

A geometric analysis of convex demixing

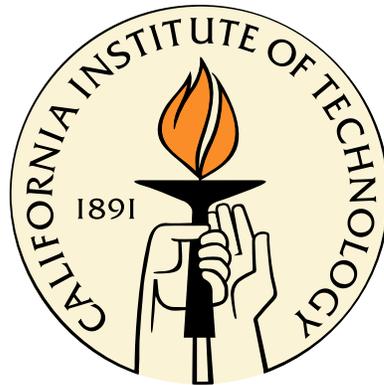
Thesis by

Michael B. McCoy

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 13, 2013)

© 2013

Michael B. McCoy

All Rights Reserved

To Anya.

Acknowledgments

Let me start by thanking my family. My father gave me an enthusiasm for knowledge and science, as well as the foolhardy belief that I really can do what I want. Without these enduring traits, the work here would not have come to fruition. My mother's love and support, and her infinite patience, let me become much more than I thought possible. My wife, Anya, gave me constant support and encouragement during my graduate studies; I could not have succeeded without her.

Second, I want to thank my friend and advisor Joel Tropp for his conscientious attention to my intellectual development. His broad perspective on research extends to many other facets of life, and this provided both confidence and freedom for me to choose my path both at Caltech and beyond. I have seen Joel extend his support and guidance generously to all of his students, which sets an aspirational standard for my own development as a mentor. Thank you.

Thanks to Alex Gittens for letting me pester him with innumerable, mostly inane, questions. These interactions provided a welcome release for the isolation that inevitably accompanies mathematical pondering. It is good to have a friend like you around. My officemate, Keenan Crane, deserves similar thanks for serving as a sounding board for half-baked ideas; his sharp wit also makes him an extraordinary friend.

To my good friends Andrew & Kelsey Homyk, and their wonderful children Natalie & Ashton, thanks for the many enjoyable evenings. My newest friends, Cole & Robin DeForest and Jessica Pfeilsticker, have made the last year spectacular. John Bruer and Trevor Fowler provided many nights of good cheer on the lawn of the Athenaeum. Guillaume Bouyt greatly enriched my first year here with his inquisitive spirit and ardent French pride.

There are many, many more people who have influenced me during my time in Pasadena. Juhwan Yoo's tenacity and persistence have opened up promising directions for my future, and I am glad to call him a friend. Thanks to Richard Chen, Simon Chien, and Brendan Farrell for many recent conversations, and to Tomek Tyranowski and Patrick Sanan for many older ones when we were just starting out. Thanks also to my food buddy Randy Clemens for his enviable prose, and to Mike & Andrea Reinhardt for their passion for life and wonderful spirits.

The administrators in the CMS department make the life of all of the students extraordinary. I personally owe many thanks to Sydney Garstang, Sheila Shull, Maria Lopez, and Jeri Chittum for their diligent and friendly help over the past six years.

Finally, I gratefully acknowledge the financial sources, under Joel's aegis, in support of my education, including ONR awards N00014-08-1-0883 and N00014-11-1002, AFOSR award FA9550-09-1-0643, and funds from a Sloan Research Fellowship.

Abstract

Demixing is the task of identifying multiple signals given only their sum and prior information about their structures. Examples of demixing problems include (i) separating a signal that is sparse with respect to one basis from a signal that is sparse with respect to a second basis; (ii) decomposing an observed matrix into low-rank and sparse components; and (iii) identifying a binary codeword with impulsive corruptions. This thesis describes and analyzes a convex optimization framework for solving an array of demixing problems.

Our framework includes a random orientation model for the constituent signals that ensures the structures are incoherent. This work introduces a summary parameter, the *statistical dimension*, that reflects the intrinsic complexity of a signal. The main result indicates that the difficulty of demixing under this random model depends only on the total complexity of the constituent signals involved: demixing succeeds with high probability when the sum of the complexities is less than the ambient dimension; otherwise, it fails with high probability.

The fact that a *phase transition* between success and failure occurs in demixing is a consequence of a new inequality in conic integral geometry. Roughly speaking, this inequality asserts that a convex cone behaves like a subspace whose dimension is equal to the statistical dimension of the cone. When combined with a geometric optimality condition for demixing, this inequality provides precise quantitative information about the phase transition, including the location and width of the transition region.

Contents

Acknowledgments	iv
Abstract	vi
1 Introduction	1
1.1 Demixing archetype	1
1.1.1 A mixed signal model	2
1.1.2 Constrained MCA demixing procedure	4
1.1.3 A probabilistic characterization of MCA	4
1.2 A recipe for demixing	6
1.2.1 Structured signals and atomic gauges	6
1.2.2 Formulating convex demixing methods	9
1.2.2.1 Multiple demixing	10
1.2.2.2 Compressed demixing	10
1.2.2.3 Lagrangian counterparts	11
1.2.3 The random alignment model of incoherence	12
1.3 Contributions	13
1.4 Outline	14
2 A survey of the literature	16
2.1 Mixed signal models	16
2.1.1 Moving beyond sparsity	18
2.2 Methods for demixing signals	18
2.2.1 Convex demixing methods	19

2.2.2	Greedy demixing methods	22
2.3	Analyses of demixing	23
2.3.1	Sparse + sparse demixing	24
2.3.2	Demixing beyond sparsity	27
2.3.3	Linear inverse problems	28
2.3.3.1	Asymptotic polytope angle computations	29
2.3.3.2	Gaussian width analyses	30
3	Mathematical preliminaries	33
3.1	Conventions	33
3.2	Convex cones	37
3.2.1	Projections and distances	40
3.2.2	Descent cones	44
3.3	The topology of convex cones	47
4	The geometry of demixing	51
4.1	Optimality conditions for constrained demixing	52
4.2	Multiple demixing	52
4.3	Compressed demixing	54
4.3.1	Linear inverse problems	57
4.4	Proof of Theorem 4.2	58
4.5	The Lagrangian formulation	61
5	Conic integral geometry	65
5.1	Conic intrinsic volumes	65
5.2	Steiner formulas	69
5.3	Kinematic formulas	71
5.3.1	Iterated kinematics	73
5.4	Rare events	76
5.4.1	Cone intersections	76
5.4.2	Closure issues	78

6	The statistical dimension	81
6.1	Basic properties	81
6.2	Explicit computations	87
6.2.1	Subspaces	87
6.2.2	Self-dual cones	88
6.2.3	Circular cones	89
6.2.4	Descent cones of the ℓ_∞ norm	93
6.2.5	A recipe for general descent cones	95
6.2.6	Descent cones of the ℓ_1 norm	96
6.2.7	Descent cones of the Schatten 1-norm	101
6.3	Relationship to the Gaussian width	105
7	Concentration of intrinsic volumes	109
7.1	Parallels with Euclidean integral geometry	110
7.2	Proof of Theorem 7.1	112
7.3	Proof of technical propositions	114
7.3.1	Concentration of projections onto cones	114
7.3.2	An inequality for chi-square random variables	117
8	Kinematic consequences	120
8.1	Iterated approximate kinematics	121
8.2	Proof of Theorem 8.2	122
8.2.1	Proof of technical lemmas	124
9	Success and failure of demixing	127
9.1	Transition for extended demixing	129
9.2	Proof of Theorem 9.2	132
10	Applications and numerical examples	137
10.1	Sparse + sparse	137
10.1.1	Strong bounds	142
10.1.2	Undersampled sparse + sparse	143

10.2 Sparse + sign	145
10.2.1 Strong bounds	148
10.2.2 Sparse + sparse + sign	149
10.3 Sparse + low-rank	150
10.4 Linear inverse problems	152
10.4.1 The sparse inverse problem	152
A A general Steiner formula	156
A.1 Consequences of Theorem A.1	157
A.1.1 The conic Wills functional	161
A.2 Proof of the general Steiner formula	164
A.2.1 A tiling induced by a cone	165
A.2.2 Polytope angles & intrinsic volumes	166
A.2.3 Proof of Theorem A.1	168
B Proofs of results in Chapter 10	171
B.1 Sparse + sparse demixing	171
B.1.1 Phase transition	171
B.1.2 Strong bound	173
B.1.3 Undersampled bounds	177
B.2 Sparse + sign demixing	177
B.2.1 Phase transition	177
B.2.2 Strong bounds	178
C Numerical details	180
C.1 Statistical dimension calculations	180
C.2 Experimental generalities	180
C.3 Sparse + sparse	181
C.3.1 Undersampled sparse + sparse	181
C.4 Sparse + sign	182
C.4.1 Sparse + sparse + sign	182

C.5 Sparse + low-rank	183
C.6 Linear inverse problems	183
Bibliography	185

List of Figures

1.1	Morphological component analysis signal model.	2
1.2	Empirical probability of success for constrained MCA.	6
1.3	An atomic gauge.	8
3.1	Descent and normal cones.	45
4.1	Geometry of optimality conditions.	53
4.2	Optimality conditions for linear inverse problems.	57
5.1	Intrinsic volumes of the nonnegative orthant.	67
6.1	Orthogonal decomposition of the sphere into subspheres.	90
6.2	Summary of statistical dimension computations.	95
10.1	Empirical probability of success for demixing two sparse vectors.	141
10.2	Empirical success of undersampled sparse + sparse demixing.	145
10.3	Demixing sparse + sign.	146
10.4	Empirical probability of demixing sparse + sparse + sign.	150
10.5	Empirical probability of success for demixing sparse + low-rank.	151
10.6	Empirical success of a sparse linear inverse problem.	153
10.7	Empirical success of a low-rank inverse problem.	155

List of Tables

6.1	Statistical dimension computations.	87
-----	---	----

List of Recipes

2.1	The <i>alternating direction method of multipliers</i>	20
2.2	The <i>recursive residual projection</i> method of Safar.	23
6.1	Bounding the statistical dimension of a descent cone.	97

Chapter 1

Introduction

Demixing is the problem of identifying multiple informative signals from a single observation of their superposition. Demixing problems arise frequently in scientific applications, including fields as diverse as geophysics [TBM79], astrophysics [SDC03], image segmentation [ESQD05], machine learning [CSPW11], robust statistics [CLMW11], and audio processing [AEJ⁺12]. The underlying observation in each of these examples is typically high-dimensional—often with millions or even billions of variables—so that effective demixing procedures require efficient computational methods.

This thesis proposes a generic computational framework for demixing based on convex optimization. By applying methods from integral geometry, we develop a general theory that characterizes the performance of our demixing approach under a probabilistic model. These results apply to a number of particular cases where the general theory predicts the outcome of numerical experiments with a high level of precision. We begin with a concrete example that illustrates the theoretical guarantees developed in this thesis.

1.1 A demixing archetype: Morphological component analysis

We introduce demixing by discussing a computational approach, called *morphological component analysis* (MCA), for identifying features in images [SDC03, ESQD05, SED05,

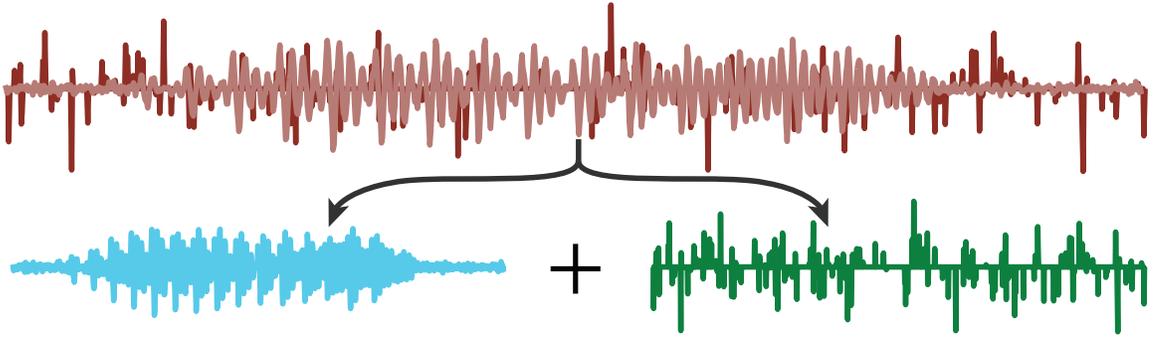


Figure 1.1: Morphological component analysis signal model. The top line shows a superposition of two signals, one consisting of a small number of tones (*blue, left*) and the other an impulsive signal (*green, right*). The goal of MCA is to identify the constituents from their superposition using knowledge about the structure in the underlying signals.

BMS06, BSFM07]. The MCA method provides a principled approach for turning prior information about the content of an image into a computational procedure that identifies its features.

The original application of MCA involves finding stars and galaxies in an astronomical image [SDC03]. There are two characteristics of stars and galaxies that make it possible to distinguish them from one another in an image of the night sky. First, stars and galaxies do not look like one another, but rather have distinct shapes or morphologies. For example, stars typically appear as localized bright points, while galaxies exhibit filamentary structure. We therefore say that stars and galaxies appear *incoherent*. The second crucial feature of stars and galaxies is their *sparsity*: there are not so many stars or galaxies in the image that they completely obscure one another. The notions of sparsity and incoherence are central to the theoretical understanding of MCA.

1.1.1 A mixed signal model

We now describe a mathematical model for signals that arise in MCA. Suppose we observe a vector $\mathbf{z}_0 \in \mathbb{R}^d$ that consists of the superposition of two structured signals:

$$\mathbf{z}_0 = \mathbf{A}\mathbf{x}^\natural + \mathbf{B}\mathbf{y}^\natural \in \mathbb{R}^d. \quad (1.1)$$

We assume that the matrices \mathbf{A} and \mathbf{B} are known, and the columns of each one encode a different type of the feature we expect to see in our image. The natural symbol \natural indicates that the unknown vectors \mathbf{x}^\natural and \mathbf{y}^\natural encode the “natural structure” we seek, while the subscripted zero reminds us that the observation \mathbf{z}_0 is the only signal to which we have access. The constituents \mathbf{x}^\natural and \mathbf{y}^\natural determine the content of the signal by selecting the columns of \mathbf{A} and \mathbf{B} that appear in \mathbf{z}_0 . We say that \mathbf{A} and \mathbf{B} are *incoherent* when the columns of \mathbf{A} are weakly correlated with the columns of \mathbf{B} , while the structures are *sparse* when \mathbf{x}^\natural and \mathbf{y}^\natural have few nonzero components.

The MCA model (1.1) appears frequently in applications. In astronomical imaging, the elements of the vector \mathbf{z}_0 correspond to pixels in an image. The columns of \mathbf{A} may consist of localized bright spots, while the columns of \mathbf{B} have wispy structures. Then the nonzero elements of \mathbf{x}^\natural determine the location of stars, while the nonzero elements of \mathbf{y}^\natural encode features of galaxies [SDC03].

Another example occurs where the signal is the sum of a small number of pure tones and impulsive noise (Figure 1.1), which models the pops and clicks that occur during the playback of music from phonograph records [SKPB12]. In this case, the columns of \mathbf{A} consist of sinusoids, while the columns of \mathbf{B} are impulse functions with very localized support. The nonzero elements of \mathbf{x}^\natural determine the frequency content of the signal, while the elements of \mathbf{y}^\natural determine the location of the pops.

Let us make a few simplifying assumptions. First, we take \mathbf{A} and \mathbf{B} to be orthonormal matrices. By a change of basis, it engenders no further loss to assume that $\mathbf{A} = \mathbf{I}$, the identity matrix. Then the observation takes the form

$$\mathbf{z}_0 = \mathbf{x}^\natural + \mathbf{Q}\mathbf{y}^\natural \tag{1.2}$$

for a known orthogonal matrix \mathbf{Q} . The restriction to orthogonal matrices is a common modeling assumption [DH01, ESQD05, SED05, HB12].

Rather than considering specific choices of \mathbf{Q} , we model incoherence by drawing \mathbf{Q} uniformly at random from the Stiefel manifold O_d of $d \times d$ orthogonal matrices. This assumption ensures incoherence because the two signals are oriented generically relative

to one another, and it appears fairly often in the context of MCA [DH01, ESQD05].

We quantify the complexity of the constituent signals in terms of their sparsity:

$$s_x := \text{nnz}(\mathbf{x}^{\natural}) \quad \text{and} \quad s_y := \text{nnz}(\mathbf{y}^{\natural}),$$

where the operator $\text{nnz}(\cdot)$ returns the number of nonzero elements of a vector. In incoherent MCA, the sparsity parameters s_x and s_y turn out to be the controlling factor that determine the identifiability of \mathbf{x}^{\natural} and \mathbf{y}^{\natural} from the observation \mathbf{z}_0 .

1.1.2 Constrained MCA demixing procedure

Morphological component analysis provides a computationally attractive method for determining the pair $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ given the superimposed observation \mathbf{z}_0 and the basis \mathbf{Q} defined by (1.2). This technique makes use of the heuristic that the ℓ_1 norm $\|\mathbf{x}\|_{\ell_1} := \sum_{i=1}^d |x_i|$ reflects the number of nonzero elements in the vector \mathbf{x} [DS89].

Assume that we have access to the side information $\alpha := \|\mathbf{y}^{\natural}\|_{\ell_1}$. A natural approach to recovering solving the demixing problem is to solve the convex program

$$\text{minimize} \quad \|\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\mathbf{y}\|_{\ell_1} \leq \alpha \quad \text{and} \quad \mathbf{x} + \mathbf{Q}\mathbf{y} = \mathbf{z}_0, \quad (1.3)$$

where the decision variables are $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. We call this approach *constrained MCA*, and we say that it succeeds if the pair $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ forms the unique optimal point of (1.3). Because the constrained MCA approach is a convex optimization program—indeed, it can even be framed as a linear program—it offers a computationally tractable approach for demixing the superimposed observation (1.2).

1.1.3 A probabilistic characterization of MCA

The analysis in this thesis provides a statistical characterization of the performance of MCA in terms of the sparsity levels s_x and s_y of the constituent signals \mathbf{x}^{\natural} and \mathbf{y}^{\natural} .

Theorem 1.1 (Success and failure of MCA). *There is a function $\psi_{\ell_1} : (0, 1) \rightarrow (0, 1)$ that generates a threshold curve $\Omega \subset (0, 1)^2$ given by*

$$\Omega := \{(\rho_x, \rho_y) : \psi_{\ell_1}(\rho_x) + \psi_{\ell_1}(\rho_y) = 1\}.$$

The curve Ω partitions the unit square into a success and failure region for MCA:

$$\begin{aligned} \left(\frac{s_x}{d}, \frac{s_y}{d}\right) \text{ lies strictly below } \Omega &\implies (1.3) \text{ succeeds with probability } \rightarrow 1 \\ \left(\frac{s_x}{d}, \frac{s_y}{d}\right) \text{ lies strictly above } \Omega &\implies (1.3) \text{ succeeds with probability } \rightarrow 0, \end{aligned}$$

where the limits are taken as the ambient dimension $d \rightarrow \infty$. At this point, the dependence of s_x and s_y on the ambient dimension d should be interpreted heuristically.

The proof appears in Section 10.1, where we restate Theorem 1.1 and state the growth regime rigorously. The function ψ_{ℓ_1} is defined in Proposition 6.14, and it is computable with standard numerical techniques (Appendix C.1).

When the ambient dimension d is large, Theorem 1.1 suggests that the probability of success of MCA (1.3) shifts from high to low near the region where the sparsity (s_x, s_y) satisfies

$$\psi_{\ell_1}(s_x/d) + \psi_{\ell_1}(s_y/d) = 1. \tag{1.4}$$

In other words, Theorem 1.1 indicates that there is a *phase transition* between the success and failure of demixing (1.3) as the sparsity pair (s_x, s_y) passes through some critical point.

Figure 1.2 presents the results of a numerical experiment designed to estimate the probability that constrained MCA succeeds. The experiment is conducted in dimension $d = 100$. For each sparsity level $s_x, s_y \in \{1, \dots, d - 1\}$, we compute the empirical probability of success of (1.3) over 25 independent trials. The yellow curve, defined by (1.4), accurately predicts the empirical 50% success region for this experiment. (Further numerical details are available in Appendix C.1.)

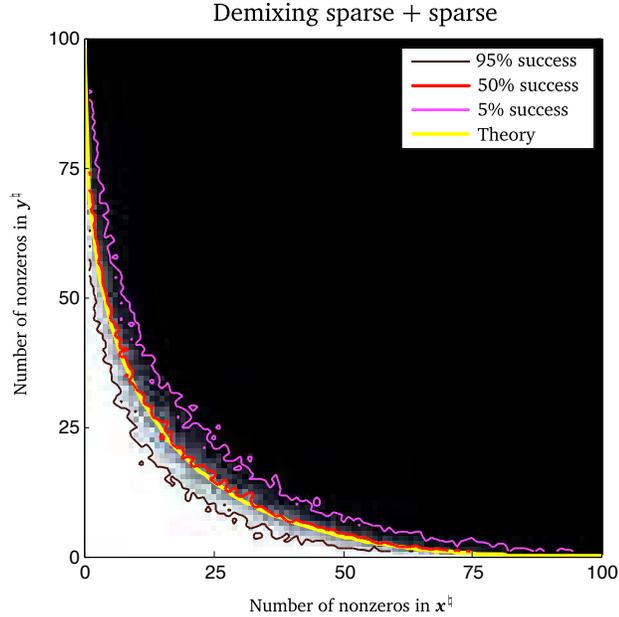


Figure 1.2: Empirical probability of success for constrained MCA. The ambient dimension is $d = 100$, and the sparsity levels vary from 1 to 100. The colormap presents the empirical probability that constrained MCA (1.3) succeeds over 25 trials. White pixels signify complete success, the black pixels represent total failures, and the gray transition region indicates a mix of successes and failures. The figure also displays the 95%, 50% and 5% success isoclines. Our theory tells us that the location of the transition region occurs (asymptotically) at the yellow curve (1.4) that lies under the 50% isocline. The details for this experiment appear in Appendix C.3.

1.2 A recipe for demixing

In this section, we present a general framework for constructing convex demixing methods. This approach develops demixing procedures arise by combining structure-inducing convex regularizers in a sensible way. Before describing the generic recipe in Section 1.2.2, we introduce the building blocks of demixing procedures, the *atomic gauges*.

1.2.1 Structured signals and atomic gauges

Natural signals often lie in a structured family whose degrees of freedom are much smaller than the ambient dimension of the signal. An image of the night sky, for example, consists of a relatively small number of points over a black background. The degrees of

freedom in such astronomical images are roughly proportional to the number of stars in the field of view, not the total number of pixels in the image.

To turn this intuition into a mathematical model, we consider natural signals that are formed by the positive sum of a few members of a known *atomic set* $\mathcal{A} \subset \mathbb{R}^d$. In the case of sparse vectors, the atomic set $\mathcal{A} = \{\pm \mathbf{e}_i \in \mathbb{R}^d : i = 1, \dots, d-1, d\}$, where \mathbf{e}_i is the i th standard basis vector. A sparse vector \mathbf{x} takes the form

$$\mathbf{x} = \sum_{i=1}^d \pm a_i \mathbf{e}_i.$$

where $a_i = 0$ for most indices i . The atomic set of one-sparse vectors comes with a natural convex function that approximates sparsity of \mathbf{x} :

$$\|\mathbf{x}\|_{\ell_1} = \sum_{i=1}^d |a_i|.$$

When the magnitude of the nonzero a_i are all comparable in magnitude, the ℓ_1 norm is proportional to the total sparsity of \mathbf{x}_0 . This well-established heuristic [CDS99, CT05, Don06a] is but one of a large number of convex penalty functions known as *atomic gauges* that measure the complexity of vectors with respect to other atomic sets [DT96, Tem03, CRPW12].

Definition 1.2 (Atomic gauge). The *atomic gauge* $f_{\mathcal{A}}(\mathbf{x})$ of a vector $\mathbf{x} \in \mathbb{R}^d$ with respect to a set $\mathcal{A} \subset \mathbb{R}^d$ is defined by

$$f_{\mathcal{A}}(\mathbf{x}) := \inf\{\lambda > 0 : \mathbf{x} \in \overline{\text{conv}}(\mathcal{A})\},$$

where $\overline{\text{conv}}(\mathcal{A})$ is the closure of the convex hull of the atomic set \mathcal{A} . We illustrate this definition in Figure 1.3.

Atomic gauges have a natural interpretation as complexity measures of atomic signals. Indeed, the atomic gauge $f_{\mathcal{A}}$ is the largest Minkowski gauge¹ such that $f_{\mathcal{A}}(\mathbf{a}) \leq 1$ for all atoms $\mathbf{a} \in \mathcal{A}$. In other words, an atomic gauge returns the maximum possible penalty

¹A Minkowski gauge is a nonnegative, positively homogeneous proper convex function.

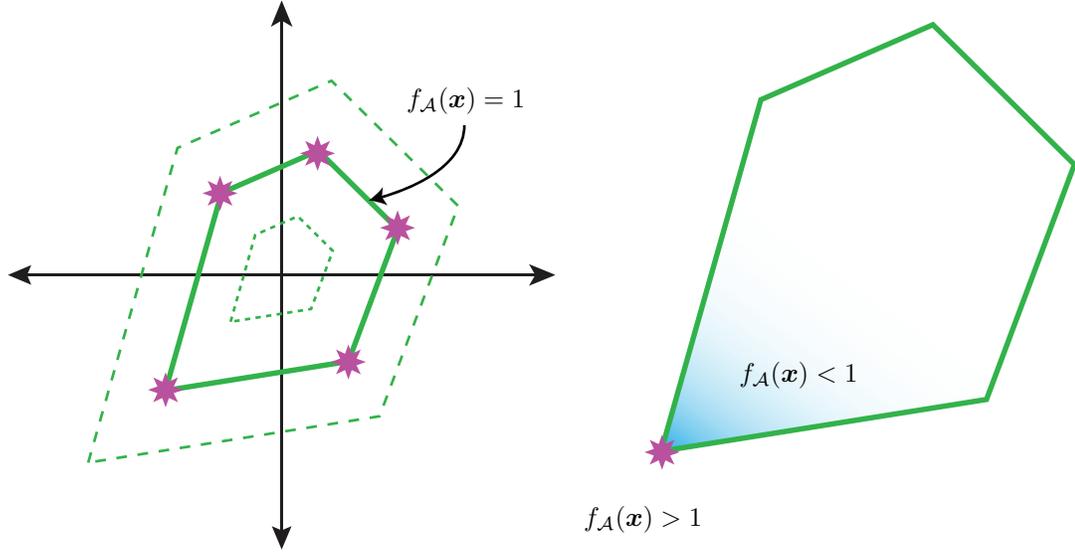


Figure 1.3: An atomic gauge. [Left] Our atomic set \mathcal{A} consists of five atoms (stars). The “unit ball” of the atomic gauge $f_{\mathcal{A}}$ is the closed convex hull of \mathcal{A} (heavy line). Other level sets (dashed lines) of the gauge are dilations of the unit ball. [Right] At an atom (star), the unit ball of $f_{\mathcal{A}}$ tends to have sharp corners. Most perturbations away from this atom increase the value of $f_{\mathcal{A}}$, so the atomic gauge tends to penalize complex atomic signals.

for deviation away from an atomic set while maintaining convexity and homogeneity, features that are important for computational efficiency and scale-invariance.

Atomic gauges are ubiquitous in the literature on inverse problems. Some common examples include

- *The ℓ_1 norm.* The ℓ_1 norm $\|\mathbf{x}\|_{\ell_1} := \sum_{i=1}^d |x_i|$ is the atomic gauge associated with the set $\{\pm \mathbf{e}_i\} \subset \mathbb{R}^d$ of signed standard basis vectors, and is widely used to promote sparsity [CDS99, CRT06a, Don06a]. The ℓ_1 norm extends to matrices as $\|\mathbf{X}\|_{\ell_1} = \sum_{i,j} |X_{ij}|$ and reflects the entrywise sparsity of a matrix [CSPW09, CPW10, CSPW11].
- *The ℓ_∞ norm.* The ℓ_∞ norm $\|\mathbf{x}\|_{\ell_\infty} := \max_{i=1,\dots,d} |x_i|$ is the atomic gauge associated with the set $\{\pm 1\}^d$ of sign vectors. We can use this norm to demix binary codewords (Section 10.2); see also [DT10a, CRPW12, MR11]. The ℓ_∞ norm has a natural extension to matrices: $\|\mathbf{X}\|_{\ell_\infty} := \sum_{i,j} |X_{ij}|$, which is the atomic gauge generated by sign matrices. It is the gauge associated to the set of rank-one matrices, and it tends to penalize the rank of a matrix [Faz02, RFP10].

Note that the Schatten 1-norm does not have a finite atomic set.

- *The operator norm.* The matrix operator norm is the maximum singular value of a matrix. On the space $\mathbb{R}^{n \times n}$ of square matrices, the operator norm is the atomic gauge associated with matrices with orthogonal rows or columns, and so it is suited for situations involving a search for orthogonal matrices [CRPW12, Prop. 3.13]. Like the Schatten 1-norm, the operator norm does not have a finite atomic set.

This list is certainly not comprehensive. A number of additional atomic gauges are discussed in [CRPW12, Sec. 2.2].

1.2.2 Formulating convex demixing methods

Atomic norms offer a principled approach for constructing structure-inducing convex functions. In this section, we describe a generic method for combining these functions into a convex program that is useful for demixing a superposition of signals. Our method generalizes an approach to demixing that appears in a number of works. For some representative examples, see [ESQD05, BSFM07, JRSR10, CSPW11].

The basic signal model mirrors that of MCA. We assume that we are given an observation consisting of the superposition of two structured signals:

$$\mathbf{z}_0 = \mathbf{x}^{\natural} + U\mathbf{y}^{\natural}, \tag{1.5}$$

where U is a known orthogonal matrix but the pair $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ is unknown. The basis U is included to model the relative orientation of the structures, and it also provides a convenient proxy for incoherence. We seek a method to demix the constituent pair $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ from the observation \mathbf{z}_0 .

Let f and g be convex complexity measures—such as atomic gauges—associated with the structures we expect to find in \mathbf{x}^{\natural} and \mathbf{y}^{\natural} . Assume further that we have access to the side information $\alpha := g(\mathbf{y}^{\natural})$. We combine these ingredients into the following

convex demixing method:

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } g(\mathbf{y}) \leq \alpha \quad \text{and} \quad \mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{z}_0, \quad (1.6)$$

where the optimization variables are $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Equation (1.6) is a convex optimization program because f and g are convex functions and the consistency constraint $\mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{z}_0$ is affine. We say that the demixing method (1.6) *succeeds* when $(\mathbf{x}^\natural, \mathbf{y}^\natural)$ is the unique optimal point of (1.6); otherwise it *fails*.

1.2.2.1 Multiple demixing

Our basic signal model involves the superposition of just two signals. A natural extension involves an observation that consists of a superposition of n signals:

$$\mathbf{z}_0 = \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i^\natural, \quad (1.7)$$

where \mathbf{U}_i are known orthogonal matrices that encode the relative orientation of the constituent vectors \mathbf{x}_i^\natural . Given convex regularizers f_i for each $i \in \{1, \dots, n-1, n\}$ and the side information $\alpha_i := f_i(\mathbf{x}_i^\natural)$ for every $i \in \{2, \dots, n-1, n\}$, we may extend the method (1.6) to

$$\begin{aligned} & \text{minimize } f_1(\mathbf{x}_1) \\ & \text{subject to } f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for } i \in \{2, \dots, n-1, n\} \quad \text{and} \quad \mathbf{z}_0 = \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i, \end{aligned} \quad (1.8)$$

where our decision variables are now $\mathbf{x}_i \in \mathbb{R}^d$ for $i \in \{1, \dots, n-1, n\}$. We analyze this model in Section 4.2.

1.2.2.2 Compressed demixing

We may further extend multiple demixing to the *undersampled* observation of superimposed signals [WGMM12]. In this setup, the observation is given by

$$\mathbf{z}_0 = \mathbf{A} \left(\sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i^\natural \right) \in \mathbb{R}^m, \quad (1.9)$$

where the constituents \mathbf{x}_i^{\natural} and rotations \mathbf{U}_i are as in the multiple demixing model (1.7) and the matrix $\mathbf{A} \in \mathbb{R}^{m \times d}$ only reveals a linear image of the superposition.

Given structure-inducing convex functions f_i and the side information α_i as in multiple demixing, we solve the compressive demixing procedure

$$\begin{aligned} & \text{minimize} && f_1(\mathbf{x}_1) \\ & \text{subject to} && f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for } i \in \{2, \dots, n-1, n\} \quad \text{and} \quad \mathbf{A}(\mathbf{x} + \mathbf{U}\mathbf{y}) = \mathbf{z}_0. \end{aligned} \tag{1.10}$$

The undersampled model (1.9) provides an interesting description of a number of realistic situations, and it is a generalization of multiple demixing because we recover (1.7) with the choice $\mathbf{A} = \mathbf{I}$. Happily, the addition of a slack variable reduces the analysis of compressive demixing (1.10) to a special case of the multiple demixing approach (1.8). See Section 4.3 for this reduction.

1.2.2.3 Lagrangian counterparts

In many practical settings, we do not have access to the side information $\alpha := g(\mathbf{y}^{\natural})$. In this case, the Lagrangian form of (1.6) is more appropriate:

$$\text{minimize} \quad f(\mathbf{x}) + \lambda g(\mathbf{y}) \quad \text{subject to} \quad \mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{z}_0, \tag{1.11}$$

where $\lambda > 0$ is a parameter that negotiates a tradeoff between the relative importance of the regularizers. The theory of Lagrange duality reveals a close connection between the penalty formulation (1.11) and its constrained cousin (1.6). Roughly speaking, knowledge of the side information is as powerful as knowing the best possible choice of λ . We leverage our geometric analysis of demixing to demonstrate this equivalence in Section 4.5. Cross validation provides a principled approach to search for the best possible λ [BDB07], but this may be computationally expensive to implement.

1.2.3 The random alignment model of incoherence

In this section, we introduce a random model that ensures that the structures we are attempting to demix are *incoherent* with one another. Incoherence captures the idea that signals are uncorrelated with one another. Indeed, demixing is impossible when the structures in the constituent signals look alike. If, for example, we observe $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{y}^{\natural}$, where both \mathbf{x}^{\natural} and \mathbf{y}^{\natural} are sparse, it is impossible consistently assign the nonzero elements of \mathbf{z}_0 to the constituents \mathbf{x}^{\natural} and \mathbf{y}^{\natural} . On the other hand, if one constituent is sparse in time and the other is sparse in frequency (Figure 1.1), then the components are typically identifiable [Tro08c].

In this work, we model incoherence by assuming that the rotation $\mathbf{U} = \mathbf{Q}$, where \mathbf{Q} is drawn randomly from the invariant Haar measure on \mathcal{O}_d . The observed signal \mathbf{z}_0 then takes the form

$$\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}. \quad (1.12)$$

We call this the *random alignment model*. We expect that this model also captures much of the qualitative behavior we seen in highly incoherent situations such as time-frequency alignments.

In the case of multiple demixing, we enforce incoherence by applying an independent random rotation to each component of the signal:

$$\mathbf{z}_0 = \sum_{i=1}^n \mathbf{Q}_i \mathbf{x}_i^{\natural} \quad (1.13)$$

where each member of the tuple $(\mathbf{Q}_1, \dots, \mathbf{Q}_n)$ is a Haar-distributed rotation independent of all the others. Again, this model ensures that the constituent signals are generically oriented relative to one another, and we expect that it sheds light on highly incoherent special cases such as mutually unbiased bases [Sch60].

1.3 Contributions

This work characterizes the success and failure regimes for the demixing method (1.6), multiple demixing (1.8), and compressed demixing (1.10) under the random incoherence model of Section 1.2.3. The ultimate outcome of this research is a new understanding of the capabilities of demixing methods that we summarize as follows:

The pairs (\mathbf{x}^\natural, f) and (\mathbf{y}^\natural, g) possess intrinsic dimensionality parameters δ_x and δ_y . The demixing method (1.6) succeeds with high probability if the sum $\delta_x + \delta_y$ is slightly less than the ambient dimension d of the observation \mathbf{z}_0 . On the other hand, the demixing method (1.6) fails with high probability if $\delta_x + \delta_y$ is slightly larger than the ambient dimension of the observation \mathbf{z}_0 .

The rigorous statement of this result forms the topic of Chapter 9, where we also find a similar characterization of multiple demixing (1.8) and compressed demixing (1.10). En route to this result, we develop a number of new ideas and results. Let us summarize these developments.

- We codify a framework for solving a number of demixing problems (Section 1.2, above) and provide a characterization of the success of demixing in terms of a configuration of convex cones (Section 3).
- We introduce a summary parameter for convex cones, the *statistical dimension*, that shares many properties with the usual affine dimension (Section 6.1). Simplifying known techniques, we provide sharp calculations of several nontrivial statistical dimensions (Section 6.2).
- We prove an inequality in conic geometry that demonstrates that geometric parameters of convex cones known as *intrinsic volumes* concentrate sharply about the statistical dimension (Theorem 7.1). This result parallels classical inequalities of Euclidean convex geometry (Section 7.1).
- As a consequence of this concentration inequality, we find an *approximate kinematic formula* that characterizes the probability that cones strike in terms of the sum of their statistical dimensions (Theorems 8.1 and 8.2).

- We leverage this kinematic formula to provide precise regions of success and failure for demixing (1.6), multiple demixing (1.8), and compressed demixing (1.10) under random alignment models (Theorems 9.1 and 9.2).
- These theoretical results are applied to a number of specific demixing methods. We provide experiments to demonstrate the accuracy of our theory for these examples in a number of settings (Section 10).

Many of these ideas first appear in the joint works [MT12, ALMT13], although the results on multiple and compressive demixing appear here for the first time. Results that are primarily due to coauthors are indicated as such when they are introduced.

1.4 Outline

Chapter 2 traces the intellectual development of demixing through the literature. In Chapter 3, we introduce the mathematical preliminaries required for our demixing analysis. Many of these facts are staples of convex analysis, although Section 3.3 introduces a topology on convex cones that does not appear in standard texts.

The mathematical exposition of this thesis begins in earnest with a geometric analysis of demixing in Chapter 4. The most important outcome of this chapter is a characterization the success of demixing in terms of the configuration of convex cones.

Motivated by this conic optimality condition for demixing, Chapter 5 recalls some key results from the field of conic integral geometry. There, we introduce geometric invariants of cones known as *intrinsic volumes*. We describe kinematic formulas that provide the precise probability that randomly oriented cones intersect, expressed in terms of the intrinsic volumes. Along the way, we dispense with some finer points—such as the probability that the intersection of two cones is a subspace—that are necessary for bringing conic integral geometry to bear on our demixing models.

In Chapter 6, we introduce a new summary parameter, the *statistical dimension*, that measures the size of a convex cone. This parameter shares many formal similarities with the usual dimension of a linear subspace. We provide explicit calculations for the

statistical dimension of a number of cones in Section 6.2.

A new general fact about convex cones is identified in Chapter 7. There, we find that the sequence of intrinsic volumes of a convex cone concentrates sharply about the statistical dimension. Chapter 8 applies these results to bound the probability that two randomly oriented cones strike in terms of the statistical dimension.

We return to demixing in Chapter 9, where we leverage the approximate kinematic formulas to characterize demixing. These results form the capstone of this work. Chapter 10 provides some numerical illustrations of the accuracy of our demixing theory.

Two appendices contain details that are somewhat tangential to the presentation. Appendix A demonstrates a new general formula in conic integral geometry. Appendix C provides the details of our numerical experiments.

Chapter 2

A survey of the literature

This section traces the development of demixing from its origins in sparse approximation to the current state of the art. We break this discussion into three broad themes: *models*, *methods*, and *analyses*. The discussion of models describes *where* superimposed signals find practical application. The methods section focuses on *how* demixing problems are solved. The analysis section summarizes prior results that ascertain *when and why* demixing is possible. This thesis focuses on a new analysis of demixing, so Sections 2.1 and 2.2 below discuss only previous work. Section 2.3 describes the antecedents to our analysis and provides direct comparisons between our results and those from the literature.

2.1 Mixed signal models

The modern history of demixing can be traced to the geophysics literature of the 1970s. Taylor, Banks & McCoy [TBM79], building on earlier work¹ of Claerbout & Muir [CM73], model an observed signal \mathbf{z}_0 as the sum

$$\mathbf{z}_0 = \mathbf{W}\mathbf{x}^{\natural} + \mathbf{y}^{\natural}, \quad (2.1)$$

¹Claerbout & Muir trace the genesis of their approach to the work of Boscovich on line fitting in the 1750s; see Plackett [Pla72]. Sheynin [She73, Sec. 1.3] traces Boscovich's method back farther still to work of Bernoulli in 1734.

where \mathbf{x}^h is a sparse spike train, the matrix \mathbf{W} is a convolutional operator, and \mathbf{y}^h is some unknown noise. The sparsity pattern of \mathbf{x}^h encodes subsurface geological structure, but no structural assumptions are imposed on the noise \mathbf{y}^h .

Sums of multiple structured signals began appearing in the signal processing community in the late 1980s. The M.S. thesis of Safar [Saf88] and the independent work of Mikhael & Spanias [MS89] both advocate representing a signal as the sum of signals that are sparse in different bases. The application these authors have in mind is efficient source coding of speech and audio signals. Typical bases that arise in this field include the standard Dirac basis, as well as fast transform bases such as the FFT, DCT, Walsh–Hadamard and (later) wavelet bases. See [BM99] for a survey of this fruitful line of inquiry.

Wavelet theory in the 1990s provided a significant boost to sparse modeling. A key result of this field is that regularity of signals implies (near) sparse representations of those signals in wavelet bases (cf. [Mal09, Sec. 9]). Mallat & Zhang argued that even more succinct representations of signals become available by combining multiple wavelet bases [MZ93].

Donoho & Huo [DH01] continued this inquiry into sparse-in-multiple-bases models, and they include a *random* rotation as a model for structural incoherence. Applications for this random rotation model appear in their work, including a robust encryption scheme and a multiple-access communications protocol.

The mixed-sparsity template found further applications in image processing under the name of *morphological component analysis* [SDC03, ESQD05, SED05]. This method, described in detail in Section 1.1, treats an image as the superposition of a small number of simple constituent elements—such as stars and galaxies in an image of the night sky—and provides a computational procedure for distinguishing them.

Research on sparse demixing models continues up to this day. The work [SKPB12] describes a number of additional applications for this model, including clipped audio signals, faulty computer memory, superresolution, and signal separation. We refer to this recent work for an impressive list of applications that involve demixing sparse signals.

2.1.1 Moving beyond sparsity

Motivated largely by the success of sparse models (including, of course, the *compressed sensing* model [CT05, Don06a]), a cadre of researchers started to investigate other types of structures amenable to demixing. An exemplary approach is the rank–sparsity decomposition of Chandrasekaran et al. [CSPW09, CSPW11], which models a matrix observation $\mathbf{M}_0 \in \mathbb{R}^{m \times n}$ as the sum

$$\mathbf{M}_0 = \mathbf{L}^\natural + \mathbf{S}^\natural, \quad (2.2)$$

where \mathbf{L}^\natural has low rank and \mathbf{S}^\natural is sparse. In machine learning, the model (2.2) naturally appears in latent variable model selection [CPW10], alignment of occluded images [PGW⁺12], and scene triangulation [ZLGM11]. The rank–sparsity model also has deep links to problems in theoretical computer science—see the discussion [CS12, p. 2002] for further details.

The rank-sparsity model spurred additional interest in matrix demixing models, especially methods for robust principal component analysis [CLMW11, XCS10a, XCS10b, MT11]. Recent work includes investigations of undersampled mixed signals, with applications to partially observed graph clustering [CJSC11a] and compressed robust principal component analysis [WGMM12, CJSC13].

2.2 Methods for demixing signals

The techniques for demixing sums of structured signals fall into two categories: convex and greedy. Convex methods, which are the focus of this thesis, pose demixing as an optimization problem whose optimal point determines the demixed constituents. Greedy methods are iterative procedures that build up the constituent signals in a piecemeal fashion.

2.2.1 Convex demixing methods

Historically, convex methods were the first² computational approaches to demixing. Taylor, Banks & McCoy [TBM79] demix the signal (2.1) by solving

$$\text{minimize } \|\mathbf{x}\|_{\ell_1} + \lambda \|\mathbf{y}\|_{\ell_1} \quad \text{subject to } \mathbf{W}\mathbf{x} + \mathbf{y} = \mathbf{z}_0, \quad (2.3)$$

where $\lambda > 0$ is a tradeoff parameter whose value is to be tuned. This approach coincides precisely with the Lagrange formulation of the MCA procedure (1.3). This basic schema proved enduring, and it remains popular today in an essentially unchanged form, cf. [DH01, ESQD05, BSFM07, WM09, SKPB12, PBS13].

Naturally, new convex regularizers led to new demixing methods. The Schatten 1-norm $\|\cdot\|_{S_1}$ gained prominence in control theory for its ability to identify low-rank matrices [MP97, Faz02]. Extending the basic schema (2.3), Chandrasekaran et al. [CSPW11], demix the rank–sparsity model (2.2) with the convex program

$$\text{minimize } \|\mathbf{L}\|_{S_1} + \lambda \|\mathbf{S}\|_{\ell_1} \quad \text{subject to } \mathbf{L} + \mathbf{S} = \mathbf{M}_0. \quad (2.4)$$

The works [XCS10a, XCS10b, MT11] further extend (2.4) by replacing the matrix ℓ_1 norm with a norm that returns the sum of the Euclidean norm of the rows of \mathbf{S} in an effort to promote row-sparsity [RKD98, RKD02]. A common modification for a partially observed model replaces the equality constraint in (2.4) by an undersampled equality constraint $\mathcal{A}(\mathbf{L} + \mathbf{S}) = \mathcal{A}(\mathbf{M}_0)$, where \mathcal{A} is a linear operator on the space of matrices that maps to a lower-dimensional vector space. See, for example, [CJSC11b, CJSC11c, WGMM12].

Numerical aspects. Off-the-shelf convex optimization software, such as the CVX front-end for SeDuMi and SDPT3 [Stu99, TTT99, GB08, GB10, TTT12], provides a convenient and quick way to prototype convex demixing methods. Such general-purpose software typically use primal-dual interior-point methods that possess strong theoretical

²Indeed, the line-fitting method of Boscovich/Bernoulli is a linear program! (See footnote on page 16.)

Recipe 2.1: The alternating direction method of multipliers for (2.7). This basic scheme proves highly effective for minimizing problems involving several nondifferentiable regularizers such as the demixing problem (2.7). Choices for the tuning parameter μ and stopping criteria, as well as numerous other details that arise in practical implementations of ADMM, are discussed in the review article [BPC⁺10].

Given: The augmented Lagrangian L_μ (2.8); a tuning parameter $\mu > 0$; a stopping criterion.

1. Set the initial points $\mathbf{x}^0 \leftarrow \mathbf{y}^0 \leftarrow \mathbf{w}^0 \leftarrow \mathbf{0}$ and the counter $k \leftarrow 0$.
2. While the stopping criterion is not satisfied, do:

- (a) Update the estimate for \mathbf{x}_* :

$$\mathbf{x}^{k+1} \leftarrow \arg \min_{\mathbf{x}} L_\mu(\mathbf{x}, \mathbf{y}^k, \mathbf{w}^k) \quad (2.5)$$

- (b) Update the estimate for \mathbf{y}_* :

$$\mathbf{y}^{k+1} \leftarrow \arg \min_{\mathbf{y}} L_\mu(\mathbf{x}^{k+1}, \mathbf{y}, \mathbf{w}^k) \quad (2.6)$$

- (c) Update the multiplier: $\mathbf{w}^{k+1} \leftarrow \mathbf{w}^k + \mu(\mathbf{x}^{k+1} + \mathbf{U}\mathbf{y}^{k+1} - \mathbf{z}_0)$.
- (d) Increment $k \leftarrow k + 1$.

Return: Final estimates $\mathbf{x}^k, \mathbf{y}^k$.

guarantees of computational efficiency [NN94]. These interior-point methods are effective for small- to moderate-scale numerical investigations, and so we use CVX for all numerical experiments in this work.

For large-scale problems, however, the computational requirements of interior-point methods can still overwhelm the available resources. In these situations, iterative methods may still be able to provide low-accuracy solutions in an acceptable amount of time. For reasons explained below, the most popular iterative algorithm for solving convex demixing problems is the *alternating direction method of multipliers* (ADMM) [GM75, GM76].

To make our discussion concrete, we consider an ADMM algorithm for solving the Lagrange version of the demixing method:

$$\text{minimize } f(\mathbf{x}) + g(\mathbf{y}) \quad \text{subject to } \mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{z}_0, \quad (2.7)$$

where the decision variables are $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and \mathbf{U} is an orthogonal matrix. We may write

the constrained demixing program (1.6) in the form (2.7) using the convex indicator function defined in (3.2), so this discussion covers all two-signal demixing methods that we study in this work. The basic ingredient in the ADMM method is the augmented Lagrangian

$$L_\mu(\mathbf{x}, \mathbf{y}, \mathbf{w}) := f(\mathbf{x}) + g(\mathbf{y}) + \langle \mathbf{w}, \mathbf{x} + \mathbf{U}\mathbf{y} - \mathbf{z}_0 \rangle + \frac{\mu}{2} \|\mathbf{x} + \mathbf{U}\mathbf{y} - \mathbf{z}_0\|^2, \quad (2.8)$$

where $\mathbf{x}, \mathbf{y}, \mathbf{w} \in \mathbb{R}^d$ and $\mu > 0$ is a tuning parameter. In each iteration, the ADMM scheme minimizes the augmented Lagrangian first over \mathbf{x} and then over \mathbf{y} , holding all other parameters fixed. It then updates the multiplier \mathbf{w} and repeats the process. See Recipe 2.1 for the full procedure.

The advantage of ADMM derives chiefly from efficient formulas for the inner minimization steps (2.5) and (2.6). After some simplifications, these subproblems reduce to computing *proximity operators* of the form

$$\text{prox}_f(\mathbf{u}) = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{1}{2} \|\mathbf{x} - \mathbf{u}\|^2.$$

Proximity operators frequently possess closed form, or easily computable, solutions that make each iteration of an ADMM algorithm relatively inexpensive. For example, when f is a convex indicator function (3.2) on a closed convex set K , the proximity operator prox_f is simply the Euclidean projection onto K . A number of formulas for proximity operators are collected in [CW05, Sec. 2.6].

ADMM algorithms possess a number of additional theoretical and practical benefits. The iterates \mathbf{x}^k and \mathbf{y}^k converge to minimizers of (2.7) even when the subproblems (2.5) and (2.6) are solved inexactly [EB92]. Moreover, a straightforward extension of this method to the multiple demixing scheme (1.8) also succeeds under some mild technical conditions [HL13]. As a practical matter, many ADMM algorithms parallelize to massive-scale distributed computations [BPC⁺10]. It is no surprise that this approach is nearly ubiquitous in the demixing literature [XCS10b, CSPW11, CLMW11, PGW⁺12, CJSC13].

Other numerical approaches. Other iterative methods for solving demixing problems

include fast proximal gradient methods. These approaches are often effective for solving linear inverse problems (the $n = 1$ case of (1.10)) [BT09, BT12]. However, applying the usual proximal gradient methods to (2.7) requires the computation of a projection or proximal operator that may not have an analytic expression. Thus, the proximal gradient approaches often require solving an additional convex subproblem at each step. The ADMM approach sidesteps this difficulty by computing two independent minimizations in (2.5) and (2.6).

2.2.2 Greedy demixing methods

In real-time systems, solving convex optimization programs may not be feasible regardless of the implementation. For this reason, greedy iterative methods may be the only option for demixing when speed is paramount.

An early greedy method explicitly for the demixing problem is the *recursive residual projection* method of Safar [Saf88, Sec. 4.3.2]. This approach, described in Recipe 2.2, foreshadows later work on matching-pursuit and its numerous variants in the field of computational harmonic analysis [MZ93, PRK93, DMA97]. The more recent CoSaMP method [NT09] was a significant breakthrough in matching-pursuit methods. CoSaMP builds a greedy representation in batches instead of one element at a time as in Recipe 2.2, which provides both practical and theoretical improvements over earlier greedy methods. Note that all of these matching-pursuit approaches are applicable for models beyond demixing signals that are sparse in different bases, but this topic is tangential to our present discussion.

There is some work on greedy demixing algorithms beyond sparse models. A notable example is the SpaRCS method for low-rank and sparse matrix identification [WSB11] which draws inspiration from both the CoSaMP method for sparse approximation and the ADMiRa greedy approach to matrix completion [LB10]. This greedy method is orders of magnitude faster than interior-point methods for the rank-sparsity demixing program (2.4), and it only performs slightly worse than the full convex solution in practice.

Recipe 2.2: The *recursive residual projection* method of Safar [Saf88]. The main idea in this method is to compute, at each stage, a “best” atom to add to the representation of the signal and update the representation. This approach is identical to the widely known matching-pursuit method of Mallat & Zhang [MZ93]. While matching-pursuit is applicable to sparse approximation problems beyond demixing, we discuss this particular method because it was developed by Safar explicitly for mixed signal models.

Given: An observation \mathbf{z}_0 ; two orthogonal bases \mathbf{U}, \mathbf{V} ; an integer $m \geq 1$.

1. Set the current residual $\mathbf{r}^{(0)} \leftarrow \mathbf{z}_0$ and current coefficients $\mathbf{c}^{(0)} \leftarrow \mathbf{0} \in \mathbb{R}^{2d}$.
2. For $i = 1, \dots, m - 1, m$, do:

- (a) Compute the basis expansions:

$$\mathbf{e} \leftarrow [\mathbf{U} \ \mathbf{V}]^t \mathbf{r}^{(i)} \in \mathbb{R}^{2d}.$$

- (b) Determine the index of the dominant component:

$$j_* \leftarrow \arg \max_{j=1, \dots, 2d-1, 2d} |e_j|.$$

- (c) Update the coefficients:

$$c_{j_*}^{(i+1)} \leftarrow c_{j_*}^{(i)} + e_{j_*}.$$

- (d) Update the residual:

$$\mathbf{r}^{(i+1)} \leftarrow \mathbf{z}_0 - [\mathbf{U} \ \mathbf{V}] \mathbf{c}^{(i+1)}.$$

Return the coefficients \mathbf{c} .

Hegde & Baraniuk [HB12] recently developed a greedy algorithm for demixing signals of the form (1.9). They assume that the constituent signals come from known manifolds such as manifold of sparse vectors or low-rank matrices. Their algorithm recovers the constituents by iterative projections onto these manifolds, and possesses provable recovery guarantees.

2.3 Analyses of demixing

The analysis of this thesis is related to a large amount of prior work. We start by reviewing previous results for demixing methods, first for the sparse + sparse model (1.1) and then for analyses that incorporate structures beyond sparse vectors such as low-rank matrices. In Section 2.3.3.1, we trace the provenance of the geometric probability tools

that form the technical core of this work, and we describe some prior applications of these concepts to the analysis of convex optimization. Section 2.3.3.2 describes an approach to understanding linear inverse problems based on Gaussian process inequalities of Gordon [Gor85, Gor87, Gor88]. We conclude each subsection with a comparison of our results in relation to the prior art.

