Fakultät für Elektrotechnik und Informationstechnik
Technische Universität München

Infinite Unions of Subspaces

With Applications to Channel Estimation in Mobile Communication Systems

Thomas Wiese

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

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzende:
Prof. Dr.-Ing. Antonia Wachter-Zeh

Prüfende der Dissertation:
1. Prof. Dr.-Ing. Wolfgang Utschick
2. Prof. Dr. Felix Krahmer

Die Dissertation wurde am 19.06.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 23.07.2021 angenommen.

ABSTRACT

Compressive sensing theory can be used to analyze linear inverse problems with constraint sets that are finite unions of subspaces. We show how the existing theory can be modified and extended to accommodate infinite union-of-subspaces constraints. These appear, for example, in models used for channel estimation in mobile communication systems that describe the propagation behavior of electromagnetic waves. We provide analyses of several algorithms and discuss conditions under which the reconstruction error can be quantified.

ZUSAMMENFASSUNG

Die Theorie der komprimierten Erfassung kann zur Analyse linearer inverser Probleme verwendet werden, deren Nebenbedingungen als endliche Vereinigungen linearer Unterräume ausgedrückt werden können. Wir erweitern diese Theorie auf Probleme, welche die Verwendung unendlich vieler Unterräume erfordern. Derartige Nebenbedingungen finden Anwendung in der Kanalschätzung von Mobilfunk-Kommunikationssystemen, in denen sie die Ausbreitung elektromagnetischer Wellen zusammenfassen. Wir analysieren verschiedene Algorithmen und geben Bedingungen, unter welchen der Rekonstruktionsfehler quantitativ bestimmt werden kann.

Acknowledgments

I would like to thank everybody who supported me in writing this thesis. This includes my wife and family and the whole MSV-team. All of you contributed to my overall PhD time being a really nice experience.
Thank you Wolfgang Utschick for inviting me to join MSV, for your genuine interest in my work and the resulting fruitful discussions, and for fostering an atmosphere of openness and willingness to discuss each and every mathematical (or typesetting) problem until solved. I cannot remember a single occasion where my attempts to confront a random colleague with a problem that was bothering me was rejected.
In particular, I would like to thank David Neumann and Lorenz Weiland with who I collaborated on a regular basis. I really enjoyed all those whiteboard discussions with you and I really appreciate that you were never tired of listening to another (non-)solution to that Vandermonde projector problem of mine.

Contents

I
INTRODUCTION

Compressive sensing is about non-adaptive data compression. Its theory provides us with tools for minimum-redundancy (compressive) measurement design and its methods are used to extract data from these measurements (this is more complicated than with classical sampling). Data compression is about exploiting structure. Structure needs to be exploited in all kinds of problems; not only those classically associated with data compression.

One such problem is channel estimation in communication systems. When very high frequencies and many antennas are used, structures are unveiled that were previously hidden. For example, the channels between different antenna elements are correlated; high sampling frequencies show that the temporal channel impulse responses contain peaks and valleys. Such observations can be explained by physical channel models in which the channel, which is an object in a high-dimensional space, can be described by a small set of parameters.

In this thesis, we consider measurements of the form

$$y = Ax + e \tag{1}$$

where $x \in \mathcal{U}$ is an element (representing the channel) of a low-dimensional subset $\mathcal{U}$ of a high-dimensional Hilbert space $\mathcal{H}$ and shall be reconstructed from y, which is an element of the measurement Hilbert space $\mathcal{H}'$. The measurement operator $A\colon \mathcal{H} \to \mathcal{H}'$ is assumed linear and $e \in \mathcal{H}'$ is measurement noise. If not much is known about the noise or if e is Gaussian noise, the reconstruction problem is best formulated as the least-squares problem

$$\hat{x} = \operatorname*{arg\,min}_{x \in \mathcal{U}} \|y - Ax\|^2. \tag{NLS}$$

This problem is nonlinear if $\mathcal{U}$ is nonlinear (not a linear subspace). For general nonlinear sets $\mathcal{U}$, problem (NLS) is a non-convex optimization problem and there are no (efficient) methods to solve this problem. However, for certain classes of nonlinear constraint sets, there *are* methods to solve (NLS). We are concerned with constraint sets $\mathcal{U}$ that are *unions of subspaces*, that is, sets that can be written as

$$\mathcal{U} = \bigcup_{t \in T} S_t, \quad S_t \subset \mathcal{H} \text{ (low.-dim.) subspace for each } t \in T. \tag{2}$$

The set T is a parameter set and t is the parameter describing the subspace. For such sets, which are still nonlinear, but also not completely arbitrary, there are efficient methods that solve (NLS) provided that we can solve the best-approximation problem (here and in the following, $\mathcal{P}_{\mathcal{U}}(z)$ denotes the *non-linear* projection of z onto $\mathcal{U}$)

$$\mathcal{P}_{\mathcal{U}}(z) = \arg\min_{x \in \mathcal{U}} \|z - x\|^2 \tag{3}$$

for general $z \in \mathcal{H}$, which is presumably easier to solve than (NLS), and provided that A is a good measurement operator.

A good measurement operator should not lose any information even when the dimension of $\mathcal{H}'$ is small compared to that of $\mathcal{H}$. Information is lost if the vector Ax becomes as small as the noise e for some not-so-small x. A condition of the form $\|Ax\| \geq c\|x\|$ for some not-too-small c is, thus, natural. A good measurement operator should also be neutral. There should not be any $x, x' \in \mathcal{U}$ of similar size for which $\|Ax\| \gg \|Ax'\|$ or vice versa.

The condition

$$(1 - \delta)\|x\|^2 \leq \|Ax\|^2 \leq (1 + \delta)\|x\|^2 \quad \forall x \in \mathcal{U}, \tag{RIP}$$

which has been introduced in [1], is called the *restricted isometry condition* and ensures that A is a good measurement operator. If $\delta = \delta(A, \mathcal{U})$ is small, then A is said to have the restricted isometry property (RIP) with respect to $\mathcal{U}$ and the smallest δ such that (RIP) is satisfied is called the restricted isometry constant (RIC) of A.

This condition has become tremendously popular as it is a very natural condition to demand of the measurement operator (as discussed above), but mostly for the following two reasons: First, there are lots of measurement operators that achieve a great level of dimensionality reduction and have the RIP. Second, the condition appears naturally in the convergence analysis of many algorithms that attempt to solve (NLS), making this problem one of the few instances of non-convex optimization problems for which a global convergence analysis is possible.

The purpose of this thesis is to extend known results regarding RIP-matrix construction and algorithm recovery analysis, which were mostly developed for *finite* unions of subspaces, to *arbitrary* unions of subspaces where T may also be infinite. In particular

- We derive versions of the iterative hard thresholding (IHT), hard thresholding pursuit (HTP), and orthogonal matching pursuit (OMP) algorithms that can be used with infinite unions of subspaces and provide convergence analyses.

- We develop a restricted isometry theorem that shows how we can obtain measurement operators that have the RIP even when there are infinitely many subspaces.

- We give two examples for infinite unions of subspaces that are relevant in communication systems – the DOA manifold and the 3GPP single-path model – and discuss measurement matrix construction, worst-case recovery conditions, and average-case simulation results.

I Introduction

The thesis is organized as follows. Chapters II–IV comprise the expository part of this thesis. While they do not contain any mathematical theorems, they are intended to give a certain overview of the main ideas of compressive sensing, its relation to parameter estimation, and the necessity for *infinite* union-of-subspaces models. Chapter II contains background material on compressive sensing. In Chapter III, we continue with a more detailed look at some unions of subspaces that appear in channel estimation for mobile communication system. In Chapter IV, we introduce several compressive-sensing algorithms and compare their performances when applied to the channel estimation problem. We also discuss some issues with small problem dimensions that are not covered by the theory we present in the later chapters.

Chapters V–X contain many mathematical theorems that have the ultimate goal of showing under what conditions the algorithms from Chapter IV find a good solution of the channel estimation problem from Chapter III. We introduce *approximate* projectors in Chapter V as we need these in Chapter VI to properly state the recovery theorems for the recovery algorithms. We only give sketches of the proofs and present some interesting consequences that can be derived from (RIP); the complete proofs are only shown in the Appendix.

The remaining chapters VII–X are concerned with random constructions of the measurement operator A in (1) that guarantee (RIP). In Chapter VII, we present a theory based on covering numbers and so-called *chaining* arguments that is useful to analyze suprema of random processes. This theory is far more general than needed to show that a random matrix has the RIP, but this generality actually makes the theory more accessible. A crucial ingredient for this theory is a point-wise concentration inequality, i.e., an inequality stating that for a random operator A, the inequality (RIP) is satisfied with high probability for any given element x. We show a principled approach for obtaining such an inequality in Chapter VIII. In Chapter IX, we state the main RIP theorem, i.e., a theorem stating under what conditions (RIP) holds *simultaneously* for all elements of a union of subspaces with high probability, and we give some examples showing how this theorem can be applied to recover various results that were previously stated in the literature. Finally, in Chapter X, we derive covering number estimates for two unions of subspaces appearing in channel estimation and state RIP theorems for these particular unions of subspaces. We give some concluding remarks in Chapter XI.

Some parts of this thesis have been or are in the process of being published in conference proceedings and journals. The results regarding the low-rank approximation of the single-cluster 3GPP covariance matrix can be found in [2]. The algorithms, their convergence analyses, and the notion of approximate projectors from Chapters IV–VI have been presented in similar form in [3–5]. The chaining-based RIP theory from Chapters VII–IX can be found in [6]. The RIP condition for the DOA manifold stated in Chapter X can also be found in [7].

1—Notation

We use B_E and ∂B_E to denote the unit ball and unit sphere, respectively, in the normed space E. The scalar product between elements x, y in a Hilbert space is denoted $\langle x, y\rangle$. The cardinality of a finite set T is denoted as $|T|$ and supp(x) denotes the index set of the nonzero elements of vectors $x \in \mathbb{R}^n$ or $x \in \mathbb{C}^n$. For a map $f\colon E \to F$, we write $f(E) = \{f(x), x \in E\} \subset F$. The expected value of a random variable X is denoted as $\mathbb{E}X$. The shorthand notation $\mathbb{P}[X \gtrless \pm u] \leq c$ summarizes the two inequalities $\mathbb{P}[X > u] \leq c$ and $\mathbb{P}[X < -u] \leq c$ so that $\mathbb{P}[X > u$ and $X < -u] \leq 2c$. The range of a linear operator A is denoted as range(A) and the linear span of a set $\mathcal{U}$ is denoted as span($\mathcal{U}$). For example, for two linear operators $A, B\colon \mathcal{H} \to \mathcal{H}'$, the set span$\{$range$(A)$, range$(B)\}$ is the smallest subspace in $\mathcal{H}'$ that contains the ranges of both operators A and B.

II
COMPRESSIVE SENSING

Compressive sensing has its roots in sparse signal processing. Sparse representations find applications in various areas of signal processing, such as compression, denoising, regularization of inverse problems, feature extraction, and so forth. Within the field of time-frequency analysis, it was found that many signals occurring in applications can be described succinctly as a superposition of only a few *time atoms* and *frequency atoms*, i.e., they can be compressed. Specifically, signals that are not sparse, i.e., compressible, in either the time domain or in the frequency domain alone, admit a sparse representation in the union of these two *incoherent* orthonormal bases. The goal is to find a representation of a signal y of the form

$$y = \sum_{i=1}^{M} \alpha_i u_i + \sum_{i=1}^{M} \beta_i v_i \tag{4}$$

with as few nonzero α_i and β_i as possible and where $\{u_i : i = 1, \dots, M\}$ and $\{v_i : i = 1, \dots, M\}$ are two incoherent orthonormal bases in $\mathbb{C}^M$. This equation can be written as

$$y = Ax = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} . \tag{5}$$

As the least-squares solution to this problem does not yield the sparsest possible solution, one cannot simply apply the pseudo-inverse of the matrix $[U, V]$ to find the optimal coefficients α, β. On the other hand, picking the sparsest of all possible combinations of basis coefficients is computationally intractable. It was found in [8] that ℓ_1-minimization, i.e., solving the problem,

$$\min_x \|x\|_1 \quad \text{s.t. } y = Ax \tag{L1}$$

is an efficient way to find a reasonably sparse solution to the representation problem. In subsequent papers, conditions on when (L1) yields a sparse solution were given in terms of the *coherence* of the orthonormal bases [9]. The coherence μ of two bases U and V is defined as

$$\mu = \max_{i,j=1,\dots,M} |\langle u_i, v_j \rangle| \tag{6}$$

and describes the maximal linear dependence between basis vectors of different bases. These results were first generalized to arbitrary numbers of bases [10, 11] and then to general matrices A [12, 13]. The coherence of a general matrix $A \in \mathbb{C}^{m\times M}$ with columns a_i is defined as

$$\mu = \max_{i,j=1,\ldots,M, i\neq j} |\langle a_i, a_j\rangle| \tag{7}$$

which is no different from (6) where one can skip evaluating the scalar product between vectors of the same orthonormal basis. It has, thus, become possible to use the sparse-approximation framework to solve arbitrary systems of linear equations under sparsity constraints and even if they are disturbed by noise [14, 15].
However, while the coherence of a matrix is a good proxy for predicting success of a sparse-approximation algorithm if the matrix consists of several orthonormal bases, it is an overly pessimistic proxy for general matrices. In fact, the recovery conditions based on coherence, which usually state that $(k-1)\mu$ must be small to recover a k-sparse signal, are strictly stronger than the condition (RIP) with $\mathcal{U} = \Sigma_k$ where

$$\Sigma_k = \{x \in \mathbb{C}^M : |\,\text{supp}(x)| \leq k\}$$

is the set of vectors with at most k nonzero entries [1] (one can show that $\delta \leq (k-1)\mu$ by using the Gersgorin circle theorem).
The RIC is a rescaled upper bound $\kappa(A, \Sigma_k)$ of the condition number of all submatrices of A with only k columns: $\kappa(A, \Sigma_k)^2 = (1 + \delta(A, \Sigma_k))(1 - \delta(A, \Sigma_k))^{-1}$ [16, 17]. The introduction of the RIP and, with it, the passage from pairs of bases to general matrices A sparked a flurry of research into *compressive sensing*. Initially, it has been found that the solution of (L1) is the correct sparse solution if (RIP) holds and that surprisingly many matrices have the RIP [1, 15, 18]: Loosely speaking, most random matrices have the RIP provided that the number of rows is larger than $ck \log M$, where M is the number of columns and c a constant independent of k and M.
After the initial works [1, 15, 18], many results were published in which exact and stable recovery of sparse or approximately sparse vectors is guaranteed if the RIC is small enough. It is now known that ℓ_1-recovery, i.e., solving (L1), is successful if $\delta(A, \Sigma_{2k}) < 1/\sqrt{2}$ [19]. On the other hand, ℓ_1-recovery can fail if $\delta(A, \Sigma_{2k}) > 1/\sqrt{2}$ [17].
In addition to ℓ_1-minimization, RIP-based recovery results were developed for a multitude of other iterative and non-iterative algorithms that find the sparsest solution to (1). These include orthogonal matching pursuit (OMP), which needs $\delta(A, \Sigma_{k+1}) < 1/(1 + \sqrt{k})$ [20], iterative hard thresholding (IHT) and hard thresholding pursuit (HTP), which need $\delta(A, \Sigma_{3k}) < 1/\sqrt{3}$ [21, 22], and Compressed Sampling Matching Pursuit (CoSaMP), which needs $\delta(A, \Sigma_{4k}) < 0.478$ [23].
To verify whether a given matrix A has the RIP for Σ_k is a difficult (NP-hard) problem by itself [24]. It can be solved, in general, only by calculating the maximal and minimal singular values of all possible submatrices with k columns. Also, no deterministic constructions of matrices are known that satisfy a given RIC condition for a fixed number of columns and with a minimal number of rows (i.e., as few

as is possible with random constructions). What is known is that partial Fourier matrices can be used for recovering k-sparse vectors $x \in \mathbb{C}^M$ with high probability if M is large and if the number of rows is greater than $ck \log M$ where c is a constant [18].
Besides Fourier matrices, some asymptotic results for certain random matrices have been proved: if we simply construct matrices A out of identically distributed Gaussian random variables, then such a matrix has the RIP with high probability as soon as the number of rows is greater than $ck \log M$, where c is some constant [1, 15, 25]. These results have been generalized, for example, to random matrices with Toeplitz structure [26], and to products of random matrices with matrices that have the RIP [27].
Any given signal $x \in \Sigma_k$ can be described perfectly by a vector of length $2k$ containing the k nonzero values of x and its locations. To find this compressed representation of any given $x \in \Sigma_k$, the vector x has to be known as one has to find the nonzero entries of x among all of its entries. In contrast, the results from compressed sensing suggest that one can also describe x by a vector y with $ck \log M$ entries by applying a linear measurement operator A to x.
It has been found that in many practical applications, such linear measurement operators can be realized in hardware in the analog domain. This is important if the physical resources are limited and the number of measurements should be kept small. Examples are pilot-based channel estimation where physical resources are the number of channel accesses or medical imaging where the physical resources are related to radiation dosages. In both cases, there is a tremendous interest in keeping low the amount of physical resources needed to reconstruct the signal of interest.

1—Compressive sensing in unions of subspaces

The question arises whether the theory developed within compressive sensing is useful for other than sparse signals. The first example is the multiple measurement vector (MMV) model, in which multiple systems of equations,

$$y_\ell = Ax_\ell + e_\ell, \quad \ell = 1, \dots, Q, \tag{8}$$

are given and where it is known that all unknown vectors x_ℓ exhibit exactly the same sparsity structure. When formulating this problem as a single, big joint recovery problem, the solution is known to have at most Q times k nonzero coefficients and, in addition, these occur in groups of size Q. This additional *block sparsity* information has been studied within the compressive sensing framework and led to the introduction of the block restricted isometry constant [28]. The block RIC (B-RIC) is defined just as the RIC with the set Σ_k of all k-sparse vectors replaced by the set $\mathcal{B}_k$ of all k-block-sparse vectors. It is relatively straightforward to extend results and algorithms for sparse recovery to block-sparse recovery [29, 30].
A further generalization was the introduction of model-based compressive sensing [31]. The idea is to incorporate any structural prior information into the com-

pressive-sensing-based reconstruction that can be expressed in terms of allowed support sets. A *model* $\mathcal{M}_k \subset \mathbb{C}^N$ is a set of vectors that are k-sparse and whose supports satisfy additional constraints. For example, one could require that between any two nonzero entries of the vector x there be at least some number of zeros, i.e., a minimum separation constraint. One can then define the *model-based RIC* (M-RIC) $\delta(A, \mathcal{M}_k)$ of a matrix A as the smallest constant that satisfies (RIP) for all vectors x in $\mathcal{M}_k$ (instead of Σ_k).
In standard compressive sensing with sparse vectors, the constraint set Σ_k can be written as

$$\Sigma_k = \bigcup_{I \subset \{1,\ldots,M\}, |I|=k} S_I \quad \text{with } S_I = \Big\{ x \in \mathbb{C}^M : \text{supp}(x) = I \Big\}. \tag{9}$$

As each S_I is a k-dimensional subspace, Σ_k is a *union of subspaces*. We can write the sets of all block-sparse signals $\mathcal{B}_k$ and that of all signals with structured sparsity also as (9) and simply restrict the allowed index sets I to more specific subsets. Other examples of finite unions of subspaces are given as

$$\mathcal{D}_k = D\Sigma_k = \Big\{ x : x = Dz, z \in \Sigma_k \Big\}, \tag{10}$$

for a *dictionary* matrix D. This is sometimes called *signal-space* compressive sensing [32]. While this may appear to be a special case of the standard compressive sensing problem with an *effective* sensing matrix AD, the two restricted isometry conditions

$$(1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2 \quad \forall x \in D\Sigma_k\,, \tag{11}$$

which is sometimes called the D-RIP condition, and

$$(1-\delta)\|z\|^2 \leq \|ADz\|^2 \leq (1+\delta)\|z\|^2 \quad \forall z \in \Sigma_k \tag{12}$$

differ drastically if D has coherent columns.
As in standard compressive sensing with sparse vectors, one cannot easily verify whether a given matrix has the RIP. However, the probabilistic results for random sensing matrices can be generalized to arbitrary unions of subspaces: A random matrix with (sub-)Gaussian entries has the RIP with high probability if the number of rows grows as the logarithm of the number of subspaces, which is consistent with the standard sparsity case (where the number of subspaces is given by $\binom{M}{k} \leq (eM/k)^k$) [21].
Compressive sensing algorithms that employ the *hard thresholding operator* H_k to account for the sparsity constraint can be trivially extended to the general union-of-subspaces setting. The key observation is that the hard thresholding operation solves the best approximation problem

$$H_k(x) = \underset{x' \in \Sigma_k}{\arg\min} \|x' - x\|^2. \tag{13}$$

Algorithms such as iterative hard thresholding (IHT), normalized IHT, hard thresholding pursuit (HTP), and compressive sampling matching pursuit (CoSaMP) can be generalized by simply replacing the set Σ_k in (13) with a general union of subspaces. The caveat is that the best approximation problem, which has to be solved repeatedly, may be much more difficult to solve than hard thresholding.
Greedy algorithms such as the orthogonal matching pursuit (OMP) can only be used with general unions of subspaces if these exhibit similar hierarchical structures as sparse signals. Convex relaxation algorithms like ℓ_1-minimization have not been studied under the viewpoint of reconstruction in general unions of subspaces.

2—Infinite unions of subspaces and parameter estimation

Unions of subspaces also appear naturally in many parameter estimation problems of the form

$$y = A \sum_{\ell=1}^{k} x_\ell f(t_\ell) + e \tag{14}$$

where $f\colon \mathbb{R} \to \mathbb{C}^M$ is a known, nonlinear function of an unknown parameter and where $A\colon \mathbb{C}^M \to \mathbb{C}^m$ is a linear measurement operator.
If we construct a dictionary matrix D out of columns $f(s_j)$ corresponding to a finite set of grid points $s_j, j = 1, \ldots, N$, we can write (14) as the linear system of equations $y = ADz + e$ with the sparsity constraint $z \in \Sigma_k$. One can then use the compressive sensing framework to solve this equation for z if AD has the standard RIP and recover the parameters t_i from the nonzero indices of z. However, if a fine grid is used to obtain a high resolution, the columns of D may become strongly correlated as f is typically a continuous function. The RIP-based theory then provides an upper limit on the resolution under which recovery of the true parameters can be guaranteed.
If, on the other hand, we are interested in recovering the signal $x = \sum_{\ell=1}^{k} x_\ell f(t_\ell)$ from $y = Ax + e$, an increase of the resolution, i.e., the coherence of the dictionary matrix D, is not an issue as long as the D-RIP condition (11) is satisfied. However, the condition $m \geq c\log$(number of subspaces) is problematic, because the number of subspaces is given by $\binom{N}{k}$ where N is the number of grid points. This number grows to infinite as the resolution is increased.
Intuitively, if the D-RIP condition (11) is already satisfied for a very fine grid, i.e.,

$$\left\|\sum_{\ell=1}^{k} x_\ell f(t_\ell)\right\|^2 \approx \left\|A\sum_{\ell=1}^{k} x_\ell f(t_\ell)\right\|^2 \tag{15}$$

for $t_1, \ldots, t_k$ on-grid, and if $\tilde{t}_1, \ldots, \tilde{t}_k$ are off-grid, but close to the grid, then

$$\tilde{x} = \sum_{\ell=1}^{k} x_\ell f(\tilde{t_\ell}) \approx \sum_{\ell=1}^{k} x_\ell f(t_\ell) \tag{16}$$

and, by continuity of the linear operator A, also

$$Ax \approx A\tilde{x} \tag{17}$$

so that if the D-RIP condition is satisfied on a fine grid, then it is probably also almost satisfied between grid points. To overcome these grid-related problems, we introduce an *infinite* union of subspaces $\mathcal{U} = \cup_{t \in T} S_t$ where each S_t is a subspace and where T is a (possibly infinite) parameter set. In this case, we can use, for example, $T = [1, 1]^k$ and

$$S_t = \text{range}\big(f(t_1), \ldots, f(t_k)\big). \tag{18}$$

As with the generalization from sparse signals to signals in finite unions of subspaces, algorithms such as IHT can be generalized without problems as long as one can calculate the best approximation operator [33]

$$\mathcal{P}_{\mathcal{U}}(x) = \arg\min_{z \in \mathcal{U}} \|x - z\|. \tag{19}$$

III
CHANNEL ESTIMATION

In this chapter, we introduce examples of infinite unions of subspaces that occur in multi-antenna communication systems (and also in radar systems). All models use the *steering vector* of a uniform linear array (ULA) as a building block. The steering vector

$$a(t) = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & \exp(it) & \dots & \exp(i(M-1)t) \end{bmatrix}^T, \; t \in [-\pi, \pi], \tag{20}$$

describes the signal received at a ULA of M antennas with half-wavelength spacing when a single planar wavefront impinges on the array from direction $\theta = \arcsin(t/\pi)$.

1—The DOA manifold

If, not one, but k harmonic signals impinge on a ULA of M antennas with half-wavelength spacing, the receive signal can be written as the superposition

$$x = \sum_{\ell=1}^{k} \alpha_\ell a(t_\ell) \tag{21}$$

with $\alpha_\ell \in \mathbb{C}$ and $t_\ell = \pi \sin(\ell\text{th angle})$. The standard direction-of-arrival (DOA) estimation problem is to recover the unknown angles from a disturbed version $y = x + e$ of the receive signal x. The additive disturbance $e \in \mathbb{C}^M$ is usually modeled as Gaussian noise.

In addition to DOA estimation, this signal model is relevant for future mobile communication systems, which use millimeter-wave frequencies and many antennas [34]. There, it is commonly referred to as the *geometric channel model* and motivated by a propagation behavior resembling that of light: electromagnetic waves at millimeter-wave frequencies do not pass through walls or bend around corners, but are rather reflected at objects [35, 36]. Thus, at least at the base station, which is usually located at some exposed location with few nearby scatterers, the channel can be modeled as the superposition of relatively few paths from distinct angular directions [34]. In contrast to DOA estimation where the goal is to recover the angles, in channel estimation the goal is to recover x from its disturbed measurements.

If we define the matrix $V(t) = [a(t_1) \dots a(t_k)] \in \mathbb{C}^{M\times k}$ and the vector $\alpha =$

$[\alpha_1 \, \ldots \, \alpha_k]^T$, we can write (21) as

$$x = \sum_{\ell=1}^{k} \alpha_\ell a(t_\ell) = V(t)\alpha. \tag{22}$$

We call the set of all possible signals x, which is given by

$$\mathcal{U}_{k,M} = \overline{\mathcal{U}'_{k,M}}, \quad \mathcal{U}'_{k,M} = \bigcup_{t \in [-\pi,\pi]^k} \text{range}\big(V(t)\big), \tag{23}$$

the DOA manifold (we take the closure for reasons we discuss later). For each t, the subspace range($V(t)$) is (at most) k-dimensional while the ambient space $\mathbb{C}^M$ has M dimensions. Consequently, the DOA manifold is a collection of low-dimensional subspaces in a high-dimensional ambient space.

2—3GPP and conditionally normal channel models

The geometric channel model from the previous section, in which the channel vector is described as a superposition of k steering vectors, assumes very high frequencies, many antennas, and the lack of scattering objects close to the antennas. The channel models proposed by the 3GPP are derived under less demanding assumptions [37]. The point scatterers in the geometric model, which lead to the equation

$$x = \sum_{\ell=1}^{k} \alpha_\ell a(t_\ell), \tag{24}$$

are replaced by *clusters* of infinitely many scatterers and an equation of the form

$$x = \sum_{\ell=1}^{k} \int_{-\Delta_\ell/2}^{\Delta_\ell/2} \alpha_\ell(\tau) a(t_\ell + \tau) d\tau. \tag{25}$$

In (25), Δ_ℓ describes the angular spread of the ℓth cluster. The path gain α_ℓ is now a function of the distance from the cluster center and is typically set as decreasing with distance. In stochastic versions of these channel models, the path angles or cluster centers t_ℓ are assumed to be uniformly distributed. The channel vectors, if conditioned on the angles t_ℓ, are assumed to be complex-normal distributed with mean zero. The covariance matrices of the channel vectors are then given as

$$\begin{aligned}\Sigma = \mathbb{E}xx^H = \mathbb{E} \sum_{\ell,j=1}^{k} \alpha_\ell \alpha_j^* a(t_\ell) a(t_j)^H \\ = \sum_{\ell,j=1}^{k} \mathbb{E}[\alpha_\ell \alpha_j^*] a(t_\ell) a(t_j)^H = \sum_{\ell=1}^{k} \sigma_\ell^2 a(t_\ell) a(t_\ell)^H \end{aligned} \tag{26}$$

for the geometric model from the previous section and

$$\Sigma = \mathbb{E}xx^H = \mathbb{E}\sum_{\ell,j=1}^{k}\int_{-\Delta_\ell/2}^{\Delta_\ell/2}\int_{-\Delta_j/2}^{\Delta_j/2}\alpha_\ell(\tau)\alpha_j^*(\tau')a(t_\ell+\tau)a(t_j+\tau')^H d\tau d\tau' \tag{27}$$

$$= \sum_{\ell,j=1}^{k}\int_{-\Delta_\ell/2}^{\Delta_\ell/2}\int_{-\Delta_j/2}^{\Delta_j/2}\mathbb{E}\big[\alpha_\ell(\tau)\alpha_j^*(\tau')\big]a(t_\ell+\tau)a(t_j+\tau')^H d\tau d\tau' \tag{28}$$

$$= \sum_{\ell=1}^{k}\int_{-\Delta_\ell/2}^{\Delta_\ell/2}\mathbb{E}\big[|\alpha_\ell(\tau)|^2\big]a(t_\ell+\tau)a(t_\ell+\tau)^H d\tau \tag{29}$$

for the 3GPP model. In both cases, we assumed independence between different paths or cluster centers, $\mathbb{E}[\alpha_\ell\alpha_j] = 0$ for $\ell \neq j$, and, in the 3GPP model, independence also between different sub-paths, $\int \mathbb{E}\alpha_\ell(\tau)\alpha_\ell^*(\tau')a(t_\ell+\tau')d\tau' = \mathbb{E}|\alpha_\ell(\tau)|^2 a(t_\ell+\tau)$.
Both models can be subsumed under the *conditionally normal* channel model

$$x\,|\,(t_1,\dots,t_k) \sim \mathcal{N}_{\mathbb{C}}(0,\Sigma_t)\,,\quad \Sigma_t = \sum_{\ell=1}^{k}\Sigma_{t_\ell} \tag{30}$$

with

$$\Sigma_{t_\ell} = \int g_\ell(\tau)a(t_\ell+\tau)a(t_\ell+\tau)^H d\tau \tag{31}$$

and with the *angular power profile* g_ℓ. In the geometric model, $g_\ell(t) = p_\ell\delta(t)$ is given as the Dirac delta function (a point measure) and in the 3GPP models,

$$g_\ell(\tau) = p_\ell\exp(-|\tau|/\Delta_\ell) \tag{32}$$

is typically a Laplace density (the factors p_ℓ can be used to describe a nonuniform power distribution between different clusters and the standard deviations Δ_ℓ are often set to two or five degrees).[1]
For given path angles or cluster centers t, the channel vector x lies in the subspace $S_t = \text{range}(\Sigma_t)$ with probability one. In the geometric channel model where the covariance matrix is given as the sum of k rank-one matrices of the form $a(t_\ell)a(t_\ell)^H$, the subspace S_t is k-dimensional and we recover the union-of-subspaces model $x \in \mathcal{U} = \cup_{t\in T}S_t$, where T is the set of all possible angle combinations.

[1] In this work, we parametrize the steering vectors by sin(angle) instead of the angle. Consequently, the angular power profile is a function of the difference between sin(angle) and sin(cluster center) instead of angle and cluster center as in the actual 3GPP channel models. This represents a mild departure from the 3GPP models, which is not very severe if only cluster centers between -60° and $+60^\circ$ are assumed. We do not attempt to quantify this additional approximation error, because also the 3GPP models do not represent a ground truth, but are rather only approximate models of reality.

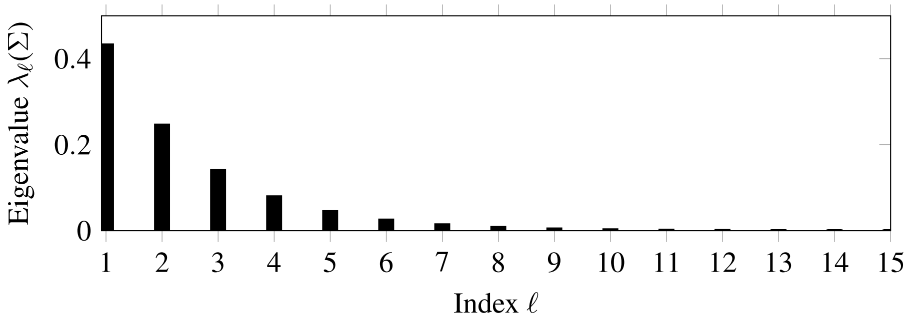

Figure 1: First 15 sorted eigenvalues of Σ_0 for $M = 64$ antennas and $\Delta = 2°$

In the 3GPP model, the covariance matrices Σ_t may be full rank and we do not gain much from exploiting the union-of-subspaces structure, because the subspaces are not low-dimensional. However, if $g_\ell(\tau) \approx 0$ for most τ, we can approximate Σ_t by a low-rank covariance matrix and recover a structure with low-dimensional subspaces: Let $t_\ell = 0$ and let

$$[\Sigma]_{m,m+n} = \frac{1}{M} \int g(\tau)[a(\tau)]_m[a^*(\tau)]_{m+n} d\tau \tag{33}$$

denote the entry in row m and column $m + n$ of one of the summands in (30). If we plug in expression (20) for the steering vectors, we obtain

$$\begin{aligned}[\Sigma]_{m,m+n} &= \frac{1}{M} \int g(\tau) \exp(-i\pi n\tau) d\tau & (34)\\ &= \frac{p_\ell}{M} \int \exp(-|\tau|/\Delta - i\pi n\tau) d\tau = \frac{2p_\ell\Delta}{M(1 + (\Delta n\pi)^2)}. & (35)\end{aligned}$$

The significant eigenvalues of this matrix are shown in Figure 1 for $M = 64$ antennas and for $\Delta = 2°$. If Δ is small, each summand in (30) can be approximated by a low-rank covariance matrix $\widetilde{\Sigma}$ (e.g., the eight strongest eigenvectors of Σ for $\Delta = 2°$ and $M = 64$). As we show in Chapter X, for covariance matrices Σ_t with $t \neq 0$, the eigenvalue distribution is the same as that of Σ_0 (as we chose the Laplacian to be a function of sin(angle) instead of the angle). Then, as for the geometric model, the channel vector x lies *approximately* in a low-dimensional subspace

$$S_t = \text{span}\{\text{range}(\widetilde{\Sigma}_{t_1}), \dots, \text{range}(\widetilde{\Sigma}_{t_k})\} \tag{36}$$

and the expected value of the squared approximation error is controlled by the largest eigenvalue of the difference matrix $\sum_{\ell=1}^{k}(\Sigma_{t_\ell} - \widetilde{\Sigma}_{t_\ell})$.

