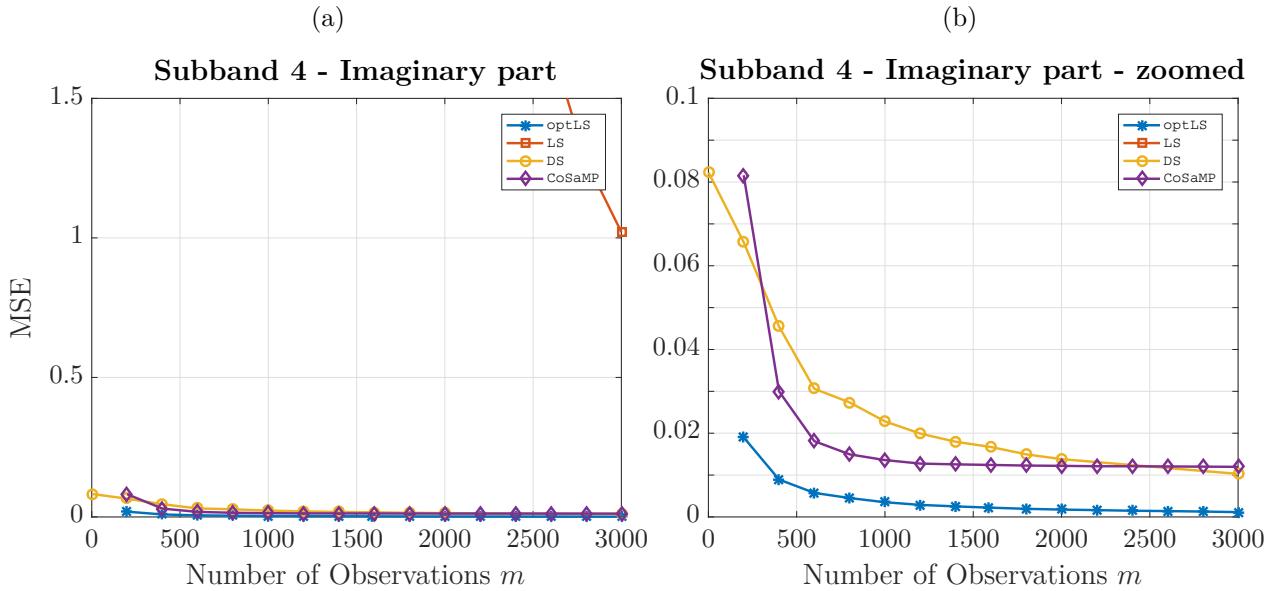


Figure 13: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of 32 subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
 (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

16 Subchannels

If the channel is divided into 16 subbands, similar results are obtained as with 32 subchannels before. The main difference is that the SNR of each subchannel is lower than it was for each subband before, as now each subband has more channel taps with more almost-zero taps. Therefore, the channel energy is lower and the noise variance is significantly higher (see Appendix A). The results are depicted in Figures 14 - 21 for $p = 3223$. For evaluation, only the first four subbands are chosen as they show the best SNR (see Appendix A). Firstly, the real part of each of the four subbands is analyzed. Afterwards the imaginary part will be evaluated.

The results for the real part of the subchannels are depicted in Figures 14 - 17. All Figures 14 - 17 show a poor performance for LS channel estimation, worse than it was for 32 subchannels. This arises from the fact, that the subchannels now have more channel taps than they had before and the noise influence is higher than for 32 subchannels. Both, DS channel estimation and CoSaMP show a better MSE close to the *oracle* based LS estimation.

For the first and second subband, depicted in Figures 14 - 15, both concepts show a similar behavior, with DS channel estimation being slightly better for a low number of observations. The reason therefore is the same as explained for the case with 32 subchannels. For channel estimation of the third and fourth subband, DS channel estimation shows a clearly better MSE in the beginning. Both concepts converge to the same MSE for a high number of observations m (see Figures 16 - 17). However, it should also be noted that the fourth subband shows a low channel energy and high noise variance, leading to a low SNR. Therefore, the same statement holds as for 32 subchannels and channel estimation of the imaginary parts of their first four

subbands. Thus, CS channel estimation might be unsuitable for this case.

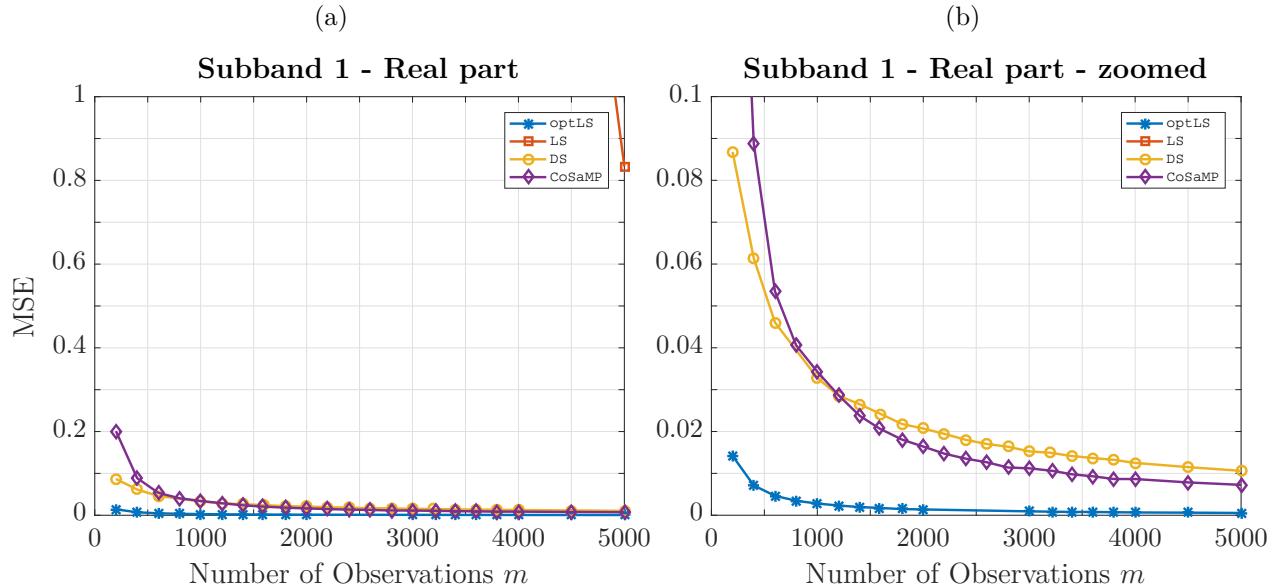


Figure 14: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of **16** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

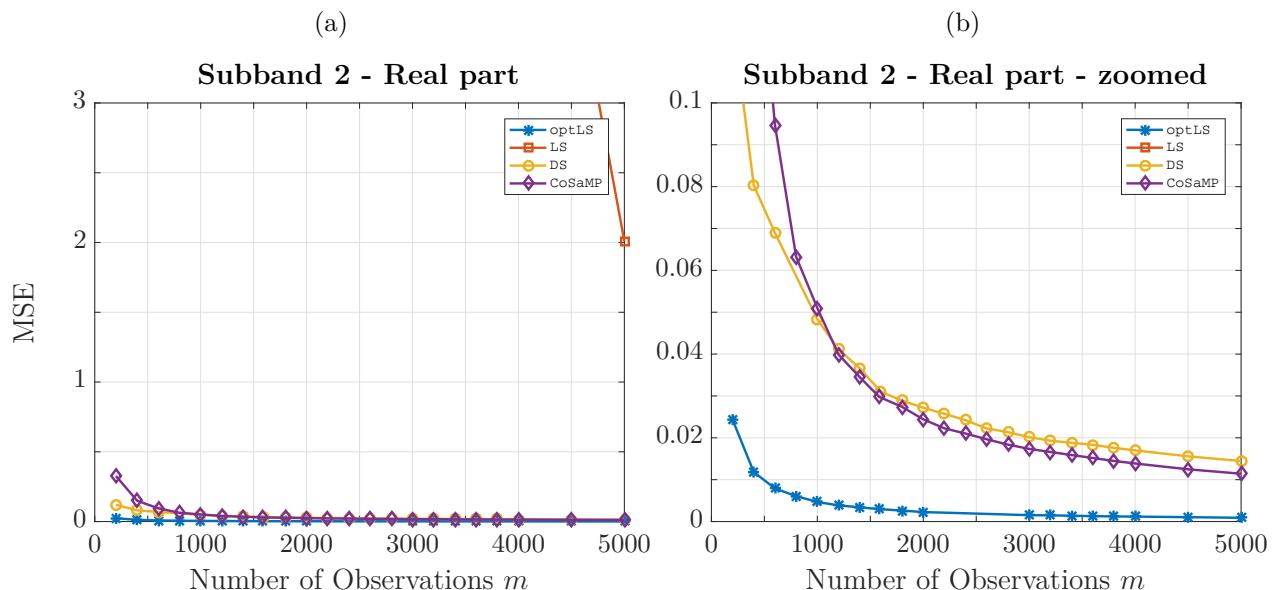


Figure 15: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of **16** subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

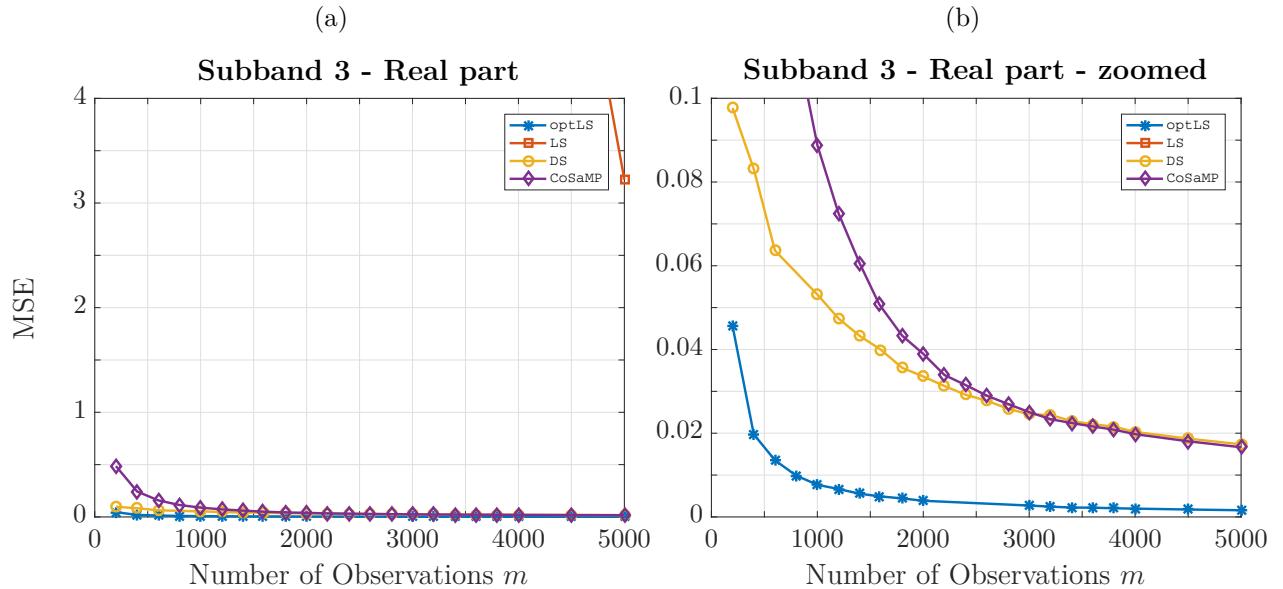


Figure 16: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of 16 subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

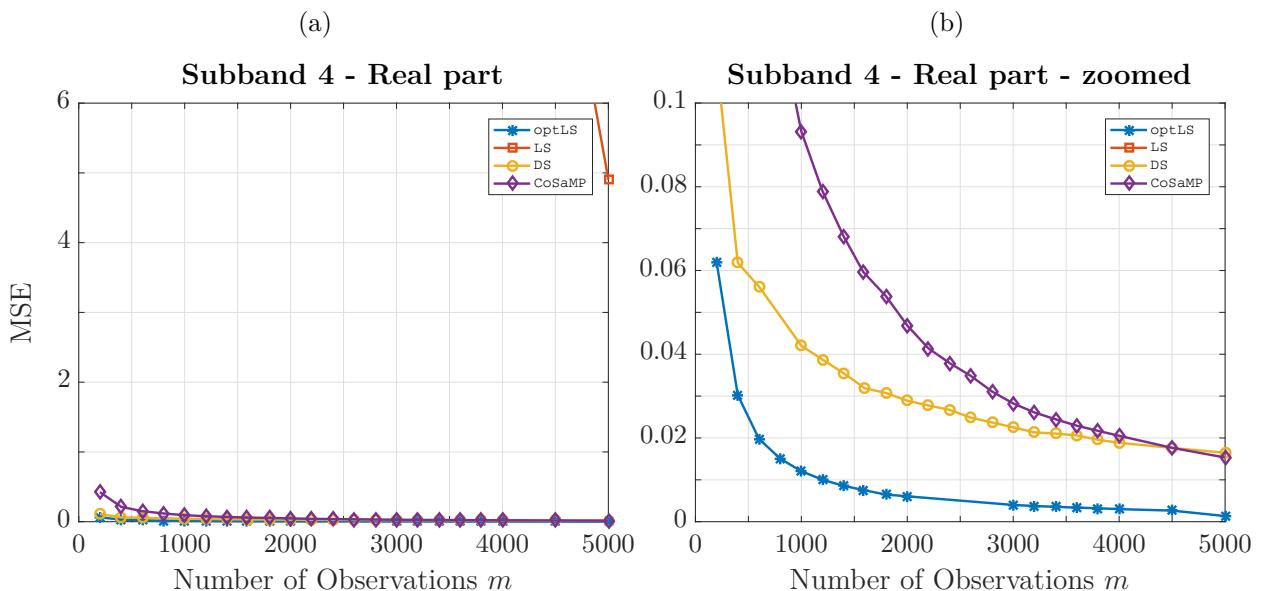


Figure 17: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of 16 subchannels. Real part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

The results for channel estimation of the imaginary part of the first four subbands are depicted in Figures 18 - 21. LS channel estimation shows a poor performance for all four subbands. CoSaMP and DS channel estimation outperform LS channel estimation significantly. CoSaMP is slightly better than DS channel estimation for the first subband (see Figure 18), whereas DS channel estimation shows a better MSE for the second subband (see Figure 19) and a low number of observations. Moreover, a clearly better performance for channel estimation of the third and fourth subband is achieved (see Figure 20 and Figure 21).