2.3.1 Sparse + sparse demixing

The work of Donoho & Huo [DH01] provided an important analysis of a sparse demixing program of the form (2.3). They defined the *mutual coherence* $\mu(\mathbf{U}, \mathbf{V})$ between two orthogonal bases $\mathbf{U}, \mathbf{V} \in \mathcal{O}_d$ by

$$\mu(\mathbf{U}, \mathbf{V}) := \|\mathbf{U}^t \mathbf{V}\|_{\ell_\infty} = \sup_{i,j=1,\dots,d} |\langle \mathbf{u}_i, \mathbf{v}_j \rangle|,$$

where \mathbf{u}_i and \mathbf{v}_j range over the columns of \mathbf{U} and \mathbf{V} . The bases are *incoherent* when the columns of \mathbf{U} are not strongly correlated with any column of \mathbf{V} , i.e., when $\mu(\mathbf{U}, \mathbf{V})$ is small. Refining the uncertainty principle analysis of [DS89], Donoho & Huo prove that an observation $\mathbf{z}_0 = \mathbf{U}\mathbf{x}^\natural + \mathbf{V}\mathbf{y}^\natural$ can be demixed using ℓ_1 minimization of the form (2.3), provided that

$$\text{nnz}(\mathbf{x}^\natural) + \text{nnz}(\mathbf{y}^\natural) < \frac{1}{2} \left(1 + \frac{1}{\mu(\mathbf{U}, \mathbf{V})} \right) \leq \frac{1}{2} (1 + \sqrt{d}), \quad (2.9)$$

where $\text{nnz}(\cdot)$ returns the number of nonzero entries of a vector [DH01, Thm. VII.1]. This result was improved slightly by Elad & Bruckstein [EB02] using similar methods. In the case where one of the bases is a *random* orthogonal matrix, this result operates in the regime where the total sparsity satisfies $\text{nnz}(\mathbf{x}^\natural) + \text{nnz}(\mathbf{y}^\natural) = O(\sqrt{d/\log(d)})$. In other words, the total degrees of freedom allowed in the vectors \mathbf{x}^\natural and \mathbf{y}^\natural is on the order of the square-root of the ambient dimension in every case, and is somewhat more restricted when the sparsity bases are random.

Tropp [Tro04] demonstrated that a variant of the greedy matching-pursuit algorithm recovers sparse vectors under condition (2.9). Further results of Tropp & Gilbert [TG07],

Tropp, Gilbert & Strauss [TGS06], and Needell & Tropp [NT09] demonstrated that it is possible to construct greedy demixing and sparse approximation algorithms that possess theoretical guarantees qualitatively similar to those available for convex methods.

Two recent works use incoherence parameters to study convex demixing methods for sparse signals when the support sets of the constituent signals are partially known. Studer et al. [SKPB12] provided incoherence guarantees for sparse + sparse demixing with recovery bounds similar to those obtained by Donoho & Huo (2.9). Pope, Bracher, & Studer [PBS13], building on the work of Tropp [Tro08a, Tro08b], improve these incoherence-based results considerably by considering sparse signals drawn from random models. In this case, the degrees of freedom in the observations need only be smaller than the ambient dimension by a logarithmic factor.

In 2009, Wright & Ma [WM09] developed recovery conditions for the sparse demixing program (2.3) under a unique set of assumptions. They assumed that the matrix W has highly correlated columns and required that the signs of the nonzero elements of the constituent signals are random. Their results show that demixing succeeds even when one of the constituents is near completely dense, provided the other is near completely sparse. While this result allows both constituents to possess sparsity levels that are proportional to the ambient dimension, the constants are far from optimal. Nguyen & Tran recently extended these results to a more standard model [NT13].

Finally, we mention that Hegde & Baraniuk consider demixing signals from incoherent manifolds via a nonconvex iterative algorithm [HB12]. In the case of two sparse signals, their success conditions closely resemble (2.9), but their methods are applicable to a wider class of signals than have been considered in the convex framework.

Comparison with this work. In Section 10.1, we analyze a specific model for demixing the sum of two sparse vectors. Our results rigorously characterize the precise recovery threshold for the program (1.3) under the random alignment model. In particular, we show that a phase transition between success and failure of (1.3) exists, and that it is possible to demix sparse vectors when the total sparsity is proportional to the dimension. This significantly improves on the previous best bounds of [WM09]

and [PBS13].

We also demonstrate a phenomenon that appears to be unheard of in the literature, a *strong bound* for demixing when the total sparsity is proportional to the dimension. Our results indicate the existence of a number $\tau > 0$ such that, with high probability in the random rotation \mathbf{Q} , the program (1.3) will succeed at demixing

$$\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$$

for *all* sufficiently sparse vectors \mathbf{x}^{\natural} and \mathbf{y}^{\natural} such that $\text{nnz}(\mathbf{x}^{\natural}) < \tau d$ and $\text{nnz}(\mathbf{y}^{\natural}) \leq \tau d$. To put this result another way, we draw the matrix \mathbf{Q} once and fix it for all time. Then with high probability, the program (1.3) will demix every possible observation \mathbf{z}_0 of the form above, so long as the sparsity in \mathbf{x}^{\natural} and \mathbf{y}^{\natural} is sufficiently small, but still proportional to the ambient dimension.

While the bound (2.9) of Donoho & Huo also provides such strong guarantees, it inherently limits the total sparsity to order \sqrt{d} . Indeed, a small incoherence parameter $\mu(\mathbf{U}, \mathbf{V})$ can *never* guarantee strong demixing bounds where the sparsity is proportional to the dimension. For example, the coherence $\mu(\mathbf{I}, \mathbf{F}) = d^{-1/2}$ between the identity matrix \mathbf{I} and the Fourier transform matrix \mathbf{F} is minimal. However, equispaced spike trains have representations with $O(\sqrt{d})$ nonzeros in both the identity and Fourier bases, which creates a fundamental identifiability problem [DS89, Sec. 7.2]. We expect that similar counterexamples are available for real case by replacing the Fourier transform with the Walsh–Hadamard basis

$$\mathbf{H} := \frac{1}{2^{k/2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes k}$$

where the exponent $\otimes k$ denotes the k th order Kronecker product. (The existence of an example of spike trains that are highly sparse in the Walsh–Hadamard basis is attributed to [EB02] by Donoho & Elad [DE03, pp. 2198], but we were unable to locate the demonstration in [EB02].) Our strong bound shows that, under generic assumptions, a convex program is capable of recovering the sum of two sparse signals in different bases

even when the sparsity is proportional to the ambient dimension. See Theorem 10.3 for the details.

2.3.2 Demixing beyond sparsity

Chandrasekaran et al. [CSPW11] provided an influential analysis of the rank-sparsity demixing model (2.2). Their approach is based on the concept of *incoherent manifolds*; let us sketch the key ideas. A rank- r matrix \mathbf{L}^\natural lies on a manifold \mathcal{L} of all rank- r matrices, and the local behavior of the Schatten 1-norm in a neighborhood of \mathbf{L}^\natural can be understood in terms of the tangent space $T_{\mathcal{L}}(\mathbf{L}^\natural)$ of the manifold \mathcal{L} at the matrix \mathbf{L}^\natural . Similarly, the ℓ_1 norm of a sparse matrix \mathbf{S}^\natural can be locally characterized by the tangent space $T_{\mathcal{S}}(\mathbf{S}^\natural)$ of a manifold \mathcal{S} of sparse matrices at \mathbf{S}^\natural .

When the tangent spaces $T_{\mathcal{L}}(\mathbf{L}^\natural)$ and $T_{\mathcal{S}}(\mathbf{S}^\natural)$ satisfy an incoherence property that ensures that they are transverse to one another, Chandrasekaran et al. show that (2.4) successfully demixes \mathbf{L}^\natural and \mathbf{S}^\natural from the observation (2.2). For typical $\mathbf{L}^\natural \in \mathbb{R}^{n \times n}$ with rank r and $\mathbf{S}^\natural \in \mathbb{R}^{n \times n}$ with s nonzeros, the incoherence conditions are in force so long as

$$r^{1/2}s = O\left(\frac{n^{3/2}}{\log(n)}\right). \quad (2.10)$$

This result implies that (2.4) succeeds in the regime where the rank r approaches n so long as the sparsity s is polynomially small in the sidelength n .

Using the “golfing” scheme developed by Gross (cf. [GLF⁺10, Gro11]), Candès et al. provide a complementary rank–sparsity analysis [CLMW11]. Their results are somewhat weaker than the approach of Chandrasekaran et al. In particular, their results do not allow the rank r to approach n no matter how small the sparsity s . Other golfing-based analyses for modified versions of rank–sparsity model (2.2) include [JRSR10, XCS10b, CJSC13]. In each case, the bounds are modestly suboptimal.

Ma & Wright [WGMM12] proposed a general method called “certificate upgrade” that demonstrates that demixing compressed measurements of the form (1.10) is possible when sufficiently strong guarantees for the uncompressed model (1.7) are available. They apply their results to show that a low-rank + sparse signal (2.2) can be

demixed even when only a few random linear measurements of the sum is available. This approach results in similar guarantees to those of [CJSC13], but the analysis applies to a much wider variety of situations.

Finally, we mention that Hegde & Baraniuk provided a rigorous analysis of their SPIN algorithm in several specific settings [HB12]. Their approach is based on a generic incoherent manifold model that shares some similarities to the analysis of [CSPW11]. Although their analysis provides general guarantees, it does not extend to demixing low-rank matrices from sparse vectors.

Comparison with this work. In Section 10.3, we consider demixing a low-rank matrix from a matrix that is sparse a random basis. Our analysis provides a precise characterization of success and failure for a standard convex demixing method under this model. This result is unique in the literature because it provides optimal bounds, but our assumption that the sparse matrix is randomly rotated is novel, and somewhat unusual, for this literature. Small-scale numerical experiments, while not conclusive, seem to indicate that the phase transition for the random models considered in [CSPW11, CLMW11] are not accurately predicted by the theory in our work. It would be interesting to understand this discrepancy.

2.3.3 Linear inverse problems

A *linear inverse problem* is the task of determining a vector \mathbf{x}^\dagger from a linearly undersampled observation $\mathbf{z}_0 = \mathbf{A}\mathbf{x}^\dagger$. This section describes some of the tools and ideas from linear inverse problems that are related to the analysis provided by this work. We start with a discussion of the geometric tools that have found applications to linear inverse problems, and describe how these tools are related to the analysis of this work. We then describe another approach for understanding linear inverse problems based on the Gaussian process inequalities of Gordon [Gor85, Gor87, Gor88]. Our work reveals a connection between these two superficially distinct approaches to linear inverse problems.

2.3.3.1 Asymptotic polytope angle computations

The geometric theory of polytope angles has proved a highly effective tool for the analysis of linear programs with random constraints. The study of polytope angles has its roots in the work Schäfli in the mid-19th century [Sch50], but made its first appearance in the analysis of optimization programs in the work of Vershik & Sporyshev [VS86] who analyzed the efficiency of the simplex method under random affine constraints. This work led to mathematical research on the neighborliness of random polytopes [VS92, AS92, BH99].

Donoho [Don04, Don06b] and Donoho & Tanner [DT05, DT09, DT10c, DT10a, DT10b] applied these earlier developments to study the behavior of several linear programs with sparsity constraints. To be precise, they considered the case where $\mathbf{x}^\natural \in \mathbb{R}^d$ is sparse and the measurement operator $\mathbf{A} = \mathbf{G} \in \mathbb{R}^{m \times d}$ has independent standard Gaussian entries. To recover the data \mathbf{x}^\natural from the observation $\mathbf{z}_0 = \mathbf{G}\mathbf{x}^\natural$, these authors considered the ℓ_1 minimization program

$$\text{minimize } \|\mathbf{x}\|_{\ell_1} \quad \text{subject to } \mathbf{G}\mathbf{x} = \mathbf{z}_0 \quad (2.11)$$

where the decision $\mathbf{x} \in \mathbb{R}^d$. We say that (2.11) *succeeds* if \mathbf{x}^\natural is the unique optimal solution to (2.11).

Donoho [Don06b] considered the asymptotic regime where $d \rightarrow \infty$ and the measurements $m = \lceil \delta d \rceil$ and sparsity $\text{nnz}(\mathbf{x}^\natural) = \lceil \rho d \rceil$ grows proportionally with d (ρ may depend on δ). Donoho demonstrated the existence of a function $\Psi: (0, 1)^2 \rightarrow \mathbb{R}$ such that, if $\Psi(\rho, \delta) < 0$, the method (2.11) succeeds with probability approaching one as the dimension d goes to infinity. This condition provided an empirically sharp lower bound on the number of measurements required for the success of (2.11) (the *weak* bound). Donoho also provided a lower bound on the number of measurements required to recover *all* sufficiently sparse vectors with a single Gaussian matrix (the *strong* bound).

In later work, Donoho & Tanner verified that if $\Psi(\rho, \delta) > 0$ and $\delta > 0$ is sufficiently small, then the success probability does not go to one as the dimension $d \rightarrow \infty$ [DT09].

However, it was not shown that the probability of success goes to zero when $\Psi(\rho) > 0$. Similar nonasymptotic guarantees appeared in [DT10b].

This polytope angle approach to analyzing (2.11) was extended by several other authors. Xu & Hassibi [XH08, XH11] used the framework to provide stability guarantees for (2.11) for all sparsity levels where $\Psi(\rho) < 0$. Khajehnejad et al. [KX09, XKAH10, KXAH10, KXAH11] consider weighted versions of (2.11), and they demonstrate that reweighting will outperform (2.11) under additional assumptions on the constituent signals.

While polytope angle methods have provided a number impressive results, the theory is inherently restricted to polyhedral settings. For more general convex programs, the theory of conic intrinsic volumes provides an appropriate generalization of polytope angles. This more modern theory has its roots in the integral-geometric investigations of Santaló [San76], and reached its current development in the thesis of Glasauer [Gla95, Gla96]. Our work appears to be the first to use this more modern theory in the context of signal processing. However, Amelunxen [Ame11] and Amelunxen & Bürgisser [AB11, AB12a, AB12b] have applied these ideas to study the condition numbers of random convex programs.

2.3.3.2 Gaussian width analyses

Finally, we discuss an apparently very different technique for analyzing convex linear inverse methods that uses a comparison principle for Gaussian processes due to Gordon [Gor85, Gor87, Gor88]. Gordon’s comparison principle provides an inequality for the probability that a randomly oriented subspace strikes a cone in terms of a geometric parameter called the *Gaussian width* (see Section 6.3). Rudelson & Vershynin [RV08] were the first to recognize that this inequality provides success guarantees for (2.11).

Stojnic [Sto09] subsequently refined this approach to provide recovery guarantees for (2.11) that matched the weak bound of Donoho. Moreover, Stojnic’s work improved Donoho’s strong bound in certain sparsity regimes. Oymak & Hassibi [OH10] used a similar approach to provide empirically sharp bounds on the number of measurements required to recover a low-rank matrix from Gaussian measurements. They also provide

a strong bound on the number of measurements required to find *all* low-rank matrices, a feat that appears unmatched in the literature.

This method was further generalized and greatly simplified by Chandrasekaran et al. [CRPW12]. These authors demonstrated that the Gaussian width method applies to regularized linear inverse problems involving all types of atomic gauges mentioned in Section 1.2.1, and they developed a general geometric framework for understanding these problems. Moreover, they showed that Gordon’s Gaussian process inequalities provide a natural stability guarantee for convex linear inverse programs.

Comparison with this work: The missing link? The fact that the Gaussian width and the asymptotic polytope angle analyses for linear inverse problems yield the same success boundary for linear inverse problems indicated that a connection between these two approaches was waiting to be explored [XH12, p. 312]. The results of this work provide one such link between these two methods through the new geometric inequality of Theorem 7.1 and the inequality (6.38) that links the Gaussian width to a dimensionality parameter for convex cones.

Our work does more than simply link these two approaches, however. Despite the fact that both of these approaches lead to empirically correct phase transition locations in the case of linear inverse problems, a rigorous justification of this fact was lacking. Our theory indicates that a phase transition for linear inverse problems must occur when the number of measurements m exceeds a computable threshold that depends on the problem parameters. In the case of ℓ_1 minimization with Gaussian measurements (2.11), the location of this threshold agrees with both the polytope angle and the Gaussian width threshold. Our result, however, comes with additional guarantees:

- This threshold provides a phase transition between near-certain success and near-certain failure as the dimension $d \rightarrow \infty$.
- In the nonasymptotic regime, our theory provides bounds on the width of the transition regime between success and failure.

Moreover, our approach is capable of providing strong bounds with much less effort

than these earlier techniques, although our simple approach sacrifices some precision.

Chapter 3

Mathematical preliminaries

This chapter introduces some of the basic notions from convex analysis that we use throughout this work. These facts span from basic definitions and conventions to technical results on the metric topology of convex cones. Along the way, we cover results from conic duality theory and polyhedral geometry. Most, if not all, of these facts are available in the literature, but no single reference is available that contains all of this material. In order to lay a strong foundation for the work that follows, we compile these facts into a number of omnibus propositions, and we provide complete proofs or explicit references to standard texts.

Section 3.1 introduces our notation and conventions. While most of our notation is standard, note that we use special notation for particular random vectors and matrices. In Section 3.2, we present a number of facts about convex cones, and we introduce the *descent cone* that plays an important role in the rest of our work. Section 3.3 contains a technical discussion of the topology of closed, convex cones, and it may be skipped on a first reading.

3.1 Conventions

Vectors and matrices. Bold lowercase letters denote vectors, while bold uppercase letters are reserved for matrices, so \mathbf{x} is a vector while \mathbf{X} is a matrix. We denote the transpose of \mathbf{X} by \mathbf{X}^t . The letter \mathbf{I} is an identity matrix, and $\mathbf{0}$ is the zero vector.

The inner product between two vectors is written $\langle \mathbf{x}, \mathbf{y} \rangle$, and the inner product

between two matrices is $\langle X, Y \rangle := \text{tr}(X^t Y)$, where $\text{tr}(\cdot)$ denotes the trace operator. The vector inner product induces the Euclidean norm $\|\mathbf{x}\|^2 := \langle \mathbf{x}, \mathbf{x} \rangle$. The matrix inner product induces the Frobenius norm $\langle X, X \rangle = \|X\|_F^2$ that provides a Euclidean structure to the space of matrices.

The vectorization operator $\text{vec}(\cdot)$ maps a matrix $X \in \mathbb{R}^{m \times n}$ to a vector $\text{vec}(X) \in \mathbb{R}^{mn}$ by stacking the columns of X from left to right into a tall vector. For two vectors $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{y} \in \mathbb{R}^{d'}$, we form the tall vector $(\mathbf{x}; \mathbf{y}) \in \mathbb{R}^{d+d'}$ by stacking \mathbf{x} on top of \mathbf{y} . The support $\text{supp}(\mathbf{x})$ of a vector $\mathbf{x} \in \mathbb{R}^d$ is the set of indices of all nonzero elements of \mathbf{x} :

$$\text{supp}(\mathbf{x}) := \{i : x_i \neq 0\}.$$

Set operations. For two sets $S, S' \subset \mathbb{R}^d$, the Minkowski sum $S + S'$ is the set of all summands formed by elements of S and S' :

$$S + S' := \{\mathbf{x} + \mathbf{x}' : \mathbf{x} \in S, \mathbf{x}' \in S'\}.$$

For a tuple $(S_i)_{i=1}^n$ of subsets of \mathbb{R}^d , we write the iterated Minkowski sum as

$$\sum_{i=1}^n S_i := \left\{ \sum_{i=1}^n \mathbf{x}_i : \mathbf{x}_i \in S_i \text{ for all } i = 1, \dots, n \right\}.$$

We denote the closure of S by \bar{S} , the linear hull of S by $\text{lin}(S)$, and the convex hull of S by $\text{conv}(S)$. For any number $\lambda \in \mathbb{R}$, we define $\lambda S := \{\lambda \mathbf{x} : \mathbf{x} \in S\}$. This notation extends to the image of set S under a linear map $A \in \mathbb{R}^{m \times d}$:

$$AS := \{A\mathbf{x} : \mathbf{x} \in S\}.$$

The indicator function $\mathbf{1}_S$ on S is

$$\mathbf{1}_S(\mathbf{x}) := \begin{cases} 1, & \mathbf{x} \in S, \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

The related function i_S defined by

$$i_S(\mathbf{x}) := -\log(\mathbf{1}_S(\mathbf{x})) = \begin{cases} 0, & \mathbf{x} \in S, \\ +\infty, & \text{otherwise,} \end{cases} \quad (3.2)$$

is also commonly referred to as an indicator function in the convex analysis literature because i_S is convex if and only if S is convex. To make the distinction between (3.1) and (3.2), we call i_S the *convex indicator function*.

Special sets. The unit Euclidean ball and unit sphere are given by

$$B_d := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\} \quad \text{and} \quad S^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = 1\}.$$

The Stiefel manifold O_d of $d \times d$ orthogonal matrices is

$$O_d := \{\mathbf{U} \in \mathbb{R}^{d \times d} : \mathbf{U}\mathbf{U}^t = \mathbf{I}\}.$$

An *orthogonal basis* for \mathbb{R}^d is an element of O_d . An orthogonal basis for a space of matrices is written in script, e.g., \mathcal{U} . For a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$, an orthogonal basis \mathcal{U} for $\mathbb{R}^{m \times n}$ can be canonically identified with a matrix $\mathbf{U} \in O_{mn}$ via the identity

$$\mathcal{U}(\mathbf{X}) = \text{vec}^{-1}(\mathbf{U} \text{vec}(\mathbf{X})).$$

The measure σ_{d-1} on the sphere S^{d-1} is induced by the usual Lebesgue measure on \mathbb{R}^d . In particular, the spherical measure σ_{d-1} is not normalized, so that the total measure of the sphere is [SW08, p. 13]

$$\sigma_{d-1}(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}. \quad (3.3)$$

Convexity. A set $K \subset \mathbb{R}^d$ is *convex* if it contains the line segment connecting each pair of its points. A convex set is a *polyhedron* if it is the intersection of a finite number of closed halfspaces. A function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\pm\infty\}$ over the extended real line is convex

if its epigraph

$$\text{epi}(f) := \{(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R} : f(\mathbf{x}) \leq t\}$$

is convex. A convex function f is *proper* if its epigraph is nonempty and $f(\mathbf{x}) > -\infty$ for all \mathbf{x} , and a convex function f is *closed* if its epigraph is closed. The *subgradient* $\partial f(\mathbf{x})$ of a convex function f at $\mathbf{x} \in \mathbb{R}^d$ is the set

$$\partial f(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^d : f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle \text{ for all } \mathbf{z} \in \mathbb{R}^d\}.$$

The domain $\text{dom}(f)$ of a convex function f on \mathbb{R}^d is given by

$$\text{dom}(f) := \{\mathbf{x} \in \mathbb{R}^d : f(\mathbf{x}) < \infty\}.$$

Probability. Probability plays a central role in our analysis of demixing. The symbol \mathbb{P} denotes the probability of an event, while \mathbb{E} returns the expectation of a random variable. The abbreviation *i.i.d.* stands for “independent and identically distributed”, while the tilde “ \sim ” indicates equality in distribution, and should be read “is distributed as.” We denote standard Gaussian vectors with the letter $\mathbf{g} \sim \text{NORMAL}(\mathbf{0}, \mathbf{I})$, while the letter \mathbf{G} represents a matrix with independent standard normal entries. We reserve the letter $\boldsymbol{\theta}$ for a uniform random variable on the Euclidean sphere S^{d-1} .

The set of orthogonal matrices O_d is a compact Lie group, and hence it has a unique invariant (Haar) probability measure. We denote elements drawn from this probability measure by \mathbf{Q} , and refer to \mathbf{Q} as a *random orthogonal basis* or a *random rotation*, with the understanding that reflections are allowed. Invariant measures have a convenient *forgetting property*.

Fact 3.1 (Forgetting property of the invariant measure.). *Let $(\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n)$ be i.i.d. random rotations in O_d . Suppose that $f : O_d^n \rightarrow \mathbb{R}$ is a measurable function that satisfies*

$$\mathbb{E} \left[\mathbb{E} [|f(\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n)| \mid \mathbf{Q}_1] \right] < \infty, \quad (3.4)$$

where the outer expectation is over \mathbf{Q}_1 , and the inner expectation is over \mathbf{Q}_i for $i \geq 2$. In

particular, condition (3.4) holds when $|f|$ is bounded. Then

$$\mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n)] = \mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_1 \mathbf{Q}_2, \dots, \mathbf{Q}_1 \mathbf{Q}_{n-1}, \mathbf{Q}_1 \mathbf{Q}_n)]. \quad (3.5)$$

We are not aware of any elementary reference for this fact as stated, and so we provide the basic proof.

Proof. The integrability condition (3.4) on f ensures that Fubini's theorem [Kal02, Thm. 1.27] applies:

$$\mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n)] = \mathbb{E}\left[\mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n) \mid \mathbf{Q}_1]\right]$$

where the outer expectation is over \mathbf{Q}_1 , and the inner expectation is over \mathbf{Q}_i for $i \geq 2$. The invariance of the measures of \mathbf{Q}_i implies that may rotate the \mathbf{Q}_i for $i \geq 2$ by an arbitrary fixed matrix without changing the inner expectation. Because \mathbf{Q}_1 is fixed relative to the inner expectation above, we have

$$\begin{aligned} \mathbb{E}\left[\mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n) \mid \mathbf{Q}_1]\right] &= \mathbb{E}\left[\mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_1 \mathbf{Q}_2, \dots, \mathbf{Q}_n) \mid \mathbf{Q}_1]\right] \\ &= \mathbb{E}[f(\mathbf{Q}_1, \mathbf{Q}_1 \mathbf{Q}_2, \dots, \mathbf{Q}_1 \mathbf{Q}_{n-1}, \mathbf{Q}_1 \mathbf{Q}_n)], \end{aligned}$$

where the second inequality is justified by another application of Fubini's theorem. \square

3.2 Convex cones

A *convex cone* $C \subset \mathbb{R}^d$ is a convex set that is positive homogeneous, so that C is a cone if and only if

$$C = \lambda C := \{\lambda \mathbf{x} : \mathbf{x} \in C\} \text{ for all } \lambda > 0.$$

We define \mathcal{C}_d as the set of all nonempty closed convex cones in \mathbb{R}^d . A (closed) *ray* $R_{\mathbf{x}} \in \mathcal{C}_d$ generated by a nonzero point $\mathbf{x} \in \mathbb{R}^d$ is the one-dimensional cone

$$R_{\mathbf{x}} := \bigcup_{\lambda \geq 0} \{\lambda \mathbf{x}\} = \{\lambda \mathbf{x} : \lambda \geq 0\}. \quad (3.6)$$

The polar C° of a cone C is the set of all outward-pointing normals to C :

$$C^\circ := \{\mathbf{y} \in \mathbb{R}^d : \langle \mathbf{y}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in C\}.$$

A *halfspace* $H_{\mathbf{x}}$ with a nonzero normal $\mathbf{x} \in \mathbb{R}^d$ is the set $H_{\mathbf{x}} := \{\mathbf{y} \in \mathbb{R}^d : \langle \mathbf{y}, \mathbf{x} \rangle \leq 0\}$. In particular, the $H_{\mathbf{x}} = R_{\mathbf{x}}^\circ$, where $R_{\mathbf{x}}$ is the ray generated by \mathbf{x} defined above. The *product* of two cones $C \subset \mathbb{R}^d$, $C' \subset \mathbb{R}^{d'}$ is the set

$$C \times C' := \{(\mathbf{x}; \mathbf{y}) : \mathbf{x} \in C, \mathbf{y} \in C'\} \subset \mathbb{R}^{d+d'}.$$

Convex cones enjoy a rich structure and duality theory that goes well beyond the relations enjoyed by simple convex sets. For future reference, we collect a number of these standard properties in the following proposition.

Proposition 3.2 (The algebra of convex cones). *Let $C, D \subset \mathbb{R}^d$ be convex cones.*

1. **Decomposition.** *The cone C has the unique orthogonal decomposition*

$$C = C_* + L_C, \tag{3.7}$$

where C_ is a pointed cone and L_C is a linear subspace.*

2. **Summation.** *The Minkowski sum satisfies*

$$C + D = \text{conv}\{C \cup D\}; \tag{3.8}$$

3. **Bipolar.** *The polar C° is a closed convex cone and $C^{\circ\circ} = \overline{C}$;*

4. **Sums.** *The polar of a sum is the intersection of the polars:*

$$(C + D)^\circ = C^\circ \cap D^\circ \quad \text{and} \quad (C \cap D)^\circ = C^\circ + D^\circ; \tag{3.9}$$

5. **Rotational covariance.** For any orthogonal matrix $U \in O_d$,

$$(UC)^\circ = UC^\circ; \quad (3.10)$$

6. **Polars of products.** For any cone $C' \subset \mathbb{R}^{d'}$, the polar of the product cone is the product of the polars:

$$(C \times C')^\circ = C^\circ \times (C')^\circ. \quad (3.11)$$

Proof. The first four properties are available in [Roc70]. For the decomposition relation (#1), we refer to p. 65, the summation property (#2) is available on p. 22, and the bipolar relation (#3) appears on p. 121, and the sum relation (#4) appears as Corollary 16.4.2.

For the rotational covariance property (#5), we have the following string of equivalences:

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle \leq 0 \quad \forall \mathbf{x} \in UC &\iff \langle U\mathbf{x}, \mathbf{y} \rangle \leq 0 \quad \forall \mathbf{x} \in C \\ &\iff \langle \mathbf{x}, U^t \mathbf{y} \rangle \leq 0 \quad \forall \mathbf{x} \in C \iff U^t \mathbf{y} \in C^\circ. \end{aligned}$$

The first equivalence is the definition of UC , the second is the characterization of the transpose, and the final equivalence follows by definition of the polar cone C° . The claim (3.10) then follows from the fact that $U = (U^t)^{-1}$ because U is orthogonal.

The final property (#6) is a standard exercise; we repeat the details for completeness. Suppose that $(\mathbf{w}; \mathbf{z}) \in (C \times C')^\circ$. Then for all $\lambda, \lambda' > 0$ and any $\mathbf{x} \in C$, $\mathbf{y} \in C'$, we have

$$0 \geq \langle (\lambda \mathbf{x}; \lambda' \mathbf{y}), (\mathbf{w}; \mathbf{z}) \rangle = \lambda \langle \mathbf{x}, \mathbf{w} \rangle + \lambda' \langle \mathbf{y}, \mathbf{z} \rangle,$$

by definition of the polar cone and homogeneity of C, C' . Taking λ or λ' to zero shows that each of the terms of the right-hand side above must be nonpositive, and hence $(\mathbf{w}; \mathbf{z}) \in C^\circ \times (C')^\circ$. On the other hand, for any $\mathbf{w} \in C^\circ$, $\mathbf{z} \in (C')^\circ$ and all $\mathbf{x} \in C$, $\mathbf{y} \in C'$, we have

$$0 \geq \langle \mathbf{x}, \mathbf{w} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle = \langle (\mathbf{x}; \mathbf{y}), (\mathbf{w}; \mathbf{z}) \rangle$$

by definition of the polar cone. Hence $(\mathbf{w}; \mathbf{z}) \in (C \times C')^\circ$, and so (3.11) holds. \square

Another useful fact is the following separating hyperplane theorem for convex cones due to Klee [Kle55, Thm. 2.5].

Fact 3.3 (Separating hyperplane theorem for convex cones). *Suppose C, C' are two closed convex cones in a finite-dimensional vector space \mathbb{E} isomorphic to \mathbb{R}^n for some integer $n \geq 1$. If $C \cap C' = \{\mathbf{0}\}$, then there exists a nonzero $\mathbf{z} \in \mathbb{E}$ such that $\mathbf{z} \in C^\circ$ and $-\mathbf{z} \in (C')^\circ$.*

We use the notation \mathbb{E} and not \mathbb{R}^n because an application of Fact 3.3 in Section 5.4.1 requires that \mathbb{E} be some arbitrary linear subspace of \mathbb{R}^n , as opposed to a coordinate subspace that might be inferred by writing \mathbb{R}^n .

In the finite-dimensional setting, Fact 3.3 is readily deduced from the usual separating hyperplane theorem for convex sets by considering the hyperplane that separates the interior of C from C' . Klee's original result is much more general, and states that the statement of Fact 3.3 holds when \mathbb{E} is any locally convex topological vector space.

3.2.1 Projections and distances

Projections onto and distances from convex cones possess a significant amount of algebraic structure. This section describes some of these remarkable properties.

Definition 3.4 (Projection and distance). For each closed cone $C \in \mathcal{C}_d$ and all $\mathbf{x} \in \mathbb{R}^d$, we define the projection $\Pi_C(\mathbf{x})$ of \mathbf{x} onto C by

$$\Pi_C(\mathbf{x}) := \arg \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|^2. \quad (3.12)$$

The distance $\text{dist}(\mathbf{x}, C)$ from \mathbf{x} to C is given by

$$\text{dist}(\mathbf{x}, C) := \inf_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|. \quad (3.13)$$

Both the projection operator and the distance function are well defined because the squared Euclidean norm is strongly convex and coercive, and hence it achieves its minimum at a unique point in every closed convex set.

We require a number of important and (mostly) well-known properties of the distance and projection functionals.

Proposition 3.5 (Projections and distances). *Let $C \in \mathcal{C}_d$ be a closed, convex cone, and let $\mathbf{x} \in \mathbb{R}^d$.*

1. **Orthogonality characterization.** *The points $\mathbf{y}, \mathbf{z} \in \mathbb{R}^d$ satisfy*

$$\mathbf{x} = \mathbf{y} + \mathbf{z}, \quad \mathbf{y} \in C, \quad \mathbf{z} \in C^\circ, \quad \text{and} \quad \mathbf{y} \perp \mathbf{z} \quad (3.14)$$

if and only if $\mathbf{y} = \Pi_C(\mathbf{x})$ and $\mathbf{z} = \Pi_{C^\circ}(\mathbf{x})$. In particular, \mathbf{x} has an orthogonal decomposition over C and C° given by

$$\mathbf{x} = \Pi_C(\mathbf{x}) + \Pi_{C^\circ}(\mathbf{x}). \quad (3.15)$$

2. **Pythagorean identity.** *The norm of \mathbf{x} satisfies*

$$\|\mathbf{x}\|^2 = \|\Pi_C(\mathbf{x})\|^2 + \|\Pi_{C^\circ}(\mathbf{x})\|^2. \quad (3.16)$$

3. **Distance as a projection.** *The distance from \mathbf{x} to C satisfies*

$$\text{dist}(\mathbf{x}, C) = \|\Pi_{C^\circ}(\mathbf{x})\|. \quad (3.17)$$

4. **Product rule.** *For $C' \in \mathcal{C}_{d'}$ and any $(\mathbf{x}; \mathbf{x}') \in \mathbb{R}^{d+d'}$, the projection onto $C \times C'$ is given by*

$$\Pi_{C \times C'}((\mathbf{x}; \mathbf{x}')) = (\Pi_C(\mathbf{x}); \Pi_{C'}(\mathbf{x}')). \quad (3.18)$$

5. **Gradient.** *The map $\mathbf{x} \mapsto \|\Pi_C(\mathbf{x})\|^2$ is everywhere differentiable with gradient given by*

$$\nabla \|\Pi_C(\mathbf{x})\|^2 = \nabla \text{dist}^2(\mathbf{x}, C^\circ) = 2\Pi_C(\mathbf{x}). \quad (3.19)$$

Proof. We provide a complete demonstration of each point.

(*Orthogonality characterization*) A necessary and sufficient condition for a point $\mathbf{y} \in C$ to minimize the strongly convex, differentiable function $f(\mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|^2$ over C set is that the gradient of f at \mathbf{y} supports C at \mathbf{y} [Roc70, p. 271]. As $\nabla f(\mathbf{y}) = 2(\mathbf{y} - \mathbf{x})$, a necessary and sufficient condition for the equality $\mathbf{y} = \Pi_C(\mathbf{x})$ is

$$\mathbf{y} \in C \quad \text{and} \quad \langle \mathbf{y} - \mathbf{x}, \mathbf{x}' - \mathbf{y} \rangle \leq 0 \quad \text{for all} \quad \mathbf{x}' \in C. \quad (3.20)$$

We will show that (3.20) is equivalent to the conditions

$$\mathbf{y} \in C, \quad \mathbf{y} - \mathbf{x} \in C^\circ, \quad \text{and} \quad \mathbf{y} \perp \mathbf{y} - \mathbf{x}. \quad (3.21)$$

Indeed, if the condition (3.21) holds, then for any $\mathbf{x}' \in C$ we have

$$\langle \mathbf{y} - \mathbf{x}, \mathbf{x}' - \mathbf{y} \rangle = \langle \mathbf{y} - \mathbf{x}, \mathbf{x}' \rangle \leq 0 \quad \text{for any} \quad \mathbf{x}' \in C$$

and so (3.20) holds. On the other hand, suppose (3.20) holds. Then homogeneity of the cone C readily implies $\mathbf{y} - \mathbf{x} \in C^\circ$ because, if $\langle \mathbf{y} - \mathbf{x}, \mathbf{x}' \rangle > 0$ for some $\mathbf{x}' \in C$, then we may force $\langle \mathbf{y} - \mathbf{x}, M\mathbf{x}' - \mathbf{y} \rangle > 0$ by taking M large enough. Hence (3.20) implies $\mathbf{y} - \mathbf{x} \in C^\circ$. But taking $\mathbf{x}' = 0$ in (3.20) yields

$$\langle \mathbf{y} - \mathbf{x}, \mathbf{y} \rangle \geq 0.$$

As we have already established $\mathbf{y} - \mathbf{x} \in C^\circ$, and the relation $\mathbf{y} \in C$ holds by construction, we must have $\langle \mathbf{y} - \mathbf{x}, \mathbf{y} \rangle = 0$. Thus, (3.20) and (3.21) are equivalent, and so $\mathbf{y} = \Pi_C(\mathbf{x})$ if and only if (3.21) holds. Applying the same argument to the polar cone reveals that $\mathbf{z} = \Pi_{C^\circ}(\mathbf{x})$ if and only if

$$\mathbf{z} \in C^\circ, \quad \mathbf{z} - \mathbf{x} \in C, \quad \text{and} \quad \mathbf{z} \perp \mathbf{z} - \mathbf{x}, \quad (3.22)$$

where we have used the fact that $C^{\circ\circ} = C$ because C is closed (see Proposition 3.2.3).

Now note that \mathbf{y} satisfies (3.21) if and only if $\mathbf{z} := \mathbf{x} - \mathbf{y}$ satisfies (3.22). Since (3.21)

and (3.22) hold if and only if $\mathbf{y} = \Pi_C(\mathbf{x})$ and $\mathbf{z} = \Pi_{C^\circ}(\mathbf{x})$, the conclusion follows.

(Pythagorean identity) For the Pythagorean identity (3.16), we apply the decomposition (3.15) and expand the Euclidean norm:

$$\|\mathbf{x}\|^2 = \|\Pi_C(\mathbf{x}) + \Pi_{C^\circ}(\mathbf{x})\|^2 = \|\Pi_C(\mathbf{x})\|^2 + 2\langle \Pi_C(\mathbf{x}), \Pi_{C^\circ}(\mathbf{x}) \rangle + \|\Pi_{C^\circ}(\mathbf{x})\|^2.$$

By orthogonality the decomposition (3.15), the inner product above is equal to zero, and hence the Pythagorean formula (3.16) holds.

(Distance as a projection) Comparing the definition (3.12) of the projection map to the definition (3.13) of the distance map, we see

$$\text{dist}(\mathbf{x}, C) = \|\mathbf{x} - \Pi_C(\mathbf{x})\| = \|\Pi_{C^\circ}(\mathbf{x})\|$$

where the second equality is (3.15). This is the claimed distance formula (3.17).

(Product rule.) For the product rule, we define $\mathbf{y}, \mathbf{z} \in \mathbb{R}^{d+d'}$ by

$$\mathbf{y} := \left(\Pi_C(\mathbf{x}); \Pi_{C'}(\mathbf{x}') \right) \quad \text{and} \quad \mathbf{z} := \left(\Pi_{C^\circ}(\mathbf{x}); \Pi_{(C')^\circ}(\mathbf{x}') \right). \quad (3.23)$$

From the orthogonal decomposition (3.15), we find the equality $(\mathbf{x}, \mathbf{x}') = \mathbf{y} + \mathbf{z}$. By virtue of the orthogonality of (3.15), the vectors \mathbf{y} and \mathbf{z} are perpendicular:

$$\langle \mathbf{y}, \mathbf{z} \rangle = \langle \Pi_C(\mathbf{x}), \Pi_{C^\circ}(\mathbf{x}) \rangle + \langle \Pi_{C'}(\mathbf{x}), \Pi_{(C')^\circ}(\mathbf{x}') \rangle = 0$$

Moreover, $\mathbf{y} \in C \times C'$ by definition. The fact that $\mathbf{z} \in (C \times C')^\circ$ is easy to verify directly.

For any $(\mathbf{q}; \mathbf{q}') \in C \times C'$, we have

$$\langle (\mathbf{q}; \mathbf{q}'), \mathbf{z} \rangle = \langle \mathbf{q}, \Pi_C(\mathbf{x}) \rangle + \langle \mathbf{q}', \Pi_{C'}(\mathbf{x}') \rangle \leq 0$$

because $\mathbf{q} \in C$ and $\mathbf{q}' \in C'$. Since \mathbf{y} and \mathbf{z} satisfy (3.14), we must have $\mathbf{y} = \Pi_{C \times C'}(\mathbf{x})$.

The claim follows upon recalling the definition (3.23) of \mathbf{y} .

(*Gradient.*) The first equality in (3.19) follows immediately from (3.17) whenever either gradient is well defined. To show the second gradient exists and that the second equality holds, we follow the proof of [RW98, Thm. 2.26]. By linearity of the gradient, we only need to demonstrate that the function

$$h(\mathbf{u}) := \text{dist}^2(\mathbf{x} + \mathbf{u}, C^\circ) - \text{dist}^2(\mathbf{x}, C^\circ) - 2\langle \Pi_C(\mathbf{x}), \mathbf{u} \rangle$$

is differentiable at $\mathbf{u} = \mathbf{0}$ and that $\nabla h(\mathbf{0}) = \mathbf{0}$. Because $\mathbf{x} - \Pi_C(\mathbf{x}) \in C^\circ$ by (3.14), we find

$$\text{dist}^2(\mathbf{x} + \mathbf{u}, C^\circ) \leq \|\mathbf{x} + \mathbf{u} - \Pi_{C^\circ}(\mathbf{x})\|^2 = \|\Pi_C(\mathbf{x}) + \mathbf{u}\|^2,$$

where the inequality follows by definition of the distance and the second equality is (3.15). After we expand the square, we find

$$h(\mathbf{u}) \leq \|\Pi_C(\mathbf{x})\|^2 - \text{dist}^2(\mathbf{x}, C^\circ) + \|\mathbf{u}\|^2 = \|\mathbf{u}\|^2 \quad \text{for every } \mathbf{u} \in \mathbb{R}^d,$$

where the second inequality follows from (3.17). Because h is the sum of convex functions, the function h is itself convex, and so it lies below its chord:

$$\frac{1}{2}(h(\mathbf{u}) + h(-\mathbf{u})) \geq h(\mathbf{0}) = 0.$$

Combining this result with the preceding display, we see that $h(\mathbf{u}) \geq -h(-\mathbf{u}) \geq -\|\mathbf{u}\|^2$, and so we conclude that $|h(\mathbf{u})| \leq \|\mathbf{u}\|^2$. Therefore,

$$\lim_{\mathbf{u} \rightarrow \mathbf{0}} \frac{|h(\mathbf{u})|}{\|\mathbf{u}\|} = 0.$$

By definition of the gradient, we have $\nabla h(\mathbf{0}) = \mathbf{0}$. This completes the proof. \square

3.2.2 Descent cones

For convex functions, local behavior determines a significant amount of global structure. The following definition encodes detailed local information about convex functions in

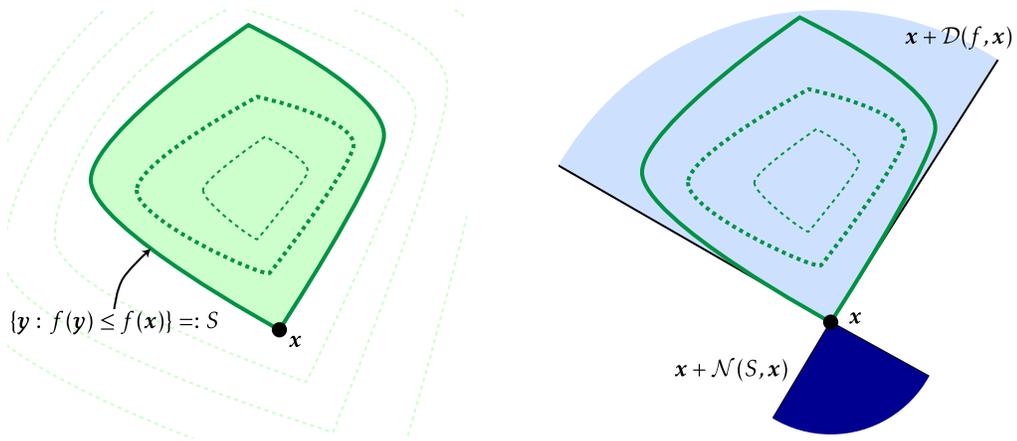


Figure 3.1: Descent and normal cones. [Left] The sublevel set S (shaded) of a convex function f (level lines) at a point \mathbf{x} . [Right] The descent cone $\mathcal{D}(f, \mathbf{x})$ (light shade) is the cone generated by the sublevel set S at \mathbf{x} (cf. (3.26)). The normal cone $\mathcal{N}(S, \mathbf{x})$ (dark shade) is polar to the descent cone.

the form of cones.

Definition 3.6 (Descent and normal cones). The *descent cone* $\mathcal{D}(f, \mathbf{x})$ of a function $f: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ at a point $\mathbf{x} \in \mathbb{R}^d$ is the cone generated by the perturbations of f at \mathbf{x} that do not increase f :

$$\mathcal{D}(f, \mathbf{x}) := \bigcup_{\lambda > 0} \{\mathbf{y} : f(\mathbf{x} + \lambda \mathbf{y}) \leq f(\mathbf{x})\}. \quad (3.24)$$

The *normal cone* $\mathcal{N}(S, \mathbf{x})$ of a set S at \mathbf{x} is the set of supporting hyperplanes to S at \mathbf{x} :

$$\mathcal{N}(S, \mathbf{x}) := \{\mathbf{y} : \langle \mathbf{y}, \mathbf{s} - \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{s} \in S\}. \quad (3.25)$$

In this definition, we allow f to take values on the extended real line. In particular, when $f(\mathbf{x}) = +\infty$, we have $\mathcal{D}(f, \mathbf{x}) = \mathbb{R}^d$.

The normal cone has a close relationship to the descent cone.

Proposition 3.7 (Properties of descent and normal cones). *Let $f: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. For any point $\mathbf{x} \in \mathbb{R}^d$, define the sublevel set $S := \{\mathbf{z} : f(\mathbf{z}) \leq f(\mathbf{x})\}$ of f at \mathbf{x} . Then*

1. **Sublevels.** The descent cone is generated by the sublevel set:

$$\mathcal{D}(f, \mathbf{x}) = \bigcup_{\lambda > 0} \lambda \cdot (S + \{-\mathbf{x}\}). \quad (3.26)$$

2. **Convexity.** The descent cone is a convex cone that contains its apex $\mathbf{0}$.

3. **Polarity.** The normal cone is polar to the descent cone:

$$\mathcal{N}(S, \mathbf{x}) = \mathcal{D}(f, \mathbf{x})^\circ. \quad (3.27)$$

In particular, the normal cone is a closed, convex cone.

4. **Minima.** The point \mathbf{x} is the unique global minimum of f if and only if

$$\mathcal{D}(f, \mathbf{x}) = \{\mathbf{0}\}, \quad \text{or equivalently, if and only if } \mathcal{N}(S, \mathbf{x}) = \mathbb{R}^d. \quad (3.28)$$

If in addition, \mathbf{x} is in the interior of $\text{dom}(f)$, the subgradient $\partial f(\mathbf{x})$ is not empty, and $\mathbf{0} \notin \partial f(\mathbf{x})$, then

5. **Subgradient.** The normal cone is generated by the subgradient:

$$\mathcal{N}(S, \mathbf{x}) = \bigcup_{\lambda \geq 0} \lambda \cdot \partial f(\mathbf{x}). \quad (3.29)$$

Proof. Again, we organize the demonstration according to the headings.

(Sublevels) We have $f(\mathbf{x} + \lambda \mathbf{y}) \leq f(\mathbf{x})$ for some $\lambda > 0$ if and only if $\lambda \mathbf{y} \in S + \{-\mathbf{x}\}$. This is of course equivalent to $\mathbf{y} \in \lambda^{-1}(S + \{-\mathbf{x}\})$. The conclusion follows by noting that $\lambda > 0$ if and only if $\lambda^{-1} > 0$ and comparing the definition (3.24) with the claim (3.26).

(Convexity) Since the sublevel sets of a convex function are convex, the set $S + \{-\mathbf{x}\}$ is convex. Convexity of $\mathcal{D}(f, \mathbf{x})$ then follows from the fact that the cone generated by a convex set is convex [Roc70, Cor. 2.6.3] and (3.26). The fact that $\mathbf{0} \in \mathcal{D}(f, \mathbf{x})$ follows from the fact that $\mathbf{x} \in S$, and hence $\mathbf{0} \in S + \{-\mathbf{x}\}$.

(Polarity) This follows readily from the definition (3.25) of the normal cone and the

characterization (3.26) of the descent cone in terms of the sublevel set S .

(*Minima*) Suppose $\mathcal{D}(f, \mathbf{x}) = \{\mathbf{0}\}$. Then for each perturbation $\mathbf{y} \neq \mathbf{0}$ about \mathbf{x} , we have $f(\mathbf{x} + \mathbf{y}) > f(\mathbf{x})$ by definition (3.24) of the descent cone. Hence \mathbf{x} is the unique global minimum of f .

On the other hand, suppose that \mathbf{x} is the unique global minimum of f . Then for each $\mathbf{y} \neq \mathbf{0}$, we have $f(\mathbf{x} + \mathbf{y}) > f(\mathbf{x})$, and hence $\mathbf{y} \notin \mathcal{D}(f, \mathbf{x})$ by definition. Because $\mathbf{0} \in \mathcal{D}(f, \mathbf{x})$ by the convexity properties (#2) of the descent cone, we see that in fact $\mathcal{D}(f, \mathbf{x}) = \{\mathbf{0}\}$.

We have shown $\mathcal{D}(f, \mathbf{x}) = \{\mathbf{0}\}$ if and only if \mathbf{x} is the unique global minimizer of f . But $\mathcal{D}(f, \mathbf{x}) = \{\mathbf{0}\}$ if and only if $\mathcal{N}(S, \mathbf{x}) = \mathbb{R}^d$ by the polarity relation (3.27) and the fact that $\mathbf{0} \in \mathcal{D}(f, \mathbf{x})$. Hence (3.28) holds.

(*Subgradient.*) This follows from [Roc70, Cor. 23.7.1] under our assumptions on $\partial f(\mathbf{x})$. □

3.3 The topology of convex cones

By identifying a cone containing the origin with its restriction to the unit sphere, we obtain a one-to-one correspondence between the set \mathcal{C}_d of nonempty closed convex cones and the set of closed geodesically connected subsets of the unit sphere (the *spherically convex bodies*). The goal of this section is to introduce the *conic Hausdorff metric* on \mathcal{C}_d that is obtained by this identification, and to show that the Euclidean projection is continuous under this metric.

The conic Hausdorff metric appears briefly in connection with the integral-geometric development of Section 5. The continuity result (Proposition 3.8) is important for a new integral-geometric formula that we develop in Appendix A. Beyond these two fine points, this section may be skipped without much loss for what follows.

Our first priority is to describe a metric on the set \mathcal{C}_d of closed convex cones that agrees with the commonly used metric on the sphere S^{d-1} . For this task, we must introduce a few pieces of notation. The *arclength distance* $\text{dist}_s(\mathbf{x}, \mathbf{y})$ between two unit

vectors $\mathbf{x}, \mathbf{y} \in S^{d-1}$ is given by

$$\text{dist}_s(\mathbf{x}, \mathbf{y}) := \cos^{-1}(\langle \mathbf{x}, \mathbf{y} \rangle).$$

We extend the arclength distance to pairs of cones C, D in the usual manner:

$$\text{dist}_s(C, D) = \inf_{\substack{\mathbf{x} \in C \cap S^{d-1} \\ \mathbf{y} \in D \cap S^{d-1}}} \text{dist}_s(\mathbf{x}, \mathbf{y}).$$

By convention, the distance is zero if $C = D = \{\mathbf{0}\}$ and infinite if precisely one of $C = \{\mathbf{0}\}$ or $D = \{\mathbf{0}\}$. The *angular extension* $\mathcal{T}_s(C, \alpha)$ of angle $\alpha > 0$ around a cone $C \in \mathcal{C}_d$ is the union of all rays within an arclength distance of α to C :

$$\mathcal{T}_s(C, \alpha) := \{\mathbf{x} \in \mathbb{R}^d : \mathbf{x} \neq \mathbf{0} \text{ and } \text{dist}_s(\mathbf{x} / \|\mathbf{x}\|, C) \leq \alpha\} \cup \{\mathbf{0}\}.$$

Note that even if a cone C is convex, its extension $\mathcal{T}_s(C, \alpha)$ may not be convex for any $\alpha > 0$. The tube around a nontrivial subspace, for example, is never convex.

The angular tube induces the *conic Hausdorff metric* $\text{dist}_{\mathcal{H}}(C, D)$ between any two cones $C, D \in \mathcal{C}_d$ via

$$\text{dist}_{\mathcal{H}}(C, D) := \inf\{\alpha \geq 0 : \mathcal{T}_s(C, \alpha) \supset D \text{ and } \mathcal{T}_s(D, \alpha) \supset C\}.$$

Throughout this work, the topology on \mathcal{C}_d is the one induced by the conic Hausdorff metric.

From the correspondence between \mathcal{C}_d and the set of spherically convex bodies, it follows that the set of polyhedral cones forms a dense subset of all closed convex cones under the conic Hausdorff metric (see [SW08, p. 252]). This fact allows us to evaluate continuous functions on all cones in \mathcal{C}_d by determining its value on a sequence of polyhedral cones and taking a limit. To apply this procedure, we must first ensure that the function is continuous in the conic Hausdorff metric. The next proposition verifies that the Euclidean projection onto a cone is continuous in the conic Hausdorff metric.

Proposition 3.8 (Continuity of the projection). *Consider a sequence $(C_i)_{i \in \mathbb{N}}$ of cones in \mathcal{C}_d such that $C_i \rightarrow C \in \mathcal{C}_d$ as $i \rightarrow \infty$. For every $\mathbf{x} \in \mathbb{R}^d$, the projection $\Pi_{C_i}(\mathbf{x}) \rightarrow \Pi_C(\mathbf{x})$ as $i \rightarrow \infty$.*

Proof. We first dispense with a trivial case. If $C = \{\mathbf{0}\}$, then $C_i = \{\mathbf{0}\}$ for all i sufficiently large by our convention about the spherical distance to the trivial cone. Therefore $\Pi_{C_i}(\mathbf{x}) = \mathbf{0}$ for all i large enough, and the claim trivially holds.