3—Spatio-temporal channel models

The channel models described in the previous sections are instances of *spatial* channel models. They describe the correlation of the receive signals between different antennas recorded at the same time. In fact, we completely ignored any

notion of time – the signal model based on steering vectors assumes a source emitting a constant sine wave of infinite duration. Therefore, we ignored that paths of different scatterers can have different delays. In communication systems, the source changes the amplitude and phase of the sine wave every so many periods. If the difference between the minimal and maximal possible path delays is larger than the period at which the source changes its signal, the path delays become important. Whether this is the case depends on the rate at which the source changes its signal – the communication bandwidth – and the propagation environment.
Let us consider a discrete-time signal model, i.e., after sampling at the Nyquist rate, in which the minimal path delay is zero (after synchronization) and the maximal path delay is $Q-1$ samples. It is common to assume that the receive signal x (at a single antenna) is given as the convolution of the transmit signal $s \in \mathbb{C}^P$ of duration P (the pilot) with the channel impulse response $h \in \mathbb{C}^Q$, which is a vector of Q samples (called *taps*) with each entry corresponding to a path delay, i.e., $x = h * s \in \mathbb{C}^{Q+P-1}$.
In wideband systems, Q can be much larger than the number of propagation paths and the resulting channel impulse response h contains many zeros [38]. If k is the maximal number of paths constituting any given channel impulse response, we have $h \in \Sigma_k$ (k-sparse signals), i.e., a union-of-subspaces model. In addition to millimeter-wave systems, long impulse responses with many zeros are also common in underwater communication systems [39, 40].
Things become interesting once we combine this *temporal* model with any of the spatial models from the previous section. If the receive antenna array is not too large, we can assume that the path delays between the source and each of the receive antennas is approximately the same (the additional propagation delay between the first and the last antenna of the receive array is assumed to be small compared to the sampling rate f_s, that is $M/2 \ll f_c/f_s$ where f_c is the carrier frequency). Consequently, if we let $h_j[n]$ denote the nth tap of the channel impulse response between the source and the jth antenna and set $h[n] = [h_1[n] \ \dots \ h_M[n]]^T$, we obtain $h[n] \in \mathcal{U}$ for each $n = 0, \dots, Q-1$, and where $\mathcal{U}$ is any of the union of subspaces from the previous section, e.g., the DOA manifold. In addition, the sparsity constraint $h_j = [h_j[0] \ \dots \ h_j[Q-1]] \in \Sigma_k$, must hold for all antennas j *simultaneously* (as the path delays are the same for all antennas, each h_j must have the same (temporal) support). The overall model is given as $h \in \mathcal{M}$ with

$$\mathcal{M} = \bigcup_{i_1,\dots,i_k \in \{0,\dots,Q-1\}, i_j \neq i_\ell \text{ if } j \neq \ell} \mathcal{U}_{i_1} \oplus \cdots \oplus \mathcal{U}_{i_k} \tag{MA-UOS}$$

where

$$\mathcal{U}_q = \{h : h[n] = 0 \text{ if } n \neq q \text{ and } h[q] \in \mathcal{U}\}. \tag{37}$$

From (37) it follows that each non-zero tap $h[n]$ must lie in the spatial union of subspaces $\mathcal{U}$ and from (MA-UOS) it follows that there may only be k nonzero taps. It is not difficult to verify that $\mathcal{M}$ is also a union of subspaces. We discuss unions of subspaces of this particular form in more detail in Chapter V.

4—Channel estimation with linear measurements

We have seen in the previous section that the per-antenna channel impulse responses can only be observed through their convolution with a pilot signal. This can be described as a linear measurement. Similarly, in millimeter-wave systems the spatial per-tap channel vector is not observed directly, but through some form of linear measurement. This is because cost and power constraints render it likely that not all antennas of the base station are connected to their own analog-to-digital converters (ADCs). Instead, an analog network with, for example, phase shifters, inverters, and combiners is used to connect the M antennas with m ADCs where m is potentially much smaller than M.

Let us describe the complete spatio-temporal measurement equation for a block-based communication system with a single-antenna transmitter and a receiver with M antennas and m ADCs. If we denote by $x[n] \in \mathbb{C}^M$ the noise-free (analog) receive signal at the antennas at time instance n, then the signal at the output of the analog network (the digital receive signal) is given by

$$y[n] = Ax[n] + e[n] \in \mathbb{C}^m \tag{38}$$

where A contains only zeros and constant-modulus entries (and is constant over time). The noise e is added to the signal x only after application of the measurement operator A, because the dominant noise sources are the ADCs (i.e., after the signal propagated through the analog network).

The communication channel is estimated during a training phase in which a *pilot signal* $s = [s[0] \, \dots \, s[P-1]] \in \mathbb{C}^P$ is transmitted that is known to the receiver. If we insert the convolutional channel model $x = h * s$, we obtain the linear measurement equation

$$y[n] = A(h * s)[n] + e[n], \; n = 0, \dots, Q + P - 1 \tag{39}$$

where Q is the length of the channel impulse response h. For white and spatially uncorrelated Gaussian noise e of known variance, the maximum likelihood channel estimate is obtained by solving

$$\hat{h}_{\text{ML}} = \arg\min_{h \in \mathcal{M}} \sum_{n=0}^{Q+P-1} \|y[n] - A(h * s)[n]\|^2 \tag{40}$$

where $\mathcal{M}$ is the union of subspaces given by (MA-UOS). This problem is of the form (NLS) with the linear measurement operator $h \mapsto (A(h * s)[n])_{n=0,\dots,Q+P-1}$.

IV
ALGORITHMS AND SIMULATION RESULTS

In this chapter, we generalize several algorithms that were developed for recovering sparse signals from compressed measurements to infinite unions of subspaces. These algorithms attempt to recover $x \in \mathcal{U}$, where $\mathcal{U}$ is a union of subspaces, from the compressed measurements $y = Ax + e$, where e is noise, by solving the nonlinear least-squares problem

$$\hat{x} = \arg\min_{x \in \mathcal{U}} J(x), \quad J(x) = \frac{1}{2}\|y - Ax\|^2. \tag{41}$$

The nonlinear constraint $x \in \mathcal{U}$ renders this problem difficult. All algorithms presented below generate a sequence of estimates x_n and residuals $r_n = y - Ax_n$. Because the direction of steepest descent (the negative gradient) of J evaluated at x_n is given by $\nabla J(x_n)\colon x_n \mapsto A^* r_n = A^*Ay - A^*Ax_n + A^*e$, the algorithms presented below use the adjoint of A to manipulate the residual.

1—Random matrices and small problem dimensions

Roughly speaking, the performance of the algorithms we present below depends on the RIP of the matrix A; the smaller the RIC, the better the results. The convergence analysis, which we present in Chapter VI, shows that the worst-case error bound can be described as a function of $\delta(A, \mathcal{U})$. We discussed in Chapter II that, for a given matrix A, it is not easy to calculate $\delta(A, \mathcal{U})$ (NP-hard problem) and that the only efficient known way to obtain a matrix A that *probably* has the RIP is by means of random sampling. The following definitions introduce some of the most common types of random matrices.

Definition 1 (Complex Gaussian and zero-inflated Steinhaus matrices)**.** *Let* $A \in \mathbb{R}^{m\times M}$ *or* $A \in \mathbb{C}^{m\times M}$ *be a random matrix with independent entries* $a_{\ell j}, \ell = 1, \ldots, m$, $j = 1, \ldots, M$.

- *If* $a_{\ell j} \sim \mathcal{N}(0, 1/m)$ *is normally distributed, then* A *is a Gaussian random matrix.*
- *If* $a_{\ell j} \sim \mathcal{N}_{\mathbb{C}}(0, 1/m)$ *is complex-normally distributed, then* A *is a complex Gaussian random matrix.*
- *If* $\mathbb{P}[a_{\ell j} = \sqrt{c/m}] = \mathbb{P}[a_{\ell j} = -\sqrt{c/m}] = 1/(2c)$ *and* $\mathbb{P}[a_{\ell j} = 0] = 1 - 1/(2c)$ *with* $1 \leq c \leq 3$*, then* A *is a zero-inflated Rademacher random matrix.*

- *If* $a_{\ell j} = \sqrt{c/m} b_{\ell j} \exp(i u_{\ell j})$ *with* $u_{\ell j} \sim \mathcal{U}[-\pi, \pi]$ *uniformly distributed and* $\mathbb{P}[b_{\ell j} = 1] = 1/c = 1 - \mathbb{P}[b_{\ell j} = 0]$ *with* $1 \leq c \leq 2$*, then* A *is a zero-inflated Steinhaus random matrix.*

In all cases, the scaling is such that $\mathbb{E}\|Ax\|^2 = \|x\|^2$. We analyze these matrices in some more detail in Chapter VIII and we show in Chapter IX under what conditions on m, M, and on the union of subspaces $\mathcal{U}$ they have the RIP. However, we will find that m and M need to be rather large if we want to guarantee that A has the RIP with high probability (in Chapter X we show some numbers that are relevant for the channel estimation problem from Chapter III).

For small problem dimensions and some simple unions of subspaces (e.g., sparse signals with small sparsity order), we can evaluate the restricted isometry constants of random matrices by brute force. But before we show some exemplary results, we discuss *weighted norms* and alternative formulations of the restricted isometry property.

All of the algorithms we present below are formulated in an abstract Hilbert space setting where $A\colon \mathcal{H} \to \mathcal{H}'$ is a continuous linear operator between Hilbert spaces. Here and in the following, all norms and scalar products are to be understood as the standard norms in these Hilbert spaces; operator norms always refer to the canonical strong operator norms. Of course, in all applications, we will then replace $\mathcal{H}$ and $\mathcal{H}'$ by some finite-dimensional spaces $\mathbb{C}^m$ and $\mathbb{C}^M$ and, if we use the standard Euclidean norm and matrix representations of A, everything works fine and we can replace A^* by the (hermitian) transpose matrix A^H. However, we can also use weighted norms in $\mathbb{C}^m$ or $\mathbb{C}^M$, i.e., for $y \in \mathbb{C}^m$, we can use

$$\|y\|_W^2 = y^H W y \tag{42}$$

for a positive definite *weighting matrix* $W \in \mathbb{C}^{m\times m}$. The algorithms and corresponding theory remain completely unaffected except that we now need to be careful *not to replace* the adjoint by the hermitian transpose as

$$\langle Ax, y\rangle_W = y^H W A x = (A^H W y)^H x = \langle x, A^H W y\rangle_{\mathbb{C}^M}. \tag{43}$$

That is, if we use $\mathcal{H} = \mathbb{C}^M$ with standard Euclidean norm and $\mathcal{H}' = \mathbb{C}^m$ with the W-weighted norm, we have to use $A^* = A^H W$ instead of A^H. For example, if we use $W = c(AA^H)^{-1}$ for some $c > 0$, we obtain $A^* = A^H W = cA^H(AA^H)^{-1} = cA^\dagger$, where $A^\dagger$ is the Moore-Penrose pseudo-inverse of A. As

$$\|Ax\|_W^2 = x^H A^H W A x = c x^H A^H (AA^H)^{-1} A x = c\|P_{\text{range}(A^H)} x\|^2, \tag{44}$$

the restricted isometry condition becomes

$$(1-\delta)\|x\|^2 \leq c\|P_{\text{range}(A^H)} x\|^2 \leq (1+\delta)\|x\|^2 \tag{45}$$

if expressed in the standard (non-weighted) norms. If the random matrix A is chosen in a way that range(A^H) is uniformly distributed on the Grasmannian manifold Gr($m, \mathbb{C}^M$) – the set of all m-dimensional subspaces in $\mathbb{C}^M$ (or $\mathbb{R}^M$) – then

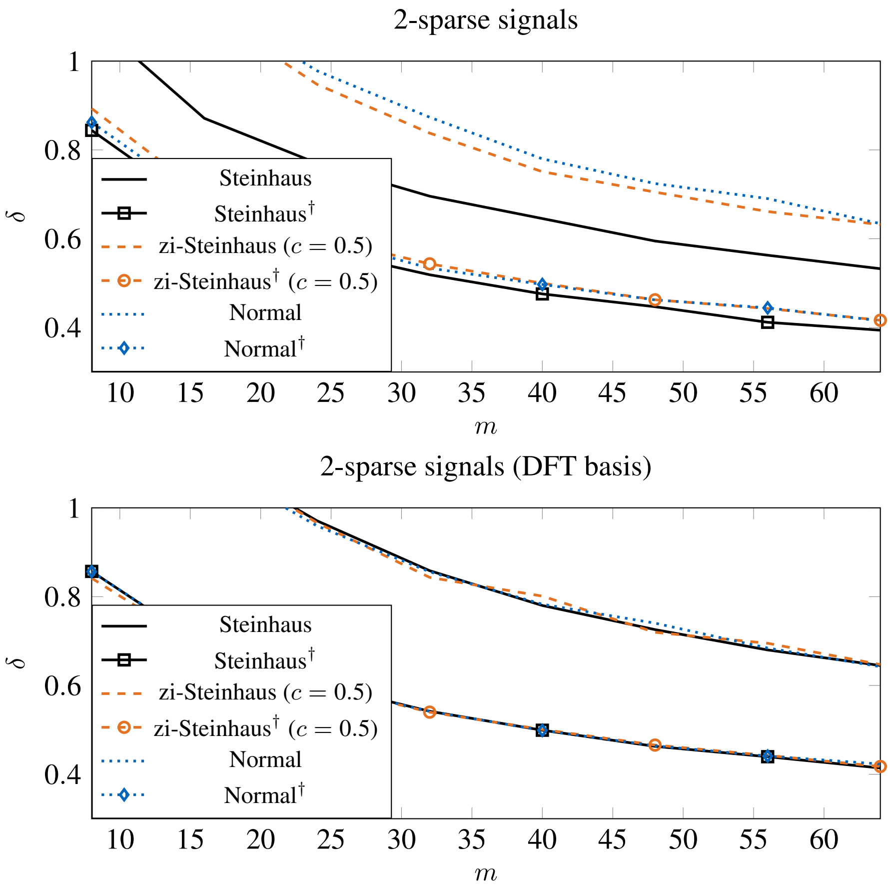

Figure 2: Mean value of the RIC for 2-sparse signals (above) and 2-sparse signals with respect to a DFT dictionary (below) over 100 realizations of different random matrices $A \in \mathbb{C}^{m\times 2m}$ and for the weighting matrix $W = I$ and $W = 2(AA^H)^{-1}$ (marked with †).

$c\mathbb{E}\|P_{\text{range}(A^H)}x\|^2 = \|x\|^2$ if $c = M/m$ (this is the case, for example, if A is a (complex) Gaussian random matrix [41]), i.e., $c = M/m$ is a good scaling for $W = c(AA^H)^{-1}$.

Figure 2 and 3 show the different simulated restricted isometry constants for 2 and 3-sparse signals in the Euclidean basis and in the discrete Fourier transform (DFT) basis when A is either a (zero-inflated) Steinhaus random matrix or a complex normal random matrix (we take the mean RIC from 100 realizations). The restricted isometry constants with respect to the weighted norms are marked with the dagger symbol †. For such small problem dimensions, the weighting matrix W leading to the pseudo inverse for the adjoint operator is clearly a preferable choice in terms of the RIC. One can also observe that non-zero-inflated Steinhaus random matrices have an edge over Gaussian random matrices for standard sparse signals (for small sparsity orders).

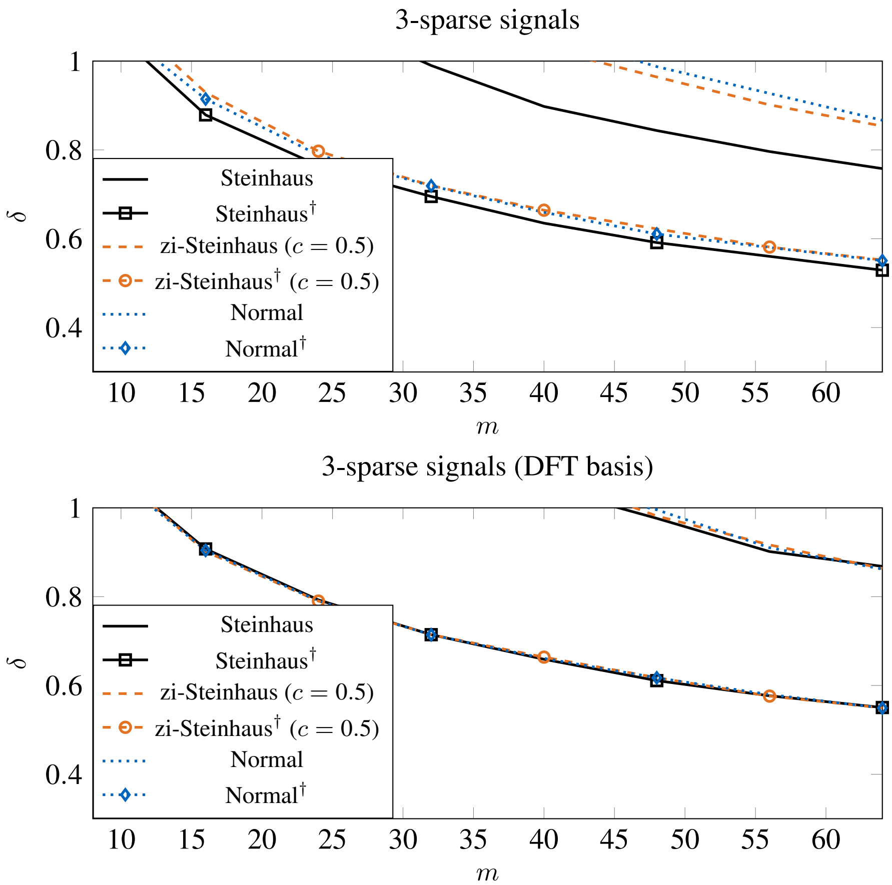

Figure 3: Mean value of the RIC for 3-sparse signals (above) and 3-sparse signals with respect to a DFT dictionary (below) over 100 realizations of different random matrices $A \in \mathbb{C}^{m\times 2m}$ and for the weighting matrix $W = I$ and $W = 2(AA^H)^{-1}$ (marked with †).

2—Projected gradient descent

One of the simplest algorithms to solve a constrained optimization problem is the projected gradient descent (PGD) method, which adds a projection step after the gradient step and is implemented as

$$z_{n+1} = x_n - \mu \nabla J(x_n)\,, \tag{46}$$

$$x_{n+1} = \mathcal{P}_{\mathcal{U}}(z_{n+1}) = \operatorname*{arg\,min}_{x\in\mathcal{U}} \|x - z_{n+1}\| \tag{47}$$

and initialized with $x_0 = 0$ and with step size μ. A particularity of this compressive sensing version of the algorithm is that the step size is usually fixed to $\mu = 1$. This unit step size is essential for the RIP-based convergence analysis in Chapter VI and cannot simply be replaced by a smaller step size and more iterations as this may prevent the algorithm from “jumping” from one subspace to the next. At the same time, however, the convergence analysis is based on the assumption that the RIC

Algorithm 1 Projected Gradient Descent (PGD)

1. Initialize $x_0 = 0$
2. For $n = 0, 1, \ldots$ (until convergence criterion is satisfied):
   i) $z_{n+1} = x_n + A^*(y - Ax_n)$
   ii) $x_{n+1} = \mathcal{Q}_{\mathcal{U}}(z_{n+1})$

of the matrix A is small and this is often not satisfied in practice, especially for small problem dimensions (but this is difficult to detect as mentioned in Chapter II, cf. [24]). In this case, a smaller step size is needed to ensure *stability* of the PGD algorithm. A summary of the PGD algorithm with the best approximation $\mathcal{P}_{\mathcal{U}}$ replaced by an approximate projector $\mathcal{Q}_{\mathcal{U}}$ – we detail what that is in the next chapter – is shown in Alg. 1.

3—Generalized hard thresholding pursuit

The projection step (47) of the projected gradient algorithm can be decomposed as

$$S_{n+1} = \underset{S \subset \mathcal{U}, S \text{ subspace}}{\arg\min} \|z_{n+1} - P_S z_{n+1}\| \tag{48}$$

$$x_{n+1} = P_S z_{n+1}, \tag{49}$$

where P_S denotes the *orthogonal* projector onto the subspace S. That is, we first find an optimal subspace and then project onto that subspace. The projected gradient algorithm is very inefficient if the subspaces of subsequent iterates are the same. We can speed up the projected gradient algorithm if we use the solution of the subspace-constrained (and thus linear) least-squares problem

$$\min_{z \in S} \|y - Az\|^2 \tag{50}$$

instead of the projection of z_{n+1} onto S. As in the PGD algorithm, we can use an approximate projector that returns an approximately optimal subspace (see next chapter). The resulting algorithm is called generalized hard thresholding pursuit (GHTP) and is described in Alg. 2.

4—Orthogonal matching pursuit

The orthogonal matching pursuit (OMP) algorithm is a greedy method that can be used to solve (41) if the overall union of subspaces $\mathcal{U}$ can be decomposed into the simpler unions of subspaces $\mathcal{U}_1, \ldots, \mathcal{U}_Q$,

$$\mathcal{U} = \bigcup_{i_1,\ldots,i_k \in \{1,\ldots,Q\}, i_j \neq i_\ell \text{ if } j \neq \ell} \mathcal{U}_{i_1} \oplus \cdots \oplus \mathcal{U}_{i_k} \tag{51}$$

Algorithm 2 Generalized Hard Thresholding Pursuit (GHTP)

1. Initialize $x_0 = 0$
2. For $n = 0, 1, \ldots$ (until convergence criterion is satisfied):
   i) $z_{n+1} = z_n + A^*(y - Az_n)$
   ii) $S_{n+1} \approx \arg\min_{S \subset \mathcal{U}, S \text{ subspace}} \|z_{n+1} - P_S z_{n+1}\|^2$
   iii) $x_{n+1} = \arg\min_{x \in S_{n+1}} \|y - Ax\|^2$

Algorithm 3 Orthogonal Matching Pursuit

1. Initialize $r_0 = y, \Lambda^0 = \emptyset$
2. For $n = 0, \ldots, k-1$:
   i) $\Lambda_{n+1} = \Lambda_n \cup \arg\max_{j \notin \Lambda_n} \|\mathcal{P}_{\mathcal{U}_j}(A^* r_n)\|^2$
   ii) $x_{n+1} = \arg\min_{x \in \oplus_{i \in \Lambda_{n+1}} \mathcal{U}_i} \|y - Ax\|^2$
   iii) $r_{n+1} = y - Ax_{n+1}$

that are orthogonal to each other, $\mathcal{U}_\ell \perp \mathcal{U}_j$ for $\ell \neq j$. The simplest case is when all $\mathcal{U}_\ell$ are subspaces, e.g., $\mathcal{U}_\ell = \text{span}\{e_\ell\}$ where e_ℓ are orthogonal basis vectors. A more sophisticated example was provided in Chapter III where each $\mathcal{U}_\ell$ is a per-block union of subspaces and the overall union of subspaces consists of signals that have only k nonzero blocks.

For given observations $y = Ax + e$, the OMP algorithm attempts to construct the optimal subspace S by successively adding subspaces S_ℓ from the simple unions of subspaces $\mathcal{U}_\ell$. This is performed by correlating the current (at step n) residual $r_n = y - Ax_n$ with all of the simple unions of subspaces, i.e. by calculating

$$\text{corr}_\ell = \|\mathcal{P}_{\mathcal{U}_\ell}(A^* r_n)\|^2 \tag{52}$$

for $\ell = 1, \ldots, Q$. The index ℓ_{max} that yields the maximal correlation is added to the current set of indices, $\Lambda_{n+1} = \Lambda_n \cup \{\ell_{\text{max}}\}$ and the new residual is calculated as $r_{n+1} = y - Ax_{n+1}$ where

$$x_{n+1} = \underset{x \in \oplus_{i \in \Lambda_{n+1}} \mathcal{U}_i}{\arg\min} \|y - Ax\|^2 \tag{53}$$

solves a union-of-subspaces constrained least-squares problem that has one level less complexity than the original problem: it is not necessary to determine the correct combination of k unions $\mathcal{U}_{i_1}, \ldots, \mathcal{U}_{i_k}$ out of $\binom{Q}{k}$ possible combinations.

Algorithm 4 Reduced-complexity Orthogonal Matching Pursuit

1. Initialize $r_0 = y, \Lambda^0 = \emptyset$
2. For $n = 0, \ldots, k-1$:
   i) $\Lambda_{n+1} = \Lambda_n \cup \arg\max_{j \notin \Lambda_n} \|\mathcal{P}_{\mathcal{U}_j}(A^* r_n)\|^2$
   ii) $z_{n+1} = \arg\min_{x \in \oplus_{i \in \Lambda_{n+1}} \text{span}\{\mathcal{U}_i\}} \|y - Ax\|^2$
   iii) $x_{n+1} = \mathcal{P}_{\oplus_{i \in \Lambda_{n+1}} \mathcal{U}_i}(z_{n+1})$
   iv) $r_{n+1} = y - Ax_{n+1}$

5—Reduced-complexity orthogonal matching pursuit

The reduced-complexity OMP algorithm is a version of the OMP algorithm in which the non-linear least-squares problem in step 2ii of Alg. 3 is replaced by a much simpler linear least-squares problem. The union-of-subspaces constraint is incorporated into the problem by projecting the solution of this linear least-squares problem onto the union of subspaces. The drawback, as we see in Chapter VI, is that a much stricter restricted isometry condition is necessary in order to prove convergence of the method.

6—Simulations for pilot-based channel estimation

Recall the measurement equation for channel estimation from Chapter III,

$$y[n] = A(h * s)[n] + e[n], \; n = 0, \ldots, Q + P - 1. \tag{54}$$

If we let $A_L = A$ denote the spatial measurement matrix and

$$A_R = \begin{bmatrix} s[0] & \ldots & s[P-1] & & \\ & \ddots & & \ddots & \\ & & s[0] & & s[P-1] \end{bmatrix} \in \mathbb{C}^{Q \times Q+P-1} \tag{55}$$

the convolution matrix for the temporal measurements and if $Y = [y[0] \ldots y[Q+P-1]]$, $E = [e[0] \ldots e[Q+P-1]]$, $H = [h[0] \ldots h[Q-1]]$, we can write (54) as

$$Y = A_L H A_R + E \in \mathbb{C}^{m \times P+Q-1}. \tag{56}$$

The columns of Y correspond to samples from different times. The columns of the channel matrix H must belong to a spatial union of subspaces $\mathcal{U}$ and the temporal sparsity constraint demands that H may have at most k nonzero columns.
Below, we show simulation results for the two examples of the spatial unions from Chapter III for a receive array with $M = 64$ antennas and $m = 32$ ADCs. We use the geometric channel model with three paths per tap and $k = 3$ out of $Q = 48$

nonzero taps and the 3GPP model with a single cluster per tap and a standard deviation of two degrees and a low-rank approximation of the (per-tap) covariance matrix of rank eight (cf. Figure 1). The pilot length is set to $P = 16$ and we use independent Steinhaus random variables for $s[p], p = 0, \ldots, P-1$ (without zero inflation, cf. Chapter VIII). For the spatial measurement matrix A_L, we also use a matrix composed of independent Steinhaus random variables, but with $p = 0.5$ so that approximately 50 per cent of the entries are zeros.

The per-block union-of-subspace problems are solved by using the Root-MUSIC algorithm [42] in the geometric model and by an exhaustive grid search in the 3GPP model (in the latter case, we need to find the optimal t in the interval $[-\pi, \pi]$). As a baseline for the comparison, we use the standard OMP algorithm with a four-times oversampled DFT dictionary for the spatial union of subspaces and a total sparsity order of nine for the geometric model (three per-tap paths times three non-zero taps) and 24 for the 3GPP model (approximation order eight times three non-zero taps). The generalized hard thresholding pursuit uses ten iterations and a unit step size. For the projected gradient algorithm, we use $\mu = 0.25$ (otherwise, the algorithm is unstable) and 20 iterations. The algorithms marked with † use the pseudo inverses of A_L and A_R instead of the hermitian transposes.

Figure 4 shows the mean squared estimation error $\|H - \hat{H}\|_F^2$ for different signal-to-noise ratios (SNRs) where H is normalized such that $\mathbb{E}\|H\|_F^2 \approx 1$ and SNR $= 1/\mathbb{E}\|E\|_F^2$ (Frobenius norms). The results are averaged over ten realizations for the sensing matrices A_L and A_R and 50 channel realizations for each pair of sensing matrices. For these examples, the reduced-complexity OMP algorithm showed a very poor performance and we did not attempt to implement the OMP algorithm (Alg. 3).

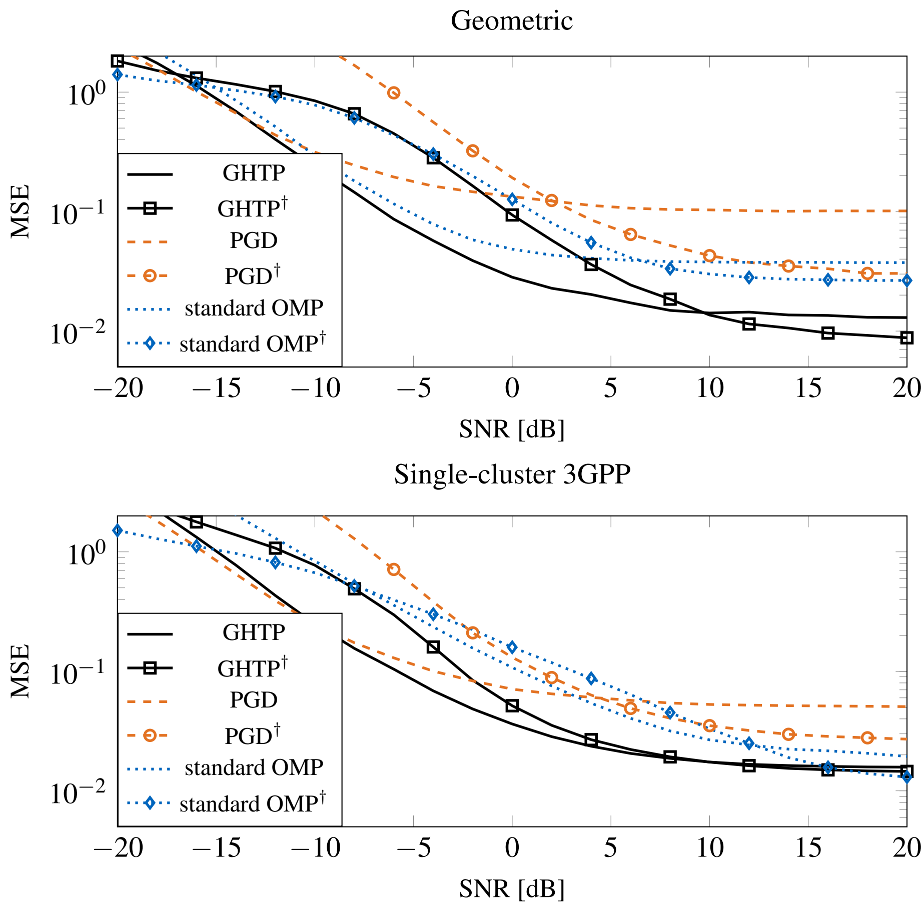

Figure 4: MSE for the geometric channel model (above) and the single-cluster 3GPP channel model (below) for different SNRs of the Root-MUSIC generalized HTP and PGD algorithms compared to the OMP algorithm with a four times oversampled DFT dictionary.

V
APPROXIMATE PROJECTORS

Let $\mathcal{U} = \cup_{t \in T} S_t$ be a *closed* union of subspaces in a Hilbert space $\mathcal{H}$. It is not hard to verify that the best-approximation operator

$$\mathcal{P}_{\mathcal{U}}(x) = \arg\min_{z \in \mathcal{U}} \|x - z\|^2 \tag{57}$$

is given by

$$\mathcal{P}_{\mathcal{U}}(x) = P_t x, \quad t = \arg\min_{t \in T} \|x - P_t x\|^2 \tag{58}$$

where P_t is the orthogonal projector onto S_t. The main difficulty with calculating $\mathcal{P}_{\mathcal{U}}(x)$ is to find the parameter t of the subspace S_t that best describes x. Even if the correct local minimum of this nonlinear optimization problem is found, there is an ultimate limit on the numerical precision by which t can be known. If, instead of t, we use an approximate value t_ε with $|t - t_\varepsilon| \le \varepsilon$, we obtain the map

$$\mathcal{Q}_{\mathcal{U}}(x) = P_{t_\varepsilon} x \tag{59}$$

which differs from $\mathcal{P}_{\mathcal{U}}$. This is the motivation for the notion of *approximate projectors*, which we introduce in Section 1. In Section 2, we show how to obtain approximate projectors in compound unions of subspaces of the form (MA-UOS) from approximate per-block projectors.

1—Notions of approximate projectors

Let us call a function $\mathcal{P}\colon \mathcal{H} \to \mathcal{U}$, which maps x onto its best approximation in $\mathcal{U}$, an *optimal projector*. For all but the most simple constraint sets $\mathcal{U}$, the calculation of the optimal projector $\mathcal{P}(x)$ is hard. In fact, if the set $\mathcal{U}$ is not closed, an optimal projector does not even exist. Thus, in view of an efficient and practical implementation of any algorithm that relies on computing $\mathcal{P}$, it is of great interest to relax the conditions on $\mathcal{P}$ such that $\mathcal{P}(x)$ can be computed in less time.
A function $\mathcal{P}'\colon \mathcal{H} \to \mathcal{U}$ that only satisfies $\mathcal{P}'(x) = x$ for $x \in \mathcal{U}$ and $\mathcal{P}'(x) \in \mathcal{U}$ for any $x \in \mathcal{H}$ is called a *projector*; the element $\mathcal{P}'(x)$ is not required to be the best approximation of x in $\mathcal{U}$. This notion of projector has been used in the context of the PGD algorithm in [33] along with the requirement $\|x - \mathcal{P}'(x)\| \le \|x - x_{\mathcal{U}}\| + \varepsilon$ for all $x_{\mathcal{U}} \in \mathcal{U}$ and $x \in \mathcal{H}$ and a small constant $\varepsilon > 0$[1]. As pointed out in [32],

[1] In this chapter, we often use a requirement such as $\|x - \mathcal{P}'(x)\| \le \|x - x_{\mathcal{U}}\| + \varepsilon$ for all $x_{\mathcal{U}} \in \mathcal{U}$, which should be thought of as $\|x - \mathcal{P}'(x)\| \le \|x - \mathcal{P}(x)\| + \varepsilon$. The reason for using the more complicated formulation with the "$\forall x \in \mathcal{U}$"-part is that we do not need to state that $\mathcal{U}$ is a closed set every time we use $\mathcal{P}(x)$ (this map is only well-defined if $\mathcal{U}$ is closed).

this relaxation guarantees that the PGD algorithm can be implemented, because such a map $\mathcal{P}'$ always exists. In [32, 43, 44] this absolute error requirement was replaced by the *relative error* requirement $\|x - \mathcal{P}'(x)\| \leq (1+\varepsilon)\|x - x_{\mathcal{U}}\|$ for all $x_{\mathcal{U}} \in \mathcal{U}, x \in \mathcal{H}$. Allowing for such a relative error opens up the possibility to use very efficient algorithms for computing $\mathcal{P}'(x)$ for certain constraint sets $\mathcal{U}$, e.g., when $\mathcal{U}$ is a finite union of sparse vectors with graph-structured sparsity [45]. In this work, we use *approximate projectors* satisfying[2]

$$\|x - \mathcal{Q}(x)\|^2 \leq \|x - P_t x\|^2 + \varepsilon^2 \|P_t x\|^2 \tag{60}$$

for all $t \in T$, where P_t is the orthogonal projector onto S_t. This is particularly useful in the infinite-union-of-subspaces model. In such a setting, it is natural to implement $\mathcal{Q}$ by first finding an (almost) optimal subspace S_t and then calculating the orthogonal projector P_t onto S_t. In contrast to the works [32,33,43,44], $\mathcal{Q}(x) = x$ for $x \in \mathcal{U}$ is not needed. Although this seems a natural requirement, it demands that the first approximation problem – finding the optimal parameter t – be solved exactly whenever $x \in \mathcal{U}$. By allowing for an inexact estimation of the parameter t, we open up the possibility to use a wider variety of algorithms for subspace estimation.