However, the imaginary part of each of the 16 subchannel exhibits a much lower SNR compared to the case considering 32 subbands. A reliable channel estimation can therefore not be guaranteed due to the same reason as was stated for 32 subchannels before, where channel estimation was performed for the imaginary part of each subchannel.

For further analysis, the imaginary part of the THz channel will not be taken into account.

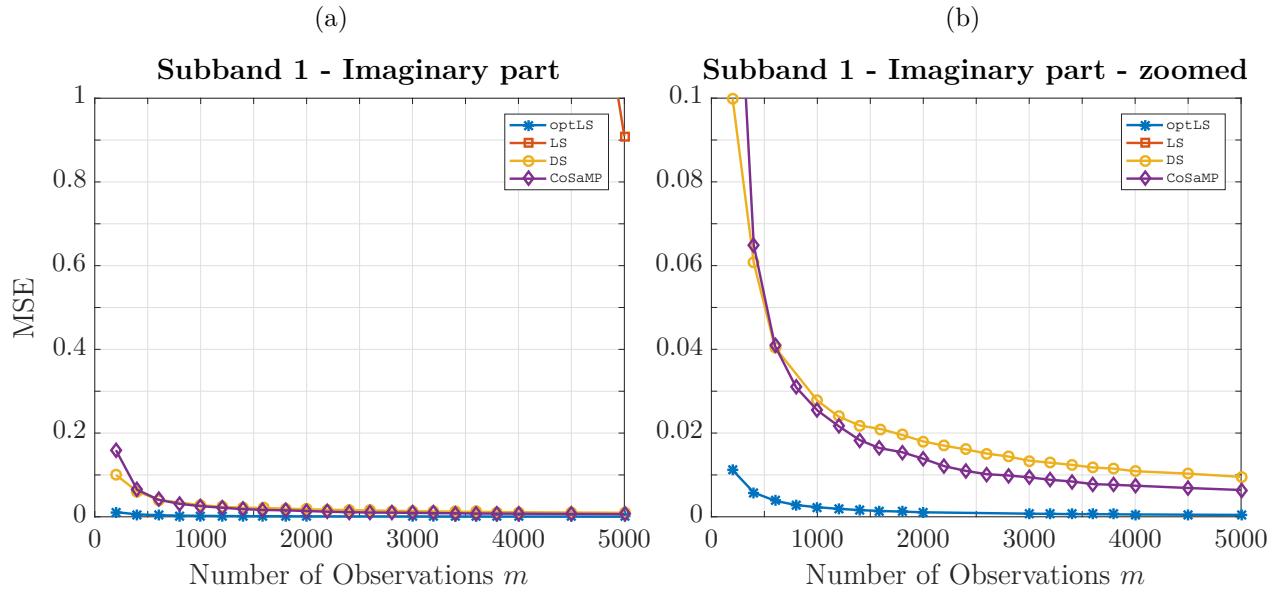


Figure 18: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of 16 subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

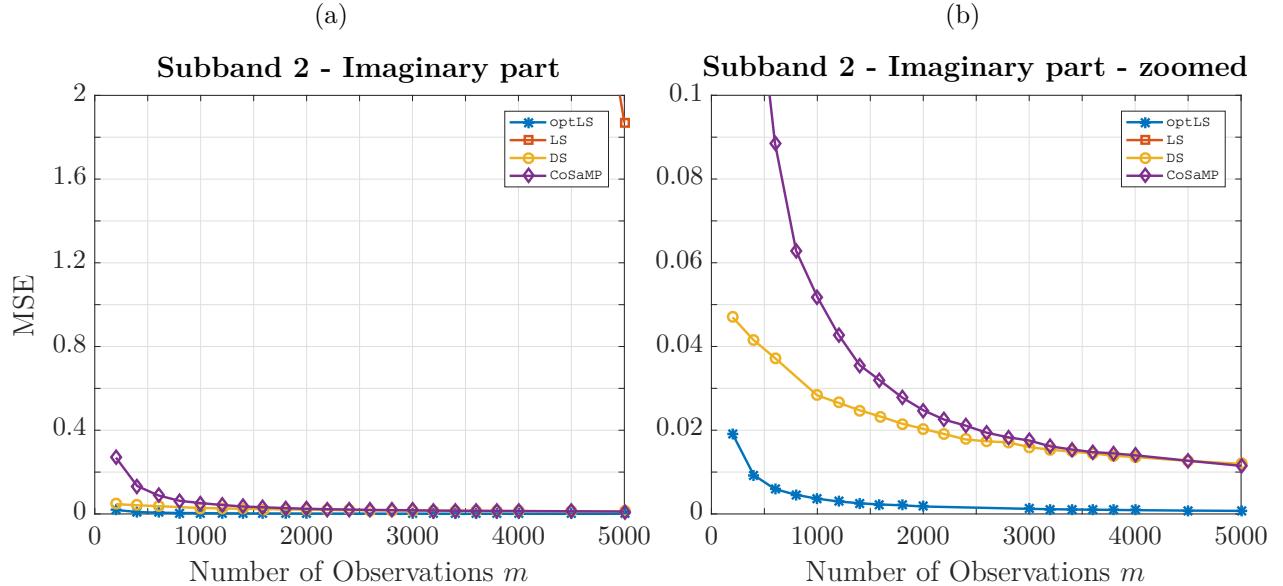


Figure 19: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of **16** subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

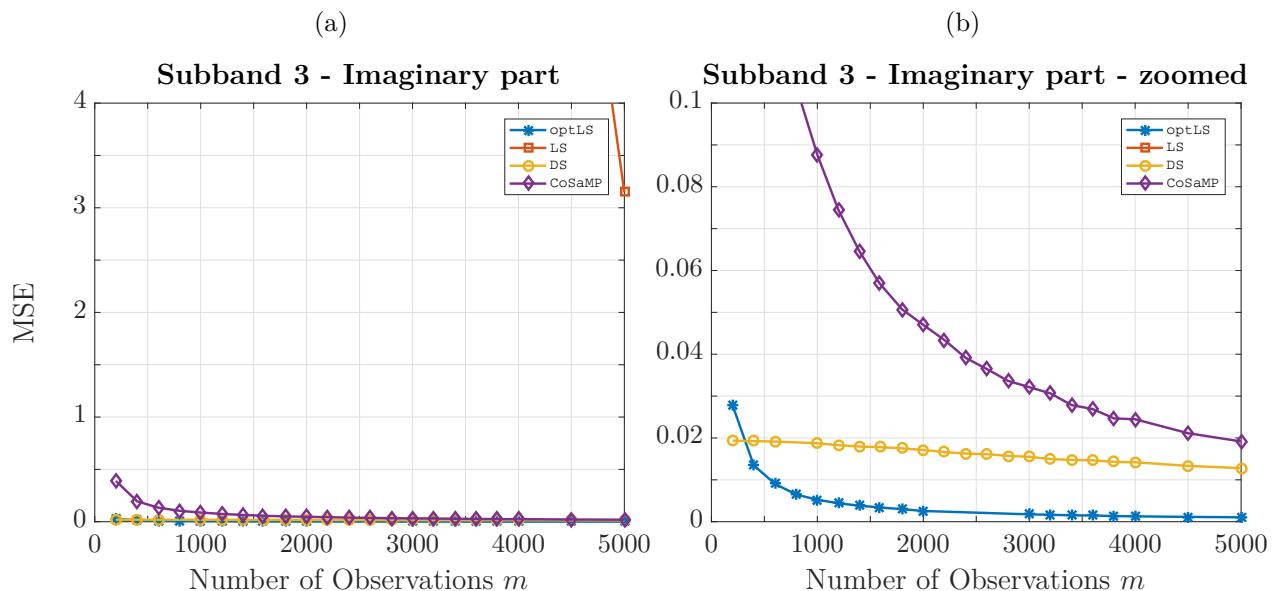


Figure 20: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of **16** subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

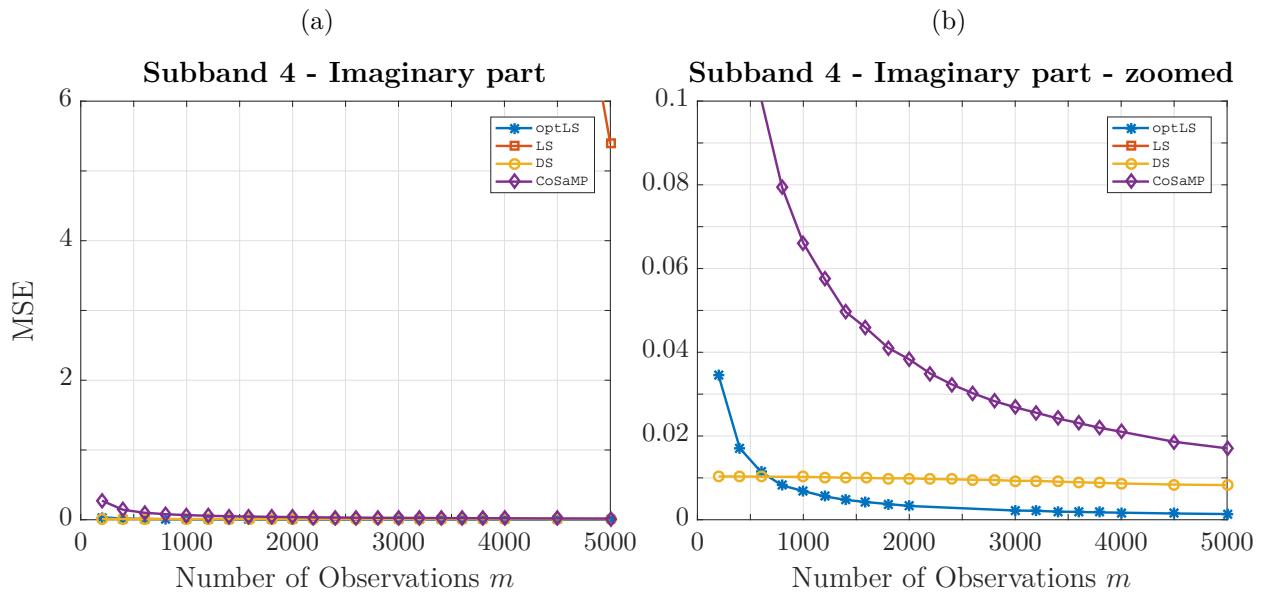


Figure 21: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of **16** subchannels. Imaginary part of the discrete-time CIR for the **entire** channel is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

3.4.2. Channel Estimation for the First 1000 Subchannel Taps

In the following channel estimation task the same parameters as before are assumed, but the channel length is reduced to $p = 1000$ channel taps. By choosing a shorter channel length either a better performance in MSE can be achieved or more insight into the importance for channel estimation of the last channel taps can be gained.

The obtained MSE for CS, LS, CoSaMP and optLS can be seen in Figures 22 - 29. As was said in Subchapter 3.4.1, the imaginary part of the discrete-time CIR has a very low channel energy and a low SNR. Therefore, a meaningful channel estimation is not achievable with the proposed methods. Thus, an analysis of the channel estimation results for the imaginary part of the discrete-time CIR will not be given.

32 Subchannels

Firstly, 1000 channel taps of the real part of the discrete-time CIR are estimated, for the case that 32 subchannels are considered. The results are shown in Figures 22 - 25.

An improvement for LS channel estimation for estimating the shorter channel compared to estimating the entire channel before can be obtained (see also Figures 6 - 13). This behavior is trivial, as the channel length is now $p = 1000$ and therefore more observations m than unknown channel taps are available for $m > p$, leading to a better performance for $m > 1000$ than before. However, both the DS estimator and the CoSaMP algorithm show a significantly better MSE close to optLS than LS estimation for $m < p$ and $m \geq p$.

The DS does not show a different performance for the estimation of 1000 channel taps compared to the estimation of the entire subchannel. A possible explanation for this behavior could be the very low energy contained in the last taps (see also Appendix B) which therefore contribute less or almost nothing to the DS channel estimation. Thus, if channel estimation is performed only for the first 1000 channel taps, no degradation in MSE is obtained, but a better computational efficiency is achieved as the length of the subchannel is only 1000.

In the case of estimating only 1000 channel taps compared to the estimation of the entire subchannel, the CoSaMP algorithm shows a similar performance for the channel estimation of the first subband and a slightly worse performance for the channel estimation of the fourth subband. This method performs worse than before for the second and third subband, leading to the assumptions that either 1000 channel taps might contain not enough information for this thresholding algorithm or the thresholding parameters were set less suitable for this case than for the case before.

