TRACTABLE RELAXATIONS AND EFFICIENT ALGORITHMIC TECHNIQUES FOR LARGE-SCALE OPTIMIZATION

BY FATMA KILINÇ-KARZAN

A Thesis
Presented to
The Academic Faculty

by

Fatma Kılınç-Karzan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology
August 2011

TRACTABLE RELAXATIONS AND EFFICIENT ALGORITHMIC TECHNIQUES FOR LARGE-SCALE OPTIMIZATION
BY FATMA KILINÇ-KARZAN

Approved by:

Professor Arkadi Nemirovski, Advisor
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor George Nemhauser
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Anatoli Juditsky
Laboratoire Jean Kuntzmann
Universite Joseph Fourier
Grenoble Cedex, France

Professor William Cook
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Alexander Shapiro
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: 08 June 2011

DEDICATION

In the beloved memory of my grandmother...

ACKNOWLEDGEMENTS

My deepest and most sincere gratitude goes to my adviser, Prof. Arkadi S. Nemirovski. Without his devoted guidance, unwavering support and generous encouragement, this thesis would not exist. All along, I was inspired with his exquisite taste in research, enthusiasm, and optimism, stimulated with his clever questions in addition to his brilliant insights brought by experience and perseverance, and nurtured with his incredible kindness and words of wisdom in research and life, whenever I needed them most. I am particularly grateful for his incredible patience and understanding as I explored new frontiers in research, the freedom and flexibility he has given, and above all his trust in my abilities. Being his student, learning from and working with him have been truly my pleasure and such an enjoyable, unforgettable privilege that his influence on me as my adviser and a role model will be everlasting. I owe him a debt of gratitude more than I can express here.

I was extremely fortunate to work with Anatoli Juditsky, who has been a tremendous source of intriguing and challenging problems. I am obliged for many interesting discussions and insights, and the resulting congenial collaborative atmosphere as well as his help in my development of a more mature researcher. I would like to thank George Nemhauser for his guidance, and financial support at the early stages of my PhD, I have learned a great deal from him. I was especially lucky to have William Cook and Alexander Shapiro as my other committee members, whose encouragement and candid advice were instrumental not only for my success during my studies, but also for steering me in the right direction in my academic pursuits and career.

In addition to my thesis committee, there have been many outstanding faculty members at Georgia Tech, who has taught me, guided me and shaped my academic life in various different ways. In particular, I would like to acknowledge Shabbir Ahmed and Martin Savelsbergh, for their collaboration on several rewarding projects and unique opportunities they have provided, Ellis Johnson, for lecturing one of the best and my most memorable special

topics class on optimization, Anton Kleywegt, for teaching me early on the importance of rigor, clarity, and attention to detail, Pınar Keskinocak, Özlem Ergün, David Goldsman, and Santanu Dey, for their remarkable kindness, availability and continuous moral support all along, and Gary Parker, Michael Thomas, and Chen Zhou, for their assistance throughout my PhD and my job search.

There are many people whose friendship, support, and encouragement have been immensely memorable. In particular, I sincerely thank to Eser Kırkızlar, Selin Enüst Çalışkan, Yonca Toker, Selçuk Uluağaç, Örsan Özener, Ali Ekici, Serhan Duran and Burak Karacık, whose presence made Atlanta feel more like home. I am indebted to Hélder Inácio and Daniel Steffy, for always keeping their doors and minds open to me, for stimulating and helpful discussions, for their perpetual attentiveness, but most importantly for their great support in dire times, and joyful company in good times. I am grateful to all my friends and colleagues at Georgia Tech for their consideration and cheerful presence, which created a convivial atmosphere that I will always feel privileged to be part of. Thank you, Doug Altner, Yao-Hsuan Chen, So Yeon Chun, Daniel Dadush, Ricardo Fukasawa, Marcos Goycoolea, Kapil Gupta, Qie He, Yi He, Michael Hewitt, Guanghui Lan, Yaxian Li, Vijay Narayanan, Amandeep Parmar, Jon 'Pete' Petersen, Feng Qiu, Claudio Santiago, Alejandro Toriello, Juan Pablo Vielma, and many others whom I shouldn't have forgotten.

I strongly felt the ceaseless faith and support of my family along the way. I deem myself fortunate to have their presence and love, and deeply appreciate all their sacrifices over many years of my studies. I am forever indebted to my husband, Orhan, for his deep love, undying support and companionship. I am especially obliged for his patient endurance against my sullen mood during the difficult stages of my PhD. Thank you for sharing both the happiest and the hardest times in my life and for everything you have endured with me!

Finally, I would like to acknowledge the partial support for this research from Office of Naval Research grant N00014-08-1-1104 and National Science Foundation grant DMS-0914785, and the John Morris and ISyE Tennenbaum Fellowships from Georgia Tech.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this thesis, we develop tractable relaxations and efficient algorithms for large-scale optimization. Our developments are motivated by a recent paradigm, *Compressed Sensing*, which covers a multitude of large-scale, sparsity-oriented convex optimization problems. Compressed sensing is focused on the recovery of sparse or well-concentrated signals from possibly noisy observations in a low-dimensional space. Nowadays, this theory is successfully utilized in many fields ranging from MRI image processing to machine learning, from biology to statistics. In the first chapter of this thesis, we provide a general introduction to compressed sensing and its applications and cover some of the earlier results.

The majority of results in compressed sensing theory rely on the ability to design/use projection matrices with good recoverability properties. In the second chapter of this thesis, we study the conditions for good recoverability properties of a sensing matrix. We propose necessary and sufficient conditions for a sensing matrix to allow for exact ℓ_1-recovery of sparse signals with at most s nonzero entries while utilizing a priori information given in the form of sign restrictions on part of the entries. We express error bounds for imperfect ℓ_1-recovery in terms of the characteristics underlying these conditions. These characteristics, although difficult to evaluate, lead to two different verifiable sufficient conditions, which can be efficiently computed via linear programming (LP) and/or semidefinite programming (SDP) and thus generate efficiently computable lower bounds on the level of sparsity, s, for which a given sensing matrix is shown to allow for exact ℓ_1-recovery. We analyze the connection between our LP- and SDP- based verifiable sufficient conditions, examine their properties, describe their limits of performance and provide numerical examples comparing them with other verifiable conditions from the literature. Even though our LP- and SDP-based relaxations are presented in CS framework, these techniques are generic and applicable in the case of disjoint bilinear programs.

In the third chapter, we study the compressed sensing *synthesis problem* – selecting the

minimum number of rows from a given matrix, so that the resulting submatrix possesses certifiably good recovery properties. Starting from the verifiable sufficient conditions, we express the synthesis problem as the problem of approximating a given matrix by a matrix of specified low rank in the uniform norm. We develop a randomized algorithm for efficient construction of rank k approximation of matrices of size $m \times n$ achieving accuracy bounds $O(1)\sqrt{\ln(mn)/k}$ which hold in expectation or with high probability. We supply a derandomized version of our approximation algorithm and provide numerical results on its performance for the synthesis problem.

Chapter 4 is dedicated to efficient first-order algorithms for large-scale, well-structured convex optimization problems. Saddle point reformulation is proven to be an effective tool to exploit problem structure for designing computationally efficient algorithms. Building upon their strength, we first demonstrate that the solutions to many large-scale problems arising from compressed sensing recovery, high-dimensional statistical inference, and machine learning can be obtained through solving a series of Bilinear Saddle Point problems (BSPs). We accelerate the solution of associated single-parametric BSP's by utilizing the Mirror Prox algorithm from [101] as a prototype and by replacing precise first order oracle (which becomes quite time-consuming in the extremely large-scale case) by its computationally cheap randomized counterpart. In the overall solution of parametric BSPs, cheap online assessment of solution quality is crucial. Our randomized algorithms come with exact guarantees on solution quality and achieves sublinear time behavior to solve large-scale parametric BSPs. Extensive simulations show that our randomized first-order methods are capable of handling very large-scale applications and improve considerably over the state-of-the-art deterministic algorithms, with benefits amplifying as the sizes of the problems grow.

In the fifth chapter, we examine a more general sparse estimation problem –estimating a signal from its undersampled observations corrupted with nuisance and stochastic noise. Instead of the standard sparse signal framework, here we work under the assumption that a priori information is presented via a block representation structure of a known linear transform of the signal, and the signal achieves a good approximation in block sparse sense

in this representation structure. There are a number of important applications where such a nontrivial sparsifying representation arises naturally such as standard image reconstruction with Total Variation regularization or finding the solution of a linear finite-difference equation with sparse right hand side ("evolution of a linear plant corrected from time to time by impulse control"). We show that an extension of the standard compressed sensing results from [79] to this framework is possible. Particularly, we introduce a family of conditions, suggest two new methods of recovery based on block-ℓ_1 minimization and study the most common cases of the block representation structure under which these estimators have efficiently verifiable guaranties of performance. We link our performance estimations to the well known results of compressed sensing by providing connections between our conditions and Restricted Isometry Property. This also establishes connections between new techniques and classical methods such as Lasso and Dantzig Selector.

We present a summary of conclusions of our study and provide future research directions in the last chapter.

CHAPTER I

INTRODUCTION

In this thesis, we develop tractable relaxations and efficient algorithms for large-scale optimization. Our developments are motivated by a recent paradigm, *Compressed Sensing*, which covers a multitude of large-scale, sparsity-oriented convex optimization problems. The traditional approach of reconstructing signals or images from measured data follows the well-known Shannon sampling theorem [126], which states that the sampling rate must be twice the highest frequency. Arisen from a recent breakthrough in signal processing, compressed sensing (also referred to as compressive sampling) is currently reshaping the way people work with large and high-dimensional data sets. Data is compressible in most cases; i.e., the number of salient features hidden in massive data is usually much smaller than its dimension and therefore can be recovered from what was previously believed to be highly incomplete measurements. Exploiting this fact, compressed sensing suggests a paradigm change by acquiring directly low-dimensional linear projections of data, possibly corrupted with noise, and then using sophisticated ℓ_1-recovery procedures for reconstruction of the original data when needed.

The idea of making inferences about the data from a few of its measurements dates back to group testing method suggested in [49] during World War II. In the late 1970s and 1980s geophysicists used ℓ_1-minimization in reflection seismology [117, 122]. In the 1990s, the work of [128] introduced the recovery of sparse Fourier spectra from a few samples in the context of Nuclear Magnetic Resonance spectroscopy and this area of research has attracted vast attention since then. Originating from [116], total variation minimization, which is closely related to compressed sensing, has been widely applied in image processing since the 1990's. In statistics, with the work of Tibshirani [123], use of ℓ_1-regularization and related methods, specifically Lasso (Least Absolute Shrinkage and Selection Operator) estimator, gained great popularity in *model selection* and *sparse estimation* areas. ℓ_1-recovery was also

proposed in computational harmonic analysis for extracting a sparse signal representation from highly overcomplete dictionaries (see [32]). Nowadays, compressed sensing theory is successfully utilized in many fields ranging from magnetic resonance image (MRI) processing to machine learning, from biology to statistics. A few of these direct application areas with references describing the setup of the problems are as follows: Imaging including MRI [21, 89, 127, 129], radar [69], error correction [27], biology including microarray gene expression studies [25, 95, 109] and even a nonparametric approach to modeling customer choice from limited data in revenue management [57]. We refer the interested reader to [11, 39] and [59] for a broader description of compressed sensing applications with details.

1.1 Overview of Compressed Sensing

1.1.1 Preliminaries and Notation

Let $x \in \mathbb{R}^n$ and $\|\cdot\|$ be any norm. For $1 \leq p < \infty$, we denote by $\|\cdot\|_p$ the usual ℓ_p-norm given by

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p},$$

and $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$. For $0 < p < 1$, $\|\cdot\|_p$ defined above becomes a quasi-norm. For a given norm $\|\cdot\|$, its conjugate norm is defined as

$$\|u\|_* = \max_x \left\{ u^T x : \ \|x\| \leq 1 \right\}.$$

For a subset $I \subset \{1, \ldots, n\}$, we denote by $x_I \in \mathbb{R}^n$ the vector which coincides with $x \in \mathbb{R}^n$ on the entries in I and is zero outside of I. While $\mathrm{Card}(I)$ represents the cardinality of the set I, the subset $\bar{I} := \{1, \ldots, n\} \setminus I$ denotes the complement of I. The kernel (nullspace) of a matrix is defined as $\mathrm{Ker}(A) = \{u : \ Au = 0\}$.

Given an $m \times n$ matrix A, we use $\|A\|_{p \to q}$ to denote the operator norm associated with norms $\|\cdot\|_p$ and $\|\cdot\|_q$ in the argument and image spaces respectively, namely:

$$\|A\|_{p \to q} = \max_x \left\{ \|Ax\|_q : \ \|x\|_p \leq 1 \right\}.$$

Unless otherwise stated, I_n denotes the $n \times n$ identity matrix and e_i is the i^{th} basic orth.

1.1.2 Sparsity and Compression

Compressed sensing is based on the empirical observation that many types of real-world signals and images have a sparse expansion of a suitable basis, for example a wavelet expansion provides sparse representation for certain types of images. This means that the expansion has only a small number of significant terms, or in other words, that the coefficient vector can be well-approximated with one having only a small number of nonzero entries.

We say that an n-dimensional signal (vector) x is *s-sparse* if it has s or fewer nonzero coordinates, i.e., $\|x\|_0 \leq s \ll n$ where

$$\|x\|_0 := \text{Card}(\text{supp}(x))$$

and $\text{supp}(x) = \{j : x_j \neq 0\}$ denotes the support of x. Although $\|\cdot\|_0$, which counts the number of nonzero coordinates of a vector, is not even a quasi-norm, it is usually referred as ℓ_0-norm.

Signals encountered in practice are often not exactly sparse, but their coefficients decay rapidly following a power law:

$$|x_{[k]}| \propto k^{(-1/q)},$$

where $x_{[k]}$ is the k-th largest in magnitude entry in x and $0 < q < 1$. This class of signals is referred to as *compressible* signals. In addition to covering sparse signals as a subclass, compressible signals exhibit the nice property that they can be well approximated by sparse signals. We will denote by x^s, the best s-sparse approximation of the vector x, which is obtained by setting to zero all but the s largest in magnitude entries in x and we will refer to $\|x - x^s\|_1$ as the *s-tail* of the vector x. For any positive integer s and $1 \leq p < \infty$, we will define the (s, p)-norm, denoted by $\|\cdot\|_{s,p}$, of a vector x as follows:

$$\|x\|_{s,p} := \|x^s\|_p.$$

Traditionally in order to compress a vector x, one may simply store its s largest entries and their locations. When reconstructing x from its compressed version the missing entries are simply set to zero leading to a small reconstruction error whenever the signal

is compressible. This observation paired with the suitable sparsifying basis plays the key role in many commonly used compression algorithms for visual and audio signals such as JPEG, JPEG-2000, MP3, and MP4. However we should emphasize that the procedure of obtaining the compressed version of x is *adaptive* and *nonlinear* since it requires the search of the largest entries of x in absolute value. In particular, the location of the nonzeros is a nonlinear type of information.

1.1.3 Compressed Sensing

The adaptive compression of a signal x by only keeping its largest coefficients as described before is certainly valid only when full information on x is available. Note that especially when the signal first has to be acquired or measured by a somewhat costly or lengthy procedure, this compression seems to be a waste of resources. At first, large efforts are made to acquire the full signal and then most of the data is thrown away in the compression process. Alternatively, in compressed sensing, one tries to avoid this waste of effort in acquisition phase as much as possible. The key objective in compressed sensing is to reconstruct a signal accurately and efficiently from a small set of *non-adaptive* (possibly noisy) *linear* measurements.

A collection of m linear measurements of a signal $x \in \mathbb{R}^n$, corresponds to applying an $m \times n$ *measurement or (sensing)* matrix A. In other words, given an $m \times n$ measurement matrix A, we only have access to a vector of *observations* (also referred as *measurements*), b, obtained in the following way

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

where x is the signal to be estimated, and if present, $e \in \mathbb{R}^m$, is the vector for observation error. The two most popular assumptions for modeling observation error, e, are as follows:

(B) e is "uncertain-but-bounded" i.e., all we know about e is that $\|e\| \leq \delta$ for a given δ and norm $\|\cdot\|$, or

(S) $e = \sigma\xi$ where ξ is a stochastic r.v. with known distribution P, and σ is the noise intensity level.

Depending on the application area, there are a number of closely related goals including the recovery of x, the recovery of the support of x or recovery of a linear transform Bx of x from the measurements b.

Regardless of the choice of noise model, the undersampled estimation problem is usually the main focus of research, i.e., $m \ll n$, where there are more columns than rows in matrix A. In the undersampled case, even when there is no noise, i.e., $e = 0$, we are dealing with a highly underdetermined system of equations and hence (1), if solvable, will have infinitely many solutions, and thus one cannot effectively distinguish the true underlying signal among them without further information.

Compressed sensing exploits the fact that there is some structure and redundancy in the majority of interesting signals –they are not pure noise. In particular, most signals are sparse, that is, they contain many coefficients close to or equal to zero, when represented in appropriate basis. Instead of working with the general class of all signals, one can achieve satisfactory results in reconstructing signals from specific signal classes such as sparse or compressible.

In this setting, one way of incorporating the a priori sparsity information in the recovery of a vector x from its observations b is simply to use it in guiding the search. In the case of uncertain-but-bounded noise model *(B)*, this approach leads to the ℓ_0-*minimization problem* given by

$$\widehat{x} = \underset{z}{\operatorname{argmin}}\{\|z\|_0 : \ \|Az - b\| \leq \delta\}, \tag{2}$$

which searches for the sparsest signal consistent with the observations within the noise level. A similar sparsity promoting estimator for the stochastic noise model *(S)* (ξ is assumed to be a Gaussian r.v. in most of the literature) is suggested as

$$\widehat{x} = \underset{z}{\operatorname{argmin}}\{\|Az - b\|_2^2 + \lambda\|z\|_0\}, \tag{3}$$

where λ is a regularization parameter depending on noise intensity level σ.

Whenever there is no observation error/noise, i.e., $e = 0$, x is s-sparse and A is one-to-one on all $2s$-sparse vectors, the unique minimizer to (2) is the original vector x (true

signal). On the other hand both of these ℓ_0-regularized minimization problems are neither stable with respect to observation errors nor easy to solve. In fact, for a general matrix A and vector b, it is NP-hard to solve (2) (see [97]).

The computational difficulty in ℓ_0-regularized optimization problems is due to the non-convex, in fact discrete, nature of the objective function. Tractable alternatives are obtained by replacing the ℓ_0-norm in the objective with ℓ_1-norm. When noise is modeled as in *(B)*, the following ℓ_1-*minimization problem*, also referred as *basis pursuit* is obtained

$$\widehat{x} = \underset{z}{\operatorname{argmin}}\{\|z\|_1 : \ \|Az - b\| \leq \delta\}. \tag{4}$$

Similarly, in the stochastic noise setting, *(S)*, there is a computationally efficient alternative to problem (3), namely the *Lasso estimator*, ℓ_1-penalized least squares from [123]:

$$\widehat{x}_L = \underset{z}{\operatorname{argmin}}\{\|Az - b\|_2^2 + \lambda\|z\|_1\}. \tag{5}$$

Another commonly used estimator in the stochastic noise setting is given by *Dantzig selector* from [29], given by

$$\widehat{x}_D = \underset{z}{\operatorname{argmin}}\{\|z\|_1 : \ \|A^T(Az - b)\|_\infty \leq \rho\}, \tag{6}$$

where ρ is a parameter of the algorithm. In high dimensional statistical inference, particularly model selection area (variable selection in linear regression models from a small number of noisy observation), both Lasso estimator and Dantzig selector have gained a lot of popularity, e.g., see [17, 18, 23, 29, 123] and also the references cited therein.

Various characterizations of conditions on the matrix A and the sparsity of the signal of interest have been proposed to guarantee that the suggested estimators $\widehat{x}$ recovers (or at least is close to) the true solution x. In general, with ℓ_1-regularization, while we attain an easy-to-solve alternative, we pay a price by requiring a stronger condition on the sensing matrices A to guarantee the perfect recovery for all signals that are "sparse enough." In the next section, we will examine some of these most commonly studied conditions from the literature.

1.2 Summary of Previous Results for ℓ_1-recovery

1.2.1 Conditions for ℓ_1-recovery

The compressed sensing theory offers strong results which state, in particular, that if the signal x is s-sparse (or compressible) and the matrix, A, possesses a certain well-defined property, then the ℓ_1-recovery of x is close to the true signal, provided the observation error, e in (1), has a small norm or noise intensity level. In this section, we will cover some of the major results in this respect.

We start by characterizing the conditions for *exact ℓ_1-recovery* when there is no noise in the observations (i.e., $e = 0$ and thus $b = Ax$). Here we are mainly interested to answer the following question:

Whether the matrix A is such that whenever the true signal x in (1) is s-sparse, the ℓ_1-recovery in (4) with $e = 0$ recovers x exactly as the unique solution.

If the answer is positive, the matrix A is said to be *s-good*.

Various sufficient, and necessary and sufficient conditions for s-goodness are proposed in the literature. Moreover the quality of the ℓ_1 estimators guaranteed by these conditions in the "imperfect settings"– when observation error is present and the signal is compressible, but not exactly s-sparse, has been an active research area.

1.2.1.1 Nullspace Property

A well known necessary and sufficient condition for s-goodness is as follows:

Definition 1.2.1 *(see [38, 81]). A matrix A is said to satisfy* Nullspace Property *at level s, where s is a positive integer, if for all $x \in \mathrm{Ker}(A)$ and for all subsets $I \subseteq \{1, \ldots, n\}$ with $\mathrm{Card}(I) \leq s$, there exists $\widehat{\gamma}_s(A) \in (0, \frac{1}{2})$ such that*

$$\sum_{i \in I} |x_i| \leq \widehat{\gamma}_s(A) \|x\|_1. \tag{7}$$

This condition has been investigated extensively. Donoho and Huo [45] proved that the matrix A is s-good if A satisfies Nullspace property with $\widehat{\gamma}_s(A) < \frac{1}{2}$. The necessity of Nullspace property with $\widehat{\gamma}_s(A) < \frac{1}{2}$ for s-goodness has been established in [42]. Note that

this condition also appears in the literature under the name of *strict s-balancedness* (see [132, 134]).

Nullspace property can be used to determine the quality of the estimate from ℓ_1-recovery under "imperfect conditions" – where the signal is approximately sparse, there is bounded noise in the observations (noise model *(B)*) and ℓ_1-minimization in (4) is approximately solved (see Proposition 3.1 in [81] or [34]):

Theorem 1.2.1 *Consider the uncertain-but-bounded noise model* (B) *with norm* $\|\cdot\|$. *Let* $x \in \mathbb{R}^n$, *and the sensing matrix* A *satisfy Nullspace property with* $\widehat{\gamma} := \widehat{\gamma}_s(A) < \frac{1}{2}$ *and* $\widehat{x}$ *be a* ν*-optimal solution to* (4), *meaning*

$$\|\widehat{x}\| \leq \min_z \{\|z\|_1 : \|Az - b\| \leq \delta\} + \nu.$$

Then

$$\|x - \widehat{x}\|_1 \leq \frac{1}{1 - 2\widehat{\gamma}} \left[4\beta\delta + 2\|x - x^s\|_1 + \nu\right] \tag{8}$$

where β *is a constant depending on only* A *and norm* $\|\cdot\|$.

On the negative side, verifying the Nullspace property turns out to be a hard optimization problem. Starting from the Nullspace property, efficiently verifiable sufficient conditions based on Linear Programming (LP) and Semidefinite Programming (SDP) relaxations are proposed in [81] and [38] respectively. In both papers, the goal is to efficiently bound $\widehat{\gamma}_s(A)$ in (7) from above. In particular, in [81], the following LP-based verifiable sufficient condition for A to be s-good is stated:

$$\text{There exists } Y \in \mathbb{R}^{m \times n} \text{ s. t. } \max_i \|(I_n - Y^T A)e_i\|_{s,1} < \frac{1}{2} \tag{9}$$

where I_n denotes the $n \times n$ identity matrix and e_i is the i^{th} basic orth. This condition can be further relaxed to obtain another verifiable sufficient condition for A to be s-good given by the following (cf. [81]):

$$\text{There exists } Y \in \mathbb{R}^{m \times n} \text{ such that } \|I_n - Y^T A\|_\infty < \frac{1}{2s} \tag{10}$$

where $\|X\|_\infty = \max_{i,j} |X_{ij}|$ and X_{ij} are the elements of the matrix X.

Unfortunately, tractability has a price: the limits of performance of the verifiable sufficient conditions from [38] and [81] have been established as follows: for an $m \times n$ matrix A, the LP-based conditions cannot verify s-goodness for levels of $s > 2\sqrt{2m}$ provided A is not "nearly square," specifically $n \geq (1 + 2\sqrt{2m})^2$ (see [81]). This bound is much worse than the $O(m/\ln(n/m))$-level of goodness bound which is theoretically achievable for random matrices. Moreover the SDP-based verifiable sufficient condition of [38], despite its significantly higher numerical complexity as compared to the LP-based condition of [81], shares the same asymptotic performance limit.

1.2.1.2 Restricted Isometry Property

Some particularly impressive results in compressed sensing literature make use of the sufficient condition for s-goodness:

Definition 1.2.2 *(see [25, 26]) An $m \times n$ matrix A satisfies the* Restricted Isometry Property *with parameters $\gamma \in (0, 1)$ and s, where s is a positive integer,* RIP(γ, s)*, if*

$$(1 - \gamma)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \gamma)\|x\|_2^2. \tag{11}$$

holds for all s-sparse $x \in \mathbb{R}^n$.

In other words, RIP(γ, s) states that A should be well conditioned when acting on signals of interest. RIP condition provides uniform recovery guarantees over *all* sparse signals of interest for many convex optimization approaches such as ℓ_1-minimization, Dantzig selector or Lasso estimator. For instance, the following result due to Candés et al. for the bounded noise model of *(B)* is well known (see Theorem 1.2 in [26] or Theorem 4.1 in [25]):

Theorem 1.2.2 *Let $x \in \mathbb{R}^n$ and $\|\cdot\|$ be the ℓ_2-norm in the ℓ_1-recovery given in (4), and let the sensing matrix A satisfy* RIP$(\gamma, 2s)$ *with $\gamma < \sqrt{2} - 1$. Then*

$$\|x - \widehat{x}\|_1 \leq \frac{2}{1 - \gamma - \sqrt{2}\gamma} \left[2\delta\sqrt{1 + \gamma}\sqrt{s} + (1 - \gamma + \sqrt{2}\gamma)\|x - x^s\|_1\right]. \tag{12}$$

Theorem 1.2.2 establishes the stability of ℓ_1-recovery with respect to uncertain-but-bounded observation error, δ, and approximate sparsity of the signals under RIP$(\gamma, 2s)$ condition.