Definition 2 (ε-approximate projector)**.** *A map $\mathcal{Q}\colon \mathcal{H} \to \mathcal{U}$ is an ε-approximate projector onto $\mathcal{U}$ if for all $x \in \mathcal{H}$ and $t \in T$, there is $s \in T$ such that*

$$\|x - \mathcal{Q}(x)\|^2 \leq \|x - P_t x\|^2 + \varepsilon^2 \|P_{s,t} x\|^2 \tag{61}$$

where $P_{s,t}$ denotes the orthogonal projector onto $\operatorname{span}\{S_t, S_s\}$.

This condition is only meaningful for $\varepsilon < 1$. The condition (61) is slightly weaker than the simpler condition (60). The additional parameter s allows for some more flexibility when designing the approximate projector. If $\mathcal{Q}(x)$ is implemented as the orthogonal projection $\mathcal{Q}(x) = P_{t_q(x)}x$, where $t_q(x)$ depends on x, then (60) is equivalent to

$$\|P_{t_q(x)}x\|^2 \geq (1 - \varepsilon^2) \sup_{t \in T} \|P_t x\|^2 \tag{62}$$

which can be seen by adding $\|P_t x\|^2 + \|P_{t_q(x)}x\|^2$ to (60). Thus, an approximate solution in the sense (62) of the optimization problem $\sup_{t\in T} \|P_t x\|^2$ yields an ε-approximate projector.

In the context of model-based compressive sensing and, more recently, general unions of subspaces, convergence of the projected gradient algorithm was shown with non-optimal projectors $\mathcal{Q}$ that satisfy

$$\|x - \mathcal{Q}(x)\| \leq c_T \|x - P_t x\| \quad \forall t \in T \tag{63}$$

$$\|\mathcal{Q}(x)\| \geq c_H \|P_t x\| \quad \forall t \in T \tag{64}$$

[2] In [32, 43, 44], what we call *projectors* are called *approximate projectors* and what we call *optimal projectors* are called *projectors*.

for some $c_T \geq 1$ and $0 < c_H \leq 1$ (if the same map $\mathcal{Q}$ is used for both, the *head* and *tail* approximations) [43, 44]. Condition (64) is the same as (62), thus, in the terminology of [43,44], an ε-approximate projector is a head approximation oracle. In a different line of works [32], compressive sensing algorithms were used with projectors that satisfy

$$\|x - \mathcal{Q}(x)\| \leq (1 + \varepsilon_1)\|x - P_t x\| \quad \forall t \in T \tag{65}$$

$$\|x - \mathcal{Q}(x)\| \leq \|x - P_t x\| + \varepsilon_2 \|P_t x\| \quad \forall t \in T\,. \tag{66}$$

The second condition is similar to (60) and the first condition, again, requires $\mathcal{Q}$ to be a projector.
The authors of [46] and of [47] consider maps that are not required to be projectors. The maps used in [46] are approximate projectors in the sense of Definition 2, but are required to be of a very specific form that is only compatible with certain group sparsity models. The maps proposed in [47] are also not required to be projectors onto the constraint set $\mathcal{U}$ and the results apply to constraint sets that are more general than unions of subspaces. In the union of subspaces setting, their map $\mathcal{Q}$ has to be of the form $\mathcal{Q} = q \circ p$ where p is a linear map. For the purpose of comparison, we can choose p as the identity map. In that case, the map $\mathcal{Q}$ needs to fulfill

$$\|\mathcal{P}(x) - \mathcal{Q}(x)\| \leq \varepsilon' \|x\| \tag{67}$$

which is alternative to (60). Furthermore, the map $\mathcal{Q}$ is required to be an optimal projector onto a set $\mathcal{U}'$, which may be different from $\mathcal{U}$. The definition of an ε-approximate projector we use does not impose such a structural constraint.

2—Approximate projectors in compound models

In Chapter III, we encountered the union of subspaces

$$\mathcal{M} = \bigcup_{i_1,\dots,i_k \in \{1,\dots,Q\}, i_j \neq i_\ell \text{ if } j \neq \ell} \mathcal{U}_{i_1} \oplus \cdots \oplus \mathcal{U}_{i_k} \tag{68}$$

where $\mathcal{U}_\ell$ are (per-block) unions of subspaces with span$\{\mathcal{U}_\ell\} \perp$ span$\{\mathcal{U}_j\}$ for $\ell \neq j$. We can write $\mathcal{M}$ as the *intersection* of two unions of subspaces as follows: Let $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_Q$ with $\mathcal{H}_\ell =$ span$\{\mathcal{U}_\ell\}$, $\ell = 1, \dots, Q$ and let P_ℓ denote the orthogonal projector onto $\mathcal{H}_\ell$. Define

$$\mathcal{B}_k = \{x \in \mathcal{H} : |\,\text{bsupp}(x)| \leq k\} \tag{69}$$

where

$$\text{bsupp}(x) = \{\ell \in \{1, \dots, Q\} : P_\ell x \neq 0\} \tag{70}$$

denotes the *block support* of x – the indices corresponding to subspaces $\mathcal{H}_\ell$ in which x is not zero. The constraint $x \in \mathcal{B}_k$ specifies that x should have "energy" in no more than k of the subspaces $\mathcal{H}_\ell$. Also define

$$\mathcal{U} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_Q. \tag{71}$$

The constraint $x \in \mathcal{U}$ specifies that each "block" $P_\ell x$ of x should lie in the (per-block) union of subspaces $\mathcal{U}_\ell$. Accordingly, $\mathcal{U}$ is given as the intersection

$$\mathcal{M} = \mathcal{B}_k \cap \mathcal{U}. \tag{72}$$

The following result shows that an approximate projector onto $\mathcal{M}$ is obtained by concatenating per-block approximate projectors with the block-thresholding operation.

Lemma 3 (Approximate projectors)**.** *With the notation from above, let $\mathcal{P}\colon \mathcal{H} \to \mathcal{B}_k$ denote the block-thresholding operator (the optimal projector onto $\mathcal{B}_k$) and let $\mathcal{Q}_\ell\colon \mathcal{H}_\ell \to \mathcal{U}_\ell$ be ε-approximate projectors onto $\mathcal{U}_\ell$. Set $\mathcal{Q}\colon \mathcal{H} \to \mathcal{U}, x \mapsto \sum_{\ell=1}^{Q} \mathcal{Q}_\ell(P_\ell x)$. Then $\mathcal{P} \circ \mathcal{Q}$ is an ε-approximate projector onto $\mathcal{M}$.*

Proof. We need to show that for each subspace $S \subset \mathcal{B}_k \cap \mathcal{U}$ there is a subspace $S' \subset \mathcal{B}_k \cap \mathcal{U}$ such that

$$\|x - \mathcal{P} \circ \mathcal{Q}(x)\|^2 \leq \|x - P_S x\|^2 + \varepsilon^2 \|P_{S,S'} x\|^2 \tag{73}$$

where $P_{S,S'}$ denotes the orthogonal projector onto $\operatorname{span}\{S, S'\}$. Any $S \subset \mathcal{B}_k \cap \mathcal{U}$ is of the form $S = \oplus_{\ell \in I} S_\ell$ with $I \subset \{1, \ldots, Q\}, |I| = k$ and $S_\ell \subset \mathcal{U}_\ell$. By the sub-optimality of $\mathcal{Q}_\ell$, for each ℓ, there is another subspace $S'_\ell \subset \mathcal{U}_\ell$ with

$$\|P_\ell x - \mathcal{Q}_\ell(P_\ell x)\|^2 \leq \|P_\ell x - P_{S_\ell} x\|^2 + \varepsilon^2 \|P_{S_\ell, S'_\ell} x\|^2. \tag{74}$$

We set $S' = \oplus_{\ell \in I} S'_\ell$. Next, as the block-thresholding operator $\mathcal{P}$ satisfies

$$\|x - \mathcal{P}(x)\|^2 \leq \sum_{\ell \notin I} \|P_\ell x\|^2 = \left\|x - \textstyle\sum_{\ell \in I} P_\ell x\right\|^2 \tag{75}$$

we use $Q_\ell(x) \in \mathcal{H}_\ell$ so that (73) is satisfied because of

$$\begin{aligned}
\|x - \mathcal{P} \circ \mathcal{Q}(x)\|^2 &= \left\|x - \mathcal{P}\left(\textstyle\sum_{\ell=1}^{Q} \mathcal{Q}_\ell(P_\ell x)\right)\right\|^2 && (76)\\
&\leq \left\|x - \textstyle\sum_{\ell \in I} \mathcal{Q}_\ell(P_\ell x)\right\|^2 && (77)\\
&= \sum_{\ell \notin I} \|P_\ell x\|^2 + \sum_{\ell \in I} \|P_\ell x - \mathcal{Q}_\ell(P_\ell x)\|^2 && (78)\\
&\leq \sum_{\ell \notin I} \|P_\ell x\|^2 + \sum_{\ell \in I} \|P_\ell x - P_{S_\ell} x\|^2 + \varepsilon^2 \|P_{S_\ell, S'_\ell} x\|^2 && (79)\\
&= \|x - P_S x\|^2 + \varepsilon^2 \|P_{S,S'} x\|^2. && (80)
\end{aligned}$$

□

Example 1 (Channel estimation)**.** Let $\mathcal{M} = \mathcal{B}_k \cap \mathcal{U}$ with $\mathcal{U}$ given by (71) and each $\mathcal{U}_\ell$ corresponding to the DOA manifold (23). This union of subspaces appears in the channel estimation problem. An optimal projector onto $\mathcal{M}$ is obtained by first optimally projecting each block $P_\ell x$ onto $\mathcal{U}_\ell$ and then applying a block-thresholding operation (zeroing all except the k blocks for which $\|P_\ell \mathcal{P}_{\mathcal{U}_\ell}(x)\|$ is largest. While the block-thresholding is a simple operation that consists only of calculating and sorting the block norms, the per-block projections require solving DOA problems. In a practical implementation of an algorithm, the per-block approximate projector can be implemented, for example, by variants of the MUSIC or ESPRIT algorithms [48, 49]. The resulting algorithms are combinations of compressive sensing algorithms (block thresholding) with arbitrary other algorithms that calculate the per-block projections (cf. *model-aware* compressive sensing [3, 50]). The "coherence limit" encountered when using standard compressive sensing methods with an oversampled DFT dictionary (see, e.g., [51]) is circumvented by the use of arbitrary algorithms for the per-block projections that do not necessarily suffer (as much) from too closely spaced angles (this is corroborated by the simulation results shown in Chapter IV where we used the Root-MUSIC algorithm for the per-block projections). □

VI
RECOVERY GUARANTEES

Having introduced compressive sensing terminology and the notion of approximate projectors, we can proceed to analyzing the convergence properties of the algorithms presented in Chapter IV. While the proofs are not particularly interesting and at times lengthy (and therefore shown in the appendix), they build on the restricted isometry calculus we present in Sect. 1 and which is interesting in its own right.

1—Restricted isometry calculus

Let $A\colon \mathcal{H} \to \mathcal{H}'$ be a continuous linear operator between general Hilbert spaces $\mathcal{H}, \mathcal{H}'$. For the following statements, we assume that A has the RIP with RIC $\delta = \delta(A, \mathcal{U}) < 1$ with respect to a union of subspaces $\mathcal{U} \subset H$.

Lemma 4. *For any subspace $S \subset \mathcal{U}$, we have $\|P_S - P_S A^* A P_S\| \leq \delta$.*

Proof. Because $P_S x \in \mathcal{U}$ for all $x \in \mathcal{H}$, it follows from

$$(1-\delta)\|P_S x\|^2 \leq \|AP_S x\|^2 \leq (1+\delta)\|P_S x\|^2 \tag{81}$$

that for all $x \in \mathcal{H}$, we have

$$|\langle (P_S - P_S A^* A P_S)x, x\rangle| = |\, \|P_S x\|^2 - \|AP_S x\|^2 \,| \leq \delta \|P_S x\|^2 \leq \delta \|x\|^2 \,. \tag{82}$$

The operator $P_S - P_S A^* A P_S$ is self-adjoint and, thus, its operator norm can be calculated as

$$\|P_S - P_S A^* A P_S\| = \sup_{x \in \mathcal{H},\, \|x\|=1} |\langle (P_S - P_S A^* A P_S)x, x\rangle| \leq \delta \tag{83}$$

which shows the assertion. □

Lemma 5. *For any two subspaces $S, T \subset \mathcal{U}$ with* span$\{S, T\} \subset \mathcal{U}$ *and $S \perp T$, we have*

$$\|P_S(I - A^* A)P_T\| = \|P_S A^* A P_T\| \leq \delta. \tag{84}$$

Proof. Let $W = \text{span}\{S, T\} \subset \mathcal{U}$ so that $P_S = P_W P_S$ and $P_T = P_W P_T$. Because $W \subset \mathcal{U}$ and $\langle P_S x, P_T y\rangle = 0$ for $x, y \in \mathcal{H}$, we obtain

$$\|P_S A^* A P_T\| = \sup_{\|x\|=\|y\|=1} |\langle P_S x, P_S A^* A P_T y\rangle| \tag{85}$$

$$= \sup_{\|x\|=\|y\|=1} |\langle P_S x, P_S A^* A P_T y\rangle - \langle P_S x, P_T y\rangle| \tag{86}$$

$$= \sup_{\|x\|=\|y\|=1} |\langle P_S x, P_W A^* A P_W P_T y\rangle - \langle P_S x, P_W P_T y\rangle| \tag{87}$$

$$= \sup_{\|x\|=\|y\|=1} |\langle P_S x, P_W (I - A^* A) P_W P_T y\rangle| \leq \delta \tag{88}$$

where the inequality follows from Lemma 4. □

Lemma 6. *For any subspace $S \subset \mathcal{U}$, we have $\|P_S A^*\| \leq \sqrt{1+\delta}$.*

Proof. Using $\|x\| = \sup_{v\in\mathcal{H}:\|v\|=1} |\langle v, x\rangle|$ and noting that $P_S v \in \mathcal{U}$, we obtain

$$\|P_S A^* e\| = \sup_{v\in\mathcal{H}:\|v\|=1} |\langle v, P_S A^* e\rangle| \tag{89}$$

$$= \sup_{v\in\mathcal{H}:\|v\|=1} |\langle A P_S v, e\rangle| \tag{90}$$

$$\leq \sup_{v\in\mathcal{H}:\|v\|=1} \|A P_S v\| \|e\| \leq \sqrt{1+\delta}\|e\| \tag{91}$$

which shows that $\|P_S A^*\| \leq \sqrt{1+\delta}$. □

2—Projected gradient descent

The following theorem shows convergence of the inexact projected gradient algorithm when used with ε-approximate projectors. We do not require $x \in \mathcal{U}$ but rather allow for a modeling error $x - x_*$ where $x_* \in \mathcal{U}$ can be chosen as $x_* = \mathcal{P}(x)$ in case the projection exists. Let $\mathcal{U}^q := \{\sum_{i=1}^q x_i : x_i \in \mathcal{U}\}$ and let $\delta_q = \delta(A, \mathcal{U}^q)$ denote the restricted isometry constants with respect to the qth order union of subspaces $\mathcal{U}^q$.

Theorem 7. *Let $y = Ax + e$ with $x \in \mathcal{H}$, a bounded linear operator $A\colon \mathcal{H} \to \mathcal{H}'$ that has the RIP with respect to $\mathcal{U}$, and a disturbance $e \in \mathcal{H}'$. Let $\mathcal{Q}\colon \mathcal{H} \to \mathcal{U}$ be an ε-approximate projector for $\mathcal{U}$ with $\varepsilon \leq 1$ and define the constants*

$$c_1 = (2 + c\varepsilon)\,\delta_3,\ c_2 = (2 + c\varepsilon)\sqrt{1+\delta_2},\ c = \frac{1}{1+\sqrt{2}}. \tag{92}$$

If the sequence x_n is generated according to the inexact projected gradient algorithm (Algorithm 1),

$$z_n = x_{n-1} + A^*(y - A x_{n-1}) \tag{93}$$

$$x_n = \mathcal{Q}(z_n) \tag{94}$$

with $x_0 = 0$*, then for each* $x_* \in \mathcal{U}$*, we have*

$$\|x_n - x\| \le \|x - x_*\| + c_1^n\|x_*\| + \frac{1-c_1^n}{1-c_1}\Big(c_2\big(\|e\| + \|A(x-x_*)\|\big) + \varepsilon\|x_*\|\Big)\,. \tag{95}$$

If $c_1 < 1$, which translates into the condition $\delta < \big(2+\varepsilon/(1+\sqrt{2})\big)^{-1}$, the sequence of residuals is bounded. If $\varepsilon = 0$, i.e., if the projection is optimal, we recover the condition $\delta < 1/2$. Finally, if, in addition, $\|e\| = 0$ and $x_* = x$, the sequence x_n converges towards x.

The full proof, which we show in Appendix 2, is based on the following argumentation: If the projection step was not necessary and if $e = 0$, i.e., if $x_{n+1} = x_n + A^*(y - Ax_n)$, we would obtain the discrete dynamical system

$$x_{n+1} - x_* = (I - A^*A)(x_n - x_*) \tag{96}$$

after subtracting x_* from both sides and inserting $y = Ax_*$. The estimation error $x_{n+1} - x_*$ converges to zero if all eigenvalues of $I - A^*A$ are inside the unit circle, i.e., if the eigenvalues of (the non-negative matrix) A^*A are larger than zero and smaller than two. As $I - A^*A$ is only applied to vectors $x_n - x_*$, where both x_* and x_n are elements of $\mathcal{U}$, it is enough to verify this eigenvalue property for such vectors, hence, the *restricted* isometry (eigenvalue) requirement. Finally, the projection is incorporated through the triangle inequality $\|x_n - x_*\| \le \|x_n - z_n\| + \|z_n - x_*\| \le 2\|z_n - x_*\|$ where the last step follows from the best-approximation property ($\|x_n - z_n\|$ is minimal, therefore smaller than $\|x_* - z_n\|$). This inequality is responsible for the factor two in front of $I - A^*A$ and the requirement that the (restricted) eigenvalues of A^*A are between $1/2$ and $3/2$.

The result is non-trivial as soon as $\varepsilon < 1 - c_1 \approx 1 - 2\delta_3$. Furthermore, if $x = x_*$ for some $x_* \in \mathcal{U}$, we obtain

$$\|x_n - x\| \le c_1^n\|x\| + \frac{1-c_1^n}{1-c_1}\big(c_2\|e\| + \varepsilon\|x\|\big)\,. \tag{97}$$

However, even if $x \in \mathcal{U}$, a tighter bound may be achieved by choosing $x_* = \alpha x$ for some $\alpha < 1$, especially if one intends an early termination of the algorithm. If we use an optimal projector ($\varepsilon = 0$), Theorem 7 is comparable to the following symmetric reformulation of [33, Theorem 2].

Theorem 8 (Theorem 2 in [33])**.** *If* $\delta_2 < 1/5$ *and if* $y = Ax + e$*, the projected gradient algorithm with step size* $\mu = 5/6$*, which is given by*

$$z_{n+1} = x_n + \mu A^*(y - Ax_n) \tag{98}$$

$$x_{n+1} = \mathcal{P}_{\mathcal{U}}(z_{n+1}) \tag{99}$$

yields a sequence x_n *with*

$$\|x_n - x\|^2 \le c_1^n\|x\|^2 + \frac{1-c_1^n}{1-c_1}c_2\|e\|^2 \tag{100}$$

where

$$c_1 = \frac{2(1+5\delta_2)}{5(1-\delta_2)} \quad \textit{and} \quad c_2 = \frac{4}{1-\delta_2}. \tag{101}$$

Proof. This follows from Theorem 2 in [33] if we set $\beta = 1+\delta_2$ and $\alpha = 1-\delta_2$. If $\delta_2 < 1/5$, then $\mu = 5/6$ is a valid step size, because $\beta < 6/5 = \mu^{-1} < \frac{3(1-\delta_2)}{2} = \frac{3\alpha}{2}$ as $3(1-\delta_2)/2 > 6/5$. We assume that $\mathcal{U}$ is a closed union of subspaces and that we can compute exact projections, i.e., $\varepsilon = 0$. Using recursion in [33, Eq. (20)], we obtain (with $x_{\mathcal{A}} = x$ and $e_{\mathcal{A}} = e$ and $c_2 = 4/\alpha$ and $c_1 = 2(\mu\alpha)^{-1} - 2$)

$$\|x - x_n\|^2 \le \left(2\left(\frac{1}{\mu\alpha} - 1\right)\right)^n \|x\|^2 + \frac{c_2(1-c_1^n)}{1-c_1}\|e\|^2. \tag{102}$$

The constant c_1 is given as

$$c_1 = 2\left(\frac{1}{\mu\alpha} - 1\right) = 2\left(\frac{6}{5(1-\delta_2)} - 1\right) = 2\left(\frac{6-5+5\delta_2}{5(1-\delta_2)}\right) = \frac{2(1+5\delta_2)}{5(1-\delta_2)}. \tag{103}$$

□

Finally, the AM-IHT algorithm proposed in [43] converges if

$$(1+c_T)\Big(\delta_3 + \sqrt{1-(c_H(1-\delta_3)-\delta_3)^2}\Big) < 1 \tag{104}$$

and where c_H and c_T are the approximation constants of the head and tail approximation oracles, respectively, as defined in (63) and (64) in Chapter V. As $c_T \ge 1$ and because the square-root term is positive, it is necessary that $\delta_3 < 1/2$. Because $(c_H(1-\delta_3)-\delta_3)^2 \le c_H^2(1-\delta_3)^2$ for $c_H \ge 1/2$ and $\delta_3 \le 1/2$, we can recover the condition $\varepsilon < 1 - 2\delta_3$, which mirrors the condition on ε in Theorem 7. Thus, in both cases the range of admissible values for ε is comparable. In contrast to the result in Theorem 7, the fixed point of the AM-IHT algorithm does not exhibit the remaining approximation error $(1-c_1)^{-1}\varepsilon\|x\|$ thanks to the projection condition (63).

3—Generalized hard thresholding pursuit

For the generalized hard thresholding algorithm, we can show the following result. As in the previous section, we define $\mathcal{U}^q := \{\sum_{i=1}^q x_i : x_i \in \mathcal{U}\}$ and let $\delta_q = \delta(A, \mathcal{U}^q)$ denote the restricted isometry constants with respect to the qth order union of subspaces $\mathcal{U}^q$. This theorem is a generalization of the result for finite unions of subspaces stated in [52].

Theorem 9. *Let $y = Ax + e$ with $x \in \mathcal{H}$, a bounded linear operator $A\colon \mathcal{H} \to \mathcal{H}'$ that has the RIP with respect to $\mathcal{U}$, and a disturbance $e \in \mathcal{H}'$. Let $\mathcal{Q}\colon \mathcal{H} \to \mathcal{U}$ be an ε-approximate projector for $\mathcal{U}$ with $\varepsilon \le 1$ and define the constants*

$$c_1 = \frac{(2+c\varepsilon)\delta_3}{1-c\delta_2},\ c_2 = \frac{(2+c\varepsilon+c)\sqrt{1+\delta_2}}{1-c\delta_2},\ c_3 = \frac{1}{1-c\delta_2},\ c = \frac{1}{1+\sqrt{2}}. \tag{105}$$

If $\delta_3 < 1/3$ and if the sequence x_n is generated according to the generalized hard thresholding pursuit (Algorithm 2) with $x_0 = 0$, then for each $x_* \in \mathcal{U}$

$$\|x_n - x\| \leq \|x - x_*\| + c_1^n \|x_*\| + \frac{1 - c_1^n}{1 - c_1} \left(c_2 \|e\| + c_2 \|A(x - x_*)\| + \varepsilon c_3 \|x_*\| \right).$$

The proof is shown in Appendix 3. We observe the similarity of the error after n iterations with that for the PGD algorithm in Theorem 7. One can see that the constants defining the convergence speed are strictly worse for the GHTP algorithm (due to the factor $1/(1 - c\delta_2)$). However, we saw in Chapter IV that this does not prevent the GHTP algorithm from outperforming the PGD algorithm.

4—Orthogonal matching pursuit

The simplest version of a union of subspaces of the form (51) is

$$\mathcal{U} = \bigcup_{i_1, \ldots, i_k \in \{1, \ldots, Q\}, i_j \neq i_\ell \text{ if } j \neq \ell} \mathcal{H}_{i_1} \oplus \cdots \oplus \mathcal{H}_{i_k} \tag{106}$$

where the components $\mathcal{H}_\ell$ are subspaces. Such an extension of the OMP algorithm has been presented in [29]: Let the vector $x \in \mathbb{C}^{QN}$ be given as

$$x = \text{vec}(x_1 \quad \ldots \quad x_Q), \quad x_\ell \in \mathbb{C}^N \text{ for } \ell = 1, \ldots, Q$$

and let bsupp(x) denote the block support of x, that is, the block indices of its nonzero components (blocks) x_ℓ. We define the set of block-sparse signals with sparsity order k as

$$\mathcal{B}_k := \{x \in \mathbb{C}^{QN} : |\,\text{bsupp}(x)| \leq k\}.$$

As for sparse signals, one can show that if $x \in \mathcal{B}_k$, then a block-version of the OMP algorithm (Alg. 3 with this particular union of subspaces) recovers x exactly if $\delta(A, \mathcal{B}_k) < (\sqrt{k} + 1)^{-1}$ (by using the proof technique from [20]). Because $\mathcal{B}_k \subset \Sigma_{kN}$, we have $\delta(A, \mathcal{B}_k) \leq \delta(A, \Sigma_{kN})$. Thus, the condition for the block-OMP algorithm is weaker than the condition $\delta(A, \Sigma_{kN}) < (\sqrt{kN} + 1)^{-1}$, which would apply if one tried to recover all entries of x individually [20].
The following theorem shows that this condition can be weakened further if more structure is known about the blocks x_ℓ and if this knowledge is exploited in the OMP algorithm.

Theorem 10 (Orthogonal matching pursuit)**.** *Assume that $y = Ax$ with $x \in \mathcal{B}_k \cap \mathcal{U}$ where $\mathcal{U}$ is given by $\mathcal{U} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_Q$ and where $\mathcal{U}_\ell$ are unions of subspaces with* span$\{\mathcal{U}_\ell\} \perp$ span$\{\mathcal{U}_j\}$ *for $j \neq \ell$. If*

$$\delta(A, \mathcal{B}_{k+1} \cap \mathcal{U}^2) < \frac{1}{\sqrt{6} + \sqrt{k-1}}$$

then the OMP algorithm (Algorithm 3) recovers x exactly after k iterations.

The proof, which we show in Appendix 4, is considerably more difficult for the case where the sets $\mathcal{U}_\ell$ are *not* subspaces. A, by now, classical argumentation in the proof is to only show that a correct index is chosen during the first iteration of the OMP algorithm and to guarantee that this index will not be chosen again in any of the following iterations. Assume that $\mathcal{U}_\ell = \mathcal{H}_\ell$ are subspaces and that $x_* \in \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_k$. It follows that

$$r_1 = y \in \text{span}\{\text{range}(AP_{\mathcal{H}_1}), \ldots, \text{range}(AP_{\mathcal{H}_k})\}. \tag{107}$$

The first index is chosen by correlating the residual r_1 with all of the subspaces $\mathcal{H}_1, \ldots, \mathcal{H}_Q$ and choosing the one that yields the maximal correlation. If a correct index is chosen in the first iteration, e.g., the first index, the residual r_2 at the second iteration of the OMP algorithm is given as $y - Ax_1$ where $x_1 \in \mathcal{H}_1$. Consequently, we still have $r_2 \in \text{span}\{\text{range}(AP_{\mathcal{H}_1}), \ldots, \text{range}(AP_{\mathcal{H}_k})\}$ and the condition about whether the OMP algorithm chooses another correct index $\ell \in \{1, \ldots, k\}$ is the same as in the first iteration. However, we have to ensure that the next index is not the same index as the one chosen in the previous iteration. This is ensured by the orthogonal projection step in the OMP algorithm. The correlation of r_n with any of the previously chosen subspaces $AP_{\mathcal{H}_1}, \ldots, AP_{\mathcal{H}_{n-1}}$ is zero and, therefore, not maximal. Thus, a correct index is found that has not been chosen before.
This step is problematic if the sets $\mathcal{U}_\ell$ are not subspaces but unions of subspaces. We have

$$r_1 = y \in \text{span}\{\text{range}(AP_{S_{*,1}}), \ldots, \text{range}(AP_{S_{*,k}})\} \tag{108}$$

at the first iteration (where $S_{*,1}$ denotes the correct subspace for the first index etc.). Then, even if the OMP algorithm chooses a correct index, say the first one, it does not necessarily choose the correct subspace $S_{*,1}$ but another subspaces $S_1 \neq S_{*,1}$ so that

$$r_2 \in \text{span}\{\text{range}(AP_{S_{*,1}}), \ldots, \text{range}(AP_{S_{*,k}}), \text{range}(AP_{S_1})\}, \tag{109}$$

i.e., the situation in the second iteration differs from that in the first iteration.

5—Reduced-complexity orthogonal matching pursuit

The following theorem shows under what conditions the reduced-complexity OMP algorithm successfully recovers original signal.

Theorem 11 (Reduced Complexity MA-OMP)**.** *Assume that $y = Ax$ with $x \in \mathcal{B}_k \cap \mathcal{U}$. Let $\delta = \delta(A, \mathcal{B}_{k+1} \cap \mathcal{U}^2)$ and $\delta' = \delta(A, \mathcal{B}_k)$. If*

$$\delta < \frac{1}{1 + u + \sqrt{(k-1)(1+u^2)}} \quad \textit{with } u = 2\sqrt{\frac{\delta + \delta'}{1 - \delta'}}$$

then the reduced-complexity OMP algorithm (Algorithm 4) recovers x exactly after k iterations. This condition is implied, for example, by the two conditions

$$\delta' \leq \frac{3}{4} \quad \textit{and} \quad \delta \leq \frac{1}{5 + 4\sqrt{k-1}}. \tag{110}$$

This condition is considerably more difficult to satisfy than the one stated in Theorem 10, but the algorithm is much easier to implement. The proof, which we show in Appendix 5, is similar to that of the OMP algorithm.

VII
Chaining in Geometrically Regular Spaces

Now that we have seen how useful the restricted isometry property is, we discuss how to obtain sensing matrices that have this property. There are several ways to proceed: First, one could guess a matrix and verify whether it has the RIP, but this has been shown to be quite difficult for non-trivial problems. More precisely, for sparse signals, the verification part of this problem is NP-hard [53,54]. Second, there are cases where a deterministic construction of matrices is possible, for example, by using equi-angular tight frames when $\mathcal{U}$ consists of sparse signals [55,56]. However, deterministic constructions encounter the *square-root bottleneck* (or almost, cf. [57]) and appear to be more difficult to use than random constructions, which is the third way to obtain RIP-matrices.
Let us assume that we have access to some random matrix generator for which the probabilistic point-wise RIP holds:

$$\mathbb{P}\left[(1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2\right] \geq 1 - 2f(\delta). \tag{P-RIP}$$

The probability is with respect to the distribution of the random matrix A and f is an upper bound for the probability that any random draw of the matrix A distorts a given vector x by more than δ. By far the most important example of such matrices are Gaussian random matrices $A \in \mathbb{R}^{m\times M}$ where each element of A is independently drawn from the normal distribution with variance $1/m$, $[A]_{ij} \sim \mathcal{N}(0, 1/m)$. As we discuss in the next chapter, for such matrices (P-RIP) holds with

$$-\log f(\delta) = m(\delta - \log(1+\delta))/2. \tag{111}$$

Importantly, the number of rows m of the matrix A appears as a linear factor in the exponent. This is a manifestation of the concentration of measure phenomenon: as we increase m, the probability that $\|Ax\| \approx \|x\|$ goes to one exponentially fast. This probability increases even fast enough that we can use the (crude) union bound to show a version of (P-RIP) which holds simultaneously for many vectors x. This is the main idea behind the probabilistic constructions of RIP matrices.
A second thing to notice is that, in this example, f is *log-concave* as a function of δ, i.e., $\log f$ is concave. This property is strongly linked to so-called *sub-exponential* random variables and we discuss this relation in the next chapter. Log-concavity of the function f is *not* a prerequisite to derive a uniform version of the probabilistic RIP, i.e., an inequality of the form

$$\mathbb{P}\left[\forall x \in \mathcal{U} \,:\, (1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2\right] \geq 1 - f_{\mathcal{U}}(\delta) \tag{112}$$

with some function $f_{\mathcal{U}}$ depending on f and $\mathcal{U}$. However, the stronger f decays as a function of δ – log-concavity means ultra-fast decay – the better the bound $f_{\mathcal{U}}$ becomes.

To show an inequality of the form (112), we first show that (P-RIP) holds for a representative (finite) sample $\mathcal{U}_\varepsilon \subset \mathcal{U}$ and then verify that nothing bad happens between sample points. The first part is readily shown by using a union-bound argument for an ε-cover of the intersection of $\mathcal{U}$ with the unit sphere. The second part is simple as well (in the light of the right argument) if $\mathcal{U}$ is given as a finite union of subspaces [25]. However, for general sets $\mathcal{U}$, this second step is considerably more difficult and usually relies on sequences of finite subsets of $\mathcal{U}$ – a process known as *chaining*. Thus, before we can show our major RIP-theorem (Theorem 28), we need to introduce the chaining technique.

In this chapter, we show how this technique can be used to prove a general concentration theorem (Theorem 20) for random processes in metric spaces with geometrically regular index sets. This theorem is much more general than the RIP Theorem, which we derive as a corollary of this theorem in Chapter IX.

We start this chapter with a discussion of ε-covers in metric spaces and the notion of geometric regularity, which describes the growth rate of the ε-covers as ε decreases (Sect. 1). Geometrically regular spaces are abundant and we give several examples in Sect. 2. In Sect. 3, we introduce the chaining technique and derive two tail bounds in very general settings. These bounds are almost trivial, but they contain the essence of the chaining technique. In Sect. 4, we present Theorem 20, a version of the general chaining theorem where we make a very severe restriction regarding the optimality of the bounds: we use ε-covers to construct chaining sequences. This is the only way that we are aware of by which the bounds that appear in the theorem can be evaluated. We briefly discuss a possible generalization of Thm. 20 to non-isotropic distributions in Sect. 5 and resolve the remaining degree of freedom in Theorem 20 – the approximation speed – in Sect. 6 by giving two possible approximation sequences. Finally, in Sect. 7, we discuss the relation of our results with so-called *generic chaining* theorems.

1—Geometric regularity

The chaining technique needs successive approximations of an uncountable index set by finite sets. In this work, we use sequences of ε-covers for these approximations and, hence, covering numbers to quantify the complexity of the index set.

Definition 12 (ε-sequences, ε-covers, ε-covering number)**.** *Let (T, d) be a metric space. A sequence of maps $(\pi_\varepsilon)_{\varepsilon>0}, \pi_\varepsilon \colon T \to T$, is called an* ε-sequence *of T, if $d(\pi_\varepsilon(t), t) \leq \varepsilon$ for all $t \in T$, $\varepsilon > 0$. The set $T_\varepsilon = \pi_\varepsilon(T)$ is called an* ε-cover *of T. A function $\overline{\pi} \colon \mathbb{R} \to \mathbb{R}$ with $|\pi_\varepsilon(T)| \leq \overline{\pi}(\varepsilon)$ for all $\varepsilon > 0$ is called an* upper bound *of π. Finally,*

$$N(T, d, \varepsilon) = \inf_{\pi\ \varepsilon\text{-sequence}} |\pi_\varepsilon(T)| \tag{113}$$

is called the ε-covering *number of* (T, d).