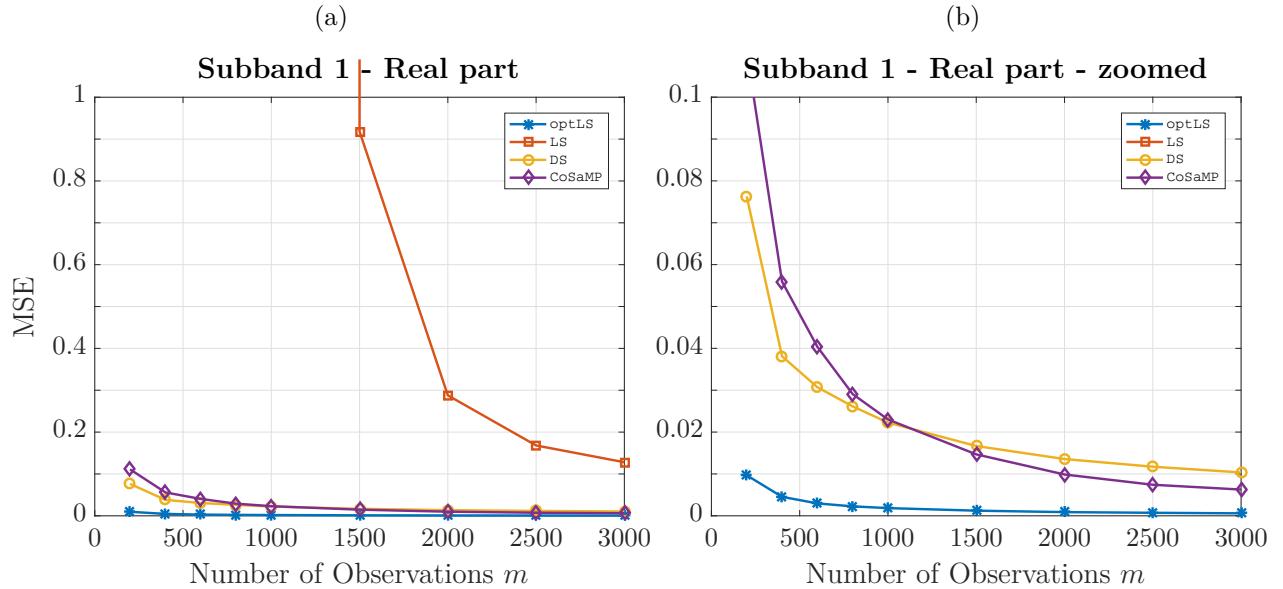


Figure 22: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of 32 subchannels. Real part of the discrete-time CIR for the first 1000 channel taps is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

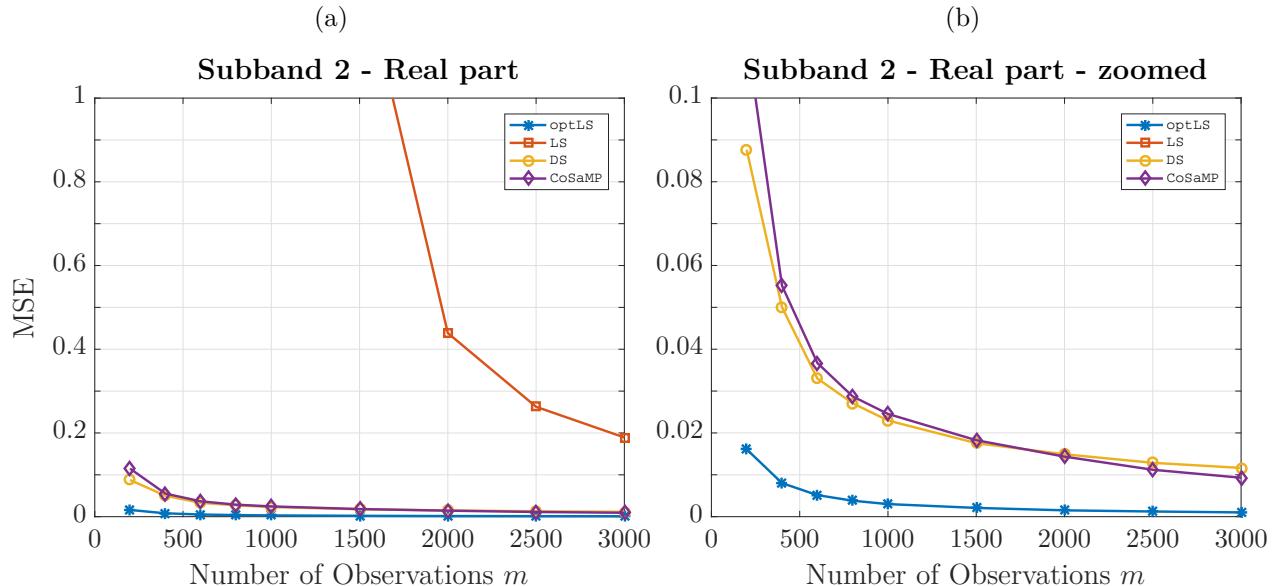


Figure 23: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of 32 subchannels. Real part of the discrete-time CIR for the first 1000 channel taps is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

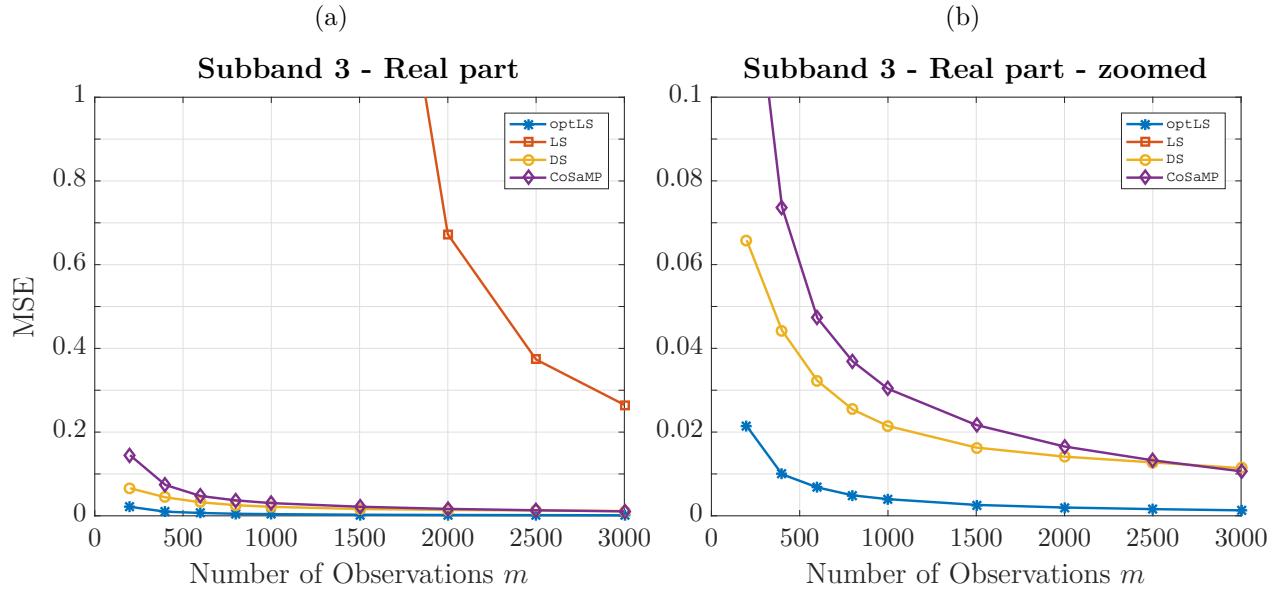


Figure 24: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of **32** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

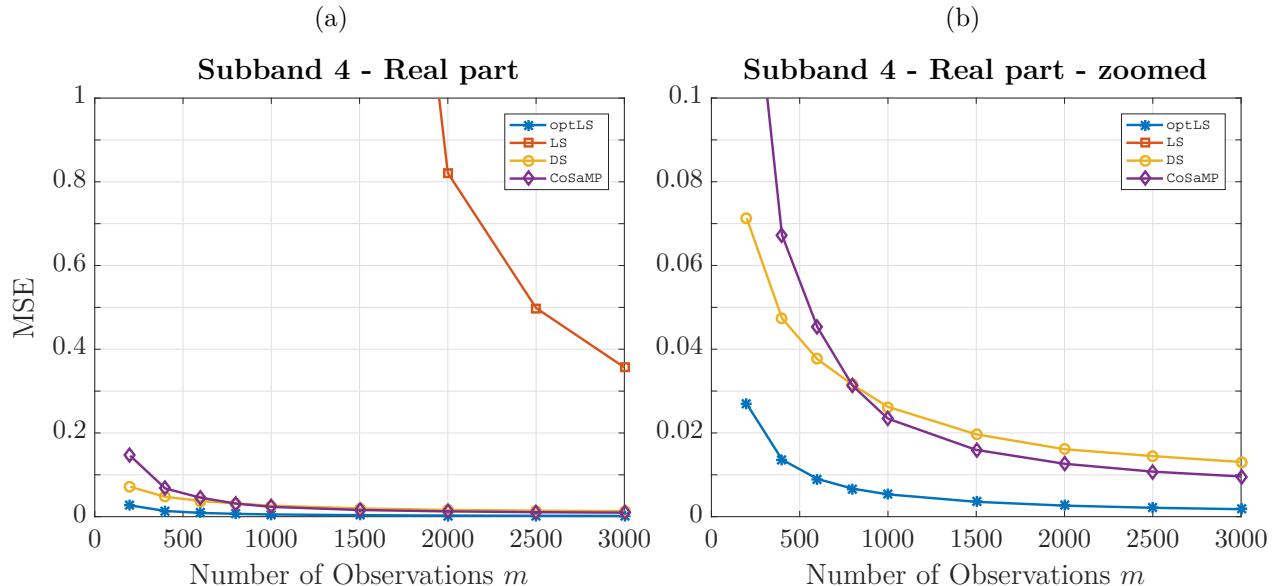


Figure 25: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of **32** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed. (b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

16 Subchannels

Secondly, 1000 channel taps of the real part of the discrete-time CIR are estimated, for the case that 16 subchannels are considered. The results are shown in Figures 26 - 29.

Similar as for the case with 32 subchannels, an improvement for LS channel estimation can be obtained for estimating the shorter subchannel compared to the entire subchannel estimation before (see also Figures 14 - 21). This arises from the same reason as in the case with 32 subchannels before. However, DS channel estimation and CoSaMP channel estimation still exhibit a better MSE.

For the estimation of 1000 channel taps, the obtained results for DS channel estimation behave similar to the case of 32 subchannels before. DS channel estimation does not show a different performance for the estimation of the shorter subchannel compared to the estimation of the entire subchannel before. This behavior leads to the same explanation, the last taps contribute less or almost nothing to the DS channel estimation because they show a very low energy (see also Appendix B). Thus, if channel estimation is performed only for the first 1000 channel taps, no degradation in MSE is obtained, but a better computational efficiency is achieved as the length of the subchannel is only 1000.

The CoSaMP algorithm shows a similar behavior for channel estimation for the first and second subband (see Figures 26 - 27). It exhibits a better performance for the third and fourth subband (see Figures 28 - 29). A possible explanation might be that the thresholding parameters and the channel length are more suitable for subband 3 and 4. However, the SNR for the third and fourth subband is very low, which might lead to an unreliable channel estimation for a thresholding algorithm.

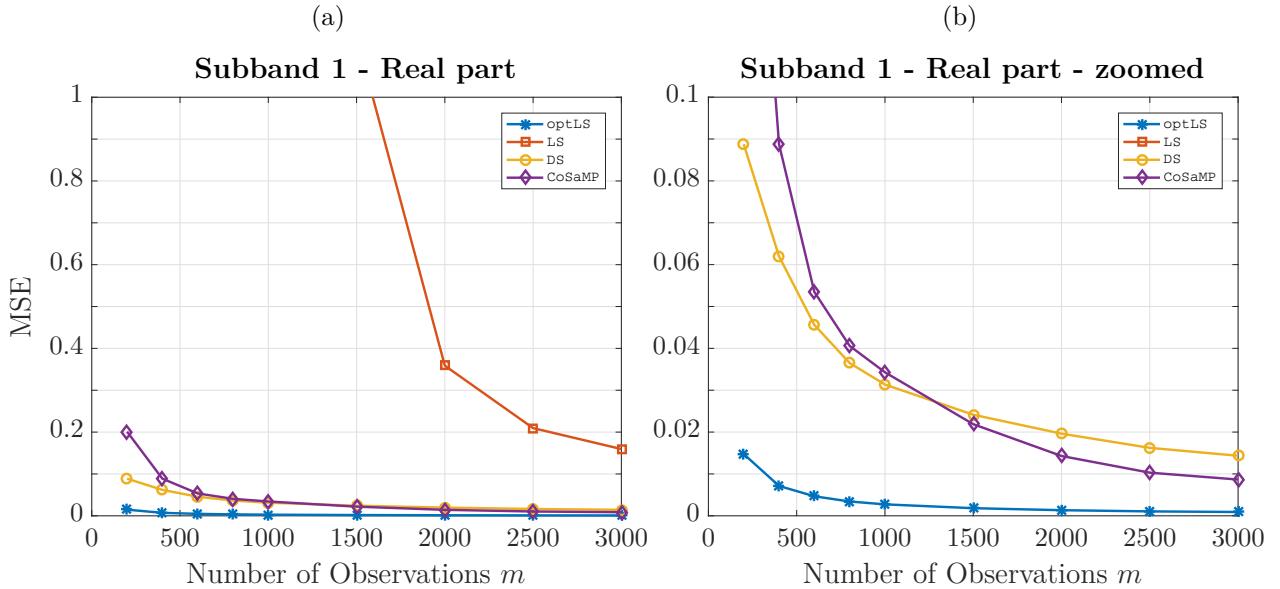


Figure 26: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the first subband out of **16** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