The RIP based results regarding levels of achievable compression in compressed sensing is even more impressive. There are $m \times n$ sensing matrices A which possess, say, the RIP$(\frac{1}{2}, s)$-property for "large" sparsity levels s as large as $O(\frac{m}{\ln(n/m)})$ (this bound is tight). For instance, for normalized (all columns have Euclidean length equal to 1) random matrices where entries are sampled from i.i.d. standard Gaussian distribution or Rademacher distribution (entries are ± 1 with equal probability), this is the case with overwhelming probability (i.e., probability $1 - O(e^{-\tau n})$ for some $\tau > 0$) (see [25, 28]). Similar results also exist for normalizations of randomly selected submatrices of the Fourier transform or other orthogonal matrices such as Hadamard.

On the negative side, random matrices are the only known matrices which possess the RIP(γ, s) property for such large s. Yet, in practical applications, random matrices are usually undesirable due to storage limitations, computational considerations, or the mismatch of such matrices with certain measurement architectures. For all known deterministic families of $m \times n$ matrices provably possessing the RIP(γ, s) property, one has $s \approx O(\sqrt{m})$ (we discuss the known deterministic constructions and achievable levels of sparsity in detail in Section 1.2.2), which is essentially worse than the bound $s = O(\frac{m}{\ln(n/m)})$ promised by the RIP based theory. Furthermore there are known lower bounds on the number of rows of a sparse matrix satisfying RIP(γ, s); particularly in [31], it is shown that for a $0-1$ matrix to satisfy RIP(γ, s), the number of rows, m should be "large," i.e., $m \geq \min\left\{\left(\frac{1-\gamma}{1+\gamma}\right)^2 s^2, \frac{1-\gamma}{1+\gamma} n\right\}$. Moreover RIP$(\gamma, s)$ itself is "intractable" (see [111] for a discussion of the hardness of this problem) – the only currently available technique to verify the property for an $m \times n$ matrix amounts to test all its $m \times s$ submatrices. In other words, given a large sensing matrix A, verifying RIP(γ, s)-property with a given $s \gg 1$ is almost impossible in a reasonable amount of time.

The link between the Nullspace property and RIP condition is studied in [34]; it is shown that whenever A satisfies the RIP$(\gamma, 3s)$, then A satisfies the Nullspace property of order $2s$ and $\widehat{\gamma}_{2s}(A) = \frac{\sqrt{1+\gamma}}{\sqrt{2(1-\gamma)+\sqrt{1+\gamma}}}$. Also in [81], it was shown that whenever matrix A is, say RIP$(1/4, m)$ (so the true level of s-goodness of A is $O(1)$[1]m), the LP-based sufficient

[1] In this thesis, we will use the notation $O(1)$ to denote absolute constants.

conditions do certify that A is $O(1)\sqrt{m}$-good –e.g., they guarantee "at least the square root of the true level of goodness."

1.2.1.3 Mutual Incoherence

To the best of our knowledge, the earliest *efficiently verifiable* condition for s-goodness offered by the existing compressed sensing theory is the sufficient condition based on the *mutual incoherence*:

Definition 1.2.3 *(see [42, 45, 90]). The* mutual incoherence *of a given matrix A is the largest absolute normalized inner product between different columns from A. It is given by*

$$\mu(A) = \max_{i \neq j} \frac{|A_i^T A_j|}{A_i^T A_i} \tag{13}$$

where A_i are columns of A (assumed to be nonzero).

The mutual incoherence aims to characterize the dependence between columns of the matrix A. For a unitary matrix, columns are pairwise orthogonal, and so the mutual incoherence is zero. For general matrices with more columns than rows, $m < n$, $\mu(A)$ is necessarily strictly positive, and one would seek for the smallest possible value so as to get as close as possible to the behavior exhibited by unitary matrices. In particular, we have the following theorem (cf. [42]):

Theorem 1.2.3 *Let A has the mutual incoherence value, $\mu(A)$, as defined in (48), then A is s-good for any sparsity level s satisfying*

$$s < \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right). \tag{14}$$

Clearly, the mutual incoherence can be easily computed even for large matrices. On the other hand, in [81], it is shown that the "level of goodness" estimate of a sensing matrix based on mutual incoherence given in Theorem 1.2.3 is usually too conservative. In particular, whenever the mutual incoherence condition from Theorem 1.2.3 verifies the sparsity level s, LP-based verifiable sufficient conditions proposed in [81] for the same sparsity level s are automatically satisfied (in particular the simple condition given in (10) is satisfied).

1.2.2 Deterministic Construction of Compressed Sensing Matrices

There is significant interest in the construction of structured sensing matrices and alternative reconstruction algorithms. One of the earliest results on deterministic construction of compressed sensing matrices with provably good recovery properties is due to [41]. Using finite fields, [41] provides deterministic constructions of cyclic 0-1-valued matrices satisfying RIP(s, γ) with $m = O(\frac{s^2 \log^2(n)}{\gamma^2})$. The analysis in [41] is based on mutual incoherence property and the resulting matrices are provably s-good for the values of $s = O(\sqrt{m})$. The recent work of [20] manage to break through the $\sqrt{m}$ "barrier" using techniques from additive combinatorics: they construct RIP matrices of order $s = O(n^{1/2+\epsilon_0})$ where $\epsilon_0 > 0$ is an unspecified "explicit constant." Note that this is still far from the order achieved by probabilistic constructions.

In recent years, building on the connection with coding theory, adjacency matrices of unbalanced expander graphs originating from [119] have gained increased popularity in compressed sensing field (see [15, 40, 62, 72, 73, 74, 112] and the references therein). An expander graph is a regular bipartite graph for which every pair of subsets of nodes in one side of the partition with sufficiently small size has a small number of colliding edges and a significant number of unique neighbors on the other side:

Definition 1.2.4 *A simple bipartite graph $G = (U, V, E)$ with vertices partitioned into two groups U and V, and edge set E, is said to be a* d-regular (s, ϵ)-unbalanced expander graph *if each vertex in the left partition, say U, has degree d and the graph is such that for any set $X \subset U$ with $|X| \leq s$, the set of neighbors $\mathcal{N}(X) \subset V$ of X has size $|\mathcal{N}(X)| \geq (1 - \epsilon)d|X|$.*

Using probabilistic techniques, the existence of left-regular (s, ϵ)-unbalanced expander graphs with n left vertices and $m = O(\frac{s \log(n/s)}{\epsilon^2})$ right vertices and the left degree $d = O(\frac{\log(n/s)}{\epsilon})$ is shown (see [15, 72] and the references therein). No explicit construction with the aforementioned parameters are known, however [15] provides an explicit construction of these matrices with $d = 2^{O(\log(\log(n)/\epsilon))}$ and $m = sd/\epsilon^{O(1)}$. Moreover explicit deterministic construction of expander graphs based on Parvaresh-Vardy codes [110] exists and in the recent work of [68], the associated explicit construction parameters are stated as

$d = O(1)(\frac{\log(n)\log(s)}{\epsilon})^{1+\frac{1}{\alpha}}$ and $m = s^{1+\alpha}d^2$ for any fixed $\alpha > 0$.

The increased interest in the adjacency matrices of expander graphs is due to their stable recovery properties, the sparseness of the associated matrices and the specialized efficient recovery algorithms. Let Φ be the adjacency matrix of a d-regular, $(2s, \epsilon)$-unbalanced expander graph with $\epsilon < \frac{1}{6}$ and define $A = \frac{\Phi}{d}$ to be the normalized adjacency matrix, then A satisfies the Nullspace property with $\widehat{\gamma}_s(A) = \frac{2\epsilon}{1-2\epsilon} < \frac{1}{2}$ (see Lemma 16 in [15]). Combined with Theorem 1.2.1 this validates ℓ_1-recovery under imperfect conditions (another validation, adjusted to the specific matrices in question, is given in [15]). Furthermore, in [15], it is demonstrated that the empirical behavior of randomly generated binary sparse matrices is consistent with the analytic performance analysis (phase transition behavior) of random Gaussian matrices. On the algorithmic side, there are specialized algorithms exploiting the combinatorial structure of expander graphs: Expander Matching Pursuit of [72] works in the noise-free setting and achieves $O(n \log \frac{n}{s})$ computational complexity; Sequential Sparse Matching Pursuit of [16] requires slightly higher (by a logarithmic factor) running time yet handles the bounded noise model of *(B)*.

1.3 Efficient Algorithms for ℓ_1-recovery

1.3.1 Convex Optimization Methods

The ℓ_1-regularized optimization problems stated in (4)-(6) are convex optimization problems, and as such, they can be solved to arbitrarily high accuracy by theoretically and practically efficient polynomial-time algorithms for convex optimization. In addition to the convexity, there is a nice transparent structure in the ℓ_1-minimization problem of (4): when $\|\cdot\|$ is a polyhedral norm such as $\|\cdot\|_\infty$ or $\|\cdot\|_1$, (4) reduces to Linear Programming (LP), and when $\|\cdot\| = \|\cdot\|_2$, the most popular choice in compressed sensing, it becomes a Conic Quadratic Programming (CQP). Problems of this type are amenable for the most advanced polynomial-time Interior Point methods (IPMs) known so far. In particular, both LP and CQP are especially well suited for IPMs, and nowadays commercial software packages such as CPLEX[35], Gurobi[67] and Mosek[4] are capable of solving LPs and CQPs to high accuracy within a moderate number of Newton-like iterations.

However in the case of large-scale LP/CQP problems with dense constraint matrices (e.g., matrix A in (4)), the practical scope of IPMs is restricted to problems with at most few thousands of decision variables. This restriction comes from the fact that IPMs have "computationally demanding" iterations unless problem's data possess favorable sparsity structure. The computational cost, e.g., the number of arithmetic operations (a.o.), of an iteration of standard IPMs grows nonlinearly, in the dense case as $O(n^3)$, with the design dimension, n, of the problem. In the problems originating from real-life decision making, the constraint matrices reflect dependencies between various elements/processes such as in production facilities, inventory and/or supply chain systems, and it is difficult to imagine such a system where "everything influences everything else." Therefore it is usual to have LPs and/or CQPs with sparsity structure in these problems. On the other hand, it is easy to arrive at dense large-scale problems in signal processing and machine learning applications, where the matrices are usually given analytically due to their "mathematical origin." Such an example is readily given in compressed sensing when A matrix in (4) is a subsampled Fourier transform.

Fast -cubic- growth of computational effort per iteration makes standard IPMs practically too expensive or even unable to handle dense problems with tens and hundreds of thousands of design variables. *Fast convergence in terms of iteration count of IPMs, does not help much when the very first iteration "lasts forever."* In situations like this, high accuracy offered by standard IPMs turns out to be computationally too expensive and it becomes necessary to have the complexity of each iteration having at most a linear growth of the design dimension n, i.e., $O(n)$. At the present level of our knowledge, the latter requirement rules out all known polynomial-time routines and, as far as constrained problems like (4) are concerned, leaves us the only option –computationally cheap "black-box-oriented" first-order methods (FOMs) like gradient descent, conjugate gradient, quasi-Newton methods with restricted memory.

1.3.1.1 *Review of First-Order Methods and Their Limits of Performance*

In order to provide sufficient background and put our results into perspective, here we review some of the literature on first-order convex programming and related complexity theory. Our review is mainly based on a summary of results relevant to us from [80], we refer the reader to [80, 103, 105] for full details on the subject and further discussions.

We restrict our attention to the convex programs of the form

$$\mathrm{Opt}(f) = \min_{x \in \mathcal{X}} f(x), \tag{15}$$

where $\mathcal{X}$ is a nonempty compact convex subset of $\mathbb{R}^n$, and f is known to belong to a given family $\mathcal{F}$ of convex and (at least) Lipschitz continuous functions[2] on $\mathcal{X}$. Clearly the feasibility and compactness of $\mathcal{X}$ combined with continuity of f implies that the optimum value in (15) is attained, i.e., (15) is solvable.

While solving (15), a FOM knows in advance what $\mathcal{X}$ and $\mathcal{F}$ are, but does not know the particular objective function $f \in \mathcal{F}$ is. It is restricted to "learn" f via subsequent calls to a *first-order oracle* – a "black-box" routine which, given as an input a point $x \in \mathcal{X}$, outputs the value $f(x)$ and a (sub)gradient $f'(x)$ of f at x[3]. For a given particular objective function f and a required accuracy $\epsilon > 0$, a FOM generates a finite sequence of *search points* $x_t \in \mathcal{X}$, $t = 1, 2, \ldots$ by calling the first-order oracle. Upon termination, the algorithm outputs an approximate solution $\widehat{x} \in \mathcal{X}$ which should be ϵ-optimal, i.e., $f(\widehat{x}) - \mathrm{Opt}(f) \leq \epsilon$. Thus a FOM, in fact, is a collection of rules for generating subsequent search points, termination criteria and building the approximate solution. These rules, in principle, can be arbitrary, with the only limitation of being *non-anticipating*, i.e., the "output" of a rule is uniquely defined by $\mathcal{X}$ and the first-order information on f accumulated before the rule is applied. Consequently, for a given FOM and $\mathcal{X}$, x_1 is independent of f, x_2 depends solely on $f(x_1), f'(x_1)$, and so on. Moreover both the termination rule and the

[2]The function $f(x)$ is said to be Lipschitz continuous with Lipschitz constant L with respect to norm $\|\cdot\|$ if

$$|f(x) - f(y)| \leq L\|x - y\|$$

holds for all x, y.

[3]Informally speaking, this setting implicitly assumes that the domain $\mathcal{X}$ is "simple" (like box, or ball, or standard simplex), while f can be complicated.

construction of approximate solution should also be non-anticipating, depending only on $\mathcal{X}$ and the previous information on f collected at the search points.

In this framework, given target accuracy ϵ, the lower complexity bound of a problem family $(\mathcal{X}, \mathcal{F})$, denoted by $\mathcal{N}_\ell(\mathcal{X}, \mathcal{F}, \epsilon)$, is defined as the minimum number of iterations, N, of any FOM capable to solve every problem of type (15) with every possible function $f \in \mathcal{F}$ within accuracy ϵ. An upper bound on complexity can be obtained by demonstrating a FOM and its associated worst case performance on the outlined problem family. Whenever there exists a FOM capable of solving all problems of type (15) with functions $f \in \mathcal{F}$ within accuracy ϵ with a worst case performance that is within an absolute factor of $\mathcal{N}_\ell(\mathcal{X}, \mathcal{F}, \epsilon)$, then that algorithm is referred as an *optimal method* for the given problem class. Limits of performance of black-box-oriented FOMs are established by Information-Based Complexity Theory in [103]. Several instructive examples for specific classes of $\mathcal{X}$ and $\mathcal{F}$ are given in [103], the following are of particular interest:

(a) [*Nonsmooth case*] Let $\mathcal{X} = \{x \in \mathbb{R}^n : \|x\|_p \leq R\}$, where $p \in \{1, 2\}$, and let $\mathcal{F}_p$ be comprised of all convex functions f which are Lipschitz continuous w.r.t. $\|\cdot\|_p$, with a given constant L. Then $\mathcal{N}_\ell(\mathcal{X}, \mathcal{F}_p, \epsilon)$ is *at least* $O(1)\min[n, L^2R^2/\epsilon^2]$. This lower complexity bound remains true when $\mathcal{F}$ is restricted to be the family of all functions of type $f(x) = \max_{1 \leq i \leq n} [\epsilon_i L x_i + a_i]$ with $\epsilon_i = \pm 1$. Moreover, an optimal FOM suggested in [103] is capable of solving all problems of the outlined type within accuracy ϵ in $O(1)(\ln(n))^{2/p-1}L^2R^2/\epsilon^2$ steps.

(b) [*Smooth case*] Let $\mathcal{X} = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$, and let $\mathcal{F}$ be comprised of all convex functions f which have Lipschitz continuous gradients w.r.t. $\|\cdot\|_2$, with a given constant L. Then $\mathcal{N}_\ell(\mathcal{X}, \mathcal{F}, \epsilon)$ is *at least* $O(1)\min[n, \sqrt{LR^2/\epsilon}]$. This lower complexity bound remains true when $\mathcal{F}$ is restricted to be the family of convex quadratic functions of form $f(x) = \frac{1}{2}x^T Ax + b^T x$ with positive semidefinite symmetric matrices A of spectral norm (maximal singular value) not exceeding L. *Nesterov's optimal algorithm for smooth convex minimization* (see [104, 106]) is capable of solving all problems of the outlined type within accuracy ϵ in $O(1)\sqrt{LR^2/\epsilon}$ steps.

(c) [*Structured nonsmooth case*] Let $\mathcal{X} = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$, and let $\mathcal{F}$ be comprised of all convex functions of the form $f(x) = \|Ax - b\|_2$ where the spectral norm of A[4] does not exceed L. Here we extend the "power" of the first-order oracle slightly and assume that at a step of a FOM we are allowed to carry out $O(1)$ matrix-vector multiplications involving A and A^T, yet we don't have direct access to A. In this setup, $\mathcal{N}_\ell(\mathcal{X}, \mathcal{F}, \epsilon)$ is *at least* $O(1)\min[n, LR/\epsilon]$. Again, this lower complexity bound is "nearly achievable:" there exist a FOM, (in particular, Nesterov's optimal algorithm from [106] as applied to the quadratic form $\|Ax - b\|_2^2$) achieves the desired accuracy in $O(1)LR/\epsilon$ steps.

In the case of large-scale convex optimization, where the problem's design dimension, n, is large, the results stated above bring us bad news: unless the number of steps exceeds n (which is of no interest in large-scale case), *a FOM can exhibit only sublinear rate of convergence if no additional structural restrictions on the objective function are imposed.* Specifically, if we let t denote the number of steps required to get accuracy ϵ, the complexity of $O(1)\frac{(\ln(n))^{1/p-1/2}LR}{\sqrt{t}}$ in the case of (a) is really slow; in the case of (b), the complexity given by $O(1)\frac{LR^2}{t^2}$ is much better but minimization of a smooth function over a simple domain is a rare commodity; and the complexity of $O(1)\frac{LR}{t}$ in (c) is "in between" (a) and (b). Therefore *FOMs are poorly suited for building high-accuracy solutions to large-scale convex problems.*

On the positive side, *for problems with favorable geometry (e.g., those in (a) – (c)), good FOMs exhibit nearly dimension-independent rate of convergence*[5], which is of utmost importance in large-scale applications.[6] Moreover, in all of the above cases, whenever $\mathcal{X}$ is simple, the implementation of the optimal algorithms nearly achieving the complexity bounds is quite simple with cheap iterations – modulo computations "hidden" in the oracle, an iteration costs just $O(n)$ a.o. where $n = \dim(\mathcal{X})$. For details of these algorithms and more discussion on the topic, we refer the reader to [80]. As a consequence, *FOMs are well*

[4] In this setting A is no longer restricted to be positive semidefinite.

[5] upto $\log(n)$ factor dependency involved in the complexity bounds

[6] On a side note, this nearly dimension-independent performance of FOMs heavily depends on the assumption $p \in \{1, 2\}$. In this nonsmooth setting, when minimization is over a box, i.e., $p = \infty$, the upper and lower complexity bounds become $O(1)n\log(LR/\epsilon)$ provided that $LR/\epsilon \geq 2$, demonstrating the significant dependence on the dimension.

suited for finding medium-accuracy solutions to large-scale convex problems with "favorable geometry."

Coming back to the origin of our interest in FOMs, it is easy to note that none of the ℓ_1-regularized optimization procedures fits into the "favorable geometry" settings described above. However a close relative given by

$$\min_z \{\|Az - b\| : \ \|z\|_1 \leq R\} \tag{16}$$

fits into the structured nonsmooth optimization over a simple domain setting as discussed in case (c). It can be shown that the optimal solution to the ℓ_1-recovery problem given in (4) can be obtained by solving a sequence of problems of the form (16) (see Section 4.2.2.2 in Chapter 4). Therefore algorithms for the structured nonsmooth setting described in (c) are fundamental in our developments. The recent development in FOMs as described in case (c), is due to Nesterov's breakthrough paper of [106]. Nesterov in [106] showed that typical problems of nonsmooth convex minimization usually can be reformulated (and this is where problems structure is exploited) as smooth (often just bilinear) convex-concave saddle point problems, and the latter can be solved by appropriate black-box oriented FOMs with $O(1/t)$ rate of convergence. For a slightly more general class of problems, utilizing the bilinear saddle point reformulation, Nemirovski in [101] suggested a new FOM – Mirror Prox (MP) algorithm – achieving the same $O(1/t)$ rate of convergence.

The deterministic first-order oracle in both algorithms of [106] and [101] for solving problems of type (16) requires just several matrix-vector multiplications (by A and A^T) plus $O(n)$ "overhead." This feature of FOMs working with only matrix-vector products rather than the full matrix as required by standard IPMs serves especially advantageous when it is not even possible to store the full matrix in memory or when A matrix admits fast routines for matrix-vector multiplication, which are available for instance in the case of subsampled Fourier matrices. Moreover, when the norm $\|\cdot\|$ in (16) is simple such as ℓ_∞-norm or ℓ_2-norm, and the problem is large-scale with dense A matrix (which is the case in many machine learning and signal processing applications), these matrix-vector multiplications dominate the computational cost of an iteration of a FOM. As the sizes of A matrix grow,

these multiplications can become prohibitively time consuming. Further acceleration of FOMs is possible by replacing precise deterministic first-order oracles, which can become too time-consuming in the extremely large-scale case, with their computationally cheap properly designed stochastic counterparts. The main idea behind the *stochastic first-order oracle* is that matrix-vector multiplications is easy to randomize –reducing to just extracting from A a row and a column, which, in the large-scale dense matrix case, can be several orders of magnitude cheaper than exact matrix-vector multiplication. This randomization, under favorable circumstances, allows for dramatic acceleration of FOMs in the extremely large-scale case.

Above results on FOMs lead to another important conclusion: unlike polynomial time IPMs, the limits of performance of FOMs heavily depend on the size R of the feasible domain; in particular, boundedness of $\mathcal{X}$ is of paramount importance, at least theoretically, for the success of FOMs. In this respect, unconstrained settings, like Lasso estimator given by (5) are less preferable than their "bounded domain" counterparts, like (16). There are a number of papers suggesting FOMs for typical compressed sensing problems and reporting good empirical results (see [2, 5, 6, 7, 12, 13, 63, 64, 130] and references therein). Some of these papers specifically deal with the unconstrained versions despite this theoretical concern.

1.3.2 Greedy Algorithms for ℓ_1-recovery

Motivated by the desire to provide a reduced complexity alternative to the ℓ_1-recovery problem, many greedy methods are suggested in the compressed sensing literature. These greedy methods include matching pursuit algorithms and iterative thresholding. The Matching Pursuit algorithm for signal recovery is first introduced in [90], several variants are proposed since then including orthogonal matching pursuit [90, 124, 125], Stagewise Orthogonal Matching Pursuit (StOMP) [48], Compressive Sampling Matching Pursuit (CoSaMP) [98], Regularized Orthogonal Matching Pursuit [99] (see, e.g., the review [21]) and iterative hard thresholding [19, 58, 60]. Majority of these approaches are aimed at calculating the support of the signal iteratively. At each iteration of the algorithm, based on successive Euclidean

projections of the signal, one or several coordinates of the signal is selected for testing. Although greedy approaches are relatively fast as compared to the exact ℓ_1-minimization, both in theory and in practice; most of them deliver smaller recoverable sparsity levels and lack provable uniform recoverability guarantees, i.e., the possibility to recover *all* sparse signals and stability w.r.t. noise and/or small perturbations of the signal. Most of the performance results on the matching pursuit algorithms rely upon the bounds on mutual incoherence $\mu(A)$ of the sensing matrix. Recently there has been efforts on analyzing different variants of matching pursuit algorithms to provide uniform recovery bounds for matrices satisfying RIP property at the expense of allowing smaller sparsity levels (see [71, 98, 99]).

1.4 Organization of the Thesis

1.4.1 Chapter 2

As discussed before, the availability of a priori sparsity information plays the key role in making compressed sensing possible. In Chapter 2, we extend the current theory by characterizing what might be achievable if a priori information beyond the sparsity of the signal, given in the form of sign restrictions on part of the entries, is available. We study the conditions for good recoverability properties of a sensing matrix in this setting, in particular our results generalize and subsume the corresponding results from [81] and [38] as our framework allows one to have no additional information on sign restrictions of the entries of the signal.

We start by proposing necessary and sufficient conditions for a sensing matrix to allow for exact ℓ_1-recovery of s-sparse signals while utilizing a priori information given in the form of sign restrictions on part of the entries (Proposition 2.2.1). We express error bounds for imperfect ℓ_1-recovery in terms of the characteristics underlying these conditions (Proposition 2.3.1). These characteristics, although difficult to evaluate, lead to two different verifiable sufficient conditions, which can be efficiently computed via LPs and/or SDPs and thus generate efficiently computable lower bounds on the level of sparsity, s, for which a given sensing matrix is shown to allow exact ℓ_1-recovery (Sections 2.4.1.3 and 2.6). Although our LP-based verifiable condition mimics those given in [81] and thus share similar limits of

performance (Proposition 2.5.1), we show that in the case when a priori sign information is available, a better SDP-based verifiable sufficient condition can be constructed which we can no longer prove to have the same limits of performance of its predecessor from [38] (see Remark in Section 2.6). Moreover, we analyze the connection between our LP- and SDP-based verifiable sufficient conditions (Proposition 2.6.1). We show that our LP-based condition has dual representation of the form close to the one of the SDP-based condition and hence they can be unified, leading to a stronger verifiable condition (Section 2.10.8).

Even though our LP- and SDP-based relaxations are presented in compressed sensing framework, these techniques are generic and applicable in the development of tractable relaxations for disjoint bilinear programs. We discuss the relation of our relaxation schemes with other tractable relaxations for disjoint bilinear programs from the literature including McCormick bounds [91], and Sherali-Adams [118] type and Lovász-Schrijver type relaxations, and we show that our LP-based bound is at least as good as the first two and our unified bound is no worse than the latter one (Section 2.9).

We also present a comparison of our LP-based verifiable condition with other verifiable conditions for s-goodness from the literature (Section 2.4.1.3 and Lemma 2.5.1), and provide numerical results indicating the value of sign information in the recovery of sparse signals (Section 2.7). We close this chapter, by proposing and analyzing a new greedy type algorithm for ℓ_1-recovery, non-Euclidean Matching Pursuit, which utilizes our LP-based sufficient conditions for goodness (Section 2.8).

1.4.2 Chapter 3

In Chapter 3, we study the compressed sensing *synthesis problem* – selecting the minimum number of rows from a given matrix, so that the resulting submatrix possesses certifiably good recovery properties. Starting from the LP-based verifiable sufficient condition given in (10), we express the synthesis problem as the problem of approximating a given matrix by a matrix of specified low rank in the uniform norm (maximum absolute values of entries in the matrix). We develop (Section 3.2.1) a randomized algorithm for efficient construction of rank k approximation of matrices of size $m \times n$ achieving accuracy bounds

$O(1)\sqrt{\ln(n)/k}$ which hold in expectation or with high probability. We supply (Section 3.2.2) a derandomized version of our approximation algorithm and provide numerical results on its performance for the synthesis problem (Section 3.2.3). We also prove that the $O(1)\sqrt{\ln(n)/k}$-accuracy bound is unimprovable up to a logarithmic factor (Proposition 3.2.4). During these developments based on the condition (10), we also establish that for certain structural matrices including subsampled Hadamard and Fourier matrices, the computational cost of verifying (10) can be further reduced (Section 3.3.4).