We will require that the covering numbers of T are finite. Such spaces are called *totally bounded*. We need some control about the growth rate of the covering numbers as ε tends to zero. A notion of growth rate that has been used in this context (cf. [58]) is the upper box-counting dimension, which is defined as

$$\dim_{\mathrm{u}}(T) = \limsup_{\varepsilon \to 0} \frac{\log(N(T, d, \varepsilon))}{\log(1/\varepsilon)}. \tag{114}$$

It is not difficult to see that sets with finite upper box-counting dimension are exactly those sets for which we can find a k-regular ε-sequence (a similar notion of regularity has also been used in [59] and [60, Definition 5.1]):

Definition 13 (k-regular ε-sequences, functions, and spaces)**.** *Let* $k > 0, \varepsilon' > 0$. *A function* $g\colon \mathbb{R} \to \mathbb{R}$ *is called* k-regular *if*

$$g(\varepsilon) \le (\varepsilon'/\varepsilon)^k g(\varepsilon') \ \ \forall \varepsilon > 0. \tag{115}$$

An ε*-sequence* π *of a metric space* (T, d) *is* k*-regular if there is a* k*-regular upper bound* $\overline{\pi}$ *of* π. *A metric space* (T, d) *is* k*-regular if there exists a* k*-regular* ε*-sequence of* T.

2—Examples of geometrically regular spaces

Euclidean balls are standard examples for geometrically regular spaces (cf., e.g., [23, App. C.2] and [61, Lemma 5.2]).

Example 2 (Euclidean balls and spheres are regular)**.** Let d denote the Euclidean distance in $\mathbb{R}^k$ or $\mathbb{C}^k$. The covering numbers of the metric spaces $(B_{\mathbb{R}^k}, d)$ and $(\partial B_{\mathbb{R}^k}, d)$ are bounded by $(2+\varepsilon)^k/\varepsilon^k$. Similarly, the covering numbers of $(B_{\mathbb{C}^k}, d)$ and $(\partial B_{\mathbb{C}^k}, d)$ are bounded by $g(\varepsilon) = (2 + \varepsilon)^{2k}/\varepsilon^{2k}$. Note that ε-covers are subsets of the respective metric spaces, i.e., elements of ε-covers of spheres have unit norm! □

Example 3 (Real sparse signals)**.** Let d denote the Euclidean distance and $\Sigma_k = \{x \in \mathbb{R}^M : |\operatorname{supp}(x)| \le k\}$ the set of k-sparse signals in $\mathbb{R}^M$. We obtain an ε-cover of $\mathcal{C} = \Sigma_k \cap \partial B_{\mathbb{R}^M}$ by taking the union of $\binom{M}{k} \le (eM/k)^k$ covers of k-dimensional unit spheres, each of which has cardinality less than $(2 + \varepsilon)^k/\varepsilon^k$. Consequently, there is an ε-sequence π of $\mathcal{C}$ with upper bound $\overline{\pi}(\varepsilon) = (eM/k)^k(2+\varepsilon)^k/\varepsilon^k$. Furthermore, by construction the set of differences $\{x - \pi_\varepsilon(x), x \in \Sigma_k\}$ also contains k-sparse vectors (we approximate k-sparse vectors by k-sparse vectors with the same support). □

Example 4 (Complex sparse signals)**.** If $\Sigma_k \subset \mathbb{C}^M$ consists of complex signals and if d is still the Euclidean distance, we can proceed as for real sparse signals except that each ε-cover of the k-dimensional complex unit ball requires $(2 + \varepsilon)^{2k}/\varepsilon^{2k}$

Geometrically regular space	bound of the covering number
Real sparse signals $\Sigma_k \subset \mathbb{R}^M$	$(eM/k)^k(2+\varepsilon)^k/\varepsilon^k$
Complex sparse signals $\Sigma_k \subset \mathbb{C}^M$	$(eM/k)^k(2+\varepsilon)^{2k}/\varepsilon^{2k}$
Finite union of real k-dim. subspaces	#subspaces $\times$ $(2+\varepsilon)^k/\varepsilon^k$
Finite union of complex k-dim. subspaces	#subspaces $\times$ $(2+\varepsilon)^{2k}/\varepsilon^{2k}$
Rank-k matrices in $\mathbb{R}^{n_1\times n_2}$	$(9/\varepsilon)^{k(n_1+n_2+1)}$
Lipschitz union of k-dim. real subspaces	$(16L)^k/\varepsilon^{2k}$
Lipschitz union of k-dim. complex subspaces	$(56L)^k/\varepsilon^{3k}$

Table VII.1: Covering number estimates.

points. Thus, there is an ε-sequence π of $\mathcal{C} = \Sigma_k \cap \partial B_{\mathbb{R}^M}$ with upper bound $\overline{\pi}(\varepsilon) = (eM/k)^k(2+\varepsilon)^{2k}/\varepsilon^{2k}$ and the set $\{x - \pi_\varepsilon(x), x \in \Sigma_k\}$ also contains k-sparse vectors (we approximate k-sparse vectors by k-sparse vectors with the same support). □

Example 5 (Finite unions of subspaces)**.** If $\mathcal{U} = \cup_{t\in T} S_t \subset \mathbb{R}^M$ (or $\mathbb{C}^M$) is a finite union of subspaces with dim$(S_t) = k$ for all t, we can proceed just as for sparse signals and construct separate ε-covers for each subspace. Consequently, there is an ε-sequence π of the set $\mathcal{C} = \mathcal{U} \cap \partial B_{\mathbb{R}^M}$ (or $\partial B_{\mathbb{C}^M}$) with upper bound $\overline{\pi}(\varepsilon) = |T|(2+\varepsilon)^{k'}/\varepsilon^{k'}$ with $k' = k$ for real subspaces and $k' = 2k$ for complex subspaces. Elements of the set $\{x - \pi(x) : x \in \mathcal{C}\}$ are also elements of $\mathcal{U}$. □

Example 6 (Low-rank matrices [62, Lemma 3.1])**.** The covering number of the set of unit-norm matrices (Frobenius norm and associated metric) in $\mathbb{R}^{n_1\times n_2}$ with rank at most k is bounded by $(9/\varepsilon)^{k(n_1+n_2+1)}$. □

Some more work is required to bound the covering numbers in infinite unions of subspaces.

Theorem 14 (Covering infinite unions of subspaces)**.** *Let T be a set and $\mathcal{U} = \cup_{t\in T} S_t$ a union of k-dimensional subspaces $S_t \subset \mathcal{H}$ of a Hilbert space $\mathcal{H}$ with orthogonal projectors P_t onto S_t. Assume that the metric space (T, d_F) is totally bounded with respect to the Finsler metric*

$$d_F(s,t) = \|P_s - P_t\| \tag{116}$$

which measures the distance of two subspaces as the operator-norm difference of their orthogonal projectors. If g_T is an upper bound for the covering numbers of T in this metric, then the covering number of $\mathcal{C} = \mathcal{U} \cap \partial B_{\mathcal{H}}$ with respect to the Hilbert space metric is bounded by

$$g(\varepsilon) = \inf_{\varepsilon_T,\varepsilon_{\mathcal{H}}>0} g_T(\varepsilon_T)(2+\varepsilon_{\mathcal{H}})^{k'}/\varepsilon_{\mathcal{H}}^{k'} \tag{117}$$

$$\text{s.t. } \varepsilon_{\mathcal{H}}^2 + \varepsilon_T^2 + (1-(1-\varepsilon_T^2)^{1/2})^2 \le \varepsilon^2. \tag{118}$$

where $k' = k$ if $\mathcal{H}$ is a real Hilbert space and $k' = 2k$ if $\mathcal{H}$ is complex.

Proof. We construct an ε-cover of $\mathcal{C}$ as follows: We first choose an ε_T-cover T_f of T of cardinality $g_T(\varepsilon_T)$ and then set

$$\mathcal{C}_\mathrm{f} = \partial B_\mathcal{H} \cap \bigcup_{t \in T_\mathrm{f}} S_t. \tag{119}$$

As T_f is finite, there is an $\varepsilon_\mathcal{H}$-cover W_f of $\mathcal{C}_\mathrm{f}$ with (cf. Example 5)

$$|W_\mathrm{f}| \leq g_T(\varepsilon_T)(2 + \varepsilon_\mathcal{H})^{k'} / \varepsilon_\mathcal{H}^{k'}. \tag{120}$$

Let us see how well we can approximate elements in $\mathcal{C}$ by elements in W_f. For any $x \in \mathcal{C}$, there is $t_0 \in T$ such that $x = P_{t_0} x$ and we can find $t \in T_\mathrm{f}$ such that

$$\|x - P_t x\| = \|(P_{t_0} - P_t)x\| \leq \|P_{t_0} - P_t\| \|x\| = d_\mathrm{F}(t_0, t) \leq \varepsilon_T. \tag{121}$$

Let $z \in W_\mathrm{f}$, $z \in S_t$, be the best approximating point of $P_t x$. By orthogonality, we obtain

$$\|x - z\|^2 \leq \varepsilon_T^2 + \|z - P_t x\|^2. \tag{122}$$

As $P_t x$ is not necessarily on the unit sphere, we do *not* simply have $\|z - P_t x\| \leq \varepsilon_\mathcal{H}$. However, if we let $x_t = P_t x$, we find that as $\|x - x_t\| \leq \varepsilon_T$ and $\|x\| = 1$, we must have

$$\|x_t\|^2 = \|x\|^2 - \|x - P_t x\|^2 \geq 1 - \varepsilon_T^2 \tag{123}$$

so that x_t is inside the unit ball and very close to the unit sphere: Let $y = x_t / \|x_t\| \in \mathcal{C}_\mathrm{f}$, then

$$\|x_t / \|x_t\| - x_t\| = 1 - \|x_t\| \leq 1 - (1 - \varepsilon_T^2)^{1/2}. \tag{124}$$

Next, note that any element $z \in B_\mathcal{H}$ must satisfy

$$\mathrm{Re}\langle z - y, y \rangle \leq 0 \tag{125}$$

because $1 \geq \|z\|^2 \geq |\langle z, y \rangle|^2 = |\langle y + z - y, y \rangle|^2 \geq (1 + \mathrm{Re}\langle z - y, y \rangle)^2$. Consequently, if $z \in W_\mathrm{f} \subset B_\mathcal{H}$ approximates y with $\|z - y\| \leq \varepsilon_\mathcal{H}$, we obtain

$$\begin{aligned}
\|z - x_t\|^2 &= \|z - y\|^2 + \|y - x_t\|^2 + 2\,\mathrm{Re}\langle z - y, y - x_t \rangle && (126)\\
&= \|z - y\|^2 + \|y - x_t\|^2 + 2\,\mathrm{Re}\langle z - y, (1 - \|x_t\|)y \rangle && (127)\\
&\leq \|z - y\|^2 + \|y - x_t\|^2 && (128)\\
&\leq \varepsilon_\mathcal{H}^2 + \left(1 - (1 - \varepsilon_T^2)^{1/2}\right)^2 && (129)
\end{aligned}$$

as $y - x_t = (1 - \|x_t\|)y$ with $1 - \|x_t\| \geq 0$. In summary, $\|x - z\|^2 \leq \varepsilon_T^2 + \varepsilon_\mathcal{H}^2 + (1 - (1 - \varepsilon_T^2)^{1/2})^2$.

We can now use those ε_T and $\varepsilon_\mathcal{H}$ that yield an ε-cover for a given ε and minimize the cardinality of that cover. As we can always replace $g(\varepsilon)$ by the closest integer $n \leq g(\varepsilon)$, which is a right-continuous operation, we can even take the infimum. □

Corollary 15 (Lipschitz union of subspaces)**.** *Let $T \subset B_{\mathbb{R}^k}$ and $\mathcal{U} = \cup_{t\in T} S_t$ a union of k-dimensional subspaces $S_t \subset \mathcal{H}$ of a Hilbert space $\mathcal{H}$ as in the previous theorem. Assume that P_t is Lipschitz-continuous with constant $L \geq 1$:*

$$\|P_t - P_s\|^2 \leq L^2 \sum_{n=1}^{k} (t_n - s_n)^2. \tag{130}$$

Then, for $\varepsilon \leq 1$, the covering number of $\mathcal{C} = \mathcal{U} \cap \partial B_{\mathcal{H}}$ is bounded by

$$g(\varepsilon) = \begin{cases} (16L)^k/\varepsilon^{2k}, & \text{if } \mathcal{H} \text{ is a real Hilbert space,} \\ (56L)^k/\varepsilon^{3k}, & \text{if } \mathcal{H} \text{ is a complex Hilbert space.} \end{cases} \tag{131}$$

Proof. First, we note that it follows from $\|P_t - P_s\| \leq Ld(t,s)$ that any ε/L-cover of $B_{\mathbb{R}^k}$ with respect to the Euclidean distance is an ε-cover of $B_{\mathbb{R}^k}$ with respect to the Finsler distance. Consequently, we can find an ε_T-cover of $B_{\mathbb{R}^k}$ with respect to the Finsler metric of size less than $g(\varepsilon_T) = L^k(2+\varepsilon_T/L)^k/\varepsilon_T^k$.

1. Real case: Let $k' = k$ and select $\varepsilon_{\mathcal{H}} = 3\varepsilon/4$ and $\varepsilon_T = 3\varepsilon/5$. In this case, $\varepsilon_{\mathcal{H}}^2 + \varepsilon_T^2 + (1-(1-\varepsilon_T^2)^{1/2})^2 \leq \varepsilon^2$ and the size of the whole cover is bounded by $g(\varepsilon)$ with

$$\sqrt[k]{g(\varepsilon)} = \frac{L(2+\varepsilon_{\mathcal{H}})(2+\varepsilon_T/L)}{\varepsilon_{\mathcal{H}}\varepsilon_T} \leq \frac{16L}{\varepsilon^2}. \tag{132}$$

2. Complex case: Let $k' = 2k$ and select $\varepsilon_{\mathcal{H}} = 4\varepsilon/5$ and $\varepsilon_T = 4\varepsilon/7$. This choice also satisfies $\varepsilon_{\mathcal{H}}^2 + \varepsilon_T^2 + (1-(1-\varepsilon_T^2)^{1/2})^2 \leq \varepsilon^2$ and the size of the whole cover is bounded by $g(\varepsilon)$ with

$$\sqrt[k]{g(\varepsilon)} = \frac{L(2+\varepsilon_{\mathcal{H}})^2(2+\varepsilon_T/L)}{\varepsilon_{\mathcal{H}}^2\varepsilon_T} \leq \frac{56L}{\varepsilon^3}. \tag{133}$$

□

The results from this section are summarized in Table VII.1. Note that for such "Lipschitz unions of subspaces", the Lipschitz constant L essentially takes the role of the ambient dimension M.

3—Chaining in topological spaces

In this section, we consider collections of uncountably many random variables $X = (X_t)_{t\in T}$ with values in a metric space E and indices in a separable topological space T. We will later apply the results to the process $(Ax)_{x\in\mathcal{C}}$, i.e., the index set T will be identified with the intersection of a union of subspaces with the unit sphere and t with a unit norm vector x in $\mathcal{C}$.
We assume that T is a separable topological space. This allows us to reason about continuity of the process X and to approximate X_t by a sequence X_{t_n} with $t_n \to t$.

These two properties are essential for analyzing the path functionals $\sup_{t \in T} \|X_t\|$. The quantitative results derived in the subsequent sections then all apply to situations where T is a totally bounded metric space (e.g., sparse vectors with unit norm in $\mathbb{R}^M$).

To understand any of the results and proofs presented here, none of the advanced concepts regarding the theory of stochastic processes are required. However, a minor disclaimer is in order: Events of the form $\sup_{t \in T} \|X_t\| > u$ for some $u \in \mathbb{R}$ are not necessarily measurable. There are several ways around this problem. The first way, which is found in chaining-related literature, e.g., [60, 63], is to define

$$\tilde{\mathbb{P}}\left[\sup_{t \in T} \|X_t\| > u\right] := \sup_{T_f \subset T,\, T_f \text{ finite}} \mathbb{P}\left[\sup_{t \in T_f} \|X_t\| > u\right] \tag{134}$$

as the *lattice supremum*. Second, we could restrict our attention to random variables with almost surely continuous paths (which is something we require anyway). In this case, it is possible to show that events involving path functionals are measurable if conditioned on the event that X is continuous (see Appendix 6). Finally, we could use the canonical extension of the probability measure $\mathbb{P}$ to an *outer measure* $\tilde{\mathbb{P}}$ on all subsets of Ω (the probability space). That is, if the event $\sup_{t \in T} \|X_t\| > u$ is contained in some measurable event $B \subset \Omega$, then $\tilde{\mathbb{P}}[\sup_{t \in T} \|X_t\| > u] \leq \mathbb{P}[B]$. As the chaining technique is all about finding a countable sequence of events that contain $\{\sup_{t \in T} \|X_t\| > 0\}$, this last interpretation is a very natural one.

The main ingredient of the chaining technique is a method to provide an approximating sequence $s \in T^{\mathbb{N}_0}, s_n \in T, n \geq 0$, with $s_n \to t$ for any given point $t \in T$. We introduce the following definition:

Definition 16 (Chaining set)**.** *Let T be a separable topological space, let $\mathcal{A} \subset T^{\mathbb{N}_0}$ be a subset of all sequences in T, and let $\mathcal{A}(t) \subset \mathcal{A}$ be such that $t = \lim_{n \to \infty} s_n$ for each $s \in \mathcal{A}(t)$. We say that $\mathcal{A}$ is a chaining set if for each $t \in T$, $\mathcal{A}(t)$ is nonempty.*

Different versions of chaining differ in terms of how the set $\mathcal{A}$ is chosen.

In the following, we will always assume that T is a separable topological space, that $\mathcal{A} \subset T^{\mathbb{N}_0}$ is a chaining set, and that $X = (X_t)_{t \in T}$ is a random process with almost surely continuous paths. Furthermore, we assume that $u = (u_n)_{n \geq 1}, u_n \colon T \times T \to \mathbb{R}$ is a sequence of functions, which we call *deviation sequence*.

Definition 17 (β_u-functional)**.** *For a given deviation sequence u and chaining set $\mathcal{A}$ in T, we define*

$$\beta_u(T, \mathcal{A}) = \sup_{t \in T} \inf_{s \in \mathcal{A}(t)} \sum_{n \geq 1} u_n(s_n, s_{n-1}). \tag{135}$$

The first theorem shows how we can decompose the supremum event into a union of more structured events.

Theorem 18 (Tail bound via chaining). *For any given chaining set $\mathcal{A}$ and deviation sequence u, we have*

$$\mathbb{P}\left[\sup_{t\in T}\inf_{s\in\mathcal{A}(t)} d(X_t, X_{s_0}) > \beta_u(T,\mathcal{A})\right] \leq \sum_{n\geq 1}\mathbb{P}\left[\exists s\in\mathcal{A} : d(X_{s_n}, X_{s_{n-1}}) > u_n(s_n, s_{n-1})\right]. \quad (136)$$

Remark 1. The right-hand-side of (136) is easier to handle than the left-hand-side if, for each n, there are only countably or even finitely many possible combinations of s_n and s_{n-1}.

Proof. Define the events

$$\mathcal{X}_n = \left\{\exists s\in\mathcal{A} : d(X_{s_n}, X_{s_{n-1}}) > u_n(s_n, s_{n-1})\right\} \quad (137)$$

$$\mathcal{X} = \{X \text{ is not continuous}\} \cup \bigcup_{n\geq 1}\mathcal{X}_n. \quad (138)$$

A union bound yields $\mathbb{P}[\mathcal{X}] \leq \sum_{n\geq 1}\mathbb{P}[\mathcal{X}_n]$ for the probability of $\mathcal{X}$. We show that if $\omega\in\mathcal{X}^c$ and if we let $x(t) = X_t(\omega)$, we have

$$\sup_{t\in T}\inf_{s\in\mathcal{A}(t)} d(x(t), x(t_0)) \leq \beta_u(T,\mathcal{A}). \quad (139)$$

Let $s\in\mathcal{A}(t)$ and repeatedly use the triangle inequality and the continuity of x to get

$$d(x(t), x(s_0)) \leq \sum_{n\geq 1} d(x(s_n), x(s_{n-1})) \quad (140)$$

as $s_n \to t$. Because $\omega\in\mathcal{X}^c$, we have $d(x(s_n), x(s_{n-1})) \leq u_n(s_n, s_{n-1})$ so that

$$d(x(t), x(s_0)) \leq \sum_{n=1} u_n(s_n, s_{n-1}). \quad (141)$$

Finally, as t and the sequence $s\in\mathcal{A}(t)$ was arbitrary, we can take the infimum and then the supremum on both sides,

$$\sup_{t\in T}\inf_{s\in\mathcal{A}(t)} d(x(t), x(s_0)) \leq \sup_{t\in T}\inf_{s\in\mathcal{A}(t)}\sum_{n\geq 1} u_n(s_n, s_{n-1}) = \beta_u(T,\mathcal{A}). \quad (142)$$

□

Theorem 19 (Concentration via chaining). *Let $X = (X_t)_{t\in T}$ have values in a* normed *space. Let $p\geq 1$ and let $\delta_0\geq 0$. Let*

$$P_0 := \mathbb{P}\left[\exists s\in\mathcal{A} : |\|X_{s_0}\|^p - 1| > \delta_0\right], \quad (143)$$

$$P_n := \mathbb{P}\left[\exists s\in\mathcal{A} : \|X_{s_n} - X_{s_{n-1}}\| > u_n(s_n, s_{n-1})\right]. \quad (144)$$

Then

$$\mathbb{P}\left[\sup_{t\in T} |\|X_t\|^p - 1| > \delta\right] \le \sum_{n\ge 0} P_n. \tag{145}$$

with $\delta = \left(\sqrt[p]{1+\delta_0} + \beta_u(T,\mathcal{A})\right)^p - 1$.

Proof. Define the events

$$\mathcal{X}_{\text{sup}} = \left\{\sup_{t\in T}\inf_{s\in\mathcal{A}(t)} d(X_t, X_{s_0}) > \beta_u(T,\mathcal{A})\right\} \tag{146}$$

$$\mathcal{X}_0 = \{\exists s \in \mathcal{A} : |\|X_{s_0}\|^p - 1| > \delta_0\} \tag{147}$$

and note that a combination of a union bound and the result from Theorem 18 yields $\mathbb{P}\left[\mathcal{X}_{\text{sup}} \cup \mathcal{X}_0\right] \le \sum_{n\ge 0} P_n$. Let $\omega \in (\mathcal{X}_{\text{sup}} \cup \mathcal{X}_0)^c$ and set $x(t) = X_t(\omega)$. In this case, we can use the triangle inequality to show

$$\|x(t)\| \le \inf_{s\in\mathcal{A}(t)} \left(\|x(s_0)\| + \|x(t) - x(s_0)\|\right) \tag{148}$$

$$\le \sup_{s\in\mathcal{A}(t)} \|x(s_0)\| + \inf_{s\in\mathcal{A}(t)} \|x(t) - x(s_0)\| \tag{149}$$

$$\le \sqrt[p]{1+\delta_0} + \beta_u(T,\mathcal{A}). \tag{150}$$

Similarly

$$\|x(t)\| \ge \sup_{s\in\mathcal{A}(t)} \left(\|x(s_0)\| - \|x(t) - x(s_0)\|\right) \tag{151}$$

$$\ge \inf_{s\in\mathcal{A}(t)} \|x(s_0)\| - \inf_{s\in\mathcal{A}(t)} \|x(t) - x(s_0)\| \tag{152}$$

$$\ge \sqrt[p]{1-\delta_0} - \beta_u(T,\mathcal{A}). \tag{153}$$

By concavity of the pth root, we have

$$\left(\sqrt[p]{1+\delta_0} + \beta_u(T,\mathcal{A})\right)^p - 1 \ge 1 - \left(\sqrt[p]{1-\delta_0} - \beta_u(T,\mathcal{A})\right)^p, \tag{154}$$

so that the event $\{\sup_{t\in T} |\|X_t\|^p - 1| > \left(\sqrt[p]{1+\delta_0} + \beta_u(T,\mathcal{A})\right)^p\}$ must be contained in $\mathcal{X}_{\text{sup}} \cup \mathcal{X}_0$. □

The key to obtaining useful bounds from Theorem 19 is to construct suitable chains together with the deviation sequence u_n that give a small value for the probability bound. This is quite obviously a very difficult problem with many degrees of freedom, not least because the summation involves the topology of the set T.

4—Concentration and covering numbers

Let (T, d) be a separable *metric* space. The following theorem is a version of Theorem 19 where we restrict the admissible sequences to such sequences $\mathcal{A}$ where the set $\{s_n : s \in \mathcal{A}\}$ is an ε_n-cover of T for each n. This restriction leads to results that are not necessarily optimal for all metric spaces (T, d). However, as shown in Chapter IX, the resulting bounds are better than the best known bounds in some important special cases.

Theorem 20 (Concentration with ε-covers)**.** *Let (T, d) be a metric space for which π is a k-regular ε-sequence with upper bound $\overline{\pi}$. Let $p \geq 1, \delta > 0$ and let $(\delta_n)_{n\geq 0}$, $(\varepsilon_n)_{n\geq 0}$ be sequences satisfying*

$$\sqrt[p]{1+\delta_0} + \sum_{n\geq 1} \varepsilon_{n-1} \sqrt[p]{1+\delta_n} \leq \sqrt[p]{1+\delta}. \tag{155}$$

Set

$$\Delta = \{(t, \pi_\varepsilon(t)) : t \in T, 0 < \varepsilon \leq \varepsilon_0\} \tag{156}$$

and let let $X = (X_t)_{t\in T}$ be an almost surely continuous process with values in a normed space that satisfies

$$\mathbb{P}[\|X_t\|^p - 1 \gtrless \pm u] \leq f(u) \;\; \forall t \in T \;\; \textit{and} \tag{157}$$

$$\mathbb{P}[\|X_t - X_s\|^p > (1+u)d(s,t)^p] \leq f(u) \;\; \forall (s,t) \in \Delta \tag{158}$$

for some monotonically decreasing function f. Then

$$\mathbb{P}\left[\sup_{t\in T} |\|X_t\|^p - 1| > \delta\right] \leq 2\overline{\pi}(\varepsilon_0) f(\delta_0) + \sum_{n\geq 1} \overline{\pi}(\varepsilon_n) f(\delta_n). \tag{159}$$

Proof. We write $\pi_n = \pi_{\varepsilon_n}$ and $T_n = \pi_n(T)$. For a given $t \in T$, we can construct a sequence as $s_n = \pi_{n+1}(t)$. Hence, by requiring $s_n \in T_n$ for all $s \in \tilde{\mathcal{A}}(t), t \in T$, we obtain a chaining set $\tilde{\mathcal{A}}$. However, for each $\tilde{s} \in \tilde{\mathcal{A}}(t)$ and each $n > 0$, we can also find a sequence $s \in \tilde{\mathcal{A}}(t)$ for which $s_{n-1} = \pi_n(s_n)$ (the convergence $s_n \to t$ is not affected). Consequently, there must be some $s \in \tilde{\mathcal{A}}(t)$ for which $s_{n-1} = \pi_{n-1}(s_n)$ for all $n \in \mathbb{N}$. Thus, we can restrict $\tilde{\mathcal{A}}$ to the set $\mathcal{A}$ that consists only of sequences s satisfying $s_{n-1} = \pi_{n-1}(s_n)$ for all $n > 0$. This restriction imposes a tree-structure (multiple disjoint trees if $|T_0| > 1$) on the chaining set and lets us use a union bound over less arguments below.

Next, let $u_n(s_n, s_{n-1}) = \sqrt[p]{1+\delta_n} d(s_n, s_{n-1})$. For $n \geq 1$, a union bound over all possible combinations of s_n and s_{n-1} (there are only $|T_n|$ combinations) yields

$$P_n = \mathbb{P}\left[\exists s \in \mathcal{A} : \|X_{s_n} - X_{s_{n-1}}\| > u_n(s_n, s_{n-1})\right] \tag{160}$$

$$= \mathbb{P}\left[\exists s \in \mathcal{A} : \|X_{s_n} - X_{s_{n-1}}\|^p > d(s_n, s_{n-1})^p (1+\delta_n)\right] \tag{161}$$

$$\leq |T_n| f(\delta_n) \leq \overline{\pi}(\varepsilon_n) f(\delta_n) \tag{162}$$

as $(s_n, s_{n-1}) \in \Delta$. For $n = 0$, a union bound yields

$$P_0 = \mathbb{P}\left[\exists s \in \mathcal{A} : |\|X_{s_0}\|^p - 1| > \delta_0\}\right] \leq 2|T_0| f(\delta_0). \tag{163}$$

We can thus apply Theorem 19 and note that

$$\beta_u(\mathcal{A}) = \sup_{t\in T} \inf_{s\in\mathcal{A}(t)} \sum_{n\geq 1} \sqrt[p]{1+\delta_n} d(s_n, s_{n-1}) \leq \sum_{n\geq 1} \sqrt[p]{1+\delta_n} \varepsilon_{n-1}. \tag{164}$$

□

5—Non-isotropic distributions

A key ingredient of Theorem 20 is the isotropic bound

$$\mathbb{P}[\|X_t - X_s\|^p > (1+u)d(s,t)^p] \leq f(u) \;\; \forall (s,t) \in \Delta \tag{165}$$

where f is independent of t and s. For non-isotropic distributions, where f depends on s and t, we could still obtain a uniform bound by taking the worst possible combination of elements s and t. However, this leads to sub-optimal results as visible in Ch. IX, Sect. 3, where we use the inequality $\|x\|_1 \leq \sqrt{k}\|x\|_2$ for k-sparse vectors from Example 11. A similar difficulty arises with matrices with sub-exponential entries where bounds depend on $\|x\|_\infty$ (cf. [64]). While it is possible, in principle, to use Theorem 19 even for non-isotropic distributions, the choice of suitable chaining sets is much more difficult.

6—Chaining set construction

We give two examples of functions f for which we can find sequences ε_n, δ_n that lead to simple expressions for the probability bounds in Theorem 20. The first one exploits an inequality that holds for log-concave functions and we attempt to calculate a nearly optimal constant for ε_0 and for this reason the proof is rather technical. The second result applies to functions $f(u) = \sigma\omega(u)^{-q}$ with a convex function ω, for example $f(u) = (1+u)^{-q}$. Such functions occur when analyzing concentration-of-measure phenomena for "power-law" or "Cauchy-type" distributions (cf., e.g., [65]).

Lemma 21 (Log-concave concentration in regular spaces)**.** *Let g be a k-regular growth function and f a log-concave function with $f(0) = 1$. Let $p \geq 1, \delta > 0, \alpha \in [0.5, 1)$. Set $\delta_0 = \alpha\delta$ and*

$$\varepsilon_0 = \frac{(1-\alpha)\delta}{p\sqrt[p]{2^{p-1}(1+2\alpha)}} \;\; \text{and } N_0 = \varepsilon_0 \sqrt[k]{g(\varepsilon_0)}. \tag{166}$$

If $\delta \leq 1 - 1/N_0$ and $g(\varepsilon_0)f(\delta_0) \leq 1/2$, there is a double sequence $(\delta_n, \varepsilon_n)_{n\geq 1}$ with

$$\sum_{n\geq 1} g(\varepsilon_n) f(\delta_n) \leq g(\varepsilon_0) f(\alpha\delta) \tag{167}$$

and

$$\sqrt[p]{1+\delta_0} + \sum_{n\geq 1} \varepsilon_{n-1} \sqrt[p]{1+\delta_n} \leq \sqrt[p]{1+\delta}. \tag{168}$$

The range for δ is large if N_0 is large. For example, for k-sparse signals in $\mathbb{R}^M$ we have $N_0 = \varepsilon_0 \sqrt[k]{g(\varepsilon_0)} \geq 2eM/k$. The proof is shown in Appendix 10.

Lemma 22 (Cauchy-type concentration in regular spaces)**.** *Let g be a k-regular growth function and $f(u) = \omega(u)^{-q}$ where $q > 1 + k/p$, $p \geq 1$, and where ω is a monotonically increasing convex function with $\omega(0) \geq 0$. Let $\alpha, \delta \in (0, 1)$ and set $\delta_0 = \alpha\delta$ and*

$$\varepsilon_0 = \frac{(1-\alpha)\delta}{2p\,(1+\zeta)},\ \zeta = 2^{1+1/k+1/p}\left(\frac{\omega(\alpha\delta)}{\omega(1)-\omega(0)}\right)^{q/k}. \tag{169}$$

Then, there is a double sequence $(\delta_n, \varepsilon_n)_{n\geq 1}$ with

$$\sum_{n\geq 1} g(\varepsilon_n) f(\delta_n) \leq g(\varepsilon_0) f(\alpha\delta) \tag{170}$$

and

$$\sqrt[p]{1+\delta_0} + \sum_{n\geq 1} \varepsilon_{n-1} \sqrt[p]{1+\delta_n} \leq \sqrt[p]{1+\delta}. \tag{171}$$

The proof is shown in Appendix 11.

7—Relation to the generic chaining

Instead of using chaining sets $\mathcal{A}$ where $T_n = \{s_n : s \in \mathcal{A}\}$ are ε_n-covers, we can use arbitrary other restrictions. In the generic chaining literature, the sets T_n have to satisfy $|T_n| \leq 2^{2^n}$ but are otherwise arbitrary. Let $\mathcal{C} \subset E \cap \partial B_E$ be a subset of unit-norm vectors in a metric space E and let $\mathcal{T}$ consist of sequences of subsets $C = (C_n)_{n\geq 0}$ with $C_n \subset \mathcal{C}$ and $|C_n| \leq 2^{2^n}$. Furthermore, assume that for each $C \in \mathcal{T}$ and each $x \in \mathcal{C}$, we can find a sequence x_n with $x_n \in C_n$ and $x_n \to x$. Then

$$\mathcal{A}_C = \{(x_n)_{n\geq 0} : x_n \in C_n \forall n\} \tag{172}$$

is a chaining set for $\mathcal{C}$.

The following corollary of Theorem 19 is similar to Theorem 4.8 in [60]. Note that we use the process $(\|Ax\|^p)_{x\in\mathcal{C}}$ instead of $(Ax)_{x\in\mathcal{C}}$ as in Theorem 20. As a result, we need the inequality (174) with the difference outside of the norm, which does not immediately follow from (173) (see also [58] and Assumption 2 in Ch. IX, Sect. 6). The conditions (173) and (174) are also known as ψ_α-continuity conditions (cf. [66]).