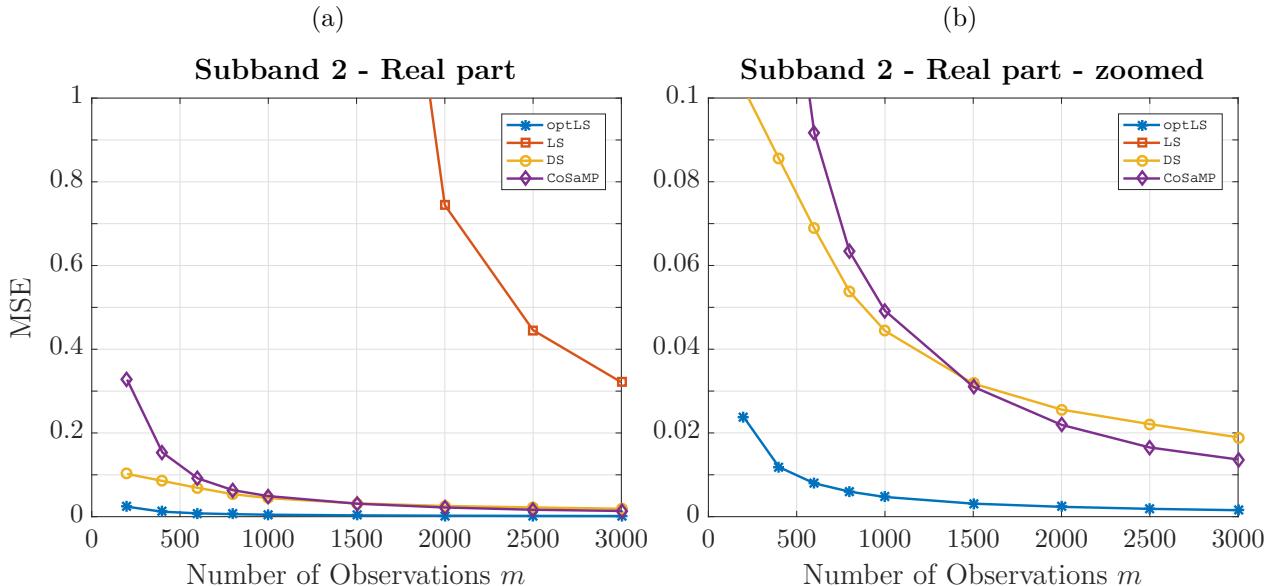


Figure 27: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the second subband out of **16** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

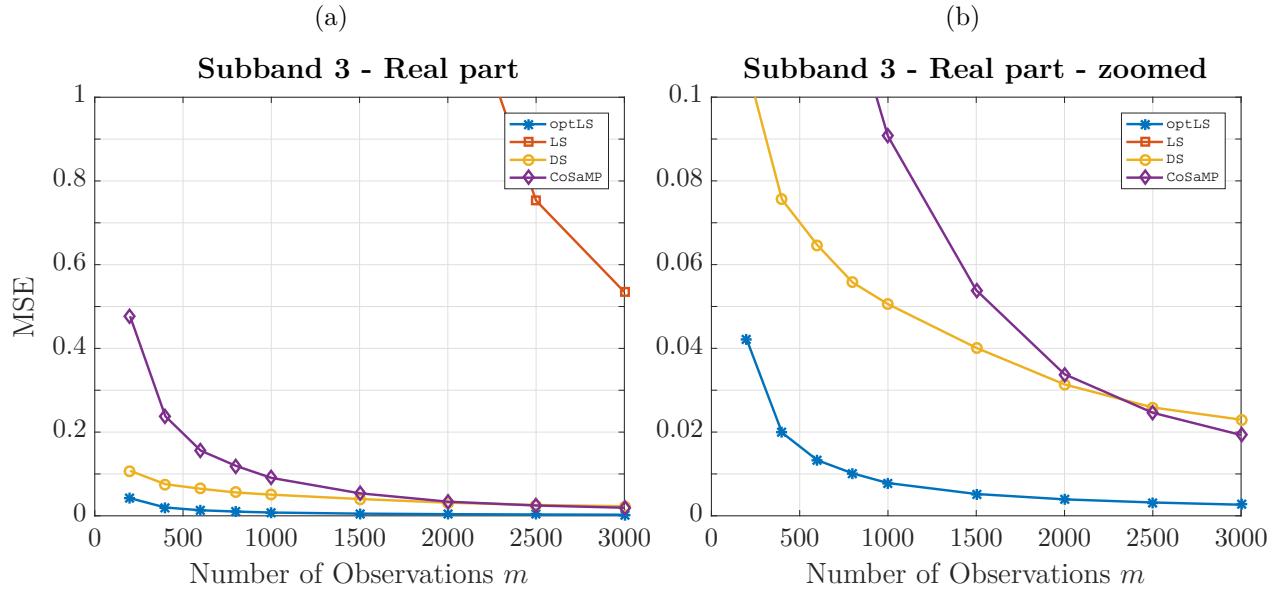


Figure 28: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the third subband out of **16** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

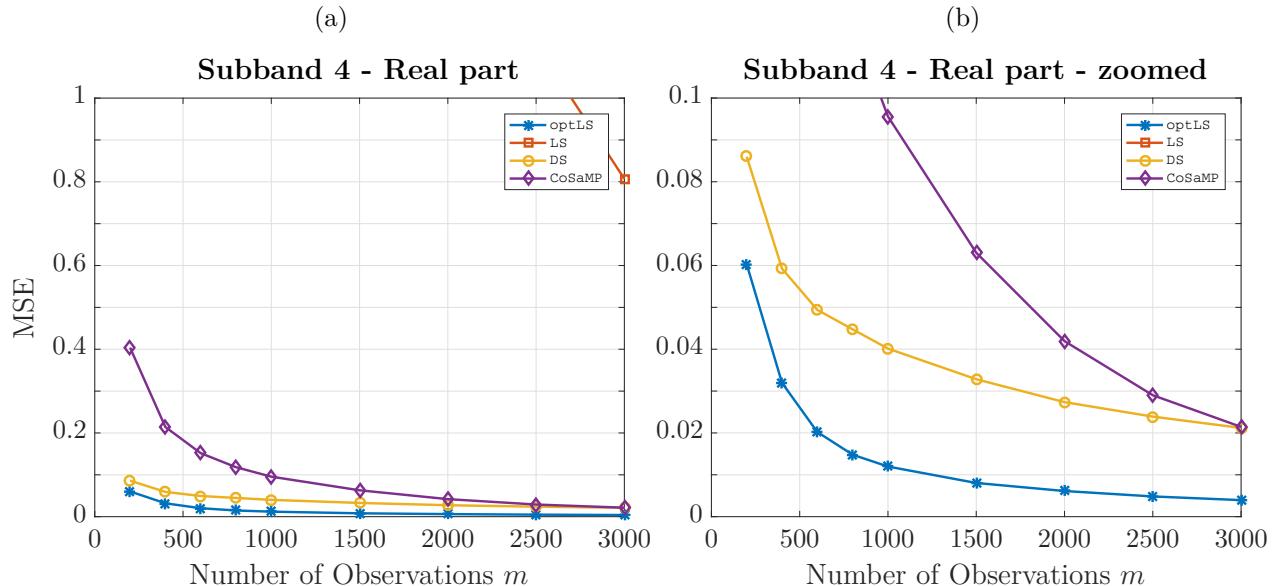


Figure 29: (a) MSE over number of observations m for the proposed estimation strategies optLS, LS, DS and CoSaMP. Result is shown for the fourth subband out of **16** subchannels. Real part of the discrete-time CIR for the first **1000** channel taps is analyzed.
(b) Same as (a) but zoomed in, showing differences between optLS, DS and CoSaMP.

3.4.3. Performance Analysis of a Stochastic Sensing Matrix

The *sensing matrix* \mathbf{A}_{TS} is designed *stochastically*, i.e. the training sequence $\{a_i\}_{i=1}^k$ is given by a sequence of i.i.d. binary random variables, taking values $+1$ and -1 with probability 0.5 each. The simulations for channel estimation by using the DS in Subchapter 3.4.1 and Subchapter 3.4.2 assumed the *same* stochastic sensing matrix for each realization. In this Subchapter the influence of *different* stochastic sensing matrices for each realization of the DS channel estimation is studied. Therefore, the first 1000 channel taps of the real part of the first subband of in total 32 subbands is analyzed and compared to the result in Subchapter 3.4.2. As mentioned before, there is no degradation in MSE if only the first 1000 channel taps instead of the entire subband are estimated by using the DS. The first subband is chosen as it exhibits the highest SNR. All other settings for DS estimation are chosen like described before.

Figure 30 shows, that no improvement in MSE can be achieved if all realizations are performed with a different *stochastic* sensing matrix. Both variants lead to the same performance.

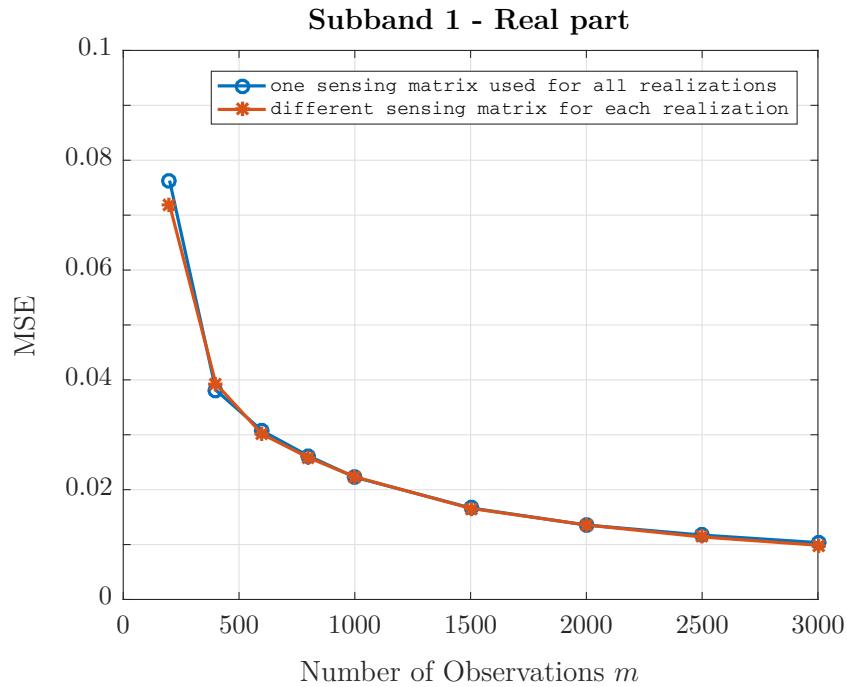


Figure 30: Channel estimation of the first **1000** channel taps of the real part of the first subband of in total **32** subbands. Channel estimation is performed by using the DS. The following two cases are depicted: one stochastic sensing matrix is used for all realizations compared to each realization is performed by using an own sensing matrix.

4. Conclusion and Future Discussion

In this work THz channel estimation by using CS was analyzed closer. Therefore, first, the channel model was investigated closer in Subchapter 2.1, showing the peculiarities of the channel and the approximately sparse structure of the discrete-time CIR. Moreover, a burst structure was proposed in Subchapter 2.2 consisting of a training sequence, a preceding GI and succeeding data symbols.

For a better understanding of the proposed channel sensing methods, the theory of compressed sensing, which is important for channel estimation, was covered shortly in Subchapter 3.1. Thereby, the importance of the $\text{RIP}(2S, \delta_{2S})$ property was pointed out. The first task in CS is to develop a suitable sensing matrix, which fulfills $\text{RIP}(2S, \delta_{2S})$. Therefore, a Toeplitz-structured matrix was proposed in Subchapter 3.2, as this seem reasonable for system identification. The Toeplitz-structure was adjusted to the chosen burst-structure. The second task if a CS framework should be used is to find a suitable channel reconstruction algorithm. For this reason the solution to a convex program, by using the DS was proposed in Subchapter 3.3.2 and a computationally more efficient and greedy method, called CoSaMP was introduced in Subchapter 3.3.3. Achievable error bounds where shown for both methods.

The proposed reconstruction methods were compared to conventional LS channel estimation and to an *oracle* based LS channel estimation approach in Subchapter 3.4. The evaluation showed that the conventional LS channel estimation is outperformed significantly in terms of MSE and number of observations. Both methods achieved performances close to the *oracle* based LS channel estimation approach, which cannot be achieved in reality, but can be seen as a target performance.

CoSaMP appears to work unstable for a low number of observations m . Therefore, two possibilities were proposed in [G⁺10] to avoid this problem. On the one hand, the hard threshold could be relaxed and adaptive thresholding could be used. On the other hand, a lower bound on the training sequence could be set, to guarantee a robust channel estimation. If CoSaMP is considered for future work, these possibilities should be taken into account to improve the performance.

As the imaginary part of the THz channel has very low channel energy, a thresholding approach like CoSaMP seems unsuitable for channel estimation. Also the usage of a DS might be unsuitable, as there are no significantly dominant taps available. According to the results in Chapter 3.4, a compressive sensing approach to estimate the imaginary part of the THz channel is not recommended.

The evaluation results show, that in comparison to CoSaMP the DS works stable for all number of observations m . There is no degradation in MSE if only 1000 channel taps are estimated, leading to the assumption that all channel taps succeeding these first, exhibit a very low channel energy and do not contribute to the estimation process. As CoSaMP exhibits an unstable behavior and is not easily adjustable to each new subchannel, the DS should be preferred for future work.

In this work, a stochastic *sensing matrix* was proposed, as this is beneficial for CS theory because high reconstruction probability can be ensured. It was shown, that the achieved MSE is the same, if either one stochastic *sensing matrix* is used for channel estimation or each realization is performed by using its own stochastic *sensing matrix*. However, random sensing matrices have some drawbacks named in [NS13]. Firstly, random matrices cannot be stored with low effort. Secondly, there are no efficient algorithms available that can verify the RIP condition. Moreover, the authors of [NS13] argued that there might be recovery problems for high signal dimensions. In future work, a deterministic sensing matrix should be developed. First ideas can be used from [NS13].

Appendix

A. THz Channel properties

The SNR of each subchannel was calculated as follows

$$\text{SNR} = 10 \log E_h + 10 \log E_a + 10 \log \left(\frac{1}{\sigma_n^2} \right),$$

with E_h being the subchannel energy for real or imaginary part, E_a the transmit symbol energy and σ^2 being the noise variance of the complex-valued subchannel. Thereby, it should be noted that $\sigma_n^2 = \frac{1}{2}\sigma^2$ was applied for real and imaginary part of the channel.