For the remainder of the proof, we assume that $C \neq \{\mathbf{0}\}$. Then the convergence $C_i \rightarrow C$ implies that $C_i \neq \{\mathbf{0}\}$ for all sufficiently large i . By restricting to sufficiently large indices i , we may assume without further loss that $C_i \neq \{\mathbf{0}\}$ for every i .

Define $\mathbf{y}_i := \Pi_{C_i}(\mathbf{x})$. We claim that the sequence $(\mathbf{y}_i)_{i \in \mathbb{N}}$ has a limit point $\mathbf{y}_* \in C$. To see that a limit point exists, note that the sequence is bounded by $\|\mathbf{y}_i\| \leq \|\mathbf{x}\|$ owing to the Pythagorean identity (3.16). The Euclidean ball of radius $\|\mathbf{x}\|$ is compact, so there exists a convergent subsequence $\mathbf{y}_{i_j} \rightarrow \mathbf{y}_*$ as $j \rightarrow \infty$.

Next, we argue that the limit point $\mathbf{y}_* \in C$. Suppose, for a contradiction, that the limit point $\mathbf{y}_* \notin C$. Then, for all j large enough, we have $\text{dist}(\mathbf{y}_{i_j}, C) > \varepsilon > 0$ from the fact that $\mathbf{y}_{i_j} \rightarrow \mathbf{y}_*$ and continuity of the distance to a cone in the Euclidean metric. But $\mathbf{y}_{i_j} \in C_{i_j}$ and $\|\mathbf{y}_{i_j}\|$ is uniformly bounded by $\|\mathbf{x}\|$, so that this inequality contradicts the fact that $C_{i_j} \rightarrow C$ in the conic Hausdorff metric. Hence $\mathbf{y}_* \in C$, as claimed.

For the next step, we show that $\mathbf{y}_* = \Pi_C(\mathbf{x})$. For any $\varepsilon > 0$ and all sufficiently large j (how large depends on ε), we have

$$\text{dist}(\mathbf{x}, C) \leq \|\mathbf{x} - \mathbf{y}_*\| \leq \|\mathbf{x} - \mathbf{y}_{i_j}\| + \|\mathbf{y}_{i_j} - \mathbf{y}_*\| \leq \text{dist}(\mathbf{x}, C_{i_j}) + \varepsilon. \quad (3.30)$$

The first inequality follows because $\mathbf{y}_* \in C$. The second is the triangle inequality, and the third arises from the definition (3.13) of the distance and the fact that $\mathbf{y}_{i_j} \rightarrow \mathbf{y}_*$. Taking $\varepsilon \rightarrow 0$, we see

$$\liminf_{j \rightarrow \infty} \text{dist}(\mathbf{x}, C_{i_j}) \geq \text{dist}(\mathbf{x}, C). \quad (3.31)$$

For the reverse inequality, define the true projection $\mathbf{z} := \Pi_C(\mathbf{x})$ with norm $\beta := \|\mathbf{z}\|$. By definition of convergence in the conic Hausdorff metric, for any $\varepsilon > 0$ there exists a

unit norm $\mathbf{z}_i^\varepsilon \in C_i$ such that $\langle \mathbf{z}_i^\varepsilon, \mathbf{z} \rangle \geq (1 - \varepsilon)\beta$ for all i large enough. (The fact that \mathbf{z}_i^ε has unit norm requires the restriction that $C_i \neq \{\mathbf{0}\}$ which, as discussed above, incurs no loss.)

By expanding the squared norm, we find that the inequality

$$\|\mathbf{z} - \beta \mathbf{z}_i^\varepsilon\|^2 = 2\beta(\beta - \langle \mathbf{z}, \mathbf{z}_i^\varepsilon \rangle) \leq 2\varepsilon\beta^2 \leq 2\varepsilon \|\mathbf{x}\|^2$$

holds for all i sufficiently large. The equality follows from the definition of β . The second inequality above follows from the Pythagorean identity (3.16). In the same manner as (3.30), we find

$$\text{dist}(\mathbf{x}, C_i) \leq \|\mathbf{x} - \beta \mathbf{z}_i^\varepsilon\| \leq \|\mathbf{x} - \mathbf{z}\| + \|\beta \mathbf{z}_i^\varepsilon - \mathbf{z}\| \leq \text{dist}(\mathbf{x}, C) + \sqrt{2\varepsilon} \|\mathbf{x}\|$$

for all i large enough. Taking $\varepsilon \rightarrow 0$, we find the inequality

$$\limsup_{i \rightarrow \infty} \text{dist}(\mathbf{x}, C_i) \leq \text{dist}(\mathbf{x}, C).$$

Since the limit superior of a subsequence is no larger than the limit superior of the entire sequence, the inequality above combined with (3.31) implies that we have the equality

$$\lim_{j \rightarrow \infty} \text{dist}(\mathbf{x}, C_{i_j}) = \text{dist}(\mathbf{x}, C).$$

Taking the limit as $\varepsilon \rightarrow 0$ and $j \rightarrow \infty$ in (3.30), we conclude that $\|\mathbf{x} - \mathbf{y}_*\| = \text{dist}(\mathbf{x}, C)$. By uniqueness of the projection, we see that the limit point $\mathbf{y}_* = \Pi_C(\mathbf{x})$. This same argument works for *any* limit point of the sequence $(\mathbf{y}_i)_{i \in \mathbb{N}}$, so we conclude that the limit point is unique. Therefore, the limit $\mathbf{y}_i \rightarrow \Pi_C(\mathbf{x})$, as claimed. \square

Remark 3.9. Amelunxen & Bürgisser [AB12a, App. A] have also considered the conic Hausdorff metric, but they did not verify the continuity of the Euclidean projection of a point. This makes Proposition 3.8 a (minor) novel contribution of this thesis.

Chapter 4

The geometry of demixing

This chapter develops a characterization of success for the convex demixing method (1.6) in terms of a configuration of cones. The basic result, Theorem 4.1, states that constrained demixing succeeds if and only if the intersection of two appropriate descent cones is trivial. This result is a corollary of a more intricate geometric characterization of multiple demixing, as discussed in Section 4.2.

We look at compressed demixing in Section 4.3, and we find that it can be cast into the framework of multiple demixing. As a corollary, we recover well-known optimality conditions for regularized linear inverse problems from the optimality conditions for demixing. In Section 4.5, we describe how the constrained formulation of demixing relates to the Lagrangian formulation often used in practice. While this result is well known in the theory of convex optimization, our proof technique is based on the geometric characterization of success we develop in this section.

The optimality condition of Theorem 4.1 appears in the joint work [MT12, Lem. 2.3] with a different proof. The rest of this section is novel, with the exception of the well-known equivalence between the Lagrange and constrained demixing programs that we discuss Section 4.5.

4.1 Optimality conditions for constrained demixing

Let us briefly recall the constrained demixing formulation from Section 1.2.2. We model an observed signal $\mathbf{z}_0 \in \mathbb{R}^d$ as the sum of two unknown signals

$$\mathbf{z}_0 = \mathbf{x}^\natural + U\mathbf{y}^\natural, \quad (4.1)$$

where $U \in O_d$ is a known orthogonal matrix that models the relative orientation of the signals. We are also given two convex functions f and g , perhaps atomic norms from Section 1.2.1, that promote the structures we expect to see in \mathbf{x}^\natural and \mathbf{y}^\natural . Given the additional side information $\alpha := g(\mathbf{y}^\natural)$, we attempt to recover the unknown vectors \mathbf{x}^\natural and \mathbf{y}^\natural by solving the *constrained demixing* program

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } g(\mathbf{y}) \leq \alpha \quad \text{and } \mathbf{x} + U\mathbf{y} = \mathbf{z}_0, \quad (4.2)$$

with decision variables $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. We say that (4.2) *succeeds* if $(\mathbf{x}^\natural, \mathbf{y}^\natural)$ is the unique optimal point (4.2); otherwise, the method *fails*. It turns out that there is a simple characterization of success and failure in terms of a configuration of descent cones.

Theorem 4.1 (Demixing optimality conditions). *The demixing method (4.2) succeeds if and only if*

$$\mathcal{D}(f, \mathbf{x}^\natural) \cap (-U\mathcal{D}(g, \mathbf{y}^\natural)) = \{\mathbf{0}\}. \quad (4.3)$$

The intersection of descent cones always contains the zero vector by Proposition 3.6.2, so this result tells us that demixing succeeds if and only if the corresponding descent cones do not share a ray. See Figure 4.1 for an illustration of the optimality condition (4.3).

4.2 Multiple demixing

Our multiple demixing approach is a natural extension from our method (4.2) for demixing two signals to demixing the superposition of an arbitrary number of signals.

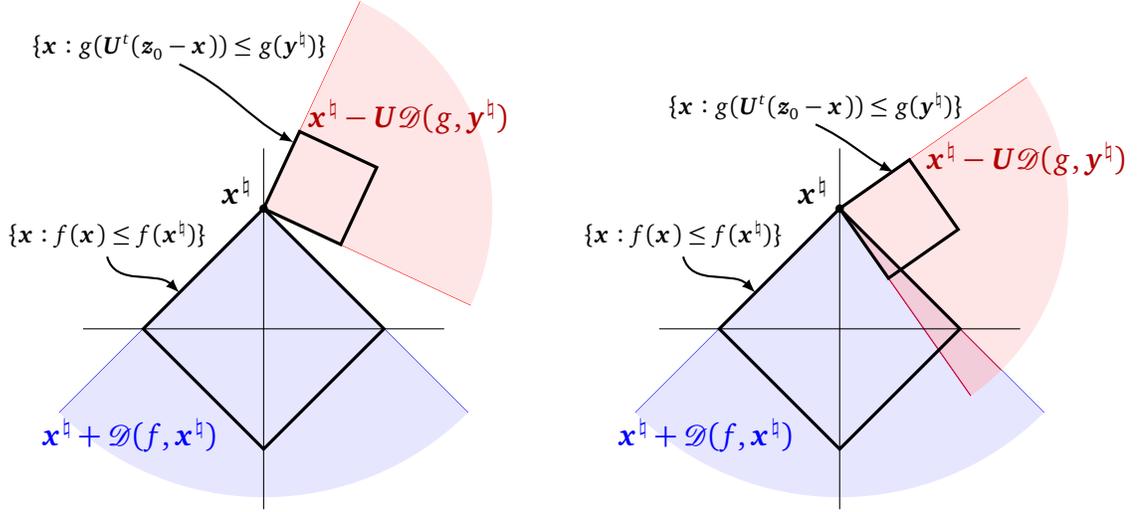


Figure 4.1: Geometry of optimality conditions. [Left] The success condition for the convex demixing method (4.3) states that the descent cone of f at \mathbf{x}_i^h has a trivial intersection with a rotated copy of the descent cone of g at \mathbf{y}_i^h . [Right] When the descent cones share a ray, the demixing method (4.2) is guaranteed to fail. This figure is drawn from the joint work [ALMT13].

In this situation, we observe a superposition of n signals:

$$\mathbf{z}_0 = \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i^h, \quad (4.4)$$

where $\mathbf{U}_i \in \mathcal{O}_d$ are known orthogonal matrices. We choose convex functions f_i that reflect the structure we expect to find in the constituents \mathbf{x}_i^h for each $i = 1, \dots, n-1, n$. Given the additional side information $\alpha_i := f_i(\mathbf{x}_i^h)$, we pose the following convex demixing method:

$$\begin{aligned} & \text{minimize} && f_1(\mathbf{x}_1) \\ & \text{subject to} && f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for } i = 2, \dots, n-1, n \quad \text{and} \quad \mathbf{z}_0 = \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i, \end{aligned} \quad (4.5)$$

where the optimization variable is the tuple $(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n)$. As in the case of basic demixing (4.2), the multiple demixing method (4.5) *succeeds* if and only if the tuple $(\mathbf{x}_1^h, \dots, \mathbf{x}_{n-1}^h, \mathbf{x}_n^h)$ is the unique optimal point of (4.5). Again, we obtain a characterization of success and failure in terms of descent cone configurations.

Theorem 4.2 (Multiple demixing optimality conditions). *The demixing method (4.5)*

succeeds if and only if

$$\mathbf{U}_i \mathcal{D}(f_i, \mathbf{x}_i^{\natural}) \cap - \left(\sum_{j \neq i} \mathbf{U}_j \mathcal{D}(f_j, \mathbf{x}_j^{\natural}) \right) = \{\mathbf{0}\} \quad \text{for each } i = 1, \dots, n-1, n. \quad (4.6)$$

In other words, multiple demixing succeeds if and only if no descent cone shares a ray with the (negative) Minkowski sum of all the others.

The proof is deferred until Section 4.4. For now, we show how the optimality conditions of basic demixing (4.2) are a special case of the multiple demixing optimality conditions.

Proof of Theorem 4.1 from Theorem 4.2. Make the identifications

$$(f_1, f_2, \mathbf{U}_1, \mathbf{U}_2, \mathbf{x}_1^{\natural}, \mathbf{x}_2^{\natural}) \leftrightarrow (f, g, \mathbf{I}, \mathbf{U}, \mathbf{x}^{\natural}, \mathbf{y}^{\natural}).$$

The $n = 2$ case of Theorem 4.2 then shows that (4.2) succeeds if and only if

$$\mathcal{D}(f, \mathbf{x}^{\natural}) \cap -\mathbf{U} \mathcal{D}(g, \mathbf{y}^{\natural}) = \{\mathbf{0}\} \quad \text{and} \quad \mathbf{U} \mathcal{D}(g, \mathbf{y}^{\natural}) \cap -\mathcal{D}(f, \mathbf{x}^{\natural}) = \{\mathbf{0}\}.$$

The equations above are equivalent, so Theorem 4.1 follows immediately. \square

4.3 Compressed demixing

Another variation on the demixing theme is a *compressed* version of demixing. In this setup, the observation $\mathbf{z}_0 \in \mathbb{R}^m$ is an undersampled version of the superimposed signal:

$$\mathbf{z}_0 = \mathbf{A} \left(\sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i^{\natural} \right), \quad (4.7)$$

where $\mathbf{A} \in \mathbb{R}^{m \times d}$ is an undersampling operator (that is, $m < d$) and the \mathbf{U}_i are known orthogonal matrices. As usual, the goal is to recover the unknown natural signals \mathbf{x}_i^{\natural} from the superimposed and undersampled observation \mathbf{z}_0 .

Given a convex function f_i that promotes the structure of \mathbf{x}_i^{\natural} and the side information $\alpha_i := f_i(\mathbf{x}_i^{\natural})$ for each $i = 1, \dots, n-1, n$, we pose the following *compressed demixing*

program

$$\begin{aligned} & \text{minimize} && f_1(\mathbf{x}_1) \\ & \text{subject to} && f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for } i = 2, \dots, n-1, n, \quad \text{and} \quad \mathbf{z}_0 = \mathbf{A} \left(\sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i \right), \end{aligned} \quad (4.8)$$

where the optimization variable is the tuple $(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n)$. The compressed demixing program *succeeds* if the tuple $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_{n-1}^{\natural}, \mathbf{x}_n^{\natural})$ is the unique optimal point of (4.8).

We can characterize the undersampled version of demixing in terms of the configuration of convex cones. To simplify the statement of the theorem, we define the cones $C_i := \mathbf{U}_i \mathcal{D}(f_i, \mathbf{x}_i^{\natural})$ for $i = 1, \dots, n-1, n$ and $C_{n+1} := \text{null}(\mathbf{A})$, the nullspace of the measurement operator \mathbf{A} .

Theorem 4.3 (Compressed demixing optimality conditions). *For each $i = 1, \dots, n, n+1$, define the cones C_i as in the prior paragraph. The compressed demixing method (4.8) succeeds if and only if*

$$C_i \cap - \left(\sum_{j \neq i} C_j \right) = \{\mathbf{0}\} \quad \text{for each } i = 1, \dots, n, n+1. \quad (4.9)$$

There is a strong structural similarity between the optimality condition (4.9) for compressed demixing and the condition (4.5) for multiple demixing. This is no coincidence. The entirety of the proof of Theorem 4.3 consists of a reduction from the compressed demixing scheme (4.8) to a special case of multiple demixing (4.5).

Proof. We start with a few definitions. Let $\hat{\mathbf{z}}_0 := \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i^{\natural}$ be the uncompressed observation and define $f_{n+1}(\mathbf{x}) := \mathbf{i}_{\text{null}(\mathbf{A})}(\mathbf{x})$, the convex indicator (3.2) of the nullspace of \mathbf{A} . Fix the rotation $\mathbf{U}_{n+1} := \mathbf{I} \in \mathbf{O}_d$ and the side information $\alpha_{n+1} := 0$. We claim that the compressed demixing problem (4.8) succeeds at demixing $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_n^{\natural})$ if and only if the multiple demixing problem

$$\begin{aligned} & \text{minimize} && f_1(\hat{\mathbf{x}}_1) \\ & \text{subject to} && f_i(\hat{\mathbf{x}}_i) \leq \alpha_i \quad \text{for } i = 2, \dots, n, n+1 \quad \text{and} \quad \hat{\mathbf{z}}_0 = \sum_{i=1}^{n+1} \mathbf{U}_i \hat{\mathbf{x}}_i \end{aligned} \quad (4.10)$$

succeeds at demixing $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_n^{\natural}, \mathbf{0})$, where the decision variables are $\hat{\mathbf{x}}_i \in \mathbb{R}^d$.

To show the equivalence of the programs (4.8) and (4.10), we demonstrate that there is a bijection between the feasible points of (4.8) and (4.10) that preserves the objective values and maps $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_n^{\natural})$ to $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_n^{\natural}, \mathbf{0})$. We show that the map defined by

$$(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \mapsto (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n, \hat{\mathbf{z}}_0 - \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i) =: (\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_n, \hat{\mathbf{x}}_{n+1}) \quad (4.11)$$

has this property.

Indeed, suppose that the tuple $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is feasible for the compressed demixing problem (4.8). Then the objective value is unchanged; that is, $f_1(\hat{\mathbf{x}}_1) = f_1(\mathbf{x}_1)$. Moreover,

$$f_i(\hat{\mathbf{x}}_i) = f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for all } i = 2, \dots, n-1, n.$$

We now verify that $f_{n+1}(\hat{\mathbf{x}}_{n+1}) \leq \alpha_{n+1}$. Our definition of $\hat{\mathbf{z}}_0$ ensures

$$\mathbf{A} \left(\hat{\mathbf{z}}_0 - \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i \right) = \mathbf{z}_0 - \mathbf{A} \left(\sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i \right) = \mathbf{0}$$

by feasibility of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$. Hence,

$$f_{n+1}(\hat{\mathbf{x}}_{n+1}) = i_{\text{null}(\mathbf{A})} \left(\hat{\mathbf{z}}_0 - \sum_{i=1}^n \mathbf{U}_i \mathbf{x}_i \right) = 0 \leq \alpha_{n+1},$$

where the inequality follows by our definition $\alpha_{n+1} := 0$. Thus, all inequality constraints in (4.10) are satisfied by the image point. The equality constraint is trivially satisfied by definition of $\hat{\mathbf{x}}_{n+1}$, and hence the map (4.10) takes feasible points of (4.8) to feasible points of (4.10) with the same objective value.

The other direction requires similar considerations. Let $(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_{n+1})$ be feasible for (4.10). Then the preimage $(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_n)$ under the map (4.11) clearly has the same objective value and satisfies all of the inequality constraints in (4.8). Moreover, the constraint

$$f_{n+1}(\hat{\mathbf{x}}_{n+1}) = i_{\text{null}(\mathbf{A})}(\hat{\mathbf{x}}_{n+1}) \leq \alpha_{n+1} = 0$$

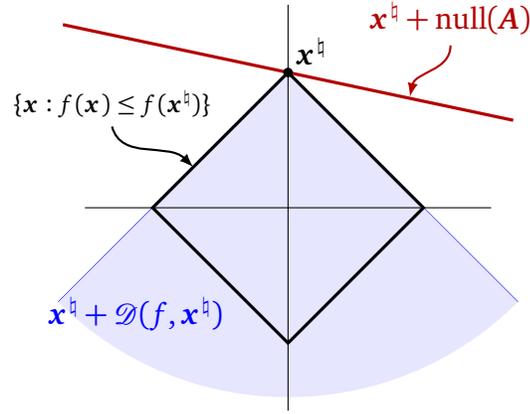


Figure 4.2: Optimality conditions for linear inverse problems. Linear inverse problems (4.13) are a special case of compressed demixing. The optimality condition for this linear inverse problems involves only the configuration of a nullspace (*line*) and a descent cone (*shaded*). The figure above is a fixture in the literature of compressed sensing; cf. [CRT06b, Fig. 2.1].

implies that $A\hat{\mathbf{x}}_{n+1} = \mathbf{0}$ by the definition (3.2) of the convex indicator function, and hence

$$A \left(\sum_{i=1}^n U_i \mathbf{x}_i \right) = A(\hat{\mathbf{z}}_0 - \hat{\mathbf{x}}_{n+1}) = \mathbf{z}_0,$$

where the final inequality uses the definition of $\hat{\mathbf{z}}_0$. Therefore the map (4.11) provides a bijection between the feasible points and objective values of the programs (4.8) and (4.10). Since (4.11) trivially takes $(\mathbf{x}_1^h, \dots, \mathbf{x}_n^h)$ to $(\mathbf{x}_1^h, \dots, \mathbf{x}_n^h, \mathbf{0})$ by definition of $\hat{\mathbf{z}}_0$, we conclude that (4.8) succeeds if and only if (4.10) succeeds at demixing $(\mathbf{x}_1^h, \dots, \mathbf{x}_n^h, \mathbf{0})$.

The claim now follows immediately from this equivalence by the optimality conditions for multiple demixing (4.6) given by Theorem 4.2 and the fact that $\mathcal{D}(i_{\text{null}(A)}, \mathbf{0}) = \text{null}(A)$. This completes the proof. \square

4.3.1 Linear inverse problems

The $n = 1$ case of the compressed demixing problem (4.8) deserves a special remark. In this case, the observation \mathbf{z}_0 consists simply of an undersampled structured vector:

$$\mathbf{z}_0 = A\mathbf{x}^h. \quad (4.12)$$

This is the setup of the generic *linear inverse problem* studied in [CRPW12]. With a convex regularizer f associated with the structure of \mathbf{x}^\natural , the corresponding demixing method reduces to

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } \mathbf{A}\mathbf{x} = \mathbf{z}_0. \quad (4.13)$$

By Theorem 4.3, the linear inverse program succeeds if and only if

$$\mathcal{D}(f, \mathbf{x}^\natural) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\},$$

where we use the fact that $\text{null}(\mathbf{A}) = -\text{null}(\mathbf{A})$ by linearity. This condition is precisely that given in [CRPW12, Prop. 2.1]. We provide the obligatory illustration of this condition in Figure 4.2.

4.4 Proof of Theorem 4.2

Before diving into the proof, we develop two facts about descent cones. The first result formalizes a basic but fundamental observation: if a convex function *eventually* decreases in some direction, then it *locally* decreases in that direction.

Proposition 4.4 (Local descent). *Let f be a convex function. Then $\mathbf{y} \in \mathcal{D}(f, \mathbf{x})$ if and only if there exists a $\lambda_0 > 0$ such that $f(\mathbf{x} + \lambda\mathbf{y}) \leq f(\mathbf{x})$ for all $\lambda \in [0, \lambda_0]$.*

Proof. The “if” part is immediate. Given any $\lambda_0 > 0$ such that $f(\mathbf{x} + \lambda_0\mathbf{y}) \leq f(\mathbf{x})$, we have $\mathbf{y} \in \mathcal{D}(f, \mathbf{x})$ by definition (3.24).

For the other direction, suppose $\mathbf{y} \in \mathcal{D}(f, \mathbf{x})$. Then the definition of a feasible cone ensures that there exists a number $\lambda_0 > 0$ for which $f(\mathbf{x} + \lambda_0\mathbf{y}) \leq f(\mathbf{x})$. Suppose $\lambda \in [0, \lambda_0]$. By the convexity of f , we have

$$\begin{aligned} f(\mathbf{x} + \lambda\mathbf{y}) &= f\left(\left(1 - \frac{\lambda}{\lambda_0}\right)\mathbf{x} + \frac{\lambda}{\lambda_0}(\mathbf{x} + \lambda_0\mathbf{y})\right) \\ &\leq \left(1 - \frac{\lambda}{\lambda_0}\right)f(\mathbf{x}) + \frac{\lambda}{\lambda_0}f(\mathbf{x} + \lambda_0\mathbf{y}) \end{aligned}$$

Applying the inequality $f(\mathbf{x} + \lambda_0\mathbf{y}) \leq f(\mathbf{x})$, we find that the last expression above is

less than $f(\mathbf{x})$. This completes the demonstration. \square

Our next result is a transformation rule for descent cones under linear maps. The result shares some similarities with common subgradient transformation rules for convex functions [Roc70, Sec. 23], but it has a slightly different flavor because the descent cone is a primal object while the subgradient is a dual object.

Proposition 4.5 (Descent cones and linear transformations.). *Let g be any convex function and define $h(\mathbf{x}) := g(\mathbf{A}\mathbf{x})$ for some matrix \mathbf{A} . Then $\mathcal{D}(h, \mathbf{x}) = \mathcal{D}(g, \mathbf{A}\mathbf{x})$.*

Proof. The inclusion $\mathbf{y} \in \mathcal{D}(h, \mathbf{x})$ occurs if and only if there exists a $\lambda > 0$ such that

$$h(\mathbf{x} + \lambda\mathbf{y}) \leq h(\mathbf{x}).$$

Expanding the definition of h , this inequality is equivalent to

$$g(\mathbf{A}(\mathbf{x} + \lambda\mathbf{y})) \leq g(\mathbf{A}\mathbf{x})$$

which occurs if and only if $\lambda\mathbf{A}\mathbf{y} \in \mathcal{D}(g, \mathbf{A}\mathbf{x})$ by linearity of \mathbf{A} . Since $\lambda > 0$ and cones are positive homogeneous, the final inclusion holds if and only if $\mathbf{A}\mathbf{y} \in \mathcal{D}(g, \mathbf{A}\mathbf{x})$. This string of equivalences establishes the claim. \square

Proof of Theorem 4.2. To avoid notational clutter, we prove the result when $\mathbf{U}_i = \mathbf{I}$ for every $i = 1, \dots, n-1, n$. This special case easily implies the general case $\mathbf{U}_i \neq \mathbf{I}$ via the linear transformation rule for descent cones given by Proposition 4.5. The details of this reduction are given at the end of the proof.

(\Rightarrow) Suppose first that the tuple $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_{n-1}^{\natural}, \mathbf{x}_n^{\natural})$ is the unique optimal point for the multiple demixing method (4.5). We must show that the intersection relations (4.6) hold. To this end, let $i \in \{1, \dots, n-1, n\}$ and suppose that

$$\mathbf{y}_i \in \mathcal{D}(f_i, \mathbf{x}_i^{\natural}) \cap - \left(\sum_{j \neq i} \mathcal{D}(f_j, \mathbf{x}_j^{\natural}) \right).$$

We will show that $\mathbf{y}_i = \mathbf{0}$. By the definition of a Minkowski sum, there exist vectors

$\mathbf{y}_j \in \mathcal{D}(f_j, \mathbf{x}_j^{\natural})$ such that

$$\mathbf{y}_i = - \sum_{j \neq i} \mathbf{y}_j. \quad (4.14)$$

Because each of the summands lies in a descent cone, Proposition 4.4 ensures the existence of a $\lambda > 0$ such that

$$f_j(\mathbf{x}_j^{\natural} + \lambda \mathbf{y}_j) \leq f_j(\mathbf{x}_j^{\natural}) \quad \text{for all } j = 1, \dots, n-1, n.$$

Moreover, we have

$$\sum_{j=1}^n \mathbf{x}_j^{\natural} + \lambda \mathbf{y}_j = \sum_{j=1}^n \mathbf{x}_j^{\natural} = \mathbf{z}_0$$

by equation (4.14) and the definition (4.4) of \mathbf{z}_0 . By combining the last two displays, we see that the tuple $(\mathbf{x}_1^{\natural} + \lambda \mathbf{y}_1, \dots, \mathbf{x}_n^{\natural} + \lambda \mathbf{y}_n)$ is feasible for (4.5) and has an objective value no larger than the minimal objective value. Since $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_n^{\natural})$ is the unique minimizer of (4.5) by assumption, we deduce in particular that $\mathbf{y}_i = \mathbf{0}$. Because the index i was arbitrary, we conclude that condition (4.6) holds.

(\Rightarrow) Suppose now that $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_{n-1}^{\natural}, \mathbf{x}_n^{\natural})$ is *not* the unique optimal point of (4.5). Then there exists a distinct tuple $(\mathbf{x}_1^{\flat}, \dots, \mathbf{x}_{n-1}^{\flat}, \mathbf{x}_n^{\flat})$ that is optimal for (4.5). Define the perturbations $\mathbf{y}_i := \mathbf{x}_i^{\flat} - \mathbf{x}_i^{\natural}$ for all $i = 1, \dots, n-1, n$. Optimality and feasibility of the \flat -tuple leads to the inequalities

$$f_i(\mathbf{x}_i^{\natural} + \mathbf{y}_i) = f_i(\mathbf{x}_i^{\flat}) \leq f(\mathbf{x}_i^{\natural}) \quad \text{for all } i = 1, \dots, n-1, n.$$

These inequalities imply the inclusion $\mathbf{y}_i \in \mathcal{D}(f_i, \mathbf{x}_i^{\natural})$ for each i by the definition (3.24) of descent cones. Since the \flat -tuple is feasible for (4.5), we also find

$$\sum_{i=1}^n \mathbf{y}_i = \sum_{i=1}^n \mathbf{x}_i^{\flat} - \mathbf{x}_i^{\natural} = \mathbf{z}_0 - \mathbf{z}_0 = \mathbf{0}.$$

The first equality is the definition of \mathbf{y}_i , while the second follows from feasibility of \natural and \flat tuples. Because the \flat -tuple is distinct from the \natural -tuple, we have $\mathbf{y}_{i_*} \neq \mathbf{0}$ for at least one index i_* . Rearranging the display above, we see that $\mathbf{y}_{i_*} = - \sum_{j \neq i_*} \mathbf{y}_j \in - \sum_{j \neq i_*} \mathcal{D}(f_j, \mathbf{x}_j^{\natural})$.

In particular,

$$\mathcal{D}(f_{i_*}, \mathbf{x}_{i_*}^{\natural}) \cap - \left(\sum_{j \neq i_*} \mathcal{D}(f_j, \mathbf{x}_j^{\natural}) \right) \neq \{\mathbf{0}\}.$$

In other words, the condition (4.6) does not hold. This completes the demonstration when $U_i = \mathbf{I}$ for each $i = 1, \dots, n-1, n$.

(Extension to general rotations) A simple change of variables extends this proof to the setting of arbitrary $U_i \in O_d$. For every $i = 1, \dots, n-1, n$, define $\hat{\mathbf{x}}_i^{\natural} := U_i \mathbf{x}_i^{\natural}$ and $h_i(\hat{\mathbf{x}}) := f(U^t \hat{\mathbf{x}})$. Then the demixing program (4.5) succeeds at identifying $(\mathbf{x}_1^{\natural}, \dots, \mathbf{x}_{n-1}^{\natural}, \mathbf{x}_n^{\natural})$ if and only if the program

$$\begin{aligned} & \text{minimize} && h_1(\hat{\mathbf{x}}_1) \\ & \text{subject to} && h_i(\hat{\mathbf{x}}_i) \leq \alpha_i \quad \text{for } i = 2, \dots, n-1, n \quad \text{and} \quad \mathbf{z}_0 = \sum_{i=1}^n \hat{\mathbf{x}}_i, \end{aligned}$$

has $(\hat{\mathbf{x}}_1^{\natural}, \dots, \hat{\mathbf{x}}_{n-1}^{\natural}, \hat{\mathbf{x}}_n^{\natural})$ as its unique optimal point. By our result above, this occurs if and only if

$$\mathcal{D}(g_i, \hat{\mathbf{x}}_i^{\natural}) \cap - \left(\sum_{j \neq i} \mathcal{D}(g_j, \hat{\mathbf{x}}_j^{\natural}) \right) = \{\mathbf{0}\} \quad \text{for all } i = 1, \dots, n-1, n.$$

Since $U^t = U^{-1}$ by orthogonality, Proposition 4.5 shows that $\mathcal{D}(g_j, \hat{\mathbf{x}}_j^{\natural}) = U_j \mathcal{D}(f_j, \mathbf{x}_j^{\natural})$. Applying this observation to the display above gives the claim (4.6). \square

4.5 The Lagrangian formulation

The results in this section allow us to interpret our conditions for the success of the constrained demixing method (4.2) as limits on, and opportunities for, its Lagrangian relative

$$\text{minimize } f(\mathbf{x}) + \lambda \cdot g(\mathbf{y}) \quad \text{subject to} \quad \text{and} \quad \mathbf{x} + U\mathbf{y} = \mathbf{z}_0. \quad (4.15)$$

As usual, we say that the Lagrange penalized formulation (4.15) succeeds if $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ is the unique optimal solution to (4.15).

Extensions of the results in this section to the multiple demixing method (4.5) follow along the same lines, but we restrict our attention to the basic demixing model to get the key ideas across. We begin with the following well-known result; it holds without any technical restrictions.

Proposition 4.6. *Suppose the Lagrange problem (4.15) succeeds for some value $\lambda > 0$. Then the constrained formulation (4.2) also succeeds.*

Proof. We check the contrapositive. Suppose that the constrained formulation (4.2) does not succeed. Then there exists a pair $(\mathbf{x}^b, \mathbf{y}^b)$ that is an optimal point for the constrained program (4.2) but is distinct from $(\mathbf{x}^h, \mathbf{y}^h)$. Because $(\mathbf{x}^h, \mathbf{y}^h)$ is feasible for the constrained method (4.2), we have the inequality $f(\mathbf{x}^b) \leq f(\mathbf{x}^h)$ by optimality of $(\mathbf{x}^b, \mathbf{y}^b)$. Feasibility of $(\mathbf{x}^b, \mathbf{y}^b)$ also implies the inequality $g(\mathbf{y}^b) \leq \alpha := g(\mathbf{y}^h)$. Combining these observations, we see

$$f(\mathbf{x}^b) + \lambda \cdot g(\mathbf{y}^b) \leq f(\mathbf{x}^h) + \lambda \cdot g(\mathbf{y}^h).$$

Hence, $(\mathbf{x}^h, \mathbf{y}^h)$ is not the unique optimal point of (4.15), so (4.15) does not succeed. \square

We have the following partial converse to Proposition 4.6 under some technical restrictions.

Proposition 4.7. *Suppose that \mathbf{x}^h is in the interior of $\text{dom}(f)$ and that \mathbf{y}^h is in the interior of $\text{dom}(g)$. Suppose further that the subgradients $\partial f(\mathbf{x}^h)$ and $\partial g(\mathbf{y}^h)$ are both nonempty and do not contain zero. If the constrained method (4.2) succeeds, then there exists a parameter $\lambda > 0$ such that $(\mathbf{x}^h, \mathbf{y}^h)$ is a solution to the Lagrange method (4.15).*

We remark that the conclusion of Proposition 4.7 provides something slightly less than a guarantee of success for the Lagrange problem because it does not guarantee that the solution to the Lagrange problem is unique. Nevertheless, it provides a near equivalence between the Lagrange problem and constrained problem that we analyze.

Proof. The key idea is the construction of a subgradient that certifies the optimality of the pair $(\mathbf{x}^h, \mathbf{y}^h)$ for the Lagrange penalized problem (4.15) for an appropriate choice of

parameter λ . As with many results in convex analysis, a separating hyperplane plays an important role.

By Theorem 4.1, the constrained problem (4.2) succeeds if and only if

$$\mathcal{D}(f, \mathbf{x}^\natural) \cap -U\mathcal{D}(g, \mathbf{y}^\natural) = \{\mathbf{0}\}.$$

This trivial intersection implies that there exists a hyperplane that separates these cones (Fact 3.3). In other words, there exists some vector $\mathbf{w} \neq \mathbf{0}$ such that

$$\langle \mathbf{w}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in \mathcal{D}(f, \mathbf{x}^\natural),$$

and moreover

$$\langle \mathbf{w}, \mathbf{y} \rangle \geq 0 \text{ for all } \mathbf{y} \in -U\mathcal{D}(g, \mathbf{y}^\natural).$$

In the language of polar cones, the first separation inequality is simply the statement that $\mathbf{w} \in \mathcal{D}(f, \mathbf{x}^\natural)^\circ$, while the second inequality is equivalent to $U^t \mathbf{w} \in \mathcal{D}(g, \mathbf{y}^\natural)^\circ$.

We will now show that \mathbf{w} generates a subgradient optimality certificate for the point $(\mathbf{x}^\natural, \mathbf{y}^\natural)$ in problem (4.15) for an appropriate choice of parameter $\lambda > 0$. At this point, we invoke our technical assumption. Since f is typical at \mathbf{x}^\natural , the polar to the feasible cone is generated by the subgradient of f at \mathbf{x}^\natural (cf. (3.29)). In particular, there exists a number $\lambda_f \geq 0$ such that $\mathbf{w} \in \lambda_f \partial f(\mathbf{x}^\natural)$. In fact, the stronger inequality $\lambda_f > 0$ holds because $\mathbf{w} \neq \mathbf{0}$. For the same reason, there exists a number $\lambda_g > 0$ such that $U^t \mathbf{w} \in \lambda_g \partial g(\mathbf{y}^\natural)$.

Define $h(\mathbf{x}) := \lambda_f f(\mathbf{x}) + \lambda_g g(U^t(\mathbf{z}_0 - \mathbf{x}))$. By standard transformation rules for subgradients [Roc70, Thms. 23.8, 23.9], we have

$$\partial h(\mathbf{x}^\natural) \supset \lambda_f \partial f(\mathbf{x}^\natural) - \lambda_g U \partial g(\mathbf{y}^\natural),$$

where $A - B := A + (-B)$ is the Minkowski sum of the sets A and $-B$. Since $\mathbf{w} \in \lambda_f \partial f(\mathbf{x}^\natural)$ and $\mathbf{w} \in \lambda_g U \partial g(\mathbf{y}^\natural)$, we see $\mathbf{0} \in \partial h(\mathbf{x}^\natural)$. By the definition of subgradients, \mathbf{x}^\natural is a global minimizer of h . Introducing the variable $\mathbf{y} = U^t(\mathbf{z}_0 - \mathbf{x})$, it follows that $(\mathbf{x}^\natural, \mathbf{y}^\natural)$ is a

global minimizer of

$$\text{minimize } f(\mathbf{x}) + \frac{\lambda_g}{\lambda_f} g(\mathbf{y}) \quad \text{subject to } \mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{z}_0.$$

This is Lagrange problem (4.15) with the parameter $\lambda = \lambda_g/\lambda_f > 0$, so we have the result. □

Chapter 5

Conic integral geometry

Convex cones characterize the optimality conditions (4.3) of demixing, so it is no surprise that a detailed study of the theory of convex cones is a central element of this thesis. The goal of this section is to introduce the theory of *conic integral geometry* that studies the invariant geometric features of convex cones and the interplay among randomly oriented convex cones.

We begin by introducing the basic geometric parameters of interest in Section 5.1 known as conic intrinsic volumes. We continue in Sections 5.2 and 5.3 with the more advanced formulas of conic integral geometry known as the *Steiner* and *kinematic* formulas. Section 5.4 considers unlikely events.

Most of the results presented in this section are standard. Minor novelties include the iterated kinematic formulas of Section 5.3.1—although the analogy with the Euclidean integral geometry makes these formulas obvious to those familiar with integral geometry. The results of Section 5.4 fill in other small gaps in the literature, and in particular, suggest that the closure assumptions usually assumed in the literature are extraneous for parts of the kinematic theory of convex cones.

5.1 Conic intrinsic volumes

Much like a convex body in Euclidean space has a volume, surface area, mean width, Euler characteristic, etc. [KR97], each convex cone has have a number of dimensionality parameters known as intrinsic volumes.

Definition 5.1. Let $C \in \mathcal{C}_d$ be a polyhedral cone. For each $i = 0, 1, \dots, d$, we define the i th (conic) intrinsic volume $v_i(C)$ of C by the probability that a Gaussian random vector projects into an i -dimensional face of C , that is

$$v_i(C) := \mathbb{P}\{\Pi_C(\mathbf{g}) \in \text{relint}(F_i) : F_i \text{ is an } i\text{-dimensional face of } C\}. \quad (5.1)$$

The uniqueness of the projection map ensures that the intrinsic volumes of polyhedral cones are well defined. The intrinsic volumes are continuous under the conic Hausdorff metric [SW08, Thm. 6.5.2], so we extend this definition to all cones via approximation.

The proof of continuity seems to be due to Glasauer [Gla95, Gla96]. Although the extension of the intrinsic volumes to nonpolyhedral cones by polyhedral approximation allows us to prove results about the intrinsic volumes via probabilistic methods, this approach has limited practical use for computing intrinsic volumes of nonpolyhedral cones. The Steiner formulas of Section 5.2 provide an alternate method for computing intrinsic volumes, and in fact the Steiner formula is frequently used to *define* conic intrinsic volumes. For simple examples, however, the probabilistic definition proves immanently useful.

Example 5.2 (Intrinsic volumes of subspaces). A subspace $L \subset \mathbb{R}^d$ of dimension k is a cone with precisely one face, so projections onto the subspace always lie in the relative interior of L . Therefore the intrinsic volumes of L form a Dirac sequence:

$$v_k(L) = \begin{cases} 1, & k = \dim(L), \\ 0, & \text{otherwise} \end{cases} \quad \text{for } k = 0, \dots, d-1, d. \quad (5.2)$$

Example 5.3 (Intrinsic volumes of the orthant [Ame11, Ex. 4.4.7]). The projection onto the nonnegative orthant $\mathbb{R}_+^d := \{\mathbf{x} : x_i \geq 0, i = 1, \dots, d\}$ is given by

$$[\Pi_{\mathbb{R}_+^d}(\mathbf{x})]_i = \max\{x_i, 0\},$$

and so $\Pi_{\mathbb{R}_+^d}(\mathbf{g})$ lies in the relative interior k -dimensional face of \mathbb{R}_+^d if and only if \mathbf{g} has

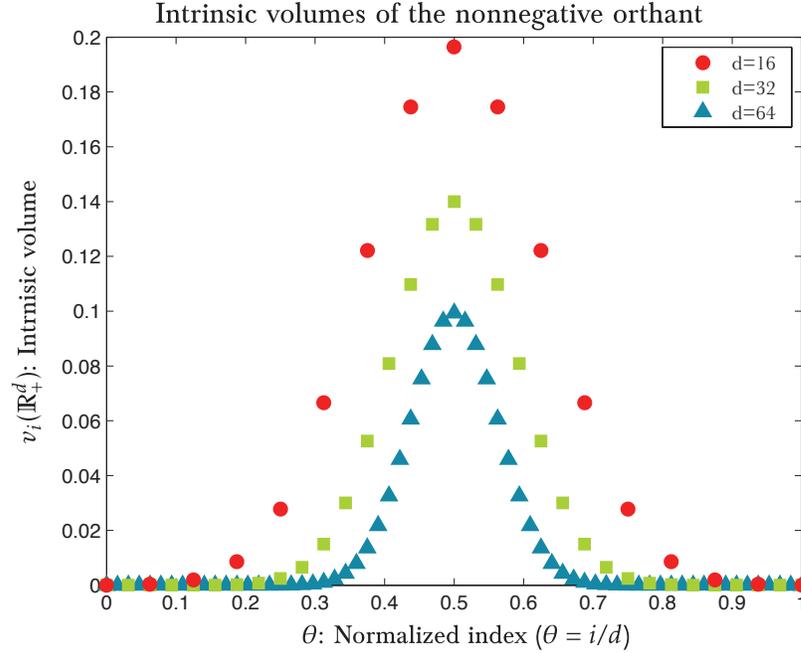


Figure 5.1: Intrinsic volumes of the nonnegative orthant. The intrinsic volumes of the nonnegative orthant are given by the normalized binomial sequence (5.3). The x -axis is the normalized index $\theta = i/d$, and in each case most of the mass of the intrinsic volumes is concentrated about $\theta = \frac{1}{2}$. The results of Section 7 show that this concentration behavior is generic.

exactly k positive values. Since the signs $\text{sgn}(g_i)$ of the elements of a Gaussian random vector are (almost surely) independent Bernoulli random variables, distribution of the number of positive entries of \mathbf{g} is given by the binomial formula. In other words,

$$v_k(\mathbb{R}_+^d) = 2^{-d} \binom{d}{k} \quad \text{for } k = 0, \dots, d-1, d. \quad (5.3)$$

See Figure 5.1 for an illustration.

The probabilistic definition of the intrinsic volumes reveals a number of fundamental properties of the intrinsic volumes.

Proposition 5.4 (Properties of the intrinsic volumes). *Let $C \in \mathcal{C}_d$ be a closed, convex cone. Then*

1. **Distribution.** *The intrinsic volumes form a probability distribution on $\{0, \dots, d -$*

$1, d\}$:

$$\sum_{k=0}^d v_k(C) = 1 \quad \text{and} \quad v_k(C) \geq 0 \quad \text{for all} \quad k = 0, \dots, d-1, d. \quad (5.4)$$

2. **Rotation invariance.** For any orthogonal matrix $U \in O_d$ and all $i = 0, \dots, d-1, d$, we have

$$v_k(C) = v_k(UC). \quad (5.5)$$

3. **Polarity reversal.** The intrinsic volumes reverse under polarity:

$$v_k(C) = v_{d-k}(C^\circ). \quad (5.6)$$

4. **Product.** For any $C' \in \mathcal{C}_d$, the intrinsic volumes of the direct product $C \times C'$ are given by

$$v_k(C \times C') = \sum_{i+j=k} v_i(C)v_j(C'). \quad (5.7)$$

5. **Gauss–Bonnet.** If C is not a subspace, then

$$\sum_{\substack{k=0 \\ k \text{ even}}}^d v_k(C) = \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_k(C) = \frac{1}{2}. \quad (5.8)$$

Proof. Suppose first that C is a polyhedral cone. Positivity of the intrinsic volume follows from positivity of probability, while the fact that the intrinsic volumes sum to one reflects the fact that the projection $\Pi_C(\mathbf{x})$ lies in the relative interior of a unique face of C . This shows the first point (5.4). Rotation invariance (5.5) follows immediately from the rotation invariance of the Gaussian measure. These two results extend to all cones by approximation with polyhedral cones owing to the continuity of intrinsic volumes under the conic Hausdorff metric.

The polarity law (5.6) appears in [SW08, Eq. (6.51)], but let us describe the basic idea. The projection $\Pi_C(\mathbf{x})$ lies in a k -dimensional face of a polyhedral cone C if and

only if¹ the projection $\Pi_{C^\circ}(\mathbf{x})$ lies in an $(d - k)$ -dimensional face of C° . This bijective correspondence leads to the polarity reversal property (5.6) for polyhedral cones, and it extends to nonpolyhedral cones via approximation.

The product rule is Corollary A.6 in Appendix A.1.1. A proof of the Gauss–Bonnet relation (5.8) is beyond the scope of this thesis; a combinatorial proof of this fact available in McMullen’s classic work on angle-sum relations for polyhedral cones [McM75]. For a geometric account of the Gauss–Bonnet relation, see [SW08, Thm. 6.5.5]. \square

The Gauss–Bonnet formula (5.8) implies a number of important results that involve alternating sums of intrinsic volumes. The following definition anticipate these results.

Definition 5.5 (Tail functionals). For a cone $C \in \mathcal{C}_d$ and index $k \in \{0, \dots, d - 1, d\}$, we define the k th *tail functional*

$$t_k(C) := v_k(C) + v_{k+1}(C) + \dots = \sum_{j=k}^d v_j(C) \quad (5.9)$$

and the k th *half-tail functional*

$$h_k(C) := v_k(C) + v_{k+2}(C) + \dots = \sum_{\substack{j=k \\ j-k \text{ even}}}^d v_j(C). \quad (5.10)$$

5.2 Steiner formulas

The classical Steiner formula gives the volume of the region around a convex body in terms of the Euclidean intrinsic volumes of the body. For most cones, the volume of the cone is infinite. However, we find a parallel to the classical case when we consider the Gaussian, rather than Lebesgue, measure of the expansion of a convex cone. Our result is stated in terms of the distance to the polar cone, because this is the claim that is relevant for our later work.

¹The bijection between k faces of C and $(d - k)$ faces of C° follows from the fact that each face F of a cone C has corresponding dual face N_F of C° such that $\Pi_C(\mathbf{x}) \in F$ if and only if $\Pi_{C^\circ}(\mathbf{x}) \in N_F$; cf. Appendix A.2.1. It is a well-known fact (see, e.g., [Pat00, p. 34]) that $\text{lin}(F + N_F) = \mathbb{R}^d$ for each face F of a polyhedral cone. For nonpolyhedral cones, it may be that $\text{lin}(F + N_F) \subsetneq \mathbb{R}^d$, in which case the probabilistic definition (5.1) of the intrinsic volumes breaks down.

Proposition 5.6 (Gaussian Steiner formula). *For any closed convex cone C and any $\varepsilon \geq 0$,*

$$\mathbb{P}\{\text{dist}^2(\mathbf{g}, C^\circ) \leq \varepsilon\} = \mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 \leq \varepsilon\} = \sum_{k=0}^d \chi_k^2(\varepsilon) \nu_k(C). \quad (5.11)$$

The function χ_k^2 is the cumulative distribution function of a chi-square random variable with k degrees of freedom:

$$\chi_k^2(\varepsilon) := \mathbb{P}\left\{\sum_{i=1}^k g_i^2 \leq \varepsilon\right\}, \quad (5.12)$$

where $g_i \sim \text{NORMAL}(0, 1)$ are i.i.d. standard Gaussian random variables for $i = 1, \dots, k - 1, k$.

While we have not seen Proposition 5.6 stated as such in the literature, this result is likely known to experts. For completeness, we provide a proof in Appendix A.1 as a corollary of a more general Steiner formula for convex cones.

Remark 5.7. Taylor [Tay06, Cor. 3.4] provides much more sophisticated formula for Gaussian measure of the expansion of a set. The formula (5.11) above is not easily reconciled with the formula of Taylor. In particular, Taylor's Gaussian tube formula involves an infinite series of functionals that have no clear connection to the conic intrinsic volumes. Moreover, we achieve our formula via elementary means—which is to say, we do not require any advanced tools from differential geometry.

A closely related result is the following Steiner formula for spherically distributed random variables. This result is a special case of [SW08, Thm. 6.5.1].

Proposition 5.8 (Spherical Steiner formula). *Let $C \subset \mathbb{R}^d$ be a closed convex cone, and let $\boldsymbol{\theta}$ be uniformly distributed over the sphere S^{d-1} . For any $\varepsilon \in [0, 1]$,*

$$\mathbb{P}\{\|\Pi_C(\boldsymbol{\theta})\|^2 \leq \varepsilon\} = \sum_{k=0}^d B_k^d(\varepsilon) \nu_k(C).$$

The function $B_k^d(\varepsilon)$ is the cumulative distribution function of a Beta($k/2, (d - k)/2$) random variable:

$$B_k^d(\varepsilon) := \mathbb{P}\{\|\Pi_{L_k}(\boldsymbol{\theta})\|^2 \leq \varepsilon\}, \quad (5.13)$$

where $\theta \sim \text{UNIFORM}(S^{d-1})$ and L_k is a k -dimensional linear subspace of \mathbb{R}^d .

The proof of Proposition 5.8 also appears in Appendix A.1, where we derive both Proposition 5.8 and 5.6 from a more general result on mean values of sufficiently nice functions of the norms of projections $\|\Pi_C(\mathbf{g})\|$ and $\|\Pi_{C^c}(\mathbf{g})\|$ of a Gaussian vector onto a cone. Our proof technique involves a tiling argument coupled with approximation by polyhedral cones. This schema appears many times in the literature, including [SW08, pp. 251–252] and [San80, p. 242].

5.3 Kinematic formulas

Kinematic formulas are a crowning achievement of the theory of integral geometry. These results answer questions such as “what is the probability that a random geometric object strikes a fixed object?” Given the characterization (4.3) of demixing in terms of the intersection of convex cones, the answer to this question clearly has a bearing on our understanding of demixing problems. The most basic kinematic formula for convex cones is a formula for the mean intrinsic volume of the intersection of two convex cones.

Fact 5.9 (The conic kinematic formula). *Let $\mathbf{Q} \in O_d$ be a Haar-distributed orthogonal matrix. Then for any closed convex cones $C, D \in \mathcal{C}_d$ and all $k = 1, \dots, d - 1, d$,*

$$\mathbb{E}[v_k(C \cap \mathbf{Q}D)] = \sum_{j=k}^d v_j(C)v_{d-j+k}(D) = v_{d+k}(C \times D). \quad (5.14)$$

For $k = 0$, we have

$$\mathbb{E}[v_0(C \cap \mathbf{Q}D)] = \sum_{j=0}^d v_j(C \times D) = 1 - t_{d+1}(C \times D). \quad (5.15)$$

An antecedent of this result appears in the classic book of Santaló [San76, Sec. IV.18.3], where it is proved via differential-geometric methods. For a proof of the result above that aligns with the spirit of our development, see [SW08, pp. 258–261]. The second formula (5.15) follows readily from the first by the fact (5.4) that the intrinsic volumes always sum to one.

By combining the conic kinematic formula with the Gauss–Bonnet theorem (5.8) we obtain the *Crofton formula* for randomly oriented convex cones.

Theorem 5.10 (Conic Crofton formula). *Suppose that $C, D \in \mathcal{C}_d$ are two closed, convex cones, at least one of which is not a subspace. Then*

$$\mathbb{P}\{C \cap \mathbf{Q}D \neq \{\mathbf{0}\}\} = 2 \sum_{\substack{k=d+1 \\ k \text{ odd}}}^{2d} \nu_k(C \times D) = 2h_{d+1}(C \times D), \quad (5.16)$$

where the half-tail functional $h_{d+1}(C \times D)$ is defined in (5.10). In particular, when C is not a subspace and $D = L_{d-m}$ is a $(d - m)$ -dimensional subspace of \mathbb{R}^d ,

$$\mathbb{P}\{C \cap \mathbf{Q}L_{d-m} \neq \{\mathbf{0}\}\} = 2 \sum_{\substack{k=m+1 \\ k \text{ odd}}}^d \nu_k(C) = 2h_{m+1}(C). \quad (5.17)$$

We provide a proof of this result because there is no detailed reference for the probabilistic interpretation given above. The only technical point missing from the standard literature is a detailed proof that the intersection between the two cones is almost surely not a subspace, a fact that is contingent upon at least one of the cones not being a subspace. We state this point as a lemma to separate the present novelty from techniques in the prior literature.

Lemma 5.11 (The intersection of cones is rarely a subspace). *Let $C, D \in \mathcal{C}_d$ be closed convex cones, at least one of which is not a subspace. Then the intersection $C \cap \mathbf{Q}D$ is almost surely not a subspace of dimension greater than zero.*

The proof is deferred until Section 5.4.1, where we collect some fine points about unlikely configurations of cones. With this result in hand, we are in a position to prove the Crofton formula.