Theorem 23 (Generic chaining version of Theorem 1)**.** *Let E, F be finite-dimensional normed spaces, $A\colon E \to F$ a random linear operator, $\mathcal{C} \subset E \cap \partial B_E$ a subset of the unit sphere in E. Let $p \geq 1, \delta \in (0,1), \alpha > 0$, and assume that the random operator A satisfies*

$$\mathbb{P}\left[\|Ax\|^p - 1 \gtrless \pm u\right] \leq \exp(-cu^\alpha) \quad \forall x \in \mathcal{C} \tag{173}$$

and

$$\mathbb{P}\left[|\|Ax\|^p - \|Ay\|^p| > u\right] \leq 2\exp(-cu^\alpha) \quad \forall x, y \in \mathcal{C} \tag{174}$$

for some $c > 0$. *Then,*

$$\mathbb{P}\left[\sup_{x\in\mathcal{C}} |\|Ax\|^p - 1| > \delta\right] \leq \sum_{n\geq 0} 2^{2^{n+1}} \exp(-c2^n u^\alpha) \tag{175}$$

with

$$\delta = \left(\sqrt[p]{1+u} + 2u\gamma_\alpha(\mathcal{C})\right)^p - 1 \tag{176}$$

and

$$\gamma_\alpha(\mathcal{C}) = \inf_{C\in\mathcal{T}} \sup_{x\in\mathcal{C}} \sum_{n\geq 0} 2^{n/\alpha} d(x, C_n). \tag{177}$$

Proof. Let $C = (C_n)_{n\geq 0} \in \mathcal{T}$. We apply Theorem 19 as follows (note that $\|Ax_n\|^p$ takes the role of X_{s_n}). First, we set $\delta_0 = u$ and note that as $|C_0| = 2$, a union bound yields

$$P_0 = \mathbb{P}\left[\exists x \in C_0 : |\|Ax\|^p - 1| > u\right] \leq 4\exp(-mu^\alpha) = f_0(u). \tag{178}$$

Let $u_n(s) = u2^{n/\alpha} d(x_n, x_{n-1})$. We use another union bound over all combinations of elements in C_n and C_{n-1} to get

$$P_n = \mathbb{P}\left[\exists x \in C_n, y \in C_{n-1} : |\|Ax\|^p - \|Ay\|^p\| > u2^{n/\alpha} d(x,y)\right] \tag{179}$$

$$\leq 2|T_n||T_{n-1}| \exp(-c2^n u^\alpha) \tag{180}$$

$$\leq 2^{2^{n+1}} \exp(-c2^n u^\alpha) = f_n(u). \tag{181}$$

An application of Theorem 19 results in

$$\mathbb{P}\left[\sup_{x\in\mathcal{C}} |\|Ax\|^p - 1| > \left(\sqrt[p]{1+u} + \beta_u(\mathcal{A}_C)\right)^p - 1\right]$$

$$\leq \sum_{n\geq 0} f_n(u) = \sum_{n\geq 0} 2^{2^{n+1}} \exp(-c2^n u^\alpha). \tag{182}$$

Finally, as the sequence $(C_n)_{n\geq 0}$ was arbitrary, we optimize over chaining sets to obtain

$$\inf_{C\in\mathcal{T}} \beta_u(\mathcal{A}_C) = \inf_{C\in\mathcal{T}} \sup_{x\in\mathcal{C}} \inf_{y\in C(x)} \sum_{n\geq 1} u2^{n/\alpha} d(y_n, y_{n-1}) \tag{183}$$

$$\leq 2u \inf_{C\in\mathcal{T}} \sup_{x\in\mathcal{C}} \sum_{n\geq 0} 2^{n/\alpha} d(x, C_n) \tag{184}$$

$$= 2u\gamma_\alpha(\mathcal{C}) \tag{185}$$

where we used

$$d(y_n, y_{n-1}) \leq d(y_n, x) + d(y_{n-1}, x) \leq d(x, C_n) + d(x, C_{n-1}). \tag{186}$$

□

It is possible to use Dudley's bound, $\gamma_2^2(T) \lesssim \log N(T, d, \varepsilon_0)$ to derive a version of Theorem 20 from Theorem 23. On the other hand, it is shown in [63] that it is possible to construct sets T and metrics d for which $\gamma_2^2(T)$ is genuinely smaller than $\log N(T, d, \varepsilon_0)$. Hence, it is not possible to derive Theorem 23 from Theorem 20. Nevertheless, for metric spaces T where no better estimate of $\gamma_\alpha(T)$ is available than a covering number estimate, Theorem 20 provides much tighter bounds than Theorem 23.

VIII
Overview of distributions with log-concave tails

In this chapter, we present an overview of inequalities of the form (P-RIP) for some common constructions of random matrices. These are derived from the convex conjugate of the cumulant generating function (CGF) of a random variable (RV). The CGF of a real-valued RV X is defined as

$$\Psi(\lambda) = \log \mathbb{E} \exp(\lambda X). \tag{187}$$

If $\Psi(\lambda)$ is non-trivial in a neighborhood of zero, $\Psi(\lambda) < \infty$ for $\lambda \in [-\varepsilon, \varepsilon]$, then X is called a subexponential RV and it is possible to derive a multitude of *qualitative* results regarding its concentration and tail decay based on this property alone [67]. For many common distributions, we not only know that the CGF Ψ is finite, but we can also calculate its values numerically. This extra information can be used to derive *quantitative* concentration and tail decay results.

To derive these bounds, we use the convex conjugate function of the CGF, which is strongly related to the exponential Markov inequality (the Chernoff bound) and its usefulness for deriving tail bounds has been recognized in the literature. In particular, the convex conjugate of the CGF is used as a *rate function* in the theory of large deviations (e.g., Cramér's theorem and its consequences such as Hoeffding's inequality, cf. [23, 68, 69]).

It is well known that the CGF Ψ is a convex function that is infinitely differentiable with $\Psi(0) = 0$ and $\frac{d}{d\lambda}\Psi(\lambda)|_{\lambda=0} = \mathbb{E}X$.

Definition 24 (Rate function). *Let c_X be a (not necessarily convex) upper bound of the CGF Ψ_X of a RV X. The convex conjugate function of c_X is defined as*

$$c_X^*(\mu) = \sup_{\lambda \geq 0} \mu\lambda - c_X(\lambda). \tag{188}$$

In the following, we call c_X^* a rate function *of X and remark that c_X^* is always convex.*

Lemma 25. *Let $X_1, \dots, X_m$ be independent RVs with common CGF upper bounds c_X and let $Z = m^{-1} \sum_{n=1}^m X_n$. Then $c_Z^*(\mu) = m c_X^*(\mu)$.*

We recall Cramér's Theorem:

Theorem 26 (Cramér's Theorem). *Let X be a RV with rate function c_X^*. Then $\mathbb{P}[X > \mu] \leq \exp\left(-c_X^*(\mu)\right)$.*

The following lemma is a version of Bennet's inequality (see, e.g., [70, Theorem 2.9]).

Distribution	$-\log f(u)$
real Gaussian or zero-inflated Rademacher	$m(u - \log(1+u))/2$
complex Gaussian or zero-inflated Steinhaus	$m(u - \log(1+u))$
Bernoulli subsampling of unitary matrix	$p\lfloor M/(kc_\infty)\rfloor((1+u)\log(1+u) - u)$

Table VIII.1: Point-wise log-concave tail bounds for some random matrices.

Lemma 27. *Let X be a random variable with $\mathbb{E}X = 0, \mathbb{E}X^2 = \sigma^2$, and $X \leq 1$ almost surely. Then*

$$\Psi_X(\lambda) \leq \log\left(1 + \sigma^2(\exp(\lambda) - \lambda - 1)\right). \tag{189}$$

If we denote the element in row m, column i, of A by a_{mi}, we can write

$$\|Ax\|^2 - 1 = \frac{1}{m}\sum_{n=1}^{m}(Z_m - 1), \tag{190}$$

$$Z_m = \sum_{i=1}^{M} |\sqrt{m}a_{mi}x_i|^2. \tag{191}$$

In the following examples, we use Lemmas 25 and 27 and Theorem 26 to show a bound of the form (P-RIP) as follows. First, we provide a CGF or an upper bound and corresponding rate function for a single term of the sum in (190). We then use Lemma 27 to show that $\Psi_{-(Z-1)}(\lambda) \leq \Psi_{Z-1}(\lambda)$ (or the upper bound), from which $c^*_{-(Z-1)}(\lambda) \geq c^*_{Z-1}(\lambda)$ follows. That is, we can use the same bound for the left tails as we use for the right tails. Finally, the concentration results follow by combining Lemma 25 and Theorem 26.
The bounds from Examples 7–11 are summarized in Table VIII.1.

Example 7 (Gaussian)**.** Let $X_1, \ldots, X_M$ be independent Gaussian RVs with unit variance and let $Z = |\sum_{i=1}^{M} X_i b_i|^2$ for some $b \in \mathbb{R}^M$ with $\|b\| = 1$. Then

$$\Psi_Z(\lambda) = -0.5\log(1 - 2\lambda)\,, \quad \lambda < 1/2 \tag{192}$$

and the rate function for the centered version

$$c^*_{Z-1}(\mu) = 0.5(\mu - \log(1+\mu)), \quad \mu \geq 0 \tag{193}$$

is a lower bound for $c^*_{-(Z-1)}$. □

Example 8 (Complex Gaussian)**.** Let $X_1, \ldots, X_M$ be independent complex Gaussian RVs, $X_i \sim \mathcal{N}_{\mathbb{C}}(0,1)$, and let $Z = |\sum_{i=1}^{M} X_i b_i|^2$ for some $b \in \mathbb{C}^M$ with $\|b\| = 1$. Then

$$\Psi_Z(\lambda) = \log\mathbb{E}\left[\exp(\lambda Z)\right] = -\log(1-\lambda)\,, \quad \lambda < 1. \tag{194}$$

The rate function of $Z-1$ is given by

$$c^*_{Z-1}(\mu) = \mu - \log(1+\mu), \quad \mu \geq 0 \tag{195}$$

and is a lower bound for $c^*_{-(Z-1)}$. □

Example 9 (Database-friendly projections [71])**.** Let X_i denote zero-inflated independent Rademacher random variables, $\mathbb{P}[X_i = \sqrt{c}] = \mathbb{P}[X_i = -\sqrt{c}] = 1/(2c)$ and $\mathbb{P}[X_i = 0] = 1 - 1/c$ for $1 \leq c \leq 3$. Let $Z = (\sum_{i=1}^M X_i b_i)^2$ for some $b \in \mathbb{R}^M$, $\|b\| = 1$. We show in Appendix 7 that $\Psi_{Z-1}(\lambda)$ and $\Psi_{-(Z-1)}(\lambda)$ can be bounded by

$$-\lambda - 0.5\log(1-2\lambda), \quad 0 \leq \lambda < 1/2, \tag{196}$$

i.e., we can use the same rate function as for Gaussian RVs. □

Example 10 (Database-friendly complex projections)**.** Let $X_i \in \mathbb{C}$ be a sequence of zero-inflated normalized Steinhaus random variables, that is,

$$X_n = \sqrt{c} B_n \exp(iU_n), \tag{197}$$

$$U_n \sim \mathcal{U}_{[-\pi,\pi]} \quad \text{(uniform distribution)}, \tag{198}$$

$$\mathbb{P}[B_n = 1] = 1/c, \quad \mathbb{P}[B_n = 0] = 1 - 1/c, \tag{199}$$

where all U_n and B_n are independent and $1 \leq c \leq 2$. Let $Z = \left|\sum_{n=1}^M X_n b_n\right|^2$ for some $b \in \mathbb{C}^M$, $\|b\| = 1$. We show in Appendix 8 that $\Psi_{Z-1}(\lambda)$ and $\Psi_{-(Z-1)}(\lambda)$ can be bounded by

$$-\lambda - \log(1-\lambda), \quad 0 \leq \lambda < 1 \tag{200}$$

and that, consequently, Z has the same rate function as a complex Gaussian RV. □

Example 11 (Structured sampling matrices)**.** Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ and let $u_1, \ldots, u_M \in \mathbb{K}^M$ be an orthonormal basis with $\max_n \|u_n\|_\infty^2 \leq c_\infty/M$. Let $X_n = B_n |\langle x, u_n\rangle|^2/p$ for some $x \in \mathbb{K}^M$ with $|\operatorname{supp}(x)| \leq k$, $\|x\| = 1$, and where B_n are independent Bernoulli RVs with $\mathbb{P}[B_n = 1] = p$. We have (with $q = 1-p$)

$$\Psi_n(\lambda) = \log \mathbb{E}\exp(\lambda X_n) = \log(q + p\exp(\lambda|\langle x, u_n\rangle|^2/p)) \tag{201}$$

and, for $Z = \sum_{n=1}^M X_n$, by independence of the B_n,

$$\Psi_Z(\lambda) = \sum_{n=1}^M \log(q + p\exp(\lambda|\langle x, u_n\rangle|^2/p)). \tag{202}$$

The normalization is chosen such that $\mathbb{E}Z = 1$. We show in Appendix 9 that

$$\Psi_{Z-1}(\lambda) \leq pM_{\text{eff}}\left(\exp\left(\frac{\lambda}{pM_{\text{eff}}}\right) - 1 - \frac{\lambda}{pM_{\text{eff}}}\right) \tag{203}$$

with

$$M_{\text{eff}} = \lfloor M/(kc_\infty) \rfloor \tag{204}$$

and rate function

$$c^*_{Z-1}(\mu) = pM_{\text{eff}}\left((1+\mu)\log(1+\mu) - \mu\right) \tag{205}$$

which is also a lower bound for $c^*_{-(Z-1)}$. □

IX
The RIP in Unions of Subspaces

In this chapter, we state our main restricted isometry theorem and show how it can be used to recover or improve known results for sparse signals, low-rank matrices, and Lipschitz unions of subspaces when measured with various types of random matrices. Even though we do not exploit any special structure of the set $\mathcal{C}$ and only exploit log-concavity of the function f, we obtain results that are competitive or surpass known results in terms of the constants involved. Two such special structures are sparse signals and low-rank matrices. In the case of sparse signals, the set $\mathcal{C}$ is given as the intersection of finitely many subspaces with the unit sphere. In [23, 25] this fact is exploited to reason about the operator norm between points of the ε-covers. In particular, a chaining argument is not necessary. Similarly for rank-k matrices, it is exploited in [62] that any sum $X_1 + X_2$ of two rank-k matrices can also be written as the sum of two other rank-k matrices X_1', X_2' with $\|X_1'\| + \|X_2'\| \leq \sqrt{2}\|X_1 + X_2\|$. Also in this case, a chaining argument is not necessary.

Theorem 28 (RIP in geometrically regular spaces)**.** *Let E, F be finite-dimensional normed spaces, $A\colon E \to F$ a random linear operator, $\mathcal{C} \subset E \cap \partial B_E$ a k-regular subset of the unit sphere in E. Let π be an ε-sequence of $\mathcal{C}$ with upper bound $\overline{\pi}$ and set*

$$\Delta = \left\{ \frac{x - \pi_\varepsilon(x)}{\|x - \pi_\varepsilon(x)\|}, x \in \mathcal{C}, 0 < \varepsilon \leq \varepsilon_0 \right\} \tag{206}$$

where ε_0 is defined below. Let $\delta \in (0,1), \alpha \in (0,1)$, and assume that the random operator A satisfies

$$\mathbb{P}\left[\|Ax\|^2 - 1 \gtrless \pm u\right] \leq f(u) \quad \forall x \in \mathcal{C} \cup \Delta. \tag{207}$$

i) If f is monotonically decreasing and log-concave with $f(0) = 1$, let

$$\varepsilon_0 = \frac{(1-\alpha)\delta}{\sqrt{8(1+2\alpha)}} \quad \textit{and } N_0 = \varepsilon_0 \sqrt[k]{\overline{\pi}(\varepsilon_0)}. \tag{208}$$

and assume that $\delta \leq 1 - 1/N_0$ and $\alpha \geq 1/2$.

ii) If $f(u) = \omega(u)^{-q}$ with a monotonically increasing convex function ω with $\omega(0) > 0$ and $q > 1 + 3k/2$, set

$$\varepsilon_0 = \frac{(1-\alpha)\delta}{4\,(1+\zeta)}, \; \zeta = 2^{1/k+3/2}\left(\frac{\omega(\alpha\delta)}{\omega(1) - \omega(0)}\right)^{q/k}. \tag{209}$$

Then,

$$\mathbb{P}\left[\sup_{x\in\mathcal{C}} |\|Ax\|^2 - 1| > \delta\right] \leq 3\overline{\pi}(\varepsilon_0) f(\alpha\delta). \tag{210}$$

Equivalently, if

$$-\log f(\alpha\delta) \geq \log \overline{\pi}(\varepsilon_0) + \log(3/\xi) \tag{211}$$

then A has the RIC δ with probability at least $1-\xi$.

Proof. This theorem follows directly from Theorem 20 combined with the sequences found in Lemma 21 or Lemma 22 where we note that we can assume $g(\varepsilon_0)f(\alpha\delta) \leq 1/3 \leq 1/2$ as otherwise the probability bound is trivial. The random process is given by $X_t = At$ for $t \in \mathcal{C}$. As A is a linear operator between finite-dimensional spaces, X is almost surely continuous. Furthermore, we have

$$\mathbb{P}\left[\|Ax - Ax'\|^2 > (1+u)\|x-x'\|^2\right] = \mathbb{P}\left[\left\|\frac{A(x-x')}{\|x-x'\|}\right\|^2 - 1 > u\right] \leq f(u) \tag{212}$$

for $(x, x') \in \Delta$ so that the property (207) implies (157) and (158). □

To establish a recovery result, we always proceed as follows: We first fix a sensing matrix that satisfies the tail and concentration conditions for a given set $\mathcal{C}$. Next, we fix α such that the result compares as best possible with some known result. We then select ε_0 slightly smaller than required by Theorem 28 to simplify the expressions and finally use the covering number estimates of the set $\mathcal{C}$ to establish the result.

1—Gaussian or Rademacher matrices and sparse signals

Let $M \geq k$ and let $\mathcal{C} = \Sigma_k \cap \partial B_{\mathbb{R}^M}$ denote real k-sparse signals with unit norm for which there are ε-covers with size less than $g(\varepsilon) = (eM/k)^k (2+\varepsilon)^k/\varepsilon^k$ (see Table VII.1). If the random matrix $A \in \mathbb{R}^{m\times M}$ has iid. Gaussian or Rademacher entries with variance $1/m$, the concentration inequality (207) holds with $\log f(u) = -m\omega(u)$ and $\omega(u) = (u - \log(1+u))/2$ (see Table VIII.1).
Let $\alpha = 9/10$ and $\varepsilon_0 = 2\delta/95$ so that $(2+\varepsilon_0)/\varepsilon_0 \leq 96/\delta$. Theorem 28 then shows that A has the RIP with constant δ with probability at least $1-\xi$ if

$$m\omega(0.9\delta) \geq k \log\left(\frac{96eM}{\delta k}\right) + \log(3/\xi) \tag{213}$$

and if $\delta \leq 1 - k/(2eM)$. If we plug in ω, we obtain the condition

$$m \geq \frac{k\log(96/\delta) + k\log(eM/k) + \log(3/\xi)}{(0.9\delta - \log(1+0.9\delta))/2}. \tag{214}$$

For Rademacher matrices, we get an improvement by a factor two (approximately) over the condition [23, Theorem 9.11],[1]

$$m \geq 8\delta^{-2}\left(9k + 2k\log(M/k) + 2\log(2/\xi)\right). \tag{215}$$

For Gaussian matrices, we compare with the condition [23, Theorem 9.27][2],

$$m \geq \frac{\left(\sqrt{2} + \sqrt{1/\log(eM/k)}\right)^2 \left(k\log(eM/k) + \log(2/\xi)\right)}{\delta + 2 - 2\sqrt{1+\delta}}. \tag{216}$$

One can verify that for $\delta \geq 0.1, M \geq 10k$, condition (214) improves upon (216) (in the asymptotic regime $M/k \to \infty$, the factor is roughly 1.65). Consequently, (214) provides the best conditions for these kinds of matrices that we are aware of.

2—Sub-Gaussian matrices and sparse signals

Let $\mathcal{C}$ be as in Sect. 1. If the random matrix $A \in \mathbb{R}^{m\times M}$ has independent sub-Gaussian rows a with $\mathbb{E}|\langle a, x\rangle|^2 = 1/m$ for each x with $\|x\| = 1$ with common sub-Gaussian parameter $\tilde{c}$, i.e., $\mathbb{E}\exp(\lambda\langle a, x\rangle) \leq \exp(\tilde{c}\lambda^2)$ for each x with $\|x\| = 1$, $\lambda \in \mathbb{R}$ (cf. [23, Def. 9.4]), then the concentration inequality (206) holds for some monotonically decreasing and log-concave function f that satisfies $f(u) \leq \exp(-mcu^2)$ for $u < 1$ with c depending on $\tilde{c}$ (cf. [23, Lemma 9.8]).
Let $\alpha = \sqrt{3/4}$ and $\varepsilon_0 = \delta/35$ so that $(2+\varepsilon_0)/\varepsilon_0 \leq 71/\delta$. Theorem 28 then shows that A has the RIP with constant δ with probability at least $1 - \xi$ if

$$m \geq 4\left(k\log\left((71eM)/(k\delta)\right) + \log(3/\xi)\right)/(3c\delta^2). \tag{217}$$

If we use $2\log(71e) \leq 11$, we obtain

$$m \geq 2\left(11k + 2k\log\left(M/(k\delta)\right) + 2\log(3/\xi)\right)/(3c\delta^2) \tag{218}$$

which is slightly worse (about a factor 1.5 for $\delta \geq 0.1$ and $M/k \geq 10$) than the condition in [23, Theorem 9.11],

$$m \geq 2\left(9k + 2k\log(M/k) + 2\log(2/\xi)\right)/(3c\delta^2). \tag{219}$$

[1] For real Gaussian or Rademacher matrices, we can achieve $C = 8$ in [23, Theorem 9.11] as $C = 2/(3\tilde{c})$ with $\tilde{c}$ such that

$$\mathbb{P}\left[|\|Ax\|^2 - 1\| > u\right] \leq 2\exp(-m\tilde{c}u^2).$$

If we use the bound $\mathbb{P}[\ldots] \leq 2\exp(-m\omega(u))$ with $\omega(u) = 0.5(u - \log(1+u))$, we can derive the factor $\tilde{c}$ from the expansion

$$\omega(u) = 0.5\,(u - \log(1+u)) \geq u^2/4 - u^3/6 \geq u^2/12 \text{ for } u \leq 1$$

so that we can use $\tilde{c} = 1/12$ and $2/(3\tilde{c}) = 8$.

[2] We substituted η in the original formulation using Eq. (9.48) in [23, Theorem 9.27].

3—Structured matrices and sparse signals

Let $\mathcal{C} = \Sigma_k \cap \partial B_{\mathbb{K}^M}$ denote k-sparse signals with unit norm in $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. If A is a unitary matrix with $\|A\|_\infty^2 \leq c_\infty/M$ and rows selected according to a Bernoulli distribution with parameter p_b, the concentration inequality (207) holds with $f(u) = \exp(-p_b M_{\text{eff}}((1+u)\log(1+u) - u))$ for each $x \in \mathcal{C}$ where $M_{\text{eff}} = \lfloor M/(kc_\infty)\rfloor$ (see Table VIII.1). As we can construct ε-covers for sparse signals for which the differences $x - \pi_\varepsilon(x)$ are also k-sparse, we have $\Delta \subset \mathcal{C}$ and it is sufficient that the concentration inequality (206) holds for $x \in \mathcal{C}$ to use Theorem 28.
We use $\varepsilon_0 = (1-\alpha)\delta/5$, which is slightly smaller than the value in Theorem 28, and for which $(2+\varepsilon_0)/\varepsilon_0 \leq 11/(\delta(1-\alpha))$. Theorem 28 then shows that A has the RIP with constant δ with probability at least $1-\xi$ if

$$p_b M_{\text{eff}} \geq \frac{k(\kappa \log(11/(\delta(1-\alpha))) + \log(eM/k)) + \log(3/\xi)}{(1+\alpha\delta)\log(1+\alpha\delta) - \alpha\delta}\ . \tag{220}$$

If M/k is an integer and if $c_\infty = 1$ (for example, if A contains discrete Fourier transform vectors), we obtain the condition

$$p_b M \geq \frac{k^2(2.4\kappa + \log(eM/k) + \log(\delta(1-\alpha))) + k\log(3/\xi)}{(1+\alpha\delta)\log(1+\alpha\delta) - \alpha\delta} \tag{221}$$

where $\kappa = 1$ if $\mathbb{K} = \mathbb{R}$ and $\kappa = 2$ if $\mathbb{K} = \mathbb{C}$. The term $p_b M$ is the expected number of rows of A. The bound is a factor $ck/\log^2(k)$ away from the best known results [72]

$$\mathbb{E}\#\,\text{rows} \geq Ck\log^2(k)\log M, \tag{222}$$

but is useful for finite values of k and M ($c, C > 0$ are some constants independent of k and M, see also [73, 74]).

4—Low-rank matrices with Gaussian measurements

Let $\mathcal{U} \subset \mathbb{R}^{n_1 \times n_2}$ denote the set of all matrices with rank at most k equipped with the Frobenius norm $\|X\|_F = \sqrt{\text{tr}(X^T X)}$. Assume that the random linear operator $\mathcal{A}\colon \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ satisfies

$$\mathbb{P}\left[\|\mathcal{A}(X)\|_2^2 - \|X\|_F^2 \lesseqgtr \pm u\right] \leq f(u) \tag{223}$$

for each $X \in \mathcal{U}$ with $\|X\|_F = 1$ and some log-concave and monotonically decreasing function f with $f(0) = 1$. If $(\mathcal{A}(X))_i = \sum_{pq} a_{ipq}[X]_{pq}$ where $[X]_{pq}$ is the element in row p and column q of X with appropriately scaled iid. Gaussian entries a_{ipq}, this is the case with $f(u) = \exp(-m(u - \log(1+u))/2)$ (see Example 7 and also [62]). As the set of unit-norm rank-k matrices is geometrically regular with $\overline{\pi}(\varepsilon) = (9/\varepsilon)^r$ and $r = (n_1 + n_2 + 1)k$ (see Table VII.1), we can use Theorem 28 with $\varepsilon_0 = \delta(1-\alpha)/5$ to obtain

$$\mathbb{P}\left[\sup_{X\in\mathcal{U}}\left|\|\mathcal{A}(X)\|_2^2 - \|X\|_F^2\right| > \delta\right] \leq 3\left(\frac{45}{\delta(1-\alpha)}\right)^{k(n_1+n_2+1)} f(\alpha\delta). \tag{224}$$

This is a quantitative version of Theorem 2.3 in [62].

5—Lipschitz unions of subspaces

Let $\mathcal{U} = \cup_{t \in T} S_t \subset \mathcal{H}$ denote a Lipschitz union of complex subspaces, that is, $\dim(S_t) = k$ for each t and

$$\|P_t - P_s\|^2 \leq L^2 \sum_{n=1}^{k} (t_n - s_n)^2 \tag{225}$$

and $T \subset B_{\mathbb{R}^k}$. Let $A\colon H \to \mathbb{C}^m$ be a random operator that fulfills the concentration inequality (207) for a monotonically decreasing and log-concave function f. Then, we can use Theorem 28 with $\alpha = 9/10$, $\varepsilon_0 = \delta/48$ to find that A has the RIP with constant δ with probability at least $1 - \xi$ if

$$m \geq \frac{k \log(56L) + 3k \log(48/\delta) + \log(3/\xi)}{-\log f(9\delta/10)}. \tag{226}$$

6—Sets with low covering dimension

The authors of [58] also present a chaining-based proof to show that a matrix A has the RIP for a subset of the unit sphere $\mathcal{C} \subset \partial B_E$. For conciseness, we only consider 2-norms and the isotropic case where $\mathbb{E}\|Ax\|^2 = 1$ holds for each $x \in \partial B_E$. Their results are based on two assumptions:

Assumption 1 (Geometric regularity)**.** *There exists a tuple $(k_{\mathit{eff}}, \varepsilon_{\mathcal{C}})$ such that there is an ε-cover $\mathcal{U}$ of $\mathcal{C}$ with*

$$|\mathcal{U}| \leq \varepsilon^{-k_{\mathit{eff}}} \text{ for all } \varepsilon \leq \varepsilon_{\mathcal{C}}. \tag{227}$$

Assumption 2 (Sub-Gaussianity)**.** *There exist constants $c_1, c_2 > 0$ such that for each $x, y \in \mathcal{C} \cup \{0\}$*

$$\mathbb{P}\left[\Big| \|Ax\|^2 - \|Ay\|^2 \Big| \geq \lambda \|x - y\|\right] \leq 2e^{-c_1 m \lambda^2} \tag{228}$$

for each $0 \leq \lambda \leq c_2/c_1$ and

$$\mathbb{P}\left[\Big| \|Ax\|^2 - \|Ay\|^2 \Big| \geq \lambda \|x - y\|\right] \leq 2e^{-c_2 m \lambda} \tag{229}$$

for each $\lambda \geq c_2/c_1$.

The first assumption is satisfied if the set $\mathcal{C}$ has a finite upper box-counting dimension, which is equivalent to assuming that $\mathcal{C}$ is k-regular. For example, if $N(\mathcal{C}, d, \varepsilon) \leq (N_0/\varepsilon)^k$, then $k_{\text{eff}} = k + s, \varepsilon_{\mathcal{C}} = N_0^{-k/s}$ is a valid pair that satisfies Assumption 1 for each $s > 0$.
The second assumption can be stated in terms of log-concavity. If we define the random process $Z_x = \|Ax\|^2$, then Assumption 2 can be written as

$$\mathbb{P}\left[d_1(Z_x, Z_y) \geq u d(x, y)\right] \leq 2f(u) \tag{230}$$

with $d_1(Z_x, Z_y) = |Z_x - Z_y|$ and the log-concave function

$$f(u) = \begin{cases} \exp(-c_1 m u^2), & \text{if } 0 \leq u \leq c_2/c_1, \\ \exp(-c_2 m u), & \text{if } c_2/c_1 \leq u\,. \end{cases} \tag{231}$$

Thus, the process $Z_x = \|Ax\|^2$ satisfies condition (207). This is in contrast to Theorem 28 where it is required that the process Ax (without the norm) satisfies condition (207) with respect to a norm in $\mathbb{R}^m$ or $\mathbb{C}^m$ and which is much simpler to show (because we have the difference $x - y$ inside the norm).
The authors then show the following theorem:

Theorem 29 (Theorem 4 in [58])**.** *If Assumptions 1 and 2 hold, then A has the RIP with RIC δ for signals in $\mathcal{C}$ with probability at least $1 - \xi$ provided that*

$$m \geq \frac{3200}{\min(c_1, c_2)\delta^2} \max\big(k_{\mathit{eff}} \log(1/\varepsilon_{\mathcal{C}}), \log(6/\xi)\big). \tag{232}$$

To see how this compares with Theorem 28, let $A \in \mathbb{R}^{m \times M}$ be a random matrix with iid. Gaussian entries so that $f(u) = \exp(-m(u - \log(1-u))/2)$ (cf. Table VIII.1) and assume that the covering number of $\mathcal{C}$ is bounded by $g(\varepsilon) = (4\sqrt{L})^k/\varepsilon^k$ (e.g., a Lipschitz union of $k/2$-dimensional real subspaces if k is even, see Sect. 5). We demonstrate in Appendix 12 that Assumption 2 is satisfied with $\min(c_1, c_2) \leq 9c/64^2$ where c is a constant (one of the constants that appears when showing equivalent definitions of subexponential random variables, see, e.g., [67]). We can further set $k_{\text{eff}} = 2k + s$, $\varepsilon_{\mathcal{C}} = (4\sqrt{L})^{-k/s}$, and let $s \to \infty$, to obtain $k_{\text{eff}} \cdot (k/s) \to k$ and

$$m \geq \frac{3200 \cdot 64^2}{9c\delta^2} \max\big(0.5k \log(16L), \log(6/\xi)\big) \tag{233}$$

from Theorem 29, which has the same asymptotic characteristics (for growing L and k) as the result stated in (226).

X
THE RIP AND CHANNEL ESTIMATION

Let us return to the question whether dimensionality reduction is possible in channel estimation for communication systems. We have seen in Chapter III that a (flat-fading) communication channel can be modeled as a vector h and that prior knowledge can be encoded as $h \in \mathcal{U}$ where two relevant examples for the union of subspaces $\mathcal{U}$ are the DOA manifold and the single-cluster 3GPP manifold. In both examples, the training equation reads as

$$y = Ah + e \tag{234}$$

where $A \in \mathbb{C}^{m \times M}$ is a matrix describing the network of phase shifters between the M antennas and the m analog-to-digital converters. As we have learned throughout this work, optimal reconstruction of h is possible if A has the restricted isometry property with respect to the union of subspaces $\mathcal{U}$. In the following two sections, we use the restricted isometry theory developed in the preceding chapters to give conditions on m and M under which a randomly chosen matrix A has the RIP.

1—DOA manifold

We introduced the DOA manifold in Chapter III as

$$\mathcal{U}_{k,M} = \overline{\mathcal{U}'_{k,M}}, \quad \mathcal{U}'_{k,M} = \bigcup_{t \in [-\pi,\pi]^k} \operatorname{range}\bigl(V(t)\bigr), \tag{235}$$

where

$$V(t) = \begin{bmatrix} a(t_1) & \dots & a(t_k) \end{bmatrix} \tag{236}$$

is a partial Vandermonde matrix and $a(t_\ell)$ are the steering vectors of a uniform linear array. If $k < M$ and if all nodes t_ℓ are distinct, then $V(t)$ has full column rank and the projector onto $S_t = \operatorname{range}(V(t))$ is given by $P_t = V(t)V(t)^\dagger$ where $V(t)^\dagger$ denotes the Moore-Penrose pseudo-inverse of $V(t)$. The key to using the general restricted isometry theorem (Theorem 28) is to establish that the operator P_t is a Lipschitz function of t, because then we can use the covering number estimate for a Lipschitz union of subspaces given in Theorem 14. This analysis is complicated by the fact that $V(t)$ is rank-deficient if $t_\ell = t_j$ for some $\ell \neq j$ and that, consequently, $V(t)$ becomes ill-conditioned whenever $t_\ell \to t_j$.
The following theorem is slightly more general than needed: we show that $V(z) = [f(z_1) \dots f(z_k)]$ with $z \in \mathbb{C}^k$ and $f(z) = [1 \ z \dots z^{M-1}]^T$ is such that $P_z =$

$V(z)V(z)^\dagger$ is Lipschitz continuous with respect to z. The DOA manifold is then equivalently given as

$$\mathcal{U}_{k,M} = \overline{\mathcal{U}'_{k,M}}, \quad \mathcal{U}'_{k,M} = \bigcup_{z\in\mathbb{C}^k : |z_\ell|=1, \ell=1,\dots,k} \operatorname{range}\big(V(z)\big), \tag{237}$$

Let

$$\mathcal{B}_R = \{z \in \mathbb{C}^k : \|z\|_\infty \le R\}, \tag{238}$$

$$\mathcal{B}'_R = \{z \in \mathbb{C}^k : \|z\|_\infty \le R, z_i \ne z_j, i \ne j\} \tag{239}$$

denote the sup-norm closed ball in $\mathbb{C}^k$ with radius R and its subset of distinct elements. We establish the following theorem, which shows that the orthogonal projectors P_z onto range($V(z)$) form a Lipschitz family of orthogonal projectors.

Theorem 30. *The function* $P\colon \mathcal{B}'_R \to \mathbb{C}^{M\times M}, z \mapsto V(z)V^\dagger(z)$*, can be extended to* $\mathcal{B}_R$ *and has the Lipschitz property*

$$\|P_z - P_y\| \le L\|z - y\|_1 \tag{240}$$

where $\|z - y\|_1 = \sum_{i=1}^k |z_i - y_i|$ *is the one-norm and*

$$L \le \frac{\sqrt{e2^{k(k-1)}M^{2k+1}}}{k!} \max(1, R^{M-2}) \,. \tag{241}$$

As will be clear from the proof, which is shown in Appendix 13, the continuous extension of the projector is *not* given in terms of the Vandermonde matrix V when multiple nodes coincide (V would be rank-deficient). Instead, if $z_1 = \cdots = z_j$ the vectors $f(z_1), \dots, f(z_j)$ have to be replaced by $f(z_1)$ and its first $j-1$ derivatives $f^{(1)}(z_1), \dots, f^{(j-1)}(z_1)$.