1.4.3 Chapter 4

Chapter 4 is dedicated to efficient first-order algorithms for large-scale, well-structured convex optimization problems. As discussed in Section 1.3.1.1, saddle point reformulation is proven to be an effective tool to exploit problem structure for designing computationally efficient algorithms. Building upon their strength, we first demonstrate that many large-scale problems arising from compressed sensing recovery (Section 4.2.2.2), high-dimensional statistical inference (Section 4.2.1.3), and machine learning (Section 4.2.2.3) can be obtained through solving a series of bilinear saddle point problems (BSPs), which we refer to as Generalized Bilinear Saddle Point Problem (GBSPP). In Section 4.4, we suggest an algorithm for solving GBSPP which reduces the problem to a *one-dimensional* root-finding for an implicitly defined function. The latter problem is solved by a Newton-type root finding routine, with the (approximate) first-order information for this routine yielded by approximately solving a single-parametric BSP. Our developments are motivated by the need for efficient sublinear time algorithms to solve large-scale GBSPPs. To achieve this, we accelerate the solution of associated single-parametric BSP's by utilizing the Mirror Prox algorithm from [101] as a prototype which we further modify by replacing precise first-order oracle (which becomes quite time-consuming in the extremely large-scale case) by its computationally cheap randomized counterpart. We provide the details of the proposed algorithms and their efficiency estimates in Section 4.3, in particular we show that our randomized algorithms have $O(1/\sqrt{t})$ rate of convergence. In our developments the stochastic oracle is constructed by randomizing the matrix-vector products, thus reducing

the complexity of the oracle from $O(mn)$ to $O(m+n)$ a.o. In this respect, our stochastic FOMs have close relatives proposed in [82, 83, 102]. The advantage of the algorithms being proposed here over those from [82, 83, 102] lies in the immediate possibility to assess, in a computationally cheap fashion, the quality of the resulting approximate solutions. This possibility is instrumental when solving parametric bilinear saddle point problems of GBSPP form. Although the deterministic algorithms to solve BSPs in this setup achieve $O(1/t)$ rate of convergence, we show that due to their high cost of iteration $O(mn)$, the randomized algorithms developed here achieving $O(1/\sqrt{t})$ rate of convergence with $O(m+n)$ cost of each iteration, outperform the deterministic ones significantly, for every fixed required accuracy, provided that the problem is large-scale. Overall for certain range of parameters, our randomized algorithms achieve sublinear-time behavior, i.e., they produce reliable solutions by inspecting a negligible part of the data (i.e., of entries in the sensing matrix A). Extensive simulations provided in Section 4.5 show that our stochastic first-order methods are capable of handling very large-scale applications and improve considerably over the state-of-the-art deterministic algorithms, with benefits amplifying as the sizes of the problems grow.

1.4.4 Chapter 5

In Chapter 5 we study a more general sparse estimation problem with stochastic noise. So far, in this thesis, we have worked with *uncertain-but-bounded noise* model of *(B)*. However, in statistical estimation framework, it is natural to have *random noise with known distribution*, as in noise model *(S)*. In addition to this, instead of studying estimation of signals, we introduce and study estimation of signals that are *block-sparse with respect to a given representation structure*. Specifically, we consider the problem of estimating a linear transform $Bx \in \mathbb{R}^N$ of a vector $x \in \mathbb{R}^n$ from the observations

$$y = Ax + u + \sigma\xi. \tag{17}$$

Here A is a given $m \times n$ sensing matrix, B is a given $N \times n$ representation matrix, and $u + \sigma\xi$ is the observation error; in this error, u is an unknown *nuisance* known to belong to a given compact convex set $\mathcal{U} \subset \mathbb{R}^m$ symmetric w.r.t. the origin, $\sigma \geq 0$ is a known noise intensity, and ξ is random noise with known distribution P. Note that the observation

model combines *(B)* and*(S)*, with u being the "uncertain-but-bounded," and $\sigma\xi$ being the "stochastic" component of the observation error.

We assume that the space $\mathbb{R}^N$ where Bx lives is represented as $\mathbb{R}^N = \mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_K}$, so that a vector $w \in \mathbb{R}^N$ is a block vector: $w = [w[1]; ...; w[K]]$ with blocks $w[k] \in \mathbb{R}^{n_k}$, $1 \le k \le K$. In particular, $Bx = [B[1]x; ...; B[K]x]$ with $n_k \times n$ matrices $B[k]$, $1 \le k \le K$. While we do not assume that the vector x is sparse in the usual sense, we do assume that the linear transform Bx to be estimated is *block sparse*, meaning that at most a given number, s, of the blocks $B[k]x$, $1 \le k \le K$, are nonzero.

We consider recovery routines based on *block-ℓ_1 minimization*, i.e., the estimate $\widehat{w}(y)$ of $w = Bx$ is $B\widehat{z}(y)$, where $\widehat{z}(y)$ is obtained by minimizing the norm $\sum_{k=1}^K \|B[k]z\|_{(k)}$ over signals $z \in \mathbb{R}^n$ with Az "fitting," in certain precise sense, the observations y. Above, $\|\cdot\|_{(k)}$ are given in advance norms on the spaces $\mathbb{R}^{n_k}$ where the blocks of Bx take their values. We refer to the given in advance collection $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ as the *representation structure*. Given such a structure and sensing matrix A, our ultimate goal is to understand how well one can recover the s-block-sparse transform Bx by appropriately implementing block-ℓ_1 minimization.

Note that in this framework, the *standard representation structure*, $B = I_n$, $n_k = 1$, $\|\cdot\|_{(k)} = |\cdot|$, $1 \le k \le K = n$, leads to the standard compressed sensing setting – recovering a sparse signal $x \in \mathbb{R}^n$ from its noisy observations (17) via ℓ_1-minimization. In this respect, our results generalize the recent work in [79].

As notation, for block vector $w = [w[1]; ...; w[K]]$, we let $L_p(w)$ be the $\|\cdot\|_p$-norm of the vector $[\|w[1]\|_{(1)}; ...; \|w[K]\|_{(K)}]$ and $L_{s,p}(w)$ is the block $\|\cdot\|_{s,p}$-norm obtained by taking the L_p-norm of the vector obtained from w by zeroing out all but the s largest in "magnitude" $\|w[k]\|_{(k)}$ blocks in w.

In Section 5.2, we introduce a parametric family of conditions, $\mathbf{Q}_{s,q}(\kappa)$ for $1 \le s \le K$ and $q \in [1, \infty]$, linking sensing matrix $A \in \mathbb{R}^{m\times n}$ and *contrast matrix* $H \in \mathbb{R}^{m\times M}$:

We say that a contrast matrix $H \in \mathbb{R}^{m\times M}$ along with a norm $\|\cdot\|$ on $\mathbb{R}^M$ satisfy

the condition $\mathbf{Q}_{s,q}(\kappa)$, if

$$\forall x \in \mathbb{R}^n : \ L_{s,q}(Bx) \leq s^{\frac{1}{q}} \|H^T Ax\| + \kappa s^{\frac{1}{q}-1} L_1(Bx).$$

We suggest (Section 5.3) two recovery routines ($\widehat{x}_{\rm reg}(\cdot)$ for regular recovery and $\widehat{x}_{\rm pen}(\cdot)$ for penalized recovery) utilizing the contrast matrices, and establish their performance guarantees under condition $\mathbf{Q}_{s,q}(\kappa)$ with $\kappa \in (0, \frac{1}{2})$. In particular, let Ξ and ρ be such that $P(\Xi) \geq 1 - \epsilon$ and $\|H^T(u + \sigma\xi)\| \leq \rho \ \forall (u \in \mathcal{U}, \xi \in \Xi)$; then under condition $\mathbf{Q}_{s,q}(\kappa)$, for all $\xi \in \Xi, u \in \mathcal{U}$, and $x \in \mathbb{R}^n$, both recovery procedures achieve the following accuracy (Theorems 5.3.1, 5.3.2):

$$L_p(B[\widehat{x}(Ax + u + \sigma\xi) - x]) \leq 4(s)^{\frac{1}{p}} \frac{2\rho + s^{-1}\upsilon_s(Bx)}{1 - 2\kappa}, \ \forall p \in [1, q],$$

where $\upsilon_s(w)$ is the "s-concentration of w," that is, the sum of magnitudes $\|w[k]\|_{(k)}$ of all *but the s largest in magnitude* blocks in w.

In Section 5.4, we study the properties of our family of conditions $\mathbf{Q}_{s,q}(\kappa)$. Similar to the Nullspace condition and RIP, the condition $\mathbf{Q}_{s,q}(\kappa)$ seems to be computationally intractable. Nonetheless, in Section 5.4.1, we establish that when all $\|\cdot\|_{(k)}$ are the uniform norms $\|\cdot\|_\infty$, then the condition $\mathbf{Q}_{s,\infty}(\kappa)$, the strongest among our family of conditions $\mathbf{Q}_{s,q}(\kappa)$ (see Observation 5.2.1), becomes "fully computationally tractable" (Proposition 5.4.1). Moreover, in Section 5.4.2, we establish the "necessity" of condition $\mathbf{Q}_{s,\infty}(\kappa)$ (Proposition 5.4.2): when error is measured in L_∞ norm and all norms $\|\cdot\|_{(k)} = \|\cdot\|_\infty$, ξ is a Gaussian r.v. and observation error "is present," then whenever the error bounds of the above form is valid for some sparsity level S, then there exists (and can be efficiently built) a contrast matrix $H \in \mathbb{R}^{m \times N}$ which, along with the norm $\|\cdot\|_\infty$ on $\mathbb{R}^N$, satisfies the condition $\mathbf{Q}_{s,\infty}$ for block-sparsity level s which is nearly as large as S: $s = O(S)$. In Section 5.4.3, under the condition that all of the norms $\|\cdot\|_{(k)} = \|\cdot\|_\pi$ with $\pi \in \{1, 2, \infty\}$, we derive verifiable sufficient alternatives for condition $\mathbf{Q}_{s,q}(\kappa)$ with general $q \in [1, \infty]$ (see Proposition 5.4.3). In the literature, mutual block-incoherence condition of [54] is the only known so far *verifiable* sufficient condition for the validity of *block-ℓ_1* recovery, and is defined specifically for the case of $B = I_n$ and all blocks having Euclidean norms, i.e., $\|\cdot\|_{(k)} = \|\cdot\|_2$. We show in Section 5.4.3.1 that

the mutual block-incoherence condition is more conservative than our *verifiable* sufficient condition for the validity of $\mathbf{Q}_{s,\infty}(\kappa)$. We establish (Proposition 5.4.5) limits of performance for our verifiable sufficient condition when $q \in \{1,2\}$, $B = I_n$, all blocks have equal size $n_k = d$ and common norms $\|\cdot\|_{(k)} = \|\cdot\|_\pi$ with $\pi \in \{1,2,\infty\}$: unless $m < 3n/4$, we cannot verify block sparsity levels s beyond $\frac{n}{2\sqrt{d(n-m)}}$.

In Section 5.5, we restrict our attention to the standard representation structure for sparse recovery and study the relation between condition $\mathbf{Q}_{s,q}(\kappa)$ and RIP. These relations allow us to establish new, and in certain cases efficiently verifiable, accuracy certificates for Lasso estimator and Dantzig selector. In addition to this, in the case of no-nuisance ($\mathcal{U} = \{0\}$) and Gaussian observation noise ($\xi \sim \mathcal{N}(0, I_m)$), for a Gaussian $m \times n$ matrix A with $\ln(m) = O(1)\ln(n)$, with overwhelming probability as m, n grow, the Dantzig selector satisfies

$$\text{Prob}\{\xi : \|x - \widehat{x}_D(Ax + \sigma\xi)\|_\infty \le O(1)\sigma\sqrt{2\ln(n/\epsilon)}\} \ge 1 - \epsilon$$

for all s-sparse x with $s \le O(1)\sqrt{m/\ln(m)}$. In contrast to this, when $s > O(1)\sqrt{m\ln(m)}$, the above error bound, for typical Gaussian A, does *not* hold for some s-sparse signals x. This establishes that the restriction $s \le O(1)\sqrt{m/\ln(m)}$ indeed is necessary (upto logarithmic factors) to achieve small recovery errors measured in the ℓ_∞-norm: when s is by a logarithmic in m factor greater than this bound, the Dantzig selector associated with a typical Gaussian sensing matrix stops to work properly.

Finally in Section 5.6, under the common norm $\|\cdot\|_{(k)} = \|\cdot\|_\infty$ for all blocks, assuming satisfiability of the (verifiable!) condition $\mathbf{Q}_{s,\infty}(\kappa)$, we develop a computationally cheap alternative, block Non-Euclidean Matching Pursuit algorithm, to the regular/penalized recoveries for signals with block sparse structure.

In the last chapter, we provide brief conclusions of the current work and outline future research directions.

CHAPTER II

VERIFIABLE SUFFICIENT CONDITIONS FOR COMPRESSED SENSING

2.1 Overview

Compressed sensing uses the most basic structural information of the signal to be recovered, its sparsity, in order to successfully recover it from a few of its observations. In practice, *a priori* information about the signal to be recovered often exists and will be beneficial if taken into account in the recovery procedure. In this chapter, we suppose that the *a priori* information about a *sparse* signal $w \in \mathbb{R}^n$ amounts to the *sign restrictions*, and is given as the subsets P_+ and P_- of $\{1, ..., n\}$, $P_+ \cap P_- = \emptyset$, such that $w_i \geq 0$ for $i \in P_+$ and $w_i \leq 0$ for $i \in P_-$. Therefore we address the following recovery problem: given an observation $y \in \mathbb{R}^m$,

$$y = Aw + u, \tag{18}$$

where $A \in \mathbb{R}^{m \times n}$ (in this context $m < n$) is a given matrix, $u \in \mathbb{R}^m$ is the *uncertain-but-bounded* observation error, assess a *sparse* signal $w \in \mathbb{R}^n$ satisfying *sign restrictions*.

A celebrated solution to the problem is given by the ℓ_1-recovery, which amounts to taking, as an estimate of w, an optimal solution $\widehat{w}$ to the optimization problem

$$\widehat{w} \in \operatorname*{Argmin}_x \{\|x\|_1 : \ \|Ax - y\| \leq \mathrm{e}, \ x_i \geq 0 \ \forall i \in P_+, \ x_i \leq 0 \ \forall i \in P_-\} \tag{19}$$

(here e is an *a priori* bound on the norm $\|u\|$ of the observation error, $\|\cdot\|$ being some norm on $\mathbb{R}^m$). When there are no sign restrictions (i.e., $P_+ = P_- = \emptyset$), we arrive at the estimator playing the central role in the compressed sensing theory. The central result here is that when signal w is s-sparse (i.e., with at most s nonzero entries) and the matrix A possesses a certain well-defined (although difficult to verify) property, then the ℓ_1-recovery $\widehat{w}$ is close to w, provided the error bound e is small (for a comprehensive survey see [25] and references therein). Our goal here is to propose efficiently verifiable sufficient conditions on A which

allow for similar 'consistency" results, with emphasis on the case where sign restrictions are present.

To outline our results and to position them with respect to what is already known, let us start with noiseless recovery (i.e., $\mathrm{e} = 0$ and $y = Aw$). Here we are interested to answer the question:

> *Whether A is such that whenever the true signal w in (18) is s-sparse and satisfies the sign constraints $w_i \geq 0$, $i \in P_+$, $w_i \leq 0$, $i \in P_-$, the ℓ_1-recovery*
>
> $$\widehat{w} \in \underset{x}{\mathrm{Argmin}} \{\|x\|_1 : \quad Ax = y,\ x_i \geq 0\ \forall i \in P_+,\ x_i \leq 0\ \forall i \in P_-\} \tag{20}$$
>
> *recovers w exactly.*

If the answer is positive, we say that A is *s-semigood*[1].

The first characterization of s-semigoodness for the case when w is nonnegative (i.e., $P_+ = \{1, ..., n\}$) was proposed in the founding paper of Donoho and Tanner [47] in terms of neighboring properties of the polytope AS, S being the standard simplex $S = \{x \in \mathbb{R}^n : x \geq 0,\ \sum_i x_i \leq 1\}$. This paper contains also several important examples of $m \times n$ matrices which are $\lfloor \frac{m}{2} \rfloor$-semigood (here $\lfloor a \rfloor$ stands for the integer part of a) and demonstrates that various types of randomly generated matrices possess this property with overwhelming probability. Extending the results from Donoho and Huo [45], an equivalent characterization of s-semigoodness has been provided in the nonnegative case by Zhang in [133, 134], where it is shown that A is s-semigood if and only if the kernel of A, $\mathrm{Ker}A$, is *strictly half s-balanced*, meaning that for any set $I \subset \{1, ..., n\}$ of cardinality $\leq s$ it holds

$$\sum_{i \in I} z_i < \sum_{i \notin I} |z_i| \ \text{ for any } \ z \in \mathrm{Ker}A \text{ such that } z_i \leq 0, \text{ for all } i \notin I. \tag{21}$$

It should be mentioned that the necessary and sufficient conditions for s-semigoodness from (7), (21) and [47, 46] share a common drawback – they seemingly cannot be verified in a computationally efficient way.

The contributions of this chapter, which follow the approach developed in [81], are as follows.

[1] We use the term "s-semigoodness" to comply with the terminology of [81], where we used the name *s-goodness* to indicate that ℓ_1-recovery as in (20) *without* the sign restrictions is exact.

1. Taking existing characterizations of (semi)goodness (7), (21) as a starting point, we develop in Section 2.2, several equivalent necessary and sufficient conditions for s-semigoodness of a matrix A in the case of general-type sign restrictions. Then in Section 2.3, we establish error bounds for inexact ℓ_1-recovery (noisy observation (18), imprecise optimization in (19), nearly-sparse true signals); these bounds are expressed in the same terms as the necessary and sufficient conditions for s-semigoodness from Section 2.2. These bounds can be seen as an extension to the sign restricted case of bounds of Section 3 in [81] and as a special case of the bounds provided in Theorem 4.1 of [134]. To the best of our knowledge, these bounds that incorporate sign information of the signal are new.

2. The major goal of this chapter is to use the LP relaxation techniques from [81] to derive novel *efficiently verifiable* sufficient conditions for s-semigoodness. These conditions allow one to build, in a computationally efficient fashion, lower bounds on the "level of s-semigoodness" of a given matrix A, that is, on the largest $s = s_*(A)$ for which A is s-semigood with respect to given $P_\pm$. Some properties of these verifiable conditions, same as limits of their performance, are studied in Sections 2.4, 2.5, where we provide also a computationally efficient scheme for upper bounding of $s_*(A)$. In Section 2.6, we develop another efficiently computable lower bound for $s_*(A)$ by applying the SDP relaxation, similar to the approach developed in [38] for the "unsigned" case $P_\pm = \emptyset$. In Section 2.7, we report on numerical experiments aimed at comparing the "power" of our LP-based sufficient conditions for s-semigoodness, their "unsigned" prototypes from [81], and conditions based on mutual incoherence. We show that incorporating the sign information can improve the bounds on the level of s-semigoodness, and that the bounds based on LP relaxations clearly outperform the bounds based on mutual incoherence.

3. It turns out that our verifiable sufficient conditions for s-semigoodness can be expressed in terms of specific properties of the *linear* recovery $\widehat{w}^{\rm lin} = Y^T y$ associated with an appropriate $m \times n$ matrix Y. In Section 2.8, we propose and justify a new

non-Euclidean *Matching Pursuit* algorithm associated with this linear recovery.

2.2 Necessary and Sufficient Conditions for s-semigoodness

Let A be an $m \times n$ matrix, let s, $1 \leq s \leq m$, be an integer, and let P_+, P_- and P_n be a partition of $\{1, \ldots, n\}$ into three non-overlapping subsets. We say that A is *s-semigood*, if for every vector w with at most s nonzero entries satisfying $w_i \geq 0$ for $i \in P_+$, and $w_i \leq 0$ for $i \in P_-$, w is the unique optimal solution to the problem

$$\mathrm{Opt} = \min_z \{\|z\|_1 : Az = Aw,\ z_i \geq 0\ \forall i \in P_+,\ z_i \leq 0\ \forall i \in P_-\}. \tag{22}$$

Our primary goals are to find necessary and sufficient and *verifiable* sufficient conditions for A to be s-semigood.

Note that without loss of generality we may assume $P_- = \emptyset$. Indeed, by replacing the partition P_+, P_-, P_n with the partition $\overline{P}_+ = P_+ \cup P_-, \overline{P}_- = \emptyset, \overline{P}_n = P_n$ and matrix A – with the matrix $\overline{A}$ obtained from A by multiplying the columns with indices $i \in P_-$ by -1, s-semigoodness of A with respect to the original sign restrictions given by $P_\pm, P_n$ is equivalent to the s-semigoodness of the new matrix $\overline{A}$ with respect to the new sign restrictions. By this reason, *we assume from now on that* $P_- = \emptyset$. Besides this, we assume without loss of generality that $P_+ = \{1, ..., p\}$ and $P_n = \{p+1, ..., n\}$ for some p. From now on, we denote by $\mathcal{P}_n$ the set of all signals satisfying the sign restrictions:

$$\mathcal{P}_n = \{w \in \mathbb{R}^n : w_i \geq 0\ \ \forall i \in P_+\}.$$

Note that since $P_- = \emptyset$, (22) simplifies to

$$\mathrm{Opt} = \min_z \{\|z\|_1 : Az = Aw,\ z_i \geq 0\ \forall i \in P_+\}. \tag{23}$$

Let us fix a norm $\|\cdot\|$ on $\mathbb{R}^n$, and let $\|\cdot\|_*$ be the conjugate norm.

Proposition 2.2.1 *Let m, n, s and P_+ be given. The following six conditions on an $m \times n$ matrix A are equivalent to each other:*

(i) *A is s-semigood;*

(ii) *For every subset* J *of* $\{1,...,n\}$ *with* $\mathrm{Card}(J) \leq s$*, and any* $x \in \mathrm{Ker}A \backslash \{0\}$ *such that* $x_i \leq 0$ *for all* $i \in P_+ \setminus J$ *one has*

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| < \sum_{i \notin J} |x_i|.$$

(iii) *There exists* $\xi \in (0,1)$ *such that for every subset* J *of* $\{1,...,n\}$ *with* $\mathrm{Card}(J) \leq s$ *and any* $x \in \mathrm{Ker}A$ *such that* $x_i \leq 0$ *for all* $i \in P_+ \setminus J$ *one has*

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| \leq \xi \sum_{i \notin J} |x_i|.$$

(iv) *There exist* $\xi \in (0,1)$ *and* $\theta \in [1,\infty)$ *such that* A *satisfies the condition* $\mathbf{SG}_s(\xi,\theta)$ *as follows:*

for every $x \in \mathrm{Ker}A$ *and every subset* J *of* $\{1,...,n\}$ *with* $\mathrm{Card}(J) \leq s$*, one has*

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| \leq \xi \left(\sum_{i \in P_n \setminus J} |x_i| + \sum_{i \in P_+ \setminus J} \psi(x_i) \right), \;\; \psi(t) = \max[-t, \theta t],$$

or, equivalently: for all $x \in \mathrm{Ker}A$*,* $\Theta(x) \leq \xi \Psi(x)$ *where*

$$\begin{array}{l} \Theta(x) := \max\limits_{\substack{J \subset \{1,...,n\}, \\ \mathrm{Card}(J) \leq s}} \left[\sum_{i \in J \cap P_+} \max[(1-\xi)x_i, (1+\theta\xi)x_i] + \sum_{i \in J \cap P_n} (1+\xi)|x_i| \right] \\ \Psi(x) := \sum_{i \in P_+} \max[-x_i, \theta x_i] + \sum_{i \in P_n} |x_i| \end{array} \tag{24}$$

(v) *There exist* $\xi \in (0,1)$*,* $\theta \in [1,\infty)$ *and* $\beta \in [0,\infty)$ *such that* A *satisfies the condition* $\mathbf{SG}_{s,\beta}(\xi,\theta)$ *as follows:*

for every $x \in \mathbb{R}^n$ *and every subset* J *of* $\{1,...,n\}$ *with* $\mathrm{Card}(J) \leq s$*, one has*

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| \leq \beta \|Ax\| + \xi \left(\sum_{i \in P_n \setminus J} |x_i| + \sum_{i \in P_+ \setminus J} \psi(x_i) \right), \;\; \psi(t) = \max[-t, \theta t].$$

(vi) *There exist* $\xi \in (0,1)$ *and* $\beta \in [0,\infty)$ *such that* A *satisfies the condition* $\mathbf{SG}_{s,\beta}(\xi)$ *as follows:*

for every $J \subset \{1,...,n\}$ *with* $\mathrm{Card}(J) \leq s$ *and any* $x \in \mathbb{R}^n$ *such that* $x_i \leq 0$ *for all* $i \in P_+ \setminus J$*, one has*

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| \leq \beta \|Ax\| + \xi \sum_{i \notin J} |x_i|.$$

We provide the proof of Proposition 2.2.1 in Section 2.10.1.

As we have already mentioned in Introduction, when $P_n = \emptyset$ or $P_+ = \emptyset$, the characterizations (i)–(iv) of s-semigoodness are not completely new. For instance, when $P_n = \emptyset$, a necessary and sufficient condition for s-semigoodness of A in the form (ii) has been established in [133] (compare (ii) to the definition (21) of half s-balancedness of $\mathrm{Ker}A$). On the other hand, the equivalent formulation of this characterization in terms of conditions $\mathbf{SG}_{s,\beta}(\xi,\theta)$ and $\mathbf{SG}_{s,\beta}(\xi)$ seems to be new. We are about to demonstrate that the latter two conditions allow to control the error of ℓ_1-recovery in the case when the vector $w \in \mathbb{R}^n$ is not s-sparse and the problem (23) is not solved to exact optimality.

2.3 Error Bounds for Imperfect ℓ_1-recovery

We have seen that the conditions provided in Proposition 2.2.1 are responsible for s-semigoodness of a sensing matrix A, that is, for the exactness of ℓ_1-recovery in the "ideal case" when the true signal w is s-sparse, there is no observation error, and the optimization problem (23) is solved to exact optimality. Below we demonstrate that these conditions control also the error of ℓ_1-recovery in the case when the signal $w \in \mathcal{P}_n$ is not exactly s-sparse, there is observation noise and problem (23) is not solved to exact optimality. The corresponding error bound (cf [81, Proposition 3.1, Theorem 3.1]) is as follows:

Proposition 2.3.1 *Let $w \in \mathcal{P}_n$ be such that $\|w - w^s\|_1 \leq \mu$, where w^s is the vector obtained from w by replacing all but the s largest in magnitude entries in w with zeros, let y be such that $\|Aw - y\| \leq \mathrm{e}$, and let, finally, x be an approximate solution to the optimization problem*

$$\mathrm{Opt} = \min_z \{\|z\|_1 : \|Az - y\| \leq \mathrm{e},\ z_i \geq 0\ \forall i \in P_+\}. \tag{25}$$

such that $\|x\|_1 \leq \mathrm{Opt} + \nu$ and $\|Ax - y\| \leq \delta$.