Proof of Theorem 5.10. The Gauss–Bonnet theorem (5.8) and the expression for the

intrinsic volumes of subspaces computed in Example 5.2 reveals

$$2 \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_k(C \cap \mathbf{Q}D) = \begin{cases} 0, & C \cap \mathbf{Q}D \text{ an even dim. subspace} \\ 2, & C \cap \mathbf{Q}D \text{ an odd dim. subspace} \\ 1, & \text{otherwise.} \end{cases} \quad (5.18)$$

Because the probability that $C \cap \mathbf{Q}D$ is an odd dimensional subspace is zero by Lemma 5.11, the sum on the left-hand side above is almost surely equal to an indicator function on the event $C \cap \mathbf{Q}D \neq \{0\}$. Upon averaging (5.18) over \mathbf{Q} , we arrive at

$$\mathbb{P}\{C \cap \mathbf{Q}D \neq \{0\}\} = \mathbb{E}[\mathbf{1}_{C \cap \mathbf{Q}D \neq \{0\}}(\mathbf{Q})] = 2 \mathbb{E} \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_k(C \cap \mathbf{Q}D) = 2 \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_{d+k}(C \times D).$$

where the final equality holds by linearity of expectation and the kinematic formula (5.14).

The right-hand side above is equal to $2h_{d+1}(C \times D)$ by the definition (5.10). This completes the demonstration of (5.16).

The specialization (5.17) of (5.16) follows from the expression for the intrinsic volumes of L_{d-m} given by Example 5.2. Indeed, the product rule (5.7) shows

$$v_k(C \times L_{d-m}) = \sum_{i+j=k} v_i(C) v_j(L_{d-m}) = v_{k+m-d}(C).$$

Applying this formula to (5.16), we find

$$\mathbb{P}\{C \cap \mathbf{Q}L_{d-m} \neq \{0\}\} = 2 \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_{k+m}(C) = 2h_{m+1}(C),$$

which completes the claim. □

5.3.1 Iterated kinematics

We now extend the kinematic formula to intersections of multiple randomly oriented cones.

Proposition 5.12 (Iterated kinematic formula). *Let $C_1, \dots, C_{n-1}, C_n \in \mathcal{C}_d$ and suppose*

$\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n \in \mathbb{R}^{n \times n}$ are i.i.d. random rotations. Then for all $k = 1, \dots, d-1, d$, we have

$$\mathbb{E}[v_k(\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 C_2 \cap \dots \cap \mathbf{Q}_n C_n)] = v_{(n-1)d+k}(C_1 \times C_2 \times \dots \times C_n). \quad (5.19)$$

For $k = 0$, we have

$$\mathbb{E}[v_0(\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 C_2 \cap \dots \cap \mathbf{Q}_n C_n)] = 1 - t_{(n-1)d+1}(C_1 \times C_2 \times \dots \times C_n). \quad (5.20)$$

Proof. Suppose first $k \geq 1$. We use induction on n . For the base case $n = 2$, we use the orthogonal invariance of the intrinsic volumes and conditional expectation to see

$$\mathbb{E}[v_k(\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 C_2)] = \mathbb{E}[v_k(\mathbf{Q}_1(C_1 \cap \mathbf{Q}_2 C_2))] = \mathbb{E}[v_k(C_1 \cap \mathbf{Q}_2)] = v_{d+k}(C_1 \times C_2)$$

where the first equality follows from the forgetting property (3.5) and the second by rotation invariance of the intrinsic volumes. The final equality is (5.14), which completes the base case.

Suppose now that $n > 2$ and the claim holds for $n-1$ cones. From the forgetting property (3.5) of the Haar measure and rotation invariance of intrinsic volumes, we find as above

$$\mathbb{E}[v_k(\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 C_2 \cap \mathbf{Q}_3 C_3 \cap \dots \cap \mathbf{Q}_n C_n)] = \mathbb{E}[v_k(C_1 \cap \mathbf{Q}_2(C_2 \cap \mathbf{Q}_3 C_3 \cap \dots \cap \mathbf{Q}_n C_n))]. \quad (5.21)$$

Conditioning on $\mathbf{Q}_3, \mathbf{Q}_4, \dots, \mathbf{Q}_n$, the conic kinematic formula (5.14) gives

$$\begin{aligned} \mathbb{E}[v_k(C_1 \cap \mathbf{Q}_2(C_2 \cap \mathbf{Q}_3 C_3 \cap \dots \cap \mathbf{Q}_n C_n))] &= \mathbb{E}[\mathbb{E}[v_k(C_1 \cap \mathbf{Q}_2(C_2 \cap \mathbf{Q}_3 C_3 \cap \dots \cap \mathbf{Q}_n C_n)) \mid \mathbf{Q}_3, \dots, \mathbf{Q}_n]] \\ &= \sum_{j=k}^d \mathbb{E}[v_j(C_2 \cap \mathbf{Q}_3 C_3 \cap \dots \cap \mathbf{Q}_n C_n)] v_{d-j+k}(C_1) \\ &= \sum_{j=k}^d v_{(n-2)d+j}(C_2 \times C_3 \times \dots \times C_n) v_{d-j+k}(C_1) \end{aligned}$$

The final relation follows from the forgetting property (3.5), rotation invariance of the

intrinsic volumes, and the induction hypothesis. We recognize the final sum above:

$$\sum_{j=k}^d v_{(n-2)d+j}(C_2 \times C_3 \times \cdots \times C_n) v_{d-j+k}(C_1) = v_{(n-1)d+k}(C_1 \times \cdots \times C_{n-1} \times C_n)$$

by the product rule (5.7) for intrinsic volumes. \square

A slight refinement of the proof of Theorem 5.10 provides an iterated Crofton formula.

Corollary 5.13 (Iterated Crofton formula). *Let $C_1, \dots, C_{n-1}, C_n \in \mathcal{C}_d$ be closed, convex cones, at least one of which is not a subspace. Suppose $\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n \in \mathcal{O}_d$ are independent random rotations. Then*

$$\mathbb{P}\{\mathbf{Q}_1 C_1 \cap \cdots \cap \mathbf{Q}_{n-1} C_{n-1} \cap \mathbf{Q}_n C_n \neq \{0\}\} = 2h_{(n-1)d+1}(C_1 \times \cdots \times C_{n-1} \times C_n). \quad (5.22)$$

Again, the heavy lifting required to prove the Crofton formula involves a demonstration that the probability that the intersection in braces in (5.22) is rarely a subspace.

Lemma 5.14. *Let $C_1, \dots, C_{n-1}, C_n \in \mathcal{C}_d$ and $\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n \in \mathbb{R}^{n \times n}$ be as in Corollary 5.13. Then the intersection $\mathbf{Q}_1 C_1 \cap \cdots \cap \mathbf{Q}_{n-1} C_{n-1} \cap \mathbf{Q}_n C_n$ is almost surely not a subspace of positive dimension.*

The proof reduces the intersection of multiple cones to the intersection of two cones and applies Lemma 5.11. The details appear in Section 5.4.1 below.

Proof of Corollary 5.13. As in the proof of Theorem 5.10, we combine the Gauss–Bonnet relation and Lemma 5.14 to find

$$\begin{aligned} \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \cdots \cap \mathbf{Q}_{n-1} C_{n-1} \cap \mathbf{Q}_n C_n \neq \{0\}\} &= 2 \sum_{\substack{k=1 \\ k \text{ odd}}}^d \mathbb{E}[v_k(\mathbf{Q}_1 C_1 \cap \cdots \cap \mathbf{Q}_{n-1} C_{n-1} \cap \mathbf{Q}_n C_n)] \\ &= 2 \sum_{\substack{k=1 \\ k \text{ odd}}}^d v_{(n-1)d+k}(C_1 \times \cdots \times C_{n-1} \times C_n), \end{aligned}$$

where the final equality follows from the iterated kinematic formula (5.19). We recognize the latter sum as the half-tail functional $h_{(n-1)d+1}(C_1 \times \cdots \times C_{n-1} \times C_n)$, which completes the proof. \square

5.4 Rare events and other fine points

We now fill in a much needed gap in the literature related to probability zero events. In Section 5.4.1, we provide the proof of Lemmas 5.11 and 5.14 that show that the intersection of cones is not likely to be a subspace. Afterward, Section 5.4.2 discusses touching probabilities and some implications for extending the kinematic theory of Section 5.3 to nonclosed cones that frequently appear in demixing applications.

5.4.1 The intersections of cones are almost never subspaces

This section provides the proof of Lemmas 5.11 and 5.14, which assert that the intersection of randomly rotated cones are almost never subspaces, so long as at least one of the cones involved is not a subspace. We start with a definition.

Definition 5.15. Two linear subspaces $L, L' \subset \mathbb{R}^d$ are said to be in *special position* when

$$\dim(L \cap L') > 0 \quad \text{and} \quad \text{lin}(L \cup L') \neq \mathbb{R}^d. \quad (5.23)$$

A result on null sets from the theory of the Haar measure on O_d asserts that subspaces in special position are highly unusual [SW08, Lem. 13.2.1].

Fact 5.16 (Special position is special). *Let $L, L' \subset \mathbb{R}^d$ be two linear subspaces and $Q \in O_d$ a random rotation. Then*

$$\mathbb{P}\{L \text{ and } QL' \text{ are in special position}\} = 0.$$

With this fact, we are able to show that cones are highly unlikely to intersect only at a subspace.

Proof of Lemma 5.11. Suppose that $C \cap \mathbf{Q}D = L$ for some linear subspace L with dimension $k > 0$. From (3.7), we have the unique orthogonal decompositions $C = L_C + C_*$ and $D = L_D + D_*$, where L_C and L_D are subspaces and C_*, D_* are pointed cones perpendicular to these subspaces. We will show that

$$\dim(L_C \cap \mathbf{Q}L_D) > 0 \quad \text{and} \quad \text{lin}(L_C \cup \mathbf{Q}L_D) \neq \mathbb{R}^d, \quad (5.24)$$

that is, L_C and L_D are in *special position*.

Since L is a linear subspace, for any $\mathbf{x} \in L$ we also have $-\mathbf{x} \in L$, so that $\pm\mathbf{x} \in C$ and $\pm\mathbf{x} \in \mathbf{Q}D$ by definition of L . Because the pointed components C_* and D_* are perpendicular to the linear components L_C and L_D , we must have $\mathbf{x} \in L_C \cap \mathbf{Q}L_D$. In other words, $L \subset L_C \cap \mathbf{Q}L_D$, so in particular

$$\dim(L_C \cap \mathbf{Q}L_D) \geq \dim(L) = k > 0.$$

Thus L_C and $\mathbf{Q}L_D$ satisfy the first condition in (5.24).

To show that $L_C \cup L_D \neq \mathbb{R}^d$, we appeal to the separating hyperplane theorem for convex cones (Fact 3.3). In the present situation, the intersection $C \cap D = L \neq \{\mathbf{0}\}$, so we cannot apply Fact 3.3 directly. Instead, we will apply it to the quotient cones $\Pi_{L^\perp}(C)$ and $\Pi_{L^\perp}(\mathbf{Q}D)$, which we claim only intersect at $\mathbf{0}$. To see this claim, note that if $\Pi_{L^\perp}(\mathbf{x}) = \Pi_{L^\perp}(\mathbf{x}')$ for $\mathbf{x} \in C$ and $\mathbf{x}' \in D$, then the difference $\mathbf{x}' - \mathbf{x} \in L$ by linearity of the projection onto a subspace. Since $L \subset C$ and C is a convex cone, the vector $\mathbf{x}' = \mathbf{x} + (\mathbf{x}' - \mathbf{x}) \in C$. In other words, $\mathbf{x}' \in C \cap D = L$, and so $\Pi_{L^\perp}(\mathbf{x}') = \mathbf{0}$, proving the claim.

The subspace L^\perp is isomorphic to \mathbb{R}^{d-k} and $\Pi_{L^\perp}(C) \cap \Pi_{L^\perp}(D) = \{\mathbf{0}\}$, so that Fact 3.3 provides a $\mathbf{z} \neq \mathbf{0}$ such that

$$\mathbf{z} \in L^\perp, \quad \mathbf{z} \in \Pi_{L^\perp}(C)^\circ, \quad \text{and} \quad -\mathbf{z} \in \Pi_{L^\perp}(\mathbf{Q}D)^\circ,$$

where the polars above are with respect to L^\perp . The first two relations imply that for any

$\mathbf{x} \in L_C$

$$\langle \mathbf{z}, \mathbf{x} \rangle = \langle \mathbf{z}, \Pi_{L^\perp}(\mathbf{x}) \rangle \leq 0.$$

Since L_C is a linear subspace, the same argument applied to $-\mathbf{x}$ gives the reverse inequality, and hence $\langle \mathbf{z}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in L_C$, which is to say $\mathbf{z} \in L_C^\perp$. Similarly, we have $\mathbf{z} \in (\mathbf{Q}L_D)^\perp$, and so we find that there is a nonzero vector $\mathbf{z} \in L_C^\perp \cap (-\mathbf{Q}L_D)^\perp$. By taking the polar of this inclusion, we find

$$\mathbb{R}^d \neq L_C + (-\mathbf{Q}L_D) = \text{lin}(L_C \cup \mathbf{Q}L_D).$$

by the polarity relationship (3.9). The final equality above is a standard fact from linear algebra. We conclude that L_C and $\mathbf{Q}L_D$ are in special position. As the set of rotations that brings L_C and L_D into special position has measure zero (Fact 5.16), we conclude that there is zero probability that $C \cap \mathbf{Q}D$ is a subspace of dimension greater than zero. \square

The multiple-cone extension of this result uses induction with the work above as a base case.

Proof of Lemma 5.14. Without loss of generality, we reorder the cones so that C_1 not a subspace. Define the random cone $D := C_2 \cap \mathbf{Q}_3 C_3 \cap \cdots \cap \mathbf{Q}_n C_n$. From the forgetting property (3.5), we have

$$\begin{aligned} \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 C_2 \cap \cdots \cap \mathbf{Q}_n C_n \text{ is not a subspace}\} &= \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 D \text{ is not a subspace}\} \\ &= \mathbb{E}\left[\mathbb{P}_{\mathbf{Q}_1}\{\mathbf{Q}_1 C_1 \cap \mathbf{Q}_2 D \text{ is not a subspace}\} \mid \mathbf{Q}_2, \dots, \mathbf{Q}_n\right] \end{aligned}$$

where the second equality follows from the law of total probability. Since C_1 is itself not a subspace, Lemma 5.11 implies that the inner probability above is equal to zero for all $\mathbf{Q}_2, \dots, \mathbf{Q}_n$. The conclusion follows immediately. \square

5.4.2 Removing closure restrictions

In our demixing application, the cones we encounter are often not closed, while the theory of conic integral geometry typically includes closure assumptions. In this section,

we show that closure assumptions do not pose an obstacle to the kinematic theory of convex cones.

Relaxing the closure assumptions requires a surprisingly deep fact about touching probabilities. Two closed convex cones *touch* if their intersection contains a ray but they may still be weakly separated by a hyperplane. In the course of a proof of the kinematic formula, Schneider & Weil prove the following result [SW08, pp. 258–260].

Fact 5.17 (Touching is not allowed). *Let $C, D \in \mathcal{C}_d$ be closed, convex cones such that both $C, D \neq \{0\}$. Then*

$$\mathbb{P}\{C \text{ touches } QD\} = 0.$$

The following result anticipates the needs of Section 9. The intersection of a cone and the *negative* image of a randomly oriented cone is related to the multiple demixing optimality condition (4.6). Moreover, all descent cones contain zero (Proposition 3.6.2), so we restrict ourselves to cones that contain zero for notational simplicity.

Proposition 5.18. *Let $C, D \subset \mathbb{R}^d$ be convex cones that contain zero but are not necessarily closed. Let $Q \in O_d$ be independent random rotations. If both $C, D \neq \{0\}$, then the sets*

$$\{Q : C \cap -QD = \{0\}\} \quad \text{and} \quad \{Q : C^\circ \cap Q^t D^\circ \neq \{0\}\}$$

coincide up to a set of Haar measure zero on $Q \in O_d$. In particular,

$$\mathbb{P}\{C \cap -QD = \{0\}\} = \mathbb{P}\{C^\circ \cap Q^t D^\circ \neq \{0\}\}.$$

Proof. Fix $U := Q$ for the moment and suppose $C \cap -UD = \{0\}$. By the separating hyperplane theorem for convex cones (Fact 3.3), there exists a nonzero vector $w \in \mathbb{R}^d$ such that

$$\langle w, x \rangle \leq 0 \text{ for all } x \in C \quad \text{and} \quad \langle w, -Uy \rangle \geq 0 \text{ for all } y \in C, \quad (5.25)$$

which is equivalent to the statement $w \in C^\circ \cap U^t D^\circ$. Therefore, we have the inclusion

$$\{\mathbf{Q} : C \cap -\mathbf{Q}D = \{\mathbf{0}\}\} \subset \{\mathbf{Q} : C^\circ \cap \mathbf{Q}^t D^\circ \neq \{\mathbf{0}\}\}.$$

For the other direction, suppose that $C^\circ \cap U^t D^\circ \neq \{\mathbf{0}\}$ for some $U \in O_d$, so that there exists a nonzero w satisfying (5.25). In this case, we either have $C \cap -UD = \{\mathbf{0}\}$, or C and $-UD$ touch, which implies

$$\{\mathbf{Q} : C^\circ \cap U^t D^\circ \neq \{\mathbf{0}\}\} \subset \{\mathbf{Q} : C \cap -\mathbf{Q}D = \{\mathbf{0}\}\} \cup \{\mathbf{Q} : C \text{ touches } \mathbf{Q}^t D\}.$$

This second set in the union has Haar measure zero by Fact 5.17. The conclusion follows by comparing the final two displayed equations above. \square

Chapter 6

The statistical dimension

In light of the Steiner and kinematic formulas of the previous section, the intrinsic volumes carry a significant amount of information about a convex cone. Unfortunately, these quantities can be vexingly difficult to calculate in situations relevant to applications. In this section, we discuss two closely related summary parameters that measure the size of convex cones, explore their interpretation as dimensionality parameters, and discuss their relationship to one another. The results in this section consists of largely of joint work that appears in [ALMT13, Sec. 4].

6.1 Basic properties

The (linear) dimension of a convex set is defined as the dimension of its linear hull. Dimension is certainly a measure of the size of a cone because it is monotonic under inclusion, and it is intrinsic in the sense that the dimension of a set is independent of the embedding dimension. Yet dimension is brittle under small changes to a set: a full dimensional cone can be very close to a ray in the Hausdorff metric, for instance. Moreover, the standard linear dimension provides no distinction between, say, the positive orthant and a halfspace.

In this section, we introduce a parameter associated with cones called the statistical dimension. We will find that this quantity shares a number of remarkable properties with the linear dimension of a subspace, but it possesses a key property that the linear dimension does not: continuity.

Definition 6.1. The *statistical dimension* $\delta(C)$ of a closed, convex cone $C \subset \mathbb{R}^d$ is given by

$$\delta(C) := \mathbb{E} [\|\Pi_C(\mathbf{g})\|^2]. \quad (6.1)$$

The next proposition highlights the many analogies between the statistical dimension and the usual linear dimension of a subspace. The only property unique to the statistical dimension is continuity.

Proposition 6.2 (Properties of the statistical dimension). *For any closed convex cones $C, D \subset \mathbb{R}^d$ and $C' \subset \mathbb{R}^n$, we have the following properties.*

1. **Continuity.** *The statistical dimension δ is continuous in the conic Hausdorff metric.*
2. **Monotonicity.** *The containment $C \subset D$ implies the inequality $\delta(C) \leq \delta(D)$.*
3. **Rotational invariance.** *For any orthogonal matrix $U \in O_d$, the statistical dimension satisfies $\delta(UC) = \delta(C)$.*
4. **Polarity.** *The polar cone satisfies $\delta(C^\circ) = d - \delta(C)$.*
5. **Direct sum.** *Direct sums split as $\delta(C \times C') = \delta(C) + \delta(C')$.*
6. **Embedding.** *The statistical dimension is independent of the embedding dimension of the cone.*

Proof. (Continuity.) Suppose a sequence $(C_i)_{i \in \mathbb{N}}$ of cones in \mathcal{C}_d has a limit $C_i \rightarrow C$ as $i \rightarrow \infty$. By Proposition 3.8 and continuity of the Euclidean norm, $\|\Pi_{C_i}(\mathbf{x})\|^2 \rightarrow \|\Pi_C(\mathbf{x})\|^2$ for every $\mathbf{x} \in \mathbb{R}^d$. The bound $\|\Pi_{C_i}(\mathbf{x})\|^2 \leq \|\mathbf{x}\|^2$ holds for all i by the Pythagorean formula (3.16), and moreover the average value $\mathbb{E}[\|\mathbf{g}\|^2] = d < \infty$. Therefore, we may apply the dominated convergence theorem [Fol99, (2.24)]:

$$\mathbb{E} [\|\Pi_{C_i}(\mathbf{g})\|^2] \rightarrow \mathbb{E} [\|\Pi_C(\mathbf{g})\|^2].$$

This is the desired conclusion.

(Monotonicity.) Suppose $C \subset C'$. Because inclusion reverses under polarity, we see

$$\|\Pi_C(\mathbf{x})\|^2 = \text{dist}^2(\mathbf{x}, C^\circ) \leq \text{dist}^2(\mathbf{x}, D^\circ) = \|\Pi_D(\mathbf{x})\|^2.$$

(*Rotational invariance.*) A direct calculation shows that $\Pi_{U^t C}(\mathbf{x}) = U \Pi_C(U^t \mathbf{x})$ for every orthogonal matrix U , rotational invariance of the statistical dimension follows from the rotational invariance of standard Gaussian vectors.

(*Polarity.*) The polarity relationship is just the Pythagorean identity (3.16) averaged over all Gaussian vectors:

$$\delta(C) + \delta(C^\circ) = \mathbb{E}[\|\Pi_C(\mathbf{g})\|^2 + \|\Pi_{C^\circ}(\mathbf{g})\|^2] = \mathbb{E}[\|\mathbf{g}\|^2] = d.$$

(*Direct sum.*) Let $\mathbf{g} = (\mathbf{g}_d; \mathbf{g}_n) \in \mathbb{R}^d \times \mathbb{R}^n$ be a standard Gaussian vector. By (3.18), the projection over the product of cones $C \times C'$ splits as $\Pi_{C \times C'}(\mathbf{g}) = (\Pi_C(\mathbf{g}_d); \Pi_{C'}(\mathbf{g}_n))$. Therefore,

$$\begin{aligned} \delta(C \times C') &= \mathbb{E}[\|(\Pi_C(\mathbf{g}_d), \Pi_{C'}(\mathbf{g}_n))\|^2] \\ &= \mathbb{E}[\|\Pi_C(\mathbf{g}_d)\|^2] + \mathbb{E}[\|\Pi_{C'}(\mathbf{g}_n)\|^2] = \delta(C) + \delta(C'), \end{aligned}$$

as claimed.

(*Embedding.*) Suppose a closed, convex cone $C \subset \mathbb{R}^d$ is embedded as a cone $D \subset \mathbb{R}^n$ in some higher dimensional space. Then there is an orthogonal matrix $U \in O_n$ such that

$$D = U(C \times \underbrace{\{\mathbf{0}\} \times \cdots \times \{\mathbf{0}\}}_{n-d \text{ times}}).$$

By points 3 and 5 above,

$$\delta(D) = \delta(C \times \{\mathbf{0}\} \times \cdots \times \{\mathbf{0}\}) = \delta(C) + (n-d)\delta(\{\mathbf{0}\}) = \delta(C),$$

which completes the claim. □

The statistical dimension has a number of equivalent characterizations.

Proposition 6.3 (Chameleon). *For any closed convex cone $C \subset \mathbb{R}^d$, we have the following alternative characterizations of the statistical dimension:*

1. **Gaussian formulation.** *The statistical dimension is defined as*

$$\delta(C) := \mathbb{E} [\|\Pi_C(\mathbf{g})\|^2], \quad \text{where } \mathbf{g} \sim \text{NORMAL}(\mathbf{0}, \mathbf{I}_{d \times d}). \quad (6.2)$$

2. **Spherical formulation.** *An equivalent definition involving spherical random variables is*

$$\delta(C) = d \mathbb{E} [\|\Pi_C(\boldsymbol{\theta})\|^2], \quad \text{where } \boldsymbol{\theta} \sim \text{UNIFORM}(\mathbb{S}^{d-1}). \quad (6.3)$$

3. **Distance formulation.** *Alternatively, we may compute the mean distance*

$$\delta(C) = \mathbb{E} [\text{dist}^2(\mathbf{g}, C^\circ)]. \quad (6.4)$$

4. **Supremum formulation.** *Another characterization involves a supremum:*

$$\delta(C) = \mathbb{E} \left[\sup_{\mathbf{x} \in C \cap \mathbb{B}_d} |\langle \mathbf{x}, \mathbf{g} \rangle|^2 \right]. \quad (6.5)$$

5. **Intrinsic volume formulation.** *Finally, the statistical dimension is the mean of the distribution of intrinsic volumes:*

$$\delta(C) = \sum_{k=0}^d k v_k(C). \quad (6.6)$$

Proof. (Gaussian formulation) This repeats the definition (6.1).

(Spherical formulation) A Gaussian vector $\mathbf{g} \in \mathbb{R}^d$ may be written as the product $\mathbf{g} = r\boldsymbol{\theta}$ of a chi-square random variable r with d degrees of freedom and an independent vector $\boldsymbol{\theta}$ uniformly distributed over \mathbb{S}^{d-1} . By homogeneity of the projection and independence of r and $\boldsymbol{\theta}$, we have

$$\delta(C) = \mathbb{E} [\|\Pi_C(r\boldsymbol{\theta})\|^2] = \mathbb{E}[r^2] \mathbb{E} [\|\Pi_C(\boldsymbol{\theta})\|^2] = d \mathbb{E} [\|\Pi_C(\boldsymbol{\theta})\|^2].$$

This is equation (6.3).

(*Distance formulation*) Equation (6.4) equality follows immediately from (3.17) and the Gaussian formulation (6.2) of the statistical dimension.

(*Supremum formulation*) For each $\mathbf{x} \in \mathbb{R}^d$, the definition (3.13) of the distance to a convex cone gives

$$\text{dist}^2(\mathbf{x}, C^\circ) = \inf_{\mathbf{y} \in C^\circ} \|\mathbf{y} - \mathbf{x}\|^2 = \left(\inf_{\mathbf{y} \in C^\circ} \|\mathbf{y} - \mathbf{x}\| \right)^2, \quad (6.7)$$

where the second equation follows by homogeneity. We manipulate the right-hand infimum using tools from convex duality theory. The Fenchel conjugate of the Euclidean norm is the convex indicator i_{B_d} on the Euclidean ball [Roc70, Thm. 13.2]. Therefore, the conjugate of the function inside the infimum above is given by

$$\sup_{\mathbf{y} \in \mathbb{R}^d} \langle \mathbf{y}, \mathbf{z} \rangle - \|\mathbf{y} - \mathbf{x}\| = i_{B_d}(\mathbf{z}) + \langle \mathbf{x}, \mathbf{z} \rangle.$$

Then the result [Roc70, Thm. 31.4], a consequence of Fenchel's duality theorem, immediately implies the equality

$$\inf_{\mathbf{y} \in C^\circ} \|\mathbf{y} - \mathbf{x}\| = \sup_{\mathbf{z} \in C} \langle \mathbf{x}, \mathbf{z} \rangle - i_{\|\mathbf{z}\| \leq 1}(\mathbf{z}) = \sup_{\mathbf{z} \in C \cap B_d} \langle \mathbf{x}, \mathbf{z} \rangle = \sup_{\mathbf{z} \in C \cap B_d} |\langle \mathbf{x}, \mathbf{z} \rangle|. \quad (6.8)$$

The second equality above follows because the maximal inner product is at least zero because $\mathbf{0} \in C \cap B_d$. Combining the display above with (6.7), we see that

$$\text{dist}^2(\mathbf{x}, C^\circ) = \sup_{\mathbf{z} \in C \cap B_d} |\langle \mathbf{x}, \mathbf{z} \rangle|^2.$$

The claim (6.5) then follows immediately from the distance characterization (6.4) of the statistical dimension.

(*Intrinsic volume formulation.*) By the integration-by-parts formula for the expectation,

$$\delta(C) = \mathbb{E}[\|\Pi_C(\mathbf{g})\|^2] = \int_0^\infty \mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 > \varepsilon\} d\varepsilon.$$

By the Gaussian Steiner formula of Proposition 5.6, the probability inside the integral is

given by the sum

$$\delta(C) = \sum_{k=0}^d v_k(C) \int_0^\infty \mathbb{P}\{\|\Pi_{L_k}(\mathbf{g})\|^2 \geq \varepsilon\} d\varepsilon = \sum_{k=0}^d v_k(C) \mathbb{E}[\|\Pi_{L_k}(\mathbf{g})\|^2],$$

where $L_k \subset \mathbb{R}^d$ is a subspace of dimension k . The equality (6.6) follows because the mean value $\mathbb{E}[\|\Pi_{L_k}(\mathbf{g})\|^2] = k$. \square

Corollary 6.4. *For any cone $C \in \mathcal{C}_d \setminus \{\{\mathbf{0}\}, \mathbb{R}^d\}$, we have*

$$\frac{1}{2} \leq \delta(C) \leq d - \frac{1}{2}.$$

Proof. Assume first C is not a subspace. Because the intrinsic volumes are nonnegative, we have

$$\delta(C) = \sum_{k=0}^d k \cdot v_k(C) \geq \sum_{\substack{k=1 \\ k \text{ even}}}^d v_k(C) = \frac{1}{2},$$

where the final equality follows by the Gauss–Bonnet relationship (5.8). The upper inequality follows when we apply the first relationship to the polar cone:

$$\delta(C) = d - \delta(C^\circ) \leq d - \frac{1}{2},$$

where the first equality above is the polarity relationship, Proposition 6.2.4. When $C = L$ is a subspace other than $\{\mathbf{0}\}$ or \mathbb{R}^d , the inequality follows from the fact that $\delta(L) = \dim(L)$ by (6.6) and the computation (5.2) of the intrinsic volumes of a subspace. \square

Remark 6.5. The intrinsic volumes, and hence the statistical dimension, are *valuations* on the set of closed convex cones in \mathbb{R}^d . That is, they satisfy the additivity properties

$$v_k(C \cup D) = v_k(C) + v_k(D) - v_k(C \cap D) \quad \text{and} \quad v_k(\{\mathbf{0}\}) = 0$$

whenever C, D and $C \cup D$ are closed, convex cones. The long-standing *spherical Hadwiger conjecture* posits that all continuous, rotationally invariant valuations on the set of closed convex cones are linear combinations of intrinsic volumes. If this conjecture holds, then

Table 6.1: Statistical dimension computations. A summary of the statistical dimension calculations in Section 6.2.

Cone	Notation	Statistical dimension	Location
Subspace	$L \subset \mathbb{R}^d$	$\dim(L)$	Sec. 6.2.1
The nonnegative orthant	\mathbb{R}_+^d	$\frac{1}{2}d$	Sec. 6.2.2
The second-order cone	\mathbb{L}^{d+1}	$\frac{1}{2}(d+1)$	Sec. 6.2.2
Symmetric positive-semidefinite matrices	$\mathbb{S}_+^{n \times n}$	$\frac{1}{4}n(n+1)$	Sec. 6.2.2

all such valuations are characterized by the value they take on linear subspaces. The spherical Hadwiger conjecture thus implies yet another characterization of the statistical dimension that expands on the analogy with the linear dimension:

If the spherical Hadwiger conjecture holds, then the statistical dimension is the unique continuous, rotationally invariant valuation on the closed convex cones such that $\delta(L) = \dim(L)$ for all linear subspaces $L \subset \mathbb{R}^d$.

The spherical Hadwiger conjecture is discussed in a number of works, including [McM93, p. 976], [KR97, Sec. 11.5], and [SW08, p. 263]. This conjecture is currently open for $d \geq 4$.

6.2 Explicit computations

We now turn to the important question of computing the statistical dimension in some special cases. In some highly symmetric cases, we obtain exact expressions for the statistical dimension, while in others we turn to an approximation technique pioneered by Stojnic [Sto09] and refined by Chandrasekaran et al. [CRPW12]. We summarize the results of this section in Table 6.1.

6.2.1 Subspaces

We first verify that the statistical dimension agrees with the usual affine dimension on a linear subspace.

Proposition 6.6 (Statistical dimension of subspaces). *The statistical dimension of a subspace $L \subset \mathbb{R}^d$ is*

$$\delta(L) = \dim(L). \quad (6.9)$$

Proof. For a k -dimensional subspace $L_k \subset \mathbb{R}^d$, the projection $\mathbf{g}_k := \mathbf{\Pi}_{L_k}(\mathbf{g})$ is a k -dimensional standard Gaussian random variable by the marginal property of Gaussian vectors. Therefore,

$$\delta(L_k) = \mathbb{E}[\|\mathbf{g}_k\|^2] = k. \quad (6.10)$$

We see that the statistical dimension of a subspace indeed agrees with its linear dimension. \square

See Remark 6.5 for an alternative approach to this calculation based on deeper considerations.

6.2.2 Self-dual cones

A cone C is self-dual if $C^\circ = -C$. The theory of self-dual cones underlies three common classes of convex optimization problems.

- **The nonnegative orthant.** The nonnegative orthant $\mathbb{R}_+^d := \{\mathbf{x} \in \mathbb{R}^d : x_i \geq 0 \text{ for } i = 1, \dots, d\}$ is a self-dual cone intimately connected to the theory of linear programming.
- **The Lorentz cone.** The Lorentz cone $\mathbb{L}^{d+1} := \{(\mathbf{x}, t) \in \mathbb{R}^{d+1} : \|\mathbf{x}\| \leq t\}$ is self dual. It is also known as the *second-order cone* or the *ice-cream cone*, and it is the fundamental object in second-order cone programs [LVBL98].
- **The positive-semidefinite cone.** The cone $\mathbb{S}^{n \times n} := \{\mathbf{X} \in \mathbb{R}_{\text{sym}}^{n \times n} : \mathbf{X} \succeq \mathbf{0}\}$ of symmetric positive-semidefinite matrices is a self-dual cone in the space of symmetric matrices. The notation $\mathbf{A} \succeq \mathbf{B}$ indicates that the difference $\mathbf{A} - \mathbf{B}$ is positive semidefinite, and $\mathbb{R}_{\text{sym}}^{n \times n}$ denotes the $\frac{1}{2}n(n-1)$ -dimensional space of symmetric $n \times n$ matrices. The positive-semidefinite cone appears in semidefinite programs written in standard form [VB96].

The statistical dimension of a self-dual cone is particularly simple to calculate; cf. [CRPW12, Cor. 3.8].

Proposition 6.7 (Self-dual cones). *Suppose $C \in \mathcal{C}_d$ is self-dual. Then the statistical dimension of C is given by*

$$\delta(C) = \frac{d}{2}.$$

Proof. By the polarity law and unitary invariance laws of Proposition 6.2, we have $d = \delta(C) + \delta(C^\circ) = 2\delta(C)$. \square

The entries for the nonnegative orthant, Lorentz cone, and the positive-semidefinite cone in Table 6.1 follow immediately from this result.

6.2.3 Circular cones

The *circular cone* $\text{Circ}_d(\alpha)$ in \mathbb{R}^d with angle $0 \leq \alpha \leq \frac{\pi}{2}$ is given by

$$\text{Circ}_d(\alpha) := \{\mathbf{x} \in \mathbb{R}^d : x_1 \geq \|\mathbf{x}\| \cos(\alpha)\}.$$

In particular, the circular cone $\text{Circ}_d(\frac{\pi}{4})$ is isometric to the Lorentz-cone \mathbb{L}^d . We have a rather precise approximation for the statistical dimension of a circular cone.

Proposition 6.8. *The statistical dimension of a circular cone satisfies*

$$\delta(\text{Circ}_d(\alpha)) = d \sin^2(\alpha) + R_d(\alpha), \tag{6.11}$$

where the remainder term $|R_d(\alpha)| \leq \frac{\pi^2}{2} < 5$.

We remark that the remainder can be improved slightly via a combinatorial argument due to D. Amelunxen (personal communication, 2012). See the remarks [ALMT13, pp. 36–37] for a sketch of the argument.

To perform our computation, we need to integrate in spherical coordinates. The following result, from [SW08, Lem. 6.5.1], gives a convenient way of parameterizing integration over the unit sphere S^{d-1} by integration over orthogonal subspheres. See Figure 6.1 for an illustration.

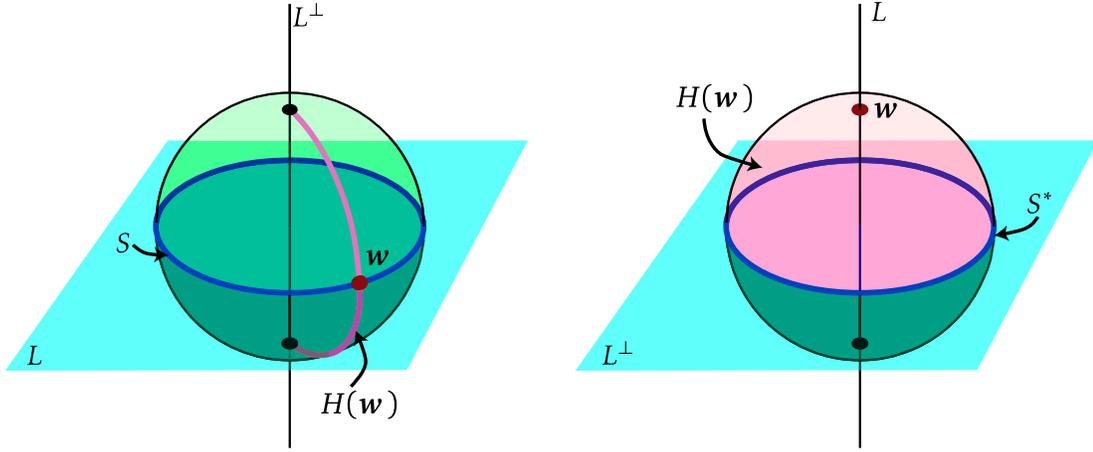


Figure 6.1: Orthogonal decomposition of the sphere into subspheres. [Left] The subspace L (light blue) intersects the sphere S^2 (green) at the equator S (blue). The perpendicular subsphere S^* (black points) is the intersection of the perpendicular subspace L^\perp (black line) with the sphere. The lune $H(\mathbf{w})$ (pink) connects the points of S^* through a point $\mathbf{w} \in S$. This decomposition is useful for integration over the sphere (Fact 6.1). [Right] With the roles of L and L^\perp reversed, the lune $H(\mathbf{w})$ becomes a hemisphere (pink) that connects S^* (blue) to a point \mathbf{w} (red) in S . (Fact 6.9).

Fact 6.9. Let $f : S^{d-1} \rightarrow \mathbb{R}$ be a nonnegative measurable function, and let L be a k -dimensional subspace of \mathbb{R}^d . Define $S := S^{d-1} \cap L$ and $S^* = S^{d-1} \cap L^\perp$ perpendicular subspheres of S^{d-1} . Then

$$\int_{S^{d-1}} f(\boldsymbol{\theta}) d\sigma_{d-1}(\boldsymbol{\theta}) = \int_S \int_{H(\mathbf{w})} \sin^{k-1}(\text{dist}_s(S^*, \mathbf{u})) f(\mathbf{u}) d\sigma_{d-k}(\mathbf{u}) d\sigma_{k-1}(\mathbf{w}). \quad (6.12)$$

where $H(\mathbf{w}) := \text{cone}(L^\perp + \{\mathbf{w}\}) \cap S^{d-1}$.

Recall that the measure σ_{d-1} on S^{d-1} is not normalized; cf. (3.3).

Proof of Proposition 6.8. We begin with an exact integral expression for the statistical dimension of the circular cone $C = \text{Circ}_d(\alpha)$. By the spherical formulation (6.3) of the statistical dimension, we express the statistical dimension as the average $\delta(C) = d \mathbb{E}[\|\Pi_C(\boldsymbol{\theta})\|^2]$, where $\boldsymbol{\theta}$ is uniformly random on S^{d-1} . Define the angle $\varphi := \varphi(\boldsymbol{\theta}) := \arccos(\theta_1)$ between $\boldsymbol{\theta}$ and the standard basis vector $\mathbf{e}_1 := (1; 0; \dots; 0)$. Basic trigonome-

try shows that the projected norm of $\boldsymbol{\theta}$ is given by

$$F(\varphi) := \|\boldsymbol{\Pi}_C(\boldsymbol{\theta})\|^2 = \begin{cases} 1, & 0 \leq \varphi < \alpha, \\ \cos^2(\varphi - \alpha), & \alpha \leq \varphi < \frac{\pi}{2} + \alpha, \\ 0, & \frac{\pi}{2} + \alpha \leq \varphi \leq \pi. \end{cases}$$

To obtain the exact statistical dimension $\delta(C)$ from (6.3), we integrate $F(\varphi)$ in polar coordinates:

$$\delta(C) = d \mathbb{E}[\|\boldsymbol{\Pi}_C(\boldsymbol{\theta})\|^2] = \frac{d}{\sigma_{d-1}(\mathbb{S}^{d-1})} \int_{\mathbb{S}^{d-1}} F(\varphi(\boldsymbol{\theta})) d\sigma_{d-1}(\boldsymbol{\theta})$$

where $\sigma_{d-1}(\mathbb{S}^{d-1})$ is the surface area of \mathbb{S}^{d-1} (3.3). We now apply Fact 6.9 with $L = \{\mathbf{x} \in \mathbb{R}^d : x_1 = 0\}$, the subspace perpendicular to \mathbf{e}_1 , to find

$$\begin{aligned} & \int_{\mathbb{S}^{d-1}} F(\varphi(\boldsymbol{\theta})) d\sigma_{d-1}(\boldsymbol{\theta}) \\ &= \int_{L \cap \mathbb{S}^{d-1}} \int_{H(\mathbf{w})} \sin^{d-2}(\text{dist}_s(\{\pm \mathbf{e}_1\}, \mathbf{u})) F(\varphi(\mathbf{u})) d\sigma_1(\mathbf{u}) d\sigma_{d-2}(\mathbf{w}), \end{aligned}$$

where $H(\mathbf{w}) := \text{cone}(\{\pm \mathbf{e}_1, \mathbf{w}\})$ is the arc from \mathbf{e}_1 to $-\mathbf{e}_1$ passing through $\mathbf{w} \in S$. (See Figure 6.1, left panel.) Since \mathbf{w} is orthogonal to \mathbf{e}_1 , we may express $\mathbf{u} \in H(\mathbf{w})$ by $\mathbf{u} = \sin(\varphi)\mathbf{w} + \cos(\varphi)\mathbf{e}_1$. Parameterizing $H(\mathbf{w})$ by φ , we find

$$\begin{aligned} \int_{\mathbb{S}^{d-1}} F(\varphi(\boldsymbol{\theta})) d\sigma_{d-1}(\boldsymbol{\theta}) &= \int_{\mathbb{S}^{d-2}} \int_0^\pi \sin^{d-2}(\varphi) F(\varphi) d\varphi d\sigma_{d-2}(\mathbf{w}) \\ &= \sigma_{d-2}(\mathbb{S}^{d-2}) \int_0^\pi \sin^{d-2}(\varphi) F(\varphi) d\varphi. \end{aligned}$$

The first equality also relies on the fact that $\text{dist}_s(\{\pm \mathbf{e}_1\}, \sin(\varphi)\mathbf{w} + \cos(\varphi)\mathbf{e}_1) = \varphi$, while the second follows from Fubini. Putting this all together, we have the *exact* formula for the statistical dimension

$$\delta(C) = d \cdot \frac{\sigma_{d-2}(\mathbb{S}^{d-2})}{\sigma_{d-1}(\mathbb{S}^{d-1})} \int_0^\pi \sin^{d-2}(\varphi) F(\varphi) d\varphi. \quad (6.13)$$

The sinusoidal kernel in (6.13) above peaks sharply near $\varphi = \pi/2$. Since F is smooth, this observation suggests replacing the function F by the constant value $F(\pi/2) = \sin(\alpha)$. Applying this procedure, we find

$$\delta(C) = d \sin^2(\alpha) + \underbrace{d \cdot \frac{\sigma_{d-2}(\mathbb{S}^{d-2})}{\sigma_{d-1}(\mathbb{S}^{d-1})} \int_0^\pi \sin^{d-2}(\varphi) [F(\varphi) - \sin^2(\alpha)] d\varphi}_{=: R_d(\alpha)},$$

where we have used the equality $\int_0^\pi \sin^{d-2}(\varphi) d\varphi = \frac{\sigma_{d-1}(\mathbb{S}^{d-1})}{\sigma_{d-2}(\mathbb{S}^{d-2})}$. The final step requires a bound on the remainder $R_d(\alpha)$. For this purpose, we use the following lemma that gives an exact remainder Taylor expansion for functions with Lipschitz derivative.

Lemma 6.10. *Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ has an L -Lipschitz derivative f' . Then there exists a function $\rho : \mathbb{R} \rightarrow \mathbb{R}$ such that, for every $x \in \mathbb{R}$, we have $|\rho(x)| \leq L$ and*

$$f(x) = f(0) + f'(0)x + \frac{\rho(x)}{2}x^2.$$

The proof, which is an easy extension of the usual Taylor formula with remainder, is included at the end of this demonstration. The function $f(x) := F(x + \frac{\pi}{2}) - \sin^2(\alpha)$ satisfies

$$f(0) = 0, \quad f'(0) = 2 \cos(\alpha) \sin(\alpha), \quad \text{and} \quad |f''(x)| \leq 4.$$

Applying Lemma 6.10 to the integrand in the definition of $R_d(\alpha)$ and applying the change of variables $x = \varphi - \pi/2$, there exists a function ρ such that $|\rho(x)| \leq 4$ and

$$\begin{aligned} \left(d \cdot \frac{\sigma_{d-2}(\mathbb{S}^{d-2})}{\sigma_{d-1}(\mathbb{S}^{d-1})} \right)^{-1} |R_d(\alpha)| &= \left| \int_{-\pi/2}^{\pi/2} \cos^{d-2}(x) \left[f'(0)x + \frac{\rho(x)}{2}x^2 \right] dx \right| \\ &\leq \frac{\sup_x |\rho(x)|}{2} \int_{-\pi/2}^{\pi/2} \cos^{d-2}(x) x^2 dx \\ &\leq \frac{\pi^2}{2} \int_{-\pi/2}^{\pi/2} \cos^{d-2}(x) \sin^2(x) dx \\ &= \frac{\pi^2}{2} \left(d \cdot \frac{\sigma_{d-2}(\mathbb{S}^{d-2})}{\sigma_{d-1}(\mathbb{S}^{d-1})} \right)^{-1} \end{aligned} \tag{6.14}$$

The first inequality above follows by noting that the linear term cancels because cosine is even, while the second follows by our bound on ρ and the basic inequality

$$x^2 \leq \frac{\pi^2}{4} \sin^2(x) \quad \text{for} \quad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}.$$

The final equality (6.14) can be verified with the aid of a computer algebra system. This is the claimed bound. \square

The only thing that remains is the proof of the Taylor expansion for Lipschitz functions.

Proof of Lemma 6.10. Define $\rho(0) = 0$. For $x \neq 0$, we set

$$\rho(x) := \frac{2}{x^2} (f(x) - f(0) - f'(0)x).$$

By construction, $f(x) = f(0) + f'(0)x + \frac{\rho(x)}{2}x^2$. To find a bound on the magnitude of ρ , we compute

$$\left| \frac{x^2}{2} \rho(x) \right| = |f(x) - f(0) - f'(0)x| = \left| \int_0^x (f'(y) - f'(0)) dy \right| \leq L \int_0^x y dy = L \frac{x^2}{2}.$$

The first equality follows by definition, while the second is the fundamental theorem of calculus. The inequality follows by the triangle inequality and the Lipschitz bound on f' . The final equality is elementary. We conclude $|\rho(x)| \leq L$ by comparing the left and right sides above. \square

6.2.4 Descent cones of the ℓ_∞ norm

The ℓ_∞ norm is simple enough to allow for a direct computation of the statistical dimension of its descent cones.

Proposition 6.11. *Let $I \subset \{1, \dots, d-1, d\}$ be a finite index set with $s = |I|$ elements, and suppose that $\mathbf{x} \in \mathbb{R}^d$ satisfies*

$$|x_i| = \|\mathbf{x}\|_{\ell_\infty} \quad \text{for} \quad i \in I \quad \text{while} \quad |x_j| < \|\mathbf{x}\|_{\ell_\infty} \quad \text{for} \quad j \in \{1, \dots, d-1, d\} \setminus I.$$

If $\mathbf{x} \neq \mathbf{0}$, then the statistical dimension of the descent cone of the ℓ_∞ norm at \mathbf{x} is given by

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x})) = d - \frac{s}{2}. \quad (6.15)$$

Proof. By homogeneity and signed permutation invariance of the ℓ_∞ norm, we may assume that $x_i = -1$ for $i = 1, \dots, s-1, s$ and that $-1 < x_j < 1$ for all indices $j > s$. We claim that

$$\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x}) = \mathbb{R}_+^s \times \mathbb{R}^{d-s}. \quad (6.16)$$

This fact implies the claim (6.15) because

$$\delta(\mathbb{R}_+^s \times \mathbb{R}^{d-s}) = \delta(\mathbb{R}_+^s) + \delta(\mathbb{R}^{d-s}) = \frac{s}{2} + (d-s) = d - \frac{s}{2},$$

where the first equality is the direct sum property from Proposition 6.2 and the second follows from our calculation for the statistical dimension of subspaces and orthants (Secs. 6.2.1 and 6.2.2).

It remains to argue that (6.16) holds. Suppose first that $\mathbf{y} \in \mathbb{R}_+^s \times \mathbb{R}^{d-s}$. Then there exists a $\lambda > 0$ that satisfies the following inequalities:

$$\begin{aligned} 0 \leq \lambda y_i \leq 1 \quad & \text{for } i = 1, \dots, s-1, s \quad \text{and} \\ \lambda |y_j| \leq 1 - |x_j| \quad & \text{for } j = s+1, \dots, d-1, d. \end{aligned} \quad (6.17)$$

For $i = 1, \dots, s-1, s$, positivity $\lambda y_i \geq 0$ is satisfied by all $\lambda > 0$ because $y_i \geq 0$ for these indices. The relation $\lambda y_i \leq 1$ for $i = 1, \dots, s-1, s$ is satisfied whenever $\lambda^{-1} \geq \max_{i=1, \dots, s} y_i$. As $|x_j| < 1$ for $j > s$, the final inequality in (6.17) requires $\lambda^{-1} \geq \min_{j=s+1, \dots, d} |y_j| / (1 - |x_j|)$. Hence the condition (6.17) is satisfied whenever λ^{-1} is sufficiently large.

Two basic consequences of the inequalities (6.17) are

$$\begin{aligned} |x_i + \lambda y_i| = |\lambda y_i - 1| \leq 1 \quad & \text{for } i = 1, \dots, s-1, s, \quad \text{and} \\ |x_j + \lambda y_j| \leq |x_j| + \lambda |y_j| \leq 1 \quad & \text{for } j = s+1, \dots, d-1, d. \end{aligned}$$

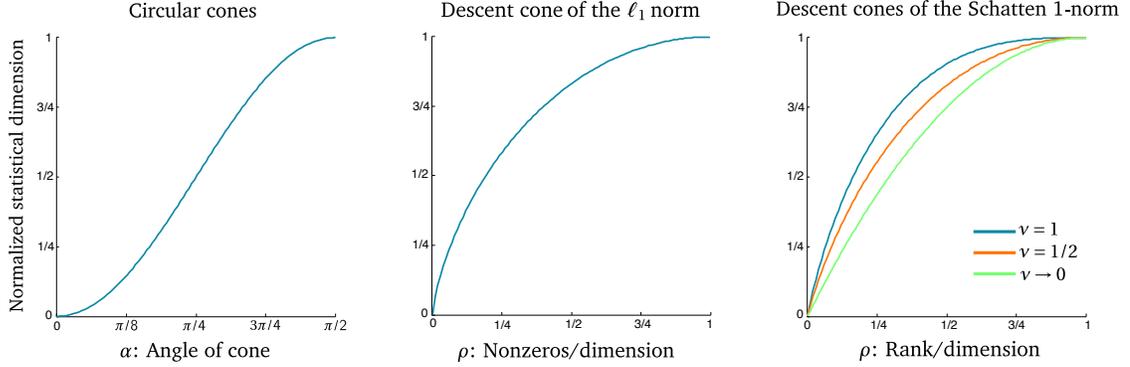


Figure 6.2: Summary of statistical dimension computations. The statistical dimension computations of Sections 6.2.3, 6.2.6, and 6.2.7 in the limit as the dimension goes to infinity. **[Left] Circular cones.** The statistical dimension $\delta(\cdot)/d$ of the circular cone $\text{Circ}_d(\alpha)$ is given by $\sin^2(\alpha)$ as $d \rightarrow \infty$. **[Center] ℓ_1 norm descent cones.** The normalized statistical dimension $\delta(\cdot)/d$ of the descent cone of the ℓ_1 norm at a $[\rho d]$ -sparse vector in \mathbb{R}^d as the $d \rightarrow \infty$. The precise formula is given by Proposition 6.14. **[Right] Schatten 1-norm descent cones.** The normalized statistical dimension $\delta(\cdot)/(mn)$ of the descent cone of the S_1 norm on $\mathbb{R}^{m \times n}$ at a matrix with rank $[\rho m]$ for several fixed aspect ratios $v = m/n$. Since the aspect ratio $v \rightarrow 0$, the limiting curve is $\rho \mapsto 2\rho - \rho^2$.

By taking a maximum over all coordinate indices, we find $\|\mathbf{x} + \lambda \mathbf{y}\|_{\ell_\infty} \leq 1 = \|\mathbf{x}\|_{\ell_\infty}$, and so $\mathbf{y} \in \mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x})$ by definition. We conclude $\mathbb{R}_+^s \times \mathbb{R}^{d-s} \subset \mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x})$.

For the other inclusion, suppose $\mathbf{y} \notin \mathbb{R}_+^s \times \mathbb{R}^{d-s}$ or equivalently, that $y_{i_*} < 0$ for some $i_* \in 1, \dots, s-1, s$. Then for any $\lambda > 0$, we have $|x_{i_*} + \lambda y_{i_*}| = 1 - \lambda y_{i_*} > 1$, and so in particular $\|\mathbf{x} + \lambda \mathbf{y}\|_{\ell_\infty} > 1 = \|\mathbf{x}\|_{\ell_\infty}$. Therefore, $\mathbf{y} \notin \mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x})$, and so we conclude that $\mathbb{R}_+^s \times \mathbb{R}^{d-s} \supset \mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{x})$. Hence the equality (6.16) holds, which completes the proof. \square

6.2.5 A recipe for general descent cones

Descent cones are an important object in this work because they link optimality conditions of convex programs to conic geometry. This section describes a technique developed by Stojnic [Sto09], and refined by Chandrasekaran et al. [CRPW12], that lets us approximate the statistical dimension of descent cones (Section 6.2.5). This technique is then applied to computing the statistical dimension of descent cones associated to the ℓ_1 norm and the Schatten 1-norm.