Corollary 31. *Let* P_t *denote the orthogonal projector onto* range($V(t)$) *with* $V(t)$ *given by* (236)*. An upper bound for the covering number of the set* $T = [-\pi, \pi]^k$ *with respect to the metric*

$$d_F(s, t) = \|P_s - P_t\| \tag{242}$$

is given by

$$\mathcal{N}(T, d_F, \varepsilon) \le \left(\frac{\pi\sqrt{e2^{k(k-1)}M^{2k+1}}}{\varepsilon(k-1)!}\right)^k \times \frac{1}{(k-1)!}. \tag{243}$$

Proof. For a given $\varepsilon > 0$, let $\varepsilon' = \varepsilon/(kL)$ and let $T'_\varepsilon = \{-\pi, -\pi + 2\varepsilon', \dots, -\pi + 2(N-1)\varepsilon'\}^k$ where $N = \lceil \pi/\varepsilon' \rceil$. There are N elements per dimension in T'_ε. As $P_s = P_t$ if s is a permutation of t, it is enough to only use those elements $t \in T'_\varepsilon$ for which $t_k \ge t_{k-1} \ge \cdots \ge t_1$. One can verify that this subset $T_\varepsilon \subset T'_\varepsilon$ only has $\binom{N-1+k}{k}$ instead of N^k elements and that

$$\binom{\lceil kL\pi/\varepsilon \rceil - 1 + k}{k} \le \frac{(kL\pi/\varepsilon)^k}{(k-1)!} = \left(\frac{\pi\sqrt{e2^{k(k-1)}M^{2k+1}}}{\varepsilon(k-1)!}\right)^k \frac{1}{(k-1)!} \tag{244}$$

for $k \geq 2$ and $M \geq k$ (this upper bound is loose but nice to manipulate).
By construction, for any given $t \in T$, we can find $s \in T'_\varepsilon$ such that $|\exp(it_\ell) - \exp(is_\ell))| \leq \varepsilon'$ for each $\ell = 1, \ldots, k$ (we can find s_ℓ such that the arc-distance along the unit circle is smaller than ε' and the straight distance between $\exp(is_\ell)$ and $\exp(it_\ell)$ is smaller than the arc-distance). Then, by Theorem 30 and with a slight abuse of notation (we apply $\exp(\cdot)$ element-wise)

$$\begin{aligned} \|P_t - P_s\| &\leq L\|\exp(it) - \exp(is)\|_1 && (245)\\ &= L\sum_{\ell=1}^{k} |\exp(it_\ell) - \exp(is_\ell)| \leq L\sum_{\ell=1}^{k} \varepsilon' \leq \varepsilon. && (246)\end{aligned}$$

Thus, T'_ε and, consequently, T_ε, are ε-covers of T and the cardinality of T_ε is bounded by (244). □

We now have an upper bound for the cardinality of an ε_T-cover of the set T with respect to the Finsler metric $d_F(t, s) = \|P_t - P_s\|$. If we combine this result with Theorem 14 to bound the covering number of the DOA manifold, we can use Theorem 28 and establish the following RIP result:

Theorem 32. *Let $A \in \mathbb{C}^{m\times M}$ be a random matrix that satisfies the point-wise concentration inequality*

$$\mathbb{P}\left[(1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2\right] \leq 2f(\delta) \tag{247}$$

for all $x \in \mathbb{C}^M$ and where f is log-concave and $f(0) = 1$. Let $\alpha \in [0.5, 1)$ and set

$$\varepsilon = \frac{(1-\alpha)\delta}{\sqrt{8(1+2\alpha)}} \tag{248}$$

and

$$N_0 = (12N_T)^{1/3}, \quad N_T = \frac{\pi\sqrt{e2^{k(k-1)}M^{2k+1}}}{(k-1)!\sqrt[k]{(k-1)!}}. \tag{249}$$

Then, if $\delta \leq 1 - 1/N_0$,

$$\mathbb{P}\left[\sup_{x\in\mathcal{U}_{k,M}} \left|\|Ax\|^2 - \|x\|^2\right| > \delta\|x\|^2\right] \leq 3(N_0/\varepsilon)^{3k} f(\alpha\delta). \tag{250}$$

Proof. The statement follows directly from Theorem 28 if we can show that there is an ε-cover $\mathcal{C}_\varepsilon \subset \mathcal{C}_{k,M} = \mathcal{U}_{k,M} \cap \partial B_{\mathbb{C}^M}$ with cardinality $|\mathcal{C}_\varepsilon| \leq (N_0/\varepsilon)^{3k}$. As shown in Theorem 14 there is an ε-cover $\mathcal{C}_\varepsilon$ of $\mathcal{C}_{k,M}$ of cardinality

$$|\mathcal{C}_\varepsilon| \leq |T_{\varepsilon_T}|(2+\varepsilon_\mathcal{H})^{2k}/\varepsilon_\mathcal{H}^{2k} \tag{251}$$

if T_{ε_T} is an ε_T-cover of T with respect to the metric $d(s,t) = \|P_s - P_t\|$ and if

$$\varepsilon_\mathcal{H}^2 + \varepsilon_T^2 + (1-(1-\varepsilon_T^2)^{1/2})^2 \leq \varepsilon^2. \tag{252}$$

By Corollary 31, there is such an ε_T-cover of T with cardinality less than $(N_T/\varepsilon)^k$ with

$$N_T = \frac{\pi\sqrt{e2^{k(k-1)}M^{2k+1}}}{(k-1)!\sqrt[k]{(k-1)!}}. \tag{253}$$

Note that for $\alpha \in [0.5, 1)$ and $\delta \leq 1$, we have $\varepsilon \leq 1/8$. For this upper bound for ε, it is not difficult to verify that $\varepsilon_{\mathcal{H}} = 5\varepsilon/6$ and $\varepsilon_T = 6\varepsilon/11$ is a valid pair for which (252) holds. If we use these values for $\varepsilon_{\mathcal{H}}$ and ε_T in (251) and exploit $\varepsilon \leq 1/8$ again, we find that with $N_0 = (12N_T)^{1/3}$, we have

$$|\mathcal{C}_\varepsilon| \leq \left(\frac{6(2+5\varepsilon/6)}{5\varepsilon}\right)^{2k} \times \left(\frac{11N_T}{5\varepsilon}\right)^k \leq \frac{(12N_T)^k}{\varepsilon^{3k}} = (N_0/\varepsilon)^{3k}. \tag{254}$$

□

For example, if A is a zero-inflated Steinhaus matrix, we have $f(\delta) = m(\delta - \log(1+\delta))$ and Theorem 32 shows that

$$m \geq \inf_{0.5 \leq \alpha < 1} \frac{k\log(12N_T) + 3k\log(\sqrt{8(1+2\alpha)}/(\delta(1-\alpha))) + \log(3/\xi)}{\alpha\delta - \log(1+\alpha\delta)} \tag{255}$$

is a sufficient condition for A to have the RIP with RIC δ with probability at least $1-\xi$ and with respect to the DOA manifold with k sources and M antennas. Crucially, if we insert N_T, we see that the right-hand side only grows as $k^2 \log M$ as a function of M (the number of antennas). The lower bound on the number of ADCs as a function of the antennas is visualized in Figure 5 for $k = 3, 4, 5$ and for $\delta = 0.5$ and $\xi = 0.01$. The number of ADCs m is smaller than the number of antennas M in the shaded area and one can observe from Figure 5 that dimensionality reduction is possible once there are more than a thousand antennas (under the assumption that the model $h \in \mathcal{U}_{k,M}$ is accurate). One should keep in mind that this result is derived using the (loose) Lipschitz bound from Theorem 30, the tail bound for Steinhaus random variables (Example 10), and a series of triangle inequalities and union bounds in the proof of Theorem 28.

2—Lipschitz continuity in 3GPP models

In the single-cluster 3GPP model, the channel vector h is conditionally normal distributed according to

$$h \,|\, t \sim \mathcal{N}_{\mathbb{C}}(0, \Sigma_t) \tag{256}$$

with covariance matrix Σ_t given by

$$\Sigma_t = \int g(\tau)a(t+\tau)a(t+\tau)^H d\tau = D_t\Sigma_0 D_t^*. \tag{257}$$

If we express the steering vector $a(t+\tau)$ as

$$a(t+\tau) = \frac{1}{\sqrt{M}}\begin{bmatrix}1 & \exp(i(t+\tau)) & \ldots & \exp(i(M-1)(t+\tau))\end{bmatrix}^T = D_t a(\tau) \tag{258}$$

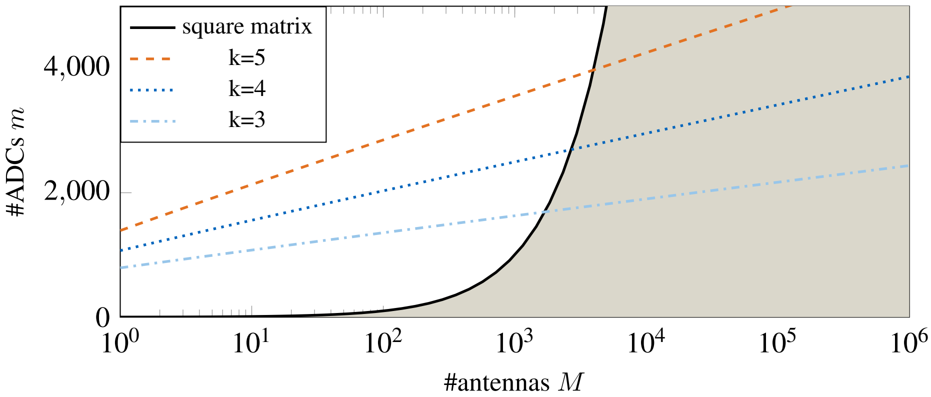

Figure 5: Minimal number of ADCs m such that in at least 99 per cent of the draws, the RIC of A is smaller than $\delta = 0.5$ for the DOA manifold with $k = 3, 4, 5$ sources (as a function of the number of antennas M). The shaded area indicates non-trivial results (wide matrix).

with the unitary diagonal matrix

$$D_t = \operatorname{diag}\big(1, \exp(it), \ldots, \exp(i(M-1)t)\big), \tag{259}$$

we can write Σ_t as

$$\Sigma_t = D_t \int g(\tau)a(\tau)a(\tau)^H d\tau \; D_t^*. \tag{260}$$

Let the eigenvalue decomposition of Σ_0 be given by $\Sigma_0 = U\Lambda U^H = D_0 U\Lambda U^H D_0^*$ with a diagonal matrix Λ. Because D_t is unitary, the eigenvalue decomposition of Σ_t is given by $\Sigma_t = U_t\Lambda U_t^H$ with $U_t = D_t U$. If we use a rank-k approximation $\widetilde{\Sigma}_0$ of Σ_0 and discard all eigenvalues that are smaller than the kth largest eigenvalue, we obtain

$$\widetilde{\Sigma}_0 = D_0\widetilde{U}\widetilde{\Lambda}\widetilde{U}^H D_0^* \quad \text{with} \quad \|\Sigma_0 - \widetilde{\Sigma}_0\| \leq \lambda_{k+1} \tag{261}$$

where λ_{k+1} is the $k+1$-th largest eigenvalue and where $\widetilde{U}$ is a tall matrix with only those eigenvectors corresponding to the k strongest eigenvalues. We immediately obtain a low-rank approximation of Σ_t by

$$\widetilde{\Sigma}_t = D_t\widetilde{U}\widetilde{\Lambda}\widetilde{U}^H D_t^*. \tag{262}$$

Let P_t denote the orthogonal projector onto range$(\widetilde{\Sigma}_t)$ = range$(D_t\widetilde{U})$. By Corollary 35, we can express the projector difference in terms of the orthonormal bases as

$$\begin{aligned} \|P_t - P_s\| &\leq \|(D_t - D_s)\widetilde{U}\| \\ &\leq \|D_t - D_s\| = \max_{0\leq n\leq M-1} |\exp(int) - \exp(ins)| \\ &= \max_{0\leq n\leq M-1} |\exp(in(t-s)) - 1| \leq (M-1)|t-s|. \end{aligned} \tag{263}$$

Consequently, the union of k-dimensional subspaces

$$\mathcal{U}_{\text{3GPP}} = \bigcup_{t\in[-\pi,\pi]} \text{range}\big(\widetilde{\Sigma}_t\big) \tag{264}$$

is Lipschitz continuous with constant $L \leq (M-1)$. We call this set the single-cluster 3GPP union of subspaces (with approximation order k and for M antennas).

Theorem 33. *Let $A \in \mathbb{C}^{m\times M}$ be a random matrix that satisfies the point-wise concentration inequality*

$$\mathbb{P}\left[(1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2\right] \leq 2f(\delta) \tag{265}$$

for all $x \in \mathbb{C}^M$ and where f is log-concave and $f(0) = 1$. Let $\alpha \in [0.5, 1)$ and set

$$\varepsilon = \frac{(1-\alpha)\delta}{\sqrt{8(1+2\alpha)}} \tag{266}$$

and

$$N_0 = 1.5\sqrt[2k+1]{\pi(M-1)} \tag{267}$$

Then, if $\delta \leq 1 - 1/N_0$,

$$\mathbb{P}\left[\sup_{x\in\mathcal{U}_{3GPP}} \big|\|Ax\|^2 - \|x\|^2\big| > \delta\|x\|^2\right] \leq 3(N_0/\varepsilon)^{2k+1} f(\alpha\delta) \tag{268}$$

Proof. The statement follows directly from Theorem 28 if we can show that there is an ε-cover $\mathcal{C}_\varepsilon \subset \mathcal{U}_{\text{3GPP}} \cap \{\|x\| = 1\}$ with cardinality $|\mathcal{C}_\varepsilon| \leq (N_0/\varepsilon)^{2k+1}$. By Theorem 14, there is such an ε-cover of cardinality

$$|\mathcal{C}_\varepsilon| \leq |T_{\varepsilon_T}|(2+\varepsilon_{\mathcal{H}})^{2k}/\varepsilon_{\mathcal{H}}^{2k} \tag{269}$$

if T_{ε_T} is an ε_T-cover of T with respect to the metric $d(s,t) = \|P_s - P_t\|$ and if

$$\varepsilon_{\mathcal{H}}^2 + \varepsilon_T^2 + (1-(1-\varepsilon_T^2)^{1/2})^2 \leq \varepsilon^2. \tag{270}$$

As $\|P_s - P_t\| \leq (M-1)|s-t|$ (cf. (263)), we can find an ε_T cover of the interval $T = [-\pi, \pi]$ with cardinality less than $\pi(M-1)/\varepsilon$. We choose $\varepsilon_{\mathcal{H}} = 8\varepsilon/9$ and $\varepsilon_T = 8\varepsilon/19$ for which (270) holds. We then have for $\varepsilon \leq 1$,

$$|\mathcal{C}_\varepsilon| \leq \left(\frac{9(2+8\varepsilon/9)}{8\varepsilon}\right)^{2k} \times \left(\frac{19\pi(M-1)}{8\varepsilon}\right) \leq (N_0/\varepsilon)^{2k+1} \tag{271}$$

with $N_0 = \frac{19}{8}\sqrt[2k+1]{\pi(M-1)}$ (where we used $\varepsilon \leq 1/8$ for $\alpha \in [0.5, 1)$ and $\delta \leq 1$). □

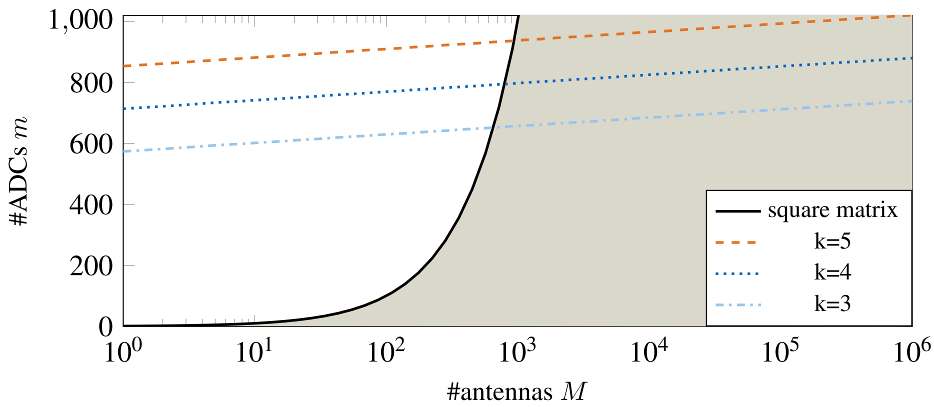

Figure 6: Minimal number of ADCs m such that in at least 99 per cent of the draws, the RIC of A is smaller than $\delta = 0.5$ for the 3GPP manifold with approximation order $k = 3, 4, 5$ (as a function of the number of antennas M). The shaded area indicates non-trivial results (wide matrix). This graph assumes a fixed approximation order k.

As in the previous section, if we use a zero-inflated Steinhaus random matrix A, we obtain that

$$m \geq \inf_{\alpha \in [0.5,1)} \frac{\log(\pi(M-1)) + (2k+1)\log(12/(\delta(1-\alpha))) + \log(3/\xi)}{\alpha\delta - \log(1+\alpha\delta)} \tag{272}$$

is a sufficient condition under which A has the RIP with RIC δ with probability at least $1-\xi$ (we used $19\sqrt{8(1+2\alpha)}/8 \leq 12$). The condition is visualized in Figure 6 for $\delta = 0.5$ and $\xi = 0.01$ and the approximation orders $k = 3, 4, 5$. However, in contrast to the DOA manifold for which the subspace dimension k – the number of paths – is independent of the number of antennas, the approximation order k in the 3GPP model should increase as the number of antennas is increased (and presumably even linearly). This is because the subspace dimension is determined by the standard deviation of the Laplace kernel defining the angular spread of a path (a cluster of micro paths). If we assume that there is an infinite number of these micro paths (and not a fixed number such as the 20 tabulated values in the 3GPP manual [37]), the subspace dimension increases (more antennas can resolve more micro paths). In this case, if k scales linearly with M, then also m must scale linearly with M and non-trivial results are only possible if $k = cM$ with $c < 2\log(12/(\delta(1-\alpha)))/(\alpha\delta - \log(1+\alpha\delta))$, i.e., for very small standard deviations of the Laplace kernel.

XI
Conclusion

We have developed a theory for signal reconstruction in unions of subspaces, which consists of several independent building blocks:

i) A convergence analysis that provides worst-case error bounds for various algorithms under a RIP condition.

ii) A chaining theory for analyzing the tail probabilities of suprema of random processes that can be used to show under what conditions matrices have the RIP. This theory requires that point-wise concentration equations for random matrices and covering-number bounds for the constraint sets are available.

iii) A result that bounds the covering number of Lipschitz unions of subspaces in terms of the Lipschitz constants of their orthogonal projectors.

iv) Upper bounds for the Lipschitz constants of two unions of subspaces that are relevant for channel estimation.

So how many ADCs do we need for channel estimation in systems with many antennas? This question was certainly the main motivation for our focus on quantitative results, especially in the parts concerned with chaining and covering numbers. We are glad to report that it is indeed possible to use the theory to answer the question. However, as can be inferred from the graphs shown in Chapter X, the answer is anything but practically relevant (at least as long as we stick to communication systems with less than a million antennas).
What is practically relevant, however, is that some parts of the theory can serve as an inspiration for practical systems. For example, we showed that zero-inflated Steinhaus random matrices achieve the same concentration results as Gaussian random matrices. About half of the entries of these matrices are zeros and each zero means one wire less between an antenna and an ADC. Similarly, for some of the algorithms we presented, it was only clear from the convergence analysis how they should be modified to properly incorporate the union-of-subspaces constraint. Even though we did not invest a lot of energy into algorithm tuning, the GHTP algorithm, for example, with a Root-MUSIC based projection shows very competitive results in the channel estimation problem stated in Chapter III.

XII
APPENDIX

1—Images of unions of subspaces

In this section, we show that if a union of subspaces $\mathcal{U}$ has the Lipschitz property $d_{\mathrm{F}}(s,t) = \|P_s - P_t\| \le L\|s-t\|$, then its image under a linear operator A also has the property provided that A does not introduce too much distortion. To show this, we first need an auxiliary result. Let P, Q be orthogonal projectors in a Hilbert space with range$(P) = S_P$ and range$(Q) = S_Q$ and let $B_{\mathcal{H}}$ denote the unit ball in $\mathcal{H}$.

Lemma 34. *The projector difference can be calculated as*

$$\|P_s - P_t\| = \max\Big(\sup_{x \in S_s \cap B_{\mathcal{H}}} \inf_{y \in S_t} \|x-y\|, \sup_{x \in S_s \cap B_{\mathcal{H}}} \inf_{y \in S_t} \|x-y\|\Big). \tag{273}$$

Proof. For arbitrary orthogonal projectors P, Q, we have

$$\|P-Q\|^2 = \sup_{x \in B_{\mathcal{H}}} \|Px - QPx + QPx - Qx\|^2 \tag{274}$$

$$= \sup_{x \in B_{\mathcal{H}}} \|Q^{\perp}Px - QP^{\perp}x\|^2 \tag{275}$$

$$\le \sup_{x \in B_{\mathcal{H}}} \|Q^{\perp}P\|^2\|Px\|^2 + \|QP^{\perp}\|^2\|P^{\perp}x\|^2 \tag{276}$$

$$\le \max\big(\|Q^{\perp}P\|^2, \|QP^{\perp}\|^2\big). \tag{277}$$

By definition of the orthogonal projectors, we obtain

$$\|Q^{\perp}P\| = \sup_{x \in B_{\mathcal{H}}} \|Px - QPx\| = \sup_{x \in S_P \cap B_{\mathcal{H}}} \inf_{y \in S_Q} \|x-y\| \tag{278}$$

and a similar equation for $\|P^{\perp}Q\|$ so that

$$\|P-Q\| \le \max\Big(\sup_{x \in S_P \cap B_{\mathcal{H}}} \inf_{y \in S_Q} \|x-y\|, \sup_{x \in S_Q \cap B_{\mathcal{H}}} \inf_{y \in S_P} \|x-y\|\Big). \tag{279}$$

For the equality, we note that

$$\sup_{x \in S_P \cap B_{\mathcal{H}}} \inf_{y \in S_Q} \|x-y\| = \sup_{x \in S_P \cap B_{\mathcal{H}}} \|x - Qx\| \quad \text{(def. of projection)} \tag{280}$$

$$= \sup_{x \in S_P \cap B_{\mathcal{H}}} \|Px - Qx\| \quad (Px = x) \tag{281}$$

$$\le \sup_{x \in B_{\mathcal{H}}} \|Px - Qx\| \tag{282}$$

$$= \|P-Q\| \tag{283}$$

and a similar calculation with the roles of P and Q reversed. □

The following corollary can be useful when orthonormal bases of the two subspaces are known.

Corollary 35. *Let $P = UU^H$ and $Q = VV^H$ be orthogonal projectors with $U, V \in \mathbb{C}^{M\times k}$. Then $\|P - Q\| \le \|U - V\|$.*

Proof. We have

$$\sup_{x\in S_P\cap B_{\mathcal{H}}} \inf_{y\in S_Q} \|x-y\| = \sup_{\alpha\in\mathbb{C}^k, \|\alpha\|=1} \inf_{\beta\in\mathbb{C}^k} \|U\alpha - V\beta\| \tag{284}$$

$$\le \sup_{\alpha\in\mathbb{C}^k, \|\alpha\|=1} \|U\alpha - V\alpha\| = \|U - V\| \tag{285}$$

and, similarly

$$\sup_{x\in S_Q\cap B_{\mathcal{H}}} \inf_{y\in S_P} \|x-y\| \le \sup_{\alpha\in\mathbb{C}^k, \|\alpha\|=1} \|V\alpha - U\alpha\| = \|U - V\|. \tag{286}$$

By Lemma 34, it follows that $\|P - Q\| \le \|U - V\|$. □

Lemma 36 (Image of a union of subspaces)**.** *Let $\mathcal{U} = \cup_{t\in T} S_t$ be a union of subspaces and let P_t denote the orthogonal projector onto S_t for $t \in T$ and let*

$$d_F(s,t) = \|P_s - P_t\| \tag{287}$$

denote the Finsler metric in T with respect to $\mathcal{U}$. If $A\colon \mathcal{H} \to \mathcal{H}'$ is a linear map with

$$\|Ax\|^2 \ge c_{min}\|x\|^2 \quad \text{and} \quad \|A(x-y)\|^2 \le c_{max}\|x-y\|^2 \tag{288}$$

for all $x, y \in \mathcal{U}$, then

$$d'_F(s,t) = \|Q_s - Q_t\| \le \frac{c_{max}}{c_{min}} d_F(s,t) \tag{289}$$

where Q_t denotes the orthogonal projector onto range(AP_t) *for $t \in T$.*

Proof. Let $y \in$ range $AP_t \cap B_{\mathcal{H}'}$. Then $y = Ax$ for some $x \in S_t$ with $\|x\| \le c_{\min}^{-1}$. Choose $y' = AP_s x$. Then

$$\|y - y'\| = \|A(P_t x - P_s x)\| \le c_{\max}\|(P_t - P_s)x\|$$
$$\le c_{\max}\|x\|\|P_t - P_s\| \le \frac{c_{\max}}{c_{\min}} d_{\mathrm{F}}(s,t). \tag{290}$$

Because y was arbitrary in range$(AP_t) \cap B_{\mathcal{H}'}$, we obtain

$$\sup_{\substack{y\in\text{range}(AP_t)\\ \|y\|\le 1}} \inf_{z\in\text{range}(AP_s)} \|y-z\| \le \sup_{\substack{y\in\text{range}(AP_t)\\ \|y\|\le 1}} \frac{c_{\max}}{c_{\min}} d_{\mathrm{F}}(s,t) = \frac{c_{\max}}{c_{\min}} d_{\mathrm{F}}(s,t). \tag{291}$$

A similar calculation holds for the roles of s and t exchanged. Thus, by Lemma 34, we have $d'_{\mathrm{F}}(s,t) \le d_{\mathrm{F}}(s,t)c_{\max}/c_{\min}$. □

2—Projected gradient descent

The proof is an adaptation of the proof for IHT in [75]. Let $x_* \in S_*$ be an arbitrary element in $\mathcal{U}$ and define the residual $r_n = x_n - x_*$ with $x_n \in S_n$. We use the following notation: $P_{*.n}, P_{*.n.n-1}, P_{*.\circ}$ denote the orthogonal projectors onto the subspaces span$\{S_*, S_n\}$, span$\{S_*, S_n, S_{n-1}\}$, and span$\{S_*, S_\circ\}$, respectively ($S_\circ \subset \mathcal{U}$ is also a subspace). We use $\delta_2 = \delta(A, \mathcal{U}^2)$ and $\delta = \delta_3 = \delta(A, \mathcal{U}^3)$ for the second and third order restricted isometry constants. By the triangle inequality, we can decompose the residual $r_n = P_{*.n} r_n$ into

$$\|P_{*.n}\, r_n\| \leq \|P_{*.n}(x_* - z_n)\| + \|P_{*.n}(x_n - z_n)\|\,. \tag{292}$$

If we use $y = Ax_* + e'$ with $e' = A(x - x_*) + e$ and expand z_n, we obtain

$$\begin{aligned} P_{*.n}(x_* - z_n) &= P_{*.n}\left(x_* - (x_{n-1} + A^*(y - Ax_{n-1}))\right) & (293)\\ &= P_{*.n}(I - A^*A)(x_* - x_{n-1}) - P_{*.n}A^*e' & (294)\end{aligned}$$

Next, we use that $r_{n-1} = x_* - x_{n-1} \in S_{*.n-1}$ and span$\{S_{*.n}, S_{*.n-1}\} \subset \mathcal{U}^3$ and apply Lemmas 4 and 6 to obtain

$$\|P_{*.n}(x_* - z_n)\| \leq \delta\|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|. \tag{295}$$

For the second term in (292), we use the sub-optimality of the approximate projection $x_n = \mathcal{Q}(z_n)$. For some subspace $S_\circ \subset \mathcal{U}$, we have

$$\begin{aligned} \|x_n - z_n\|^2 = \|\mathcal{Q}(z_n) - z_n\|^2 &\leq \|P_* z_n - z_n\|^2 + \varepsilon^2\|P_{*.\circ} z_n\|^2 & (296)\\ &\leq \|x_* - z_n\|^2 + \varepsilon^2\|P_{*.\circ} z_n\|^2 & (297)\end{aligned}$$

where second inequality follows because $P_* z_n$ is a better approximation of z_n in S_* than x_*. By subtracting the terms $\|(I - P_{*.n})(x_n - z_n)\|^2 = \|(I - P_{*.n})(x_* - z_n)\|^2$ from both sides, we obtain

$$\begin{aligned} \|P_{*.n}(x_n - z_n)\|^2 &\leq \|P_{*.n}(x_* - z_n)\|^2 + \varepsilon^2\|P_{*.\circ} z_n\|^2 & (298)\\ &\leq \|P_{*.n}(x_* - z_n)\|^2 + \varepsilon^2\left(\|P_{*.\circ}(x_* - z_n)\| + \|x_*\|\right)^2 & (299)\end{aligned}$$

where the second inequality follows from the triangle inequality. By applying Lemma 4 as in (295), we get (this also works with $P_{*.n}$ replaced by $P_{*.\circ}$)

$$\begin{aligned} &\|P_{*.n}(x_n - z_n)\|^2 & (300)\\ &\leq \left(\delta\|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right)^2 + \varepsilon^2\left(\delta\|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\| + \|x_*\|\right)^2 & (301)\\ &\leq \left((\gamma_L + \gamma_R\varepsilon)\left(\delta\|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \gamma_R\varepsilon\|x_*\|\right)^2 & (302)\end{aligned}$$

where the second inequality follows from Lemma 42 with $\{\gamma_L, \gamma_R\} = \{1, c_\gamma\}$ and $c_\gamma = 1/(1+\sqrt{2})$. While we do not know which of γ_L and γ_R is smaller than one, we know that $\gamma_L + \gamma_R \varepsilon \leq 1 + c_\gamma \varepsilon$ (because $\varepsilon \leq 1$) so that by taking the square root and using $\gamma_R \leq 1$, we get

$$\|P_{*.n}(x_n - z_n)\| \leq \left(1 + c_\gamma \varepsilon\right) \left(\delta \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \varepsilon \|x_*\|. \quad (303)$$

We can plug this result together with (295) into the triangle inequality (292) to obtain

$$\|r_n\| \leq \left(2 + c_\gamma \varepsilon\right) \left(\delta \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \varepsilon \|x_*\|. \quad (304)$$

Thus, we get

$$\|r_n\| \leq c_1 \|r_{n-1}\| + c_2 \|e'\| + \varepsilon \|x_*\| \quad (305)$$

with

$$c_1 = \left(2 + \varepsilon/(1+\sqrt{2})\right)\delta \quad \text{and} \quad c_2 = \left(2 + \varepsilon/(1+\sqrt{2})\right)\sqrt{1+\delta_2}. \quad (306)$$

For $x_0 = 0$, this is recursively found to yield

$$\|r_n\| \leq c_1^n \|x_*\| + \frac{1-c_1^n}{1-c_1}\left(c_2\|e'\| + \varepsilon\|x_*\|\right). \quad (307)$$

We arrive at (95) by noting that

$$\|x_n - x\| \leq \|x_n - x_*\| + \|x - x_*\| = \|r_n\| + \|x - x_*\| \quad (308)$$

and expanding

$$\|e'\| = \|e + A(x - x_*)\| \leq \|e\| + \|A(x - x_*)\|. \quad (309)$$

3—Generalized hard thresholding pursuit

This proof is an adaptation of the proof in [52, Theorem 3.8]. Let $x_* \in S_*$ be an arbitrary element in $\mathcal{U}$ and define the residual $r_n = x_n - x_*$ with $x_n \in S_n$. We use the following notation: $P_{*.n}, P_{*.n.n-1}, P_{*.\circ}$ denote the orthogonal projectors onto the subspaces $S_{*.n} = \text{span}\{S_*, S_n\}$, $S_{*.n.n-1} = \text{span}\{S_*, S_n, S_{n-1}\}$, and $S_{*.\circ} = \text{span}\{S_*, S_\circ\}$, respectively ($S_\circ \subset \mathcal{U}$ is also a subspace). Next, let $P_{*|*.n} = P_{*.n}P_n^\perp$ and $P_{n|*.n} = P_{*.n}P_*^\perp$ denote the orthogonal projectors onto the orthogonal complements of S_n and S_* in $S_{*.n}$, respectively. Finally, $\delta_2 = \delta(A, \mathcal{U}^2)$ and $\delta_3 = \delta(A, \mathcal{U}^3)$ are the second- and third-order restricted isometry constants. Also, we write $e' = e + A(x - x_*)$ so that $y = Ax_* + e'$ with $x_* \in \mathcal{U}$.
Since $r_n \in S_{*.n}$, we can write $r_n = P_{*|*.n}r_n + P_n r_n$ and use the orthogonal decomposition

$$\|r_n\|^2 = \|P_n r_n\|^2 + \|P_{*|*.n} r_n\|^2. \quad (310)$$

3.1 First bound

Because x_n minimizes $\|y - Az\|^2$ with $z \in S_n$, the orthogonality principle states that the error $y - Ax_n$ is orthogonal to range(AP_n), i.e.,

$$\langle Ax_n - y, Az\rangle = 0 \quad \text{for } z \in S_n. \tag{311}$$

Since $y = Ax_* + e'$, we can rewrite (311) as

$$\langle A^*A(\underbrace{x_n - x_*}_{=r_n}), z\rangle = \langle A^*e', z\rangle \quad \text{for } z \in S_n. \tag{312}$$

If we use $z = P_n r_n \in S_n$, we obtain

$$\|P_n r_n\|^2 = \langle r_n, P_n r_n\rangle - \langle A^*Ar_n, P_n r_n\rangle + \langle A^*Ar_n, P_n r_n\rangle \tag{313}$$
$$= \langle (I - A^*A)r_n, P_n r_n\rangle + \langle A^*e', P_n r_n\rangle \tag{314}$$
$$= \langle P_n(I - A^*A)P_{*.n} r_n, P_n r_n\rangle + \langle P_n A^*e', P_n r_n\rangle \tag{315}$$
$$\leq \delta_2\|r_n\|\|P_n r_n\| + \sqrt{1+\delta_1}\|e'\|\|P_n r_n\| \tag{316}$$

where, in the last step, we applied Lemmas 4 and 6 from Chapter VI to handle $P_n(I - A^*A)P_{*.n}$ (noting that span$\{S_n, S_{*.n}\} = S_{*.n} \subset \mathcal{U}^2$) and $P_n A^*e'$. This can be simplified to

$$\|P_n r_n\| \leq \delta_2\|r_n\| + \sqrt{1+\delta_1}\|e'\|. \tag{317}$$

3.2 Second bound

Bounding the term $\|P_{*|*.n} r_n\|$ requires some work. First, we use that $P_{*|*.n} x_n = 0$ so that $P_{*|*.n} r_n = P_{*|*.n} x_*$. The triangle inequality thus yields

$$\|P_{*|*.n} r_n\| = \|P_{*|*.n} x_*\| \leq \|P_{*|*.n}(x_* - z_n)\| + \|P_{*|*.n} z_n\|. \tag{318}$$