1. *If A satisfies the condition $\mathbf{SG}_{s,\beta}(\xi,\theta)$ with some $\xi \in (0,1)$, $\beta \in [0,\infty)$ and $\theta \in [1,\infty)$, then*

$$\|x - w\|_1 \leq \frac{1+\xi}{1-\xi}\nu + \frac{2(1+\xi\theta)}{1-\xi}\mu + \frac{2\beta}{1-\xi}(\mathrm{e} + \delta). \tag{26}$$

2. *If A satisfies the condition $\mathbf{SG}_{s,\beta}(\xi)$ with some $\xi \in (0,1)$, $\beta \in [0,\infty)$, then*

$$\|x - w\|_1 \le \frac{1+\xi}{1-\xi}\nu + \frac{2(1+\beta\alpha)}{1-\xi}\mu + \frac{2\beta}{1-\xi}(\mathrm{e}+\delta). \tag{27}$$

where α stands for the maximum of $\|\cdot\|$-norms of the columns in A.

For proof, see Section 2.10.2.

2.4 Verifiable Conditions for s-semigoodness

In this section, our goal is to demonstrate that condition $\mathbf{SG}_{s,\beta}(\xi,\theta)$ from Proposition 2.2.1 leads to efficiently computable lower and upper bounds on the level of s-semigoodness.

2.4.1 Lower Bounding the Level of s-semigoodness

2.4.1.1 Origin of Verifiable Sufficient Condition

The essence of obtaining a lower bound on the level of s-semigoodness is in building a verifiable sufficient condition for the validity of (24), see Proposition 2.2.1.iv. By positive homogeneity of degree 1 of the convex functions Θ, Ψ participating in (24), the latter condition is equivalent to verifying

$$\mathrm{Opt} := \max_x \{\Theta(x) : Ax = 0, x \in X\} \le \xi, \text{ where } X = \{x : \Psi(x) \le 1\}. \tag{28}$$

A verifiable sufficient condition for (28) is basically the same as an efficiently computable upper bound for Opt; the sufficient condition for the validity of (28) associated with such a bound merely states that the bound is $\le \xi$. Now observe that from the origin of Ψ (see (24)) it is clear that X has a moderate number, N, of readily available extreme points $x^1, ..., x^N$ (in the case of (24), $N = 2n$), so that the only difficulty in computing Opt exactly comes from linear constraints $Ax = 0$. The standard way to circumvent this difficulty and to efficiently bound Opt from above is to use the Lagrange relaxation: for any $v \in \mathbb{R}^m$,

$$\begin{aligned}\mathrm{Opt} &= \max_{x\in X}\left\{\Theta(x) + v^T Ax : Ax = 0, x \in X\right\}\\ &\le \max_x\left\{\Theta(x) + v^T Ax : x \in X\right\} = \max_{1\le i\le N}[\Theta(x^i) + v^T Ax^i],\end{aligned}$$

and hence the efficiently computable *Lagrange relaxation bound* $\inf_v \max_{1\le i\le N}[\Theta(x^i) + v^T Ax^i]$ is an upper bound on Opt. Unfortunately, in our situation this bound can be very

poor; e.g., when X is symmetric with respect to the origin and Θ is even (as it happens in (24) when $P_+ = \emptyset$), it is immediately seen that the bound becomes the trivial bound $\text{Opt} \le \max_{x\in X} \Theta(x) = \max_i \Theta(x^i)$. In order to strengthen the relaxation, we pass to the Fenchel-type representation of Θ

$$\Theta(x) = \sup_u \left[[Pu+q]^T x - \Theta_*(u)\right]$$

with a proper convex function Θ_*; such a representation, even with $Pu + p \equiv u$, exists whenever Θ is a proper convex function (and can be easily found for Θ we are interested in). We now have for any $Y \in \mathbb{R}^{m\times n}$, $v \in \mathbb{R}^m$,

$$\begin{aligned}
\text{Opt} &= \max_x \{\Theta(x) : Ax = 0, x \in X\} \\
&= \sup_{x,u} \left\{[Pu+p]^T x - \Theta_*(u) : Ax = 0, x \in X\right\} \\
&= \sup_{x,u} \left\{[Pu+p]^T [x - Y^T Ax] + v^T Ax - \Theta_*(u) : Ax = 0, x \in X\right\} \\
&\le \sup_{x,u} \left\{[Pu+p]^T [x - Y^T Ax] + v^T Ax - \Theta_*(u) : x \in X\right\} \\
&= \max_{1\le i\le N} \underbrace{\sup_u \left\{[Pu+p]^T [x^i - Y^T Ax^i] + v^T Ax^i - \Theta_*(u)\right\}}_{:=\Theta_i(Y,v)},
\end{aligned}$$

so that the condition

$$\exists (Y \in \mathbb{R}^{m\times n}, v \in \mathbb{R}^m) : \Theta_i(Y,v) \le \xi,\ 1 \le i \le N, \tag{29}$$

is sufficient for the validity of (28). Note that the functions Θ_i, by their origin, are convex, so that the condition (29) is efficiently verifiable, provided that $\Theta_i(\cdot)$ are efficiently computable. Whenever $\Theta_*(u)$ admits a polyhedral representation, the condition (29) can be verified by solving a linear program, therefore we refer to this procedure and the corresponding bound as *linear programming based condition*.

2.4.1.2 Tractable Relaxations for Disjoint Bilinear Programming

It should be stressed that the outlined scheme of Section 2.4.1.1 can be applied to bounding from above the optimal value of a whatever problem of the form (28) with a convex polytope X and a proper convex objective Θ; all what matters is that X is given as $\text{Conv}\{x^1, ..., x^N\}$ and Θ is efficiently computable. Note also that when X is a polytope given by list of

M linear inequalities, we can efficiently represent it as the intersection of M-dimensional standard simplex and an affine plane, so that the outlined scheme is applicable to a whatever problem of maximizing an efficiently computable proper convex function under a (finite) system of linear inequality and equality constraints. Therefore we believe that the study of this new bounding scheme in the following general setting is important.

Here we consider the problem of bounding from above the quantity

$$\mathrm{Opt} = \max_{x,u} \left\{ x^T[Pu+p] : x \in X, Ax = 0, u \in U \right\}, \quad X = \mathrm{Conv}\{x^1, ..., x^N\}, \tag{30}$$

where $x^i \in \mathbb{R}^n$, the set $\{x \in X : Ax = 0\}$ is nonempty, and $U \subset \mathbb{R}^n$ is a computationally tractable compact convex set which contains the origin in its interior.

In this setting the linear programming based relaxation scheme corresponds to

$$\mathrm{Opt}^+ = \inf_{Y,v} \max_{1 \le i \le N} \left[\max_{u \in U} [(I - Y^T A)x^i]^T [Pu+p] + v^T A x^i \right], \tag{31}$$

and we have seen that $\mathrm{Opt} \le \mathrm{Opt}^+$.

In addition to its simple derivation, Opt^+ has a meaningful interpretation given by the following

Proposition 2.4.1 *Whenever X is given as $\mathrm{Conv}\{x^1, ..., x^N\}$ and $U \subset \mathbb{R}^n$ is a computationally tractable compact convex set which contains the origin in its interior*

$$\mathrm{Opt}^+ = \max_V \left\{ \mathrm{Tr}(V) : \ \exists \bar{x}, \bar{u} : [\bar{x}, \bar{u}, V] \in \mathcal{W}, AV = 0, A\bar{x} = 0 \right\},$$

where $\mathcal{W} = \mathrm{Conv}\{[x, u, x(Pu+p)^T] : x \in X, u \in U\}$. Moreover $\mathcal{W}$ is a computationally tractable set given by

$$\mathcal{W} = \left\{ (x,u,V) : \begin{array}{l} \exists (t_i, w^i) : x = \sum_{i=1}^N t_i x^i, u = \sum_{i=1}^N w^i, V = \sum_{i=1}^N x^i [Pw^i + t_i p]^T, \\ \sum_{i=1}^N t_i = 1, \phi(w^i) \le t_i, \forall i = 1, \ldots, N \end{array} \right\},$$

where $\phi(\cdot)$ is the Minkowski functional of U, meaning $U = \{u : \phi(u) \le 1\}$.

For proof, see Section 2.10.3.

The disjoint bilinear optimization problem given in (30) has broad range of applications and developing tractable relaxations for particular forms of it has been studied extensively.

One of the most commonly studied form is the case when the set U is given as a polytope. In the polytope case, we present the relation of our bounding schema and other relaxations such as McCormick Bounds, Sherali-Adams Relaxation and Lovász-Schrijver Relaxation in Section 2.9. In particular, we show that when U is a polytope, the bound produced by Opt^+ is at least as good as the one given by McCormick Bounds and Sherali-Adams Relaxation, and a simple addition to our bounding schema makes it at least as strong as the Lovász-Schrijver Relaxation.

2.4.1.3 Verifiable Sufficient Conditions for s-semigoodness by Linear Programming

The simple and general construction of tractable linear programming based relaxation for disjoint bilinear programs presented in Section 2.4.1.1 can be combined with the condition $\mathbf{SG}_{s,\beta}(\xi,\theta)$ from Proposition 2.2.1 to obtain an efficiently computable lower bound on the level of s-semigoodness. In the case we are interested in, the extreme points of X are the $2n$ vectors $-e_i$ for $1 \le i \le n$, e_i for $i \in P_n$, and $\theta^{-1}e_i$ for $i \in P_+$, where e_i is the i-th basic orth. Implementing the outlined bounding scheme and adding additional restrictions to get a control over β, we arrive at the following verifiable sufficient condition, $\mathbf{VSG}_s(\xi,\theta,\rho,\sigma)$.

Let

$$\mathcal{U}_s = \{u \in \mathbb{R}^n : \ \|u\|_1 \le s, \ \|u\|_\infty \le 1\},$$

so that $\mathcal{U}_s$ is the convex hull of all $\{-1,0,1\}$ vectors with at most s nonzero entries, and for $x \in \mathbb{R}^n$, let $\|x\|_{s,1}$ be the sum of the s largest magnitudes of entries in x, or, equivalently,

$$\|x\|_{s,1} = \max_{u \in \mathcal{U}_s} u^T x.$$

Let

$$(D_\theta[x])_i = \begin{cases} [1+\theta\xi]\max[x_i,0], & i \in P_+ \\ (1+\xi)|x_i|, & i \notin P_+ \end{cases}, \qquad \Phi(x) = \|D_\theta[x]\|_{s,1}.$$

Suppose $\xi \in [0,1)$, $\theta \in [1,\infty)$ and $\rho,\sigma \in [0,\infty)$ are given. Consider the following condition on an $m \times n$ matrix A:

$\mathbf{VSG}_s(\xi,\theta,\rho,\sigma)$: *There exist $m \times n$ matrix $Y = [y_1,...,y_n]$ and a vector $v \in \mathbb{R}^m$*

such that

$$\begin{array}{rcll} \Phi_s(-C_i[Y,A]) + (A^T v)_i & \leq & \xi, \ 1 \leq i \leq n & (a) \\ \Phi_s(C_i[Y,A]) - (A^T v)_i & \leq & \xi, \ i \notin P_+ & (b) \\ \Phi_s(C_i[Y,A]) - (A^T v)_i & \leq & \theta\xi, \ i \in P_+ & (c) \\ \|y_i\|_* & \leq & \sigma, \ 1 \leq i \leq n & (d) \\ \|v\|_* & \leq & \rho & (e) \end{array} \tag{32}$$

where $C_i[Y,A]$ is the i-th column of the matrix $I - Y^T A$.

Observe that this condition is verifiable, since (32) is a system of explicit convex constraints on Y and v.

Proposition 2.4.2 *Let A satisfy $\mathbf{VSG}_s(\xi,\theta,\rho,\sigma)$ with some $\xi \in [0,1)$, $\theta \in [1,\infty)$, and $\rho,\sigma \in [0,\infty)$. Then A satisfies $\mathbf{SG}_{s,\beta}(\xi,\theta)$ with*

$$\beta = \rho + \sigma \max_{k_+,k_n} \left\{ k_+(1+\theta\xi) + k_n(1+\xi) : \begin{array}{l} 0 \leq k_+ \leq \mathrm{Card}(P_+) \\ 0 \leq k_n \leq \mathrm{Card}(P_n) \\ k_+ + k_n \leq s \end{array} \right\} \leq \rho + \sigma s(1+\theta\xi). \tag{33}$$

In particular, A is s-semigood.

For proof, see Section 2.10.4.

Some comments are in order.

Effect of increasing β,θ,ξ. The condition $\mathbf{SG}_{s,\beta}(\xi,\theta)$ appearing in Proposition 2.2.1.v clearly is "monotone" in the parameters β,θ,ξ: whenever A satisfies this condition and $\beta' \geq \beta$, $\theta' \geq \theta$ and $\xi' \geq \xi$, A satisfies the condition $\mathbf{SG}_{s,\beta'}(\xi',\theta')$ as well. Proposition 2.4.2 offers a verifiable sufficient condition for the validity of $\mathbf{SG}_{s,\beta}(\xi,\theta)$, specifically,

$\mathbf{VSG}^*_{s,\beta}(\xi,\theta)$: $\exists Y, v$ ρ, σ satisfying (32) and the relation $\rho + \sigma s(1+\theta\xi) \leq \beta$.

A natural question is, whether this verifiable condition possesses the same monotonicity properties as the "target" condition $\mathbf{SG}_{s,\beta}(\xi,\theta)$. In the case of the affirmative answer, in order to conclude that A is s-semigood, we could check the validity of $\mathbf{VSG}^*_{s,\beta}(\xi,\theta)$ for appropriately large values of β,θ and a close to one value of $\xi < 1$; if the condition

is satisfied, A is s-semigood, and error bounds from Proposition 2.3.1 take place. Were the condition $\mathbf{VSG}^*_{s,\beta}(\xi,\theta)$ "not monotone," to justify the s-semigoodness of A via this condition would require a problematic and time-consuming search in the space of parameters β,θ,ξ. Fortunately, the condition $\mathbf{VSG}^*_{s,\beta}(\xi,\theta)$ indeed is monotone:

Proposition 2.4.3 *Let A satisfy $\mathbf{VSG}^*_{s,\beta}(\xi,\theta)$, and let Y,v,σ,ρ be the corresponding certificate, that is, $\rho+\sigma s(1+\theta\xi)\le\beta$ and Y,v,σ,ρ satisfy (32). Then A satisfies $\mathbf{VSG}^*_{s,\beta'}(\xi',\theta')$ whenever $\beta'\ge\beta$, $\theta'\ge\theta$ and $\xi'\in(\xi,1)$, the certificate being (Y',v,σ,ρ), where the columns Y'_i of Y' are multiplies of the columns Y_i of Y, namely,*

$$Y'_i=a_iY_i;\ \ [0,1]\ni a_i=\begin{cases}(1+\xi\theta)/(1+\xi'\theta'), & i\in P_+\\ (1+\xi)/(1+\xi'), & i\in P_n\end{cases}$$

For proof, see Section 2.10.6.

Relation to the sufficient condition for s-goodness from [81] and the Restricted Isometry Property. The verifiable sufficient condition for s-goodness from [81] requires from an $m\times n$ matrix A the existence of $\gamma<1/2$ and $Y=[y_1,...,y_n]\in\mathbb{R}^{m\times n}$ such that

$$\|C_i[Y,A]\|_{s,1}\le\gamma,\ \ \text{for all } 1\le i\le n,$$

Setting $\theta=1$ and $\xi=\frac{\gamma}{1-\gamma}$ (so that $\xi<1$ and $\gamma=\frac{\xi}{1+\xi}$) and taking into account that in the case of $\theta=1$ we have $\Phi_s(z)\le(1+\xi)\|z\|_{s,1}$, the latter condition implies that

$$\Phi_s(\pm C_i[Y,A])\le(1+\xi)\gamma=\xi,\ \ \forall i,$$

that is, it implies the validity of $\mathbf{VSG}_s(\xi,1,0,\sigma)$, provided that σ is large enough, specifically, $\sigma\ge\|y_i\|_*$ for all i.

As it was shown in [81], when A satisfies the *Restricted Isometry Property* $\mathrm{RIP}(\delta,k)$ with parameters $\delta\in(0,1)$, $k>1$, the above sufficient condition for s-goodness is satisfied with $\gamma=1/3$ for s as large as $O(1)(1-\delta)\sqrt{k}$; as a result, a $\mathrm{RIP}(\delta,k)$-matrix satisfies $\mathbf{VSG}_s(\frac12,1,0,\sigma)$ provided that σ is large enough and $s\le O(1)(1-\delta)\sqrt{k}$. Since for large m,n, $m<n$, typical random matrices possess, with overwhelming probability, property $\mathrm{RIP}(\frac12,k)$ with k as large as $O(1)m/\ln(n/m)$, we see that our verifiable sufficient condition

for s-semigoodness can certify the latter property for s as large as $O(1)\sqrt{m/\ln(n/m)}$, provided that the matrix in question is "good enough".

2.4.2 Upper Bounding the Level of s-semigoodness

Here we address the issue of bounding from above the maximal $s = s_*(A)$ for which A is s-semigood. The construction to follow is motivated by item (iv) of Proposition 2.2.1. A necessary and sufficient condition for the s-semigoodness of A is the existence of $\xi < 1$ and $\theta \geq 1$ such that for all $x \in \mathrm{Ker}A$ and any set I of indices with $\mathrm{Card}(I) \leq s$

$$\sum_{i \in I \cap P_+} \max[(1-\xi)x_i, (1+\theta\xi)x_i] + \sum_{i \in I \cap P_n} (1+\xi)|x_i| \leq \xi \Psi(x)$$

where

$$\Psi(x) = \sum_{i \in P_+} \max[-x_i, \theta x_i] + \sum_{i \in P_n} |x_i|, \tag{34}$$

or, equivalently,

> (!) for every $x \in \mathrm{Ker}A$ and every vector v with at most s nonzero entries and nonzero entries v_i belonging to $[1-\xi, 1+\xi\theta]$ if $i \in P_+$ and belonging to $[-1-\xi, 1+\xi]$ if $i \in P_n$, one has
>
> $$v^T x \leq \xi \Psi(x).$$

Observe that the convex hull of the vectors v in question is exactly the set

$$\mathcal{U}^{\xi,\theta} = \left\{ v \in \mathbb{R}^n : \begin{array}{c} 0 \leq v_i \leq 1+\theta\xi,\ i \in P_+,\ |v_i| \leq 1+\xi,\ i \in P_n, \\ \sum_{i \in P_+} \frac{v_i}{1+\theta\xi} + \sum_{i \in P_n} \frac{|v_i|}{1+\xi} \leq s \end{array} \right\}.$$

Recalling that $P_+ = \{1, ..., p\}$, setting $q = n - p = \mathrm{Card}(P_n)$ and

$$\mathcal{U} = \{u \in \mathbb{R}^n : \ \|u\|_1 \leq s,\ \|u\|_\infty \leq 1,\ u_i \geq 0 \text{ for } i \in P_+\} \tag{35}$$

we see that

$$\mathcal{U}^{\xi,\theta} = C^{\xi,\theta}\mathcal{U}, \quad \text{where} \quad C^{\xi,\theta} = \left[\begin{array}{c|c} (1+\xi\theta)I_p & 0 \\ \hline 0 & (1+\xi)I_q \end{array}\right]. \tag{36}$$

The condition (!) now reads

$$\max_{v\in\mathcal{U}^{\xi,\theta}} v^Tx \le \xi\Psi(x) \ \text{ for all } x\in \mathrm{Ker}A.$$

Setting $\mathcal{X}=\{x\in \mathrm{Ker}A:\ \Psi(x)\le 1\}$ the latter condition, by homogeneity reason, is the same as

$$\mathrm{Opt}=\mathrm{Opt}(\xi,\theta):=\max_{v,x}\Big\{v^Tx:\ v\in\mathcal{U}^{\xi,\theta},\ x\in\mathcal{X}\Big\}\le\xi; \tag{37}$$

recall that A is s-semigood if and only if there exist $\theta\ge 1$ and $\xi<1$ such that (37) takes place.

We can use (37) in order to bound $s_*(A)$ from above, as follows. In order to certify that $s_*(A)<s$ for a given s (s is the input to our algorithm), we fix a large θ and a close to one $\xi<1$ (these are the parameters of the algorithm) and run the iterations

$$u_0\in\mathcal{U}^{\xi,\theta}\mapsto x_1\in \mathrm{Argmax}_{x\in\mathcal{X}}u_0^Tx\mapsto u_1\in\mathrm{Argmax}_{u\in\mathcal{U}^{\xi,\theta}}u^Tx_1\mapsto ...$$

initiating them by a picked at random vertex u_0 of $\mathcal{U}^{\xi,\theta}$. Note that the quantities $u_i^Tx_i$, $i=1,2,...$ clearly form a nondecreasing sequence of lower bounds on Opt. We terminate the outlined iterations when the progress in the bounds – the difference $u_i^Tx_i-u_{i-1}^Tx_{i-1}$ – falls below a given small threshold, and we run this process a predetermined number of times from different randomly chosen starting points. As a result, we get a set of lower bounds on Opt of the form u^Tx, where u is a vertex of $\mathcal{U}^{\xi,\theta}$ and $x\in\mathcal{X}$. If our goal were merely to certify that (40) is not valid for given s,θ,ξ, we could terminate this process at the first step, if any, when the current lower bound u^Tx becomes $>\xi$ (cf. [81, Section 4.1]). We, however, want to certify that $s>s_*(A)$, or, which is the same by Proposition 2.2.1.iv, that (40) fails to be true for *all* θ and all $\xi<1$, and not only for those θ,ξ we have selected for our test. To overcome this difficulty, we accompany every step $u\mapsto x\in\mathrm{Argmax}_{x\in\mathcal{X}}u^Tx$ by an additional computation as follows. In our process, u is an extreme point of $\mathcal{U}^{\xi,\theta}$, that is, a point with $s_u\le s$ nonzero entries, let the set of indices of these entries be I. Setting

$\epsilon_i = \mathrm{sign}(u_i)$, we solve the following LP problem

$$\max_x \left\{ \sum_{i\in I\cap P_+} x_i + \sum_{i\in I\cap P_n} \epsilon_i x_i : \left\{ \begin{array}{l} x_i \le 0,\ i \in P_+\backslash I \\ Ax = 0 \\ \sum_{i\notin I} |x_i| \le 1 \end{array} \right\} \right.$$

If the optimal value in this problem is ≥ 1, we terminate our test and claim that A is not s-good; by Proposition 2.2.1.ii, this indeed is the case.

As applied to a given input s, the outlined test either terminates with a valid claim "$s > s_*(A)$", or terminates with no conclusion at all, in which case we could pass to testing a larger value of s.

2.5 *Limits of Performance of LP-based Sufficient Conditions for s-semigoodness*

Unfortunately, the condition in question, same as its predecessor from [81], can*not* certify s-semigoodness of an $m \times n$ matrix in the case of $s > O(1)\sqrt{m}$, unless the matrix is "nearly square". The precise statement is as follows (cf. [81, Proposition 4.2]):

Proposition 2.5.1 *Let*

$$n > 2(2\sqrt{2m}+1)^2 \tag{38}$$

and let $\xi < 1, \theta \ge 1, \sigma \ge 0, \rho \ge 0$, an integer s and an $m \times n$ matrix A be such that A satisfies $\mathbf{VSG}_s(\xi,\theta,\rho,\sigma)$. Then

$$s \le 2\sqrt{2m}+1. \tag{39}$$

For proof, see Section 2.10.5.

The results from Proposition 2.5.1 show that our verifiable sufficient conditions can only certify s-semigoodness of an $m \times n$ matrix at a suboptimal rate of $s \le O(1)\sqrt{m}$, unless the matrix is "nearly square". In fact this verifiable bound can still give a very poor impression on the true largest $s = s_*(A)$ for which A is s-semigood. An instructive example in this direction is as follows. Consider the case of $P_+ = \{1, ..., n\}$, let $m = 2d+1$ be odd, and let the rows of A be comprised of the values of basic trigonometric polynomials

$$p_0(\phi) \equiv 1,\ \ p_{2i-1}(\phi) = \cos(i\phi),\ \ p_{2i}(\phi) = \sin(i\phi),\ \ 1 \le i \le d,$$

taken along the regular grid $\phi_j = 2\pi j/n$, $0 \le j < n$, so that $A_{ij} = p_i(\phi_j)$, $0 \le i < m$, $0 \le j < n$ (we enumerate rows and columns starting with 0 rather than with 1). It is well known [30, 47] that in this case A is s-semigood for $s = d$. In contrast to this, when A is not "nearly square", specifically, when $n > 4\pi d$, A can satisfy the condition $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ only for $s \le 2$, no matter how large θ, σ, ρ are and how close to 1, $\xi < 1$ is. The validity of this claim is readily given by the following

Lemma 2.5.1 *For any positive integer d, let $n \ge 4\pi d$, and A be the matrix obtained from the basic trigonometric polynomials as described in Section 2.5, then the condition $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ can hold true for $s \le 2$ only.*

For proof, see Section 2.10.7.

2.6 Verifiable Sufficient Conditions for s-semigoodness by Semidefinite Relaxation

Following d'Aspremont and El Ghaoui [38], we are about to derive another verifiable sufficient condition for s-semigoodness, now - via semidefinite relaxation. The construction to follow is motivated by the development in the beginning of Section 2.4.2, according to which s-semigoodness of A is implied by the validity of (37) for $\theta > 1$ and $\xi < 1$.