The following result provides a rather general technique for developing an upper bound on the statistical dimension.

Fact 6.12 (The statistical dimension of a descent cone). *Let f be a proper convex function and $\mathbf{x} \in \mathbb{R}^d$ be a point such that $\partial f(\mathbf{x})$ is nonempty, compact, and does not contain the origin. Define the map $F: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ by*

$$F: \tau \mapsto \mathbb{E}[\text{dist}^2(\mathbf{g}, \tau \cdot \partial f(\mathbf{x}))] \quad \text{for } \tau \geq 0. \quad (6.18)$$

Then we have the upper bound

$$\delta(\mathcal{D}(f, \mathbf{x})) \leq \inf_{\tau \geq 0} F(\tau). \quad (6.19)$$

The map F is strictly convex, continuous at $\tau = 0$, and differentiable for $\tau > 0$. It achieves its minimum at a unique point.

The inequality (6.19) was first observed by Stojnic [Sto09] and later extended by Chandrasekaran et al. [CRPW12]. The remaining claims are primarily the work of J. Tropp. The complete argument is available in our joint work [ALMT13, Prop. 4.4].

Fact 6.12 indicates that the best possible bound can be found simply by taking a derivative and setting it equal to zero. To make this observation formal, we write the computational procedure as Recipe 6.1. It turns out that the recipe yields surprisingly good results in a number of interesting cases.

Fact 6.13 (Error bound for recipe). *Let f be a norm on \mathbb{R}^d , and suppose $\mathbf{x} \neq \mathbf{0}$. Then*

$$\left| \delta(\mathcal{D}(f, \mathbf{x})) - \inf_{\tau \geq 0} F(\tau) \right| \leq \frac{2 \sup \{ \|\mathbf{s}\| : \mathbf{s} \in \partial f(\mathbf{x}) \}}{f(\mathbf{x} / \|\mathbf{x}\|)}, \quad (6.20)$$

where the function F is defined in (6.18).

The proof of Fact 6.13 is also due primarily to J. Tropp, and hence we refer to [ALMT13, Thm. 4.5] for the details.

6.2.6 Descent cones of the ℓ_1 norm

The ℓ_1 norm commonly appears in demixing methods where one of the underlying signals is sparse; morphological component analysis (1.3) is one particular example.

Recipe 6.1: Bounding the statistical dimension of a descent cone. This recipe is justified by Fact 6.12, which states that the minimal value of $F(\tau)$ is an upper bound on the statistical dimension of the descent cone $\delta(\mathcal{D}(f, \mathbf{x}))$. When f is a norm, Fact 6.13 provides an explicit error bound on the quality of the approximation obtained in this manner.

Assume that f is a proper convex function on \mathbb{R}^d

Assume that the subdifferential $\partial f(\mathbf{x})$ is nonempty, compact, and does not contain the origin

1. Compute the subdifferential $S = \partial f(\mathbf{x})$.
2. Find an expression for $F(\tau) := \mathbb{E}[\text{dist}^2(\mathbf{g}, \tau S)]$.
3. Determine the unique solution, if it exists, to the stationary equation $F'(\tau) = 0$.
4. If a stationary point τ_* exists, then $\delta(\mathcal{D}(f, \mathbf{x})) \leq F(\tau_*) = \inf_{\tau \geq 0} F(\tau)$.
5. Otherwise, the bound is vacuous: $\delta(\mathcal{D}(f, \mathbf{x})) \leq F(0) = d$.

The next result applies Recipe 6.1 and Fact 6.13 to generate an approximation to the statistical dimension of the ℓ_1 norm at sparse vectors with an explicit error bound. See Figure 6.2[center] for an illustration of this result.

Proposition 6.14 (Descent cones of the ℓ_1 norm). *Let \mathbf{x} be a vector in \mathbb{R}^d with $s \geq 1$ nonzero entries. Then the normalized statistical dimension of the descent cone of the ℓ_1 norm at \mathbf{x} satisfies the bounds*

$$\psi_{\ell_1}(s/d) - \frac{2}{\sqrt{sd}} \leq \frac{\delta(\mathcal{D}(\|\cdot\|_1, \mathbf{x}))}{d} \leq \psi_{\ell_1}(s/d). \quad (6.21)$$

The function $\psi_{\ell_1} : (0, 1] \rightarrow [0, 1]$ is a weakly increasing function given by

$$\psi_{\ell_1}(\rho) := (1 - \rho) \sqrt{\frac{2}{\pi}} \frac{e^{-\tau_*^2/2}}{\tau_*}, \quad 0 < \rho < 1 \quad \text{and} \quad \psi(1) = 1, \quad (6.22)$$

where $\tau_* = \tau_*(\rho)$ is the unique solution to the stationary equation

$$\int_{\tau_*}^{\infty} \frac{e^{-u^2/2}}{u^2} du = \sqrt{\frac{\pi}{2}} \cdot \frac{\rho}{1 - \rho}. \quad (6.23)$$

Proof. Because the ℓ_1 norm is invariant under signed permutations, we may assume

that $\mathbf{x} = (x_1, \dots, x_{s-1}, x_s, 0, \dots, 0)^t$ and moreover $x_i > 0$ for all $i = 1, \dots, s-1, s$. The subgradient of the ℓ_1 norm at \mathbf{x} is given by

$$\partial \|\mathbf{x}\|_{\ell_1} = \left\{ \mathbf{u} \in \mathbb{R}^d : u_i = 1 \text{ for } 1 \leq i \leq s \text{ and } |u_i| \leq 1 \text{ for } s+1 \leq i \leq d \right\}. \quad (6.24)$$

A direct computation shows that the distance between a Gaussian vector \mathbf{g} and the subgradient $\partial \|\mathbf{x}\|_{\ell_1}$ is given by

$$\text{dist}^2(\mathbf{g}, \tau \cdot \partial \|\mathbf{x}\|_{\ell_1}) = \sum_{i=1}^s (g_i - \tau)^2 + \sum_{i=s+1}^d \text{Pos}^2(|g_i| - \tau),$$

where $\text{Pos}(t) := 0 \vee t$ returns the positive part of a real number $t \in \mathbb{R}$. Since each summand above depends only on a single component of \mathbf{g} , we can compute the mean distance to the subgradient exactly:

$$\begin{aligned} F(\tau) &:= \mathbb{E}[\text{dist}^2(\mathbf{g}, \tau \cdot \partial \|\mathbf{x}\|_{\ell_1})] \\ &= s(1 + \tau^2) + (d-s) \sqrt{\frac{2}{\pi}} \left((1 + \tau^2) \int_{\tau}^{\infty} e^{-u^2/2} du - \tau e^{-\tau^2/2} \right). \end{aligned} \quad (6.25)$$

By (6.19), we have $F(\tau) \geq \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}))$ for all $\tau \geq 0$. By virtue of Fact 6.12, the function $F(\tau)$ is also strictly convex and differentiable.

When $s = d$, it is trivial to compute that $\inf_{\tau \geq 0} F(\tau) = F(0) = d$. For $s < d$, we differentiate the expression (6.25), we find that τ_* is the unique stationary point of $F(\tau)$ if and only if τ_* satisfies the stationary equation

$$(d-s) \left(\frac{1}{\tau_*} e^{-\tau_*^2/2} - \int_{\tau_*}^{\infty} e^{-u^2/2} du \right) = \sqrt{\frac{\pi}{2}} s$$

A simple application of integration-by-parts shows that the stationary equation above is equivalent to (6.23). Moreover, we may simplify our expression for $F(\tau_*)$ by solving for the integral in the display above. With this substitution and some basic algebra, we find

$$d\psi_{\ell_1}(s/d) := \inf_{\tau \geq 0} F(\tau) = F(\tau_*) = (d-s) \sqrt{\frac{2}{\pi}} \frac{e^{-\tau_*^2/2}}{\tau_*},$$

This proves the upper bound in (6.21) upon normalizing by the dimension d .

It is easy to show that $\psi_{\ell_1}(\rho) \in [0, 1]$ for $\rho \in (0, 1]$. Indeed, $F(\tau) \geq 0$ by its definition (6.25) as the average of a positive quantity. But we also have $d\psi_{\ell_1}(\rho) \leq F(0) = d$ because ψ_{ℓ_1} is defined as the infimal value of F over $\tau \geq 0$. This shows that $0 \leq \psi_{\ell_1}(\rho) \leq 1$.

We now argue that ψ_{ℓ_1} is increasing due to a geometric fact: the subgradients of the ℓ_1 norm satisfy an inclusion property. Suppose $\mathbf{x}' = (x'_1, \dots, x'_{s'}, 0, \dots, 0)^t \in \mathbb{R}_+^d$ has $s' > s$ nonzeros. Then by the subgradient calculation (6.24), we see

$$\partial \|\mathbf{x}\|_{\ell_1} \supset \partial \|\mathbf{x}'\|_{\ell_1},$$

and hence $\text{dist}(\mathbf{g}, \partial \|\mathbf{x}\|_{\ell_1}) \leq \text{dist}(\mathbf{g}, \partial \|\mathbf{x}'\|_{\ell_1})$. Therefore $F(\tau)$ is weakly increasing in the sparsity level $s = \lceil \rho d \rceil$ for each τ , and so $\psi_{\ell_1}(\rho)$ is also increasing. (In Lemma 6.15 below, we will find that ψ_{ℓ_1} is actually *strictly* increasing.)

The lower bound on ψ_{ℓ_1} follows from Fact 6.13 after some reductions. First, note that while the ℓ_1 norm depends on the value of the nonzero elements of \mathbf{x} , while the subgradient, and hence the descent cone $\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x})$, depends only on the number of nonzero elements. Therefore, we may choose the s -sparse vector \mathbf{x} that yields the most favorable bound in (6.20). Taking $\mathbf{x} = (1, \dots, 1, 0, \dots, 0)^t$, we have $\|\mathbf{x}\|_{\ell_1} = \sqrt{s}$. In addition,

$$\sup\{\|\mathbf{u}\| : \mathbf{u} \in \partial \|\mathbf{x}\|_{\ell_1}\} = \sqrt{d},$$

so that the error in our approximation is no larger than

$$\left| \frac{\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}))}{d} - \psi_{\ell_1}(\rho) \right| \leq \frac{1}{\sqrt{sd}}.$$

This gives the claim (6.21). □

The next result lists some technical facts that are important for rigorously demonstrating several interesting features of demixing problems involving the ℓ_1 norm discussed in Section 10.

Lemma 6.15. *The function ψ_{ℓ_1} defined in Proposition 6.14 is analytic on $(0, 1)$, strictly increasing, and satisfies*

$$\psi_{\ell_1}(\rho) > \rho \quad \text{for } 0 < \rho < 1 \quad \text{and} \quad \lim_{\rho \rightarrow 0} \psi_{\ell_1}(\rho) = 0. \quad (6.26)$$

Moreover,

$$\psi_{\ell_1}(\rho) < 1 \quad \text{for } 0 < \rho < 1 \quad \text{and} \quad \lim_{\rho \rightarrow 1} \psi_{\ell_1}(\rho) = 1. \quad (6.27)$$

Proof. We begin by demonstrating the properties in (6.26). We achieve the inequality by approximating an integral:

$$\frac{e^{-\tau_*^2/2}}{\tau_*} = e^{-\tau_*^2/2} \int_{\tau_*}^{\infty} \frac{1}{u^2} du > \int_{\tau_*}^{\infty} \frac{e^{-u^2/2}}{u^2} du = \sqrt{\frac{\pi}{2}} \cdot \frac{\rho}{1-\rho}.$$

The first relation is basic calculus, the inequality follows from the fact that $e^{-u^2/2}$ is a strictly decreasing function, and the final equality is the definition (6.23) of τ_* . The claim $\psi_{\ell_1}(\rho) > \rho$ follows by applying the inequality above to the definition (6.21) of ψ_{ℓ_1} .

We now consider the limit in (6.26). As ρ goes to 0, the right-hand side of the stationary equation (6.23) also tends to zero, which implies in particular that $\tau_* \rightarrow \infty$ as $\rho \rightarrow 0$. Therefore

$$\lim_{\rho \rightarrow 0} \psi_{\ell_1}(\rho) = \sqrt{\frac{2}{\pi}} \lim_{\rho \rightarrow 0} \left((1-\rho) \frac{e^{-\tau_*^2/2}}{\tau_*} \right) = 0,$$

as claimed.

The analytic implicit function theorem [KP13, Thm. 6.1.2] implies that the $\tau_* = \tau_*(\rho)$ defined implicitly by (6.23) is an analytic function for $\rho \in (0, 1)$. The analyticity of ψ_{ℓ_1} then follows from its definition (6.22) because the composition of analytic functions is analytic.

The properties of ψ_{ℓ_1} listed in (6.26) imply that $\psi_{\ell_1}(\rho)$ is not constant in a neighborhood of $\rho = 0$. Because $\psi_{\ell_1}(\rho)$ is analytic over $(0, 1)$, the fact that it is not a constant function means that it is not constant in any neighborhood of its domain [AF03,

Thm. 3.2.6], i.e., ψ_{ℓ_1} is not flat on any open subset of $(0, 1)$. Since ψ_{ℓ_1} is an increasing of ρ by Proposition 6.14, it must in fact be *strictly* increasing.

Finally, we consider the claim (6.27). The fact that ψ_{ℓ_1} is strictly increasing lets us easily deduce that $\psi_{\ell_1}(\rho) < 1$ for $\rho \in (0, 1)$. Indeed, we know that $\psi_{\ell_1}(\rho) \leq 1$ for all $\rho \in (0, 1)$ by Proposition 6.14. Since ψ_{ℓ_1} is *strictly* increasing, it cannot achieve $\psi_{\ell_1}(\rho) = 1$ for any $\rho < 1$. But $\psi_{\ell_1}(\rho) \geq \rho$, so we also have $\lim_{\rho \rightarrow 1} \psi_{\ell_1}(\rho) = 1$. This completes the final claim. \square

6.2.7 Descent cones of the Schatten 1-norm

Demixing problems involving low-rank matrices often take the Schatten 1-norm as a regularizer. The next result provides an asymptotically sharp computation for the statistical dimension of the Schatten 1-norm at a low-rank matrix.

Proposition 6.16 (Descent cones of the S_1 norm). *Let $(X(r, m, n))_{n \in \mathbb{N}}$ be a sequence of matrices $X(r, m, n) \in \mathbb{R}^{m \times n}$ with rank r . Suppose that $r, m, n \rightarrow \infty$ with limiting ratios $r/m \rightarrow \rho \in (0, 1)$ and $m/n \rightarrow \nu \in (0, 1]$. Then*

$$\frac{\delta(\mathcal{D}(\|\cdot\|_{S_1}, X(r, m, n)))}{mn} \rightarrow \psi_{S_1}(\rho, \nu). \quad (6.28)$$

The function $\psi_{S_1} : (0, 1) \times (0, 1] \rightarrow [0, 1]$ is defined as

$$\psi_{S_1}(\rho, \nu) := \rho\nu + (1 - \rho\nu) \left[\rho(1 + \tau_*^2) + (1 - \rho) \int_{a_- \vee \tau_*^2}^{a_+} (\sqrt{x} - \tau_*)^2 \cdot \varphi(x) dx \right], \quad (6.29)$$

The integration limits are $a_{\pm} := 1 \pm \sqrt{y}$, and the kernel φ is a probability density given by

$$\varphi(x) := \frac{1}{2\pi y x} \sqrt{(x - a_-)(a_+ - x)} \quad \text{for } x \in [a_-, a_+],$$

where the constant $y := (\nu - \rho\nu)/(1 - \rho\nu)$. The parameter τ_* in (6.29) is the unique solution to equation

$$\int_{a_- \vee \tau_*^2}^{a_+} \left(\frac{\sqrt{x}}{\tau_*} - 1 \right) \cdot \varphi(x) dx = \frac{\rho}{1 - \rho}. \quad (6.30)$$

See Figure 6.2[right] for a visualization of the curve (6.29) as a function of ρ for several choices of aspect ratio ν .

Proof of Proposition 6.16. For convenience, we abbreviate $\mathbf{X} = \mathbf{X}(r, m, n)$ with the understanding that r, m, n are parameters associated with \mathbf{X} . Because the Schatten 1-norm is unitarily invariant, we may assume that

$$\mathbf{X} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (6.31)$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{r \times r}$ is a diagonal matrix consisting of the singular values \mathbf{X} . The subgradient of the Schatten 1-norm at \mathbf{X} is then given by [Wat92, Ex. 2]

$$\partial \|\mathbf{X}\|_{S_1} = \left\{ \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{pmatrix} : \sigma_1(\mathbf{W}) \leq 1 \right\},$$

where $\sigma_1(\mathbf{W})$ returns the maximum singular value of \mathbf{W} and the block decomposition corresponds to the decomposition (6.31) of \mathbf{X} . Let $\mathbf{G} \in \mathbb{R}^{m \times n}$ be a Gaussian matrix whose entries are independent $\text{NORMAL}(0, 1)$ random variables. Then

$$\text{dist}^2(\mathbf{G}, \tau \partial \|\mathbf{X}\|_{S_1}) = \inf_{\substack{\tau \geq 0 \\ \sigma_1(\mathbf{W}) \leq 1}} \left\{ \left\| \begin{pmatrix} \mathbf{G}_{11} - \tau \mathbf{I} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{0} \end{pmatrix} \right\|_F^2 + \|\mathbf{G}_{22} - \tau \mathbf{W}\|_F^2 \right\}.$$

A standard argument in matrix analysis shows that the minimum over \mathbf{W} above occurs when \mathbf{G}_{22} and \mathbf{W} share the same spectrum (cf. [Bha97, Prob. III.16.13]). This reduces the minimization to a series of one-parameter minimizations that we solve explicitly:

$$\inf_{\sigma_1(\mathbf{W}) \leq 1} \|\mathbf{G}_{22} - \tau \mathbf{W}\|_F^2 = \sum_{i=1}^{m-r} \inf_{0 \leq w_i \leq 1} (\sigma_i(\mathbf{G}_{22}) - \tau w_i)^2 = \sum_{i=1}^{m-r} \text{Pos}(\sigma_i(\mathbf{G}_{22}) - \tau)^2, \quad (6.32)$$

where $\sigma_i(\mathbf{G}_{22})$ is the i th singular value of \mathbf{G}_{22} and $\text{Pos}(t) = 0 \vee t$ returns the positive

part of a number $t \in \mathbb{R}$. We also compute the average value

$$\begin{aligned} \mathbb{E} \left[\left\| \begin{pmatrix} \mathbf{G}_{11} - \tau \mathbf{I} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{0} \end{pmatrix} \right\|_F^2 \right] &= \mathbb{E}[\|\mathbf{G}_{11} - \tau \mathbf{I}\|_F^2] + \mathbb{E}[\|\mathbf{G}_{12}\|_F^2] + \mathbb{E}[\|\mathbf{G}_{21}\|_F^2] \\ &= r(m + n - r + \tau^2). \end{aligned} \quad (6.33)$$

By Fact 6.12, we have the upper bound $\delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{X})) \leq \inf_{\tau \geq 0} F(\tau)$, where $F(\tau)$ is given by

$$\begin{aligned} F(\tau) &:= \mathbb{E} [\text{dist}^2(\mathbf{G}, \tau \partial \|\mathbf{X}\|_{S_1})] \\ &= r(m + n - r + \tau^2) + \mathbb{E} \left[\sum_{i=1}^{m-r} \text{Pos}(\sigma_i(\mathbf{G}_{22}) - \tau)^2 \right]. \end{aligned} \quad (6.34)$$

The second equality is derived from (6.32) and (6.33). In principle it is possible to evaluate (6.34) using the exact singular value density of Gaussian matrices [And84, p. 534]. Instead of pursuing this technically daunting task, we turn to a classical result from random matrix theory.

Fact 6.17 (Marčenko–Pastur [MP67]). *Let $\mathbf{Z} \in \mathbb{R}^{p \times q}$ be a standard Gaussian matrix and suppose $f: \mathbb{R} \rightarrow \mathbb{R}$ is continuous. Let $\mathbf{S} := q^{-1} \mathbf{Z} \mathbf{Z}^t$ be the sample covariance matrix, and suppose that the aspect ratio $p/q \rightarrow y \in (0, 1]$ as $q \rightarrow \infty$. Then*

$$\mathbb{E} \left[\frac{1}{p} \sum_{i=1}^p f(\lambda_i(\mathbf{S})) \right] \rightarrow \int_{a_-}^{a_+} f(x) \varphi(x) dx \quad \text{as } q \rightarrow \infty, \quad (6.35)$$

where $\lambda_i(\mathbf{S})$ denotes the i th eigenvalue of \mathbf{S} . The kernel φ is defined by

$$\varphi(x) := \frac{1}{2\pi y x} \sqrt{(x - a_-)(a_+ - x)} \quad \text{for } x \in [a_-, a_+],$$

and the integration limits are $a_{\pm} := (1 \pm \sqrt{y})^2$.

In the literature, the Marčenko–Pastur result often is stated with a boundedness condition on the function f . The result above follows easily from the classical result by the almost sure convergence of the extreme eigenvalues of the sample covariance

matrix [Gem80, Sil85] and the almost sure weak convergence of the spectral density to the Marčenko–Pastur distribution [BS10, Thm. 5.8].

In order to apply (6.35) to the expectation in the definition (6.34) of F , we must make the appropriate renormalization. Since the squared singular values of a matrix are the eigenvalues of its Gram matrix, we have

$$\frac{1}{\sqrt{n-r}} \sigma_i(\mathbf{G}_{22}) = \sqrt{\lambda_i(\mathbf{S})} \quad \text{where} \quad \mathbf{S} := \frac{1}{(n-r)} \mathbf{G}_{22} \mathbf{G}_{22}^t.$$

With this identification, Fact 6.17 implies

$$\begin{aligned} \frac{1}{mn} \mathbb{E} \left[\sum_{i=1}^{m-r} \text{Pos}(\sigma_i(\mathbf{G}_{22}) - (n-r)^{1/2} \tau)^2 \right] \\ \rightarrow (1-\rho)(1-\rho\nu) \int_{a_-}^{a_+} \text{Pos}(\sqrt{x} - \tau)^2 \varphi(x) dx \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Combining this limit with the definition (6.34) of $F(\tau)$, we find

$$\frac{F(\sqrt{n-r}\tau)}{mn} \rightarrow \rho\nu + (1-\rho\nu) \left[\rho(1+\tau^2) + (1-\rho) \int_{a_- \vee \tau^2}^{a_+} (\sqrt{x} - \tau)^2 \varphi(x) dx \right]$$

as $n \rightarrow \infty$. The second derivative test shows that the limiting expression above is strictly convex in τ , and so the minimal value of the expression on the right above is achieved for the unique τ_* that satisfies the stationary equation

$$\int_{a_- \vee \tau_*^2}^{a_+} \left(\frac{\sqrt{x}}{\tau_*} - 1 \right) \varphi(x) dx = \frac{\rho}{1-\rho}.$$

This stationary equation corresponds to (6.30). Applying the formula for the limiting value of $F(\tau)$, we see

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{X}))}{mn} &\leq \liminf_{n \rightarrow \infty} \frac{F(\sqrt{n-r}\tau)}{mn} \\ &= \rho\nu + (1-\rho\nu) \left[\rho(1+\tau_*^2) + (1-\rho) \int_{a_- \vee \tau_*^2}^{a_+} (\sqrt{x} - \tau_*)^2 \varphi(x) dx \right], \quad (6.36) \end{aligned}$$

where the inequality follows from Fact 6.2. The equality follows because we are taking the limit of convex functions, so that pointwise convergence ensures *epigraphical convergence* [RW98, Thm. 7.17], which in turn implies convergence of the infimal values [RW98, Thm. 7.33].

For the lower bound, we will make use of the error estimate (6.20). First note that

$$\left\| \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{pmatrix} \right\|_{\mathbb{F}}^2 \leq m \quad \text{for all } \sigma_1(\mathbf{W}) \leq 1,$$

where the decomposition aligns with the decomposition (6.31) of \mathbf{X} above. Therefore, the numerator of (6.20) is no larger than $2\sqrt{m}$. For the denominator, note that the subgradient $\partial\|\mathbf{X}\|_{S_1}$ depends only on the rank of \mathbf{X} , and not the particulars of the singular values. Hence, we may choose the most favorable \mathbf{X} , which in this case amounts to taking $\mathbf{X} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, a truncated identity matrix. This leads to the bound

$$\left\| \frac{\mathbf{X}}{\|\mathbf{X}\|_{S_1}} \right\|_{S_1} = \frac{1}{\sqrt{r}} \left\| \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right\|_{S_1} = \sqrt{r}.$$

Our error bound (6.20) delivers the inequality

$$\frac{1}{mn} \left| \delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{X})) - \inf_{\tau \geq 0} F(\tau) \right| \leq 2\sqrt{\frac{1}{mrn^2}} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

This estimate, combined with (6.36), gives the claim. \square

6.3 Relationship to the Gaussian width

A summary parameter closely associated with the statistical dimension is the *Gaussian width* of a set. The Gaussian width is the central geometric parameter in the Gaussian process inequalities of Gordon [Gor85, Gor87, Gor88]. These inequalities were first applied to the analysis of convex optimization programs in [RV08], which inspired a number of other works in the same vein [Sto09, OH10, CRPW12]. This section describes the close relationship between the Gaussian width and the statistical dimension and

argues that these parameters are essentially interchangeable.

Definition 6.18. The *Gaussian width* $w(S)$ of a set $S \subset \mathbb{R}^d$ is given by

$$w(S) := \mathbb{E}[\sup_{\mathbf{x} \in S} \langle \mathbf{g}, \mathbf{x} \rangle]. \quad (6.37)$$

As noted in [CRPW12, Sec. 3.2], the Gaussian width is proportional to the classical mean width of a set, making it a common fixture in the geometric analysis of convex bodies. However, for unbounded sets, and cones in particular, the Gaussian width takes an infinite value. There are two natural ways to extend the Gaussian width to cones. The first, more common, approach is to restrict the cone to the unit sphere, while the second is the restriction to the solid unit ball.

It turns out that it makes rather little difference which convention we use to define the Gaussian width of a cone. Moreover, for most purposes, the statistical dimension $\delta(C)$ is a sufficiently accurate estimate of the (squared) Gaussian width.

Theorem 6.19. *For any convex cone $C \subset \mathbb{R}^d$, we have the string of inequalities*

$$w^2(C \cap S^{d-1}) \leq w^2(C \cap B_d) \leq \delta(C) \leq w^2(C \cap S^{d-1}) + 1. \quad (6.38)$$

We note that the first two inequalities follow immediately from [CRPW12, (30) and (33)], but we provide a full proof for completeness. We require the following bound on Lipschitz functions of Gaussian random variables.

Fact 6.20 (Variance of a Lipschitz function). *Let $H : \mathbb{R}^d \rightarrow \mathbb{R}$ be a function that is L -Lipschitz with respect to the Euclidean norm. Then $\text{Var}(H(\mathbf{g})) \leq L^2$, where \mathbf{g} is a standard normal vector.*

This fact follows immediately from the Gaussian Poincaré inequality [Bec89, Eq. (3)] with $p = 1$.

Proof of Theorem 6.19. The first inequality follows immediately from the containment $C \cap S^{d-1} \subset C \cap B_d$ and the definition (6.37) of the Gaussian width. For the second

relation in (6.38), we have the string of relations

$$w^2(C \cap B_{d-1}) = \mathbb{E} \left[\sup_{\mathbf{x} \in C \cap B_d} \langle \mathbf{x}, \mathbf{g} \rangle \right]^2 = \mathbb{E} [\inf_{\mathbf{x} \in C^\circ} \|\mathbf{x} - \mathbf{g}\|]^2 \leq \mathbb{E} [\text{dist}^2(\mathbf{g}, C^\circ)]$$

The first equality is the definition (6.37) of the Gaussian width, while the second relation is (6.8) from the proof of Proposition 6.3.4. The final relation combines Jensen's inequality with the definition (3.13) of the distance to a cone. We recognize the right-hand side as $\delta(C)$ by the distance characterization (6.4) of the statistical dimension, which demonstrates the second inequality in (6.38).

The final bound in (6.38) requires slightly more effort. Define the random variable $Z := \sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{g} \rangle$. We claim Z is a 1-Lipschitz function of \mathbf{g} with mean $\mathbb{E}[Z] = w(C \cap S^{d-1})$. To see this, consider $\mathbf{y}, \mathbf{z} \in \mathbb{R}^d$. We may assume without loss that C is closed because the closure operation does not affect the supremum. Then there exists an $\mathbf{x}_* \in C \cap S^{d-1}$ such that

$$\sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{y} \rangle - \sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{x}_*, \mathbf{y} \rangle - \sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{z} \rangle \leq \langle \mathbf{x}_*, \mathbf{y} - \mathbf{z} \rangle \leq \|\mathbf{y} - \mathbf{z}\|,$$

where we used the fact that \mathbf{x}_* is unit norm for the last inequality. The same argument yields the analogous bound for the negative difference, and the 1-Lipschitz property follows immediately. We bound the variance of Z using Gaussian concentration:

$$\mathbb{E}[Z^2] - w(C \cap S^{d-1})^2 = \mathbb{E}[(Z - \mathbb{E}[Z])^2] \leq 1, \quad (6.39)$$

where the inequality follows from the bound on the variance of a Lipschitz function of a Gaussian, Fact 6.20.

The claim follows immediately after we demonstrate that $\mathbb{E}[Z^2] \geq \delta(C)$. To verify this inequality, note that positivity of Z^2 implies

$$\mathbb{E}[Z^2] \geq \mathbb{E} \left[\left(\sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{g} \rangle \right)^2 \cdot \mathbf{1}(\mathbf{g} \notin C^\circ) \right].$$

We claim the right-hand side above is in fact equal to $\delta(C)$. Indeed, for any $\mathbf{y} \notin$

C° , we have the equality $\sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{y} \rangle = \sup_{\mathbf{x} \in C \cap B_d} \langle \mathbf{x}, \mathbf{y} \rangle$ by a basic homogeneity argument. Moreover, when $\mathbf{y} \in C^\circ$, we have $\sup_{\mathbf{x} \in C \cap B_d} \langle \mathbf{x}, \mathbf{y} \rangle = 0$. By combining these two observations, we see that

$$\mathbb{E} \left[\left(\sup_{\mathbf{x} \in C \cap S^{d-1}} \langle \mathbf{x}, \mathbf{g} \rangle \right)^2 \cdot \mathbf{1}(\mathbf{g} \notin C^\circ) \right] = \mathbb{E} \left[\left(\sup_{\mathbf{x} \in C \cap B_d} \langle \mathbf{x}, \mathbf{g} \rangle \right)^2 \right] = \delta(C),$$

where the final equality follows from the supremum characterization (6.5) of the statistical dimension. We conclude that $\mathbb{E}[Z^2] \geq \delta(C)$. Applying this result to (6.39), we arrive at the upper inequality of (6.38). \square

Chapter 7

Concentration of intrinsic volumes

In this section, we prove a new concentration inequality involving the intrinsic volumes of convex cones. The results of this section are joint work with collaborators, which appear in [ALMT13, Thm. 6.1]. Before stating the result, we introduce some summary parameters that simplify the exposition of our main result. For any cone $C \in \mathcal{C}_d$, we define the *transition width*

$$\omega(C) := \sqrt{\delta(C) \wedge \delta(C^\circ)}, \quad (7.1)$$

where the wedge \wedge returns the minimum of two numbers. Further, introduce the concentration function $p_C(\lambda)$ which is defined by

$$p_C(\lambda) := 4 \exp\left(-\frac{\lambda^2/8}{\omega(C)^2 + \lambda}\right). \quad (7.2)$$

Then we have the following result.

Theorem 7.1 (Concentration of intrinsic volumes). *Let $C \in \mathcal{C}_d$ and fix any $\lambda \geq 0$. Let k_+, k_- be indices such that*

$$k_- \leq \delta(C) - \lambda + 1 \quad \text{and} \quad k_+ \geq \delta(C) + \lambda.$$

Then we have the tail bounds

$$t_{k_-}(C) \geq 1 - p_C(\lambda) \quad \text{and} \quad t_{k_+}(C) \leq p_C(\lambda). \quad (7.3)$$

The schema of this proof is due essentially to M. Lotz, but each of the coauthors of [ALMT13] played a substantial role in developing the current result.¹

The bound (7.3) demonstrates that the tail functionals $t_k(C)$ change from near one to near zero over a range of $O(\omega(C))$ indices, and that this transition region occurs near the statistical dimension $\delta(C)$. The width $\omega(C) \leq \sqrt{d/2}$ for any cone $C \in \mathcal{C}_d$ by the polarity rule of Proposition 6.2.4, which implies that essentially all of the intrinsic volumes are negligible in high dimensions. Given the important role of intrinsic volumes in determining the geometric properties of convex cones, the concentration property provides a rigorous justification for the qualitative observation that a convex cone C with statistical dimension $\delta(C)$ sometimes behaves like a subspace of dimension $[\delta(C)]$.

In the next section, we discuss some of the parallels between the bound (7.3) and classical inequalities in Euclidean integral geometry. The full proof of Theorem 7.1 appears in Section 7.2.

7.1 Parallels with Euclidean integral geometry

Because of analogies between conic and Euclidean integral geometry, Schneider & Weil [SW08, p. 263] express interest in inequalities among the conic intrinsic volumes v_k and the tail functionals t_k and h_k . To the best of our knowledge, there are only two other nontrivial inequalities relating these quantities. The first is a direct consequence of the classical spherical isoperimetric inequality, which provides that, among all cones with some fixed value of $v_d(C)$, the smallest value of $v_{d-1}(C)$ is attained precisely when C is a circular cone. The other inequality² between conic intrinsic volumes asserts that circular cones maximize $v_0(C)$ over all cones with $v_d(C)$ fixed [GHS03].

Theorem 7.1, coupled with the interlacing inequalities of Lemma 8.3, provides a rich family of inequalities relating the conic intrinsic volumes to the tail functionals. Let us

¹This author independently proved a very similar result by means of a Gaussian width inequality of Gordon [Gor85, Gor87, Gor88], but the approach we present here is more intrinsic to the theory of conic integral geometry. One novelty in our current presentation is the use of the Gaussian Steiner formula (Proposition 5.6).

²This second inequality can also be derived as a straightforward consequence of spherical isoperimetry, at least for the case of convex cones.

highlight one intriguing analogy between our inequalities and the classical inequalities for convex bodies Euclidean space.

Corollary 7.2. *For any closed cone $C \in \mathcal{C}_d$, we have the upper bound*

$$v_k(C) \leq s_k(C) := 2 \exp\left(-\frac{(k - \delta(C))^2/8}{\omega(C)^2 + |k - \delta(C)|}\right) \quad \text{for every } k = 0, \dots, d-1, d. \quad (7.4)$$

If $C \notin \{\mathbf{0}, \mathbb{R}^d\}$, then the sequence $(s_k(C))_{k=0}^d$ is strictly log-concave:

$$s_k(C)^2 > s_{k-1}(C)s_{k+1}(C) \quad \text{for every } k = 1, \dots, d-2, d-1.$$

Proof. For $k > \delta(C)$, we set $\lambda = k - 1 - \delta(C)$ and apply the first bound of (7.3) to find

$$v_k(C) \leq h_k(C) \leq \frac{1}{2}t_{k-1}(C) \leq \frac{1}{2}p_C(\lambda),$$

where the first inequality follows by the definition (5.10) of tail functionals, the second is the interlacing inequality of Lemma 8.3, and the final inequality is Theorem 7.1. When $k \leq \delta(C)$, we set $\lambda = \delta(C) - k$ and use the polarity relationship (5.6) to find

$$v_k(C) = v_{d-k}(C^\circ) \leq \frac{1}{2}t_{d-k-1}(C^\circ) \leq \frac{1}{2}p_{C^\circ}(\lambda),$$

where the inequalities follow as above. The claim (7.4) follows from the definition (7.2) of $p_C(\lambda)$ and the fact that $\omega(C^\circ) = \omega(C)$.

It remains to argue the log-concavity of the bound. Corollary 6.4 immediately implies $\omega^2(C) \geq \frac{1}{2}$ when C is a nontrivial cone, so the map $u \mapsto u^2/(\omega^2(C) + |u|)$ is strictly log-concave. The concavity of the exponent implies that the sequence $k \mapsto s_k(C)$ is log-concave. \square

Corollary (7.4) implies that the sequence of conic intrinsic volumes behaves like a log-concave sequence as conjectured in [Ame11, Conj. 4.4.16]. The corresponding log-concave behavior of the Euclidean intrinsic volumes is an important consequence of the deep Alexandrov–Fenchel inequalities. Notable special cases of this classical result

include the usual isoperimetric inequality for convex bodies, Minkowski's inequality, and Urysohn's inequality [Sch93, Ch. 6]. Theorem 7.1 provides a system of inequalities for conic intrinsic volumes that parallel these deep classical results.

7.2 Proof of Theorem 7.1

The proof of inequality (7.3) requires two auxiliary propositions. The proofs are technical and have little geometric content, and hence are postponed until Section 7.3.

Proposition 7.3 (Bounds on the chi-square cdf). *For every integer $k \geq 1$,*

$$\frac{1}{2} \leq \chi_k^2(k) \leq \chi_1^2(1) < 0.683, \quad (7.5)$$

where $\chi_k^2(\cdot)$ is the cumulative distribution function of a chi-square random variable with k degrees of freedom (5.12). The lower bound also holds for $k = 0$. Moreover, $\chi_j^2(k)$ is decreasing in j for each fixed argument k .

The proof, available in Section 7.3.2, depends on an integral approximation argument. The second result is a concentration inequality for the projection of a Gaussian vector onto a convex cone.

Proposition 7.4 (Deviation of conic projections). *For any $C \in \mathcal{C}_d$ and $\lambda \geq 0$, we have the bounds*

$$\mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 \geq \delta(C) + \lambda\} \leq q_C(\lambda) \quad \text{and} \quad \mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 \leq \delta(C) - \lambda\} \leq q_C(\lambda)$$

where

$$q_C(\lambda) := \exp\left(-\frac{\lambda^2/8}{\delta(C) + \lambda}\right).$$

The proof uses a standard moment comparison approach. See Section 7.3.1 for the details. We are now ready for the proof of our main result.

Proof of Theorem 7.1. Suppose $k_+ \geq \delta(C) + \lambda$. We have the estimate

$$\begin{aligned} t_{k_+}(C) &= \sum_{j=k_+}^d v_j(C) \left([1 - \chi_j^2(k_+)] + \chi_j^2(k_+) \right) \\ &\leq \sum_{j=0}^d v_j(C) [1 - \chi_j^2(k_+)] + \chi_{k_+}^2(k_+) \sum_{j=k_+}^d v_j(C) \\ &\leq \mathbb{P} \{ \|\mathbf{\Pi}_C(\mathbf{g})\|^2 > k_+ \} + 0.7 \sum_{j=k_+}^d v_j(C), \end{aligned}$$

where the first inequality follows from the positivity of intrinsic volumes and monotonicity of $\chi_j^2(k)$ given by Proposition 7.3. The final inequality follows from the Gaussian Steiner formula (5.11) and the upper bound (7.5). We recognize the final sum as $t_{k_+}(C)$. By subtracting $0.7t_{k_+}(C)$ from both sides above and applying the fact that $k_+ \geq \delta(C) + \lambda$, we have

$$t_{k_+}(C) \leq \frac{1}{0.3} \mathbb{P} \{ \|\mathbf{\Pi}_C(\mathbf{g})\|^2 > \delta(C) + \lambda \} \leq 4q_C(\lambda)$$

by Proposition 7.4.

The bound $q_C(\lambda)$ is best when $\delta(C)$ is small. However, when $\delta(C) \approx d$, the polarity rule of Proposition 6.2.4 implies $\delta(C^\circ) \ll d$. A dual version of the same argument above provides a bound on the tail functional $t_{k_+}(C)$ in terms of $q_{C^\circ}(\lambda)$:

$$\begin{aligned} t_{k_+}(C) &= \sum_{j=k_+}^d v_j(C) \left(\chi_{d-j}^2(d - k_+) + [1 - \chi_{d-j}^2(d - k_+)] \right) \\ &\leq \sum_{j=0}^d v_j(C) \chi_{d-j}^2(d - k_+) + [1 - \chi_{d-k_+}^2(k_+)] \sum_{j=k_+}^d v_j(C) \\ &\leq \sum_{j=0}^d v_j(C^\circ) \chi_j^2(d - k_+) + 0.5 \sum_{j=k_+}^d v_j(C). \end{aligned}$$

The final inequality requires the fact that $v_j(C) = v_{d-j}(C^\circ)$, owing to the polarity property of intrinsic volumes (5.6), and it also uses the lower bound on the chi-square cumulant from Proposition 7.3. The final sum is equal to $t_{k_+}(C)$, and so by rearranging the inequality and applying the Gaussian Steiner formula again we obtain the bound

$$\sum_{j=k_+}^d v_j(C) \leq 2\mathbb{P} \{ \|\mathbf{\Pi}_{C^\circ}(\mathbf{g})\|^2 \leq d - \delta(C) - \lambda \} \leq 2q_{C^\circ}(\lambda).$$

The final inequality above requires the polarity property $\delta(C^\circ) = d - \delta(C)$ from Proposition 6.2 and the lower deviation inequality of Proposition 7.4. Combining this estimate with our previous bound, we find

$$t_{k_+}(C) \leq 4 [q_C(\lambda) \wedge q_{C^\circ}(\lambda)] = p_C(\lambda). \quad (7.6)$$

This is the first bound of (7.3). The second inequality follows from the first by duality. Indeed, we have $v_j(C) = v_{d-j}(C^\circ)$ by (5.6), so that

$$1 - t_{k_-}(C) = \sum_{j=0}^{k_-} v_j(C) = \sum_{j=d-k_-}^d v_j(C^\circ) = t_{d-k_-}(C^\circ)$$

Applying the polarity relation $\delta(C) = d - \delta(C^\circ)$ of Proposition 6.2 and our assumption on k_- , we see that

$$d - k_- \geq d - \delta(C) + \lambda = \delta(C^\circ) + \lambda.$$

Combining our last two displays yields the result

$$1 - t_{k_-}(C) = t_{d-k_-}(C^\circ) \leq p_{C^\circ}(\lambda) = p_C(\lambda),$$

where the inequality follows from the bound (7.6) developed earlier. \square

7.3 Proof of technical propositions

The rest of this section is devoted to the proof of our technical lemmas.

7.3.1 Concentration of projections onto cones

The concentration inequality for conic projections is our next task. The proof below is due to J. Tropp, and appears in the joint work [ALMT13].

Proof of Proposition 7.4. We prove the upper deviation inequality first. By exponentiat-

ing and applying Markov's inequality, the Laplace transform methods provides

$$\mathbb{P} \{ \|\Pi_C(\mathbf{g})\|^2 \geq \delta(C) + \lambda \} \leq e^{-\xi\lambda} \mathbb{E} e^{\xi(\|\Pi_C(\mathbf{g})\|^2 - \delta(C))}, \quad (7.7)$$

valid for all $\xi > 0$. We wish to apply a moment bound from [Bog98, Cor. 1.7.9].

Fact 7.5. *Let $F: \mathbb{R}^d \rightarrow \mathbb{R}$ be a weakly differentiable function such that $\mathbb{E}[F(\mathbf{g})] = 0$ and*

$$\mathbb{E} [|F(\mathbf{g})|^2 + \|\nabla F(\mathbf{g})\|^2] < \infty.$$

Then the moment bound

$$\mathbb{E} e^{\xi F(\mathbf{g})} \leq \left(\mathbb{E} e^{(\xi/4)\|\nabla F(\mathbf{g})\|^2} \right)^{2\xi/(1-2\xi)} \quad \text{holds for all } 0 < \xi < \frac{1}{2}. \quad (7.8)$$

To apply this result to our present situation, we first verify that the regularity conditions of Fact 7.5 hold for the exponent appearing in (7.7). Writing $F(\mathbf{x}) := \|\Pi_C(\mathbf{x})\|^2 - \delta(C)$, we see that $F(\mathbf{g})$ is a zero-mean random variable bounded by

$$|F(\mathbf{x})|^2 \leq 2 \|\Pi_C(\mathbf{x})\|^4 + 2\delta^2(C) \leq 2 \|\mathbf{x}\|^4 + 2\delta^2(C).$$

Moreover, by virtue of (3.19), F is differentiable and $\|\nabla F(\mathbf{x})\|^2 = 4 \|\Pi_C(\mathbf{x})\|^2 \leq 4 \|\mathbf{x}\|^2$.

Because all moments of a chi-square random variable are finite, we see

$$\mathbb{E} [|F(\mathbf{g})|^2 + \|\nabla F(\mathbf{g})\|^2] \leq 2\delta(C) + 2\mathbb{E}[\|\mathbf{g}\|^4] + 4\mathbb{E}[\|\mathbf{g}\|^2] < \infty.$$

With the regularity conditions out of the way, we may apply Fact 7.5 to the expectation in (7.7) to find

$$\mathbb{E} e^{\xi F(\mathbf{g})} \leq \left(\mathbb{E} e^{(\xi/4)\|\nabla F(\mathbf{g})\|^2} \right)^{2\xi/(1-2\xi)} \quad \text{whenever } 0 < \xi < \frac{1}{2}.$$

The formula $\|\nabla F(\mathbf{x})\|^2 = 4(F(\mathbf{x}) + \delta(C))$ follows from the formula (3.19) for the

gradient of the projection onto a cone, so we have

$$\mathbb{E} e^{\xi F(\mathbf{g})} \leq \exp\left(\frac{2\xi^2 \delta(C)}{1-2\xi}\right) (\mathbb{E} e^{\xi F(\mathbf{g})})^{2\xi/(1-2\xi)}.$$

Since $2\xi/(1-2\xi) < 1$ for $0 < \xi < \frac{1}{4}$, the inequality above implies

$$\mathbb{E} e^{\xi F(\mathbf{g})} \leq \exp\left(\frac{2\xi^2 \delta(C)}{1-4\xi}\right) \quad \text{for } 0 < \xi < \frac{1}{4}.$$

Applying this bound to our Laplace transform inequality (7.7) yields our deviation inequality for the upper tail:

$$\mathbb{P}\{\|\mathbf{\Pi}_C(\mathbf{g})\|^2 \geq \delta(C) + \lambda\} \leq \inf_{0 < \xi \leq 1/4} \exp\left(-\xi\lambda + \frac{2\xi^2 \delta(C)}{1-4\xi}\right) \leq \exp\left(\frac{-\lambda^2/8}{\delta(C) + \lambda}\right),$$

where the final bound follows by the specific choice $\xi = (\lambda/4)/(\delta(C) + \lambda)$.

The lower tail requires a similar, but slightly more intricate, argument. In the same manner as above, our first step is an application of the Laplace transform method:

$$\mathbb{P}\{\|\mathbf{\Pi}_C(\mathbf{g})\|^2 \leq \delta(C) - \lambda\} = \mathbb{P}\{\delta(C) - \|\mathbf{\Pi}_C(\mathbf{g})\|^2 \geq \lambda\} \leq e^{-\xi\lambda} \mathbb{E} e^{-\xi F(\mathbf{g})},$$

valid for all $\xi > 0$, where the function F is as defined above. Fact 7.5 yields the moment bound

$$\mathbb{E} e^{-\xi F(\mathbf{g})} \leq \left(\mathbb{E} e^{\xi \|\mathbf{\Pi}_C(\mathbf{g})\|^2}\right)^{2\xi/(1-2\xi)} \quad \text{whenever } 0 < \xi < \frac{1}{2}. \quad (7.9)$$

Due to the negative exponent on the left-hand side above, the upper and lower portions of this inequality are not comparable. To get a handle on the right-hand side, we apply Fact 7.5 to the term in parenthesis to find

$$\mathbb{E} e^{\xi \|\mathbf{\Pi}_C(\mathbf{g})\|^2} = e^{\xi \delta(C)} \mathbb{E} e^{\xi F(\mathbf{g})} \leq e^{\xi \delta(C)} \left(\mathbb{E} e^{\xi \|\mathbf{\Pi}_C(\mathbf{g})\|^2}\right)^{2\xi/(1-2\xi)}$$

for all $0 < \xi < \frac{1}{2}$. Solving this relation reveals

$$\mathbb{E} e^{\xi \|\mathbf{\Pi}_C(\mathbf{g})\|^2} \leq \exp\left(\frac{\xi(1-2\xi)}{1-4\xi} \delta(C)\right) \quad \text{for } 0 < \xi < \frac{1}{4}.$$

By combining this inequality with (7.9), we find the same upper bound as before:

$$\mathbb{P} \{ \|\mathbf{\Pi}_C(\mathbf{g})\|^2 \leq \delta(C) - \lambda \} \leq \exp \left(-\xi\lambda + \frac{2\xi^2}{1-4\xi} \delta(C) \right) \quad \text{whenever } 0 < \xi < \frac{1}{4}.$$

The specific choice $\xi = (\lambda/4)/(\delta(C) + \lambda)$ completes the proof. \square

7.3.2 An inequality for chi-square random variables

Recall that a chi-square random variables with k degrees of freedom has the cumulative distribution given by

$$\chi_k^2(\varepsilon) := \mathbb{P} \{ \|\mathbf{g}_k\|^2 \leq \varepsilon \} = \frac{1}{2^{k/2} \Gamma(\frac{k}{2})} \int_0^\varepsilon x^{k/2-1} e^{-x/2} dx \quad \text{for all } k \geq 1, \quad (7.10)$$

where \mathbf{g}_k is a k -dimensional Gaussian vector and $\varepsilon \geq 0$. It is consistent with this definition to set $\chi_0^2(\varepsilon) := 1$ for all $\varepsilon \geq 0$.

Proof of Proposition 7.3. The lower bound in (7.5) is an immediate consequence of the mean–median–mode inequality for the gamma distribution [GM77, Sec. 3(a)] when $k \geq 1$. When $k = 0$, the lower bound holds trivially by our definition.

The upper bound requires more work. For $k = 1$, the inequality is trivial. For integers $2 \leq k \leq 8$, the bound (7.5) may be verified numerically. For $k \geq 9$, our strategy is to apply a second-order approximation to the exponent in the integral

$$1 - \chi_k^2(k) = \frac{1}{2^{k/2} \Gamma(\frac{k}{2})} \int_k^\infty \exp \left(\left(\frac{k}{2} - 1 \right) \log(x) - \frac{x}{2} \right) dx \quad (7.11)$$

and combine the resulting expression with Stirling's formula. For this purpose, we develop a quadratic lower bound for the logarithm. For $x \geq 0$, we compute

$$\frac{d}{dx} \left(\log(1+x) - x + \frac{x^2}{2} \right) = \frac{x^2}{x+1} \geq 0.$$

Integrating this relationship yields the inequality $\log(1+x) - x + \frac{x^2}{2} \geq \log(1) = 0$ for

all $x \geq 0$. We bootstrap this simple inequality to find the more elaborate relation

$$\log(x+k) = \log(k) + \log\left(1 + \frac{x-k}{k}\right) \geq \log(k) + \frac{x-k}{k} - \frac{(x-k)^2}{2k^2}$$

which holds for all $x \geq k > 0$. Since $k/2 > 1$ for $k \geq 9$, we may apply this inequality to the exponent of the integrand in (7.11):

$$\begin{aligned} \int_k^\infty x^{k/2-1} e^{-x/2} dx &\geq k^{k/2-1} \int_k^\infty \exp\left(\left(\frac{k}{2}-1\right)\left(\frac{x-k}{k} - \frac{(x-k)^2}{2k^2}\right) - \frac{x}{2}\right) dx \\ &= k^{k/2-1} e^{-k/2} \int_0^\infty \exp\left(\left(\frac{1}{2k^2} - \frac{1}{4k}\right)y^2 - \frac{y}{k}\right) dy. \end{aligned}$$

The equality follows by the change of variables $y = x - k$ and algebraic simplification. Dropping the nonnegative term $y^2/(2k^2)$ in the exponent of the integrand above, we continue our estimate

$$\begin{aligned} \int_k^\infty x^{k/2-1} e^{-x/2} dx &\geq k^{k/2-1} e^{-k/2} \int_0^\infty \exp\left(-\frac{y}{k} - \frac{1}{4k}y^2\right) dy \\ &= \sqrt{\pi}k^{(k-1)/2} e^{1/k-k/2} \operatorname{erfc}\left(\sqrt{\frac{1}{k}}\right). \end{aligned} \quad (7.12)$$

The final relation (7.12) can be verified with the aid of a computer algebra system.

An inequality form of Stirling's formula [DLMF, OLBC10, 5.6.1] provides a bound on the normalization constant appearing in (7.11):

$$2^{k/2} \Gamma\left(\frac{k}{2}\right) < 2\sqrt{\pi}k^{(k-1)/2} e^{1/(6k)-k/2}. \quad (7.13)$$

Applying the estimates (7.12) and (7.13) to the expression (7.11), we find

$$1 - \chi_k^2(k) > \frac{e^{5/(6k)}}{2} \operatorname{erfc}\left(\sqrt{\frac{1}{k}}\right) > \frac{1}{2} \operatorname{erfc}\left(\frac{1}{3}\right) > 1 - \chi_1^2(1),$$

where the second inequality follows from $e^{5/(6k)} > 1$ and the fact that $k \geq 9$. The final inequality can be verified numerically, which completes the proof of the upper bound

in (7.5).

Finally, we argue that $\chi_j^2(k) \geq \chi_{j+1}^2(k)$. Let $\mathbf{g} \in \mathbb{R}^{j+1}$ be a Gaussian vector. Then

$$\begin{aligned}\chi_j^2(k) &= \mathbb{P} \left\{ \sum_{i=1}^j g_i^2 \leq k \right\} = \mathbb{P} \left\{ \|\mathbf{g}\|^2 - g_{j+1}^2 \leq k \right\} \\ &\geq \mathbb{P} \left\{ \|\mathbf{g}\|^2 \leq k \right\} = \chi_{j+1}^2(k),\end{aligned}$$

where the inequality follows from the fact that $g_{j+1}^2 \geq 0$. This completes the proof. \square

Chapter 8

Kinematic consequences

This chapter develops the consequence of Theorem 7.1 most important for demixing: the sum of statistical dimensions controls the probability that cones strike. Many of the ideas in this section are drawn from joint work [ALMT13], but the iterated approximate kinematic formula (Theorem 8.2) appears here for the first time.

Theorem 8.1 (Approximate kinematics). *Let $C, C' \in \mathcal{C}_d$ be two closed, convex cones, and let $\mathbf{Q} \in \mathbf{O}_d$ be a random rotation. For any $\lambda > 0$, the following implications hold:*

$$\begin{aligned} \delta(C) + \delta(C') \leq d - 2\lambda &\implies \mathbb{P}\{C \cap \mathbf{Q}C' \neq \{\mathbf{0}\}\} \leq p_C(\lambda) + p_{C'}(\lambda); \\ \delta(C) + \delta(C') \geq d + 2\lambda &\implies \mathbb{P}\{C \cap \mathbf{Q}C' \neq \{\mathbf{0}\}\} \geq 1 - p_C(\lambda) - p_{C'}(\lambda). \end{aligned}$$

The function $p_C(\lambda)$ is defined in (7.2).