The first term is easy (see also (293) in Appendix 2): If P is any of the projectors $P_{*|*.n}, P_{n|*.n}, P_{*.\circ}$, we get

$$\|P(z_n - x_*)\| = \|P(I - A^*A)r_{n-1} + PA^*e'\| \tag{319}$$
$$\leq \delta_3\|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\| \tag{320}$$

as span$\{\text{range}(P), r_{n-1}\} \subset \mathcal{U}^3$ in all cases.
For the second term, we note that range($P_{*|*.n} + P_n$) = range($P_{n|*.n} + P_*$) so that

$$\|P_{*|*.n} z_n\|^2 + \|P_n z_n\|^2 = \|P_{n|*.n} z_n\|^2 + \|P_* z_n\|^2 \tag{321}$$
$$\Leftrightarrow \|z_n - P_* z_n\|^2 + \|P_{*|*.n} z_n\|^2 = \|z_n - P_n z_n\|^2 + \|P_{n|*.n} z_n\|^2 \tag{322}$$

as $P_{*|*.n} \perp P_n$ and $P_{n|*.n} \perp P_*$. By the approximate optimality of the subspace S_n, there is a subspace $S_\circ$ with

$$\|z_n - P_n z_n\|^2 \leq \|z_n - P_* z_n\|^2 + \varepsilon^2\|P_{*.\circ} z_n\|^2. \tag{323}$$

Combined with (322), this inequality yields

$$\|P_{*|*.n} z_n\|^2 \leq \varepsilon^2 \|P_{*.\circ} z_n\|^2 + \|P_{n|*.n} z_n\|^2. \tag{324}$$

Then, using $\|P_{*.\circ} z_n\| \leq \|P_{*.\circ}(x_* - z_n)\| + \|x_*\|$ and $P_{n|*.n} z_n = P_{n|*.n}(z_n - x_*)$ and applying (320), we obtain

$$\|P_{*|*.n} z_n\|^2 \leq \varepsilon^2 \left(\|x_*\| + a\right)^2 + a^2 \tag{325}$$

with $a = \delta_3 \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|$. We then use Lemma 42 to get the complicated-looking

$$\|P_{*|*.n} z_n\|^2 \leq a^2 + \varepsilon^2 \left(a + \|x_*\|\right)^2 \leq \left(\left(\gamma_L + \gamma_R \varepsilon\right) a + \gamma_R \varepsilon \|x_*\|\right)^2 \tag{326}$$

with $\{\gamma_L, \gamma_R\} \in \{1, c_\gamma\}$ and $c_\gamma = 1/(1+\sqrt{2})$. While we do not know which of γ_L and γ_R is smaller than one, we know that $\gamma_L + \gamma_R \varepsilon \leq 1 + c_\gamma \varepsilon$ (because $\varepsilon \leq 1$) so that by taking the square root and using $\gamma_R \leq 1$, we get

$$\|P_{*|*.n} z_n\| \leq \left(1 + c_\gamma \varepsilon\right) \left(\delta_3 \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \varepsilon \|x_*\|. \tag{327}$$

Finally, it follows from (318) and (320) that

$$\|P_{*|*.n} r_n\| \leq (2 + c_\gamma \varepsilon) \left(\delta_3 \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \varepsilon \|x_*\|. \tag{328}$$

3.3 Synthesis

If we insert (317) and (328) into (310), we obtain another sum of squares:

$$\|r_n\|^2 \leq \left(\delta_2 \|r_n\| + \sqrt{1+\delta_2}\|e'\|\right)^2 \tag{329}$$

$$+ \left((2 + c_\gamma \varepsilon) \left(\delta_3 \|r_{n-1}\| + \sqrt{1+\delta_2}\|e'\|\right) + \varepsilon \|x_*\|\right)^2. \tag{330}$$

Let $\xi = 2 + c_\gamma \varepsilon$ and $\phi = \sqrt{1+\delta_2}\|e'\|$ so that

$$\|r_n\|^2 \leq \left(\gamma'_L \left(\delta_2 \|r_n\| + \phi\right) + \gamma'_R \left(\xi \left(\delta_3 \|r_{n-1}\| + \phi\right) + \varepsilon \|x_*\|\right)\right)^2 \tag{331}$$

with $\{\gamma'_L, \gamma'_R\} = \{1, c_\gamma\}$ (see Lemma 42, we do not know which of γ'_l and γ'_R is smaller than one) so that

$$\|r_n\| \leq \frac{(\gamma'_L + \gamma'_R \xi)\phi + \gamma'_R \left(\xi \delta_3 \|r_{n-1}\| + \varepsilon \|x_*\|\right)}{1 - \gamma'_L \delta_2}. \tag{332}$$

For $\delta_2 \leq 1/3$ (which is satisfied as $\delta_2 \leq \delta_3 < 1/3$) and $\xi \geq 2$, we have[1]

$$\frac{1 + c_\gamma \xi}{1 - \delta_2} \leq \frac{c_\gamma + \xi}{1 - c_\gamma \delta_2} \quad \Rightarrow \quad \frac{\gamma'_L + \gamma'_R \xi}{1 - \gamma'_L \delta_2} \leq \frac{c_\gamma + \xi}{1 - c_\gamma \delta_2} \tag{333}$$

[1] These formulas can be easily verified numerically.

and

$$\frac{c_\gamma}{1-\delta_2} \le \frac{1}{1-c_\gamma\delta_2} \quad \Rightarrow \quad \frac{\gamma'_R}{1-\gamma'_L\delta_2} \le \frac{1}{1-c_\gamma\delta_2}. \tag{334}$$

If we insert these upper bounds into (332), we get

$$\|r_n\| \le \frac{(c_\gamma+\xi)\phi + \xi\delta_3\|r_{n-1}\| + \varepsilon\|x_*\|}{1-c_\gamma\delta_2} = c_1\|r^{n-1}\| + c_2\|e'\| + \varepsilon c_3\|x_*\| \tag{335}$$

with

$$c_1 = \frac{\left(2+c_\gamma\varepsilon\right)\delta_3}{1-c_\gamma\delta_2}, \quad c_2 = \frac{\left(2+c_\gamma\varepsilon+c_\gamma\right)\sqrt{1+\delta_2}}{1-c_\gamma\delta_2}, \quad c_3 = \frac{1}{1-c_\gamma\delta_2} \tag{336}$$

which is the desired statement of the theorem.

4—Orthogonal matching pursuit

Let $\Lambda = \{i_1^*, \dots, i_k^*\}$ be such that $x_* \in S_* = \oplus_{i\in\Lambda} V_i^*$ with $V_i^* \subset \mathcal{U}_i$, i.e., Λ contains the true indices and V_i^* are the true subspaces. Similarly, let $\Lambda_n = \{i_1, \dots, i_n\}$ denote the sequence of indices found by the OMP algorithm and let $x_n \in S_n = \oplus_{i\in\Lambda_n} V_i^n$ with $V_i^n \subset \mathcal{U}_i$. We show by recursion that for $n = 1, \dots, k$, we have $\Lambda_n \subset \Lambda$, i.e., the OMP algorithm always selects correct indices.
For the first index, this is particularly simple. If $j \notin \Lambda$, we have

$$\|\mathcal{P}_{\mathcal{U}_j}(A^*Ax_*)\| = \sup_{S_j\subset\mathcal{U}_j} \|P_{S_j}A^*AP_*x_*\| \le \delta\|x_*\| \tag{337}$$

by Lemma 5 and as $S_* and S_j$ are orthogonal. Here and in the following, we use $\delta = \delta(A, \mathcal{W}_{k+1})$ with the union of subspaces

$$\mathcal{W}_{k+1} = \bigcup_{i_1,\dots,i_{k+1}\in\{1,\dots,P\}} \bigoplus_{j=1,\dots,k+1} \mathcal{U}_{i_j}^2 \tag{338}$$

where the union is over indices with $i_j \ne i_{j'}$ for $j \ne j'$ and $\mathcal{U}_{i_j}^2 = \{x : x = x_1 + x_2, x_1, x_2 \in \mathcal{U}_{i_j}\}$. For indices $i \in \Lambda$, we have

$$k \max_{i\in\Lambda} \|\mathcal{P}_{\mathcal{U}_i}\left(A^*Ax_*\right)\|^2 \ge \sum_{i\in\Lambda} \|P_{V_i^*}A^*Ax_*\|^2 = \|P_{S_*}A^*Ax_*\|^2 \ge (1-\delta)^2\|x_*\|^2. \tag{339}$$

Consequently, if $(1-\delta)/\sqrt{k} > \delta \Leftrightarrow \delta < 1/(1+\sqrt{k})$, then a correct index is chosen at the first iteration.
For the subsequent iterations, let us introduce the subspaces

$$S_{*.n} = \oplus_{i\in\Lambda_n} \operatorname{span}\{V_i^*, V_i^n\} \quad \text{and} \quad S_{*\backslash n} = \oplus_{i\in\Lambda\backslash\Lambda_n} V_i^* \tag{340}$$

so that $S_{*.n} \oplus S_{*\backslash n} = \operatorname{span}\{S_n, S_*\}$ with corresponding orthogonal projectors $P_{*.n}$ and $P_{*\backslash n}$. The subspaces are chosen such that

$$x_* - x_n = P_{*.n}(x_* - x_n) + P_{*\backslash n}x_*, \tag{341}$$

$$x_* = P_{*.n}x_* + P_{*\backslash n}x_*. \tag{342}$$

First, we note that because of step 2ii of Alg. 3, the intermediate solution x_n is such that

$$\|y - Ax_n\|^2 \leq \|y - AP_{*.n}x_*\|^2 = \|Ax_* - AP_{*.n}x_*\|^2 = \|AP_{*\setminus n}x_*\|^2 \quad (343)$$

where the inequality follows, because $P_{*.n}x_*$ is a valid candidate point in the minimization.[2] We use this property to bound the error in subspaces of already detected unions of subspaces, $P_{*.n}(x_* - x_n)$, by the residual error, $P_{*\setminus n}x_*$, in subspaces with indices that are not yet in Λ_n. We use the restricted isometry property of A together with (343) in

$$(1-\delta)\|x_* - x_n\|^2 \leq \|A(x_* - x_n)\|^2 \quad (344)$$

$$= \|y - Ax_n\|^2 \leq \|AP_{*\setminus n}x_*\|^2 \leq (1+\delta)\|P_{*\setminus n}x_*\|^2. \quad (345)$$

Consequently

$$\|P_{*.n}(x_* - x_n)\|^2 = \|x_* - x_n\|^2 - \|P_{*\setminus n}x_*\|^2 \quad (346)$$

$$\leq \left(\frac{1+\delta}{1-\delta} - 1\right)\|P_{*\setminus n}x_*\|^2 = \frac{2\delta}{1-\delta}\|P_{*\setminus n}x_*\|^2. \quad (347)$$

Now, let $S_j \subset \mathcal{U}_j$ be an arbitrary subspace with $j \notin \Lambda$. By assumption, we have $x_* - x_n \perp S_j$ and $\text{span}\{x_* - x_n, S_j\} \subset \mathcal{W}_{k+1}$ so that

$$\|\mathcal{P}_{\mathcal{U}_j}\big(A^*A(x_* - x_n)\big)\| = \sup_{S_j \subset \mathcal{U}_j} \|P_{S_j}A^*A(x_* - x_n)\| \quad (348)$$

$$\leq \delta\|x_* - x_n\| \leq \delta\sqrt{\frac{1+\delta}{1-\delta}}\|P_{*\setminus n}x_*\| \quad (349)$$

yields an upper bound for the bad indices. The first inequality is due to Lemma 5 and the second one follows from (345).
On the other hand, for correct indices, we have

$$(k-n)\max_{i \in \Lambda\setminus\Lambda_n} \|\mathcal{P}_{\mathcal{U}_i}\big(A^*A(x_* - x_n)\big)\|^2 \geq \sum_{i \in \Lambda\setminus\Lambda_n} \|\mathcal{P}_{\mathcal{U}_i}\big(A^*A(x_* - x_n)\big)\|^2 \quad (350)$$

$$\geq \sum_{i \in \Lambda\setminus\Lambda_n} \|P_{V_i^*}A^*A(x_* - x_n)\|^2 \quad (351)$$

$$= \|P_{*\setminus n}A^*A(x_* - x_n)\|^2. \quad (352)$$

We use the decomposition (341), the inverse triangle inequality, and Lemma 5 to obtain

$$\|P_{*\setminus n}A^*A(x_* - x_n)\| \geq \|P_{*\setminus n}A^*AP_{*\setminus n}x_*\| - \|P_{*\setminus n}A^*AP_{*.n}(x_* - x_n)\| \quad (353)$$

$$\geq (1-\delta)\|P_{*\setminus n}x_*\| - \delta\|P_{*.n}(x_* - x_n)\| \quad (354)$$

$$\geq (1-\delta)\|P_{*\setminus n}x_*\| - \delta\sqrt{\frac{2\delta}{1-\delta}}\|P_{*\setminus n}x_*\| \quad (355)$$

[2]The projector $P_{*.n}$ only zeros some blocks of x_*: $P_{*.n}x_* = P_{\oplus_{i\in\Lambda_n}V_i^*}x_* \in \oplus_{i\in\Lambda_n}\mathcal{U}_i$.

where the last inequality follows from (347). By combining (355) and (352) with (349), we see that we always select a correct index if

$$\sqrt{k-n}\sqrt{\frac{1+\delta}{1-\delta}} < \delta^{-1} - 1 - \sqrt{\frac{2\delta}{1-\delta}} \tag{356}$$

for $n = 1, \ldots, k-1$. This bound holds for all $n \geq 1$ if it holds for $n = 1$. In the remaining part of the proof, we show that this slightly complicated and implicit bound for δ is implied by the stronger bound $\delta \leq (\sqrt{6} + \sqrt{k-1})^{-1}$ (a comparison of the bound for small values of k is shown in Figure 7). Thus, let $n = 1$ and $\delta^{-1} = 1 + u^2$ and re-write (356) as

$$\sqrt{k-1}\sqrt{\frac{u^2+2}{u^2}} \leq u^2 - \sqrt{\frac{2}{u^2}} \quad \Leftrightarrow \quad \xi\sqrt{1+2/u^2} \leq u^2 - \sqrt{2}/u \tag{357}$$

with $\xi = \sqrt{k-1}$. As $\sqrt{1+2/u^2} \leq 1 + 1/u^2$ (we use the concavity of $\sqrt{\cdot}$ and a Taylor approximation), this condition is met if

$$\xi + \xi/u^2 \leq u^2 - \sqrt{2}/u. \tag{358}$$

Let $u^2 = \xi + x$, then we write (358) as

$$x \geq \xi/u^2 + \sqrt{2}/u = \frac{\xi}{\xi+x} + \sqrt{\frac{2}{\xi+x}}. \tag{359}$$

For $k = 2, \ldots, 9$, one can numerically verify that this inequality is met if $x \geq \sqrt{6} - 1$. For $k \geq 9$, the derivative of the right-hand side of (359) with respect to ξ is negative for $x \geq \sqrt{6} - 1$ so that (359) also holds for such k.[3] As the resulting condition $\delta \leq (\sqrt{6} + \sqrt{k-1})^{-1}$ is always stronger than the condition $\delta \leq (1 + \sqrt{k})$, which was necessary for the first iteration of the OMP algorithm, we find that the OMP algorithm successfully recovers x_* if

$$\delta \leq \frac{1}{\sqrt{6} + \sqrt{k-1}}. \tag{360}$$

5—Reduced-complexity OMP

Let $\Lambda = \{i_1^*, \ldots, i_k^*\}$ be such that $x_* \in S_* = \oplus_{i \in \Lambda} V_i^*$ with $V_i^* \subset \mathcal{U}_i$, i.e., Λ contains the true indices and V_i^* are the true subspaces. Similarly, let $\Lambda_n =$

[3] We have

$$\frac{d}{d\xi}\left(\frac{\xi}{\xi+x} + \sqrt{\frac{2}{\xi+x}}\right) = \frac{x}{(\xi+x)^2} - \frac{1}{\sqrt{2(\xi+x)^3}} \propto \frac{\sqrt{2}x}{\sqrt{\xi+x}} - 1.$$

If $\xi \geq 2x^2 - x$, the derivative is negative. If $x \geq \sqrt{6} - 1$, this condition is met if $\xi \geq \sqrt{9-1}$, i.e., if $k \geq 9$.

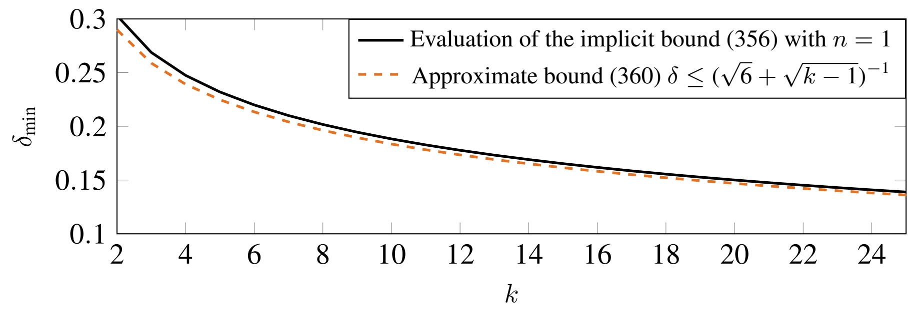

Figure 7: Comparison of the implicit bound for δ given by (356) with the explicit, but approximate, bound (360).

$\{i_1, \dots, i_n\}$ denote the sequence of indices found by the reduced-complexity OMP algorithm and let $x_n \in S_n = \oplus_{i \in \Lambda_n} V_i^n$ with $V_i^n \subset \mathcal{U}_i$. We show by recursion that for $n = 1, \dots, k$, we have $\Lambda_n \subset \Lambda$, i.e., the reduced-complexity OMP algorithm always selects correct indices.

For the first index, this is particularly simple: if $j \notin \Lambda$, we have

$$\|\mathcal{P}_{\mathcal{U}_j}(A^* A x_*)\| = \sup_{S_j \subset \mathcal{U}_j} \|P_{S_j} A^* A P_* x_*\| \le \delta \|x_*\| \tag{361}$$

by Lemma 5, where here and in the following, we use $\delta = \delta(A, \mathcal{W}_{k+1})$ and $\delta' = \delta(A, \mathcal{W}'_k)$ with the unions of subspaces

$$\mathcal{W}_{k+1} = \bigcup_{i_1, \dots, i_{k+1} \in \{1, \dots, P\}, \text{ all } i_j \text{ distinct}} \bigoplus_{j=1,\dots,k+1} \mathcal{U}_{i_j}^2 \tag{362}$$

$$\mathcal{W}'_k = \bigcup_{i_1, \dots, i_k \in \{1, \dots, P\} \text{ all } i_j \text{ distinct}} \bigoplus_{j=1,\dots,k} \operatorname{span}\{\mathcal{U}_{i_j}\}. \tag{363}$$

For indices $i \in \Lambda$, we have

$$k \max_{i \in \Lambda} \|\mathcal{P}_{\mathcal{U}_i}\big(A^* A x_*\big)\|^2 \ge \sum_{i \in \Lambda} \|P_{V_i^*} A^* A x_*\|^2 \tag{364}$$

$$= \|P_{S_*} A^* A x_*\|^2 \ge (1-\delta)^2 \|x_*\|^2. \tag{365}$$

Consequently, if $(1-\delta)/\sqrt{k} > \delta \Leftrightarrow \delta < 1/(1+\sqrt{k})$, then a correct index is chosen at the first iteration.

For the subsequent iterations, things become substantially more complicated. Let

$$S_{*.n} = \oplus_{i \in \Lambda_n} \operatorname{span}\{\mathcal{U}_i\} \quad \text{and} \quad S_{*\setminus n} = \oplus_{i \in \Lambda \setminus \Lambda_n} V_i^* \tag{366}$$

with corresponding orthogonal projectors $P_{*.n}$ and $P_{*\setminus n}$. Note that the definition of $S_{*.n}$ differs from the proof for the OMP algorithm (because of the span$\{\cdot\}$ operation). The subspaces are chosen such that

$$x_* - x_n = P_{*.n}(x_* - x_n) + P_{*\setminus n} x_*, \tag{367}$$

$$x_* = P_{*.n} x_* + P_{*\setminus n} x_*. \tag{368}$$

First, let $S_j \subset \mathcal{U}_j$ be an arbitrary subspace with $j \notin \Lambda$. By assumption, we have $x_* - x_n \perp S_j$ and $\operatorname{span}\{x_* - x_n, S_j\} \subset \mathcal{W}_{k+1}$ so that

$$\|\mathcal{P}_{\mathcal{U}_j}\big(A^*A(x_* - x_n)\big)\| = \sup_{S_j \subset \mathcal{U}_j} \|P_{S_j} A^*A(x_* - x_n)\| \leq \delta \|x_* - x_n\| \tag{369}$$

yields an upper bound for the incorrect indices. The inequality follows from Lemma 5.
On the other hand, for correct indices, we have

$$(k-n) \max_{i \in \Lambda \setminus \Lambda_n} \|\mathcal{P}_{\mathcal{U}_i}\big(A^*A(x_* - x_n)\big)\|^2 \geq \sum_{i \in \Lambda \setminus \Lambda_n} \|\mathcal{P}_{\mathcal{U}_i}\big(A^*A(x_* - x_n)\big)\|^2 \tag{370}$$

$$\geq \sum_{i \in \Lambda \setminus \Lambda_n} \|P_{V_i^*} A^*A(x_* - x_n)\|^2 \tag{371}$$

$$= \|P_{*\setminus n} A^*A(x_* - x_n)\|^2. \tag{372}$$

If we use the decomposition (367) and the inverse triangle inequality, we obtain

$$\|P_{*\setminus n} A^*A(x_* - x_n)\| \geq \|P_{*\setminus n} A^*A P_{*\setminus n} x_*\| - \|P_{*\setminus n} A^*A P_{*.n}(x_* - x_n)\| \tag{373}$$

$$\geq (1-\delta)\|P_{*\setminus n} x_*\| - \delta \|P_{*.n}(x_* - x_n)\|. \tag{374}$$

If we combine this inequality with (372), we obtain

$$\max_{i \in \Lambda \setminus \Lambda_n} \|\mathcal{P}_{\mathcal{U}_i}\big(A^*A(x_* - x_n)\big)\| \geq \frac{(1-\delta)\|P_{*\setminus n} x_*\| - \delta \|P_{*.n}(x_* - x_n)\|}{\sqrt{k-n}} \tag{375}$$

Thus, if we compare (369) and (375), we see that we always select correct indices if

$$(1-\delta)\|P_{*\setminus n} x_*\| > \delta \|P_{*.n}(x_* - x_n)\| + \delta\sqrt{k-n}\|x_* - x_n\|. \tag{376}$$

Below, we show the inequality

$$\|P_{*.n}(x_* - x_n)\| \leq u \|P_{*\setminus n} x_*\| \text{ with } u = 2\sqrt{\frac{\delta + \delta'}{1-\delta'}} \tag{377}$$

If we use this bound in (376) and use the orthogonal decomposition

$$\|x_* - x_n\|^2 = \|P_{*.n}(x_* - x_n)\|^2 + \|P_{*\setminus n} x_*\|^2 \leq (1+u^2)\|P_{*\setminus n} x_*\|^2 \tag{378}$$

we can simplify the condition (376) to

$$1 - \delta > u\delta + \delta\sqrt{k-n}\sqrt{1+u^2} \quad \Leftrightarrow \quad \delta < \frac{1}{1 + u + \sqrt{(k-n)(1+u^2)}} \tag{379}$$

and this condition holds for all $n \geq 1$ if it holds for $n = 1$.

To show (377), we note that $x_* - z_n \in \mathcal{W}'_k$ (z_n is the not-yet projected solution of the LS problem), that $P_{*.n}x_* \in S_{*.n}$, and that z_n minimizes the expression $\|y - Az\|^2$ over all $z \in S_{*.n}$ (step 2ii of Alg. 4) so that

$$
\begin{aligned}
(1-\delta')\|x_* - z_n\|^2 &\leq \|A(x_* - z_n)\|^2 && (380)\\
&= \|y - AP_{*.n}z_n\|^2 && (381)\\
&\leq \|y - AP_{*.n}x_*\|^2 && (382)\\
&= \|AP_{*\setminus n}x_*\|^2 \leq (1+\delta)\|P_{*\setminus n}x_*\|^2. && (383)
\end{aligned}
$$

Finally, by the triangle inequality and because x_n is a better approximation of z_n as x_* (step 2iii of Alg. 4)

$$
\begin{aligned}
\|P_{*.n}(x_* - x_n)\| &\leq \|P_{*.n}(x_* - z_n)\| + \|P_{*.n}(x_n - z_n)\| && (384)\\
&\leq 2\|P_{*.n}(x_* - z_n)\| && (385)\\
&= 2\sqrt{\|x_* - z_n\|^2 - \|P_{*\setminus n}x_*\|^2} && (386)\\
&\leq 2\sqrt{(1+\delta)/(1-\delta') - 1}\|P_{*\setminus n}x_*\| && (387)\\
&= 2\sqrt{(\delta+\delta')/(1-\delta')}\|P_{*\setminus n}x_*\| && (388)
\end{aligned}
$$

which shows (377).

6—On measurability

Lemma 37. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, (T, d) a separable metric space, and (E, d) a metric space. Let $X = (X_t)_{t\in T}$ denote a random process on Ω with index set (T, d) and values in (E, d) that is almost surely continuous. Let $f\colon E \times E \to \mathbb{R}$ be a continuous function, $T_0 \subset T$ a countable subset, and $\Gamma \in \mathcal{A}$ satisfying $\mathbb{P}[\Gamma] = 1$ and $X_t(\omega)$ is continuous for $\omega \in \Gamma$. Then, the event*

$$
A = \left\{ \sup_{t\in T} \inf_{s\in T_0} f(X_t, X_s) > u \right\} \cap \Gamma \qquad (389)
$$

is measurable.

Proof. First, we note that for $s, t \in T$ and open subsets $W \subset E \times E$, the sets

$$
\{\omega \in \Omega : (X_t(\omega), X_s(\omega)) \in W\} \qquad (390)
$$

are clearly measurable as they generate the σ-algebra on the set of functions from T to E (the usual cylinder sets). Because f is continuous, also the sets

$$
\{\omega \in \Omega : f(X_t(\omega), X_s(\omega)) \in U\} \qquad (391)
$$

are measurable if U is open. Let $T_n \subset T$ be a sequence of countable subsets such that each T_n is an ε_n-cover of T with $\varepsilon_n \to 0$ (such a family must exist by the separability assumption). The set

$$B = \bigcup_{n\geq 1} \bigcup_{t\in T_n} \bigcap_{s\in T_0} \{f(X_t, X_s) > u\} \cap \Gamma \tag{392}$$

is also measurable as a countable union of a countable intersection of measurable sets. We show that $A = B$. The direction $B \subset A$ is clear: If $\omega \in B$, there must be some $n \geq 1$ and $t_n \in T_n$ such that $f(X_t(\omega), X_s(\omega) > u$ for all $s \in T_0$. In this case, $\sup_{t\in T} \inf_{s\in T_0} f(X_t, X_s) \geq \inf_{s\in T_0} f(X_{t_n}, X_s) > u$.
For the reverse direction, let $\omega \in A$. There must be $t_0 \in T$ and $\varepsilon > 0$ such that $f(X_{t_0}(\omega), X_s(\omega)) \geq u+\varepsilon$ for all $s \in T_0$. As f and $X(\omega)$ are continuous, there must be a neighborhood of t_0, i.e., some $\delta > 0$, such that $f(X_t(\omega), X_s(\omega)) \geq u + \varepsilon/2$ for all $s \in T_0$ and t with $d(t, t_0) \leq \delta$. Let n be large enough such that $\varepsilon_n \leq \delta$ and such that we can find $t_n \in T_n$ with $d(t_n, t_0) \leq \delta$. For this t_n, we have $f(X_{t_n}(\omega), X_s(\omega)) \geq u + \varepsilon/2 > u$ for all $s \in T_0$. Thus, $\omega \in B$, which completes the proof. □

7—CGF bound for database friendly projections

Let X_j denote zero-inflated independent Rademacher random variables, $\mathbb{P}[X_j = \sqrt{c}] = \mathbb{P}[X_j = -\sqrt{c}] = 1/(2c)$ and $\mathbb{P}[X_j = 0] = 1 - 1/c$ for $1 \leq c \leq 3$. Let $Z = \left(\sum_{j=1}^m X_j b_j\right)^2$ for some $b \in \mathbb{R}^M$, $\|b\| = 1$. It is shown in [71] that $\mathbb{E}Z^p \leq (2p)!/(2^p p!)$ for each $p \in \mathbb{N}$ (for $c = 1$ and $c = 3$, but their proof works just as well for $1 \leq c \leq 3$, see also [23, Theorem 8.5]). Consequently,

$$\mathbb{E}Z^p \leq \frac{(2p)!}{2^p p!} = (2p-1)!! = 1 \times 3 \times \cdots \times (2p-1) = \mathbb{E}Q^p \tag{393}$$

where $Q \sim \chi_1^2$ is chi-squared distributed with a single degree of freedom. By the monotone convergence theorem, we have

$$\mathbb{E}\exp(\lambda Z) = \sum_{p\geq 0} \lambda^p \mathbb{E}Z^p/p! \leq \sum_{p\geq 0} \lambda^p \mathbb{E}Q^p/p! = \mathbb{E}\exp(\lambda Q) \tag{394}$$

for $\lambda \geq 0$ so that taking logarithms, we obtain $\Psi_Z(\lambda) \leq \Psi_Q(\lambda)$ so that

$$\Psi_{Z-1}(\lambda) \leq -\lambda - 0.5\log(1-2\lambda), \ 0 \leq \lambda < 1/2. \tag{395}$$

For the lower tail, we use that $-(Z-1) \leq 1$ and

$$\sigma^2 = \mathbb{E}[(-(Z-1))^2] = \mathbb{E}Z^2 - 1 \leq \mathbb{E}Q^2 - 1 \leq 2 \tag{396}$$

and apply Lemma 27 to obtain, for $\lambda < 1/2$,

$$\Psi_{-(Z-1)}(\lambda) \leq \log(2\exp(\lambda) - 2\lambda - 1) \tag{397}$$
$$\leq 2\,(\exp(\lambda) - \lambda - 1) \tag{398}$$
$$= 2\sum_{p\geq 2} \lambda^p/p! \tag{399}$$
$$\leq 0.5\sum_{p\geq 2}(2\lambda)^p/p! \tag{400}$$
$$\leq -\lambda + 0.5\sum_{p\geq 1}(2\lambda)^p/p \tag{401}$$
$$= -\lambda - 0.5\log(1-2\lambda). \tag{402}$$

8—CGF bound for Steinhaus sums

Let $X_j \in \mathbb{C}$ be a sequence of zero-inflated normalized Steinhaus random variables, that is,

$$X_j = \sqrt{c}B_j\exp(iU_j), \tag{403}$$
$$U_j \sim \mathcal{U}_{[-\pi,\pi]} \quad \text{(uniform distribution)}, \tag{404}$$
$$\mathbb{P}[B_j = 1] = 1/c, \quad \mathbb{P}[B_j = 0] = 1 - 1/c, \tag{405}$$

where all U_j and B_j are independent and $1 \leq c \leq 2$. Let $Z = \left|\sum_{j=1}^M X_j b_j\right|^2$ for some $b \in \mathbb{C}^M$, $\|b\| = 1$. For $p, q \geq 0$, we have $\mathbb{E}B_j^{p+q} = 1/c$ and, thus,

$$\mathbb{E}[X_j^p(X_j^q)^*] = \frac{c^{(p+q)/2}}{c}\mathbb{E}[\exp(i(p-q)U_j)] = \begin{cases} c^{p-1}, & p = q > 0, \\ 1, & p = q = 0, \\ 0, & \text{otherwise.}\end{cases} \tag{406}$$

For $\mathbb{E}[Z^p]$, we use the multinomial theorem in

$$\mathbb{E}[Z^p] = \mathbb{E}\left|\sum_{j=1}^M X_j b_j\right|^{2p} = \sum_{\substack{k_1+\cdots+k_M=p\\ \ell_1+\cdots+\ell_M=p}} \frac{(p!)^2\prod_{j=1}^M b_j^{k_j}(b_j^{\ell_j})^*\mathbb{E}[X_j^{k_j}(X_j^{\ell_j})^*]}{k_1!\cdots k_M!\ell_1!\cdots\ell_M!} \tag{407}$$

$$= p!\sum_{k_1+\cdots+k_m=p}\frac{p!\prod_{j=1}^M |b_j|^{2k_j}c^{\max(k_j-1,0)}}{(k_1!\cdots k_M!)^2}. \tag{408}$$

Note that for $k_1 + \cdots + k_M = p$, we have

$$\frac{1}{p!} \leq \prod_{j=1}^M \frac{c^{\max(k_j-1,0)}}{k_j!} \leq 1 \tag{409}$$

and that

$$\sum_{k_1+\cdots+k_M=p} \frac{p!\prod_{j=1}^{M}|b_j|^{2k_j}}{k_1!\cdots k_M!} = \|b\|^{2p} = 1 \tag{410}$$

so that

$$1 \leq \mathbb{E}[Z^p] \leq p!. \tag{411}$$

A moment expansion for the CGF of Z then yields

$$\exp\left(\Psi_Z(\lambda)\right) = \sum_{p\geq 0} \frac{\lambda^p}{p!}\mathbb{E}[Z^p] \leq \frac{1}{1-\lambda}\,, \quad \lambda \geq 0 \tag{412}$$

from which

$$\Psi_{Z-1}(\lambda) \leq -\lambda - \log(1-\lambda), \quad \lambda \geq 0 \tag{413}$$

gives us the same bound we already encountered for complex Gaussian RVs. For the lower tail, we use that $-(Z-1) \leq 1$ and

$$\mathbb{E}[(-(Z-1))^2] = \mathbb{E}Z^2 - 1 \leq 1 \tag{414}$$

and apply Lemma 27 to obtain, for $\lambda < 1$,

$$\begin{aligned}\Psi_{-(Z-1)}(\lambda) &\leq \log(\exp(\lambda)-\lambda) \leq \exp(\lambda) - 1 - \lambda \\ &= \sum_{p\geq 2}\lambda^p/p! \leq -\lambda + \sum_{p\geq 1}\lambda^p/p = -\lambda - \log(1-\lambda).\end{aligned} \tag{415}$$

9—CGF bound for structured matrices

We have

$$\Psi_Z(\lambda) = \sum_{n=1}^{M} \log\left(q + p\exp(\lambda|\langle x, u_n\rangle|^2/p)\right). \tag{416}$$

Let us calculate an upper bound for Ψ_Z that does not depend on x. We exploit that $1 = \|x\|^2 = \sum_{n=1}^{M}|\langle x, u_n\rangle|^2$, $\|x\|_1 \leq \sqrt{k}\|x\|_2$, and

$$|\langle x, u_n\rangle|^2 \leq |\|x\|_1\|u_n\|_\infty|^2 \leq kc_\infty/M =: M_{\text{eff}}^{-1}. \tag{417}$$

As each term $f\colon \xi \mapsto \log(q + p\exp(\lambda\xi))$ is convex with $f(0) = 0$, we have

$$\begin{aligned}\Psi_Z(\lambda) &\leq \sum_{n=1}^{M} M_{\text{eff}}|\langle x, u_n\rangle|^2 \log\left(q + p\exp(\lambda/(pM_{\text{eff}}))\right) && (418)\\ &= M_{\text{eff}}\log\left(q + p\exp(\lambda/(pM_{\text{eff}}))\right) && (419)\\ &\leq pM_{\text{eff}}\left(\exp(\lambda/(pM_{\text{eff}})) - 1\right). && (420)\end{aligned}$$

For the lower tail, we use that

$$\Psi_{-(Z-1)}(\lambda) = \Psi_{Z-1}(-\lambda) = \lambda + pM_{\text{eff}}\left(\exp\left(-\lambda/(pM_{\text{eff}})\right) - 1\right) \tag{421}$$

$$\leq -\lambda + pM_{\text{eff}}\left(\exp\left(\lambda/(pM_{\text{eff}})\right) - 1\right) \tag{422}$$

where the inequality is quickly derived from

$$e^{-x} - 1 + x = \sum_{p\geq 2}(-x)^p/p! \leq \sum_{p\geq 2} x^p/p! = e^x - 1 - x. \tag{423}$$

We calculate the rate function corresponding to the upper bound

$$c(\lambda) = pM_{\text{eff}}\left(\exp\left(\frac{\lambda}{pM_{\text{eff}}}\right) - 1 - \frac{\lambda}{pM_{\text{eff}}}\right). \tag{424}$$

We obtain $\lambda = pM_{\text{eff}}\log(\mu+1)$ as a solution to the first-order optimality condition

$$0 = \frac{d}{d\lambda}\left(\mu\lambda - c(\lambda)\right) = \mu - \exp\left(\frac{\lambda}{pM_{\text{eff}}}\right) + 1 \tag{425}$$

so that

$$c^*_{Z-1}(\mu) = \sup_{\lambda\geq 0} \mu\lambda - c(\lambda) = pM_{\text{eff}}\left((1+\mu)\log(1+\mu) - \mu\right) \tag{426}$$

is a rate function for $Z-1$ and $-(Z-1)$.