32 Subchannels

Table 2: Channel details for the real part of 32 subbands.

n	1	2	3	4	5	6	7	8
σ_n^2	0.198	0.303	0.433	0.587	0.756	0.951	1.168	1.408
E_a	1	1	1	1	1	1	1	1
E_h	1.174	1.106	1.076	1.043	1.06	1.037	1.024	1.018
SNR	7.7	5.6	4.0	2.5	1.5	0.4	-0.6	-1.4
n	9	10	11	12	13	14	15	16
σ_n^2	1.691	1.963	2.536	2.600	3.288	3.443	3.834	4.684
E_a	1	1	1	1	1	1	1	1
E_h	1.013	1.011	1.009	1.006	1.003	1.003	1.002	1.002
SNR	-2.2	-2.9	-4.0	-4.1	-5.2	-5.4	-5.8	-6.7
n	17	18	19	20	21	22	23	24
σ_n^2	74.661	11.809	6.367	6.795	6.8	7.321	8.234	18.944
E_a	1	1	1	1	1	1	1	1
E_h	1.294	1.005	1.001	1.001	1.002	1.002	1.002	1.011
SNR	-17.6	-10.7	-8.0	-8.3	-8.3	-8.6	-9.1	-12.7
n	25	26	27	28	29	30	31	32
σ_n^2	27.209	10.554	10.544	11.068	11.74	13.07	14.754	29.212
E_a	1	1	1	1	1	1	1	1
E_h	1.018	1.002	1.002	1.002	1.002	1.001	1	1.008
SNR	-14.3	-10.2	-10.2	-10.4	-10.7	-11.2	-11.7	-14.6

Table 3: Channel details for the imaginary part of 32 subbands.

n	1	2	3	4	5	6	7	8
σ_n^2	0.198	0.303	0.433	0.587	0.756	0.951	1.168	1.408
E_a	1	1	1	1	1	1	1	1
E_h	0.055	0.086	0.081	0.082	0.033	0.034	0.027	0.018
SNR	-5.6	-5.5	-7.3	-8.5	-13.6	-14.5	-16.3	-18.9
n	9	10	11	12	13	14	15	16
σ_n^2	1.691	1.963	2.536	2.600	3.288	3.443	3.834	4.684
E_a	1	1	1	1	1	1	1	1
E_h	0.012	0.007	0.009	0.003	0.010	0.001	0.001	0.002
SNR	-21.4	-24.8	-24.7	-29.7	-25.3	-34.6	-34.8	-33.6
n	17	18	19	20	21	22	23	24
σ_n^2	74.661	11.809	6.367	6.795	6.8	7.321	8.234	18.944
E_a	1	1	1	1	1	1	1	1
E_h	2.649	0.072	0	0.002	0	0	0.001	0.138
SNR	-14.5	-22.2	-43.8	-35.8	-46.1	-46.3	-41.0	-21.4
n	25	26	27	28	29	30	31	32
σ_n^2	27.209	10.554	10.544	11.068	11.74	13.07	14.754	29.212
E_a	1	1	1	1	1	1	1	1
E_h	0.236	0	0	0	0	0.001	0.001	0.034
SNR	-20.6	-47.0	-47.6	-52.5	-48.6	-42.6	-42.2	-29.3

16 Subchannels

Table 4: Channel details for the real part of 16 subbands.

n	1	2	3	4	5	6	7	8
σ_n^2	0.246	0.515	0.872	1.327	1.889	2.592	3.557	4.578
E_a	1	1	1	1	1	1	1	1
E_h	1.083	1.086	1.057	1.029	1.013	1.005	1.004	1.000
SNR	6.4	3.2	0.8	-1.1	-2.7	-4.1	-5.5	-6.6

n	9	10	11	12	13	14	15	16
σ_n^2	94.783	6.744	7.375	33.687	11.173	11.243	14.136	26.226
E_a	1	1	1	1	1	1	1	1
E_h	1.867	1.002	1.001	1.174	1.000	1.001	1.004	1.025
SNR	-17.1	-8.3	-8.7	-14.6	-10.5	-10.5	-11.5	-14.1

Table 5: Channel details for the imaginary part of 32 subbands.

n	1	2	3	4	5	6	7	8
σ_n^2	0.246	0.515	0.872	1.327	1.889	2.592	3.557	4.578
E_a	1	1	1	1	1	1	1	1
E_h	0.123	0.048	0.019	0.010	0.005	0.003	0.003	0.006
SNR	-3.0	-10.3	-16.5	-21.1	-25.4	-29.5	-31.4	-29.2

n	9	10	11	12	13	14	15	16
σ_n^2	94.783	6.744	7.375	33.687	11.173	11.243	14.136	26.226
E_a	1	1	1	1	1	1	1	1
E_h	4.517	0.001	0.001	0.991	0.001	0.000	0.012	0.107
SNR	-13.2	-37.9	-41.2	-15.3	-40.2	-45.0	-30.8	-23.9

B. THz Discrete Time CIR

Some discrete-time CIR of the THz channel exhibit low channel energy. This can be seen in the following Subchapters.

32 Subchannels

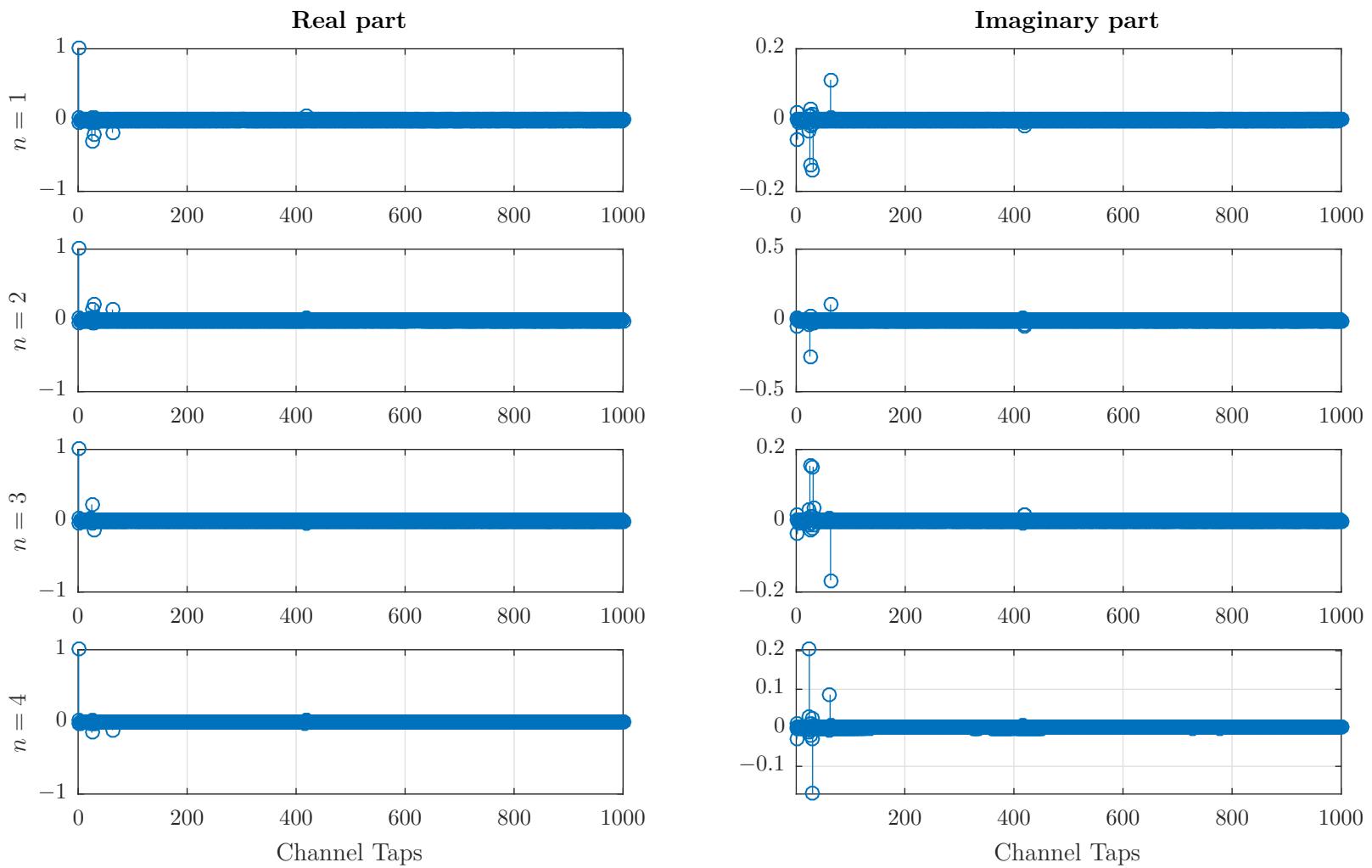


Figure 31: Discrete-time CIR of the first four subbands of the THz channel, with n denoting each subband, for in total 32 subbands.

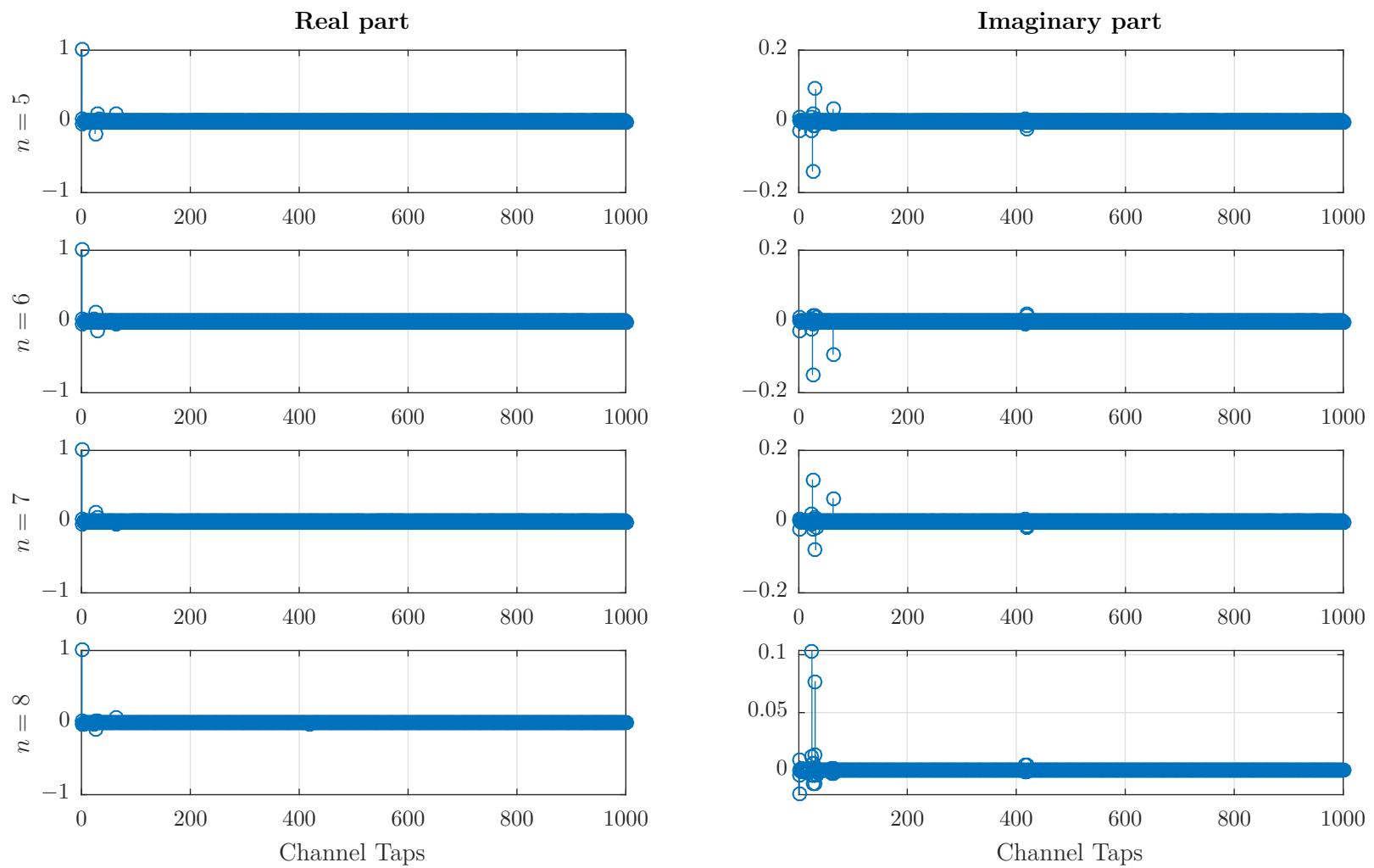


Figure 32: Discrete-time CIR of subbands 5, 6, 7 and 8 of the THz channel, with n denoting each subband, for in total 32 subbands.