This bound is a corollary of a more intricate bound presented in Section 8.1. See Section 8.1 for the proof of this theorem from our more sophisticated result.

The interpretation of Theorem 8.1 is simple. When the total statistical dimension of two cones is slightly less than the ambient dimension, there are very few rotations that bring the cones into contact. On the other hand, when the total statistical dimension is a bit larger than the ambient dimension, almost every rotation brings the cones into contact. This interpretation mirrors the familiar fact about generically oriented subspaces in linear algebra.

Section 8.1 below presents a more general version of 8.1 involving an arbitrary number of cones and proves Theorem 8.1 as a consequence. The rest of the chapter

is devoted to the proof of our general result. We wait until Chapter 9 to explore the implications of these bounds for demixing problems.

8.1 Iterated approximate kinematics

The iterated kinematic formulas of Section 5.3.1 allow us to prove an approximate kinematic formula for an arbitrary number of cones. This result, while complicated at first sight, is a straightforward generalization of Theorem 8.1. Moreover, as a consequence of the geometric optimality conditions for multiple demixing (Theorem 4.2), intersection probabilities involving multiple cones are crucial for applications.

Theorem 8.2. *Let $C_1, \dots, C_{n-1}, C_n \in \mathcal{C}_d$ be closed, convex cones, and let L be an arbitrary m -dimensional subspace of \mathbb{R}^d . Define the total statistical dimension of the cones $\Delta := \sum_{i=1}^n \delta(C_i)$, and suppose $\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}, \mathbf{Q}_n \in \mathbf{O}_d$ are i.i.d. random rotations. Then, for any $\lambda > 0$, the following implications hold:*

$$\Delta + m \leq n(d - \lambda) \implies \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \dots \cap \mathbf{Q}_n C_n \cap L \neq \{\mathbf{0}\}\} \leq \sum_{i=1}^n p_{C_i}(\lambda); \quad (8.1)$$

$$\Delta + m \geq n(d + \lambda) \implies \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \dots \cap \mathbf{Q}_n C_n \cap L \neq \{\mathbf{0}\}\} \geq 1 - \sum_{i=1}^n p_{C_i}(\lambda). \quad (8.2)$$

The function $p_C(\lambda)$ is defined in (7.2).

The proof of this result is the topic of Section 8.2. For now, let us demonstrate how Theorem 8.2 follows readily from this new result.

Proof of Theorem 8.1. An application of the forgetting property (3.5) shows that

$$\mathbb{P}\{C \cap \mathbf{Q}C' \neq \{\mathbf{0}\}\} = \mathbb{P}\{\mathbf{Q}_1 C \cap \mathbf{Q}_2 C' \neq \{\mathbf{0}\}\}$$

for i.i.d. random rotations \mathbf{Q}_1 and \mathbf{Q}_2 . Make the identifications

$$(C_1, C_2, L) \leftrightarrow (C, C', \mathbb{R}^d).$$

The implications (8.1) and (8.2) of Theorem 8.2 are then equivalent to the claims of Theorem 8.1. \square

8.2 Proof of Theorem 8.2

The proof requires two lemmas. The first is an interlacing inequality for the tail functionals of the intrinsic volumes.

Lemma 8.3 (Interlacing). *For every convex cone $C \in \mathcal{C}_d$ that is not a subspace, the inequalities*

$$h_k(C) \geq \frac{1}{2} t_k(C) \geq h_{k+1}(C) \quad (8.3)$$

hold for every $k = 0, \dots, d - 1, d$.

While the proof of Lemma 8.3, which appears in Section 8.2.1, requires only the Crofton formula and a basic geometric observation, the interlacing inequality is rather deep. Indeed, it was observed by Dennis Amelunxen (personal communication) that taking $k = 0$ readily yields the Gauss–Bonnet relationship (5.8). The second lemma lets us control the size of the tails of the intrinsic volumes of the product of cones by the tails of the constituent cones.

Lemma 8.4 (Tail functional bound). *With notation as in Theorem 8.2, we have*

$$t_{\lceil \Delta + m + n\lambda \rceil}(C_1 \times \cdots \times C_{n-1} \times C_n \times L) \leq \sum_{i=1}^n t_{\lceil \delta(C_i) + \lambda \rceil}(C_i). \quad (8.4)$$

The proof utilizes a simple probabilistic argument; see Section 8.2.1. With these two results in hand, we are now in a position to prove the approximate kinematic formula.

Proof of Theorem 8.2. We assume that at least one of the cones C_i is not a subspace so that an application of the Crofton formula (5.22) is justified. At the end, we dispense with the trivial case where all C_i are subspaces.

Suppose $\Delta + m \leq n(d - \lambda)$. By the forgetting property (3.5), the probability in (8.1) is unchanged if we replace L with the randomly oriented subspace QL . The iterated

Crofton formula (5.22) gives the probability

$$\begin{aligned} \mathbb{P}\{\mathbf{Q}_1 C_1 \cap \dots \cap \mathbf{Q}_{n-1} C_{n-1} \cap \mathbf{Q}_n C_n \cap \mathbf{Q}L \neq \{\mathbf{0}\}\} &= 2h_{nd+1}(C_1 \times \dots \times C_n \times L) \\ &\leq t_{nd}(C_1 \times \dots \times C_n \times L), \end{aligned}$$

where the upper bound follows from the interlacing inequality (8.3). By our assumption on the total statistical dimension Δ , we have $nd \geq \lceil \Delta + m + n\lambda \rceil$, and so

$$t_{nd}(C_1 \times \dots \times C_n \times L) \leq t_{\lceil \Delta + m + n\lambda \rceil}(C_1 \times \dots \times C_n \times L) \leq \sum_{i=0}^n t_{\lceil \delta(C) + \lambda \rceil}(C_i),$$

where the first inequality follows by positivity of the intrinsic volumes and the second follows from Lemma 8.4. We complete the proof of the implication (8.1) with an application of Theorem 7.1:

$$t_{\lceil \delta(C) + \lambda \rceil}(C_i) \leq p_{C_i}(\lambda) \quad \text{for each } i = 1, \dots, n-1, n.$$

The second implication (8.2) involves a dual version of the preceding argument. Suppose $\Delta + m \geq n(d + \lambda)$. As above, we replace L with $\mathbf{Q}L$, where \mathbf{Q} is a random rotation independent of everything else. Then the conic Crofton formula (5.17) gives

$$\begin{aligned} \mathbb{P}\{C_1 \cap \mathbf{Q}_2 C_2 \cap \dots \cap \mathbf{Q}_n C_n \cap \mathbf{Q}L \neq \{\mathbf{0}\}\} &= 2h_{nd+1}(C_1 \times \dots \times C_n \times L) \\ &\geq t_{nd+1}(C_1 \times \dots \times C_n \times L), \end{aligned}$$

where the final inequality is the interlacing result (8.3). Since the intrinsic volumes sum to one (5.4), we have

$$\begin{aligned} t_{nd+1}(C_1 \times \dots \times C_n \times L) &= 1 - \sum_{k=0}^{nd} v_{(n+1)d-k}(C_1 \times \dots \times C_n \times L) \\ &= 1 - \sum_{k=d}^{(n+1)d} v_k(C_1^\circ \times \dots \times C_n^\circ \times L^\perp) = 1 - t_d(C_1^\circ \times \dots \times C_n^\circ \times L^\perp) \end{aligned}$$

where the second equality follows from the polarity rule (5.6) and the last relation is the definition (5.9) of the tail functional. Since $\lceil nd - \Delta + d - m + \lambda \rceil \geq d$ by the assumption

on $\Delta + m$, we have the following lower bound for the final term above:

$$1 - t_d(C_1^\circ \times \cdots \times C_n^\circ \times L^\perp) \geq 1 - t_{\lceil nd - \Delta + d - m + n\lambda \rceil}(C_1^\circ \times \cdots \times C_n^\circ \times L^\perp). \quad (8.5)$$

The statistical dimension adds under direct products (Proposition 6.2.5), so we may express the statistical dimension of the product cone in terms of Δ :

$$\delta(C_1^\circ \times \cdots \times C_n^\circ \times L^\perp) = \sum_{i=1}^n \delta(C_i^\circ) + \delta(L) = (nd - \Delta) + (d - m),$$

where the second inequality follows by the polarity law (Proposition 6.2.4). Applying Lemma 8.4, we reach the bound

$$t_{\lceil nd - \Delta + (d - m) + n\lambda \rceil}(C_1^\circ \times \cdots \times C_n^\circ \times L^\perp) \leq \sum_{i=1}^n t_{\lceil \delta(C_i^\circ) + \lambda \rceil}(C_i^\circ) \leq \sum_{i=1}^n p_{C_i^\circ}(\lambda).$$

The first inequality follows from the tail bound of Lemma 8.4. The final bound is a consequence of the concentration of the intrinsic volumes, Theorem 7.1. Combining this inequality with (8.5) and the fact that $p_{C^\circ}(\lambda) = p_C(\lambda)$ yields the second implication (8.2).

(Degenerate case.) We now return to the outstanding case where all of the cones C_i are linear subspaces, then the claim is a consequence of the fact that the intersection of randomly oriented subspaces L, L' has dimension $\dim(L \cap QL') = \max\{0, \dim(L) + \dim(L') - d\}$ almost surely. Therefore, if $\Delta + m \leq nd$, the probability in (8.1) is zero; otherwise, it is equal to one. \square

8.2.1 Proof of technical lemmas

We now establish the technical lemmas required for the approximate kinematic formulas of Theorem 8.1.

Proof of Lemma 8.3. Assume first that $k < d$. Let $L_{d-k} \subset \mathbb{R}^d$ be a $(d - k - 1)$ -dimensional subspace and let L_{d-k-1} be a $(d - k - 1)$ -dimensional subspace contained in L_{d-k} . By

the Crofton formula (5.17), we have

$$\begin{aligned} h_k(C) &= \mathbb{P}\{C \cap \mathbf{Q}L_{d-k-1} \neq \{\mathbf{0}\}\} \\ &\geq \mathbb{P}\{C \cap \mathbf{Q}L_{d-k} \neq \{\mathbf{0}\}\} = h_{k+1}(C), \end{aligned}$$

which shows that the left-hand side of (8.3) is greater than the right-hand side. The tail function $t_k(C)$ is the average of half-tail functionals:

$$\frac{1}{2}t_k(C) = \frac{1}{2}[h_k(C) + h_{k+1}(C)].$$

Since a midpoint lies between its endpoints, we see $h_k(C) \geq \frac{1}{2}t_k(C) \geq h_{k+1}(C)$. Therefore (8.3) holds for all $k < d$. For $k = d$, relation (8.3) is equivalent to the string $v_d(C) \geq \frac{1}{2}v_d(C) \geq 0$, which holds by positivity of the intrinsic volumes. \square

The second lemma provides a bound on the tail functionals of a product cones by the sum of bounds of the constituent cones.

Proof of Lemma 8.4. Let L be an m -dimensional subspace of \mathbb{R}^d . For any $C \in \mathcal{C}_{d'}$ and index $k = 0, \dots, d' - 1, d'$, we have the equality

$$v_{k+m}(C \times L) = \sum_{i+j=k+m} v_i(C)v_j(L) = v_k(C)$$

by the product formula (5.7) and the expression (5.2) for the intrinsic volumes of subspaces. Since m is an integer, we find

$$t_{\lceil \Delta+m+n\lambda \rceil}(C_1 \times \cdots \times C_n \times L) = t_{\lceil \Delta+n\lambda \rceil}(C_1 \times \cdots \times C_n) \quad (8.6)$$

We use a probabilistic method to bound the second tail functional above. Let $\{X_i\}_{i=1}^n$ be independent integer-valued random variables with probability distribution defined by the sequence of intrinsic volumes:

$$\mathbb{P}\{X_i = k\} = v_k(C_i).$$

The probability distribution of the sum $Y := \sum_{i=1}^n X_i$ is the convolution of the distributions of the set of random variables $\{X_i\}_{i=1}^n$. Since the convolution of intrinsic volume sequences is equal to the intrinsic volume sequence of the product cone (5.7), we see

$$\mathbb{P}\{Y = k\} = v_k(C_1 \times \cdots \times C_{n-1} \times C_n).$$

Therefore

$$\mathbb{P}\{Y \geq \Delta + n\lambda\} \leq \mathbb{P}\{X_i \geq \delta(C_i) + \lambda \text{ for some } i\} \leq \sum_{i=1}^n \mathbb{P}\{X_i \geq \delta(C_i) + \lambda\}.$$

The first inequality holds because the first event implies the second, while the second inequality is the union bound. Writing the inequality above in terms of the distributions of Y and X_i yields

$$t_{\lceil \Delta + n\lambda \rceil}(C_1 \times \cdots \times C_n) \leq \sum_{i=1}^n t_{\lceil \delta(C_i) + \lambda \rceil}(C_i).$$

The result follows upon combining this bound with (8.6). □

Chapter 9

Success and failure of demixing

This chapter demonstrates that recovery bounds for the demixing method are controlled by the total statistical dimension of all descent cones under the random alignment model. We introduce our result in the context of the two-signal demixing framework introduced in Section 1.2.2 under the random model for incoherence given in Section 1.2.3. Let us review the setup. We observe the superposition

$$\mathbf{z}_0 = \mathbf{x}^\natural + \mathbf{Q}\mathbf{y}^\natural, \quad (9.1)$$

where \mathbf{Q} is a random, but known, matrix that models the incoherence between \mathbf{x}^\natural and \mathbf{y}^\natural . Given structure-inducing convex functions f and g and the side information $\alpha := g(\mathbf{y}^\natural)$, we demix the observation by solving

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } g(\mathbf{y}) \leq \alpha \quad \text{and} \quad \mathbf{x} + \mathbf{Q}\mathbf{y} = \mathbf{z}_0. \quad (9.2)$$

This approach succeeds if and only if the pair $(\mathbf{x}^\natural, \mathbf{y}^\natural)$ is the unique optimal solution to (9.2). For this simple demixing model, our main result tells us that demixing succeeds with high probability, so long as the ambient dimension exceeds the total statistical dimension of the corresponding descent cones by a small amount. (A similar result is available in the joint work [ALMT13, Thm. III].)

Theorem 9.1 (Recovery bounds for standard demixing). *Fix a probability tolerance $\eta \in (0, 1)$, and suppose the observation \mathbf{z}_0 is drawn from the random alignment model (9.1).*

Define the maximum transition width

$$\omega_v := \omega(\mathcal{D}(f, \mathbf{x}^h)) \vee \omega(\mathcal{D}(g, \mathbf{y}^h)), \quad (9.3)$$

where the transition width $\omega(\cdot)$ is defined in (7.1). Define the constant $R_\eta := \log(8/\eta)$ and the bandwidth

$$\beta := 16R_\eta + \sqrt{8R_\eta} \omega_v.$$

Then

$$\delta(\mathcal{D}(f, \mathbf{x}^h)) + \delta(\mathcal{D}(g, \mathbf{y}^h)) \leq d - \beta \quad \implies \quad (1.6) \text{ succeeds with probability } \geq 1 - \eta;$$

$$\delta(\mathcal{D}(f, \mathbf{x}^h)) + \delta(\mathcal{D}(g, \mathbf{y}^h)) \geq d + \beta \quad \implies \quad (1.6) \text{ succeeds with probability } \leq \eta.$$

The proof of Theorem 9.1 appears in Section 9.1, where it is derived as a corollary of a general result for multiple demixing. Theorem 9.1 explains a number of qualitative features of demixing problems, including the width and taper of the transition region between success and failure (cf. Figure 1.1). We now discuss some general features of this result; applications to specific demixing models appear in concert with numerical experiments in Section 10.

The most important parameter that appears in the theorem above is the sum of the statistical dimensions. When the total statistical dimension of the descent cones is a little less than the ambient dimension, demixing succeeds. On the other hand, if the sum of the statistical dimensions is a little larger than the ambient dimension, demixing fails.

The definition of “a little” is encoded in the bandwidth parameter β . The bandwidth β depends logarithmically on the probability tolerance η and linearly on the maximum transition width ω_v . Because the transition width is no larger than $\sqrt{d/2}$, the transition width grows no faster than

$$\beta = O\left(\log\left(\frac{1}{\eta}\right)\sqrt{d}\right) \quad \text{as } \eta \rightarrow 0, \quad d \rightarrow \infty.$$

Thus, for moderate values of η , the bandwidth is much smaller than the ambient dimension d . This is sufficient to explain the phase transition observed in many demixing problems.

In some cases, however, the approximation $\omega_v = O(\sqrt{d})$ is too coarse because the maximum transition width is much smaller than \sqrt{d} . This occurs in particular when statistical dimension is very small or very close to the ambient dimension d . In these cases, the phase transition between a high probability of success and failure occurs over a very narrow region. This observation explains the taper seen in the experiment where we demix sparse vectors (Figure 1.2).

Finally, we note that extraordinarily small values of η are useful for achieving stronger guarantees for demixing. See Sections 10.1 and 10.2 for examples.

9.1 Transition for extended demixing

The result mentioned above for the demixing of two signals is a consequence of a more general result that addresses the multiple demixing scenario (1.7) as well as the compressed demixing model (1.10). In this scenario, we assume that the observation $\mathbf{z}_0 \in \mathbb{R}^m$ is given by the sum

$$\mathbf{z}_0 = A \left(\sum_{i=1}^n \mathbf{Q}_i \mathbf{x}_i^{\natural} \right), \quad (9.4)$$

where $A \in \mathbb{R}^{m \times d}$ is a measurement matrix with full row rank¹ and the \mathbf{Q}_i are i.i.d. random rotations. Given convex functions f_i for $i = 1, \dots, d-1, d$ and the side information $\alpha_i := f_i(\mathbf{x}_i^{\natural})$, we may demix the observation (9.4) by solving the problem

$$\begin{aligned} & \text{minimize} && f_1(\mathbf{x}_1) \\ & \text{subject to} && f_i(\mathbf{x}_i) \leq \alpha_i \quad \text{for } i = 2, \dots, n-1, n, \quad \text{and} \quad \mathbf{z}_0 = A \left(\sum_{i=1}^n \mathbf{Q}_i \mathbf{x}_i \right). \end{aligned} \quad (9.5)$$

Section 4.3 outlines the geometric optimality conditions for this model when the matrices \mathbf{Q}_i are fixed. Our goal now is to understand the statistical properties of

¹We may reduce an arbitrary $A \in \mathbb{R}^{m \times d}$ to one with full row rank by eliminating redundant rows. In this scenario, Theorem 9.2 still holds if we interpret m as the row rank of the measurement operator.

the compressed demixing model under the random incoherence model (9.4). The following result shows that the probability that the demixing method (9.5) succeeds is characterized by the statistical dimension certain descent cones.

Theorem 9.2. *Fix a probability tolerance $\eta \in (0, 1)$. With the same notation as above, define the width parameter*

$$\omega_{\vee} := \max_{i=1, \dots, n-1, n} \left\{ \omega \left(\mathcal{D}(f_i, \mathbf{x}_i^{\natural}) \right) \right\},$$

the total statistical dimension $\Delta := \sum_{i=1}^n \delta(\mathcal{D}(f_i, \mathbf{x}_i^{\natural}))$, and the parameter $R := \log(4n/\eta)$.

Then

$$m \geq \Delta + n(8R + \sqrt{2R}\omega_{\vee}) \implies (9.5) \text{ succeeds with probability } \geq 1 - \eta; \quad (9.6)$$

$$m \leq \Delta - n(8R + \sqrt{2R}\omega_{\vee}) \implies (9.5) \text{ succeeds with probability } \leq \eta. \quad (9.7)$$

In other words, when the number of measurements slightly exceeds the statistical dimension, the demixing procedure (9.5) succeeds with high probability, while when the number of measurements is slightly less than the total statistical dimension, the demixing method (9.5) fails with high probability. Thus, a *phase transition* between success and failure exists for compressed demixing problems.

Let us discuss the width phase transition given by Theorem 9.2. The parameter R grows logarithmically in the number of cones n and the probability tolerance η^{-1} , so the width of the transition between success and failure is not much larger than $n\omega_{\vee}$ for moderate values of η and n :

$$n(8R + \sqrt{2R}\omega_{\vee}) = O\left(n\omega_{\vee} \log(n\eta^{-1})\right) \quad \text{as } n \rightarrow \infty, \quad \eta \rightarrow 0.$$

The linear dependence of the transition width and the number of cones is suboptimal. Variance bounds suggest that the transition width should be controlled by the root-mean-square $(\sum \omega^2(\mathcal{D}(f_i, \mathbf{x}_i^{\natural})))^{1/2}$. We expect that a refined bound is available that shows that the transition width is on the order of $\sqrt{n}\omega_{\vee}$. In most applications, however, the

number of constituents n is relatively small, so such a refinement does not give much practical advantage.

Most of the hard work involved in proving Theorem 9.2 has already been completed. Let us sketch the remaining argument in a few sentences. We start with the basic optimality condition (4.9) for compressed demixing. While this condition involves $n + 1$ independent conditions, dualizing the optimality conditions yields—up to a set of measure zero—a single condition for the intersection of randomly rotated convex cones (Proposition 5.18). Then it is just a matter of applying the iterated approximate kinematic formula of Section 8.1 and checking that our transition width parameter ω_V results in the claimed probability estimates.

The full details of this reduction are in Section 9.2. Let us first show how Theorem 9.2 implies Theorem 9.1.

Proof of Theorem 9.1 from Theorem 9.2. Theorem 9.1 is essentially the $n = 2$ and $m = d$ case of Theorem 9.2. Let us describe how the statistics of the demixing method (9.2) can be recast to match this case.

Define the rotated observation $\hat{\mathbf{z}}_0 := \mathbf{Q}_* \mathbf{z}_0$, where \mathbf{Q}_* is a random orthogonal matrix independent of \mathbf{Q} . Then (9.2) succeeds for a given \mathbf{Q} if and only if

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } g(\mathbf{y}) \leq \alpha \quad \text{and} \quad \mathbf{Q}_*(\mathbf{x} + \mathbf{Q}\mathbf{y}) = \hat{\mathbf{z}}_0 \quad (9.8)$$

succeeds for *every* possible $\mathbf{Q}_* \in O_d$ because the equality constraints are equivalent. Applying independence of \mathbf{Q} and \mathbf{Q}_* , we see

$$\begin{aligned} \mathbb{P}\{(9.2) \text{ succeeds at demixing } \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}\} \\ = \mathbb{P}\{(9.8) \text{ succeeds at demixing } \mathbf{Q}_*(\mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural})\}. \end{aligned}$$

By the forgetting property (3.5), the probability on the right-hand side above is equal to the probability that

$$\text{minimize } f(\mathbf{x}) \quad \text{subject to } g(\mathbf{y}) \leq \alpha \quad \text{and} \quad \mathbf{Q}_1 \mathbf{x} + \mathbf{Q}_2 \mathbf{y} = \hat{\hat{\mathbf{z}}}_0,$$

where $\hat{\mathbf{z}}_0 := \mathbf{Q}_1 \mathbf{x}^{\natural} + \mathbf{Q}_2 \mathbf{y}^{\natural}$ for two i.i.d. random rotations \mathbf{Q}_1 and \mathbf{Q}_2 . With the identifications

$$(f, g, \mathbf{I}) \leftrightarrow (f_1, f_2, \mathbf{A}),$$

the program above is a special case of (9.5) with $n = 2$ constituents and $m = d$ measurements. The conclusion follows immediately from Theorem 9.2. \square

A final remark is in order before we move on to the proof of Theorem 9.2.

Remark 9.3. While our signal model (9.4) assumes a fixed matrix \mathbf{A} , the proof of Theorem 9.2 reveals that the only relevant modeling assumption is that the constituents \mathbf{x}^{\natural} and the nullspace $\text{null}(\mathbf{A})$ are oriented randomly relative to one another before being combined into the observation \mathbf{z}_0 . Theorem 9.2 holds just as well when we fix $\mathbf{Q}_1 = \mathbf{I}$ but take $\mathbf{A} = \mathbf{G} \in \mathbb{R}^{m \times d}$ where \mathbf{G} is a Gaussian matrix. In other words, our demixing results hold for the observation model

$$\mathbf{z}_0 = \mathbf{G} \left(\mathbf{x}_1^{\natural} + \sum_{i=2}^n \mathbf{Q}_i \mathbf{x}_i^{\natural} \right),$$

so long as the equality restriction in (9.5) is updated to reflect this modified observation model.

9.2 Proof of Theorem 9.2

We start by setting up some notation. Define the cones $C_i := \mathcal{D}(f_i, \mathbf{x}_i^{\natural})$ for $i = 1, \dots, n - 1, n$ and $C_{n+1} := \text{null}(\mathbf{A})$, and let \mathbf{Q}_{n+1} be another rotation independent of all the others. Define the total statistical dimension of the polars

$$\Delta^\circ := \sum_{i=1}^n \delta(C_i^\circ) = nd - \Delta, \quad (9.9)$$

where the second equality follows from the polarity relation for the statistical dimension (Proposition 6.2.4) and our definition of C_i . Throughout most of the proof, we assume that $C_i \neq \{\mathbf{0}\}$ for at least two indices $i \in \{1, \dots, n, n + 1\}$. This restriction arises from

our desire to apply Proposition 5.18, whose assumptions require two nonzero cones. We deal with the degenerate case where all but one of the cones is trivial at the end.

By Theorem 4.3, compressed demixing succeeds with the same probability that

$$\mathbf{Q}_i C_i \cap - \left(\sum_{j \neq i} \mathbf{Q}_j C_j \right) = \{\mathbf{0}\} \quad \text{for all } i = 1, \dots, n, n+1 \quad (9.10)$$

where we are justified in applying an independent random rotation \mathbf{Q}_{n+1} to the cone C_{n+1} by the forgetting property (3.5) and rotational invariance of the condition (9.10).

For any i such that $C_i = \{\mathbf{0}\}$, the intersection in (9.10) is always equal to $\{\mathbf{0}\}$, so that we can eliminate these cases from consideration. Our assumption that at least two of the cones do not equal $\{\mathbf{0}\}$ ensures that

$$\sum_{j \neq i} \mathbf{Q}_j C_j \neq \{\mathbf{0}\} \quad \text{for all } i = 1, \dots, n, n+1.$$

Thus, for every i such that $C_i \neq \{\mathbf{0}\}$, we may apply Proposition 5.18 to find that the events

$$\left\{ \mathbf{Q}_i C_i \cap - \left(\sum_{j \neq i} \mathbf{Q}_j C_j \right) = \{\mathbf{0}\} \right\} \quad \text{and} \quad \left\{ \mathbf{Q}_1^t C_1^\circ \cap \dots \cap \mathbf{Q}_n^t C_n^\circ \neq \{\mathbf{0}\} \right\} \quad (9.11)$$

coincide except on a set of measure zero. Because a finite union of sets of measure zero still has measure zero, we have

$$\begin{aligned} \mathbb{P}\{(9.5) \text{ succeeds}\} &= \mathbb{P}\left\{ \mathbf{Q}_i C_i \cap - \left(\sum_{j \neq i} \mathbf{Q}_j C_j \right) = \{\mathbf{0}\} \text{ for all } i \text{ such that } C_i \neq \{\mathbf{0}\} \right\} \\ &= \mathbb{P}\left\{ \mathbf{Q}_1^t C_1^\circ \cap \dots \cap \mathbf{Q}_{n+1}^t C_{n+1}^\circ \neq \{\mathbf{0}\} \right\} = \mathbb{P}\left\{ \mathbf{Q}_1^t C_1^\circ \cap \dots \cap \mathbf{Q}_n^t C_n^\circ \cap C_{n+1}^\circ \neq \{\mathbf{0}\} \right\}, \end{aligned} \quad (9.12)$$

where the final inequality follows, as usual, from the forgetting property (3.5). This equality is ripe for an application of the approximate kinematic formula developed in Section 8.1.

Consider the case (9.6) where $m \geq \Delta + 8R + \sqrt{2R}\omega_\vee$. By our definition (9.9) of Δ° , we have

$$\Delta^\circ + m \geq n(d + \lambda_*),$$

where the parameter λ_* is given by

$$\lambda_* := 8R + \sqrt{2R}\omega_\nu \quad (9.13)$$

The cone $C_{n+1}^\circ = \text{null}(\mathbf{A})^\perp$ is a subspace of dimension m , so the approximate kinematic formula (8.2) gives the lower bound

$$\mathbb{P}\{\mathbf{Q}_1^t C_1^\circ \cap \cdots \cap \mathbf{Q}_n C_n^\circ \cap C_{n+1}^\circ \neq \{\mathbf{0}\}\} \geq 1 - \sum_{i=1}^n p_{C_i}(\lambda_*) \geq 1 - n \max_{i=1, \dots, n-1, n} p_{C_i}(\lambda_*). \quad (9.14)$$

The following lemma provides a bound on the maximum in (9.14).

Lemma 9.4. *With the notation from Theorem 9.2 and with λ_* defined in (9.13), we have*

$$\max_{i=1, \dots, n-1, n} p_{C_i}(\lambda_*) \leq \frac{\eta}{n} \quad \text{for each } i = 1, \dots, n-1, n.$$

This lemma requires only a straightforward concavity bound; the proof appears at the end of this section.

By Lemma 9.4, we see that the right-hand side of (9.14) is no smaller than $1 - \eta$, and hence the probability that (9.5) succeeds is at least $1 - \eta$ by (9.12). This proves the first implication (9.6).

For the second claim (9.7), suppose that $m \leq \Delta - (8R + \sqrt{2R}\omega_\nu)$. Then

$$\Delta^\circ + m \leq n(d - \lambda_*),$$

so applying the upper bound (8.1) from the approximate kinematic formula yields the inequality

$$\mathbb{P}\{\mathbf{Q}_1^t C_1^\circ \cap \cdots \cap \mathbf{Q}_n C_n^\circ \cap C_{n+1}^\circ\} \leq \sum_{i=1}^n p_{C_i}(\lambda_*) \leq n \max_{i=1, \dots, n-1, n} p_{C_i}(\lambda_*) \leq \eta.$$

The final inequality follows from Lemma 9.4. The conclusion (9.7) follows when we combine the estimate above with the expression (9.12) for the probability that demixing succeeds.

(*The degenerate case*) In the degenerate case, all but one of the cones C_i is equal to the trivial cone $\{\mathbf{0}\}$. In other words, there is an index $i_* \in \{1, \dots, n, n-1\}$ such that

$$C_i = \{\mathbf{0}\} \quad \text{for all } i \neq i_*.$$

This implies that

$$\mathbf{Q}_i C_i = \{\mathbf{0}\} \quad \text{or} \quad \sum_{j \neq i} \mathbf{Q}_j C_j = \{\mathbf{0}\}$$

for every $i = 1, \dots, n, n+1$. In particular, the optimality condition (9.10) always holds, so that the demixing method (9.5) always succeeds.

Since demixing succeeds with probability one when in this degenerate setting, we only need to check that the left-hand side of the implication (9.7) never holds. If $i_* = n+1$, then we must have $C_i = \{\mathbf{0}\}$ for all $i = 1, \dots, n-1, n$. In particular, the total statistical dimension $\Delta = 0$. Since the integer $m \geq 0$, we have $m > \Delta - n(8R + \sqrt{2R}\omega_\vee)$, so the left-hand side of the implication (9.7) does not hold.

On the other hand, if $i_* \neq n+1$, then $C_{n+1} = \text{null}(\mathbf{A}) = \{\mathbf{0}\}$. Since $\mathbf{A} \in \mathbb{R}^{m \times d}$ has full row rank by assumption, a trivial nullspace implies that $m = d$. But we also have

$$\Delta = \sum_{i=1}^n \delta(C_i) = \max_{i=1, \dots, n-1, n} \delta(C_i) = \delta(C_{i_*}) \leq d.$$

Combining these observations, we find the inequality

$$m = d > \Delta - n(8R + \sqrt{2R}\omega_\vee),$$

and so again the left-hand side of the implication (9.7) does not hold. \square

Proof of Lemma 9.4. By solving a quadratic equation, we see that

$$p_{C_i}(\lambda_i) := 4 \exp\left(-\frac{\lambda_i^2/8}{\omega(C_i)^2 + \lambda_i}\right) = \frac{\eta}{n} \iff \lambda_i := 4R + \sqrt{16R^2 + 2R\omega^2(C_i)},$$

Since $p_{C_i}(\cdot)$ is a decreasing function of $\lambda \geq 0$, we only need argue $\lambda_i \leq \lambda_*$ for each $i = 1, \dots, n-1, n$. This follows immediately from the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for

$a, b \geq 0$:

$$\lambda_i \leq 8R + \sqrt{2R\omega^2(C_i)} \leq 8R + \sqrt{2R}\omega_v = \lambda_*$$

The final inequality follows from the fact that $\omega(C_i) \leq \omega_v$.

□

Chapter 10

Applications and numerical examples

This chapter provides concrete applications of the theory developed earlier. We first tackle the most classical problem, demixing two sparse signals, in Section 10.1. We consider strong guarantees for this sparse + sparse model in Section 10.1.1 and an undersampled variant in Section 10.1.2. We discuss demixing sparse and sign vectors in Section 10.2 as a model for a secure communications protocol that is robust to sparse errors. Section 10.3 describes demixing a low-rank matrix from a matrix that is sparse in a random basis; this is a stylized model for data that appears in the fields of machine learning and robust statistics. We close by considering an application of our theory to regularized linear inverse problems in Section 10.4. In every case, we find that the statistical dimension accurately predicts the phase transition observed in our numerical experiments.

10.1 Sparse + sparse

Demixing two sparse vectors is the foundational problem in demixing. As described in the introduction, this model appears in morphological component analysis (MCA), where the nonzero components in the vectors indicate the presence of certain features in the signal. Numerous other applications for demixing two sparse signals are described in [SKPB12].

Assume that our observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural} \in \mathbb{R}^d$ is the sum of a sparse signal \mathbf{x}^{\natural} and

a randomly rotated sparse signal $\mathbf{Q}\mathbf{y}^{\natural}$. For concreteness, we define the sparsity levels

$$s_x := \text{nnz}(\mathbf{x}^{\natural}) \quad \text{and} \quad s_y := \text{nnz}(\mathbf{y}^{\natural}),$$

where $\text{nnz}(\cdot)$ returns the number of nonzero elements of a vector. The atomic gauge associated to a sparse signal is the ℓ_1 norm, so we demix the observation to the convex program

$$\text{minimize} \quad \|\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\mathbf{y}\|_{\ell_1} \leq \alpha \quad \text{and} \quad \mathbf{x} + \mathbf{Q}\mathbf{y} = \mathbf{z}_0, \quad (10.1)$$

where $\alpha := \|\mathbf{y}^{\natural}\|_{\ell_1}$ is known side information.

Let us discuss how our theory provides a way to compute the location of the phase transition of the demixing method (10.1) under the random orientation model. Theorem 9.1 indicates that (10.1) succeeds with high probability when

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural})) \lesssim d,$$

while it fails with high probability when the inequality above is reversed. Proposition 6.14 provides an approximate formula for the statistical dimension of the descent cones by the identification $d\psi_{\ell_1}(s_x/d) \approx \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural}))$. Hence, we expect that a phase transition between success and failure of (10.1) occurs in the region where the sparsity levels s_x and s_y satisfy

$$\psi_{\ell_1}\left(\frac{s_x}{d}\right) + \psi_{\ell_1}\left(\frac{s_y}{d}\right) = 1. \quad (10.2)$$

The following result provides a rigorous statement of this intuitive argument.

Theorem 10.1 (Phase transitions in sparse + sparse demixing). *Consider the implicit equation*

$$\psi_{\ell_1}(\rho_x) + \psi_{\ell_1}(\rho_y) = 1 \quad (10.3)$$

where ψ_{ℓ_1} is defined in Proposition 6.14. Then:

1. (Well-defined curve) For each $\rho_x \in (0, 1)$, there exists a unique $\rho_y \in (0, 1)$ that

satisfies (10.3). This value of ρ_y is a continuous function of ρ_x , and moreover

$$\rho_y \rightarrow 1 \text{ as } \rho_x \rightarrow 0 \text{ and } \rho_y \rightarrow 0 \text{ as } \rho_x \rightarrow 1. \quad (10.4)$$

2. (Phase transition) Fix a probability tolerance $\eta \in (0, 1)$. For each pair (ρ_x, ρ_y) satisfying (10.3), there exist parameters $C_x, C_y > 0$ and an integer $d_0 \in \mathbb{N}$ such that the following implications hold whenever $d \geq d_0$:

(a) If

$$\frac{s_x}{d} \leq \rho_x - \frac{C_x}{\sqrt{d}} \quad \text{and} \quad \frac{s_y}{d} \leq \rho_y - \frac{C_y}{\sqrt{d}}, \quad (10.5)$$

then (10.1) succeeds with probability at least $1 - \eta$.

(b) On the other hand, if

$$\frac{s_x}{d} \geq \rho_x + \frac{C_x}{\sqrt{d}} \quad \text{and} \quad \frac{s_y}{d} \geq \rho_y + \frac{C_y}{\sqrt{d}}, \quad (10.6)$$

then (10.1) succeeds with probability at most η .

The triple (C_x, C_y, d_0) depends only ρ_x , ρ_y , and η .

The proof of this result appears in Appendix B.1 below, and it requires little more than a Taylor series expansion of ψ_{ℓ_1} coupled with Theorem 9.1. The statement above may appear pedantic, it rigorously demonstrates two important points.

1. The phase transition occurs as a function of the sparsity levels, not just as a function of the statistical dimensions.
2. Since $\rho_y \rightarrow 1$ as $\rho_x \rightarrow 0$ (10.4), the demixing method (10.1) will succeed even when one of the constituents is nearly completely dense, so long as the other is sufficiently sparse. This result holds in the proportional growth regime, meaning that the sparsity of each constituents can be a nonnegligible fraction of the ambient dimension d .

The first point is completely new to the literature on demixing, although the joint work [MT12] provided strong numeric evidence that suggested that a sparsity-

parameterized phase transition occurs along the curve Ω . The second point was demonstrated under an alternative model by Wright & Ma [WM09], and confirmed under a model similar to our own random orientation model more recently by [NT13]. Neither of these works identifies a sharp boundary where this phenomenon occurs.

We demonstrate the accuracy of our theory with a numerical experiment. Figure 10.1 shows¹ the results of a numeric experiment for demixing sparse vectors in dimension $d = 100$. Further experimental details are in Appendix C.3. The black region denotes a high empirical probability of failure, while the light region indicates near-certain success. Three curves mark the empirical 95%, 50% and 5% success regions. Underneath these empirical curves, we plot the level set determined by the implicit equation (10.2). See Appendix C.1 for details on how we compute ψ_{ℓ_1} . The correspondence between the 50% empirical success curves and the theoretical level curve that determines the phase transition is remarkable.

Finally, we provide the promised proof of the Theorem 1.1 from Section 1.1.3. We first restate our claim in a rigorous asymptotic growth setting.

Theorem 10.2 (Success and failure of MCA). *There is a function $\psi_{\ell_1} : (0, 1) \rightarrow (0, 1)$ that generates a threshold curve $\Omega \subset (0, 1)^2$ given by*

$$\Omega := \{(\rho_x, \rho_y) : \psi_{\ell_1}(\rho_x) + \psi_{\ell_1}(\rho_y) = 1\}.$$

Suppose the sparsity levels s_x and s_y grow with d in such a way that $s_x/d \rightarrow \rho_x^ \in [0, 1)$ and $s_y/d \rightarrow \rho_y^* \in [0, 1)$ as $d \rightarrow \infty$. The curve Ω partitions the unit square into a success and failure region for MCA:*

(ρ_x^, ρ_y^*) lies strictly below $\Omega \implies$ (1.3) succeeds with probability $\rightarrow 1$; and*

(ρ_x^, ρ_y^*) lies strictly above $\Omega \implies$ (1.3) succeeds with probability $\rightarrow 0$,*

where the limits are taken as the ambient dimension $d \rightarrow \infty$.

¹This figure reproduces Figure 1.1 from the introduction. Similar figures also appear in the joint works [MT12, ALMT13].

Demixing sparse + sparse

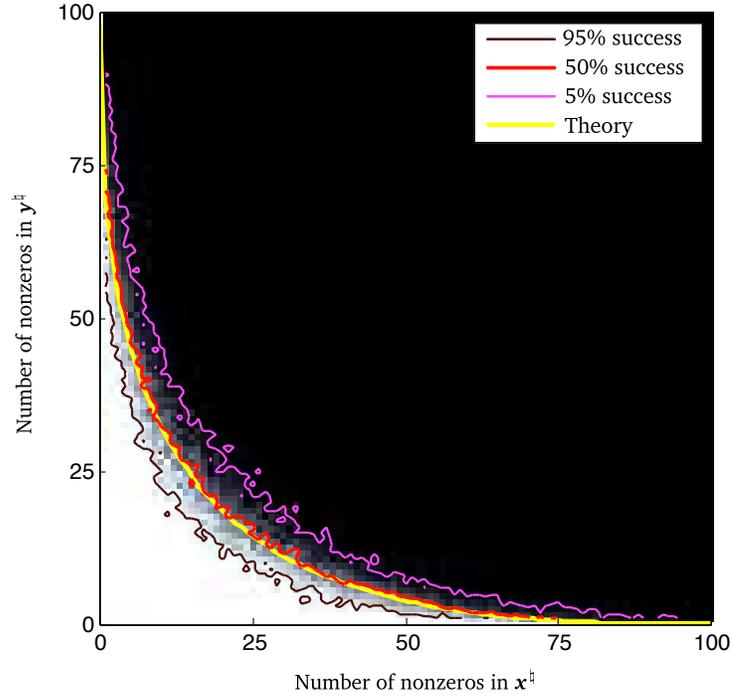


Figure 10.1: Empirical probability of success for demixing two sparse vectors. The ambient dimension is $d = 100$, and the sparsity levels vary from 1 to 100. The colormap presents the empirical probability that (10.1) succeeds over 25 trials. White pixels indicate complete success, the black region consists of total failures, and the gray transition region displays a mix of successes and failures. The figure also displays the 95%, 50% and 5% success isoclines. Our theory tells us that the location of the transition region occurs at the yellow curve that lies under the 50% isocline. (This figure also appears in Section 1.1.) See Appendix C.3 for the experimental details.

Proof of Theorem 10.2 from Theorem 10.1. Suppose that ρ_x^* and ρ_y^* lie strictly below the curve $\Omega := \{(\rho_x, \rho_y) : \psi_{\ell_1}(\rho_x) + \psi_{\ell_1}(\rho_y) = 1\}$. Then there exists a $(\rho_x, \rho_y) \in \Omega$ and an $\varepsilon > 0$ such that

$$\frac{s_x}{d} \leq \rho_x - \varepsilon \quad \text{and} \quad \frac{s_y}{d} \leq \rho_y - \varepsilon$$

for all d large enough. Hence, for any $\eta > 0$, the relation (10.5) holds for all large enough d . Therefore, Theorem 10.1 implies that (1.3) succeeds with probability at least $1 - \eta$ for all large enough d . The first limit then follows by taking $\eta \rightarrow 0$. The second limit is proved in the same manner. \square

10.1.1 Strong bounds

Let us address another situation considered in the literature on demixing. Suppose that we draw the rotation \mathbf{Q} at random, but fix it for all time. Is it possible to guarantee that, for every sufficiently sparse pair $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$, the demixing method (10.1) succeeds given the observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$? Answers to such questions are known as *strong bounds* in the compressed sensing literature [Don04, Don06b, DT09, XH12]. It turns out that our theory is capable of providing an affirmative answer.

Theorem 10.3 (Strong bounds for sparse + sparse demixing). *Choose a probability tolerance $\eta \in (0, 1)$ and draw a random rotation $\mathbf{Q} \in \mathcal{O}_d$. There exists a parameter $\rho_S > 0$ and integer d_0 , each depending only on η , such that for each $d \geq d_0$, the following holds with probability at least $(1 - \eta)$:*

The demixing program (10.1) with observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$ succeeds for every pair of vectors $\mathbf{x}^{\natural}, \mathbf{y}^{\natural} \in \mathbb{R}^d$ with $\text{nnz}(\mathbf{x}^{\natural}) \leq \rho_S d$ and $\text{nnz}(\mathbf{y}^{\natural}) \leq \rho_S d$.

The proof of the strong bound relies on a union bound over a large number of success characterizations of the form 4.1. The demonstration appears in Appendix B.1. At this point we make two remarks.

1. Our proof technique, coupled with a numerical computation of the bound (B.10), yields the value $\rho_S \approx 0.0036$ for very large d and fixed η . This value appears to be suboptimal by at least a factor of 10. In joint work [MT12], we showed numerically that this bound can be improved to approximately $\rho_S \approx 0.037$ for large d . Our prior work requires a detailed asymptotic analysis of the sequence $\{v_i(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural}))\}$ of intrinsic volumes in high dimensions. Even with this heavy lifting, this earlier work did not *prove* that the strong bound exists, but rather provided strong numerical evidence for its existence. The present work provides a more authoritative result, at the expense of a weaker estimate on the value of ρ_S .
2. Our proof technique does not inherently require that the probability of success η remain fixed as the dimension grows. A more careful analysis reveals that the probability tolerance can grow at any rate such that $\log(1/\eta) = o(d)$ as

$d \rightarrow \infty$. (This alternative assumption preserves asymptotic limit of the crucial estimate (B.7).)

10.1.2 Undersampled sparse + sparse

In a refinement of the sparse + sparse model considered above, we consider when it is possible to demix two sparse constituents \mathbf{x}^{\natural} and \mathbf{y}^{\natural} given an *undersampled* observation of their superposition:

$$\mathbf{z}_0 := \mathbf{G}(\mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}).$$

With the side information $\alpha := \|\mathbf{y}_0\|_{\ell_1}$, we solve the compressed demixing program

$$\text{minimize } \|\mathbf{x}\|_{\ell_1} \quad \text{subject to } \|\mathbf{y}\|_{\ell_1} \leq \alpha \quad \text{and} \quad \mathbf{G}(\mathbf{x} + \mathbf{Q}\mathbf{y}) = \mathbf{z}_0. \quad (10.7)$$

This approach succeeds if $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ is the unique optimal point of (10.7).

Before stating a rigorous guarantee for this method, we provide an informal discussion that shows how to compute the location of the phase transition for this problem. Denote the total statistical dimension $\Delta := \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural}))$, and let $s_x := \text{nnz}(\mathbf{x}^{\natural})$ and $s_y := \text{nnz}(\mathbf{y}^{\natural})$ be the sparsity levels of the constituent signals. By Theorem 9.2 (cf. Remark 9.3), we have the implications

$$\begin{aligned} \Delta \leq m - C\sqrt{d} &\implies (10.7) \text{ succeeds with high probability, while} \\ \Delta \geq m + C\sqrt{d} &\implies (10.7) \text{ fails with high probability.} \end{aligned} \quad (10.8)$$

Proposition 6.14 guarantees that we have the formula

$$\frac{1}{d}\Delta = \psi_{\ell_1}\left(\frac{s_x}{d}\right) + \psi_{\ell_1}\left(\frac{s_y}{d}\right) + o\left(\frac{1}{\sqrt{d}}\right),$$

where ψ_{ℓ_1} is defined in (6.22). Combining this approximation with (10.8), we find that a phase transition between success and failure of (10.7) occurs in the neighborhood of (s_x, s_y) satisfying

$$\psi_{\ell_1}\left(\frac{s_x}{d}\right) + \psi_{\ell_1}\left(\frac{s_y}{d}\right) = \frac{m}{d}. \quad (10.9)$$

This heuristic discussion encompasses most of the key features of the following rigorous result.

Theorem 10.4 (Phase transition for undersampled sparse + sparse). *For any $\mu \in (0, 1]$, there exist parameters $\rho_x, \rho_y \in (0, 1)$ that solve the implicit equation*

$$\psi_{\ell_1}(\rho_x) + \psi_{\ell_1}(\rho_y) = \mu. \quad (10.10)$$

Fix a probability tolerance $\eta > 0$. Let the number of measurements $m = \lceil \mu d \rceil$ for some $\mu \in (0, 1]$. For each pair (ρ_x, ρ_y) that satisfies (10.10), there exist parameters $C_x, C_y > 0$ and an integer $d_0 \in \mathbb{N}$ such the following implications hold whenever $d \geq d_0$:

1. *If*

$$\frac{s_x}{d} \leq \rho_x - \frac{C_x}{\sqrt{d}} \quad \text{and} \quad \frac{s_y}{d} \leq \rho_y - \frac{C_y}{\sqrt{d}} \quad (10.11)$$

then (10.7) succeeds with probability at least $1 - \eta$.

2. *On the other hand, if*

$$\frac{s_x}{d} \geq \rho_x + \frac{C_x}{\sqrt{d}} \quad \text{and} \quad \frac{s_y}{d} \geq \rho_y + \frac{C_y}{\sqrt{d}} \quad (10.12)$$

then (10.7) succeeds with probability at most η .

The triple (C_x, C_y, d_0) depends only on ρ_x, ρ_y and η .

The proof appears in Appendix B.1 below; the argument closely parallels the proof of Theorem 10.1. Again, we find that there is a phase transition in the success of undersampled demixing parameterized by the sparsity.

We illustrate the accuracy of this claim with a numerical experiment in dimension $d = 200$ (Figure 10.2). The colormaps display the empirical probability that (10.7) succeeds at demixing two sparse vectors for various numbers of measurements. The yellow curve marks the sparsity pairs (s_x, s_y) satisfying (10.9). The agreement with the 50% success contour (red) is very good even for the moderate dimensional value $d = 200$. A full description of the experiment is available in Appendix C.3.1.

Undersampled sparse + sparse demixing

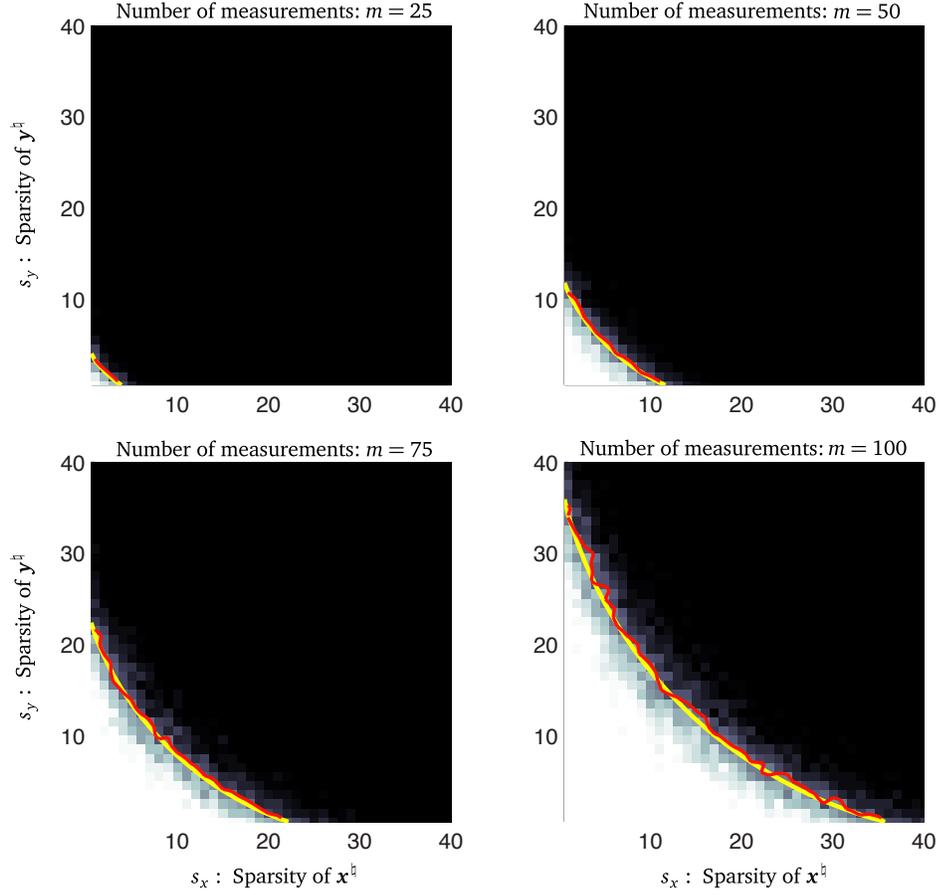


Figure 10.2: Empirical success of undersampled sparse + sparse demixing. The ambient dimension $d = 200$, and we demix two sparse signals \mathbf{x}^h and \mathbf{y}^h after taking $m = 25, 50, 75, 100$ random measurements with the demixing program (10.7). The light region denotes a high empirical probability of success, while the dark region denotes high empirical probability of failure. The red contour is the empirical 50% success line, which falls on top of the line $\psi_{\ell_1}(s_x/d) + \psi_{\ell_1}(s_y/d) = m/d$, where s_x and s_y are the number of nonzero elements of \mathbf{x}^h and \mathbf{y}^h . Further experimental details are available in Appendix C.3.1.

10.2 Sparse + sign

We now consider a demixing method that serves as a stylized protocol for a secure and robust communications scheme. Suppose we wish to securely transmit a d -bit message $\mathbf{y}^h \in \{\pm 1\}^d$ over a communications channel in the presence of sparse noise \mathbf{x}^h .

Wyner [Wyn79a, Wyn79b] proposed an intriguing transmission scheme that offers essentially perfect security. The transmitted signal is a random rotation $\mathbf{Q}\mathbf{y}^h$ of the message \mathbf{y}^h . Without the presence of noise, we may simply reconstruct the original

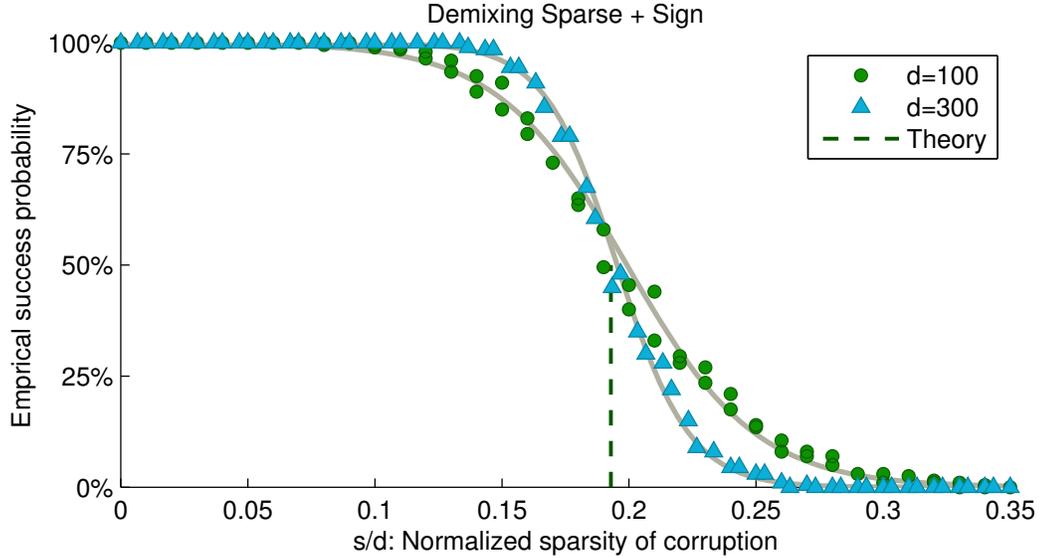


Figure 10.3: Demixing sparse + sign. The markers show the empirical probability of success for demixing a sparse vector from a sign vector. Circles are for dimension $d = 100$, while triangles denote dimension $d = 300$. The gray curves are the corresponding logistic fits. The horizontal axis gives the normalized sparsity s/d of \mathbf{x}^{\natural} . The dashed line occurs the point where $\psi_{\ell_1}(s/d) = \frac{1}{2}$, where ψ_{ℓ_1} is defined by (6.22).

message via the “matched filter” inversion $\hat{\mathbf{y}} = \mathbf{Q}^t(\mathbf{Q}\mathbf{y}^{\natural})$. However, the presence of corruptions eliminates the prospect of perfect reconstruction with this method.