Let, as before,

$$\mathcal{X} = \{x \in \mathrm{Ker}A : \ \Psi(x) \le 1\} \ \text{ and } \ \mathcal{U}^{\xi,\theta} = \{C^{\xi,\theta}u : \ u \in \mathcal{U}\},$$

where $\Psi, \mathcal{U}$ and $C^{\xi,\theta}$ are defined in, respectively, (34), (35) and (36). The condition (37) is equivalent to

$$\max_{u,x} \Big\{(C^{\xi,\theta}u)^T x : \ u \in \mathcal{U}, \ x \in \mathcal{X}\Big\} \le \xi. \tag{40}$$

Observe that for $x \in \mathcal{X}$, $u \in \mathcal{U}$ the matrices $Z = xx^T$, $V = xu^T$ and $Q = uu^T$ satisfy the

relations

$$
\begin{array}{ll}
\multicolumn{2}{l}{\exists t \in \mathbb{R}^n, R \in \mathbf{S}^{2n}, \Lambda \in \mathbf{S}^{2n}:} \\
(a) & G = \left[\begin{array}{c|c|c} 1 & x^T & u^T \\ \hline x & Z & V \\ \hline u & V^T & Q \end{array}\right] \succeq 0; \\
(b) & \left\{\begin{array}{l} Q = \underbrace{\left[\begin{array}{c|c} I_n & -I_n \end{array}\right]}_{:=L} \underbrace{\left[\begin{array}{c|c} R^{11} & R^{12} \\ \hline R^{12} & R^{11} \end{array}\right]}_{:=R} L^T, \\ 0 \le R_{ij} \le \frac{1}{2}, \quad R \succeq 0, \quad R^{12} = [R^{12}]^T, \mathrm{Tr}(R) \le s, \\ \sum_{i,j} R_{ij} \le s^2, \quad R^{12}_{ij} = 0 \; \forall i,j \in P_+; \end{array}\right. \\
(c) & Z = \underbrace{\left[\begin{array}{c|c|c|c} -I_p & 0 & \frac{1}{\theta} I_p & 0 \\ \hline 0 & -I_q & 0 & I_q \end{array}\right]}_{:=F} \Lambda F^T, \quad 0 \le \Lambda_{ij}, \quad \Lambda \succeq 0, \quad \sum_{i,j} \Lambda_{ij} \le 1; \\
(d_1) & \left\{\begin{array}{ll} -t_i \le V_{ij} \le \frac{t_i}{\theta}, & \forall i,j \in P_+ \\ |V_{ij}| \le t_i, & \text{otherwise}; \end{array}\right. \\
(d_2) & \sum_{j \in P_+} \max[-V_{ij}, \theta V_{ij}] + \sum_{j \in P_n} |V_{ij}| \le s t_i, \forall i \in P_+; \\
(d_3) & \sum_j |V_{ij}| \le s t_i, \forall i \in P_n; \\
(d_4) & \sum_i t_i \le 1; \\
(e) & AZA^T = 0 \\
(f) & x \in \mathcal{X}, \quad u \in \mathcal{U}.
\end{array}
\tag{41}
$$

Besides this,

$$
u^T (C^{\xi,\theta})^T x = \mathrm{Tr}(C^{\xi,\theta} V).
$$

Indeed, the latter relation, same as (41.a), (41.e) and (41.f), is evident. To verify

(41.b), let $u_+ = \max[u, 0]$, $u_- = \max[-u, 0]$, where max is acting coordinate-wise. Then

$$\begin{aligned} Q &= L \left[\begin{array}{c|c} u_+u_+^T & u_+u_-^T \\ \hline u_-u_+^T & u_-u_-^T \end{array}\right] L^T = L \left[\begin{array}{c|c} u_-u_-^T & u_-u_+^T \\ \hline u_+u_-^T & u_+u_+^T \end{array}\right] L^T \\ &= \underbrace{L \left[\begin{array}{c|c} \frac{1}{2}[u_+u_+^T + u_-u_-^T] & \frac{1}{2}[u_+u_-^T + u_-u_+^T] \\ \hline \frac{1}{2}[u_-u_+^T + u_+u_-^T] & \frac{1}{2}[u_-u_-^T + u_+u_+^T] \end{array}\right]}_{R} L^T, \end{aligned}$$

and the matrix R we have just defined clearly satisfies all requirements from (41.b). To verify (41.c), observe that the extreme points of the set $\mathcal{X}^+ = \{x : \Psi(x) \le 1\} \supset \mathcal{X}$ are the vectors $\pm e_i$, $i > p$, and $-e_i, \theta^{-1}e_i$, $i \le p$, so that $x = F\lambda$ with $\lambda \in \mathbb{R}^{2n}_+$, $\sum_i \lambda_i \le 1$; setting $\Lambda = \lambda\lambda^T$, we satisfy (41.$c$). To satisfy (41.$d$), it suffices to set $t_i = |x_i|$ for all $i > p$ and $t_i = \max[-x_i, \theta x_i]$ for $i \le p$ and to take into account that $\max[-V_{ij}, \theta V_{ij}] \ge |V_{ij}|$ for all i, j due to $\theta \ge 1$, and that $u_i \ge 0$ for $i \in P_+$.

It follows that a sufficient condition for (40) is

$$\mathrm{Opt}^{\xi,\theta} := \max_{\substack{Z, Q \in \mathbf{S}^n,\ R, \Lambda \in \mathbf{S}^{2n}, \\ V \in \mathbb{R}^{n\times n},\ t \in \mathbb{R}^n}} \left\{ \mathrm{Tr}(C^{\xi,\theta}V) : \ (41) \text{ is satisfied} \right\} \le \xi. \tag{42}$$

The optimization problem in (42) clearly reduces to a semidefinite maximization program $\mathcal{S}$; by weak duality, the optimal value in the semidefinite dual $\mathcal{D}$ to $\mathcal{S}$ is $\ge \mathrm{Opt}^{\xi,\theta}$. It follows that the efficiently verifiable condition

$$\mathrm{Opt}(\mathcal{D}) \le \xi$$

is a sufficient condition for s-semigoodness of A. Note that the above construction depends on $\theta \ge 1$ and $\xi < 1$ as parameters.

Remark. Consider the case of $P_+ = \emptyset$, where $\mathcal{X} = \{x \in \mathbb{R}^n : \|x\|_1 \le 1, Ax = 0\} \supset \mathcal{Z} = \{x \in \mathbb{R}^n : \|x\|_1 \le 1\}$. In this case, the standard semidefinite relaxation of the set

$\mathcal{C}_* = \mathrm{Conv}\{xx^T : x \in \mathcal{Z}\}$ is

$$\mathcal{C} = \left\{ Z : \ Z \succeq 0, \sum_{i,j} |Z_{ij}| \leq 1 \right\}$$

(cf. [38]). Note that (41.c) uses another semidefinite relaxation of $\mathcal{C}_*$, namely,

$$\mathcal{C}' = \left\{ Z : \exists \Lambda \in \mathbf{S}^{2n} : \begin{array}{l} \Lambda \succeq 0, \Lambda_{i,j} \geq 0 \ \forall i,j, \ \sum_{i,j} \Lambda_{ij} \leq 1 \\ Z = [I_n, -I_n]\Lambda[I_n, -I_n]^T \end{array} \right\}.$$

It is immediately seen that $\mathcal{C}_* \subset \mathcal{C}' \subset \mathcal{C}$; a surprising fact is that the second of these inclusions is strict. Thus, the relaxation of $\mathcal{C}_*$ given by $\mathcal{C}'$ is less conservative than the standard relaxation given by $\mathcal{C}$. As observed by A. d'Aspremont (private communication), the relaxation $\mathcal{C}'$ can be further improved, namely, by replacing $\mathcal{C}'$ with

$$\mathcal{C}^+ = \left\{ Z : \exists \Lambda = \left[\begin{array}{cc} \Lambda^{11} & \Lambda^{12} \\ \Lambda^{21} & \Lambda^{22} \end{array} \right] \in \mathbf{S}^{2n} : \begin{array}{l} \Lambda^{\mu\nu} \in \mathbb{R}^{n\times n}, \ \Lambda \succeq 0, \ \Lambda_{i,j} \geq 0 \ \forall i,j \\ \sum_{i,j} \Lambda_{ij} \leq 1, \ \Lambda^{12}_{ii} = 0, \ 1 \leq i \leq n \\ Z = [I_n, -I_n]\Lambda[I_n, -I_n]^T \end{array} \right\}.$$

Note that this idea can be used to improve the semidefinite relaxation given by $\mathcal{C}$ as well. Specifically, the matrix R as built in the justification of (41) clearly satisfies $(R^{12})_{ii} = 0$, $1 \leq i \leq n$, and we can add these linear constraints on R to (41.b). Similarly, when representing a vector $x \in \mathcal{X}^+$ as $F\lambda$ with $\lambda \in \mathbb{R}^{2n}_+$, $\sum_i \lambda_i \leq 1$, see the justification of (41), we clearly can ensure that $\lambda_i \lambda_{n+i} = 0$, $1 \leq i \leq n$, that is, the matrix Λ we have built in fact satisfies $\Lambda_{i,n+i} = \Lambda_{n+i,i} = 0$, $1 \leq i \leq n$, and we can add these linear constraints on Λ to (41.c).

Proposition 2.6.1 *If* $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ *with* $\rho = \sigma = \infty$ *holds, then* $\mathrm{Opt}^{\xi,\theta} \leq \xi$.

For proof, see Section 2.10.8.

Although Proposition 2.6.1 states that the verifiable sufficient conditions based on semidefinite programming are at least as good as the ones based on linear programming, i.e., $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$, in terms of their computational cost, conditions based on linear programming are far more advantageous.

2.7 Numerical Results

In order to compare the performance of the proposed bounds on the maximal $s = s_*(A)$ for which a given matrix, A, is s-semigood, with the bounds known from the literature, we present some preliminary numerical results for relatively small sensing matrices. Our goal is to see if the sign information on a signal allows to improve the bounds for $s_*(A)$ as compared to the bounds on the largest $s = s_0(A)$ for which A is s-good.

We generate four sets of random matrices, which are normalizations (all columns scaled to be of $\|\cdot\|_2$-norm 1) of (a) Rademacher matrices (i.i.d. entries taking values ± 1 with probabilities 0.5), (b) Gaussian matrices (i.i.d. $\mathcal{N}(0,1)$ entries), (c) Fourier matrices — $m \times n$ submatrices of the matrix of $n \times n$ Discrete Fourier Transform, and (d) Hadamard matrices — $m \times n$ submatrices of the $n \times n$ Hadamard matrix[2]; in the cases (c,d), the m rows comprising the submatrix were drawn at random from the n rows of the "parent" matrix. For each type, we set the number of columns to $n = 128$ and $n = 256$ and vary the number of rows, $m = 0.5n, \ldots, 0.95n$.

We bound from below the value $s_0(A)$ using the bound $s[\mu]$ by mutual incoherence and the bounds $s[\alpha_1]$ and $s[\alpha_s]$, computed through the LP-based verifiable sufficient conditions for s-goodness (see [81, Section 6]).

The lower bound on $s_*(A)$ is computed by invoking condition $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$, where $\rho = \sigma = \infty$ and θ is set to once for ever fixed "large enough" value, and ξ is set to 0.9999, see Section 2.4.1.3 and Propositions 2.4.2, 2.4.3. Note that given a matrix Y, and setting $v = 0$, one can compute the largest s satisfying (32) and thus ensuring the validity of $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$. We first compute the best lower bound $\underline{s}$ on $s_*(A)$ given by the Y-matrices generated when bounding $s_0(A)$. Then we compute the "improved" lower bound for $s_*(A)$ as follows: we check whether the condition $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ holds true for $s = \underline{s} + 1$, if it is the case, check whether this condition holds true for $s = \underline{s} + 2$, and so on.

While the outlined lower bounds on $s_*(A)$ and $s_0(A)$ are efficiently computable via LP (when $\sigma = \rho = \infty$, the sufficient condition is easily checked by solving a Linear Programming

[2]The Hadamard matrix H_d, $d = 0, 1, 2, ...$, has order $2^d \times 2^d$ and is given by the recurrence $H_0 = 1$, $H_{d+1} = [H_d, H_d; H_d, -H_d]$.

Table 1: Comparison of efficiently computable bounds on $s_*(A)$, $n = 128$, *Fourier* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
64	2	3	3	6	3	14	0.4	70.0	22.2	258.9	44.9
64	2	3	3	5	3	14	0.2	61.7	13.7	249.4	52.5
76	1	3	3	5	3	21	0.3	163.1	12.0	239.6	22.6
76	2	4	4	6	4	23	0.5	142.2	10.3	237.7	21.7
88	2	4	4	7	4	20	0.2	132.9	31.5	175.4	68.3
88	2	5	5	8	5	28	0.2	63.2	31.7	259.4	25.1
102	3	6	6	11	6	32	0.2	69.9	39.2	223.8	28.3
102	2	5	5	9	5	25	0.5	70.2	35.9	255.1	48.7
114	3	6	6	11	6	34	0.5	52.7	43.7	249.2	57.0
114	3	7	7	12	7	33	0.1	69.4	42.9	228.3	56.1
120	3	7	7	14	7	40	0.2	64.7	42.6	255.3	30.5
120	2	6	6	12	7	34	0.2	79.5	39.6	494.2	29.2

Table 2: Comparison of efficiently computable bounds on $s_*(A)$, $n = 128$, *Hadamard* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
64	2	4	4	6	4	8	0.1	156.3	11.0	237.6	57.6
64	2	3	3	3	4	27	0.1	59.3	1.7	211.4	14.1
76	2	5	5	7	5	7	0.4	76.0	13.6	230.1	23.4
76	3	5	5	7	5	19	0.1	70.4	12.7	253.6	24.5
88	3	6	6	7	6	7	0.0	72.9	15.1	191.1	29.1
88	3	6	6	7	6	9	0.0	99.1	15.2	519.3	43.7
102	4	8	8	13	8	16	0.2	65.1	55.1	240.4	36.2
102	4	9	9	15	9	23	0.1	65.9	37.6	290.1	39.8
114	6	13	13	15	13	30	0.1	125.6	17.5	272.6	37.6
114	6	13	13	15	13	20	0.3	128.6	16.8	276.1	17.8
120	7	15	15	15	15	15	0.3	129.7	5.8	179.3	46.6
120	7	15	15	15	15	23	0.1	129.1	0.6	178.2	36.2

program), the sizes of the resulting LPs are rather large. For instance, when A is $m \times n$, the LP associated with (32) has a $(2n^2 + 2n + 1) \times ((m + 2n)(n + 1) + 2)$ constraint matrix (compared to $(2n^2 + n) \times (n(m + n + 1) + 1)$ constraint matrices arising when computing lower bounds for $s_0(A)$). For instance, for $m = 230$ and $n = 256$, bounding $s_*(A)$ results in an LP program of the size $131,585 \times 190,696$, while computing a lower bound on $s_0(A)$ requires solving an LP problem of size $131,328 \times 124,673$. In all the computations, we used the state-of-the-art commercial LP solver `mosekopt` [4].

The upper bounds on $s_*(A)$ and on $s_0(A)$ are computed by the techniques from Section 2.4.2 and [81, Section 4.1].

The results of our experiments and related CPU times are presented in Tables 1-8. The computations were carried out on a single core of an 8-core Intel Xeon E5520@2.27GHz CPU Linux workstation.

Table 3: Comparison of efficiently computable bounds on $s_*(A)$, $n = 128$, *Rademacher* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
64	1	3	3	7	4	20	3.2	240.6	23.7	690.0	20.9
64	1	3	3	7	4	21	2.9	258.4	24.6	752.7	20.4
76	1	4	4	8	5	29	4.1	226.1	32.7	755.0	23.5
76	1	4	4	9	5	22	3.8	250.2	32.1	830.1	23.2
88	1	5	5	10	6	36	4.7	124.1	63.8	935.6	38.0
88	1	5	5	12	6	37	4.4	115.3	41.7	850.0	28.9
102	1	7	7	15	8	42	5.4	88.9	71.7	927.3	33.4
102	1	6	7	15	8	45	5.2	172.0	70.7	884.0	33.4
114	1	9	11	19	13	51	5.8	222.2	47.1	1492.2	25.0
114	1	9	11	18	13	50	5.8	193.2	46.3	1505.5	50.6
120	2	12	14	22	18	51	6.4	235.6	44.4	2236.1	27.7
120	1	12	14	22	18	52	5.9	231.5	43.4	2192.3	28.1

Table 4: Comparison of efficiently computable bounds on $s_*(A)$, $n = 128$, *Gaussian* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
64	1	3	3	6	4	20	3.0	189.9	39.5	904.7	21.6
64	1	3	3	7	4	21	3.1	189.0	25.5	994.0	21.8
76	1	4	4	9	5	34	3.8	217.4	32.9	937.0	23.3
76	1	4	4	9	5	25	3.5	207.6	32.2	781.5	23.9
88	1	5	5	12	6	37	4.2	123.8	44.3	896.1	29.6
88	1	5	5	12	6	36	4.5	128.8	42.6	857.2	28.4
102	1	6	7	16	8	46	5.2	163.2	48.6	988.2	31.7
102	1	6	7	15	8	42	5.0	168.8	52.5	990.9	44.8
114	1	10	11	19	13	49	6.3	336.9	53.8	1547.5	24.0
114	1	9	10	18	13	50	6.3	212.2	50.3	1546.1	36.5
120	1	12	14	23	18	52	5.8	213.9	48.0	2455.3	26.4
120	1	13	15	23	18	52	5.7	199.8	45.2	1768.0	38.6

Table 5: Comparison of efficiently computable bounds on $s_*(A)$, $n = 256$, *Fourier* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
128	3	5	5	12	5	47	0.8	1054.0	146.0	3114.4	172.9
128	3	5	5	11	5	32	0.9	986.0	169.4	2891.5	311.5
152	2	6	6	11	6	49	1.1	898.5	252.5	3680.2	179.6
152	3	6	6	11	6	53	1.3	899.3	161.7	3836.7	183.5
178	2	6	6	12	6	47	1.1	866.5	228.6	3976.0	294.0
178	3	7	7	16	7	42	0.7	484.8	365.2	3216.8	416.9
204	4	8	8	17	8	67	1.0	828.5	235.4	3829.7	209.2
204	3	7	7	15	7	65	1.1	906.8	220.2	3914.4	197.4
230	4	10	10	21	10	70	1.1	1879.9	300.5	4287.6	384.6
230	4	9	9	20	9	65	1.0	856.6	286.5	4040.2	362.0
242	5	11	11	26	11	89	1.7	1425.1	290.5	6444.1	513.0
242	4	10	10	19	10	75	1.2	1920.6	265.3	4069.1	232.8

Table 6: Comparison of efficiently computable bounds on $s_*(A)$, $n = 256$, *Hadamard* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
128	3	5	5	7	5	8	0.2	1148.1	77.8	3007.0	68.5
128	2	5	5	7	5	7	0.3	1297.1	73.4	2894.4	116.8
152	3	7	7	7	7	58	0.3	1224.4	47.9	3997.0	186.8
152	4	7	7	13	7	58	0.2	1205.8	245.0	3962.6	310.4
178	4	9	9	15	9	70	0.2	1269.8	238.9	4828.2	212.0
178	4	9	9	15	9	19	0.3	1340.7	271.1	4923.3	342.8
204	4	12	12	15	12	16	0.5	2908.1	131.2	6409.9	385.4
204	5	12	12	15	12	16	0.4	2996.7	148.9	5507.9	253.9
230	8	18	18	31	19	31	0.3	1860.1	250.8	9046.7	331.1
230	8	18	18	31	18	39	0.4	2100.2	282.8	4081.3	396.8
242	12	26	26	31	27	31	0.3	2015.1	92.7	7478.2	176.2
242	12	26	26	31	26	31	0.3	1976.7	116.8	3597.9	412.0

Table 7: Comparison of efficiently computable bounds on $s_*(A)$, $n = 256$, *Rademacher* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
128	1	5	5	14	5	53	27.8	1253.1	171.6	3388.7	124.8
128	1	5	5	15	5	48	27.8	1361.5	191.1	3291.6	123.4
152	2	6	6	18	7	65	38.4	1426.3	322.7	9592.1	136.3
152	1	6	6	19	7	66	38.3	1183.0	218.9	9146.3	139.0
178	2	7	8	25	9	78	44.2	2819.1	258.9	8032.1	225.8
178	2	7	8	24	9	78	41.8	2481.7	256.0	8306.3	168.2
204	2	10	11	32	12	92	51.1	1434.2	291.8	9738.5	209.3
204	2	10	11	30	12	90	50.8	1316.6	448.3	9146.8	345.4
230	2	14	16	41	19	107	61.8	2422.9	302.7	15235.2	162.2
230	2	14	16	39	19	107	61.7	2466.2	624.0	15578.4	161.9
242	2	20	23	47	27	116	64.8	3929.4	269.2	19828.7	178.1
242	2	19	23	47	27	111	68.0	4242.4	277.8	20506.7	270.5

Table 8: Comparison of efficiently computable bounds on $s_*(A)$, $n = 256$, *Gaussian* matrices

	Unsigned				**Nonnegative**		**CPU time (s)**				
	LBs on $s_0(A)$			UB	LB	UB	Unsigned			Nonnegative	
m	$s[\mu]$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_1]$	$s[\alpha_s]$	$\bar{s}$	$s[\alpha_s]$	$\bar{s}$
128	1	5	5	14	5	44	28.2	852.1	172.4	3283.2	114.7
128	1	4	5	15	5	52	27.7	1913.9	177.7	3712.0	124.6
152	2	6	6	19	7	58	35.4	981.0	214.1	8433.5	392.8
152	1	6	6	19	7	58	38.9	1004.0	242.6	8231.7	373.3
178	2	7	8	24	9	79	43.0	2164.4	393.9	10294.7	368.2
178	2	7	8	25	9	77	47.6	2390.3	263.1	9548.8	374.0
204	2	10	11	32	12	88	58.0	1363.6	293.3	11496.7	274.1
204	2	10	11	32	12	91	51.7	1218.4	293.4	12497.2	529.5
230	2	14	17	41	19	102	70.4	3200.9	339.7	18771.3	431.6
230	2	14	16	39	19	106	61.5	2118.4	485.4	18959.5	435.0
242	2	19	22	46	27	113	73.6	2212.8	277.4	26874.6	269.2
242	2	20	23	47	27	112	65.3	2995.2	426.7	21308.7	191.7

The results in Tables 1-8 merit some comments. We observe that our LP-based efficiently computable lower bounds on $s_0(A)$ and $s_*(A)$ clearly outperform the bounds based on mutual incoherence. We notice that for Fourier and Hadamard matrices, the lower bounds on $s_*(A)$ and $s_0(A)$ are nearly always the same, except for three Hadamard instances with $m = 64$, $m = 230$ and $m = 242$ and one Fourier instance with $m = 120$. On the other hand, for Gaussian and Rademacher matrices, as the number of rows m approaches the number of columns n, the difference between the best certified lower bounds on $s_*(A)$ and on $s_0(A)$ increases (for the sizes we have considered, this difference attains 5 for the Gaussian matrix with $m = 242$). While for Gaussian, Rademacher and Fourier matrices, the upper bounds on $s_*(A)$ become loose (they are twice or three times higher than the upper bounds on $s_0(A)$), these bounds become tighter in the case of Hadamard matrices. Further, for some matrices the lower and the upper bound on $s_0(A)$ match (e.g., the Hadamard matrix with $m = 152$), what allows to identify the exact value of $s_0(A)$. Moreover, we have observed samples of smaller random Hadamard matrices (with $n = 128$ and $m = 120$) for which the lower bounds and upper bounds on both $s_*(A)$ and $s_0(A)$ coincide, which implies $s_*(A) = s_0(A)$ in these cases.

2.8 Matching Pursuit Algorithm

The Matching Pursuit algorithm for signal recovery has been first introduced in [90] and is motivated by the desire to provide a reduced complexity alternative to the ℓ_1-recovery problem. Several implementations of Matching Pursuit has been proposed in the compressed sensing literature (see, e.g., the review [21]). All of them are based on successive Euclidean projections of the signal and the corresponding performance results rely upon the bounds on mutual incoherence $\mu(A)$ of the sensing matrix. We are about to show that the LP-based verifiable sufficient conditions from the previous section can be used to construct a specific version of the Matching Pursuit algorithm which we refer to as *Non-Euclidean Matching Pursuit (NEMP) algorithm*.

Suppose that we have in our disposal $\tau, \tau_\pm \geq 0$ and a matrix $Y = [y_1, ..., y_n]$, such that

$$\begin{array}{lll} (a) & -\tau_- \leq [I - Y^T A]_{ij} \leq \tau_+, & \forall i \in P_+, \ \forall j, \\ (b) & -\tau \leq [I - Y^T A]_{ij} \leq \tau, & \forall i \in P_n, \ \forall j, \\ (c) & \|y_j\|_* \leq \sigma, & \forall j. \end{array} \tag{43}$$

Consider a signal $w \in \mathcal{P}_n$ such that $\|w - w^s\|_1 \leq \mu$, where w^s is the vector obtained from w by replacing all but s largest magnitudes of entries in w with zeros, and let y and δ be such that $\|Aw - y\| \leq \delta$.

Suppose that

$$\rho = s \max\{\tau_+, \tau_-, \tau\} < 1. \tag{44}$$

To simplify notation, we denote $\max[a, b]$ by $a \vee b$. Consider the following iterative procedure:

Algorithm 1

1. *Initialization: Set* $v^{(0)} = 0$, $\alpha_0 = \frac{\|Y^T y\|_{s,1} + s\sigma\delta + \mu}{1 - \rho}$.

2. *Step* k, $k = 1, 2, ...$: *Given* $v^{(k-1)} \in \mathbb{R}^n$ *and* $\alpha_{k-1} \geq 0$, *compute*

   (a) $u = Y^T(y - Av^{(k-1)})$ *and* n *segments*

   $$S_i = \begin{cases} [u_i - \tau_- \alpha_{k-1} - \sigma\delta, \, u_i + \tau_+ \alpha_{k-1} + \sigma\delta], & i \in P_+, \\ [u_i - \tau\alpha_{k-1} - \sigma\delta, \, u_i + \tau\alpha_{k-1} + \sigma\delta], & i \in P_n. \end{cases}$$

   Define $\Delta \in \mathbb{R}^n$ *by setting*

   $$\Delta_i = \begin{cases} [u_i - \tau_- \alpha_{k-1} - \sigma\delta]_+, & i \in P_+, \\ [u_i - \tau\alpha_{k-1} - \sigma\delta]_+, & i \in P_n, \ u_i \geq 0, \\ -[|u_i| - \tau\alpha_{k-1} - \sigma\delta]_+, & i \in P_n, \ u_i < 0 \end{cases}$$

   (here $[a]_+ = \max[0, a]$*).*

   (b) *Set* $v^{(k)} = v^{(k-1)} + \Delta$ *and*

   $$\alpha_k = s[2\tau \vee (\tau_- + \tau_+)]\alpha_{k-1} + 2s\sigma\delta + \mu. \tag{45}$$

   and loop to step $k + 1$.

3. The approximate solution found after k iterations is $v^{(k)}$.

Proposition 2.8.1 *Assume that $w_i \geq 0$ for $i \in P_+$, (44) takes place, and that $\|w - w^s\|_1 \leq \mu$ with a known in advance value of μ. Then the approximate solution $v^{(k)}$ and the value α_k after the k-th step of Algorithm 1 satisfy*

$$(a_k) \quad \textit{for all } i \;\; v_i^{(k)} \in \mathrm{Conv}\{0; w_i\}, \qquad (b_k) \quad \|w - v^{(k)}\|_1 \leq \alpha_k.$$

For proof, see Section 2.10.9.