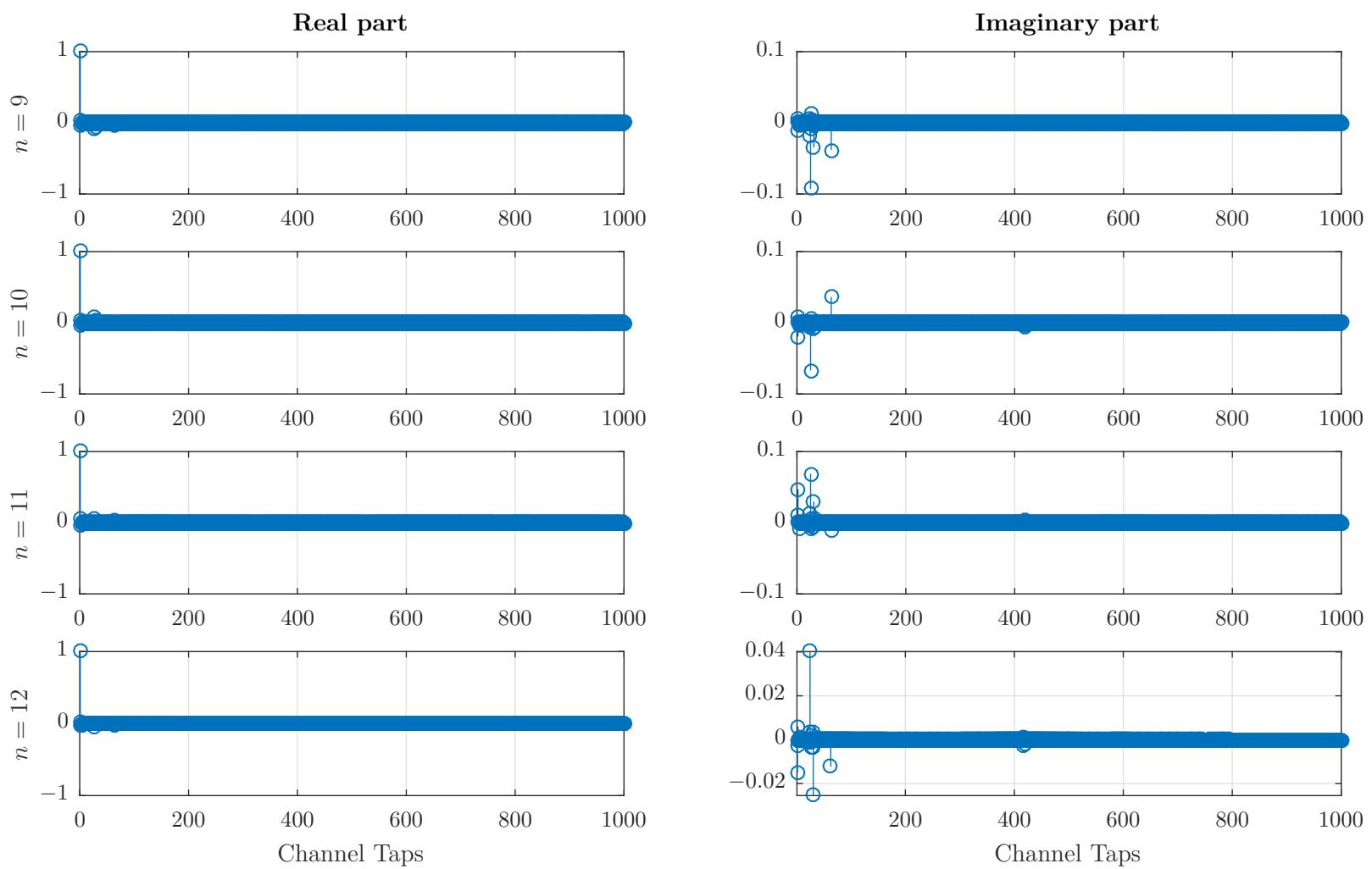


Figure 33: Discrete-time CIR of subbands 9, 10, 11 and 12 of the THz channel, with n denoting each subband, for in total 32 subbands.

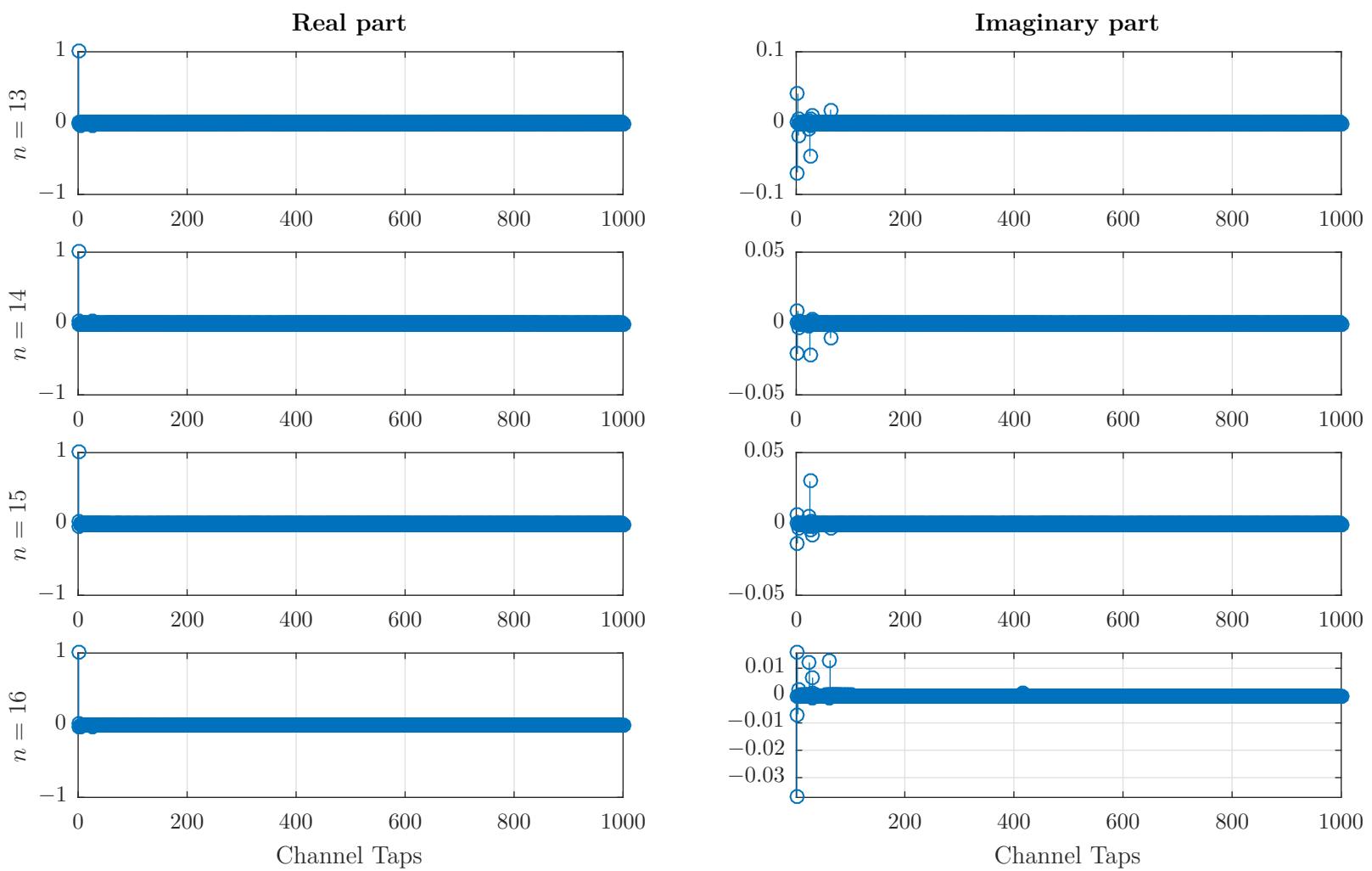


Figure 34: Discrete-time CIR of subbands 13, 14, 15 and 16 of the THz channel, with n denoting each subband, for in total 32 subbands.

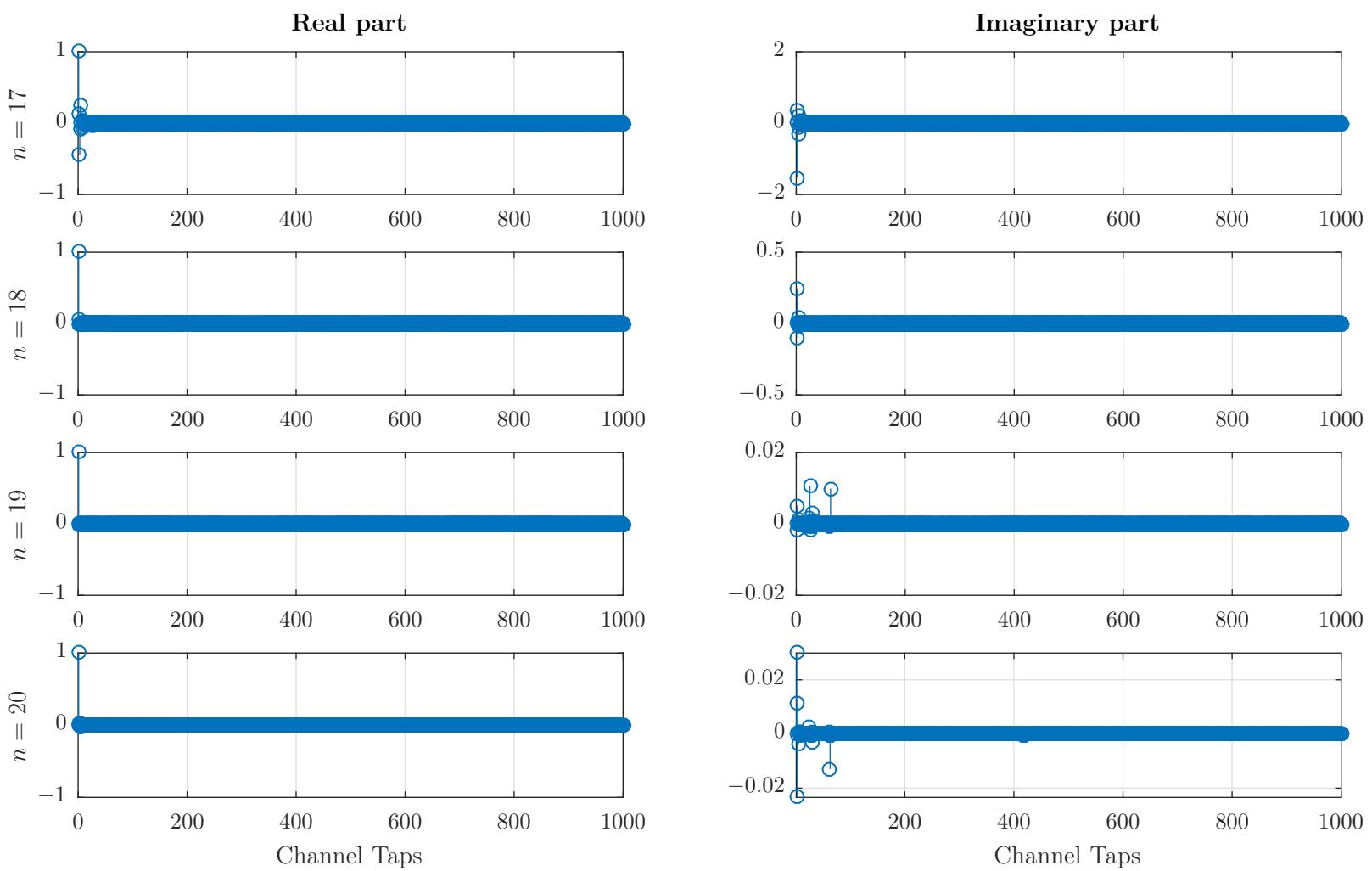


Figure 35: Discrete-time CIR of subbands 17, 18, 19 and 20 of the THz channel, with n denoting each subband, for in total 32 subbands.

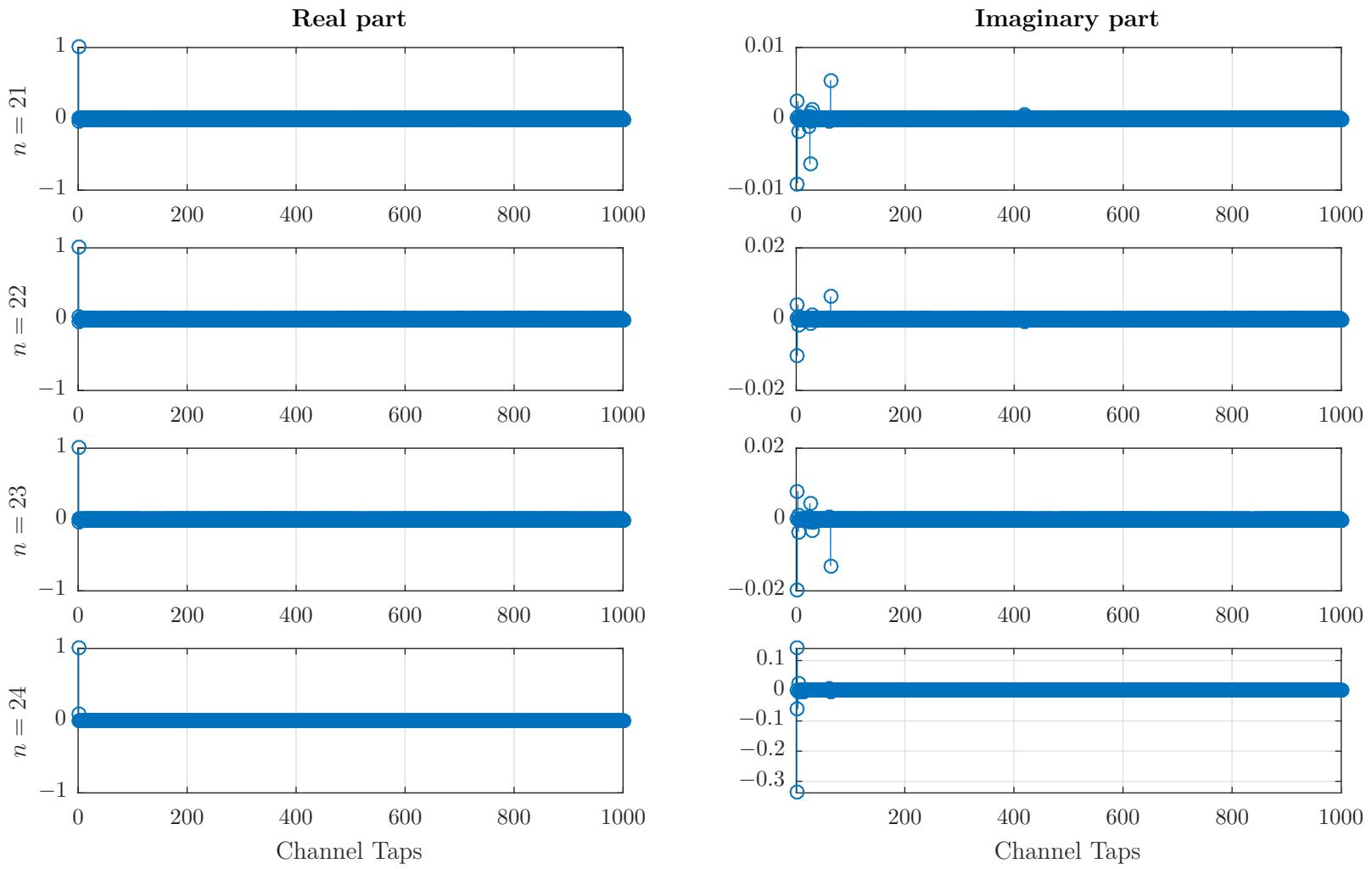


Figure 36: Discrete-time CIR of subbands 21, 22, 23 and 24 of the THz channel, with n denoting each subband, for in total 32 subbands.