If we know that the noise is sparse, however, demixing methods offer a way around this predicament. Model the corrupted observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$, where \mathbf{x}^{\natural} is an unknown sparse vector and $\mathbf{y}^{\natural} \in \{\pm 1\}^d$. We demix this observation by solving

$$\text{minimize } \|\mathbf{x}\|_{\ell_1} \quad \text{subject to } \|\mathbf{y}\|_{\ell_\infty} \leq 1 \quad \text{and} \quad \mathbf{x} + \mathbf{Q}\mathbf{y} = \mathbf{z}_0 \quad (10.13)$$

and declare success if $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural})$ forms the unique optimal point to (10.13).

Since we know *a priori* that $\|\mathbf{y}^{\natural}\|_{\ell_\infty} = 1$, this situation is ideally suited for an application of Theorem 9.1. The statistical dimension of the descent cone of the ℓ_∞ norm at \mathbf{y}^{\natural} is

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{y}^{\natural})) = \frac{d}{2}$$

by Proposition 6.11. By Proposition 6.14, the statistical dimension of the ℓ_1 norm at \mathbf{x}^{\natural}

is given by

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) = d \psi_{\ell_1}\left(\frac{s}{d}\right) + O(\sqrt{d}) \quad \text{as } d \rightarrow \infty,$$

where $s = \text{nnz}(\mathbf{x}^{\natural})$ is the sparsity of the corruption \mathbf{x}^{\natural} . Thus Theorem 9.1 suggests that demixing succeeds so long as

$$\psi_{\ell_1}\left(\frac{s}{d}\right) \leq \frac{1}{2} - \frac{C}{\sqrt{d}}$$

and fails when

$$\psi_{\ell_1}\left(\frac{s}{d}\right) \geq \frac{1}{2} + \frac{C}{\sqrt{d}},$$

where $C > 0$ is some universal constant. Again, we are able to make this heuristic argument completely rigorous.

Theorem 10.5 (Phase transition for sparse + sign demixing). *There exists a unique $\rho_* > 0$ that satisfies the stationary equation*

$$\psi_{\ell_1}(\rho_*) = \frac{1}{2}. \tag{10.14}$$

Fix a probability tolerance $\eta \in (0, 1)$. Then there is a parameter $C > 0$ and integer $d_0 \in \mathbb{N}$ such that for each $d \geq d_0$, the following implications hold:

1. *If*

$$\frac{\text{nnz}(\mathbf{x}^{\natural})}{d} \leq \rho_* - \frac{C}{\sqrt{d}}, \tag{10.15}$$

then (10.13) succeeds with probability at least $1 - \eta$.

2. *On the other hand, if*

$$\frac{\text{nnz}(\mathbf{x}^{\natural})}{d} \geq \rho_* + \frac{C}{\sqrt{d}},$$

then (10.13) succeeds with probability at most η .

The parameters C is a universal constant, and d_0 depends only on η .

The proof is very similar to the proof of the theorems in Section 10.1. The details appear in Appendix B.2.

Figure 10.3 displays the results of a numerical experiment that illustrates the accuracy of our theory in ambient dimension $d = 100$ and $d = 300$. Numerical details are available in Appendix C.4.

10.2.1 Strong bounds

We now describe how it is possible to achieve strong bounds for the sparse + sign demixing program of the form given in Section 10.1.1. In this case we ask the question: Is it possible to recover a binary codeword $\mathbf{y}^{\natural} \in \{\pm 1\}^d$ from the observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$ for *every* sufficiently sparse corruption \mathbf{x}^{\natural} ? This question arises when the sparse noise depends on the matrix \mathbf{Q} including adversarial corruptions and nonlinear phenomena such as clipped signals [AEJ⁺12]. Our theory is able to provide a rigorous guarantee even for this more restrictive setting.

Theorem 10.6 (Strong transition for sparse + sign demixing). *Choose a probability tolerance $\eta \in (0, 1)$ and draw a random rotation $\mathbf{Q} \in \mathcal{O}_d$. There exists a parameter $\rho_s > 0$ and integer $d_0 \in \mathbb{N}$ such that, for any $d \geq d_0$, the following holds with probability at least $(1 - \eta)$:*

Fix $\mathbf{y}^{\natural} \in \{\pm 1\}^d$. The demixing program (10.13) with observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural}$ succeeds for every sparse vector \mathbf{x}^{\natural} with $\text{nnz}(\mathbf{x}^{\natural}) \leq \rho_s d$.

The proof again is very similar to the strong bound for sparse + sparse demixing from Section 10.1.1 above. The details appear in Appendix B.2. Numerical computations based on our proof technique indicate that the parameter $\rho_s \approx 0.0034$ for fixed η and very large d . Again, this bound appears suboptimal when compared against the numerical calculation in the joint work [MT12], where it was found that $\rho_s \approx 0.018$. However, our current result provides a rigorous guarantee that the previous work was unable to achieve.

10.2.2 Sparse + sparse + sign

What if we add another sparse vector into the mix? That is, suppose our observation is given by

$$\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}_y \mathbf{y}^{\natural} + \mathbf{Q}_w \mathbf{w}^{\natural},$$

where \mathbf{x}^{\natural} and \mathbf{y}^{\natural} are sparse vectors and $\mathbf{w}^{\natural} \in \{\pm 1\}^d$ is a sign vector. Given the side information $\alpha := \|\mathbf{y}^{\natural}\|_{\ell_1}$, we demix our observation by solving the multiple demixing problem

$$\begin{aligned} & \text{minimize} && \|\mathbf{x}\|_{\ell_1} \\ & \text{subject to} && \begin{cases} \|\mathbf{y}\|_{\ell_1} \leq \alpha, \\ \|\mathbf{w}\|_{\ell_\infty} \leq 1, \quad \text{and} \\ \mathbf{x} + \mathbf{Q}_y \mathbf{y} + \mathbf{Q}_w \mathbf{w} = \mathbf{z}_0. \end{cases} \end{aligned} \tag{10.16}$$

In our usual manner, we say that this approach succeeds if the tuple $(\mathbf{x}^{\natural}, \mathbf{y}^{\natural}, \mathbf{w}^{\natural})$ is the unique optimal point of (10.16).

Let $s_x := \text{nnz}(\mathbf{x}^{\natural})$ and $s_y := \text{nnz}(\mathbf{y}^{\natural})$ be the sparsity levels of \mathbf{x}^{\natural} and \mathbf{y}^{\natural} . By applying the multiple demixing result of Theorem 9.2 as in the previous section, we find that a phase transition between the success and failure of (10.16) occurs around the region of sparsity pairs (s_x, s_y) such that

$$\psi_{\ell_1}\left(\frac{s_x}{d}\right) + \psi_{\ell_1}\left(\frac{s_y}{d}\right) = \frac{1}{2}. \tag{10.17}$$

Figure 10.4 illustrates this prediction with a numerical experiment in dimension $d = 100$. Continuing the theme, we see that the values of (s_x, s_y) satisfying the equality (10.17) closely align with the empirical 50% success contour. The experimental details are available in Appendix C.4.1. A rigorous result similar to the theorems in Section 10.1 is readily accessible to our methods. We omit the statement because it reflects our heuristic discussion above but offers no further surprises.

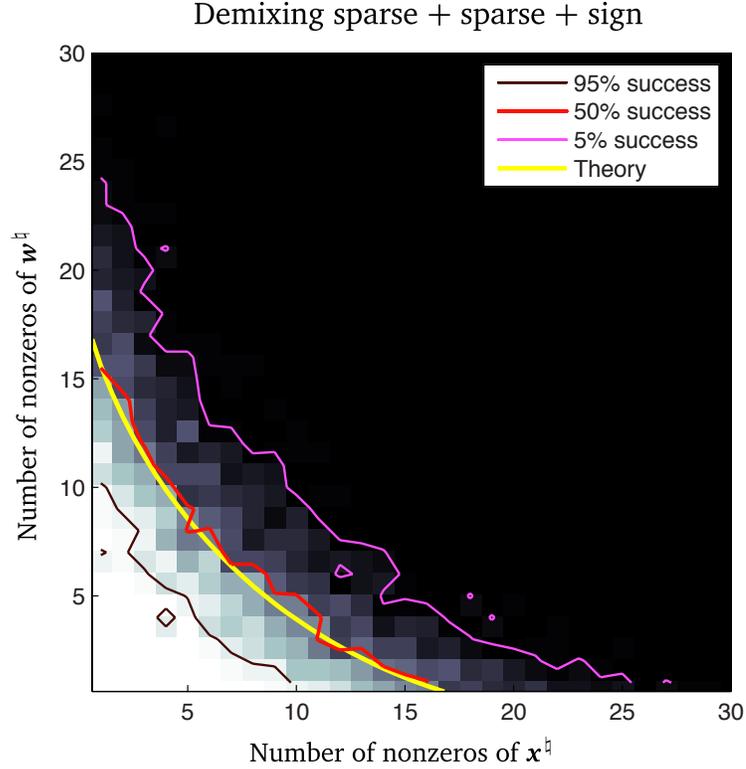


Figure 10.4: Empirical probability of demixing sparse + sparse + sign. This experiment tests the demixing method (10.16). The ambient dimension $d = 100$, and the sparsity levels of the constituents \mathbf{x}^b and \mathbf{y}^b are varied between one and 30. The colormap illustrates the empirical probability of success. The 95%, 50%, and 5% empirical success contours are shown, as is the smooth curve is the line where $\psi_{\ell_1}(s_x/d) + \psi_{\ell_1}(s_y/d) = \frac{1}{2}$, where s_x and s_y denote the sparsities of \mathbf{x}^b and \mathbf{y}^b and ψ_{ℓ_1} is defined in (6.22). See Appendix C.4.1 for numerical details, and Section 10.2.2 for a discussion.

10.3 Sparse + low-rank

We now consider a more sophisticated example where we are given a *matrix* observation $\mathbf{Z}_0 = \mathbf{X}^b + \mathcal{Q}(\mathbf{Y}^b) \in \mathbb{R}^{d \times d}$ where \mathbf{X}^b has low rank, \mathbf{Y}^b is sparse and \mathcal{Q} is a random rotation on the matrix space $\mathbb{R}^{d \times d}$. Such an observation provides a stylized model for applications in latent variable selection [GSPW09, CPW10, GSPW11] and robust principal component analysis [CLMW11]. In the first setting, the low-rank structure of \mathbf{X}^b models the confounding effect of latent variables, while \mathbf{Y}^b represents a sparse dependency structure suitable for sparse graphical models. In the context of principal component analysis, \mathbf{X}^b encodes a linear model with a few degrees of freedom, while \mathbf{Y}^b represents a corruption that is sparse in a random basis. We seek to recover the pair

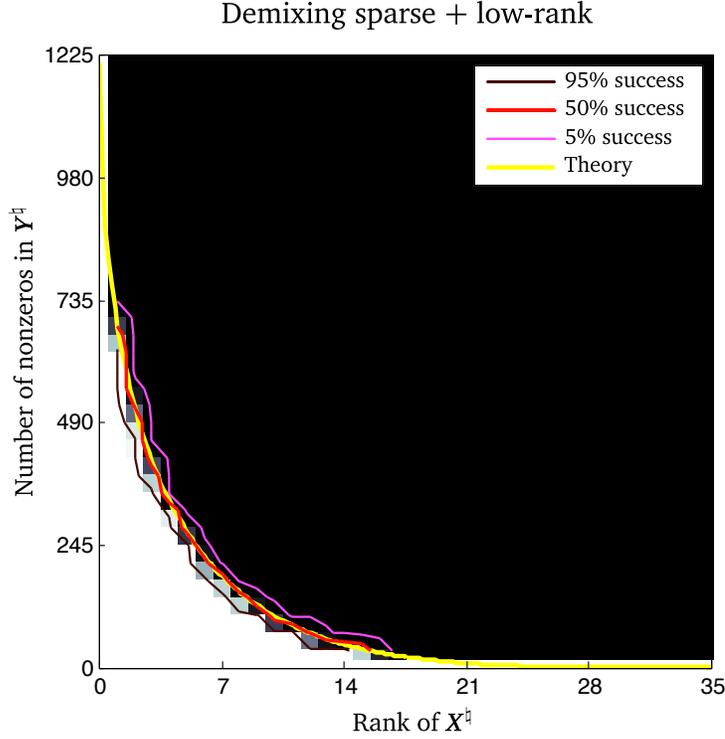


Figure 10.5: Empirical probability of success for demixing sparse + low-rank. We fix the sidelength $n = 35$, and explore sparsity levels $s = 1, n, 2n, \dots, n^2$ and ranks $r = 1, 2, \dots, n$ with the sparse + low-rank model described in Section 10.3. The colormap denotes the empirical probability of success over 25 trials. Overlaid on the colormap are the empirical 95%, 50% and 5% success curves. The smooth yellow curve gives the asymptotic location of the phase transition; its accuracy at predicting the location of the transition is quite good.

$(\mathbf{X}^d, \mathbf{Y}^d)$ from the observation \mathbf{Z}_0 and the basis \mathcal{Q} .

Following the demixing recipe of Section 1.2, we attempt to demix the observation by solving the convex program

$$\text{minimize } \|\mathbf{X}\|_{S_1} \quad \text{subject to } \|\mathbf{Y}\|_{\ell_1} \leq \alpha \quad \text{and} \quad \mathbf{X} + \mathcal{Q}(\mathbf{Y}) = \mathbf{Z}_0, \quad (10.18)$$

where $\alpha := \|\mathbf{Y}^d\|_{\ell_1}$ is known side information. Theorem 9.1 suggests that demixing will succeed when

$$\Delta := \delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{X}^d)) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{Y}^d)) \leq d - C\sqrt{d},$$

but will fail when the total statistical dimension $\Delta \geq d + C\sqrt{d}$.

We test this prediction with a numerical experiment. We fix the sidelength $n = 35$ and consider observations $\mathbf{Z}_0 \in \mathbb{R}^{n \times n}$ generated from the model described above. Figure 10.5 shows the empirical probability of success for a range of ranks r and sparsity levels s . (See Appendix C.4 for the full details.) The yellow curve gives the values of (r, s) such that

$$\psi_{\ell_1}\left(\frac{s}{n^2}\right) + \psi_{s_1}\left(\frac{r}{n^2}\right) = 1,$$

where $\psi_{\ell_1}(sn^{-2}) \rightarrow n^{-2}\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{Y}^{\natural}))$ and $\psi_{s_1}(rn^{-2}) \rightarrow n^{-2}\delta(\mathcal{D}(\|\cdot\|_{s_1}, \mathbf{X}^{\natural}))$ as $n \rightarrow \infty$ by Propositions 6.14 and 6.16. By Theorem 9.1, we expect a phase transition to occur, asymptotically as $n \rightarrow \infty$, near the yellow curve. As in previous sections, we find that this prediction is quite accurate even for the moderate value $n = 35$ used in this experiment.

10.4 Linear inverse problems

Demixing problems where the observation $\mathbf{z}_0 = \mathbf{G}\mathbf{x}^{\natural}$ consists of a single constituent are known as *linear inverse problems*. In this section, we describe how Theorem (9.5) with $n = 1$ implies that a phase transition for this model occurs as the number of measurements passes through a critical point. We focus here on two specific constituents, namely sparse vectors and low-rank matrices.

10.4.1 The sparse inverse problem

Suppose that our constituent signal $\mathbf{x}^{\natural} \in \mathbb{R}^d$ is sparse and the observation is given by $\mathbf{z}_0 = \mathbf{G}\mathbf{x}^{\natural}$ for some Gaussian matrix $\mathbf{G} \in \mathbb{R}^{m \times d}$. We solve for \mathbf{x}^{\natural} using the convex program

$$\text{minimize } \|\mathbf{x}\|_{\ell_1} \quad \text{subject to } \mathbf{G}\mathbf{x} = \mathbf{z}_0. \quad (10.19)$$

Theorem 9.2 with $n = 1$ implies (cf. Remark 9.3) that this approach succeeds with high probability when $\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) \leq m - O(\sqrt{d})$ and fails with high probability for $\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) \geq m + O(\sqrt{d})$. By our approximation bound (6.21), the phase transition

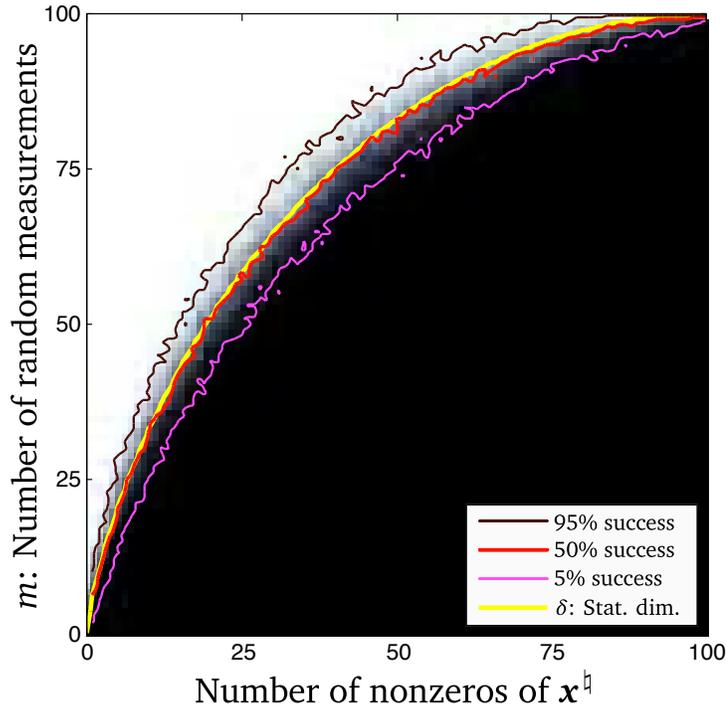
Compressed sensing with ℓ_1 minimization

Figure 10.6: Empirical success of a sparse linear inverse problem. We fix the dimension $d = 100$ and test the linear inverse method (10.19) for sparsity $s = 1, 2, \dots, d$ and number of measurements $m = 1, 2, \dots, d$. The colormap gives the empirical probability of success over 50 trials. The yellow curve is the theoretical (asymptotic) phase transition given by (10.20).

occurs in the region where

$$\psi_{\ell_1}\left(\frac{s_x}{d}\right) = \frac{m}{d}. \quad (10.20)$$

Figure 10.6 compares this prediction to a numerical experiment where the ambient dimension $d = 100$. We test (10.19) for each sparsity level $s_x = 1, \dots, d - 1, d$. The colormap displays the empirical probability of success, along with the 95%, 50% and 5% empirical success curves. We also show the numerically computed sparsity levels satisfying (10.20). Once again, the correspondence between our theory and the 50% success curve is quite good.

The low-rank inverse problem We now consider the observation model $\mathbf{z}_0 = \mathcal{G}(\mathbf{X}^\natural) \in \mathbb{R}^m$ where $\mathbf{X} \in \mathbb{R}^{n \times n}$ is a low-rank matrix and $\mathcal{G}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ is a Gaussian² measurement operator. Since the Schatten 1-norm is naturally associated with the variety of low-rank matrices, we solve for \mathbf{X}^\natural with the program

$$\text{minimize } \|\mathbf{X}\|_{S_1} \quad \text{subject to } \mathcal{G}(\mathbf{X}) = \mathbf{z}_0. \quad (10.21)$$

In the same manner as above, Theorem 9.2 with $n = 1$ implies that this method undergoes a phase transition from success to failure as the number of measurements goes through the region $m \approx \delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{x}^\natural))$. Proposition (6.28) shows that the statistical dimension $\delta(\mathcal{D}(\|\cdot\|_{S_1}, \mathbf{x}^\natural)) \approx n^2 \psi_{S_1}(\rho)$ for large n , which indicates that the phase transition region occurs for measurement levels that satisfy

$$\psi_{S_1} \left(\frac{r}{n}, 1 \right) = \frac{m}{n^2}. \quad (10.22)$$

Figure 10.7 shows the results of a numerical experiment where we fix the side length $n = 30$ and vary the rank of \mathbf{X}^\natural and the number of measurement m . The colormap indicates the empirical probability of success for this program, and we plot the 95%, 50% and 5% empirical success curves. Underneath these curves, we plot the values of (r, m) that satisfy the implicit equation (10.22). The theoretical curves again provide a very good prediction for the location of the empirical 50% success probability curve.

²We define a Gaussian operator on the matrix space $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ via the identification $\mathcal{G}(\mathbf{X}) := \mathbf{G} \text{vec}(\mathbf{X})$ for a standard Gaussian matrix $\mathbf{G} \in \mathbb{R}^{m \times n^2}$.

Low-rank matrix recovery via Schatten 1-norm minimization

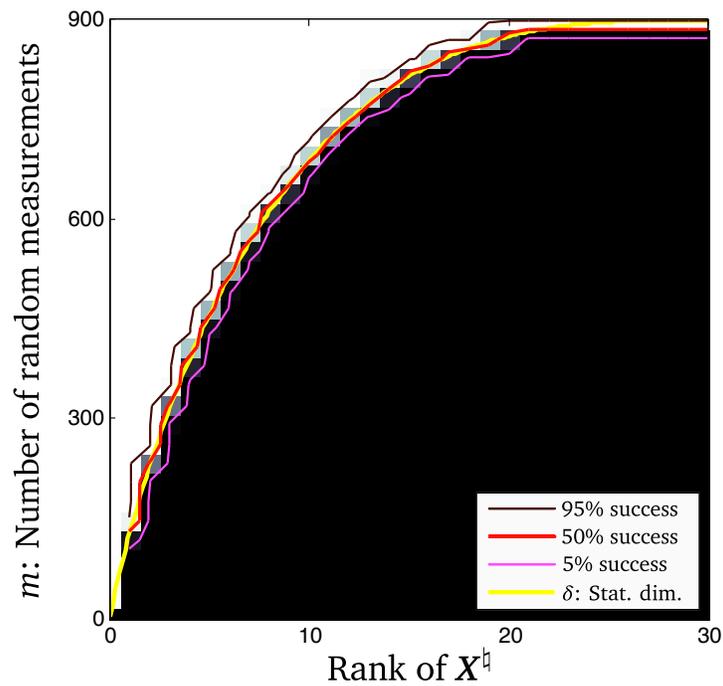


Figure 10.7: Empirical success of a low-rank inverse problem. We set the matrix sidelength $n = 30$. We test the capability of (10.21) to recover an $n \times n$ rank- r matrix from m Gaussian measurements. We repeat the experiment 50 times for each rank $r = 1, 2, \dots, n$ and number of measurements $m = 1, n - 1, 2n - 2, \dots, n^2$. The yellow curve is determined by (10.22). Further details for both of these experiments are available in Appendix C.6.

Appendix A

A general Steiner formula

This section contains a general Steiner formula for convex cones. As a special case, we recover both the Gaussian Steiner formula (Proposition 5.6) and the classical spherical Steiner formula (Proposition 5.8). In short, we show that the mean value of a sufficiently nice function of the projected norm of a Gaussian vector onto a cone is a linear combination of the intrinsic volumes of the cone.

Let us take a moment to introduce some notation. Given a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, we define the average value

$$\varphi_f(C) := \mathbb{E} [f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|)]. \quad (\text{A.1})$$

The rotational invariance of the Gaussian measure implies that φ_f is rotation-invariant, but in fact much more is true under some mild integrability conditions on f .

Theorem A.1 (General Steiner formula). *Let $C \in \mathcal{C}_d$ be a polyhedral cone, and suppose the function $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ satisfies $\mathbb{E} [|f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|) |] < \infty$. Then the function φ_f defined in (A.1) has the decomposition*

$$\varphi_f(C) = \sum_{k=0}^d \varphi_f(L_k) v_k(C) \quad (\text{A.2})$$

where L_k is a k -dimensional subspace of \mathbb{R}^d and $v_k(C)$ are the intrinsic volumes (5.1) of C . If $C \in \mathcal{C}_d$ is an arbitrary closed convex cone, f is bounded, and the map

$$\mathbf{x} \mapsto f(\|\Pi_C(\mathbf{x})\|, \|\Pi_{C^\circ}(\mathbf{x})\|)$$

is almost everywhere continuous on \mathbb{R}^d , then (A.2) also holds for this C .

While this result does not seem to be available in the literature, our proof is based on standard arguments that will certainly be familiar to experts. Indeed, a similar proof was used in the context of two-dimensional hyperbolic integral geometry in [San80, p. 242], but with fewer integrability considerations.

We demonstrate in Section A.1 that the Gaussian and spherical Steiner formulas follow from Theorem A.1. A new conic analog of the classical Euclidean Wills functional appears in Section A.1.1. The proof of Theorem A.1 appears in Section A.2.

A.1 Consequences of Theorem A.1

A number of results follow from the Theorem A.1. As a first application, we recover the Gaussian Steiner formula, Proposition 5.6. For convenience, we restate the claim here.

Proposition A.2 (Gaussian Steiner formula). *For an arbitrary closed convex cone $C \in \mathcal{C}_d$ and any $\varepsilon \geq 0$,*

$$\mathbb{P} \{ \|\Pi_C(\mathbf{g})\|^2 \leq \varepsilon \} = \sum_{k=0}^d \chi_k^2(\varepsilon) v_k(C). \quad (\text{A.3})$$

where $\chi_k^2(\varepsilon)$ is the cumulative distribution function (5.12) of a chi-squared random variable.

Proof. When $\varepsilon = 0$, we have $\|\Pi_C(\mathbf{g})\|^2 \leq \varepsilon = 0$ if and only if $\mathbf{g} \in C^\circ$, and hence the left-hand side of (A.3) is equal to $v_0(C)$. On the other hand, $\chi_0^2(0) = 1$ but $\chi_k^2(0) = 0$ for all $k \geq 1$, so the right-hand side of (A.3) is also equal to $v_0(C)$, verifying this special case.

For $\varepsilon > 0$, we define the function $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ by

$$f(a, b) := \begin{cases} 1, & a^2 \leq \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

We now verify that the conditions of Theorem A.1 hold for this f . By definition, f is

bounded, and moreover

$$\mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 \leq \varepsilon\} = \mathbb{E}[f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|)].$$

We must demonstrate that the map

$$h(\mathbf{x}) := f(\|\Pi_C(\mathbf{x})\|, \|\Pi_{C^\circ}(\mathbf{x})\|)$$

is continuous almost everywhere. Since the function $f(a, b)$ is continuous except where $a^2 = \varepsilon$, the set of (potential) discontinuities of h is given by

$$\text{discont}(h) := \{\mathbf{x} : \|\Pi_C(\mathbf{x})\|^2 = \varepsilon\} = \{\mathbf{x} : \text{dist}^2(\mathbf{x}, C^\circ) = \varepsilon\}$$

where the second equality follows from the formula (3.17) for the distance to a convex cone. Since $\varepsilon > 0$, the set $\text{discont}(h)$ is given by the boundary of a convex set:

$$\text{discont}(h) = \text{bdy}(C^\circ + \sqrt{\varepsilon}B_d),$$

where $\text{bdy}(\cdot)$ denotes the boundary map. (The equality above fails for $\varepsilon = 0$, which explains why we treat the $\varepsilon = 0$ as a special case above.) A consequence of the separating hyperplane theorem and the Lebesgue density theorem implies that the boundary of every convex set has Lebesgue measure zero [Lan86]. We conclude that h is continuous almost everywhere. By Theorem A.1, we have the representation

$$\mathbb{P}\{\|\Pi_C(\mathbf{g})\|^2 \leq \varepsilon\} = \sum_{k=0}^d v_k(C) \mathbb{E}[f(\|\Pi_{L_k}(\mathbf{g})\|, \|\Pi_{L_k^\perp}(\mathbf{g})\|)] = \sum_{k=0}^d v_k(C) \chi_k^2(\varepsilon),$$

where the last equality follows immediately from the definition of f and χ_k^2 . This completes the proof. \square

A natural variant of Proposition A.2 replaces the Gaussian vector \mathbf{g} with a spherically distributed random variable $\boldsymbol{\theta}$.

Proposition A.3 (Spherical Steiner formula). *Let $C \in \mathcal{C}_d$ be a closed, convex cone, and*

let $\boldsymbol{\theta}$ be uniformly distributed over the sphere S^{d-1} . For any $\varepsilon \in [0, 1]$,

$$\mathbb{P}\{\|\boldsymbol{\Pi}_C(\boldsymbol{\theta})\|^2 \leq \varepsilon\} = \sum_{k=0}^d v_k(C) B_k^d(\varepsilon) \quad (\text{A.4})$$

where $B_k^d(\varepsilon)$ is the cumulative distribution function (5.13) of a $\text{Beta}(k/2, (d-k)/2)$ random variable.

The proof, while very similar to the Gaussian case, contains enough additional technical wrinkles that we present it in full.

Proof. We first consider the boundary cases where $\varepsilon \in \{0, 1\}$. When $\varepsilon = 0$, we have $\|\boldsymbol{\Pi}_C(\boldsymbol{\theta})\|^2 \leq \varepsilon = 0$ if and only if $\boldsymbol{\theta} \in C^\circ$, so that the left-hand side of (A.4) is equal to $v_0(C)$. For the same reason, $B_k^d(0)$ is equal to zero if and only if $k = 0$, so the left- and right-hand sides of (A.4) indeed agree.

In the case where $\varepsilon = 1$, the left-hand side of (A.4) is equal to one. Since $B_k^d(1) = 1$ for all k , the right-hand side of (A.4) is given by $\sum_{k=0}^{d-1} v_k(C) = 1$ by the fact (5.4) that the intrinsic volumes sum to one. Therefore, the equality in (A.4) also holds for $\varepsilon = 1$.

We have now reached the interesting case where $0 < \varepsilon < 1$. Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by the indicator function

$$f(a, b) = \begin{cases} 1, & a^2 \leq \frac{\varepsilon}{1-\varepsilon} b^2, \\ 0, & \text{otherwise.} \end{cases}$$

A uniformly distributed variable $\boldsymbol{\theta}$ is equal in distribution to $\mathbf{g} / \|\mathbf{g}\|$ for a standard Gaussian vector \mathbf{g} , so that

$$\begin{aligned} \mathbb{P}\{\|\boldsymbol{\Pi}_C(\boldsymbol{\theta})\|^2 \geq \varepsilon\} &= \mathbb{P}\{\|\boldsymbol{\Pi}_C(\mathbf{g})\|^2 \geq \varepsilon \|\mathbf{g}\|^2\} = \mathbb{P}\left\{\|\boldsymbol{\Pi}_C(\mathbf{g})\|^2 \geq \frac{\varepsilon}{1-\varepsilon} \|\boldsymbol{\Pi}_{C^\circ}(\mathbf{g})\|^2\right\} \quad (\text{A.5}) \\ &= \mathbb{E}[f(\|\boldsymbol{\Pi}_C(\mathbf{g})\|, \|\boldsymbol{\Pi}_{C^\circ}(\mathbf{g})\|)] \end{aligned}$$

where the second equality follows from the Pythagorean identity (3.16).

We now verify the hypothesis of Theorem A.1. First, we note that f is bounded in magnitude by definition. In order to verify that (A.4) holds for all closed convex cones,

we must only verify that the map h defined by

$$h(\mathbf{x}) := f(\|\Pi_C(\mathbf{x})\|, \|\Pi_{C^\circ}(\mathbf{x})\|)$$

is continuous almost everywhere. This demonstration, while not difficult, requires some enterprise.

The function $f(a, b)$ is continuous except where $a(1 - \varepsilon) = \varepsilon b$ and the projection map Π_C is continuous everywhere, so the set of (potential) discontinuities of h is given by

$$\text{discont}(h) := \{\mathbf{x} : \|\Pi_C(\mathbf{x})\|^2(1 - \varepsilon) = \varepsilon \|\Pi_{C^\circ}(\mathbf{x})\|^2\}.$$

The set $\text{discont}(h)$ is a (nonconvex) cone, so we may verify that the Lebesgue measure of $\text{discont}(h)$ is zero by checking that the spherical Lebesgue measure of the restriction $\text{discont}(h) \cap S^{d-1}$ is zero. This restricted set is of the form

$$\text{discont}(h) \cap S^{d-1} = \{\boldsymbol{\theta} \in S^{d-1} : \|\Pi_C(\boldsymbol{\theta})\|^2 = \varepsilon\},$$

which is seen by reversing the homogeneity argument (A.5) above. But the norm of the projection satisfies $\|\Pi_C(\boldsymbol{\theta})\|^2 = \text{dist}^2(\boldsymbol{\theta}, C^\circ)$, so that the set of restricted discontinuities can be written as the boundary of a convex set restricted to the sphere:

$$\text{discont}(h) \cap S^{d-1} = \{\mathbf{x} : \text{dist}^2(\mathbf{x}, C^\circ) = \varepsilon\} \cap S^{d-1} = \text{bdy}(\{\mathbf{x} : \text{dist}(\mathbf{x}, C^\circ) \leq \varepsilon\}) \cap S^{d-1},$$

where $\text{bdy}(\cdot)$ denotes the boundary map. Note that the boundary interpretation requires the assumption $\varepsilon > 0$. The boundary of a convex set has Lebesgue measure zero [Lan86], so the set $\text{discont}(h) \cap S^{d-1}$ also has measure zero, and by homogeneity the set $\text{discont}(h)$ has measure zero.

Because f is bounded and h is continuous almost everywhere, Theorem A.1 implies that (A.4) holds for any closed cone C and all $0 < \varepsilon < 1$. Since we checked the case $\varepsilon \in \{0, 1\}$ previously, we are done. \square

A.1.1 The conic Wills functional

This section introduces a conic analog of the classical *Wills functional* from Euclidean geometry. For a compact convex set $K \subset \mathbb{R}^d$, the Euclidean Wills functional is given by

$$\mathcal{W}(K) := \int_{\mathbb{R}^d} e^{-\pi \operatorname{dist}^2(x, K)} dx.$$

The Wills functional gained some attention in the mid-20th century due to Wills' conjecture that the value $\mathcal{W}(K)$ provides an upper bound on the number of integer lattice points contained in the convex body K [Wil73]. The conjecture holds in dimensions two and three [Nos48, Ove75] and in some other special cases [HW74]. Wills' conjecture was ultimately disproved by Hadwiger, who provided a counterexample in dimension $d = 441$ [Had75]. Betke & Henk [BH93] showed that the crosspolytope provides a counterexample¹ in dimension 207. Despite these counterexamples, the Wills functional continues to have intrinsic geometric interest. For example, Vitale [Vit96] demonstrated a deep connection between the Wills functional and Gaussian processes.

In this section, we define a new conic analog of the Wills functional, and we use our general Steiner formula to give an expression for its value as a polynomial in the conic intrinsic volumes. As a corollary, we obtain an elegant proof of the product rule (5.7) for conic intrinsic volumes.

Definition A.4. For any $C \in \mathcal{C}_d$, we define the *conic Wills functional* $W_C(\lambda)$ by

$$W_C(\lambda) := \lambda^d \mathbb{E} \exp \left(-\frac{\lambda^2 - 1}{2} \operatorname{dist}^2(\mathbf{g}, C) \right). \quad (\text{A.6})$$

The awkward exponent and leading polynomial are justified by the simple expression for the conic Wills functional given by the following theorem.

Theorem A.5 (Conic Wills functional). *For any $C \in \mathcal{C}_d$ and $\lambda > 0$, the conic Wills*

¹Betke & Henk's work is also notable because it provided formulas for the internal and external angles of the crosspolytope (see Section A.2.2 below) based on work of Ruben [Rub60]. Later work on this subject led to precise neighborliness investigations that eventually brought Donoho to phase transition calculations for ℓ_1 minimization.

functional is given by

$$W_C(\lambda) = \sum_{k=0}^d \lambda^k \nu_k(C). \quad (\text{A.7})$$

Proof. Assume first that $\lambda \geq 1$. We will extend the formula to all $\lambda > 0$ by a standard analytic continuation argument at the end of the proof. Define the function $f_\lambda: \mathbb{R}_+^2 \rightarrow \mathbb{R}$ by

$$f_\lambda(a, b) := \exp\left(-\frac{\lambda^2 - 1}{2} b^2\right)$$

For each $\lambda \geq 1$, the function f_λ is bounded by one over \mathbb{R}_+^2 . Moreover, f_λ is continuous, and hence the map $\mathbf{x} \mapsto f_\lambda(\|\Pi_C(\mathbf{x})\|, \|\Pi_{C^\circ}(\mathbf{x})\|)$ is continuous on \mathbb{R}^d . By Theorem A.1 and the fact (3.17) that $\text{dist}^2(\mathbf{x}, C) = \|\Pi_{C^\circ}(\mathbf{x})\|^2$, we find the expression

$$\lambda^{-d} W_C(\lambda) = \mathbb{E}\left[f\left(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|\right)\right] = \sum_{k=0}^d \nu_k(C) \mathbb{E}\left[f\left(\|\Pi_{L_k}(\mathbf{g})\|, \|\Pi_{L_k^\perp}(\mathbf{g})\|\right)\right] \quad (\text{A.8})$$

The coefficients in this final expression can be computed exactly. By rotational invariance of the Gaussian measure, we have

$$\begin{aligned} \mathbb{E}\left[f\left(\|\Pi_{L_k}(\mathbf{g})\|, \|\Pi_{L_k^\perp}(\mathbf{g})\|\right)\right] &= \mathbb{E}\exp\left(-\frac{\lambda^2 - 1}{2} \|\Pi_{L_{d-k}}(\mathbf{g})\|^2\right) \\ &= \mathbb{E}\exp\left(-\frac{\lambda^2 - 1}{2} \sum_{i=0}^{d-k} g_i^2\right) \end{aligned}$$

where the last equality follows from the fact that the projection of a Gaussian vector onto a $(d - k)$ -dimensional subspace is standard Gaussian vector on that subspace. But the coordinates g_i of a standard Gaussian vector are independent, so

$$\mathbb{E}\exp\left(-\frac{\lambda^2 - 1}{2} \sum_{i=0}^{d-k} g_i^2\right) = \left(\mathbb{E}\exp\left(-\frac{\lambda^2 - 1}{2} g^2\right)\right)^{d-k},$$

where $g \in \mathbb{R}$ is a standard one-dimensional Gaussian random variable. The inner moment generating function then given by

$$\mathbb{E}\exp\left(-\frac{\lambda^2 - 1}{2} g^2\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\lambda^2 g^2/2} dg = \frac{1}{\lambda}.$$

To summarize, our calculation shows that

$$\mathbb{E}\left[f\left(\|\Pi_{L_k}(\mathbf{g})\|, \|\Pi_{L_k^\perp}(\mathbf{g})\|\right)\right] = \lambda^{k-d}.$$

Expression (A.7) follows upon applying this expression to (A.8), which shows the claim for all $\lambda \geq 1$.

(Analytic continuation) We will show that the expectation in the definition (A.6) of the conic Wills functional is finite for all $\lambda > 0$. This fact immediately implies that the formula (A.7) holds for all $\lambda > 0$ by a standard analytic continuation argument, cf. [AF03, pp. 124–125].

The integrability condition is easy to check. For any $\lambda > 0$, we apply the Pythagorean theorem (3.16) and the distance formula (3.17) to find

$$\frac{\lambda^2 - 1}{2} \text{dist}^2(\mathbf{g}, C) + \frac{1}{2} \|\mathbf{g}\|^2 = \frac{\lambda^2}{2} \|\Pi_{C^c}(\mathbf{g})\|^2 + \|\Pi_C(\mathbf{g})\|^2 \geq \frac{\lambda^2}{2} \|\mathbf{g}\|^2. \quad (\text{A.9})$$

Therefore,

$$\mathbb{E} \exp\left(-\frac{\lambda^2 - 1}{2} \text{dist}^2(\mathbf{g}, C)\right) \leq \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \exp\left(-\frac{\lambda}{2} \|\mathbf{g}\|^2\right) d\mathbf{g} = \frac{1}{\lambda^d},$$

where the inequality follows by the definition of the Gaussian measure and (A.9). This expression above evidently finite for $\lambda > 0$. \square

As a corollary, we obtain an elegant proof of the product rule (5.7).

Corollary A.6. *For any $C \in \mathcal{C}_d$ and $C' \in \mathcal{C}_{d'}$, the intrinsic volumes of the product $C \times C' \in \mathcal{C}_{d+d'}$ are given by*

$$v_k(C \times C') = \sum_{i+j=k} v_i(C)v_j(C').$$

Proof. Let $\mathbf{g} \in \mathbb{R}^d$ and $\mathbf{g}' \in \mathbb{R}^{d'}$ be independent standard Gaussian vectors so that the concatenation $(\mathbf{g}; \mathbf{g}')$ is a standard Gaussian vector in $\mathbb{R}^{d+d'}$. Then by definition of the

Wills functional,

$$\begin{aligned}
W_{C \times C'}(\lambda) &= \lambda^{d+d'} \mathbb{E} \exp \left(-\frac{\lambda^2 - 1}{2} \text{dist}^2(\mathbf{g}; \mathbf{g}', C \times C') \right) \\
&= \lambda^d \mathbb{E} \exp \left(-\frac{\lambda^2 - 1}{2} \text{dist}^2(\mathbf{g}, C) \right) \times \lambda^{d'} \mathbb{E} \exp \left(-\frac{\lambda^2 - 1}{2} \text{dist}^2(\mathbf{g}', C') \right) \\
&= W_C(\lambda) \times W_{C'}(\lambda).
\end{aligned}$$

The second equality follows by Fubini's theorem; we may justify this application by noting that the exponential is dominated by a constant for all $\lambda \geq 1$. Applying Theorem A.5 and the multiplication rule for polynomial coefficients, we have

$$W_C(\lambda) \times W_{C'}(\lambda) = \sum_{k=0}^{d+d'} \lambda^k \sum_{i+j=k} v_i(C) v_j(C').$$

Comparing the coefficients of λ^k above to those in the expression

$$W_{C \times C'}(\lambda) = \sum_{k=0}^d \lambda^k v_k(C \times C')$$

yields the claim. □

Remark A.7. In the Euclidean case, Schneider & Weil [SW08, pp. 610] attribute the schema of the proof above to Hadwiger [Had75].

A.2 Proof of the general Steiner formula

Before the diving into the proof of Theorem A.1, we develop a fundamental decomposition of \mathbb{R}^d induced by the facial structure of a cone (Section A.2.1) and the *polytope angle* characterization of conic intrinsic volumes (Section A.2.2).

A.2.1 A tiling induced by a cone

Recall that a *face* F of a closed cone C is a set consisting of the maximizers of some linear function over C . In other words, a face is of the form

$$F = \{\mathbf{x} \in C : \langle \mathbf{x}, \mathbf{y} \rangle = \sup_{\mathbf{z} \in C} \langle \mathbf{z}, \mathbf{y} \rangle\}.$$

In this case, we say \mathbf{y} is an outward normal to C that supports F . Note that the face F above is nonempty if and only if $\mathbf{y} \in C^\circ$ because the supremum is $+\infty$ for $\mathbf{y} \notin C^\circ$. Given a face F of a cone C , the *normal face*

$$N_F := \text{lin}(F)^\perp \cap C^\circ \tag{A.10}$$

is the face of the polar cone C° consisting of all the outward normals to C that support the face F . The following result, an amplification of an observation of McMullen [McM75, Lem. 3], provides a tiling of \mathbb{R}^d as an orthogonal sum of the faces and normal faces of C .

Proposition A.8. *Suppose C is a closed convex cone. Then we have the orthogonal decomposition $\Pi_C^{-1}(\text{relint}(F)) = \text{relint}(F) + N_F$, and moreover \mathbb{R}^d can be written as the disjoint union*

$$\mathbb{R}^d = \bigsqcup_{F \text{ a face of } C} (\text{relint}(F) + N_F). \tag{A.11}$$

Proof. Equation (3.15) gives a unique orthogonal decomposition of \mathbf{x} over C and C° , so that in particular

$$\Pi_C^{-1}(\text{relint}(F)) \supset \text{relint}(F) + N_F.$$

We now show the reverse inclusion. Let F be any face of C . For $\mathbf{x} \in \Pi_C^{-1}(\text{relint}(F))$, we have the orthogonal decomposition (3.15)

$$\mathbf{x} = \Pi_C(\mathbf{x}) + \Pi_{C^\circ}(\mathbf{x}).$$

Since $\Pi_C(\mathbf{x}) \in \text{relint}(F)$ by assumption and $\Pi_{C^\circ}(\mathbf{x}) \in C^\circ$, we only need argue that

$\Pi_{C^\circ}(\mathbf{x}) \in \text{lin}(F)^\perp$ by definition of N_F .

We take a perturbative approach. Let $\mathbf{h} \in \text{lin}(F)$. Because $\Pi_C(\mathbf{x})$ is in the relative interior of F , there is an $\varepsilon > 0$ such that both $\Pi_C(\mathbf{x}) + \varepsilon\mathbf{h}$ and $\Pi_C(\mathbf{x}) - \varepsilon\mathbf{h}$ lie in F . By the orthogonality of the decomposition (3.15), we see

$$\langle \Pi_{C^\circ}(\mathbf{x}), \pm\varepsilon\mathbf{h} \rangle = \langle \Pi_{C^\circ}(\mathbf{x}), \Pi_C(\mathbf{x}) \pm \varepsilon\mathbf{h} \rangle \leq 0$$

where the inequality holds by the definition of a polar cone. In other words, $\Pi_{C^\circ}(\mathbf{x})$ must be orthogonal to \mathbf{h} . Since \mathbf{h} was an arbitrary member of $\text{lin}(F)$, we have $\Pi_{C^\circ}(\mathbf{x}) \in \text{lin}(F)^\perp$, and hence $\Pi_{C^\circ}(\mathbf{x}) \in N_F$. Therefore $\mathbf{x} = \Pi_C(\mathbf{x}) + \Pi_{C^\circ}(\mathbf{x}) \in \text{relint}(F) + N_F$, which shows the second claimed inclusion. Hence $\Pi_C^{-1}(\text{relint}(F)) = \text{relint}(F) + N_F$, as claimed.

To finish the proof, note that the uniqueness of the projection map implies that every point $\mathbf{x} \in \mathbb{R}^d$ maps to some *unique* face of C . Therefore, the entire space \mathbb{R}^d is given by the disjoint union

$$\mathbb{R}^d = \bigsqcup_{F \text{ a face of } C} \Pi_C^{-1}(\text{relint}(F)) = \bigsqcup_{F \text{ a face of } C} (\text{relint}(F) + N_F),$$

where the second equality follows from the demonstration above. \square

A.2.2 Polytope angles & intrinsic volumes

The *solid angle* of a cone C is defined by

$$\angle(C) := \int_C e^{-\pi\|y\|^2} dy, \tag{A.12}$$

where the integral is taken with respect to the Lebesgue measure on the linear hull of C . By requiring integration over the linear hull of the cone, we ensure that the solid angle is an *intrinsic* measure of the size of a cone, that is, $\angle(C)$ does not depend on the dimension of the space that C is embedded in, but only on the intrinsic dimensionality of C . When C is full dimensional (that is $\text{lin}(C) = \mathbb{R}^d$ say) the change of variables

$\mathbf{x} = \sqrt{2\pi}\mathbf{y}$ shows that we may interpret the solid angle probabilistically:

$$\angle(C) = \frac{1}{(2\pi)^{d/2}} \int_C e^{-\|\mathbf{x}\|^2/2} d\mathbf{x} = \mathbb{P}\{\mathbf{g} \in C\} = \mathbb{P}\{\boldsymbol{\theta} \in C \cap S^{d-1}\}, \quad (\text{A.13})$$

where \mathbf{g} is a standard Gaussian vector in \mathbb{R}^d and $\boldsymbol{\theta}$ is uniform on the sphere S^{d-1} . Both interpretations of the solid angle $\angle(C)$ prove useful below.

The intrinsic volumes of polyhedral cones have an important representation in terms of polytope angles, defined as follows. For a face F of a polyhedral cone C , the *internal* and *external angles* of F at C are simply the solid angles of the cones F and N_F . For comparison with the classical literature, we note that the internal angle is often denoted by $\beta(0, F)$, while the external angle is $\gamma(F, C)$; however, we will not use this notation in this work.

Fact A.9 (Polytope angles). *Let $C \in \mathcal{C}_d$ be a polyhedral cone. Then*

$$v_k(C) = \sum_{\substack{F \text{ a } k\text{-dim} \\ \text{face of } C}} \angle(F) \angle(N_F) \quad (\text{A.14})$$

This result is classical [McM75], but we repeat the easy demonstration because it illustrates an important technique that appears below.

Proof. By the definition (5.1) of the intrinsic volumes we have the formula and the tiling of \mathbb{R}^d given in Proposition (A.11),

$$v_k(C) = \mathbb{P}\{\Pi_C(\mathbf{x}) \in \text{a } k\text{-dim. face of } C\} = \sum_{\substack{F \text{ a } k\text{-dim} \\ \text{face of } C}} \frac{1}{(2\pi)^{d/2}} \int_{\Pi_C^{-1}(\text{relint}(F))} e^{-\|\mathbf{x}\|^2/2} d\mathbf{x},$$

where we have used the fact that polyhedral cones have only a finite number of faces. The orthogonal decomposition of the tiles in (A.11) and an application of Fubini's theorem lets us write the integral above as a product:

$$\frac{1}{(2\pi)^{d/2}} \int_{\Pi_C^{-1}(\text{relint}(F))} e^{-\|\mathbf{x}\|^2/2} d\mathbf{x} = \int_{\text{relint}(F)} \int_{N_F} e^{-\pi(\|\mathbf{w}\|^2 + \|\mathbf{u}\|^2)} d\mathbf{w} d\mathbf{u} = \angle(F) \angle(N_F),$$

where we used the change of variables $\mathbf{x} = 2\pi(\mathbf{u} + \mathbf{w}) \in \Pi_C^{-1}(C)$ for some (uniquely

determined) $\mathbf{u} \in \text{relint}(F)$ and $\mathbf{w} \in N_F$. The second equality follows from the definition (A.12) of the solid angle and the fact that the boundary of a convex set has measure zero [Lan86]. The claim (A.14) follows from the two displayed equations above. \square

A.2.3 Proof of Theorem A.1

Suppose that C is a polyhedral cone and that the expectation

$$\mathbb{E} [|f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|)|] < \infty.$$

Define the random variables $\mathbf{u} := \Pi_C(\mathbf{g})$ and $\mathbf{w} := \Pi_{C^\circ}(\mathbf{g})$. Although \mathbf{u} and \mathbf{w} are dependent random variables, we will find that they behave like independent Gaussian vectors when we restrict \mathbf{g} to a single tile from Proposition A.8. This fact is the heart of the proof.

Since C is a polyhedral cone, the tiling (A.11) given by Proposition A.8 contains only a finite number of terms. The boundary $F \setminus \text{relint}(F)$ of a convex cone has Lebesgue measure zero [Lan86, Thm. 1], so that

$$\varphi_f(C) = \mathbb{E} [f(\|\mathbf{u}\|, \|\mathbf{w}\|)] = \sum_{k=0}^d \sum_{\substack{F \text{ a } k\text{-dim} \\ \text{face of } C}} \mathbb{E} [f(\|\mathbf{u}\|, \|\mathbf{w}\|) \cdot \mathbf{1}_F(\mathbf{u})], \quad (\text{A.15})$$

where $\mathbf{1}_F$ is the indicator function (3.1) on the face F . The orthogonality of the tiles in (A.11) implies that we may integrate over F and N_F independently:

$$\mathbb{E} [f(\|\mathbf{u}\|, \|\mathbf{w}\|) \cdot \mathbf{1}_F(\mathbf{u})] = \frac{1}{(2\pi)^{d/2}} \int_F \int_{N_F} f(\|\mathbf{u}\|, \|\mathbf{w}\|) e^{-(\|\mathbf{u}\|^2 + \|\mathbf{w}\|^2)/2} d\mathbf{w} d\mathbf{u},$$

where our integrability assumption justifies the application of Fubini's theorem [Fol99, Thm. 2.37]. Note that the integrals above are taken with respect to the Lebesgue measure on the linear hulls of F and N_F . It turns out that the integral above is proportional to the product of the solid angles of the face F and the normal face N_F .

Let k denote the dimension of the linear hull of F . Integration in polar coordi-

nates [Fol99, Thm. 2.49] yields

$$\begin{aligned} \mathbb{E}[f(\|\mathbf{u}\|, \|\mathbf{w}\|) \cdot \mathbf{1}_F(\mathbf{u})] &= \int_{F \cap \mathbb{S}^{k-1}} d\sigma_{k-1} \int_{N_F \cap \mathbb{S}^{d-k-1}} d\sigma_{d-k-1} \\ &\quad \times \frac{1}{(2\pi)^{d/2}} \int_0^\infty \int_0^\infty f(s, t) s^{k-1} t^{d-k-1} e^{-(s^2+t^2)/2} ds dt. \end{aligned}$$

Because the solid angle of a cone is equal to the probability that a spherically distributed random variable lands in the cone (A.13), the first two integrals above are proportional to the solid angle of the faces:

$$\int_{F \cap \mathbb{S}^{k-1}} d\sigma_{k-1} = \sigma_{k-1}(\mathbb{S}^{k-1}) \angle(F) \quad \text{and} \quad \int_{N_F \cap \mathbb{S}^{d-k-1}} d\sigma_{d-k-1} = \sigma_{d-k-1}(\mathbb{S}^{d-k-1}) \angle(N_F).$$

But the same argument as before, with F replaced by $\text{lin}(F)$ and N_F replaced by $\text{lin}(F)^\perp$, delivers the identity

$$\begin{aligned} \frac{\sigma_{k-1}(\mathbb{S}^{k-1}) \sigma_{d-k-1}(\mathbb{S}^{d-k-1})}{(2\pi)^{d/2}} \int_0^\infty \int_0^\infty f(s, t) s^{k-1} t^{d-k-1} e^{-(s^2+t^2)/2} ds dt \\ = \mathbb{E}[f(\|\mathbf{\Pi}_{\text{lin}(F)}(\mathbf{g})\|, \|\mathbf{\Pi}_{\text{lin}(F)^\perp}(\mathbf{g})\|)]. \end{aligned}$$

The expression expectation is invariant under rotations, so it depends only on the dimension k of the face F . Therefore,

$$\mathbb{E}[f(\|\mathbf{u}\|, \|\mathbf{w}\|) \cdot \mathbf{1}_F(\mathbf{u})] = \angle(F) \angle(N_F) \mathbb{E}[f(\|\mathbf{\Pi}_{L_k}(\mathbf{g})\|, \|\mathbf{\Pi}_{L_k^\perp}(\mathbf{g})\|)] = \angle(F) \angle(N_F) \varphi_f(L_k)$$

for some k -dimensional subspace L_k . Applying this relation to (A.15) reveals

$$\varphi_f(C) = \sum_{k=0}^d \varphi_f(L_k) \sum_{\substack{F \text{ a } k\text{-dim} \\ \text{face of } C}} \angle(F) \angle(N_F) = \sum_{k=0}^d \varphi_f(L_k) \nu_k(C),$$

where the final equality follows by (A.14). This completes the claim when C is polyhedral.