Let

$$\lambda = s[2\tau \vee (\tau_- + \tau_+)];$$

if $\lambda < 1$, then also $\rho < 1$, so that Proposition 2.8.1 holds true. Furthermore, by (45) the sequence α_k converges exponentially fast to the limit $\alpha_\infty := \frac{2s\sigma\delta+\mu}{1-\lambda}$:

$$\alpha_k = \lambda^k[\alpha_0 - \alpha_\infty] + \alpha_\infty.$$

Note that when $P_+ = \emptyset$, we can set $\tau_- = \tau_+ = 0$ to obtain $\lambda = 2s\tau$; in the case of $P_n = \emptyset$, by setting $\tau = 0$, we have $\lambda = s(\tau_- + \tau_+)$.

The bottom line is: if the optimal value in the convex program

$$\mathrm{Opt} = \min_{\tau, \tau_\pm, Y} \left\{ s[2\tau \vee (\tau_- + \tau_+)] : \begin{array}{rc} -\tau_- \leq [I - Y^T A]_{ij} \leq \tau_+, & \forall i \in P_+, \; \forall j \\ -\tau \leq [I - Y^T A]_{ij} \leq \tau, & \forall i \in P_n, \; \forall j \\ \tau, \tau_\pm \geq 0 & \end{array} \right\}$$

is < 1, the above procedure, as yielded by an optimal solution to the latter problem, possesses the following properties:

1. All approximations $v^{(k)}$, $k = 0, 1, ...$ of w are supported on the support of w;
2. For $i \in P_+$, $v_i^{(k)} \geq 0$ are nondecreasing in k and are $\leq w_i$ for all k;
3. For $i \in P_n$,
   - if $w_i > 0$, then $0 \leq v_i^{(k)} \leq w_i$ and $v_i^{(k)}$ are nondecreasing in k;
   - if $w_i < 0$, then $w_i \leq v_i^{(k)} \leq 0$ and $v_i^{(k)}$ are nonincreasing in k;

4. As k grows, the upper bound α_k on the ℓ_1-error of approximating w by $v^{(k)}$ goes exponentially fast to

$$\alpha_\infty = \frac{2s\sigma\delta + \mu}{1 - \mathrm{Opt}}.$$

Let now $\xi \in [0, 1)$, $\sigma \geq 0$ and $\theta \geq 1$ and suppose that an $m \times n$ matrix A satisfies the following condition:

$\overline{\mathbf{VSG}}_s(\xi, \sigma, \theta)$: *There exists $m \times n$ matrix $Y = [y_1, ..., y_n]$ such that $\|y_i\|_* \leq \sigma$ for all i and*

$$\begin{array}{ll} -\frac{\xi}{(1+\xi)s} \leq [I - Y^T A]_{ij} \leq \frac{\xi}{(1+\xi)s} & \forall i \notin P_+, \ \forall j, \\ -\frac{\xi}{(1+\xi\theta)s} \leq [I - Y^T A]_{ij} \leq \frac{\xi}{(1+\xi\theta)s} & \forall i \in P_+, \ \forall j \notin P_+, \\ -\frac{\xi}{(1+\xi\theta)s} \leq [I - Y^T A]_{ij} \leq \frac{\xi\theta}{(1+\xi\theta)s} & \forall i, j \in P_+. \end{array} \tag{46}$$

Observe that (46) is a system of convex inequalities in Y. Further, $\overline{\mathbf{VSG}}_s(\xi, \sigma, \theta)$ certainly implies $\mathbf{VSG}_s(\xi, \theta, 0, \sigma)$, and is therefore sufficient condition for s-semigoodness of the matrix A.

When $\overline{\mathbf{VSG}}_s(\xi, \sigma, \theta)$ is satisfied with $\xi \in (0, 1)$ and $\theta > 1$, by taking

$$\tau_- = \frac{\xi}{(1 + \xi\theta)s}, \quad \tau_+ = \frac{\xi\theta}{(1 + \xi\theta)s} \quad \text{and} \quad \tau = \frac{\xi}{(1 + \xi)s},$$

we obtain

$$\lambda = \max\left(\frac{\xi + \xi\theta}{1 + \xi\theta}, \frac{2\xi}{1 + \xi}\right) < 1. \tag{47}$$

Combining this condition with Proposition 2.8.1 gives:

Corollary 2.8.1 *Suppose that A satisfies the condition $\overline{\mathbf{VSG}}_s(\xi, \sigma, \theta)$ with certain $\xi \in (0, 1)$, $\sigma \geq 0$ and $\theta \geq 1$. Let $w \in \mathcal{P}_n$ be a vector with $\|w - w^s\|_1 \leq \mu$ where w^s is the vector obtained from w by replacing all but s largest in magnitude entries in w with zeros, and let y be such that $\|Aw - y\| \leq \delta$. Then the approximate solution $v^{(t)}$ found by Algorithm 1 after t iterations satisfies $v_i^{(t)} \geq 0$ for all $i \in P_+$ and*

$$\|w - v^{(t)}\|_1 \leq \frac{2s\sigma\delta + \mu}{1 - \lambda} + \lambda^t \left[\frac{\|Y^T y\|_{s,1} + s\sigma\delta + \mu}{1 - \rho} - \frac{2s\sigma\delta + \mu}{1 - \lambda}\right],$$

where λ is given by (47) and $\rho = \frac{\xi\theta}{1+\xi\theta}$.

It should be noted the NEMP algorithm has several drawbacks as compared with the ℓ_1-recovery. First, the pursuit algorithm requires a priori knowledge of several parameters (σ, Y, τ, τ_-, τ_+, s and μ). Second, the value $(1-\lambda)^{-1}(2s\sigma\delta+\mu)$ is a conservative upper bound on the error of the ℓ_1-recovery, but the error bound in Corollary 2.8.1 is exact. On the other hand, the NEMP algorithm can be an interesting option if the ℓ_1-recovery is to be used repeatedly on the observations obtained with the same sensing matrix A; the numerical complexity of the pursuit algorithm for a given matrix A may only be a fraction of that of the ℓ_1-recovery, especially when used on high-dimensional data.

Our concluding remark is on the condition

$$\frac{\mu(A)}{1+\mu(A)} < \frac{1}{2s}, \tag{48}$$

where $\mu(A)$ is the mutual incoherence of A (see (13)). This condition is usually used in order to establish convergence results for the Matching Pursuit algorithms (see, e.g., [44, 53, 22]). As it is immediately seen, when $\mu(A)$ is well defined (i.e., all columns in A are nonzero), the matrix $Y=[y_1,...,y_n]$ with the columns

$$y_i = \frac{A_i}{(1+\mu(A))A_i^TA_i}$$

satisfies for all $i=1,...,m$ and $j=1,...,n$ the relations

$$|[I-Y^TA]_{ij}| \le \frac{\mu(A)}{1+\mu(A)}.$$

In the case of (48), setting $\theta=1$ and specifying ξ from the relation $\frac{\xi}{1+\xi}=\frac{s\mu(A)}{1+\mu(A)}$, we get $0<\xi<1$ and meet all inequalities in (46). It follows that Y certifies the validity of the condition $\overline{\mathbf{VSG}}_s(\xi,\sigma,1)$ with the outlined ξ and with all $\sigma \ge \max\limits_i \frac{\|A_i\|_*}{(1+\mu(A))\|A_i\|_2^2}$, and thus the above Y can be readily used in Matching Pursuit. Note that in the situation in question Corollary 2.8.1 recovers some results from [22, 44, 53].

2.9 Appendix: Connections to Other Tractable Relaxations for Disjoint Bilinear Programs

Without loss of generality, here we will assume that P is the identity matrix and $p=0$ (an equivalent problem can be defined by redefining the u variables and the set U). We will

examine various relaxations for the following problem:

$$\text{Opt} = \max_{x,u} \left\{ x^T u : x \in X, Ax = 0, u \in U \right\}, \quad X = \text{Conv}\{x^1, ..., x^N\}.$$

Recall that our LP-based relaxation technique of Section 2.4.1.2 leads to the following tractable relaxation

$$\text{Opt} \leq \text{Opt}^+ = \max_V \{\text{Tr}(V) : \ \exists \bar{x}, \bar{u} : [\bar{x}, \bar{u}, V] \in \mathcal{W}, AV = 0, A\bar{x} = 0\}$$

where $\mathcal{W} = \text{Conv}\{[x, u, xu^T] : x \in X, u \in U\}$.

Suppose that the associated polytopes X, U have the following inequality representation $X = \{x : \ Bx \leq b, \ \ell_x \leq x \leq \beta_x\}$ and $U = \{u : \ Cu \leq c, \ \ell_u \leq u \leq \beta_u\}$.

McCormick Relaxation: One of the earliest results on bilinear terms is due to McCormick [91]. In [91] a single bilinear term is considered and the set $\mathcal{B}_{ij} = \{(x_i, u_j, z_{ij}) : \ z_{ij} = x_i u_j, \ell_{x_i} \leq x_i \leq \beta_{x_i}, \ell_{u_j} \leq u_j \leq \beta_{u_j}\}$ is studied. Note that $\mathcal{B}$ is a nonconvex set, using the bounds on the variables, the following set of valid relations can be derived:

$$\begin{aligned}
0 &\leq (\beta_{x_i} - x_i)(\beta_{u_j} - u_j) = \beta_{x_i}\beta_{u_j} - \beta_{u_j}x_i - \beta_{x_i}u_j + x_i u_j \\
0 &\leq (\beta_{x_i} - x_i)(u_j - \ell_{u_j}) = -\beta_{x_i}\ell_{u_j} + \ell_{u_j}x_i + \beta_{x_i}u_j - x_i u_j \\
0 &\leq (x_i - \ell_{x_i})(u_j - \ell_{u_j}) = \ell_{x_i}\ell_{u_j} - \ell_{u_j}x_i - \ell_{x_i}u_j + x_i u_j \\
0 &\leq (x_i - \ell_{x_i})(\beta_{u_j} - u_j) = -\ell_{x_i}\beta_{u_j} + \beta_{u_j}x_i + \ell_{x_i}u_j - x_i u_j
\end{aligned}$$

By replacing the bilinear term $x_i u_j$ with z_{ij} in the above inequalities, the following convex (in fact linear) relaxation of $\mathcal{B}$ is introduced in [91]:

$$\mathcal{B}_{ij}^M = \left\{ (x_i, u_j, z_{ij}) : \begin{array}{r} z_{ij} \geq \beta_{u_j}x_i + \beta_{x_i}u_j - \beta_{x_i}\beta_{u_j}, \\ z_{ij} \leq \ell_{u_j}x_i + \beta_{x_i}u_j - \beta_{x_i}\ell_{u_j}, \\ z_{ij} \geq \ell_{u_j}x_i + \ell_{x_i}u_j - \ell_{x_i}\ell_{u_j}, \\ z_{ij} \leq \beta_{u_j}x_i + \ell_{x_i}u_j - \ell_{x_i}\beta_{u_j}, \\ \ell_{x_i} \leq x_i \leq \beta_{x_i}, \ \ell_{u_j} \leq u_j \leq \beta_{u_j} \end{array} \right\}$$

Later on Al-Khayyal and Falk in [3] showed that $\mathcal{B}_{ij}^M$ defines the convex hull of $\mathcal{B}_{ij}$.

In our framework, a relaxation based on the principles of McCormick can be build by defining a matrix V where the entries V_{ij} will replace the bilinear terms $x_i u_j$ and introducing the inequalities given in $\mathcal{B}_{ij}^M$ to relate the matrix variables V with the vectors x and u in addition to the original constraints stating $x \in X$ and $u \in U$. Therefore the McCormick relaxation for our problem will be:

$$\text{Opt}^M := \max \left\{ \text{Tr}(W) : Ax = 0,\ x \in X,\ u \in U,\ (x_i, u_j, W_{ij}) \in \mathcal{B}_{ij}^M\ \forall i, j \right\}.$$

Sherali-Adams Relaxation: Generalizing the ideas from McCormick relaxation, Sherali-Adams relaxation is built as follows:

Step 1 Generate the nonlinear system:

$$(Ax)(x - \ell_x)^T = 0,\ (Ax)(\beta_x - x)^T = 0$$

$$(Ax)(u - \ell_u)^T = 0,\ (Ax)(\beta_u - u)^T = 0$$

$$0 \leq (b - Bx)(x - \ell_x)^T,\ 0 \leq (b - Bx)(\beta_x - x)^T,$$

$$0 \leq (x - \ell_x)(x - \ell_x)^T,\ 0 \leq (x - \ell_x)(\beta_x - x)^T,$$

$$0 \leq (\beta_x - x)(x - \ell_x)^T,\ 0 \leq (\beta_x - x)(\beta_x - x)^T,$$

$$0 \leq (c - Cu)(u - \ell_u)^T,\ 0 \leq (c - Cu)(\beta_u - u)^T,$$

$$0 \leq (u - \ell_u)(u - \ell_u)^T,\ 0 \leq (u - \ell_u)(\beta_u - u)^T,$$

$$0 \leq (\beta_u - u)(u - \ell_u)^T,\ 0 \leq (\beta_u - u)(\beta_u - u)^T,$$

$$0 \leq (b - Bx)(u - \ell_u)^T,\ 0 \leq (b - Bx)(\beta_u - u)^T,$$

$$0 \leq (c - Cu)(x - \ell_x)^T,\ 0 \leq (c - Cu)(\beta_x - x)^T,$$

$$0 \leq (x - \ell_x)(u - \ell_u)^T,\ 0 \leq (\beta_x - x)(u - \ell_u)^T,$$

$$0 \leq (x - \ell_x)(\beta_u - u)^T,\ 0 \leq (\beta_x - x)(\beta_u - u)^T$$

Step 2 Define symmetric matrices Z, Q and matrix W. Linearize the system by substituting Z_{ij} for $x_i x_j$; Q_{ij} for $u_i u_j$ for all $i = 1, \ldots, n$ and for all $j \geq i$; and W_{ij}

for $x_i u_j$ for all i, j. For ease of reference we will define

$$M_x = \left\{ (x, Z) : \begin{array}{l} -BZ + bx^T + Bx(\ell_x)^T - b(\ell_x)^T \geq 0, \\ BZ - bx^T - Bx(\beta_x)^T + b(\beta_x)^T \geq 0, \\ (x_i, x_j, Z_{ij}) \in \mathcal{B}_{ij}^M \quad \forall i, j \\ Z \text{ is a symmetric matrix} \end{array} \right\}$$

$$M_u = \left\{ (u, Q) : \begin{array}{l} -CQ + cu^T + Cu(\ell_u)^T - c(\ell_u)^T \geq 0, \\ CQ - cu^T - Cu(\beta_u)^T + c(\beta_u)^T \geq 0, \\ (u_i, u_j, Q_{ij}) \in \mathcal{B}_{ij}^M \quad \forall i, j \\ Q \text{ is a symmetric matrix} \end{array} \right\}$$

$$M_w = \left\{ (x, u, W) : \begin{array}{l} -BW + bu^T + Bx(\ell_u)^T - b(\ell_u)^T \geq 0, \\ BW - bu^T - Bx(\beta_u)^T + b(\beta_u)^T \geq 0, \\ -CW^T + cx^T + Cu(\ell_x)^T - c(\ell_x)^T \geq 0, \\ CW^T - cx^T - Cu(\beta_x)^T + c(\beta_x)^T \geq 0, \\ (x_i, u_j, W_{ij}) \in \mathcal{B}_{ij}^M \quad \forall i, j \end{array} \right\}$$

Noting that $AZ = Ax(\ell_x)^T$ and $AZ = Ax(\beta_x)^T$ and $Ax = 0$, we get $AZ = 0$, similarly $AW = Ax(\ell_u)^T = 0$, and therefore the bound from Sherali-Adams relaxation is given by:

$$\text{Opt}^{SA} := \max \left\{ \text{Tr}(W) : \begin{array}{l} Ax = 0, \quad AZ = 0, \quad AW = 0, \quad x \in X, \quad u \in U, \\ (x, Z) \in M_x, \ (u, Q) \in M_u, \ (x, u, W) \in M_w \end{array} \right\}.$$

Note that in the above relaxation whenever $x \in X$ and $Ax = 0$, it is trivial to construct Z such that $(x, Z) \in M_x$ and $AZ = 0$ (just define $Z = xx^T$), therefore the constraints $AZ = 0, \ (x, Z) \in M_x$ are redundant whenever $Ax = 0, \ x \in X$ are enforced. Similarly, the constraint $(u, Q) \in M_u$ is redundant given $u \in U$ is enforced. Therefore Sherali-Adams relaxation can be stated as

$$\text{Opt}^{SA} = \max \left\{ \text{Tr}(W) : \begin{array}{l} Ax = 0, \quad AW = 0, \\ x \in X, \ u \in U, \ (x, u, W) \in M_w \end{array} \right\}.$$

Clearly $\text{Opt}^{SA} \leq \text{Opt}^{M}$, since Sherali-Adams relaxation already contains all of the constraints from the McCormick relaxation. Moreover we can compare the quality of Sherali-Adams relaxation with that of our LP-based bound Opt^{+}.

Lemma 2.9.1 $\text{Opt}^{+} \leq \text{Opt}^{SA}$.

Proof: Let $\{u_1, \ldots, u_L\}$ be the set of extreme points of U. Let's consider an optimal solution to Opt^{+}, say V^*. Since $V^* \in \mathcal{V}$, $AV^* = 0$ and there exists x^*, u^* such that $Ax^* = 0$ and $(x^*, u^*, V^*) \in \mathcal{W}$. From the definition of $\mathcal{W}$, we know that there exists convex combination weights $\alpha^*_{kl} \geq 0$ with $\sum_k \sum_l \alpha^*_{kl} = 1$ such that $x^* = \sum_k \sum_l \alpha^*_{kl} x^k$, $u^* = \sum_k \sum_l \alpha^*_{kl} u^l$ and $V^* = \sum_k \sum_l \alpha^*_{kl} x^k (u^l)^T$. In order to finish the proof, it suffices to show that $(x^*, u^*, V^*) \in M_w$. Let's consider the first constraint in M_w

$$
\begin{aligned}
& -BV^* + b(u^*)^T + B(x^*)(\ell_u)^T - b(\ell_u)^T \\
= \; & -B \sum_k \sum_l \alpha^*_{kl} x^k (u^l)^T + b \sum_k \sum_l \alpha^*_{kl} (u^l)^T + B \sum_k \sum_l \alpha^*_{kl} x^k (\ell_u)^T - b(\ell_u)^T \sum_k \sum_l \alpha^*_{kl} \\
= \; & \sum_k \sum_l \alpha^*_{kl} [-Bx^k (u^l)^T + b(u^l)^T + Bx^k (\ell_u)^T - b(\ell_u)^T] \\
= \; & \sum_k \sum_l \alpha^*_{kl} (b - Bx^k)(u^l - \ell_u)^T \geq 0
\end{aligned}
$$

where the last inequality follows from the fact that $Bx^k \leq b$ due to $x^k \in X$; $u^l \geq \ell_u$ due to $u^l \in U$ and $\alpha^*_{kl} \geq 0$ for all k, l. In a similar fashion, it can be shown that (x^*, u^*, V^*) satisfies the rest of the inequalities except the ones coming from McCormick relaxation in M_w. For a given i, j to see that $(x^*_i, u^*_j, V^*_{ij}) \in \mathcal{B}^M_{ij}$ is also satisfied, consider the first inequality from $\mathcal{B}^M_{ij}$

$$
\begin{aligned}
& V^*_{ij} - \beta_{u_j} x^*_i - \beta_{x_i} u^*_j + \beta_{x_i} \beta_{u_j} \\
= \; & \sum_k \sum_l \alpha^*_{kl} x^k_i u^l_j - \beta_{u_j} \sum_k \sum_l \alpha^*_{kl} x^k_i - \beta_{x_i} \sum_k \sum_l \alpha^*_{kl} u^l_j + \beta_{x_i} \beta_{u_j} \sum_k \sum_l \alpha^*_{kl} \\
= \; & \sum_k \sum_l \alpha^*_{kl} [x^k_i u^l_j - \beta_{u_j} x^k_i - \beta_{x_i} u^l_j + \beta_{x_i} \beta_{u_j}] \geq 0
\end{aligned}
$$

where again the last inequality follows the fact that $x^k \in X$, $u^l \in U$ and $\alpha^*_{kl} \geq 0 \; \forall k, l$. Similarly the other inequalities in $\mathcal{B}^M_{ij}$ are also satisfied by the solution (x^*, u^*, V^*). Hence (x^*, u^*, V^*) is a feasible solution in Sherali-Adams relaxation, proving that $\text{Opt}^{+} \leq \text{Opt}^{SA}$.

□

Note that Opt^+ as compared to Opt^{SA} also has the advantage of being somewhat more efficiently computable especially when X has small number of extreme points. On the other hand, Sherali-Adams relaxation adds many new variables and inequalities and usually one would start out with only McCormick inequalities and try to separate the rest of the violated inequalities in M_w afterwards. Moreover, it is very unlikely to have $\text{Opt}^+ = \text{Opt}^{SA}$ in general since the inequalities in Opt^{SA} only use bound information but the corresponding feasible region in Opt^+ is obtained considering the convex hull of the extreme points.

Lovász-Schrijver Relaxation: A Lovász-Schrijver type relaxation for disjoint bilinear programs can be built in the same way as Sherali-Adams relaxation but in addition to the given constraints in M_x, M_u, M_w, we introduce an additional requirement connecting all of the variables x, u, Z, Q, W by stating that the matrix

$$\begin{bmatrix} 1 & x^T & u^T \\ x & Z & W \\ u & W^T & Q \end{bmatrix}$$

has to be positive semidefinite.

Unfortunately this new requirement of positive semidefiniteness makes it harder to compare our LP-based bound Opt^+ with the bound from Lovasz-Schrijver relaxation. On the other hand, we can simply note that in the unified bound given by Opt^*, we already have the very same requirement. Therefore we can conclude that Opt^* is at least as good as the bound we can obtain from a Lovász-Schrijver type relaxation for our problem.

2.10 Proofs of Chapter 2

2.10.1 Proof of Proposition 2.2.1

(i)⇒(ii): Let A be s-semigood, and let, in contrast to what is stated by (ii), J be a subset of $\{1, ..., n\}$ with $\text{Card}(J) \leq s$ and $x \in \text{Ker}A \backslash \{0\}$ be such that $x_i \leq 0$ for all $i \in P_+ \setminus J$ and

$$\sum_{i \in J \cap P_+} x_i + \sum_{i \in J \cap P_n} |x_i| \geq \sum_{i \notin J} |x_i|.$$

Let $I = (J \cap P_n) \cup \{i \in J \cap P_+ : x_i \geq 0\}$ so that $I \subseteq J$. From the construction of I, we

have $x_i \le 0$ for $i \in J \setminus I$ implying that $x_i \le 0$ for $i \in P_+ \setminus I$. Further,

$$\sum_{i\in I\cap P_+} x_i + \sum_{i\in I\cap P_n} |x_i| = \sum_{i\in J\cap P_+} x_i - \sum_{i\in J\setminus I} x_i + \sum_{i\in J\cap P_n} |x_i|$$
$$\ge \sum_{i\notin J} |x_i| - \sum_{i\in J\setminus I} x_i = \sum_{i\notin J} |x_i| + \sum_{i\in J\setminus I} |x_i| = \sum_{i\notin I} |x_i|.$$

Hence I also violates the condition in (ii). Setting $u_i = x_i$ when $i \in I$ and $u_i = 0$ otherwise and setting $v = u - x$, we have $u_i \ge 0$ for any $i \in I \cap P_+$, $u_i = 0$ for any $i \in P_+ \setminus I$, and $v_i \ge 0$ for $i \in P_+ \setminus I$, $v_i = 0$ for $i \in I \cap P_+$ and $\sum_i |u_i| \ge \sum_i |v_i|$. In addition, $Au = Av$ due to $Ax = 0$, and u is s-sparse; finally, $u \ne v$ due to $x \ne 0$. We see that the s-sparse vector $u \in \mathcal{P}_n$ is not the unique solution to

$$\min_z \left\{ \sum_i |z_i| : \ Az = Au, \quad z_i \ge 0 \ \forall i \in P_+ \right\},$$

which is a desired contradiction.

(ii)⇒(iii): Let A satisfy (ii). Let $\mathcal{J}$ be the family of all subsets J of $\{1, ..., n\}$ of cardinality $\le s$. For $J \in \mathcal{J}$, let

$$X_J = \{x \in \mathrm{Ker} A : \|x\|_1 = 1, \ x_i \le 0 \ \ \forall i \in P_+ \setminus J\}.$$

Assuming that $X_J \ne \emptyset$, let $x \in X_J$. By (ii), we have

$$\sum_{i\in J\cap P_+} x_i + \sum_{i\in J\cap P_n} |x_i| < \sum_{i\notin J} |x_i|.$$

We claim that $\sum_{i\notin J} |x_i| > 0$.

> Indeed, otherwise $x_i \ne 0$ implies that $i \in J$. Let I_+ and I_- be the subsets of J such that $x_i > 0$ for $i \in I_-$ and $x_i < 0$ for $i \in I_+$. At least one of these sets is nonempty due to $x \ne 0$. W.l.o.g. we can assume that $\sum_{i\in I_+} x_i \ge \sum_{i\in I_-} |x_i|$ (otherwise we could replace x with $-x$ and swap I_+ and I_-). Applying (ii) to x and to I_+ in the role of J, we should have
>
> $$\sum_{i\in I_+\cap P_+} x_i + \sum_{i\in I_+\cap P_n} |x_i| = \sum_{i\in I_+} x_i < \sum_{i\notin I_+} |x_i| = \sum_{i\in I_-} |x_i|,$$
>
> which is not the case. This contradiction shows that $\sum_{i\notin J} |x_i| > 0$ whenever $x \in X_J$.

From our claim it follows that the function

$$\frac{\sum_{i\in J\cap P_+} x_i + \sum_{i\in J\cap P_n} |x_i|}{\sum_{i\notin J} |x_i|}$$

is continuous on X_J and is < 1 at every point of this set. Since X_J is compact, we conclude that when $J \in \mathcal{J}$ is such that $X_J \neq \emptyset$, there exists $\xi_J < 1$ such that

$$\sum_{i\in J\cap P_+} x_i + \sum_{i\in J\cap P_n} |x_i| \leq \xi_J \sum_{i\notin J} |x_i| \text{ for any } x \in X_J.$$

Setting $\xi = \max\limits_{J\in\mathcal{J}: X_J\neq\emptyset} \xi_J$, we clearly ensure the validity of (iii). The implication (ii)⇒(iii) is proved.