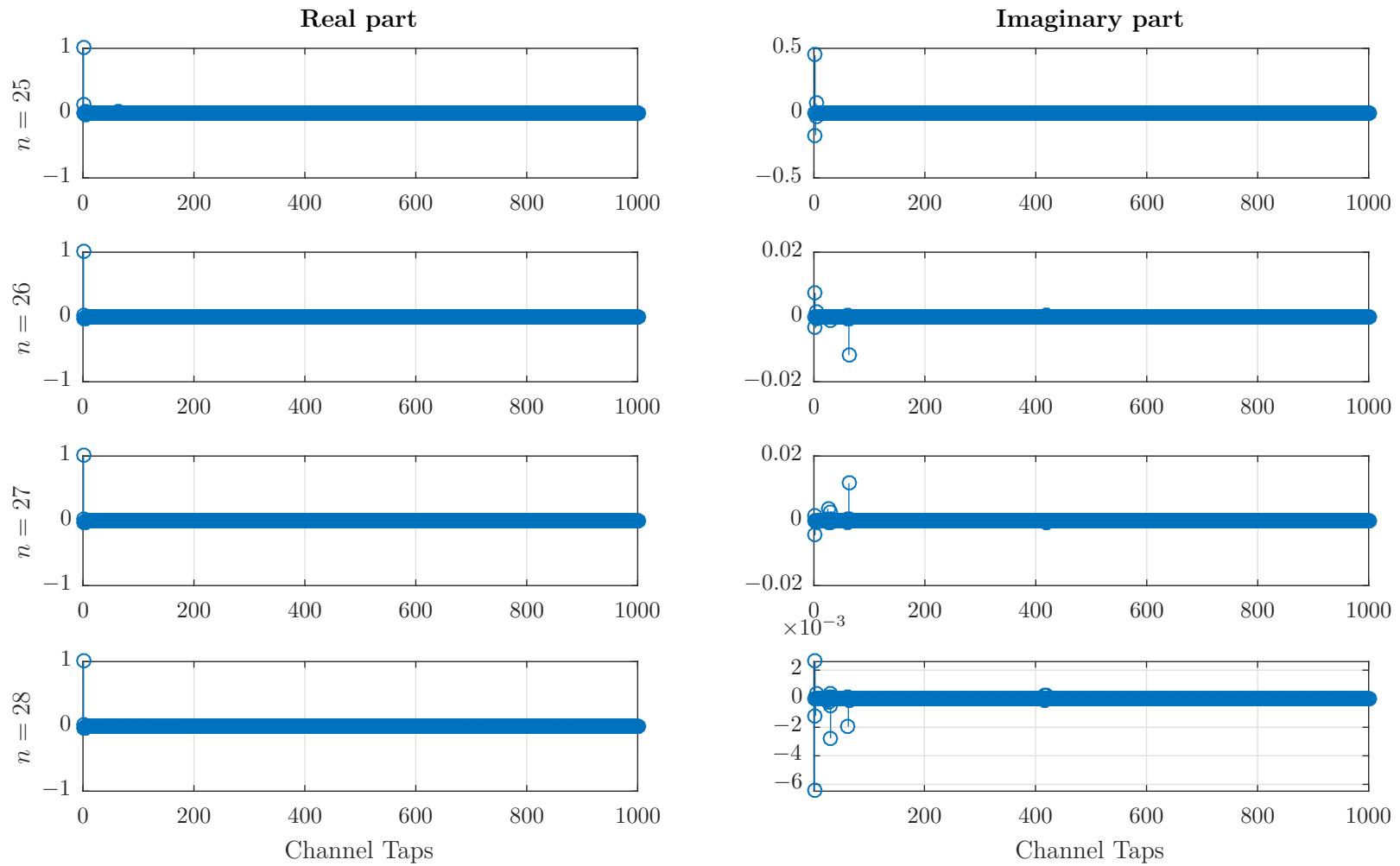


Figure 37: Discrete-time CIR of subbands 25, 26, 27 and 28 of the THz channel, with n denoting each subband, for in total 32 subbands.

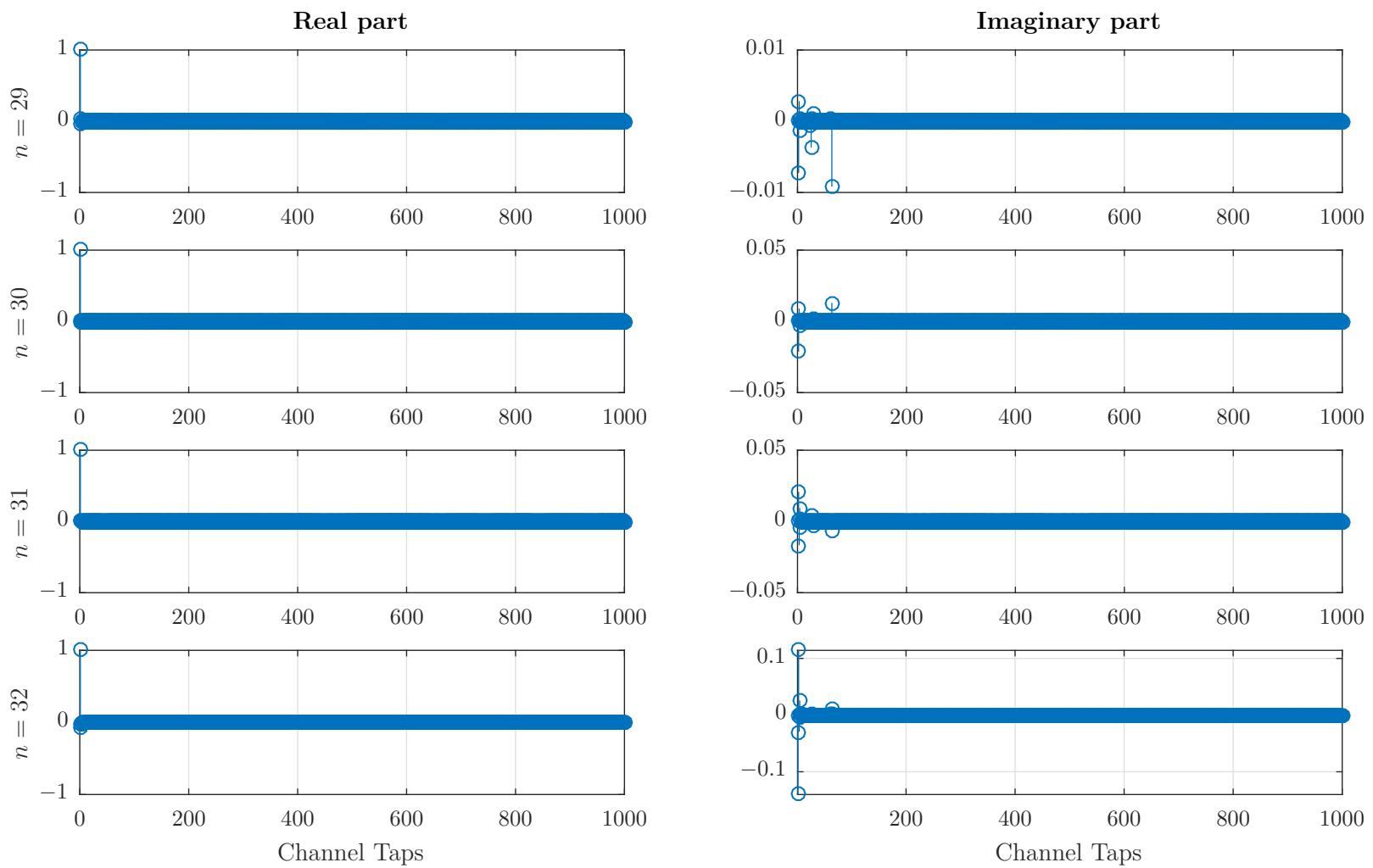


Figure 38: Discrete-time CIR of subbands 29, 30, 31 and 32 of the THz channel, with n denoting each subband, for in total 32 subbands.

16 Subchannels

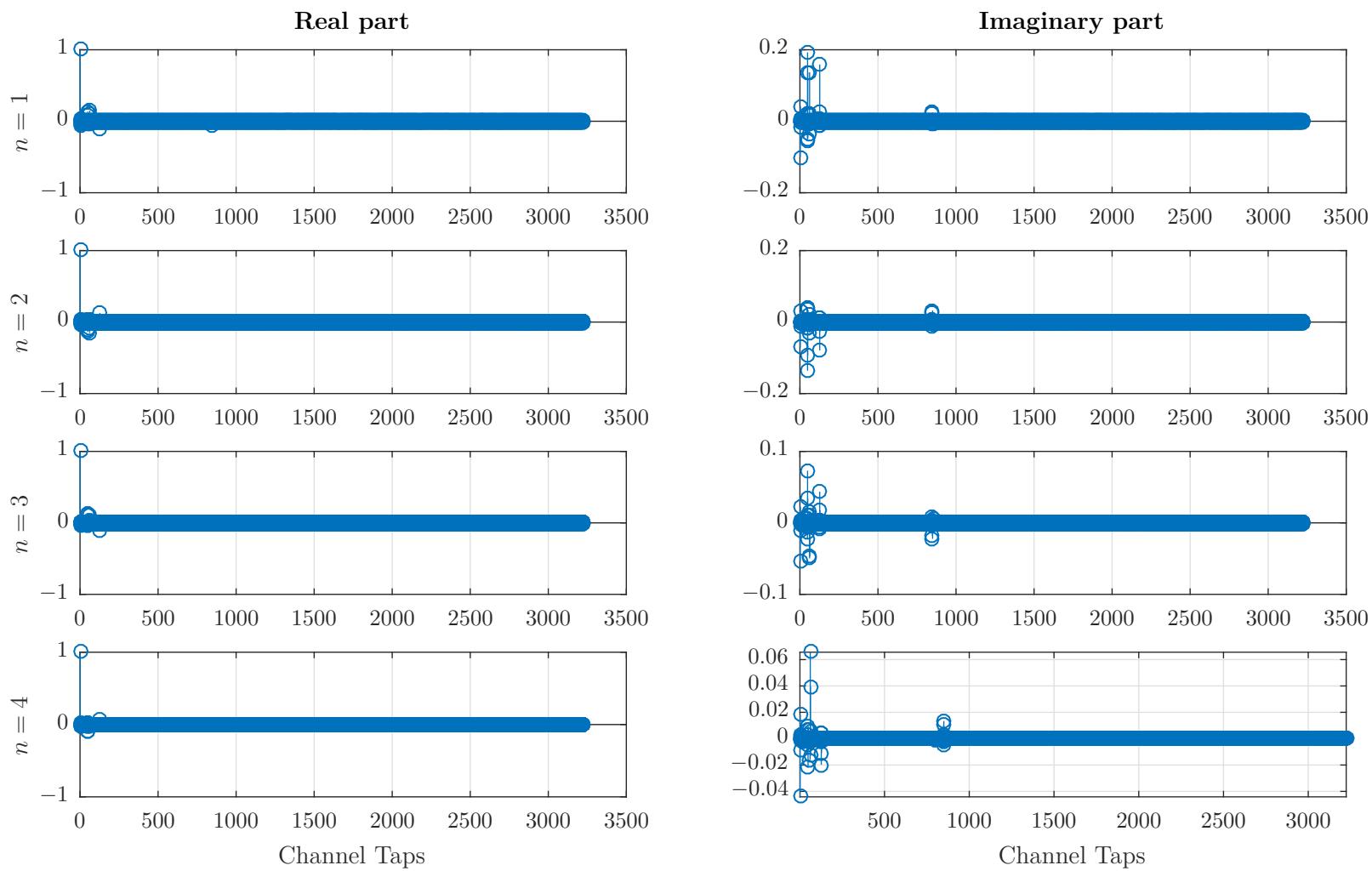


Figure 39: Discrete-time CIR of subbands 1, 2, 3 and 4 of the THz channel, with n denoting each subband, for in total 16 subbands.