10—Proof of Lemma 21

Define the sequences

$$\delta_n = (1+n)\alpha\delta = (1+n)\delta_0 \tag{427}$$

$$\varepsilon_n = \varepsilon_0 g(\varepsilon_0)^{-n/k} \tag{428}$$

for $n \geq 1$. Let ω be a convex function with $\omega(0) = 0$, then $\omega(nx) \geq n\omega(x)$ for $n \geq 1$. This can be seen from

$$\omega((nx)/n) \leq \omega(nx)/n + \omega(0)(n-1)/n = \omega(nx)/n. \tag{429}$$

For a log-concave function with $f(0) = 1$, we have $f(x) = \exp(-\omega(x))$ with $\omega(0) = 0$ and, hence, $f(nx) \leq \exp(-n\omega(x)) = f(x)^n$. Because $1 + n \geq 1$ for all n, we can use this inequality to get

$$f(\delta_n) = f((1+n)\delta_0) \leq f(\delta_0)^{1+n} \leq f(\delta_0)\left(2g(\varepsilon_0)\right)^{-n} \tag{430}$$

and, by the geometric regularity of g,

$$g(\varepsilon_n) \leq (\varepsilon_0/\varepsilon_n)^k g(\varepsilon_0) = g(\varepsilon_0)^{n+1} \tag{431}$$

such that $g(\varepsilon_n)f(\delta_n) \le g(\varepsilon_0)f(\delta_0)2^{-n}$ and

$$\sum_{n\ge1} g(\varepsilon_n)f(\delta_n) \le g(\varepsilon_0)f(\delta_0). \tag{432}$$

As concerns the other summation (168), we have

$$\sum_{n\ge1} \varepsilon_{n-1}\sqrt[p]{1+\delta_n} = \varepsilon_0\left(\sqrt[p]{1+2\alpha\delta}+\eta\right) \tag{433}$$

with

$$\eta = \sum_{n\ge1} g(\varepsilon_0)^{-n/k}\sqrt[p]{1+(2+n)\alpha\delta}. \tag{434}$$

To show (168), which can be reformulated as

$$\varepsilon_0\left(\sqrt[p]{1+2\alpha\delta}+\eta\right) \le \sqrt[p]{1+\delta} - \sqrt[p]{1+\alpha\delta}, \tag{435}$$

we use the following inequalities that follow from the concavity and a first-order approximation of the function $x \mapsto \sqrt[p]{x}$:

$$\sqrt[p]{1+\delta} - \sqrt[p]{1+\alpha\delta} \ge \frac{(1-\alpha)\delta}{p\sqrt[p]{(1+\delta)^{p-1}}} \tag{436}$$

$$\sqrt[p]{1+2\alpha} - \sqrt[p]{1+2\alpha\delta} \ge \frac{2(1-\delta)\alpha}{p\sqrt[p]{(1+2\alpha)^{p-1}}} \tag{437}$$

$$\sqrt[p]{1+2\alpha\delta+n\alpha\delta} - \sqrt[p]{1+2\alpha\delta} \le \frac{n\alpha\delta}{p\sqrt[p]{(1+2\alpha\delta)^{p-1}}}\ . \tag{438}$$

First, by (437), the inequality $\sqrt[p]{1+2\alpha\delta}+\eta \le \sqrt[p]{1+2\alpha}$ is implied by the stricter condition

$$\eta \le \frac{2(1-\delta)\alpha}{p\sqrt[p]{(1+2\alpha)^{p-1}}}. \tag{439}$$

Then, if (439) is satisfied, also (435) is fulfilled because

$$\varepsilon_0\left(\sqrt[p]{1+2\alpha\delta}+\eta\right) \le \varepsilon_0\sqrt[p]{1+2\alpha} \overset{(166)}{\le} \frac{(1-\alpha)\delta}{p\sqrt[p]{(1+\delta)^{p-1}}} \overset{(436)}{\le} \sqrt[p]{1+\delta} - \sqrt[p]{1+\alpha\delta}. \tag{440}$$

To show (439), we need to calculate the series (434), which determines η. If we use the abbreviation $\xi = g(\varepsilon_0)^{-1/k}$, we obtain

$$\eta = \sum_{n\ge1} \xi^n\sqrt[p]{1+(2+n)\alpha\delta} \tag{441}$$

$$\overset{(438)}{\le} \sum_{n\ge1} \xi^n\left(\sqrt[p]{1+2\alpha\delta} + \frac{n\alpha\delta}{p\sqrt[p]{(1+2\alpha\delta)^{p-1}}}\right) \tag{442}$$

$$= \xi\left(\frac{\sqrt[p]{1+2\alpha\delta}}{1-\xi} + \frac{\alpha\delta}{p\sqrt[p]{(1+2\alpha\delta)^{p-1}}(1-\xi)^2}\right). \tag{443}$$

Inserting ε_0 from (166) into ξ results in

$$\xi = g(\varepsilon_0)^{-1/k} = \frac{(1-\alpha)\delta}{N_0 p \sqrt[p]{(1+2\alpha)2^{p-1}}} \le \frac{1}{4pN_0} \le \frac{1}{4} \tag{444}$$

for $\alpha \ge 1/2$ and $p, N_0 \ge 1$. Following up on (443), we obtain

$$\eta \le \xi \left(\frac{\sqrt[p]{1+2\alpha\delta}}{1-\xi} + \frac{\alpha\delta}{p\sqrt[p]{(1+2\alpha\delta)^{p-1}}(1-\xi)^2} \right) \tag{445}$$

$$\le \xi \left(\frac{4}{3}\sqrt[p]{1+2\alpha\delta} + \frac{16}{9}\frac{\alpha\delta}{p\sqrt[p]{(1+2\alpha\delta)^{p-1}}} \right) \le \frac{2(1-\alpha)\delta}{N_0 p \sqrt[p]{2^{p-1}}} \tag{446}$$

where, in the last step, we used that

$$\frac{\alpha\delta}{p\sqrt[p]{(1+2\alpha)(1+2\alpha\delta)^{p-1}}} \le \frac{1}{3p} \text{ and } \sqrt[p]{\frac{1+2\alpha\delta}{1+2\alpha}} \le 1 \tag{447}$$

and $4/3 + 16/(27p) \le 2$. Thus, condition (439) is satisfied if

$$\frac{2(1-\alpha)\delta}{N_0 p \sqrt[p]{2^{p-1}}} \le \frac{2(1-\delta)\alpha}{p\sqrt[p]{(1+2\alpha)^{p-1}}}, \tag{448}$$

i.e., if

$$\frac{(1-\alpha)\delta}{\alpha N_0(1-\delta)} \left(\frac{1+2\alpha}{2} \right)^{(p-1)/p} \le 1. \tag{449}$$

Because of the condition $\delta \le 1 - 1/N_0$ and because $((1+2\alpha)/2)^{(p-1)/p} \le (1 + 2\alpha)/2$ for all $p \ge 1$ if $\alpha \ge 1/2$, this is satisfied if

$$1 \ge \frac{(1-\alpha)(1+2\alpha)}{2\alpha} = \frac{1+\alpha-2\alpha^2}{\alpha+\alpha}. \tag{450}$$

Because $\alpha \ge 1/2$, we have

$$\frac{1+\alpha-2\alpha^2}{\alpha+\alpha} \le \frac{1+\alpha-1/2}{1/2+\alpha} = 1 \tag{451}$$

which concludes the proof.

11—Proof of Lemma 22

Let $\delta_n = 2^n - 1$ and set

$$\varepsilon_n = \varepsilon_0 \sqrt[k]{\frac{2^n f(\delta_n)}{f(\delta_0)}} \tag{452}$$

for $n \ge 1$. By the geometric regularity of g, we have

$$g(\varepsilon_n) \le (\varepsilon_0/\varepsilon_n)^k g(\varepsilon_0) = \frac{g(\varepsilon_0)f(v_0)}{2^n f(v_n)} \tag{453}$$

so that $g(\varepsilon_n)f(\delta_n) \leq g(\varepsilon_0)f(\delta_0)2^{-n}$ and, thus,

$$\sum_{n\geq 1} g(\varepsilon_n)f(\delta_n) \leq g(\varepsilon_0)f(\delta_0)\sum_{n\geq 1} 2^{-n} = g(\varepsilon_0)f(\delta_0). \tag{454}$$

As concerns the other summation (171), we have

$$\sum_{n\geq 1} \varepsilon_{n-1}\sqrt[p]{1+\delta_n} = \varepsilon_0\left(\sqrt[p]{2}+\eta\right) \tag{455}$$

with

$$\eta = \frac{1}{\sqrt[k]{f(\delta_0)}}\sum_{n\geq 1}\sqrt[k]{2^n f(\delta_n)}\sqrt[p]{2^{n+1}}. \tag{456}$$

We now observe that the monotonicity and convexity of ω yield[4]

$$\omega(\delta_n) \geq 2^{n-1}\left(\omega(1)-\omega(0)\right). \tag{459}$$

If we insert this expression into (456), we obtain

$$\eta \leq \frac{1}{\sqrt[k]{f(\delta_0)}}\sum_{n\geq 1}\sqrt[k]{\frac{2^n}{(2^{n-1}(\omega(1)-\omega(0)))^q}}\sqrt[p]{2^{n+1}} \tag{460}$$

$$= \sqrt[k]{2^{q+k/p}\left(\frac{\omega(\delta_0)}{\omega(1)-\omega(0)}\right)^q}\sum_{n\geq 1}\left(\sqrt[k]{2^{-(q-1-k/p)}}\right)^n \tag{461}$$

$$= \sqrt[k]{2^{q+k/p}\left(\frac{\omega(\delta_0)}{\omega(1)-\omega(0)}\right)^q}\frac{2^{-(q-1-k/p)/k}}{1-2^{-(q-1-k/p)/k}} \tag{462}$$

$$= \sqrt[p]{4}\left(\frac{\omega(\delta_0)}{\omega(1)-\omega(0)}\right)^{q/k}\frac{\sqrt[k]{2}}{1-2^{-(q-1-k/p)/k}} \tag{463}$$

$$\leq 2^{1+1/k+1/p}\left(\frac{\omega(\delta_0)}{\omega(1)-\omega(0)}\right)^{q/k}\sqrt[p]{2} \tag{464}$$

if $q > 1+k/p$. With the choice of ε_0, we achieve

$$\varepsilon_0\left(\sqrt[p]{2}+\eta\right) \leq \frac{(1-\alpha)\delta}{p\sqrt[p]{2^{p-1}}} \leq \sqrt[p]{1+\delta}-\sqrt[p]{1+\alpha\delta} \tag{465}$$

as in the proof of Lemma 21 (see (436)).

---

[4]For $\alpha = 2^{1-n}, x = 2^n - 1 = v_n, y = 0$, we obtain for $n \geq 1$

$$\begin{aligned}\omega(1) &\leq \omega(2-2^{1-n}) = \omega(\alpha x+(1-\alpha)y) \leq \alpha\omega(x)+(1-\alpha)\omega(y)\\ &= 2^{1-n}\omega(\delta_n)+(1-\alpha)\omega(0) \leq 2^{1-n}\omega(\delta_n)+\omega(0)\end{aligned} \tag{457}$$

from which

$$\omega(\delta_n) \geq 2^{n-1}(\omega(1)-\omega(0)) \tag{458}$$

follows.

12—Sub-Gaussian norm of a Gaussian random variable

The sub-Gaussian norm of a Gaussian random variable is given by

$$\|X\|_{\Psi_2} = \inf\left\{t > 0 : \mathbb{E}\exp(X^2/t^2) \le 2\right\} \tag{466}$$
$$= \inf\left\{t > \sqrt{2} : (1 - 2/t^2)^{-1/2} \le 2\right\} = \sqrt{8/3}. \tag{467}$$

If A has normalized iid. real Gaussian entries, $a^T x$ is Gaussian and $\|a^T x\|_{\Psi_2} = \sqrt{8/3}$. Then, going through the proof of [58, Theorem 8] we observe that $c_1 = c/(64\Lambda_2^4)$ with $\Lambda_2^4 = \|a^T x\|_{\Psi_2}^4 = 64/9$ and an unspecified constant c (which appears when showing the equivalence of the different characterizing properties of sub-exponential random variables).

13—Lipschitz continuity of the Vandermonde projector

Let $A\colon [0,1] \to \mathbb{C}^{M\times k}, t \mapsto A(t)$, be a continuously differentiable matrix-valued function with full column rank for all t and with derivative $\dot{A} = \frac{d}{dt}A(t)$. It is well known (cf. [76]) that the derivative of the projector $P = AA^\dagger$ is given by $\dot{P} = P^\perp \dot{A}A^\dagger P + P(\dot{A}A^\dagger)^H P^\perp$ where $P^\perp = I - P$ and where $(\cdot)^H$ denotes the conjugate transpose of a matrix. By the fundamental theorem of calculus, we have

$$\|P(1) - P(0)\| = \left\| \int_0^1 \dot{P}(t)dt \right\| \le \sup_t \|\dot{P}(t)\|$$
$$\le \sup_t \|P^\perp(t)\dot{A}(t)A^\dagger(t)P(t)\| \le \sup_t \|\dot{A}(t)A^\dagger(t)\| \tag{468}$$

where the supremum is over t in the interval [0, 1] and where the second inequality follows from (with $B = P^\perp \dot{A}A^\dagger P$)

$$\|(P^\perp BP + PB^H P^\perp)x\|^2 = \|P^\perp BPx\|^2 + \|PB^H P^\perp x\|^2 \quad (\text{as } PP^\perp = 0)$$
$$\le \|B\|^2\|Px\|^2 + \|B^H\|^2\|P^\perp x\|^2 = \|B\|^2\|x\|^2. \tag{469}$$

An upper bound is obtained by using $\|\dot{A}A^\dagger\| \le \|\dot{A}\|\|A^\dagger\|$. The main difficulty is that when $A = V$ is the Vandermonde matrix, a uniform bound for $\|V^\dagger\|$ cannot exist, because $V(z)$ becomes rank deficient as $z_i \to z_j$. A straightforward proof would require a minimum separation condition $|z_i - z_j| \ge \varepsilon$. Our principal contribution is to show that Newton's divided differences yield a re-parametrization of range(V) that avoids such a condition.
Define Newton's divided differences by (cf. [77])

$$f[z_1,\ldots,z_j] = \sum_{\ell=1}^{j} \frac{f(z_\ell)}{\prod_{i=1,\ldots,j, i\neq\ell}(z_\ell - z_i)} \tag{470}$$

and the matrix $W(z) = \big[f[z_1] \dots f[z_1,\dots,z_k]\big]$. It follows from (470) that $W(z) = V(z)\Delta(z)$ with an upper triangular matrix $\Delta(z)$ with

$$\det \Delta(z) = \prod_{j=1}^{k}\prod_{i=1}^{j-1}(z_j - z_i)^{-1} \neq 0 \tag{471}$$

if $z_i \neq z_j$ when $i \neq j$. Thus, the two projectors $P_W = WW^\dagger$ and $P_V = VV^\dagger$ coincide: $P_W(z) = P_V(z)$ if $z \in \mathcal{B}'_R$.
To see what advantages the divided differences offer, let $z' \in \mathcal{B}_R \setminus \mathcal{B}'_R$ denote a problematic node with, e.g., $z'_1 = z'_2$, and let $(z^n)_{n\in\mathbb{N}} \subset \mathcal{B}'_R$ be a sequence with $z^n \to z'$. As the first two columns of $V(z')$, which are given by $f(z'_1) = f(z'_2)$, coincide, the projector $P_V(z')$ is rank-deficient. Hence, $P_V(z')$ cannot be the limit of the sequence $P_V(z^n)$, because the projectors $P_V(z^n)$ are not rank-deficient. In contrast, the first two columns of the Newton matrix $W(z')$ are given by $f[z'_1] = f(z'_1)$ and $f[z'_1, z'_2] = \dot{f}(z'_1)$, and these vectors are linearly independent. Consequently, $P_W(z')$ is not rank-deficient and it turns out that $P_W(z') = \lim_{n\to\infty} P_V(z^n)$. Thus, the correct way to define the projector for such z' is to use P_W instead of P_V. The relationship between Newton's differences with repeated nodes and the derivatives follows from the Hermite-Genocchi theorem stated below.
We now proceed from the single-parameter formula (468) to the more general statement in Theorem 30. First, we define $\rho = \max(1, R)$ with the sole purpose of simplifying the problem to the special case $R = 1$. Let $\xi\colon [0,1] \to B_1$ be a continuously differentiable curve with $y = \rho\xi(0)$ and $z = \rho\xi(1)$. Because $W(\rho z) = \Lambda W(z)$ with a matrix $\Lambda = \operatorname{diag}(1, \rho, \dots, \rho^{M-1})$ that is independent of z, we can restrict our attention to B_1 in the following sections. Using $A(t) = W(\rho\xi(t))$ and $P(t) = P_W(\rho\xi(t))$ in (468), we obtain

$$\|P(1) - P(0)\| \leq \sup_t \Big\|\frac{d}{dt} W\big(\rho\xi(t)\big)\Big\|\Big\|W^\dagger\big(\rho\xi(t)\big)\Big\| \tag{472}$$

$$= \sup_t \Big\|\Lambda \frac{d}{dt} W\big(\xi(t)\big)\Big\|\Big\|\Big(\Lambda W\big(\xi(t)\big)\Big)^\dagger\Big\| \tag{473}$$

$$\leq \kappa(\Lambda) \sup_t \Big\|\frac{d}{dt} W\big(\xi(t)\big)\Big\|\Big\|W^\dagger\big(\xi(t)\big)\Big\| \tag{474}$$

where $\kappa(\Lambda) = \sigma_{\max}(\Lambda)/\sigma_{\min}(\Lambda) = \rho^{M-1}$ is the condition number of Λ. Next, we show how to bound the derivative and pseudo-inverse of W in B_1 where $\|z\|_\infty \leq 1$.

Theorem 38 (see [78])**.** *The smallest singular value of a matrix $A \in \mathbb{C}^{k\times k}$ satisfies*

$$\sigma_{min}(A) \geq \left(\frac{k-1}{k}\right)^{\frac{k-1}{2}} |\det A| \frac{\min_n r_n}{\prod_{n=1}^{k} r_n} \tag{475}$$

where r_n denotes the Euclidean norm of the nth row of A.

The following theorem is a direct corollary of the Hermite-Genocchi theorem applied to the function $z \mapsto z^n$ (assuming $\|z\|_\infty \leq 1$, see Sec. 14 below).

Lemma 39. *The nth entry of $f[z_1, \ldots, z_j]$ satisfies*

$$f_n[z_1, \ldots, z_j] = 0 \qquad \text{if } n < j, \tag{476}$$
$$f_n[z_1, \ldots, z_j] = 1 \qquad \text{if } n = j, \tag{477}$$
$$|f_n[z_1, \ldots, z_j]| \leq \binom{n-1}{j-1} \qquad \text{if } n \geq j. \tag{478}$$

For the row norms r_n of W, we obtain

$$r_n^2 = \sum_{j=1}^{k} |f_n[z_1, \ldots, z_j]|^2 \leq \sum_{j=1}^{n} \binom{n-1}{j-1}^2 \leq 2^{2(n-1)}. \tag{479}$$

Define U as the first k rows of W. As $\|Wx\| \geq \|Ux\|$ implies $\sigma_{\min}(W) \geq \sigma_{\min}(U)$, we can use Theorem 38 with U:

$$\|W^\dagger\|^2 \leq \frac{1}{\sigma_{\min}(U)^2} \leq e \prod_{n=1}^{k} 2^{2(n-1)} = e2^{k(k-1)} \tag{480}$$

as $(k/(k-1))^{k-1} \leq e$ and because also $\min_n r_n(U) = 1$ and $\det(U) = 1$ by virtue of Lemma 39.
Another consequence of the Hermite-Genocchi theorem is the following expression for the derivative of the divided differences and the ensuing bound for the Frobenius norm (assuming $\|z\|_\infty \leq 1$, see Sec. 14 below).

Lemma 40. *The derivative of the jth column of $W(z)$ with respect to the ℓth coordinate z_ℓ is given by*

$$\frac{\partial f[z_1, \ldots, z_j]}{\partial z_\ell} = \begin{cases} f[z_1, \ldots z_j, z_\ell], & \text{if } \ell \leq j, \\ 0, & \text{if } \ell > j. \end{cases} \tag{481}$$

Furthermore,

$$\left\| \frac{\partial W(z)}{\partial z_\ell} \right\|_F \leq \frac{M^{k+1/2}}{k!}. \tag{482}$$

By the chain rule, we obtain

$$\begin{aligned} \left\| \frac{d}{dt} W(\xi(t)) \right\|_F &= \left\| \sum_{\ell=1}^{k} \frac{\partial W(z)}{\partial z_\ell} [\dot{\xi}(t)]_\ell \right\|_F \\ &\leq \sum_{\ell=1}^{k} |[\dot{\xi}(t)]_\ell| \left\| \frac{\partial W(z)}{\partial z_\ell} \right\|_F \leq \|\dot{\xi}(t)\|_1 \frac{M^{k+1/2}}{k!}. \end{aligned} \tag{483}$$

Theorem 30 follows from (474), (480), (483), by setting $\xi(t) = y/\rho + t(z-y)/\rho$, for which $\|\dot{\xi}(t)\|_1 = \|z-y\|_1/\rho$ holds, and noting that the remaining expressions do not depend on t.

14—Hermite-Genocchi Theorem

The following version of the Hermite-Genocchi theorem for holomorphic functions is a straightforward generalization of the version for real functions (cf., e.g., [79]). The theorem can be used to define the differences with repeated arguments.

Theorem 41. *Let $f\colon U \to \mathbb{C}$ be holomorphic on an open subset $U \subset \mathbb{C}$ with jth complex derivative $f^{(j)}$. If the convex hull of the nodes satisfies* $\mathrm{conv}(z_0, \dots, z_j) \subset U$*, we have*

$$f[z_0, \dots, z_j] = \int_{\tau_j} f^{(j)}\left(\sum_{i=0}^{j} t_i z_i\right) d\lambda(t) \tag{484}$$

where $\tau_j = \{t \geq 0 : \sum_{i=0}^{j} t_i = 1\}$ is the unit simplex and $d\lambda$ the Lebesgue measure on τ_j.

The first consequence of this theorem is that the order of the nodes in $f[z_0, \dots, z_j]$ is irrelevant. Moreover, by exchanging integration and differentiation in (484), one can show that the partial derivative with respect to the ℓth node is given by [80]

$$\frac{\partial}{\partial z_\ell} f[z_0, \dots, z_j] = f[z_0, \dots, z_j, z_\ell] \tag{485}$$

if $\ell \leq j$, i.e., the node z_ℓ is repeated. Lemma 39 follows by observing that for $f(z) = z^n$, we have $f^{(j)}(z) = 0$ for $n < j$ and $f^{(j)}(z) = n!/(n-j)!\, z^{n-j}$ for $n \geq j$. Because the volume of the simplex is given by $\int_{\tau_j} d\lambda = 1/j!$, we obtain the bound

$$\begin{aligned} |f[z_0, \dots, z_j]| &\leq \int_{\tau_j} \left| \frac{n!}{(n-j)!} \left(\sum_{i=0}^{j} t_i z_i\right)^{n-j} \right| d\lambda(t) \\ &\leq \frac{n! \cdot \mathrm{vol}(\tau_j)}{(n-j)!} \sup_{t \in \tau_j} \left| \sum_{i=0}^{j} t_i z_i \right|^{n-j} \leq \binom{n}{j} \|z\|_\infty^{n-j} \end{aligned} \tag{486}$$

for $n \geq j$. If $n = j$, we even have equality, because in this case $f^{(n)}(z) = j!$ is constant so that $\int_{\tau_j} f^{(j)} d\lambda = j! \int_{\tau_j} d\lambda = 1$.
Furthermore, Lemma 40 follows by applying (485) to all entries of the vector $f[z_1, \dots, z_j]$. Using Lemma 39 with j replaced by $j+1$, we can bound its Euclidean norm by

$$\begin{aligned} \|f[z_1, \dots, z_j, z_\ell]\|^2 &\leq \sum_{n=1}^{M} \binom{n-1}{j}^2 \leq \binom{M-1}{j} \sum_{n=1}^{M} \binom{n-1}{j} \\ &= \binom{M-1}{j} \binom{M}{j+1} = \frac{j+1}{M} \binom{M}{j+1}^2 \end{aligned} \tag{487}$$

if $j \geq \ell$ and zero otherwise. From this, we obtain the bound on the Frobenius norm of the matrix $\partial W(z)/\partial z_\ell$ stated in Lemma 40:

$$\sum_{j=\ell}^{k} \|f[z_1, \ldots, z_j, z_\ell]\|^2 \leq \sum_{j=\ell}^{k} \frac{j+1}{M} \binom{M}{j+1}^2$$
$$\leq \sum_{j=\ell}^{k} \frac{(j+1)M^{2j+1}}{(j+1)!^2} \leq \sum_{j=\ell}^{k} \frac{(k+1)M^{2k+1}}{(k+1)!^2} \leq \frac{M^{2k+1}}{k!^2}. \quad (488)$$

15—Sum of squares

The following simple result can sometimes be used to achieve slightly tighter bounds than the trivial bound $a^2 + b^2 \leq (a+b)^2$ for non-negative a, b even when it is not known whether $a \geq b$ or $b \geq a$.

Lemma 42. *Let $a, b \geq 0$. Then*

$$a^2 + b^2 \leq (\gamma_L a + \gamma_R b)^2 \quad (489)$$

and $\{\gamma_L, \gamma_R\} = \{1, 1/(1+\sqrt{2})\}$, i.e., one of γ_L and γ_R is smaller than one.

Proof. Let $a \geq b$ and set $x = b/c$ for some $c \geq 1$. Then $a \geq cx = b$ and

$$(a+x)^2 = a^2 + 2ax + x^2 \geq a^2 + (2c+1)x^2 = a^2 + \frac{2c+1}{c^2} b^2. \quad (490)$$

If we set $c = 1 + \sqrt{2}$, we obtain $(2c+1)/c^2 = 1$ and get

$$a^2 + b^2 \leq (a+x)^2 = \left(a + \frac{b}{1+\sqrt{2}}\right)^2. \quad (491)$$

The same holds for the roles of a and b reversed so that we get

$$a^2 + b^2 \leq (\gamma_L a + \gamma_R b)^2 \quad (492)$$

where one of γ_L and γ_R is one and the other value is $1/(1+\sqrt{2})$ depending on whether $a \geq b$ or $b \geq a$. □

Bibliography

[1] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[2] T. Wiese, L. Weiland, and W. Utschick, "Low-rank approximations for spatial channel models," in *20th International ITG Workshop on Smart Antennas (WSA)*, Mar. 2016.

[3] ——, "Exact recovery of structured block-sparse signals with model-aware orthogonal matching pursuit," in *17th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2016, pp. 1–5.

[4] ——, "Inexact projected gradients on unions of subspaces," in *IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 236–240.

[5] M. Koller, T. Wiese, and W. Utschick, "A generalized hard thresholding pursuit on infinite unions of subspaces," in *22nd International ITG Workshop on Smart Antennas (WSA)*, Mar. 2018.

[6] T. Wiese and W. Utschick, "Quantitative bounds for restricted isometries in geometrically regular spaces," 2020, submitted.

[7] T. Wiese, D. Neumann, and W. Utschick, "Lipschitz continuity of the Vandermonde projector," 2020, in preparation.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[9] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[10] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.

[11] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.

[12] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[13] J.-J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1341–1344, Jun. 2004.

[14] ——, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3601–3608, Oct. 2005.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via lq-minimization for 0<q<=1," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, 2009.

[17] M. E. Davies and R. Gribonval, "Restricted isometry constants where ℓ^p sparse recovery can fail for $0 \ll p \leq 1$," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2203–2214, May 2009.

[18] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[19] T. T. Cai and A. Zhang, "Sparse representation of a polytope and recovery of sparse signals and low-rank matrices," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 122–132, Jan. 2014.

[20] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3654–3656, Jun. 2012.

[21] T. Blumensath and M. E. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1872–1882, Apr. 2009.

[22] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.

[23] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer New York, 2013.

[24] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, Feb. 2014.

[25] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[26] J. Haupt, W. U. Bajwa, G. Raz, and R. Nowak, “Toeplitz compressed sensing matrices with applications to sparse channel estimation,” *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.

[27] H. Rauhut, K. Schnass, and P. Vandergheynst, “Compressed sensing and redundant dictionaries,” *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008.

[28] Y. C. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[29] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, “Block-sparse signals: Uncertainty relations and efficient recovery,” *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[30] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[31] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[32] M. A. Davenport, D. Needell, and M. B. Wakin, “Signal space CoSaMP for sparse recovery with redundant dictionaries,” *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6820–6829, Oct. 2013.

[33] T. Blumensath, “Sampling and reconstructing signals from a union of linear subspaces,” *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4660–4671, Jul. 2011.

[34] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, “MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?” *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 110–121, Dec. 2014.

[35] A. Maltsev, R. Maslennikov, A. Sevastyanov, A. Khoryaev, and A. Lomayev, “Experimental investigations of 60 GHz WLAN systems in office environment,” *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1488–1499, Oct. 2009.

[36] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, “Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 1850–1859, Apr. 2013.

[37] 3GPP, “Spatial channel model for multiple input multiple output (MIMO) simulations (release 12),” 3rd Generation Partnership Project (3GPP), TR 25.996 V12.0.0, 2014.

[38] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: Channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 383–395, Oct. 2007.

[39] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.

[40] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE J. Ocean. Eng.*, vol. 32, no. 4, pp. 927–939, Oct. 2007.

[41] A. T. James, "Normal multivariate analysis and the orthogonal group," *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, 1954. [Online]. Available: http://www.jstor.org/stable/2236512

[42] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'83)*, vol. 8, Apr. 1983, pp. 336–339.

[43] C. Hegde, P. Indyk, and L. Schmidt, "Approximation algorithms for model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 5129–5147, Sep. 2015.

[44] ——, "Fast recovery from a union of subspaces," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 4394–4402.

[45] ——, "A nearly-linear time framework for graph-structured sparsity," in *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, 2015, pp. 928–937.

[46] P. Jain, N. Rao, and I. Dhillon, "Structured sparse regression via greedy hard thresholding," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 1516–1524.

[47] R. Giryes, Y. C. Eldar, A. M. Bronstein, and G. Sapiro, "Tradeoffs between convergence speed and reconstruction accuracy in inverse problems," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1676–1690, Apr. 2018.

[48] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[49] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[50] L. Weiland, T. Wiese, and W. Utschick, "Coherent MIMO radar imaging with model-aware block sparse recovery," in *6th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2015, pp. 425–428.

[51] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Comm. Pure Appl. Math.*, vol. 67, pp. 906–956, Apr. 2014.

[52] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.

[53] A. Bandeira, E. Dobriban, D. Mixon, and W. Sawin, "Certifying the restricted isometry property is hard," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3448–3450, Jun. 2012.

[54] J. Weed, "Approximately certifying the restricted isometry property is hard," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5488–5497, Aug. 2018.

[55] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4, pp. 918–925, 2007.

[56] A. Bandeira, M. Fickus, D. Mixon, and P. Wong, "The road to deterministic matrices with the restricted isometry property," *J. Fourier Anal. Appl.*, vol. 19, pp. 1123–1149, 2013.

[57] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, and D. Kutzarova, "Explicit constructions of rip matrices and related problems," *Duke Math. J.*, vol. 159, no. 1, pp. 145–185, Jul. 2011. [Online]. Available: https://doi.org/10.1215/00127094-1384809

[58] G. Puy, M. E. Davies, and R. Gribonval, "Recipes for stable linear embeddings from Hilbert spaces to $\mathbb{R}^m$," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2171–2187, Apr. 2017.

[59] W. Mantzel and J. Romberg, "Compressed subspace matching on the continuum," *Information and Inference: A Journal of the IMA*, vol. 4, no. 2, Jun. 2015.

[60] S. Dirksen, "Dimensionality reduction with subgaussian matrices: A unified theory," *Found. Comput. Math.*, vol. 16, no. 5, pp. 1367–1396, 2016.

[61] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," in *Compressed Sensing: Theory and Applications*, Y. C. Eldar and G. Kutyniok, Eds. Cambridge University Press, May 2012, pp. 210–268.

[62] E. J. Candès and Y. Plan, "Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2342–2359, Apr. 2011.

[63] M. Talagrand, *Upper and Lower Bounds of Stochastic Processes*. Springer, 2014.

[64] R. Adamczak, A. E. Litvak, A. Pajor, and N. Tomczak-Jaegermann, "Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling," *Constr. Approx.*, vol. 34, pp. 61–88, 2011.

[65] F. Barthe, P. Cattiaux, and C. Roberto, "Concentration for independent random variables with heavy tails," *Applied Mathematics Research eXpress*, vol. 2005, no. 2, pp. 39—60, Mar. 2005.

[66] S. Dirksen, "Tail bounds via generic chaining," *Electron. J. Probab.*, vol. 20, no. 53, 2015.

[67] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

[68] R. van Handel, "Probability in high dimension," 2016, [Online]. Available: https://www.princeton.edu/~rvan/ORF570.pdf.

[69] M. Fradelizi, M. Madiman, and L. Wang, "Optimal concentration of information content for log-concave densities," in *High Dimensional Probability VII*, C. Houdré, D. M. Mason, P. Reynaud-Bouret, and J. Rosiński, Eds. Springer International Publishing, 2016, pp. 45–60.

[70] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford Scholarship Online, May 2013.

[71] D. Achlioptas, "Database-friendly random projections," in *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ser. PODS '01. New York, NY, USA: ACM, 2001, pp. 274–281. [Online]. Available: http://doi.acm.org/10.1145/375551.375608

[72] I. Haviv and O. Regev, "The restricted isometry property of subsampled Fourier matrices," in *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2014–2016*, B. Klartag and E. Milman, Eds. Cham: Springer International Publishing, 2017, pp. 163–179.

[73] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[74] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Comm. Pure Appl. Math.*, vol. 61, no. 8, pp. 1025–1045, Nov. 2007.

[75] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.

[76] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1110–1121, May 1991.

[77] G. M. Phillips, *Interpolation and Approximation by Polynomials*. Springer, 2003.

[78] Y. Hong and C.-T. Pan, "A lower bound for the smallest singular value," *Linear Algebra and its Applications*, vol. 172, pp. 27–32, 1992.

[79] L. Filipsson, "Complex mean-value interpolation and approximation of holomorphic functions," *Journal of Approximation Theory*, vol. 91, pp. 244–278, 1997.

[80] K. E. Atkinson, *An Introduction to Numerical Analysis*. Wiley, 1988.