(iii)⇒(i): Let (iii) take place; let us prove that A is s-semigood. Thus, let u with $u_i \geq 0$ for all $i \in P_+$ be s-sparse; we should prove that u is the unique optimal solution to the problem

$$\min_z \left\{ \sum_i |z_i| : Az = Au,\ z_i \geq 0\ \forall i \in P_+ \right\}.$$

Assume, on the contrary to what should be proved, that the latter problem has an optimal solution v different from u, and let $x = u - v$, so that $x \in \mathrm{Ker}A$ and $x \neq 0$. Setting $I = \{i : u_i \neq 0\}$, we have $\mathrm{Card}(I) \leq s$ and $x_i \leq 0$ when $i \in P_+ \setminus I$, whence by (iii)

$$\sum_{i\in I\cap P_+} x_i + \sum_{i\in I\cap P_n} |x_i| \leq \xi \sum_{i\notin I} |x_i| = \xi \sum_{i\notin I} |v_i|,$$

whence also

$$\underbrace{\sum_{i\in I\cap P_+} u_i + \sum_{i\in I\cap P_n} |u_i|}_{=\sum_{i\in I}|u_i|} \leq \underbrace{\sum_{i\in I\cap P_+} v_i + \sum_{i\in I\cap P_n} |v_i|}_{=\sum_{i\in I}|v_i|} + \xi \sum_{i\notin I} |v_i|. \tag{49}$$

Since $\sum_i |v_i| \leq \sum_i |u_i| = \sum_{i\in I} |u_i|$ due to the origin of v, (49) implies that $\sum_{i\notin I} |v_i| = 0$, that is, both u and v are supported on I, so that x is supported on I as well. Now let $I_+ = \{i \in I \cap P_+ : x_i \geq 0\}$, $I_- = \{i \in I \cap P_+ : x_i < 0\}$ and $I_n = I \cap P_n$. Replacing, if necessary, x with $-x$ and swapping I_+ and I_-, we can assume that $\sum_{i\in I_+} x_i = \sum_{i\in I_+} |x_i| \geq \sum_{i\in I_-} |x_i|$. Applying (iii) to x and to $I_+ \cup I_n$ in the role of J, we get

$$\sum_{i\in I_+} x_i + \sum_{i\in I_n} |x_i| \leq \xi \sum_{i\in I_-} |x_i|,$$

thereby $\sum_{i\in I_+} x_i = \sum_{i\in I_n} |x_i| = \sum_{i\in I_-} |x_i| = 0$ due to $\sum_{i\in I_+} x_i \geq \sum_{i\in I_-} |x_i|$. Thus, $x = 0$, which is a desired contradiction.

We have proved that the properties (i) – (iii) of A are equivalent to each other.

(iii)⇔(iv): The implication (iv)⇒(iii) is evident. Let us prove the inverse implication. Thus, let A satisfy (iii) (and thus – (i) – (ii) as well), and let $\xi' \in (\xi, 1)$. Let, as above, $\mathcal{J}$ be the family of all subsets J of $\{1, ..., n\}$ of cardinality $\leq s$. Let $X = \{x \in \mathrm{Ker}A : \|x\|_1 = 1\}$, and let $J \in \mathcal{J}$. Let $x \in X$. We claim that there exists a neighborhood U_x of x in X and $\theta_{J,x} \in [1, \infty)$ such that for any $u \in U_x$ and $\theta \geq \theta_{J,x}$ it holds

$$\sum_{i\in J\cap P_+} u_i + \sum_{i\in J\cap P_n} |u_i| \leq \xi' \left(\sum_{i\in P_n\setminus J} |u_i| + \sum_{i\in P_+\setminus J} \max[-u_i, \theta u_i] \right). \tag{50}$$

The claim is clearly true when there exists $i \in P_+ \setminus J$ such that $x_i > 0$. Now assume that $x_i \leq 0$ for $i \in P_+ \setminus J$. Then $\sum_{i\notin J} |x_i| > 0$. Indeed, otherwise $x_i = 0$ for all $i \notin J$, which combines with s-semigoodness of A and the relation $Ax = 0$ to imply that $x = 0$ (since assuming $x \neq 0$, we have $x = u - v$ with s-sparse $u \geq 0, v \geq 0$ with non-overlapping supports, and $Au = Av$ due to $Ax = 0$, which of course contradicts the s-semigoodness of A), while x definitely is nonzero (since $\|x\|_1 = 1$ due to $x \in X$). Now, since $x \in \mathrm{Ker}A$ and $x_i \leq 0$, $i \in P_+ \setminus J$, we have

$$\sum_{i\in J\cap P_+} x_i + \sum_{i\in J\cap P_n} |x_i| \leq \xi \sum_{i\notin J} |x_i| < \xi' \sum_{i\notin J} |x_i|$$

where the first inequality is due to (iii), and the second – due to $\sum_{i\notin J} |x_i| > 0$. The concluding strict inequality clearly implies the validity of (50) with $\theta = 1$, provided that U_x is a small enough neighborhood of x. Thus, our claim is true.

From the validity of our claim, extracting from the covering $\{U_x\}_{x\in X}$ of the compact set X a finite subcovering, we conclude that there exists $\theta_J \in [1, \infty)$ such that

$$\forall (x \in X,\ \theta \geq \theta_J) : \sum_{i\in J\cap P_+} x_i + \sum_{i\in J\cap P_n} |x_i| \leq \xi' \left(\sum_{i\in P_n\setminus J} |x_i| + \sum_{i\in P_+\setminus J} \max[-x_i, \theta x_i] \right).$$

Setting $\theta = \max_{J\in\mathcal{J}} \theta_J$, we see that A satisfies $\mathbf{SG}_s(\xi', \theta)$.

(iv)⇒(v): Let A satisfy $\mathbf{SG}_s(\xi,\theta)$ for certain $\xi\in(0,1)$, $\theta\in[1,\infty)$ and let $\|\cdot\|$ be a norm on $\mathbb{R}^m$. Let, further, P be the orthogonal projector of $\mathbb{R}^n$ on $\mathrm{Ker}A$. Then clearly with a properly chosen C one has

$$\|Px-x\|_1\le C\|Ax\|$$

for any $x\in\mathbb{R}^n$. Now let J be a subset of $\{1,...,n\}$ of cardinality $\le s$, $x\in\mathbb{R}^n$ and $u=Px$. We have

$$\begin{aligned}
&\sum_{i\in J\cap P_+}x_i+\sum_{i\in J\cap P_n}|x_i|\le\sum_{i\in J\cap P_+}u_i+\sum_{i\in J\cap P_n}|u_i|+\sum_{i\in J}|u_i-x_i|\\
&\le\ \xi\left[\sum_{i\in P_n\setminus J}|u_i|+\sum_{i\in P_+\setminus J}\max[-u_i,\theta u_i]\right]+\sum_{i\in J}|u_i-x_i|\\
&\le\ \xi\left[\sum_{i\in P_n\setminus J}[|x_i|+|u_i-x_i|]+\sum_{i\in P_+\setminus J}[\max[-x_i,\theta x_i]+\theta|x_i-u_i|]\right]+\sum_{i\in J}|u_i-x_i|\\
&\le\ \xi\left[\sum_{i\in P_n\setminus J}|x_i|+\sum_{i\in P_+\setminus J}\max[-x_i,\theta x_i]\right]+\max[1,\theta\xi]\|x-u\|_1\\
&\le\ \xi\left[\sum_{i\in P_n\setminus J}|x_i|+\sum_{i\in P_+\setminus J}\max[-x_i,\theta x_i]\right]+\max[1,\theta\xi]C\|Ax\|,
\end{aligned}$$

so that A satisfies $\mathbf{SG}_{s,\beta}(\xi,\theta)$ with $\beta=\max(1,\theta\xi)C$. The implication (iv)⇒(v) is proved.

(v)⇒(vi)⇒(iii): These implications are evident. □

2.10.2 Proof of Proposition 2.3.1

Let I be the support of w^s, $\bar I$ be the complement of I in $\{1,...,n\}$, and let $z=w-x$. We denote $I_+=\{i\in I:z_i\ge0\}$, $\bar I_+=\{i\in\bar I:z_i\ge0\}$, and $I_-=I\setminus I_+$, $\bar I_-=\bar I\setminus\bar I_+$. Observe that w is a feasible solution to (25), so that

$$\|x\|_1\le\|w\|_1+\nu. \tag{51}$$

Obviously, $|x_i| - |w_i| \geq -|z_i|$ and $|x_i| - |w_i| \geq |z_i| - 2|w_i|$. Now using $x_i, w_i \geq 0\ \forall i \in P_+$, and $z_i \geq 0\ \forall i \in I_+$, we get

$$\begin{array}{rcl}
\nu & \geq & \sum_i [|x_i| - |w_i|] \qquad [\text{by (51)}] \\
& \geq & \sum_{i \in I_+ \cap P_+} \underbrace{(x_i - w_i)}_{=-z_i} + \sum_{i \in I_- \cap P_+} \underbrace{(x_i - w_i)}_{=-z_i = |z_i|} + \sum_{i \in \bar{I}_- \cap P_+} \underbrace{(x_i - w_i)}_{=-z_i = |z_i|} + \sum_{i \in \bar{I}_+ \cap P_+} \underbrace{(x_i - w_i)}_{=-z_i \geq -w_i} \\
& & + \sum_{i \in P_n} (|x_i| - |w_i|) \\
& \geq & - \sum_{i \in I_+ \cap P_+} z_i + \sum_{i \in I_- \cap P_+} |z_i| + \sum_{i \in \bar{I}_- \cap P_+} |z_i| - \sum_{i \in \bar{I}_+ \cap P_+} w_i \\
& & - \sum_{i \in I \cap P_n} |z_i| + \sum_{i \in \bar{I} \cap P_n} (|z_i| - 2|w_i|),
\end{array}$$

or, equivalently,

$$\begin{array}{l}
\sum_{i \in I_- \cap P_+} |z_i| + \sum_{i \in \bar{I}_- \cap P_+} |z_i| + \sum_{i \in \bar{I} \cap P_n} |z_i| \\
\leq \nu + \sum_{i \in I_+ \cap P_+} z_i + \sum_{i \in I \cap P_n} |z_i| + \sum_{i \in \bar{I}_+ \cap P_+} w_i + 2 \sum_{i \in \bar{I} \cap P_n} |w_i|.
\end{array} \tag{52}$$

On the other hand, we have

$$\|Az\| = \|Aw - Ax\| \leq \|Aw - y\| + \|Ax - y\| \leq \mathrm{e} + \delta. \tag{53}$$

Then by condition $\mathbf{SG}_{s,\beta}(\xi, \theta)$ with $(I_+ \cap P_+) \cup (I \cap P_n)$ in the role of J, we get

$$\begin{array}{l}
\underbrace{\sum_{i \in I_+ \cap P_+} z_i + \sum_{i \in I \cap P_n} |z_i|}_{:=\kappa} \leq \beta \|Az\| + \xi \left[\sum_{i \in \bar{I} \cap P_n} |z_i| + \sum_{i \in (\bar{I} \cap P_+) \cup (I_- \cap P_+)} \psi(z_i) \right] \\
\kappa \leq \beta \|Az\| + \xi \Big[\underbrace{\sum_{i \in \bar{I} \cap P_n} |z_i| + \sum_{i \in I_- \cap P_+} |z_i| + \sum_{i \in \bar{I}_- \cap P_+} |z_i| + \theta \sum_{i \in \bar{I}_+ \cap P_+} z_i}_{:=\tau(\theta)} \Big]
\end{array} \tag{54}$$

Let us derive a bound on $\tau(\theta)$. Now (52) implies, independently of whether $\mathbf{SG}_{s,\beta}(\xi, \theta)$ is or is not true, the first inequality in the following chain:

$$\begin{array}{rcl}
\tau(\theta) & \leq & \nu + \sum_{i \in I_+ \cap P_+} z_i + \sum_{i \in I \cap P_n} |z_i| + \sum_{i \in \bar{I}_+ \cap P_+} w_i + 2 \sum_{i \in \bar{I} \cap P_n} |w_i| + \theta \sum_{i \in \bar{I}_+ \cap P_+} z_i \\
& \leq & \nu + \kappa + (1 + \theta) \sum_{i \in \bar{I}_+ \cap P_+} w_i + 2 \sum_{i \in \bar{I} \cap P_n} |w_i| \qquad [\text{since } w_i \geq z_i \text{ for } i \in P_+] \\
& \leq & \nu + \kappa + (1 + \theta)\mu, \qquad [\text{since } \theta \geq 1 \text{ and } \textstyle\sum_{i \in \bar{I}} |w_i| \leq \mu],
\end{array} \tag{55}$$

and, in particular,

$$\tau(1) = \sum_{i \in I_- \cap P_+} |z_i| + \sum_{i \in \bar{I}} |z_i| \leq \nu + \kappa + 2\mu. \tag{56}$$

Combining (53), (54) and (55), we obtain

$$\kappa \le \beta(\mathrm{e}+\delta) + \xi\left[\nu + \kappa + (1+\theta)\mu\right],$$

and thereby,

$$\kappa = \sum_{i\in I_+\cap P_+} z_i + \sum_{i\in I\cap P_n} |z_i| \le \frac{\beta(\mathrm{e}+\delta) + \xi(\nu + (\theta+1)\mu)}{1-\xi}.$$

Summing up the latter inequality and (56), we obtain

$$\begin{aligned}\|z\|_1 &= \sum_{i\in I\cap P_n} |z_i| + \sum_{i\in I_+\cap P_+} z_i + \left[\sum_{i\in I_-\cap P_+} |z_i| + \sum_{i\in \bar{I}} |z_i|\right] \le \nu + 2\mu + 2\kappa \\ &\le \nu + 2\mu + \frac{2\beta(\mathrm{e}+\delta) + 2\xi(\nu+(\theta+1)\mu)}{1-\xi} = \frac{1+\xi}{1-\xi}\nu + \frac{2(1+\xi\theta)}{1-\xi}\mu + \frac{2\beta}{1-\xi}(\mathrm{e}+\delta),\end{aligned}$$

which is (26).

To show (27) observe that increasing e to $\mathrm{e}' = \mathrm{e} + \alpha\mu$, we can think that the true signal underlying the observation y is w^s rather than w; note that (51) implies that

$$\|x\|_1 \le \|w^s\|_1 + \nu', \ \nu' = \nu + \mu. \tag{57}$$

We can now repeat the reasoning which follows (51), with (57) in the role of (51), w^s in the role of w, e' in the role of e and 0 in the role of μ, thus arriving at the following analogy of the bound (26):

$$\|x - w^s\|_1 \le \frac{1+\xi}{1-\xi}\nu' + \frac{2\beta}{1-\xi}(\mathrm{e}' + \delta),$$

whence

$$\|x - w\|_1 \le \frac{1+\xi}{1-\xi}\nu' + \frac{2\beta}{1-\xi}(\mathrm{e}' + \delta) + \mu,$$

which is nothing but (27). □

2.10.3 Proof of Proposition 2.4.1

Let $\phi(u)$ be the Minkowski function of U, that is, a positively homogeneous, of order 1, function on $\mathbb{R}^n$ such that $U = \{u : \phi(u) \le 1\}$, let $\mathcal{U}$ be the cone $\{(u,t) : \phi(u) \le t\}$. Note that $\mathcal{U}$ is a closed pointed convex cone with a nonempty interior, and its dual cone is

$$\mathcal{U}_* = \{(\omega,\gamma) : \phi_*(-\omega) \le \gamma\}, \ \phi_*(\omega) = \max\left\{\omega^T u : u \in U\right\}.$$

Now,

$$\begin{aligned}
\mathrm{Opt}^+ &= \inf_{Y,v} \max_{1\le i\le N} \left[\max_{u\in U}[(I-Y^TA)x^i]^T[Pu+p] + v^TAx^i\right] \\
&= \inf_{Y,v,\tau}\left\{\tau:\ \max_{u\in U} u^TP^T[I-Y^TA]x^i + [p^T(I-Y^TA)+v^TA]x^i \le \tau,\ \text{ for } 1\le i\le N\right\} \\
&= \inf_{Y,v,\tau}\left\{\tau:\ \phi_*(-P^T[I-Y^TA]x^i) + [p^T(I-Y^TA)+v^TA]x^i \le \tau,\ \text{ for } 1\le i\le N\right\} \\
&= \inf_{Y,v,\tau,\gamma}\left\{\tau:\ \begin{array}{l}\phi_*(-P^T[I-Y^TA]x^i)\le\gamma_i,\ \text{ for } 1\le i\le N\\ \gamma_i + [p^T(I-Y^TA)+v^TA]x^i\le\tau,\ \text{ for } 1\le i\le N\end{array}\right\}
\end{aligned}$$

Since the constraints $\phi_*(-P^T[I-Y^TA]x^i)\le\gamma_i$ imply exactly that $(P^T[I-Y^TA]x^i,\gamma_i)\in\mathcal{U}_*$, Opt^+ is the optimal value of a conic minimization problem. Moreover it is immediately seen that this problem is strictly feasible and bounded, so that the dual problem is solvable with the optimal value Opt^+, which amounts to

$$\begin{aligned}
\mathrm{Opt}^+ &= \max_{w^i,t_i}\left\{\sum_i \mathrm{Tr}(x^i[Pw^i+t_ip]^T):\ \begin{array}{l}A\sum_{i=1}^N t_ix^i = 0\\ A\left[\sum_{i=1}^N x^i[Pw^i+t_ip]\right]=0\\ \phi(w^i)\le t_i,\ \text{ for } 1\le i\le N\\ \sum_{i=1}^N t_i = 1\end{array}\right\} \\
&= \max_{V\in\mathcal{V}}\{\mathrm{Tr}(V)\},
\end{aligned}$$

where

$$\mathcal{V} = \left\{V=\sum_{i=1}^N x^i[Pw^i+t_ip]^T:\ \ \phi(w^i)\le t_i,\ \sum_{i=1}^N t_i = 1,\ A\sum_{i=1}^N t_ix^i = 0,\ AV=0\right\}.$$

Note that $\mathcal{V}$ is a computationally tractable convex compact set. Moreover the set $\mathcal{V}$ admits a simple interpretation. Specifically, setting

$$\mathcal{W} = \mathrm{Conv}\left\{[x,u,x[Pu+p]^T]:\ x\in X,\ u\in U\right\},$$

we have

$$\mathcal{V} = \left\{V:\ \exists\bar{x},\bar{u}:[\bar{x},\bar{u},V]\in\mathcal{W},\ AV=0,\ A\bar{x}=0\right\}.$$

Indeed, if $V\in\mathcal{V}$, that is, $V=\sum_{i=1}^N x^i[Pw^i+t_ip]^T$ with $\phi(w^i)\le t_i$, $\sum_{i=1}^N t_i=1$ and $AV=0$, $A\sum_{i=1}^N t_ix^i=0$, then $w^i=t_iu^i$ with $u^i\in U$, so that, setting $\bar{x}=\sum_{i=1}^N t_ix^i$ and

$\bar{u} = \sum_{i=1}^N t_i u^i$, we have

$$[\bar{x}, \bar{u}, V] = \left[\sum_{i=1}^N t_i x^i, \sum_{i=1}^N t_i u^i, \sum_{i=1}^N t_i x^i [Pu^i + p]^T\right] \in \mathcal{W}.$$

Vice versa, if $[\bar{x}, \bar{u}, V] \in \mathcal{W}$ and $A\bar{x} = 0, AV = 0$, then $[\bar{x}, \bar{u}, V] = \sum_{k=1}^K \lambda_k [\hat{x}^k, \hat{u}^k, \hat{x}^k [P\hat{u}^k + p]^T]$ with $\hat{u}^k \in U$, $\hat{x}^k \in X$ and nonnegative λ_k summing up to 1. Representing $\hat{x}^k = \sum_{i=1}^N \mu_{ki} x^i$ with nonnegative μ_{ki}, $\sum_{i=1}^N \mu_{ki} = 1$, we have

$$\begin{array}{l} [\bar{x}, V] = \sum_{k=1}^K \lambda_k [\hat{x}^k, \hat{x}^k [P\hat{u}^k + p]^T] = \sum_{k=1}^K \sum_{i=1}^N \lambda_k \mu_{ki} [x^i, x^i [P\hat{u}^k + p]^T] \\ = \left[\sum_{i=1}^N t_i x^i, \sum_{i=1}^N x^i [Pw^i + t_i p]^T\right], \end{array}$$

where $w^i = \sum_{k=1}^K \lambda_k \mu_{ki} \hat{u}^k$, $t_i = \sum_{k=1}^K \lambda_k \mu_{ki}$. Clearly $\sum_{i=1}^N t_i = 1$ and since $\phi(\hat{u}^k) \leq 1$ and $\phi(\cdot)$ is a convex function, we have $\phi(w^i) \leq t_i$. Thus, $V = \sum_{i=1}^N x^i [Pw^i + t_i p]^T$ with $\phi(w^i) \leq t_i$ and t_i summing up to 1 and such that $A \sum_{i=1}^N t_i x^i = 0$, that is, $V \in \mathcal{V}$. □

2.10.4 Proof of Proposition 2.4.2

Let A satisfy $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$, and let $Y = [y_1, ..., y_n]$ and v satisfy (32). Let, further, $I \subset \{1, ..., n\}$ be such that $\mathrm{Card}(I) \leq s$, and let $x \in \mathbb{R}^n$. Let $u \in \mathbb{R}^n$ be given by

$$u_i = \begin{cases} 1 + \theta\xi, & i \in P_+ \cap I,\ x_i \geq 0 \\ 1 - \xi, & i \in P_+ \cap I,\ x_i < 0 \\ (1 + \xi)\mathrm{sign}(x_i), & i \in P_n \cap I \\ 0, & i \notin I \end{cases}.$$

Note that u has at most s nonzero entries, the entries of u with indices from P_+ belong to $[0, 1 + \theta\xi]$, and the modulae of entries in u with indices from P_n are $\leq 1 + \xi$, so that

$u^T z \le \Phi_s(z)$ for all z. We have

$$\begin{aligned}
u^T[I - Y^T A]x &= \sum_i u^T C_i[Y,A]x_i = \sum_{i:x_i\ge 0} u^T C_i[Y,A]x_i + \sum_{i:x_i<0} u^T[-C_i[Y,A]]|x_i| \\
&\le \sum_{i:x_i\ge 0} \Phi_s(C_i[Y,A])x_i + \sum_{i:x_i<0} \Phi_s(-C_i[Y,A])|x_i| \qquad [\text{since } u^T z \le \Phi_s(z)] \\
&\le \sum_{i:x_i\ge 0, i\notin P_+} [\xi + (A^T v)_i]x_i + \sum_{i:x_i\ge 0, i\in P_+} [\theta\xi + (A^T v)_i]x_i + \sum_{i:x_i<0} [\xi - (A^T v)_i]|x_i| \quad [\text{by (32)}] \\
&= \xi\left[\sum_{i:x_i\ge 0, i\notin P_+} x_i + \theta \sum_{i:x_i\ge 0, i\in P_+} x_i + \sum_{i:x_i<0} |x_i|\right] + x^T A^T v \\
&= \xi\left[\sum_{i\in P_+} \max[-x_i, \theta x_i] + \sum_{i\in P_n} |x_i|\right] + x^T A^T v,
\end{aligned}$$

whence

$$u^T[I - Y^T A]x \le \xi\left[\sum_{i\in P_+} \max[-x_i, \theta x_i] + \sum_{i\in P_n} |x_i|\right] + \rho\|Ax\| \tag{58}$$

(recall that $\|v\|_* \le \rho$). On the other hand, recalling the definition of u and that $\|y_i\|_* \le \sigma$, we have

$$\begin{aligned}
u^T[I - Y^T A]x &= u^T x - \sum_{i\in I} u_i y_i^T Ax \\
&= \sum_{i\in I\cap P_+} \max[(1-\xi)x_i, (1+\theta\xi)x_i] + (1+\xi)\sum_{i\in I\cap P_n} |x_i| - \sum_{i\in I} u_i y_i^T Ax \\
&\ge \sum_{i\in I\cap P_+} \max[(1-\xi)x_i, (1+\theta\xi)x_i] + (1+\xi)\sum_{i\in I\cap P_n} |x_i| \\
&\quad \underbrace{-\sigma\left[\sum_{i\in I\cap P_+} (1+\theta\xi) + \sum_{i\in I\cap P_n} (1+\xi)\right]}_{\le \beta-\rho} \|Ax\|.
\end{aligned}$$

Combining the resulting inequality with (58), we get

$$\sum_{i\in I\cap P_+} [x_i + \xi\max[-x_i, \theta x_i]] + (1+\xi)\sum_{i\in I\cap P_n} |x_i| \le \beta\|Ax\| + \xi\left[\sum_{i\in P_+} \max[-x_i, \theta x_i] + \sum_{i\in P_n} |x_i|\right]$$

with β given by (33), or, equivalently,

$$\sum_{i\in I\cap P_+} x_i + \sum_{i\in I\cap P_n} |x_i| \le \beta\|Ax\| + \xi\left[\sum_{i\in P_+\setminus I} \max[-x_i, \theta x_i] + \sum_{i\in P_n\setminus I} |x_i|\right].$$

The latter relation holds true for every $x \in \mathbb{R}^n$ and for every set $I \subset \{1, ..., n\}$ of cardinality $\le s$, so that A satisfies $\mathbf{SG}_{s,\beta}(\xi, \theta)$. □

2.10.5 Proof of Proposition 2.5.1

Proof is based on the following

Lemma 2.10.1 *Let Z be a $\nu \times \nu$ matrix of rank m, $s > 1$ be a positive integer, and $\delta_i \in (0,1]$, $1 \le i \le \nu$, be such that for the columns C_i of the matrix $I_\nu - Z$ it holds $\|C_i\|_{s,1} \le 1 - \delta_i$. Assume that*

$$\nu > (2\sqrt{2m} + 1)^2. \tag{59}$$

Then

$$s \le 2\sqrt{2m} + 1. \tag{60}$$

Proof of the lemma. Let $\sigma_i = Z_{ii}$, and let γ_i be the sum of $s-1$ largest magnitudes of the entries in C_i with indices different from i. We have

$$1 - \sigma_i + \gamma_i \le \|C_i\|_{s,1} \le 1 - \delta_i,$$

consequently $\sigma_i \ge \delta_i + \gamma_i > 0$. Let us set $\lambda_i = \frac{1}{\sigma_i}$, and let $\bar{Z}$ be the matrix with the columns $\bar{Z}_i = \lambda_i Z_i$, where Z_i is the i-th column in Z. Note that $\bar{Z}$ is of the same rank m as Z, and that $\bar{Z}_{ii} = 1$ for all i. Recalling that $\gamma_i < \sigma_i$, we have also

$$\|\bar{Z}_i\|_{s-1,1} = \lambda_i \|Z_i\|_{s-1,1} \le \lambda_i[\gamma_i + \sigma_i] \le 2\lambda_i\sigma_i = 2.$$

Now let $\bar{s} = \min[s-1, \lfloor \nu^{1/2} \rfloor]$, so that $\bar{s} \ge 1$ due to $s > 1$. We have $\|\bar{Z}_i\|_{\bar{s},1} \le \|\bar{Z}_i\|_{s-1,1} \le 2$ and $\bar{s}^2 \le \nu$. From the latter inequality and due to $\|\bar{Z}_i\|_2^2 \le \max\{1, \nu\bar{s}^{-2}\}\|\bar{Z}_i\|_{\bar{s},1}^2$ (cf. the proof of [81, Proposition 4.2]), it follows that $\|\bar{Z}_i\|_2^2 \le 4\nu\bar{s}^{-2}$. We conclude that $\|\bar{Z}\|_2^2 \le 4\nu^2\bar{s}^{-2}$, where for a matrix B, $\|B\|_2$ is the Frobenius norm of B. Setting $H = \frac{1}{2}[\bar{Z} + \bar{Z}^T]$, we have therefore $\|H\|_2^2 \le 4\nu^2\bar{s}^{-2}$. On the other hand, $\mathrm{Tr}(H) = \sum_{i=1}^{\nu} \bar{Z}_{ii} = \nu$, while $\mathrm{rank}(H) \le 2m$, whence, denoting by μ_i, $1 \le i \le p \le 2m$, the nonzero eigenvalues of H, we have

$$\|H\|_2^2 = \sum_{i=1}^{p} \mu_i^2 \ge (\sum_{i=1}^{p} \mu_i)^2/p = (\mathrm{Tr}(H))^2/p \ge \nu^2/(2m).$$

We arrive at the inequality $4\nu^2\bar{s}^{-2} \ge \|H\|_2^2 \ge \nu^2/(2m)$, thereby

$$\bar{s}^2 \le 8m. \tag{61}$$

Assuming that $\bar{s} = \lfloor \nu^{1/2} \rfloor$, (61) says that $\nu \le (2\sqrt{2m}+1)^2$, which is impossible. The only other option is that $\bar{s} = s-1$, and we arrive at (60). □

Lemma 2.10.1 ⇒ Proposition 2.5.1: Let Y, v satisfy (32). Consider first the case when $\nu := \mathrm{Card}(P_n) \ge n/2$. Denoting by $\widehat{C}_i$ the ν-dimensional vector comprised of the last ν entries in $C_i = C_i[Y, A]$ (i.e., entries with indices from P_n). By (32), for every $i \in P_n$ and for every set $I \subset P_n$ with $\mathrm{Card}(I) \le s$ we have

$$\textstyle\sum_{j\in I}(1+\xi)|[C_i]_j| \le \Phi_s(-C_i) \le \xi - (A^T v)_i, \quad \sum_{j\in I}(1+\xi)|[C_i]_j| \le \Phi_s(C_i) \le \xi + (A^T v)_i,$$

thus for any $i \in P_n$,

$$2(1+\xi)\|\widehat{C}_i\|_{s,1} \le \Phi_s(-C_i) + \Phi_s(C_i) \le 2\xi,$$

so that $\|\widehat{C}_i\|_{s,1} < 1/2$. We see that the South-Eastern $\nu \times \nu$ submatrix Z of $Y^T A$ satisfies the premise of Lemma 2.10.1, while the size ν of Z satisfies (59) due to (38) and $\nu \ge n/2$. Applying the lemma, we arrive at (39).