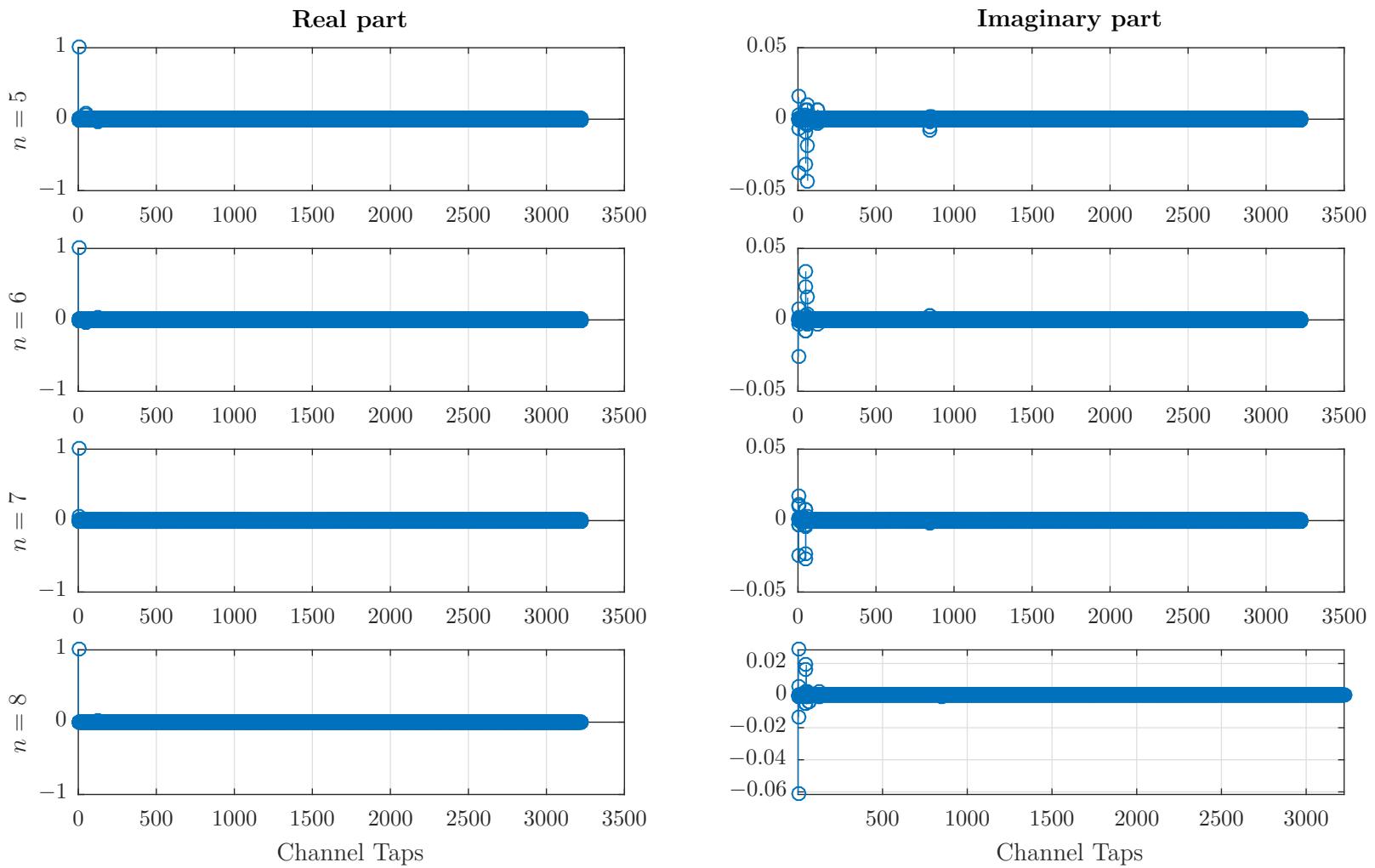


Figure 40: Discrete-time CIR of subbands 5, 6, 7 and 8 of the THz channel, with n denoting each subband, for in total 16 subbands.

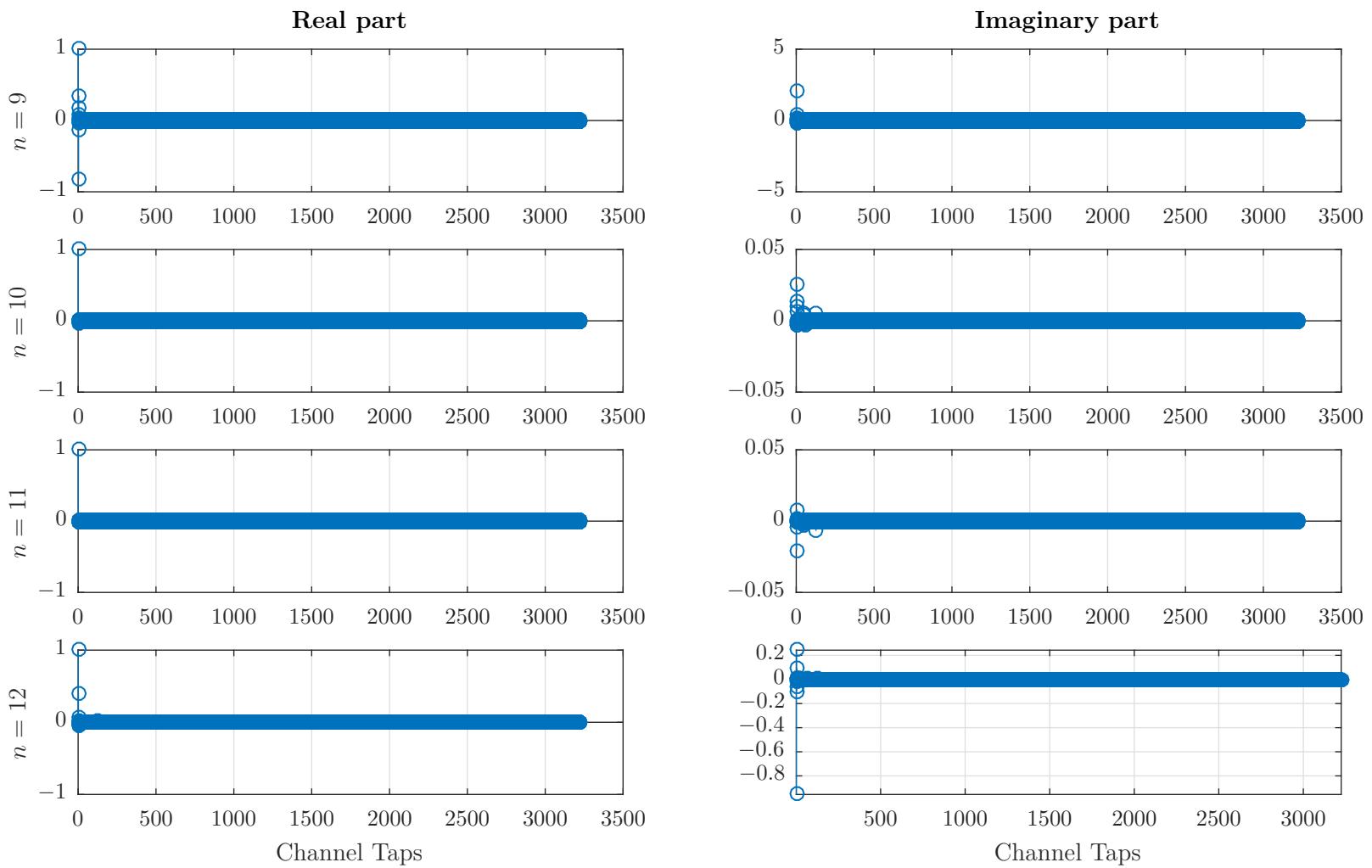


Figure 41: Discrete-time CIR of subbands 9, 10, 11 and 12 of the THz channel, with n denoting each subband, for in total 16 subbands.

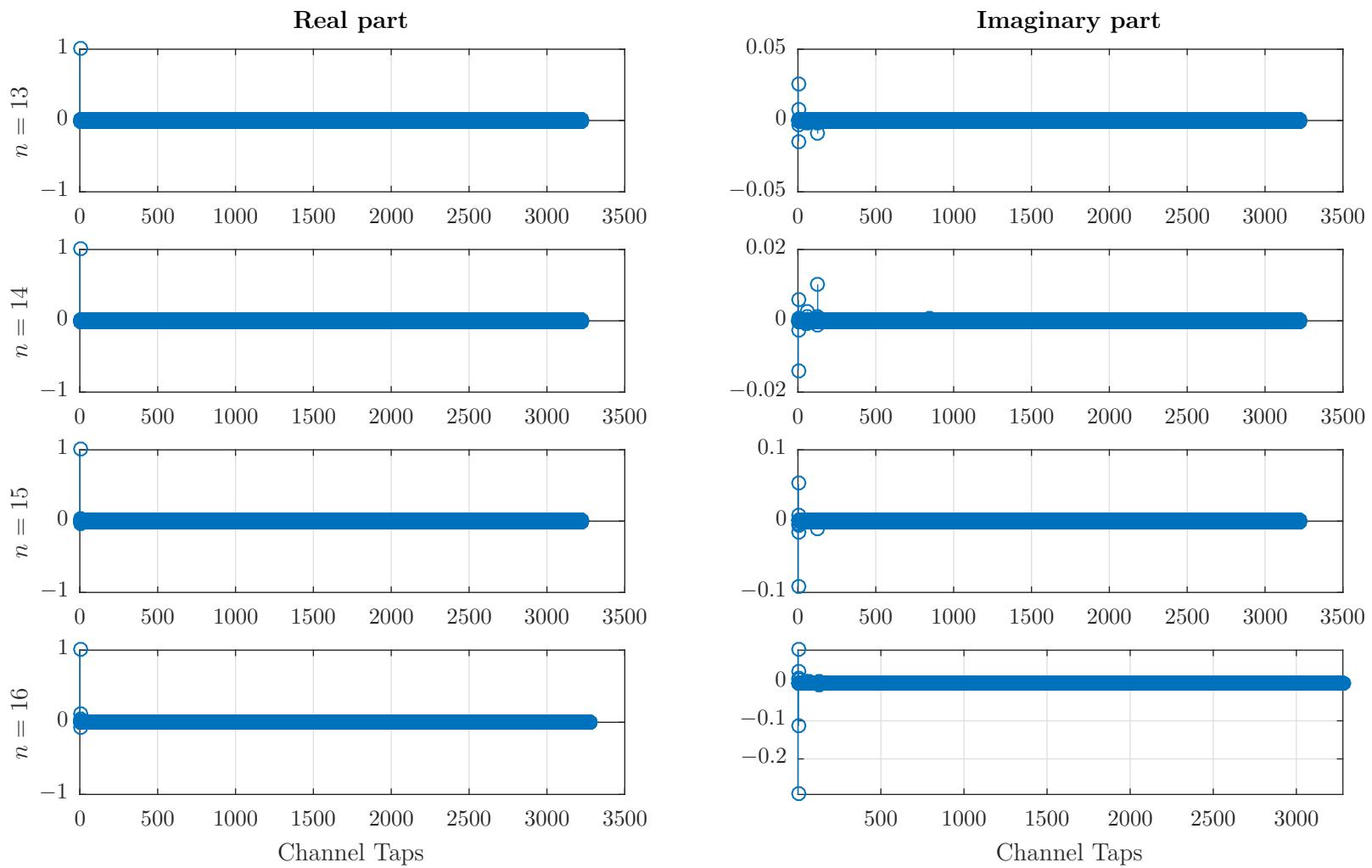


Figure 42: Discrete-time CIR of subbands 13, 14, 15 and 16 of the THz channel, with n denoting each subband, for in total 16 subbands.

References

- [A⁺14] I. F. Akyildiz et al. Terahertz band: Next frontier for wireless communications. *Physical Communication*, 12:16–32, 2014.
- [B⁺07] W. U. Bajwa et al. Toeplitz-structured compressed sensing matrices. In *Statistical Signal Processing, 2007. SSP'07. IEEE/SP 14th Workshop on*, pages 294–298. IEEE, 2007.
- [B⁺10a] W. U. Bajwa et al. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proceedings of the IEEE*, 98(6):1058–1076, 2010.
- [B⁺10b] C. R. Berger et al. Application of compressive sensing to sparse channel estimation. *IEEE Communications Magazine*, 48(11), 2010.
- [Baj08] W. U. Bajwa. Compressed channel sensing. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, pages 5–10. IEEE, 2008.
- [Baj09] W. U. Bajwa. New information processing theory and methods for exploiting sparsity in wireless systems. *The university of Wisconsin-Madison*, 2009.
- [C⁺10] R. Chartrand et al. Introduction to the issue on compressive sensing. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):241–243, 2010.
- [Can08] E. J. Candes. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9-10):589–592, 2008.
- [CR07] E. Candes and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23(3):969, 2007.
- [CT07] E. Candes and T. Tao. The dantzig selector: Statistical estimation when p is much larger than n. *The Annals of Statistics*, pages 2313–2351, 2007.
- [FR13] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*, volume 1. Birkhäuser Basel, 2013.
- [G⁺10] G. Gui et al. Sparse multipath channel estimation using compressive sampling matching pursuit algorithm. *arXiv preprint arXiv:1005.2270*, 2010.
- [Ger16] W. Gerstacker. *Equalization and Adaptive Systems for Digital Communications*. University of Erlangen-Nueremberg, Institute for Digital Communications, WS 2015/16.
- [H⁺10] J. Haupt et al. Toeplitz compressed sensing matrices with applications to sparse channel estimation. *IEEE transactions on information theory*, 56(11):5862–5875, 2010.
- [Kam11] K.-D. Kammayer. *Nachrichtenübertragung*. Vieweg + Teubner, 2011.
- [M⁺16a] A. Moldovan et al. Data rate maximization for terahertz communication systems using finite alphabets. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–7. IEEE, 2016.
- [M⁺16b] A. Moldovan et al. Statistical channel modeling for THz communications. *accepted for presentation at EMN Meeting on Terahertz, San Sebastian, Spain*, 2016.

- [NS13] T. L. Nguyen and Y. Shin. Deterministic sensing matrices in compressive sensing: a survey. *The Scientific World Journal*, 2013, 2013.
- [NT09] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.
- [SB15] Z. Shakeri and W. U. Bajwa. Deterministic selection of pilot tones for compressive estimation of mimo-ofdm channels. In *Information Sciences and Systems (CISS), 2015 49th Annual Conference on*, pages 1–6. IEEE, 2015.
- [Stu11] B. L. Sturm. Algorithm power hour: Compressive sampling matching pursuit (cosamp), 2011. http://media.aau.dk/null_space_pursuits/2011/07/algorithm-power-hour-compressive-sampling-matching-pursuit-cosamp.html. Accessed: 2017-05-30.