For a general closed convex C , the proof proceeds by approximation with polyhedral cones. Assume f is bounded and that the map $\mathbf{x} \mapsto f(\|\mathbf{\Pi}_C(\mathbf{x})\|, \|\mathbf{\Pi}_{C^\circ}(\mathbf{x})\|)$ is continuous

Lebesgue almost everywhere. Let $(C_i)_{i \in \mathbb{N}}$ be a sequence of polyhedral cones converging to C in the conic Hausdorff metric. For each \mathbf{x} , continuity of the projection map under the conic Hausdorff metric (Proposition 3.8) implies the convergence of the projected norms:

$$\lim_{i \rightarrow \infty} \|\Pi_{C_i}(\mathbf{x})\| = \|\Pi_C(\mathbf{x})\| \quad \text{and} \quad \lim_{i \rightarrow \infty} \|\Pi_{C_i^\circ}(\mathbf{x})\| = \|\Pi_{C^\circ}(\mathbf{x})\|.$$

Since f is bounded, the map $\mathbf{x} \mapsto f(\|\Pi_C(\mathbf{x})\|, \|\Pi_{C^\circ}(\mathbf{x})\|)$ is integrable with respect to the Gaussian measure. The convergence of the norms, together with the integrability and almost everywhere continuity of the map, imply that

$$f(\|\Pi_{C_i}(\mathbf{g})\|, \|\Pi_{C_i^\circ}(\mathbf{g})\|) \rightarrow f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|) \text{ in probability as } i \rightarrow \infty$$

by the continuous mapping theorem [Kal02, Lem. 4.3]. But convergence in probability implies convergence in distribution, so for $i \rightarrow \infty$, we have

$$\varphi_f(C_i) = \mathbb{E}[f(\|\Pi_{C_i}(\mathbf{g})\|, \|\Pi_{C_i^\circ}(\mathbf{g})\|)] \rightarrow \mathbb{E}[f(\|\Pi_C(\mathbf{g})\|, \|\Pi_{C^\circ}(\mathbf{g})\|)] = \varphi_f(C).$$

The boundedness of f implies that the map $\mathbf{x} \mapsto f(\|\Pi_{C_i}(\mathbf{x})\|, \|\Pi_{C_i^\circ}(\mathbf{x})\|)$ corresponding to the polyhedral cone C_i is also integrable with respect to the Gaussian measure on \mathbb{R}^d . Therefore, we may apply the first part of Theorem A.1 to the polyhedral cones $(C_i)_{i \in \mathbb{N}}$ to find

$$\lim_{i \rightarrow \infty} \varphi_f(C_i) = \lim_{i \rightarrow \infty} \sum_{k=0}^d \varphi_f(L_k) v_k(C_i) = \sum_{k=0}^d \varphi_f(L_k) v_k(C)$$

where the second equality follows by the continuity of intrinsic volumes. Thus (A.2) holds, completing the proof. \square

Remark A.10. The boundedness condition on f is only used to ensure that the integral $\varphi_f(C)$ is well defined for every convex cone C . Any other condition on f that ensures integrability for all polyhedral cones will lead the same conclusion.

Appendix B

Proofs of results in Chapter 10

We collect the proof of the rigorous results from Chapter 10 below. The proofs simply combine our general results on phase transitions in demixing programs from Section 9 with the estimates for the statistical dimensions of descent cones in Section 6.2.

B.1 Sparse + sparse demixing

This section presents the proof of the results presented in Section 10.1. We first prove that a phase transition occurs in sparse + sparse demixing (Theorem 10.1), then provide a proof of Theorem 10.3 concerning strong guarantees for sparse + sparse demixing. Finally, we prove Theorem 10.4 which shows that a phase transition occurs sparse + sparse undersampled demixing.

B.1.1 Phase transition

We now rigorously establish the existence of a sparsity-parameterized phase transition in sparse + sparse demixing.

Proof of Theorem 10.1. We split the demonstration according to the headings.

(Well-defined curve) The fact that ρ_y is well-defined for each $\rho_x \in (0, 1)$ follows easily from the properties of ψ_{ℓ_1} listed in Lemma 6.15. Since $\psi_{\ell_1}(\rho)$ is an analytic, strictly increasing function of ρ that maps $(0, 1)$ to $(0, 1)$, the inverse $\psi_{\ell_1}^{-1}: (0, 1) \rightarrow (0, 1)$ is a

well-defined analytic function [AF03, Thm. 5.3.3]. Hence, the parameter ρ_y defined by

$$\rho_y := \psi_{\ell_1}^{-1}(1 - \psi_{\ell_1}(\rho_x))$$

is a continuous function of ρ_x over the interval $(0, 1)$. By definition, the pair (ρ_x, ρ_y) satisfies the implicit equation (10.3).

It only remains to show that the limiting values (10.4) hold. When $\rho_x \rightarrow 0$, we have $\psi_{\ell_1}(\rho_x) \rightarrow 0$ by (6.26). This implies that $\psi_{\ell_1}^{-1}(1 - \psi_{\ell_1}(\rho_x)) \rightarrow 1$ by (6.27). The same properties imply that as $\rho_x \rightarrow 1$, the function $\psi_{\ell_1}(\rho_x) \rightarrow 1$, and so $\psi_{\ell_1}^{-1}(1 - \psi_{\ell_1}(\rho_x)) \rightarrow 0$.

(*Phase transition*) With the notation R_η from Theorem 9.1, we define the parameter

$$C(\rho) := \frac{\sqrt{8R_\eta} + 2}{\psi'_{\ell_1}(\rho)} \quad (\text{B.1})$$

The parameter $C(\rho)$ is well defined and $C(\rho) > 0$ because ψ_{ℓ_1} is a strictly increasing analytic function by Lemma 6.15. Moreover, the definition depends only on ρ and η .

Now we define $C_x := C(\rho_x)$ and $C_y := C(\rho_y)$, and suppose that s_x and s_y satisfy (10.5). From our error estimate (6.21) and the monotonicity of ψ_{ℓ_1} , we see

$$\begin{aligned} \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural})) &\leq d \left(\psi_{\ell_1}\left(\rho_x - \frac{C_x}{\sqrt{d}}\right) + \psi_{\ell_1}\left(\rho_y - \frac{C_y}{\sqrt{d}}\right) + \frac{2}{\sqrt{d}} \right) \\ &= d - \sqrt{d} \left(C_x \psi'_{\ell_1}(\rho_x) + C_y \psi'_{\ell_1}(\rho_y) - 2 \right) + O(1) \\ &= d - 2\sqrt{8R_\eta d} - 2\sqrt{d} + O(1). \end{aligned} \quad (\text{B.2})$$

The second equality follows by a Taylor series expansion of the analytic function ψ_{ℓ_1} , while the final equality is the definition of C_x and C_y .

For all d sufficiently large, we have the inequality

$$2\sqrt{8R_\eta d} \geq 16R_\eta + \sqrt{8R_\eta d} \geq \beta,$$

where β is the bandwidth parameter from Theorem 9.1, and the final inequality also

requires the observation that $d > \omega_v^2$, where ω_v is defined in (9.3).

Applying this inequality to the bound (B.2) on the statistical dimension, we find the inequality

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural})) \leq d - \beta$$

whenever d is sufficiently large. The success condition of Theorem 9.1 implies that method (10.1) succeeds with probability at least $1 - \eta$ so long as d is large enough.

For s_x and s_y satisfying (10.6), the same argument as above shows that

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural})) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural})) \geq d + \beta$$

for all d sufficiently large. The failure condition of Theorem 9.1 then implies that (10.1) succeeds with probability at most η . \square

B.1.2 Strong bound

This section provides strong guarantees for the demixing method (10.1) under the random alignment model.

Proof of Theorem 10.3. For any ρ , we define the sparse family

$$\mathcal{S}(\rho) := \{\mathbf{x} \in \mathbb{R}^d : \text{nnz}(\mathbf{x}) \leq \rho d\}.$$

By the optimality condition (4.3), the method (10.1) succeeds for *every* two constituent vectors $\mathbf{x}^{\natural}, \mathbf{y}^{\natural} \in \mathcal{S}(\rho)$ if and only if

$$\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural}) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^{\natural}) = \{\mathbf{0}\} \quad \text{for all } \mathbf{x}^{\natural}, \mathbf{y}^{\natural} \in \mathcal{S}(\rho). \quad (\text{B.3})$$

A priori, this condition appears to involve an infinite number of intersections. The following lemma reduces (B.3) to a manageably finite number of intersections.

Lemma B.1. *Suppose two vectors $\mathbf{x}^\natural, \mathbf{x}^\sharp \in \mathbb{R}^d$ satisfy*

$$\text{supp}(\mathbf{x}^\natural) \subset \text{supp}(\mathbf{x}^\sharp) \quad \text{and} \quad \text{sgn}(x_i^\natural) = \text{sgn}(x_i^\sharp) \quad \text{for all } i = 1, \dots, d-1, d.$$

Then $\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \subset \mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\sharp)$.

The proof of this lemma appears after we complete the current demonstration. Lemma B.1 implies that we only need to check the intersection of the minimally sparse vectors in our family $\mathcal{S}(\rho)$ in (B.3). In other words, condition (B.3) is equivalent to the statement

$$\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural) = \{\mathbf{0}\} \quad \text{for all } \mathbf{x}^\natural, \mathbf{y}^\natural \in \mathcal{S}_0(\rho), \quad (\text{B.4})$$

where $\mathcal{S}_0(\rho)$ is the reduced sparse family

$$\mathcal{S}_0(\rho) := \{\mathbf{x} : \text{nnz}(\mathbf{x}) = \lfloor \rho d \rfloor \quad \text{and} \quad x_i \in \{0, \pm 1\}\}. \quad (\text{B.5})$$

By applying the union bound, we find an inequality for the failure probability:

$$\begin{aligned} & \mathbb{P}\{\text{(B.4) does not hold}\} \\ & \leq |\mathcal{S}_0(\rho)|^2 \times \max_{\mathbf{x}^\natural, \mathbf{y}^\natural \in \mathcal{S}_0(\rho)} \mathbb{P}\{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural) \neq \{\mathbf{0}\}\}. \end{aligned} \quad (\text{B.6})$$

The remainder of the proof demonstrates that we can drive the upper bound (B.6) to η by choosing $\rho > 0$ sufficiently small and d sufficiently large.

Define the strong probability $\eta_s := \eta |\mathcal{S}_0|^{-2}$. The number of elements in \mathcal{S}_0 is precisely equal to the number of number of ways to choose $\lfloor \rho d \rfloor$ nonzero locations from d indices and then assign all $\lfloor \rho d \rfloor$ sign patterns to these nonzero locations:

$$|\mathcal{S}_0| = 2^{\lfloor \rho d \rfloor} \binom{d}{\lfloor \rho d \rfloor}.$$

A sharp-remainder form of Stirling's formula [DLMF, OLBC10, Sec. 5.6.1] provides a

bound for η_s , valid uniformly over $\rho \in (0, 1)$:

$$\frac{1}{d} \log(\eta_s) = \frac{1}{d} \log(\eta) - 2(\rho \log(2) + H(\rho)) + O\left(\frac{\log(d)}{d}\right) \quad \text{as } d \rightarrow \infty, \quad (\text{B.7})$$

where $H(\rho)$ is the base-e bit-entropy given by

$$H(\rho) := \rho \log\left(\frac{1}{\rho}\right) + (1 - \rho) \log\left(\frac{1}{1 - \rho}\right). \quad (\text{B.8})$$

By symmetry of the ℓ_1 norm, we have $\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)) = \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural))$ for any $\mathbf{x}^\natural, \mathbf{y}^\natural \in \mathcal{S}_0(\rho)$. By Proposition 6.14, we have

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural)) \leq 2d\psi_{\ell_1}(\rho). \quad (\text{B.9})$$

Each probability appearing in (B.6) is precisely equal to the probability that the demixing method (10.1) succeeds for a single pair of $\lfloor \rho d \rfloor$ sparse vectors \mathbf{x}^\natural and \mathbf{y}^\natural by the optimality condition (4.3). By combining Theorem 9.1 with the estimate (B.9), we find that $\mathbb{P}\{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural) \neq \{\mathbf{0}\}\} \leq \eta_s$ whenever

$$2\psi_{\ell_1}(\rho) + \frac{1}{d} \left(16R_{\eta_s} + \sqrt{8R_{\eta_s}} \omega_\vee\right) \leq 1 \quad (\text{B.10})$$

where the parameters $R_{\eta_s} := \log(8/\eta_s)$ and $\omega_\vee := \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)) \vee \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural))$ are drawn from the statement of Theorem 9.1.

The expression (B.7) for the strong probability η_s shows that

$$\frac{1}{d} R_{\eta_s} = O\left(\rho + H(\rho) + \frac{1}{d} \log\left(\frac{1}{\eta}\right) + \frac{\log(d)}{d}\right) \quad \text{as } d \rightarrow \infty.$$

Basic calculus shows that $H(\rho) \rightarrow 0$ as $d \rightarrow \infty$, so we can force R_{η_s} arbitrarily close to zero by taking d sufficiently large and ρ sufficiently small.

Upon applying Proposition 6.14 to the definition (9.3) of ω_\vee , we find that $\omega_\vee \leq \sqrt{d\psi_{\ell_1}(\rho)}$. Since $\psi_{\ell_1}(\rho) \rightarrow 0$ as $\rho \rightarrow 0$ by Lemma 6.15, we find that every term in (B.10) can be made arbitrarily small for sufficiently small ρ and large d . Theorem 9.1

then implies that for any $\mathbf{x}^\natural, \mathbf{y}^\natural \in \mathcal{S}_0(\rho)$, we have

$$|\mathcal{S}_0|^2 \times \mathbb{P}\{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural) \neq \{\mathbf{0}\}\} \leq |\mathcal{S}_0|^2 \eta_S = \eta$$

when $\rho > 0$ is small enough and d is large enough. Since the quantity above is an upper bound on the failure probability (B.6), the proof is complete. \square

Proof of Lemma B.1. By the invariance of the ℓ_1 norm under signed permutations, we may assume that $\mathbf{x}^\natural, \mathbf{x}^\sharp \in \mathbb{R}_+^d$ and that the support sets are given by

$$\text{supp}(\mathbf{x}^\natural) = \{1, \dots, i-1, i\} \subset \{1, \dots, j-1, j\} = \text{supp}(\mathbf{x}^\sharp)$$

for some integers $i \leq j$. From the expression (6.24) for the subgradient of the ℓ_1 norm at these sparse vectors, we have the subgradient inclusion

$$\partial \|\mathbf{x}^\natural\|_{\ell_1} \supset \partial \|\mathbf{x}^\sharp\|_{\ell_1}.$$

By (3.29), the normal cones inherit this inclusion property:

$$\mathcal{N}(S^\natural, \mathbf{x}^\natural) \supset \mathcal{N}(S^\sharp, \mathbf{x}^\sharp),$$

where $S^\natural := \{\mathbf{y} : \|\mathbf{y}\|_{\ell_1} \leq \|\mathbf{x}^\natural\|_{\ell_1}\}$ and $S^\sharp := \{\mathbf{y} : \|\mathbf{y}\|_{\ell_1} \leq \|\mathbf{x}^\sharp\|_{\ell_1}\}$ are the sublevel sets of the ℓ_1 norm at \mathbf{x}^\natural and \mathbf{x}^\sharp . Since a normal cone is polar to a descent cone (3.27), we have

$$\overline{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)} \subset \overline{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\sharp)},$$

where we have applied the bipolar property (Proposition 3.2.3) and the fact that inclusion reverses under polarity. But the ℓ_1 norm is a polyhedral function, so the descent cones above are generated by a polyhedral set (3.26), and in particular, both descent cones are closed. Hence, the displayed equation above is equivalent to the statement $\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \subset \mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\sharp)$. \square

B.1.3 Undersampled bounds

Finally, we provide a proof that a phase transition occurs in the undersampled sparse + sparse demixing model of Section 10.1.2.

Proof of Theorem 10.4. The proof follows the schema of the proof of Theorem 10.1 above, so we move briskly. The fact that there always exist solutions to (10.10) follows from the fact that $\psi_{\ell_1}(\rho) \rightarrow 0$ as $\rho \rightarrow 0$.

Let $\rho_x, \rho_y > 0$ satisfy (10.10), and suppose that the sparsity levels s_x and s_y satisfy the (10.11). With $C(\rho)$ defined as in (B.1), we follow the approach leading up to (B.2) to find

$$\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)) + \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural)) \leq m - 2\sqrt{8R_\eta d} - 2\sqrt{d} + O(1).$$

As above, we find that for sufficiently large d

$$2\sqrt{8R_\eta d} \geq 16R_\eta + \sqrt{8R_\eta d} \geq 16R_\eta + \sqrt{8R_\eta} \omega_\vee,$$

where $\omega_\vee^2 = \max\{\delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)), \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{y}^\natural))\} < d$. Hence, when d is sufficiently large, the total statistical dimension satisfies the conditions of Theorem 9.2 with $n = 2$. This is the first implication. The second implication follows in the same manner as the first. We do not repeat the details. \square

B.2 Sparse + sign demixing

This section provides the proofs of Theorems 10.5 and 10.6 that establish rigorous guarantees for sparse + sign demixing.

B.2.1 Phase transition

The proof below establishes that sparsity-parameterized phase transition occurs in sparse + sign demixing.

Proof of Theorem 10.5. The proof follows along the lines of the proof of Theorem 10.1.

The equation $\psi_{\ell_1}(\rho_*) = \frac{1}{2}$ uniquely determines ρ_* because Lemma 6.15 implies that $\psi_{\ell_1}(\rho)$ is a bijection from $(0, 1)$ to $(0, 1)$. Let $C(\rho)$ be as defined in (B.1), and set the parameter $C := C(\rho_*)$. Because ρ_* is determined uniquely regardless of the problem parameters, C is a universal constant independent of everything else.

Suppose $s := \text{nnz}(\mathbf{x}^\natural)$ satisfies (10.15). Following the same steps that lead to (B.2), we find

$$\begin{aligned} \delta(\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural)) &\leq \frac{d}{2} - \sqrt{d}(C\psi_{\ell_1}(\rho_*) - 1) + O(1) \\ &= \frac{d}{2} - 2\sqrt{8R_\eta d} - 2\sqrt{d} + O(1). \end{aligned}$$

As above, we find that the inequality $2\sqrt{8R_\eta d} \geq 16R_\eta + \sqrt{8R_\eta}\omega_\nu$ holds for all sufficiently large d . From the fact that $\delta(\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{y}^\natural)) = d/2$ (Proposition 6.11), Theorem 9.1 implies that (10.13) succeeds with probability at least $(1-\eta)$ for sufficiently large d . This completes the demonstration of the first point. The second point follows in a completely analogous manner, so we omit the details. \square

B.2.2 Strong bounds

The next argument establishes strong guarantees for sparse + sign demixing.

Proof of Theorem 10.6. We follow the approach of the proof of the sparse + sparse strong bound in Section B.1 above. Define the sparse family

$$\mathcal{S}(\rho) := \{\mathbf{x} \in \mathbb{R}^d : \text{nnz}(\mathbf{x}) \leq \rho d\}.$$

Then (10.13) succeeds for all $\mathbf{x}^\natural \in \mathcal{S}(\rho)$ if and only if

$$\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^\natural) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{y}^\natural) = \{\mathbf{0}\} \quad \text{for all } \mathbf{x}^\natural \in \mathcal{S}(\rho).$$

It follows from Lemma B.1 that this is equivalent to

$$\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural}) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{y}^{\natural}) = \{\mathbf{0}\} \quad \text{for all } \mathbf{x}^{\natural} \in \mathcal{S}_0(\rho). \quad (\text{B.11})$$

where $\mathcal{S}_0(\rho)$ is the reduced sparse family defined in (B.5). The union bound shows that the probability that the equation above does not hold is bounded above by

$$\mathbb{P}\{\text{(B.11) does not hold}\} \leq |\mathcal{S}_0(\rho)| \times \max_{\mathbf{x}^{\natural} \in \mathcal{S}_0(\rho)} \mathbb{P}\{\mathcal{D}(\|\cdot\|_{\ell_1}, \mathbf{x}^{\natural}) \cap -\mathbf{Q}\mathcal{D}(\|\cdot\|_{\ell_\infty}, \mathbf{y}^{\natural}) \neq \{\mathbf{0}\}\}.$$

We can force the bound above to be less than η for some $\rho > 0$ and all sufficiently large d in precisely the same manner as in the proof of Theorem 10.3. We omit the repetitive details. \square

Appendix C

Numerical details

In this chapter, we describe the details of our numerical procedures for computing the approximate statistical dimensions ψ_{ℓ_1} and ψ_{S_1} of Sections 6.2.6 and 6.2.7, as well as the numerical experiments from Chapter 10. All code and experiments utilize the MATLAB environment for scientific computing.

C.1 Statistical dimension calculations

Propositions 6.14 and 6.16 define two functions ψ_{ℓ_1} and ψ_{S_1} implicitly. We determine the value of these functions numerically using a two-step procedure. First, we solve the stationary equations (6.23) and (6.30) using the numerical root finding procedure `fzero`. The function `erfc` is used to compute the value of the exponential integral in (6.23), while we use `quadgk` for the integral appearing in (6.30).

With the stationary points in hand, we compute (6.22) and (6.29) using the `erfc` function or the numeric quadrature function `quadgk`. This procedure appears highly stable for values of ρ in the region $10^{-5} < \rho < 1 - 10^{-5}$.

C.2 Experimental generalities

All of the actual optimization is performed with the MATLAB package CVX [GB08, GB10] with the default settings. When drawing random rotations from O_d , we follow the simple, stable approach described in [Mez07]. All random variables are drawn independently

of all others.

The colormaps in our figures indicate the empirical probability of success of the experiments. In each case, black indicates total failure, while white indicates complete success. The varying shades of gray indicate mixtures of success and failures. We use the contour function to generate all empirical success contour lines.

We generate the phase transition curves using the method described in the previous section for computing ψ_{ℓ_1} and ψ_{S_1} . The formulas for these curves accompany the corresponding discussion in Section 10.

C.3 Sparse + sparse

This section describes the experiment shown in Figures 1.1 and 10.1 for demixing two sparse vectors. We fix the ambient dimension $d = 100$, and for each sparsity pair $(s_x, s_y) \in \{0, \dots, 99, 100\}^2$, we perform the following steps 25 times:

1. Draw $\mathbf{x}^\natural, \mathbf{y}^\natural \in \mathbb{R}^{100}$ with s_x and s_y nonzero elements. The locations and signs of the nonzero elements are random, but their magnitude is fixed to one.
2. Draw a random rotation $\mathbf{Q} \in O_d$.
3. Solve the convex optimization program (10.1) with $\mathbf{z}_0 = \mathbf{x}^\natural + \mathbf{Q}\mathbf{y}^\natural$. Denote the optimal pair $\mathbf{x}_*, \mathbf{y}_*$.
4. Declare success if $\|\mathbf{x}_* - \mathbf{x}^\natural\|_{\ell_\infty} < 10^{-4}$.

C.3.1 Undersampled sparse + sparse

In the case of Figure 10.2, we set the ambient dimension $d = 200$. For each number of measurements $m \in \{25, 50, 75, 100\}$ and sparsity levels $s_x, s_y \in \{1, \dots, d - 1, d\}$, we repeat the following procedure 35 times:

1. If $s_x + s_y > m$, declare failure because the number of degrees of freedom in the signal outstrips the number of measurements.
2. Otherwise, draw $\mathbf{x}^\natural, \mathbf{y}^\natural \in \mathbb{R}^d$ as above.

3. Draw a random rotation $\mathbf{Q} \in \mathcal{O}_d$ and a Gaussian matrix $\mathbf{G} \in \mathbb{R}^{m \times d}$.
4. Solve (10.7) with $\mathbf{z}_0 := \mathbf{G}(\mathbf{x}^{\natural} + \mathbf{Q}\mathbf{y}^{\natural})$ for an optimal pair $(\mathbf{x}_*, \mathbf{y}_*)$.
5. Declare success if $\|\mathbf{x}^{\natural} - \mathbf{x}_*\|_{\ell_\infty} < 10^{-5}$.

C.4 Sparse + sign

We now relate the details of the experiment shown in Figure 10.1. For each dimension $d = 100$ and $d = 300$, we take 70 equally spaced sparsity values $s \in [0, 0.35d]$, and we repeat the following procedure 200 times:

1. Draw a sparse vector \mathbf{x}^{\natural} with s nonzero elements that have magnitude one but have random locations and signs.
2. Choose a sign vector $\mathbf{y}^{\natural} \in \{\pm 1\}^d$ uniformly at random.
3. Generate a random rotation $\mathbf{Q} \in \mathcal{O}_d$ independently from all of the others.
4. Solve (10.13) for the optimal pair $(\mathbf{x}_*, \mathbf{y}_*)$.
5. Declare success if $\|\mathbf{y}_* - \mathbf{y}^{\natural}\|_{\ell_\infty} < 10^{-4}$.

C.4.1 Sparse + sparse + sign

The experiment of Figure 10.4 involves demixing three vectors. The dimension $d = 100$, and for each $s_1, s_2 \in \{1, \dots, d-1, d\}$, we repeat the following procedure 35 times:

1. Generate sparse vectors \mathbf{x}^{\natural} and \mathbf{y}^{\natural} with s_1 and s_2 nonzero elements, in the same manner as described in Section C.3, and choose $\mathbf{w} \in \{\pm 1\}^d$ uniformly at random.
2. Draw two random rotations \mathbf{Q}_y and \mathbf{Q}_w and generate the observation $\mathbf{z}_0 = \mathbf{x}^{\natural} + \mathbf{Q}_y\mathbf{y}^{\natural} + \mathbf{Q}_w\mathbf{w}^{\natural}$ and the side information $\alpha := \|\mathbf{y}\|_{\ell_1}$.
3. Solve (10.16) for the optimal tuple $(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*)$.
4. Declare success if $\|\mathbf{x}_1^* - \mathbf{x}_1^{\natural}\|_{\ell_\infty} < 10^{-5}$ and $\|\mathbf{x}_2^* - \mathbf{x}_2^{\natural}\|_{\ell_\infty} < 10^{-5}$.

C.5 Sparse + low-rank

We now describe the experiment that leads to Figure 10.5. We fix the sidelength $n = 35$, and for each rank $r = 1, \dots, n - 1, n$ and sparsity level $s = 1, n, \dots, n^2 - n + 1, n^2$, we repeat the following procedure 25 times:

1. Draw a low-rank matrix $\mathbf{X}^\natural = \mathbf{Q}_L \mathbf{\Lambda} \mathbf{Q}_R \in \mathbb{R}^{n \times n}$ with rank r , where $\mathbf{\Lambda}$ is a diagonal matrix that satisfies $\Lambda_{ii} = 1$ for $i = 1, \dots, r - 1, r$ and $\Lambda_{ii} = 0$ otherwise. The left and right operators \mathbf{Q}_L and \mathbf{Q}_R are random rotations in \mathcal{O}_d that determine the singular vectors of \mathbf{X}^\natural .
2. Generate a sparse matrix \mathbf{Y}^\natural with s nonzero entries with random signs at random locations. The magnitude of the nonzero entries is equal to one.
3. Generate a random rotation \mathcal{Q} for the matrix space $\mathbb{R}^{d \times d}$ and set the observation $\mathbf{Z}_0 = \mathbf{X}_0 + \mathcal{Q}(\mathbf{Y}_0)$.
4. Solve (10.18) for the optimal points $(\mathbf{X}_*, \mathbf{Y}_*)$.
5. Declare success if $\|\mathbf{X}_* - \mathbf{X}^\natural\|_{\ell_\infty} < 10^{-4}$.

C.6 Linear inverse problems

This section describes the two linear inverse problems from Figures 10.6 and 10.7. Both figures appear in the joint work [ALMT13]. In the compressed sensing example (left panel), the ambient dimension is fixed at $d = 100$. For each number of measurements $m = 1, \dots, d - 2, d - 1$ and each $s = 1, \dots, d - 2, d - 1$, we repeat the following 50 times:

1. Draw a sparse vector \mathbf{x}^\natural with s nonzero entries whose locations and signs are random, but whose magnitudes are all equal to one.
2. Generate a standard Gaussian matrix $\mathbf{G} \in \mathbb{R}^{m \times d}$ independently of \mathbf{x}^\natural and set $\mathbf{z}_0 = \mathbf{G}\mathbf{x}_0$.
3. Solve (10.19) for the optimal point \mathbf{x}_* .
4. Declare success if $\|\mathbf{x}_* - \mathbf{x}^\natural\| < 10^{-5}$.

Figure 10.6 shows the empirical probability that this method succeeds.

For the experiment of Figure 10.7, we fix the sidelength $n = 30$. For each rank $r = 1, \dots, 29, 30$ and each number of measurements $m = 1, 29, \dots, n^2 - n + 1, n^2$, we repeat the following procedure 50 times:

1. If $r \geq \lceil \sqrt{m} \rceil + 1$, we declare failure and move on to the next case.
2. Draw a rank- r matrix $\mathbf{X}^{\natural} = \mathbf{Q}_L \mathbf{Q}_R^t$, where the matrices \mathbf{Q}_L and \mathbf{Q}_R are drawn from the Haar measure on the Stiefel manifold on the set of $n \times r$ orthogonal matrices.
3. Draw a Gaussian measurement matrix $\mathbf{G} \in \mathbb{R}^{m \times n^2}$ and set $\mathbf{z}_0 = \mathbf{G} \text{vec}(\mathbf{X}^{\natural})$.
4. Solve (10.21) for the optimal point \mathbf{X}_* .
5. Declare success if $\|\mathbf{X}_* - \mathbf{X}^{\natural}\| < 10^{-5}$.

Bibliography

- [AB11] Dennis Amelunxen and Peter Bürgisser. Probabilistic analysis of the Grassmann condition number. arXiv:1112.2603, 2011.
- [AB12a] Dennis Amelunxen and Peter Bürgisser. A coordinate-free condition number for convex programming. *preprint*, June 2012. arXiv:1105.4049v1.
- [AB12b] Dennis Amelunxen and Peter Bürgisser. Intrinsic volumes of symmetric cones. *preprint*, 2012. arXiv:1205.1863.
- [AEJ⁺12] Amir Adler, Valentin Emiya, Maria G. Jafari, Michel Elad, Rémi Gribonval, and Mark D. Plumbley. Audio inpainting. *IEEE Trans. Audio, Speech, and Lang. Process.*, 20(3):922–932, March 2012.
- [AF03] Mark J. Ablowitz and Athanassios S. Fokas. *Complex variables: introduction and applications*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, second edition, 2003.
- [ALMT13] Dennis Amelunxen, Martin Lotz, Michael B. McCoy, and Joel A. Tropp. Living on the edge: A geometric theory of phase transitions in convex optimization. *preprint*, March 2013. arXiv:1303.6672.
- [Ame11] Dennis Amelunxen. *Geometric analysis of the condition of the convex feasibility problem*. Dissertation, Universität Paderborn, 2011.
- [And84] Theodore W. Anderson. *An introduction to multivariate statistical analysis*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, second edition, 1984.

- [AS92] Fernando Affentranger and Rolf Schneider. Random projections of regular simplices. *Discrete Comput. Geom.*, 7(3):219–226, 1992.
- [BDB07] Petros Boufounos, Marco F. Duarte, and Richard G. Baraniuk. Sparse signal reconstruction from noisy compressive measurements using cross validation. In *IEEE/SP 14th Workshop on Statistical Signal Processing*, pages 299–303, August 2007.
- [Bec89] William Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.*, 105(2):397–400, 1989.
- [BH93] Ulrich Betke and Martin Henk. Intrinsic volumes and lattice points of crosspolytopes. *Monatsh. Math.*, 115(1-2):27–33, 1993.
- [BH99] Károly Böröczky, Jr. and Martin Henk. Random projections of regular polytopes. *Arch. Math. (Basel)*, 73(6):465–473, December 1999.
- [Bha97] Rajendra Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
- [BM99] Albert P. Berg and Wasfy B. Mikhael. A survey of mixed transform techniques for speech and image coding. In *IEEE Trans. Circuits Syst.*, volume 4, pages 106–109 vol.4, 1999.
- [BMS06] Jérôme Bobin, Yassir Moudden, and Jean-Luc Starck. Morphological diversity and source separation. *IEEE Trans. Signal Process.*, 13(7):409–412, 2006.
- [Bog98] Vladimir I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.
- [BPC⁺10] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed Optimization and Statistical Learning via the Alternating

Direction Method of Multipliers. *Foundations and Trends[®] in Machine Learning*, 3(1):1–122, 2010.

- [BS10] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.
- [BSFM07] Jérôme Bobin, Jean-Luc Starck, Jalal Fadili, and Yassir Moudden. Sparsity and morphological diversity in blind source separation. *IEEE Trans. Image Process.*, 16(11):2662–2674, November 2007.
- [BT09] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.
- [BT12] Amir Beck and Marc Teboulle. Smoothing and first order methods: a unified framework. *SIAM J. Optim.*, 22(2):557–580, 2012.
- [CDS99] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20(1):33–61, 1999.
- [CJSC11a] Yudong Chen, Ali Jalali, Sujay Sanghavi, and Constantine Caramanis. Clustering partially observed graphs via convex optimization. In *International Symposium on Information Theory (ISIT)*, 2011.
- [CJSC11b] Yudong Chen, Ali Jalali, Sujay Sanghavi, and Constantine Caramanis. Low-rank matrix recovery from errors and erasures. *preprint*, 2011. arXiv:1104.0354.
- [CJSC11c] Yudong Chen, Ali Jalali, Sujay Sanghavi, and Constantine Caramanis. Low-rank matrix recovery from errors and erasures. In *International Symposium on Information Theory (ISIT)*, pages 2313 –2317, August 2011.

- [CJSC13] Yudong Chen, Ali Jalali, Sujay Sanghavi, and Constantine Caramanis. Low-rank matrix recovery from errors and erasures. *IEEE Trans. Inform. Theory.*, 2013. To appear.
- [CLMW11] Emmanuel J. Candès, Xiadong Li, Yi Ma, and John Wright. Robust principal component analysis? *J. Assoc. Comput. Mach.*, 58(3):1–37, May 2011.
- [CM73] Jon F. Claerbout and Francis Muir. Robust modeling with erratic data. *Geophysics*, 38(5):826–844, 1973.
- [CPW10] Venkat Chandrasekaran, Pablo A. Parrilo, and Alan S. Willsky. Latent variable graphical model selection via convex optimization. In *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1610–1613, October 2010.
- [CRPW12] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The convex geometry of linear inverse problems. *Found. Comput. Math.*, 12(6):805–849, 2012.
- [CRT06a] Emmanuel J. Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.
- [CRT06b] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, 2006.
- [CS12] Emmanuel J. Candès and Mahdi Soltanolkotabi. Discussion: Latent variable graphical model selection via convex optimization. *Ann. Stat.*, 40(4):1997–2004, 2012.
- [CSPW09] Venkat Chandrasekaran, Sujay Sanghavi, Pablo A. Parrilo, and Alan S. Willsky. Sparse and low-rank matrix decompositions. In *SYSID 2009*, Saint-Malo, France, July 2009.

- [CSPW11] Venkat Chandrasekaran, Sujay Sanghavi, Pablo A. Parrilo, and Alan S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM J. Optim.*, 21(2):572–596, 2011.
- [CT05] Emmanuel J. Candès and Terence Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51(12):4203–4215, 2005.
- [CW05] Patrick L. Combettes and Valérie R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, 4(4):1168–1200 (electronic), 2005.
- [DE03] David L. Donoho and Michael Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization. *Proc. Natl. Acad. Sci. USA*, 100(5):2197–2202 (electronic), 2003.
- [DH01] David L. Donoho and Xiaoming Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inform. Theory*, 47(7):2845–2862, August 2001.
- [DLMF] NIST Digital Library of Mathematical Functions. <http://dlmf.nist.gov/>, Release 1.0.5 of 2012-10-01. Online companion to [OLBC10].
- [DMA97] Geoffrey Davis, Stéphane G. Mallat, and Marco Avellaneda. Adaptive greedy approximations. *Constructive Approximation*, 13(1):57–98, 1997.
- [Don04] David L. Donoho. Neighborly polytopes and sparse solutions of underdetermined linear equations. Technical report, Stanford University, 2004.
- [Don06a] David L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.
- [Don06b] David L. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete Comput. Geom.*, 35(4):617–652, 2006.

- [DS89] David L. Donoho and Philip B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 49(3):906–931, June 1989.
- [DT96] Ronald A. DeVore and Vladimir N. Temlyakov. Some remarks on greedy algorithms. *Adv. Comput. Math.*, 5(2-3):173–187, 1996.
- [DT05] David L. Donoho and Jared Tanner. Neighborliness of randomly projected simplices in high dimensions. *Proc. Natl. Acad. Sci. USA*, 102(27):9452–7, July 2005.
- [DT09] David L. Donoho and Jared Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.*, 22(1):1–53, 2009.
- [DT10a] David L. Donoho and Jared Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete Comput. Geom.*, 43(3):522–541, 2010.
- [DT10b] David L. Donoho and Jared Tanner. Exponential bounds implying construction of compressed sensing matrices, error-correcting codes, and neighborly polytopes by random sampling. *IEEE Trans. Inform. Theory*, 56(4):2002–2016, 2010.
- [DT10c] David L. Donoho and Jared Tanner. Precise undersampling theorems. *Proc. IEEE*, 98(6):913–924, June 2010.
- [EB92] Jonathan Eckstein and Dimitri P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 55(3, Ser. A):293–318, 1992.
- [EB02] Michael Elad and Alfred M. Bruckstein. A generalized uncertainty principle and sparse representation in pairs of bases. *IEEE Trans. Inform. Theory*, 48(9):2558–2567, 2002.

- [ESQD05] Michael Elad, Jean-Luc Starck, Philippe Querre, and David L. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Appl. Comput. Harmon. Anal.*, 19(3):340–358, November 2005.
- [Faz02] Maryam Fazel. *Matrix rank minimization with applications*. Dissertation, Stanford University, Stanford, CA, 2002.
- [Fol99] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.
- [GB08] Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, London, 2008. http://stanford.edu/~boyd/graph_dcp.html.
- [GB10] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. <http://cvxr.com/cvx>, October 2010.
- [Gem80] Stuart Geman. A limit theorem for the norm of random matrices. *Ann. Probab.*, 8(2):252–261, 1980.
- [GHS03] Fuchang Gao, Daniel Hug, and Rolf Schneider. Intrinsic Volumes and Polar Sets in Spherical Space. *Math. Notae*, 41:159–176, 2003.
- [Gla95] Stefan Glasauer. *Integralgeometrie konvexer Körper im sphärischen Raum-rischen Raum*. Dissertation, University of Freiburg, 1995.
- [Gla96] Stefan Glasauer. Integral geometry of spherically convex bodies. *Diss. Summ. Math.*, 1(1-2):219–226, 1996.

- [GLF⁺10] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, Oct 2010.
- [GM75] Roland Glowinski and Americo Marrocco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires. *Rev. Française Automat. Informat. Recherche Opérationnelle RAIRO Analyse Numérique*, 9(R-2):41–76, 1975.
- [GM76] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Comp. Math. Appl.*, 2:17–40, 1976.
- [GM77] Richard A. Groeneveld and Glen Meeden. The mode, median, and mean inequality. *Amer. Statist.*, 31(3):120–121, 1977.
- [Gor85] Yehoram Gordon. Some inequalities for Gaussian processes and applications. *Israel J. Math.*, 50(4):265–289, 1985.
- [Gor87] Yehoram Gordon. Elliptically contoured distributions. *Probab. Theory Related Fields*, 76(4):429–438, 1987.
- [Gor88] Yehoram Gordon. On Milman's inequality and random subspaces which escape through a mesh in \mathbf{R}^n . In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 84–106. Springer, Berlin, 1988.
- [Gro11] David Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inform. Theory*, 57(3):1548–1566, 2011.
- [Had75] Hugo Hadwiger. Das Wills'sche Funktional. *Monatsh. Math.*, 79:213–221, 1975.

- [HB12] Chinmay Hegde and Richard G. Baraniuk. Signal recovery on incoherent manifolds. *IEEE Trans. Inform. Theory*, 58(12):7204–7214, December 2012.
- [HL13] Mingyi Hong and Zhi-Quan Luo. On the linear convergence of the alternating direction method of multipliers. *preprint*, 2013. arXiv:1208.3922.
- [HW74] Hugo Hadwiger and Jörg M. Wills. Gitterpunktanzahl konvexer Rotationkörper. *Math. Ann.*, 208:221–232, 1974.
- [JRSR10] Ali Jalali, Pradeep Ravikumar, Sujay Sanghavi, and Chao Ruan. A dirty model for multi-task learning. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 964–972. NIPS, 2010.
- [Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [Kle55] Victor L. Klee, Jr. Separation properties of convex cones. *Proc. Amer. Math. Soc.*, 6(2):313–318, 1955.
- [KP13] Steven G. Krantz and Harold R. Parks. *The implicit function theorem*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. History, theory, and applications, Reprint of the 2003 edition.
- [KR97] Daniel A. Klain and Gian-Carlo Rota. *Introduction to geometric probability*. Lezioni Lincee. [Lincei Lectures]. Cambridge University Press, Cambridge, 1997.
- [KX09] Amin Khajehnejad and Weiyu Xu. Weighted ℓ_1 minimization for sparse recovery with prior information. In *International Symposium on Information Theory (ISIT)*, pages 483–487, 2009.
- [KXAH10] Amin Khajehnejad, Weiyu Xu, A. Salman Avestimehr, and Babak Hassibi. Improved sparse recovery thresholds with two-step reweighted ℓ_1 mini-

- mization. In *International Symposium on Information Theory (ISIT)*, pages 1603–1607, Austin, TX, 2010.
- [KXAH11] Amin Khajehnejad, Weiyu Xu, A. Salman Avestimehr, and Babak Hassibi. Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models. *IEEE Trans. Signal Process.*, 59(5):1985–2001, 2011.
- [Lan86] Robert Lang. A note on the measurability of convex sets. *Arch. Math. (Basel)*, 47(1):90–92, 1986.
- [LB10] Kiryung Lee and Yoram Bresler. ADMiRA: atomic decomposition for minimum rank approximation. *IEEE Trans. Inform. Theory*, 56(9):4402–4416, 2010.
- [LVBL98] Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284(1-3):193–228, 1998. ILAS Symposium on Fast Algorithms for Control, Signals and Image Processing (Winnipeg, MB, 1997).
- [Mal09] Stéphane G. Mallat. *A wavelet tour of signal processing*. Elsevier/Academic Press, Amsterdam, third edition, 2009. The sparse way, With contributions from Gabriel Peyré.
- [McM75] Peter McMullen. Non-linear angle-sum relations for polyhedral cones and polytopes. *Math. Proc. Cambridge Philos. Soc.*, 78(02):247, October 1975.
- [McM93] Peter McMullen. Valuations and dissections. In *Handbook of convex geometry, Vol. A, B*, pages 933–988. North-Holland, Amsterdam, 1993.
- [Mez07] Francesco Mezzadri. How to generate random matrices from the classical compact groups. *Notices Amer. Math. Soc.*, 54(5):592–604, 2007.
- [MP67] Vladimir A. Marčenko and Leonid A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

- [MP97] Mehran Mesbahi and George P. Papavassilopoulos. On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Trans. Automat. Control*, 42(2):239–243, 1997.
- [MR11] Olvi L. Mangasarian and Benjamin Recht. Probability of unique integer solution to a system of linear equations. *European J. Oper. Res.*, 214(1):27–30, October 2011.
- [MS89] Wasfy B. Mikhael and Andreas S. Spanias. Accurate representation of time-varying signals using mixed transforms with applications to speech. *IEEE Trans. Circuits Syst. I. Regul. Pap.*, 36(2):329–331, 1989.
- [MT11] Michael B. McCoy and Joel A. Tropp. Two proposals for robust PCA using semidefinite programming. *Elec. J. Statist.*, 5:1123–1160, 2011.
- [MT12] Michael B. McCoy and Joel A. Tropp. Sharp recovery bounds for convex deconvolution, with applications. *preprint*, 2012. arXiv:1205.1580v1.
- [MZ93] Stéphane G. Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, 1993.
- [NN94] Yurii Nesterov and Arkadi Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.
- [Nos48] M. Nosarzewska. Évaluation de la différence entre l’aire d’une région plane convexe et le nombre des points aux coordonnées entières couverts par elle. *Colloquium Math.*, 1:305–311, 1948.
- [NT09] Deanna Needell and Joel A. Tropp. CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.*, 26(3):301–321, 2009.

- [NT13] Nam H. Nguyen and Trac D. Tran. Exact recoverability from dense corrupted observations via ℓ_1 -minimization. *IEEE Trans. Inform. Theory*, 59(4):2017–2035, 2013.
- [OH10] Samet Oymak and Babak Hassibi. New null space results and recovery thresholds for matrix rank minimization. *preprint*, 2010. arXiv:1011.6326.
- [OLBC10] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010. Print companion to [DLMF].
- [Ove75] Theo Overhagen. Zur Gitterpunktanzahl konvexer Körper im 3-dimensionalen euklidischen Raum. *Math. Ann.*, 216(3):217–224, 1975.
- [Pat00] Gábor Pataki. The geometry of semidefinite programming. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenbergh, editors, *Handbook of semidefinite programming: Theory, Algorithms, and Applications*, volume 27 of *Internat. Ser. Oper. Res. Management Sci.*, pages 29–65. Kluwer Acad. Publ., Boston, MA, 2000.
- [PBS13] Graeme Pope, Annina Bracher, and Christoph Studer. Probabilistic recovery guarantees for sparsely corrupted signals. *IEEE Trans. Inform. Theory* (to appear), 2013.
- [PGW⁺12] Yigang Peng, A. Ganesh, J. Wright, Wenli Xu, and Yi Ma. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2233–2246, 2012.
- [Pla72] Robin L. Plackett. Studies in the history of probability and statistics. XXIX. The discovery of the method of least squares. *Biometrika*, 59:239–251, 1972.
- [PRK93] Y.C. Pati, R. Rezaifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: recursive function approximation with applications to wavelet

- decomposition. In *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 40–44, 1993.
- [RFP10] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.
- [RKD98] Bhaskar D. Rao and Kenneth Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *Proceedings of the 8th IEEE Digital Signal Processing Workshop*, 1998.
- [RKD02] Bhaskar D. Rao and Kenneth Kreutz-Delgado. Basis selection in the presence of noise. *Signals, Systems & Computers*, 1:752–756, November 2002.
- [Roc70] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [Rub60] Harold Ruben. On the geometrical moments of skew-regular simplices in hyperspherical space, with some applications in geometry and mathematical statistics. *Acta Math.*, 103:1–23, 1960.
- [RV08] Mark Rudelson and Roman Vershynin. On sparse reconstruction from Fourier and Gaussian measurements. *Comm. Pure Appl. Math.*, 61(8):1025–1045, 2008.
- [RW98] R. Tyrrell Rockafellar and Roger J-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1998.
- [Saf88] Felix G. Safar. *Signal Compression and Reconstruction using Multiple Bases Representation*. PhD thesis, Virginia Polytechnic Institute and State University (VirginiaTech), 1988.
- [San76] Luis A. Santaló. *Integral geometry and geometric probability*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. With a

foreword by Mark Kac, *Encyclopedia of Mathematics and its Applications*, Vol. 1.

- [San80] Luis A. Santaló. Notes on the integral geometry in the hyperbolic plane. *Portugal. Math.*, 39(1-4):239–249 (1985), 1980. Special issue in honor of António Monteiro.
- [Sch50] Ludwig Schläfli. *Theorie der Vielfachen Kontinuität (1852), Gesammelte mathematische Abhandlungen. Band I*. Verlag Birkhäuser, Basel, 1950.
- [Sch60] Julian Schwinger. Unitary operator bases. *Proc. Nat. Acad. Sci. U.S.A.*, 46:570–579, 1960.
- [Sch93] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.
- [SDC03] Jean-Luc Starck, David L. Donoho, and Emmanuel J. Candès. Astronomical image representation by the curvelet transform. *Astronom. Astrophys.*, 398(2):785–800, 2003.
- [SED05] Jean-Luc Starck, Michael Elad, and David L. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. Image Process.*, 14(10):1570–1582, October 2005.
- [She73] Oscar B. Sheynin. R. J. Boscovich’s work on probability. *Arch. History Exact Sci.*, 9(4-5):306–324, 1972/73.
- [Sil85] Jack W. Silverstein. The smallest eigenvalue of a large-dimensional Wishart matrix. *Ann. Probab.*, 13(4):1364–1368, 1985.
- [SKPB12] Christoph Studer, Patrick Kuppinger, Graeme Pope, and Helmut Bölcskei. Recovery of sparsely corrupted signals. *IEEE Trans. Inf. Theory*, pages 3115–3130, May 2012.

- [Sto09] Mihailo Stojnic. Various thresholds for ℓ_1 -optimization in compressed sensing. *preprint*, 2009. arXiv:0907.3666.
- [Stu99] Jos F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12:625–653, 1999. Special issue on Interior Point Methods (CD supplement with software).
- [SW08] Rolf Schneider and Wolfgang Weil. *Stochastic and Integral Geometry*. Springer series in statistics: Probability and its applications. Springer, 2008.
- [Tay06] Jonathan E. Taylor. A Gaussian kinematic formula. *Ann. Probab.*, 34(1):122–158, 2006.
- [TBM79] Howard L. Taylor, Stephen C. Banks, and John F. McCoy. Deconvolution with the ℓ_1 norm. *Geophysics*, 44(1):39, 1979.
- [Tem03] Vladimir N. Temlyakov. Nonlinear methods of approximation. *Found. Comput. Math.*, 3(1):33–107, 2003.
- [TG07] Joel A. Tropp and Anna C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inform. Theory*, 53(12):4655–4666, 2007.
- [TGS06] Joel A. Tropp, Anna C. Gilbert, and Martin Strauss. Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Processing*, 86(572–588), April 2006.
- [Tro04] Joel A. Tropp. Greed is good: algorithmic results for sparse approximation. *IEEE Trans. Inform. Theory*, 50(10):2231–2242, 2004.
- [Tro08a] Joel A. Tropp. Norms of random submatrices and sparse approximation. *C. R. Math. Acad. Sci. Paris*, 346(23-24):1271–1274, 2008.

- [Tro08b] Joel A. Tropp. On the conditioning of random subdictionaries. *Appl. Comput. Harmon. Anal.*, 25(1):1–24, 2008.
- [Tro08c] Joel A. Tropp. On the linear independence of spikes and sines. *J. Fourier Anal. Appl.*, 14:838–858, 2008.
- [TTT99] Kim-Chuan Toh, Michael J. Todd, and Reha H. Tütüncü. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optim. Methods Softw.*, 11/12(1-4):545–581, 1999. Interior point methods.
- [TTT12] Kim-Chuan Toh, Michael J. Todd, and Reha H. Tütüncü. On the implementation and usage of SDPT3—a Matlab software package for semidefinite-quadratic-linear programming, version 4.0. In *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *Internat. Ser. Oper. Res. Management Sci.*, pages 715–754. Springer, New York, 2012.
- [VB96] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996.
- [Vit96] Richard A. Vitale. The Wills functional and Gaussian processes. *Ann. Probab.*, 24(4):2172–2178, 1996.
- [VS86] Anatoly M. Vershik and Piotr V. Sporyshev. An asymptotic estimate for the average number of steps in the parametric simplex method. *USSR Comput. Maths. Math. Phys.*, 26(3):104–113, 1986.
- [VS92] Anatoly M. Vershik and Piotr V. Sporyshev. Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem. *Selecta Math. Soviet.*, 11(2):181–201, 1992. Selected translations.
- [Wat92] G. Alistair Watson. Characterization of the subdifferential of some matrix norms. *Linear Algebra Appl.*, 170:33–45, 1992.
- [WGMM12] John Wright, Arvind Ganesh, Kerui Min, and Yi Ma. Compressive principal component pursuit. *preprint*, 2012. arXiv:1202.4596v1.

- [Wil73] Jörg M. Wills. Zur Gitterpunktanzahl konvexer Mengen. *Elem. Math.*, 28:57–63, 1973.
- [WM09] John Wright and Yi Ma. Dense error correction via l_1 -minimization. *IEEE Trans. Inform. Theory*, pages 3033–3036, April 2009.
- [WSB11] Andrew E. Waters, Aswin C. Sankaranarayanan, and Richard Baraniuk. SpaRCS: Recovering low-rank and sparse matrices from compressive measurements. In J. Shawe-Taylor, R.S. Zemel, P. Bartlett, F.C.N. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 1089–1097, 2011.
- [Wyn79a] Aaron D. Wyner. An analog scrambling scheme which does not expand bandwidth, Part I: Discrete time. *IEEE Trans. Inform. Theory*, 25(3):261–274, May 1979.
- [Wyn79b] Aaron D. Wyner. An analog scrambling scheme which does not expand bandwidth, Part II: Continuous time. *IEEE Trans. on Inform. Theory*, 25(4):415–425, 1979.
- [XCS10a] Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust PCA via outlier pursuit. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 2496–2504. NIPS, 2010.
- [XCS10b] Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust PCA via outlier pursuit. *IEEE Trans. Inform. Theory*, to appear, pages 1–24, 2010. arXiv:1010.4237.
- [XH08] Weiyu Xu and B. Hassibi. Compressed sensing over the grassmann manifold: A unified analytical framework. In *46th Annual Allerton Conference on Communication, Control, and Computing*, pages 562–567, sept. 2008.

- [XH11] Weiyu Xu and Babak Hassibi. Precise stability phase transitions for ℓ_1 minimization: A unified geometric framework. *IEEE Trans. Inform. Theory*, 57(10):6894–6919, October 2011.
- [XH12] Weiyu Xu and Babak Hassibi. Fundamental thresholds in compressed sensing: a high-dimensional geometry approach. In *Compressed sensing*, pages 305–347. Cambridge Univ. Press, Cambridge, 2012.
- [XKAH10] Weiyu Xu, Amin Khajehnejad, A. Salman Avestimehr, and Babak Hassibi. Breaking through the thresholds: an analysis for iterative reweighted ℓ_1 minimization via the Grassmann angle framework. In *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pages 5498–5501, March 2010.
- [ZLGM11] Zhengdong Zhang, Xiao Liang, Arvind Ganesh, and Yi Ma. TILT: Transform invariant low-rank textures. In Ron Kimmel, Reinhard Klette, and Akihiro Sugimoto, editors, *Computer Vision – ACCV 2010*, volume 6494 of *Lecture Notes in Computer Science*, pages 314–328. Springer Berlin Heidelberg, 2011.