Now consider the case when $\mathrm{Card}(P_n) < n/2$, that is, $\nu := \mathrm{Card}(P_+) \ge n/2$. By (32), setting $C_i = C_i[Y, A]$, for every set $I \subset P_+$ with $\mathrm{Card}(I) \le s$ and every $i \in P_+$ we have

$$\begin{aligned}\textstyle\sum_{j\in I}(1+\theta\xi)\max[-[C_i]_j, 0] &\le \Phi_s(-C_i) \le \xi - (A^T v)_i,\\ \textstyle\sum_{j\in I}(1+\theta\xi)\max[[C_i]_j, 0] &\le \Phi_s(C_i) \le \theta\xi + (A^T v)_i,\end{aligned}$$

whence

$$\sum_{j\in I}|[C_i]_j| \le \frac{\xi(1+\theta)}{1+\theta\xi} < 1.$$

Since the latter inequality holds true for every subset I of P_+ with $\mathrm{Card}(I) \le s$, when denoting by $\bar{C}_i$ the part of C_i comprised of the first ν entries (those with indexes from P_+), we have for all $i \in P_+$:

$$\|\bar{C}_i\|_{s,1} < 1.$$

Now the proof can be completed exactly as in the previous case, with the North-Western $\nu \times \nu$ submatrix of $Y^T A$ in the role of Z. □

2.10.6 Proof of Proposition 2.4.3

Let $Y = [Y_1, ..., Y_n], v, \sigma, \rho$ certify the validity of $\mathbf{VSG}^*_{s,\beta}(\xi, \theta)$, and let $\beta' \geq \beta$, $\theta' \geq \theta$ and $\xi' \in [\xi, 1)$. Let us set

$$\lambda = \frac{1+\theta\xi}{1+\theta'\xi'}, \ \mu = \frac{1+\xi}{1+\xi'}.$$

so that $\lambda, \mu \in [0, 1]$, and let Y' be as in the assertion to be proved, that is, the columns of Y' are multiples of those of Y: $Y'_i = \lambda Y_i$ when $i \in P_+$ and $Y'_i = \mu Y_i$ otherwise. All we need to prove is that (Y', v, σ, ρ) certify the validity of $\mathbf{VSG}^*_{s,\beta'}(\xi', \theta')$, and this immediately reduces to verification of the following fact:

Lemma 2.10.2 *Let i, $1 \leq i \leq n$, be fixed, and let $z \in \mathbb{R}^n$ for any $I \subset \{1, ..., n\}$ of cardinality s satisfy the relations*

$$\begin{aligned}
&(a) \quad (1+\theta\xi) \sum_{j \in P_+ \cap I} \max[z_j - \delta_{ij}, 0] + (1+\xi) \sum_{j \in P_n \cap I} |z_j - \delta_{ij}| + (Av)_i \leq \xi, \\
&(b) \quad (1+\theta\xi) \sum_{j \in P_+ \cap I} \max[\delta_{ij} - z_j, 0] + (1+\xi) \sum_{j \in P_n \cap I} |z_j - \delta_{ij}| - (Av)_i \\
&\qquad \leq \eta = \begin{cases} \theta\xi, & i \in P_+, \\ \xi, & i \in P_n, \end{cases}
\end{aligned} \tag{62}$$

where $\delta_{ij} = \begin{cases} 0, & j \neq i, \\ 1, & i = j. \end{cases}$ *Then for every set $I \subset \{1, ..., n\}$ of cardinality s we have*

$$\begin{aligned}
&(a) \quad (1+\theta'\xi') \sum_{j \in P_+ \cap I} \max[\lambda z_j - \delta_{ij}, 0] + (1+\xi') \sum_{j \in P_n \cap I} |\mu z_j - \delta_{ij}| + (Av)_i \leq \xi', \\
&(b) \quad (1+\theta'\xi') \sum_{j \in P_+ \cap I} \max[\delta_{ij} - \lambda z_j, 0] + (1+\xi') \sum_{j \in P_n \cap I} |\mu z_j - \delta_{ij}| - (Av)_i \\
&\qquad \leq \eta_+ = \begin{cases} \theta'\xi', & i \in P_+, \\ \xi', & i \in P_n. \end{cases}
\end{aligned} \tag{63}$$

Proof. Taking into account the definition of λ, μ, in the case of $i \notin I$ the relations (63) are readily given by (62), hence we can assume $i \in I$. Consider two possible cases: $i \in P_+ \cap I$ and $i \in P_n \cap I$.

The case of $i \in P_+ \cap I$. In this case (62) reads:

$$\begin{array}{ll}(a) & (1+\theta\xi)\max[z_i-1,0]+(1+\theta\xi)\sum\limits_{j\in P_+\cap I, j\neq i}\max[z_j,0]\\ & +(1+\xi)\sum\limits_{j\in P_n\cap I}|z_j|+(Av)_i\leq\xi,\\ (b) & (1+\theta\xi)\max[1-z_i,0]+(1+\theta\xi)\sum\limits_{j\in P_+\cap I, j\neq i}\max[-z_j,0]\\ & +(1+\xi)\sum\limits_{j\in P_n\cap I}|z_j|-(Av)_i\leq\theta\xi,\end{array}\tag{64}$$

and our goal is to verify that then

$$\begin{array}{ll}(a) & (1+\theta'\xi')\max[\lambda z_i-1,0]\\ & +\overbrace{(1+\theta'\xi')\lambda}^{=1+\theta\xi}\sum\limits_{j\in P_+\cap I, j\neq i}\max[z_j,0]+\overbrace{(1+\xi')\mu}^{=1+\xi}\sum\limits_{j\in P_n\cap I}|z_j|+(Av)_i\leq\xi',\\ (b) & (1+\theta'\xi')\max[1-\lambda z_i,0]\\ & \underbrace{+(1+\theta\xi)\sum\limits_{j\in P_+\cap I, j\neq i}\max[-z_j,0]+(1+\xi)\sum\limits_{j\in P_n\cap I}|z_j|-(Av)_i}_{:=R}\leq\theta'\xi'.\end{array}\tag{65}$$

We have $\lambda z_i-1\leq\lambda(z_i-1)$ due to $\lambda\leq 1$, consequently

$$\max[\lambda z_i-1,0]\leq\max[\lambda(z_i-1),0]=\lambda\max[z_i-1,0],$$

and therefore (65.a) follows from (64.a) due to $(1+\theta'\xi')\lambda=1+\theta\xi$ and $\xi'\geq\xi$. It remains to verify (65.b). Assume, first, that $\lambda z_i\leq 1$. From (64.b) it follows that

$$(1+\theta\xi)[1-z_i]+R\leq(1+\theta\xi)\max[1-z_i,0]+R\leq\theta\xi,$$

implying $z_i\geq\frac{1+R}{1+\theta\xi}$ and therefore

$$1-\lambda z_i\leq 1-\frac{1+R}{1+\theta'\xi'}=\frac{\theta'\xi'-R}{1+\theta'\xi'}.$$

Since we are in the case $1-\lambda z_i\geq 0$, we arrive at

$$(1+\theta'\xi')\max[1-\lambda z_i,0]+R=(1+\theta'\xi')[1-\lambda z_i]+R\leq(1+\theta'\xi')\frac{\theta'\xi'-R}{1+\theta'\xi'}+R=\theta'\xi',$$

as required in (65.b). The case of $1-\lambda z_i\leq 0$ is trivial, since here the left hand side in (65.b) clearly is $\leq$ the left hand side in (64.b), while $\theta'\xi'\geq\theta\xi$, so that (65.b) is readily given by (64.b). Thus, when $i\in P_+\cap I$, (65) follows from (64).

The case of $i \in P_n \cap I$. In this case (62) means that

$$\begin{array}{ll}(a) & (1+\theta\xi)\sum\limits_{j\in P_+\cap I, j\neq i}\max[z_j,0]+(1+\xi)|1-z_i|+(1+\xi)\sum\limits_{j\in P_n\cap I,j\neq i}|z_j|+(Av)_i\leq\xi,\\ (b) & (1+\theta\xi)\sum\limits_{j\in P_+\cap I}\max[-z_j,0]+(1+\xi)|1-z_i|+(1+\xi)\sum\limits_{j\in P_n\cap I,j\neq i}|z_j|-(Av)_i\leq\xi,\end{array}\tag{66}$$

and our goal is to verify that then

$$\begin{array}{ll}(a) & (1+\theta'\xi')\sum\limits_{j\in P_+\cap I,j\neq i}\max[\lambda z_j,0]\\ & \quad+(1+\xi')|1-\mu z_i|+(1+\xi')\mu\sum\limits_{j\in P_n\cap I,j\neq i}|z_j|+(Av)_i\leq\xi',\\ (b) & (1+\theta'\xi')\sum\limits_{j\in P_+\cap I}\max[-\lambda z_j,0]\\ & \quad+(1+\xi')|1-\mu z_i|+(1+\xi')\sum\limits_{j\in P_n\cap I,j\neq i}|\mu z_j|-(Av)_i\leq\xi'.\end{array}\tag{67}$$

Comparing (66.a) with (67.a), and (66.b) with (67.b), we see that all we need in order to derive (67) from (66) is to verify the following statement: if $(1+\xi)|1-z|\leq\xi+a$, then $(1+\xi')|1-\mu z|\leq\xi'+a$. This is immediate: assuming $(1+\xi)|1-z|\leq\xi+a$, the premises in the following two implication chains hold true:

$$\begin{array}{l}(1+\xi)[1-z]\leq\xi+a\Rightarrow z\geq\frac{1-a}{1+\xi}\Rightarrow\mu z\geq\frac{1-a}{1+\xi'}\Rightarrow 1-\mu z\leq 1-\frac{1-a}{1+\xi'}=\frac{\xi'+a}{1+\xi'}\\ \qquad\Rightarrow(1+\xi')[1-\mu z]\leq\xi'+a,\\ (1+\xi)[z-1]\leq\xi+a\Rightarrow z\leq 1+\frac{\xi+a}{1+\xi}\Rightarrow\mu z\leq\frac{1+2\xi+a}{1+\xi'}\Rightarrow\mu z-1\leq\frac{2\xi-\xi'+a}{1+\xi'}\\ \qquad\Rightarrow(1+\xi')[\mu z-1]\leq 2\xi-\xi'+a\Rightarrow(1+\xi')[\mu z-1]\leq\xi'+a,\end{array}$$

while the resulting inequalities in these chains lead to the desired conclusion $(1+\xi')|1-\mu z|\leq\xi'+a$. □

2.10.7 Proof of Lemma 2.5.1

Let L be the $n\times n$ permutation matrix corresponding to the cyclic shift $e_j\mapsto e_{j_+}$, $j_+=(j+1)\bmod n$, of the standard basic orths $e_0,...,e_{n-1}$ in $\mathbb{R}^n$, and R be the $m\times m$ orthogonal block-diagonal matrix with the North-Western block 1 and d additional 2×2 diagonal blocks $\begin{bmatrix}\cos(2\pi i/n) & -\sin(2\pi i/n)\\ \sin(2\pi i/n) & \cos(2\pi i/n)\end{bmatrix}$, $1\leq i\leq d$. Denoting by A_j the j-th column of A, $0\leq j\leq n-1$, we clearly have $RA_j=A_{j_+}$, hence $A=RAL^{-1}$ and therefore also

$A = R^i A L^{-i}$ for $1 \le i \le n$. Now assume that Y, v satisfy (32) for certain $\xi < 1$, $\theta \ge 1$, ρ, σ. Then

$$\max_i \left[\Phi_s(-C_i[Y,A]) + \Phi_s(C_i[Y,A])\right] \le \xi(1+\theta),$$

in this way, it is immediately seen, $\max_i \|C_i[Y,A]\|_{s,1} \le \kappa := \frac{\xi(1+\theta)}{1+\theta\xi} < 1$, or, which is the same,

$$\Gamma(I - Y^T A) \le \kappa < 1,$$

where $\Gamma(Z)$ is the maximum of the $\|\cdot\|_{s,1}$-norms of columns of $Z \in \mathbb{R}^{n\times n}$. Observe that Γ is a convex function which is symmetric in the sense that $\Gamma(PZP^T) = \Gamma(Z)$ whenever P is a permutation matrix. Now let $\bar{Y} = \frac{1}{n}\sum_{i=1}^n R^{-i} Y L^i$. Since $L^n = I_n$, $R^{-n} = I_m$, we have $R^{-1}\bar{Y}L = \bar{Y}$. We claim that

$$\Gamma(I - \bar{Y}^T A) \le \kappa.$$

Indeed, we have

$$\begin{aligned}
\Gamma(I - \bar{Y}^T A) &= \Gamma(\frac{1}{n}\sum_{i=1}^n [I - L^{-i} Y^T R^i A]) \\
&\le \frac{1}{n}\sum_{i=1}^n \Gamma(I - L^{-i} Y^T R^i A) \qquad \text{[since } \Gamma \text{ is convex]} \\
&= \frac{1}{n}\sum_{i=1}^n \Gamma(L^{-i}\left[I - Y^T [R^i A L^{-i}]\right] L^i) \\
&= \frac{1}{n}\sum_{i=1}^n \Gamma(I - Y^T A) \qquad \text{[since } \Gamma \text{ is symmetric and } R^i A L^{-i} = A] \\
&= \Gamma(I - Y^T A)
\end{aligned}$$

Now let

$$y_j(\phi) = \bar{Y}_{0j} + \sum_{i=1}^d [\bar{Y}_{2i-1,j}\cos(i\phi) + \bar{Y}_{2i,j}\sin(i\phi)].$$

We have $R^{-1}\bar{Y}L = \bar{Y}$, that is, $R^{-1}\bar{Y} = \bar{Y}L^{-1}$. In other words, the columns $\bar{Y}_j$ of $\bar{Y}$ satisfy the relation $\bar{Y}_j = R\bar{Y}_{j_-}$, where $j_- = (j-1) \bmod n$. This is nothing but $y_j(\phi) \equiv y_{j_-}(\phi - \delta)$, $\delta = 2\pi/n$, whence $y_j(\phi) = y_0(\phi - j\delta)$. Observe that the j-th column in $\bar{Y}^T A$ has the entries

$$\bar{Y}_i^T A_j = y_i(j\delta) = y_0((j-i)\delta), \quad 0 \le i \le n-1,$$

meaning that the columns in the matrix $I - \bar{Y}^T A$ are cyclic shifts of each other (so that the $\|\cdot\|_{s,1}$-norms of all columns are the same), and the zero column is comprised of the

values of the trigonometric polynomial $1 - y_0(\phi)$ on the grid $G = \{\phi_j = \frac{2\pi j}{n} : 0 \le j < n\}$. Assuming $s > 1$, when denoting by γ the sum of $s-1$ largest magnitudes of entries in the $(n-1)$-dimensional vector $\{y_0(\phi_i)\}_{i=1}^{n-1}$, we have

$$1 - y_0(0) + \gamma \le \|C_0[\bar{Y}, A]\|_{s,1} \le \kappa < 1,$$

thereby $\mu := y_0(0) > \gamma$. Now let $M = \max_{0 \le \phi \le 2\pi} |y_0(\phi)|$, and let $\bar{\phi} \in \mathrm{Argmax}_\phi |y_0(\phi)|$, so that $y_0'(\bar{\phi}) = 0$. By Bernstein theorem, we have $|y_0''(\phi)| \le d^2 M$ for all ϕ, whence $|y_0(\phi)| \ge M/2$ when $|\phi - \bar{\phi}| \le 1/d$, so that

$$\mathrm{Card}\{j : |y_0(\phi_j)| \ge M/2\} > \frac{n}{\pi d} - 1.$$

It follows that $\gamma \ge \min\left[s-1, \frac{n}{\pi d} - 2\right] M/2$, while $\mu = y_0(0) \le M$. Thus, the relation $\mu > \gamma$ implies that

$$\min[s-1, \frac{n}{\pi d} - 2] < 2,$$

that is, $s \le 2$ provided that $n \ge 4\pi d$. □

2.10.8 Proof of Proposition 2.6.1

We will we consider the more general problem from Section 2.4.1.2 of bounding from above the quantity given in (30):

$$\mathrm{Opt} = \max_{x,u} \left\{x^T[Pu + p] : x \in X, Ax = 0, u \in U\right\}, \quad X = \mathrm{Conv}\{x^1, ..., x^N\},$$

where $x^i \in \mathbb{R}^n$, the set $\{x \in X : Ax = 0\}$ is nonempty, and $U \subset \mathbb{R}^n$ is a computationally tractable compact convex set which contains the origin in its interior. Note that, the only role of ρ and σ in the linear programming based verifiable sufficient condition $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ is to get a control over β; and in the case of $\rho = \sigma = \infty$, $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ becomes equivalent to computing the following upper bound on Opt given by Proposition 2.4.1:

$$\mathrm{Opt}^+ = \max_V \left\{\mathrm{Tr}(V) : \ \exists \bar{x}, \bar{u} : [\bar{x}, \bar{u}, V] \in \mathcal{W}, AV = 0, A\bar{x} = 0\right\},$$

where $\mathcal{W} = \mathrm{Conv}\{[x, u, x(Pu + p)^T] : x \in X, u \in U\}$.

To examine the connection with SDP based condition. Let $\mathcal{X} = \{x \in X : Ax = 0\}$. Given $x \in \mathcal{X}$ and $u \in U$, consider the positive semidefinite matrix

$$\Delta(x,u) := [1;x;[Pu+p]][1;x;[Pu+p]]^T = \left[\begin{array}{c|c|c} 1 & x^T & [Pu+p]^T \\ \hline x & xx^T & x^T[Pu+p]^T \\ \hline [Pu+p] & [Pu+p]x^T & [Pu+p][Pu+p]^T \end{array}\right]$$

The convex hull $\mathcal{K}_*$ of these matrices is contained in every set of the form

$$\mathcal{K} = \left\{ \Delta = \left[\begin{array}{c|c|c} 1 & x^T & [Pu+p]^T \\ \hline x & Z & V \\ \hline [Pu+p] & V^T & Q \end{array}\right] : \Delta \succeq 0, AZ = 0, [x,u,V] \in \mathcal{W}, (*) \right\},$$

where $(*)$ is a set of efficiently computable convex constraints on Δ which are valid for matrices $\Delta(x,u)$ given by $x \in \mathcal{X}$, $u \in U$. When $\Delta \succeq 0$ and $AZ = 0$, we automatically have $Ax = 0, AV = 0$, that is,

$$\mathcal{K}_* \subset \left\{ \Delta = \left[\begin{array}{c|c|c} 1 & x^T & [Pu+p]^T \\ \hline x & Z & V \\ \hline [Pu+p] & V^T & Q \end{array}\right] : \Delta \succeq 0, AZ = 0, V \in \mathcal{V}, (*) \right\}$$

where $\mathcal{V} = \{V : \exists \bar{x}, \bar{u} : [\bar{x}, \bar{u}, V] \in \mathcal{W}, A\bar{x} = 0, AV = 0\}$. Let us denote

$$\mathrm{Opt}^* := \max_{Z,Q,V,x,u} \left\{ \mathrm{Tr}(V) : \Delta = \left[\begin{array}{c|c|c} 1 & x^T & [Pu+p]^T \\ \hline x & Z & V \\ \hline [Pu+p] & V^T & Q \end{array}\right] \succeq 0, AZ = 0, V \in \mathcal{V}, (*) \right\},$$

Then Opt^* is efficiently computable and it follows that $\mathrm{Opt} \le \mathrm{Opt}^* \le \mathrm{Opt}^+$.

In our particular case, for the derivation of verifiable sufficient conditions, $p = 0$ and $P = C^{\xi,\theta}$ which is defined in (36). The extreme points of X are the $2n$ vectors $-e_i$ for $1 \le i \le n$, e_i for $i \in P_n$, and $\theta^{-1}e_i$ for $i \in P_+$, where e_i is the i-th basic orth. Moreover $U = \mathcal{U}$ as defined in (35). Our LP based verifiable sufficient condition $\mathbf{VSG}_s(\xi,\theta,\rho,\sigma)$ with $\rho = \sigma = \infty$ is exactly $\mathrm{Opt}^+ \le \xi$. In addition to this, our SDP bound given in (42) is at least as good as Opt^* without any inequalities included in $(*)$. Note that in our case $\Delta = HGH^T$ where $H = \left[\begin{array}{c|c} I_{n+1} & \\ \hline & C^{\xi,\theta} \end{array}\right]$, under this connection it is clear that the objective functions in

two SDPs are the same. Furthermore, (41.a), i.e., $G \succeq 0$, holds if and only if $\Delta \succeq 0$, with the same transformation, (41.b) and (41.c) correspond to constraints in $(*)$, (41.d) together with (41.f) characterize the set $\mathcal{W}$ and (41.e) is equivalent to $AZ = 0$ in Opt^* .

Hence $\mathrm{Opt}^{\xi,\theta} \leq \mathrm{Opt}^* \leq \mathrm{Opt}^+ \leq \xi$ where the last inequality holds whenever $\mathbf{VSG}_s(\xi, \theta, \rho, \sigma)$ with $\rho = \sigma = \infty$ holds. □

2.10.9 Proof of Proposition 2.8.1

Let us proceed by induction. First, let us show that (a_{k-1}, b_{k-1}) implies (a_k, b_k). Thus, assume that (a_{k-1}, b_{k-1}) holds true. Let $z^{(k-1)} = w - v^{(k-1)}$. By (a_{k-1}), $z^{(k-1)}$ is supported on the support of w and is such that $z_i^{(k-1)} \geq 0$ for $i \in P_+$. Note that

$$\begin{aligned} z^{(k-1)} - u &= w - v^{(k-1)} - Y^T(y - Av^{(k-1)}) = (I - Y^T A)(w - v^{(k-1)}) - Y^T e \\ &= (I - Y^T A)z^{(k-1)} - Y^T e, \end{aligned}$$

where $e = y - Aw$ with $\|Y^T e\|_\infty \leq \sigma\delta$ due to (43.c). Then by (43.a,b) for any $i \in P_+$,

$$-\tau_- \left[\sum_{j\in P_+} z_j^{(k-1)} + \sum_{j\in P_n} |z_j^{(k-1)}|\right] - \sigma\delta \leq z_i^{(k-1)} - u_i \leq \tau_+ \left[\sum_{j\in P_+} z_j^{(k-1)} + \sum_{j\in P_n} |z_j^{(k-1)}|\right] + \sigma\delta,$$

consequently,

$$-\gamma_- := -\tau_-\alpha_{k-1} - \sigma\delta \leq z_i^{(k-1)} - u_i \leq \gamma_+ := \tau_+\alpha_{k-1} + \sigma\delta. \tag{68}$$

We conclude that for any $i \in P_+$ the interval $S_i = [u_i - \gamma_-, u_i + \gamma_+]$ of the width

$$\ell_+ = [\tau_- + \tau_+]\alpha_{k-1} + 2\sigma\delta,$$

covers $z_i^{(k-1)}$. In the same way for any $i \in P_n$

$$-\gamma := -\tau\alpha_{k-1} - \sigma\delta \leq z_i^{(k-1)} - u_i \leq \tau\alpha_{k-1} + \sigma\delta = \gamma,$$

so that the interval $S_i = [u_i - \gamma, u_i + \gamma]$ of the width

$$\ell = 2\tau\alpha_{k-1} + 2\sigma\delta,$$

covers $z_i^{(k-1)}$ when $i \in P_n$.

Recalling that $z_i^{(k-1)} \geq 0$ for $i \in P_+$, the closest to 0 point of S_i is

$$\begin{array}{rclll} \widetilde{\Delta}_i & = & [u_i - \gamma_-]_+ & \text{for } i \in P_+, & \qquad \widetilde{\Delta}_i = [u_i - \gamma]_+ \quad \text{for } i \in P_n, \ \ u_i \geq 0, \\ \widetilde{\Delta}_i & = & -[|u_i| - \gamma]_+ & \text{for } i \in P_n, \ \ u_i < 0, & \end{array}$$

that is, $\widetilde{\Delta}_i = \Delta_i$ for all i. Since the segment S_i covers $z_i^{(k-1)}$ and Δ_i is the closest to 0 point in S_i, while the width of S_i is at most $\ell \vee \ell_+$, we clearly have

$$(a) \quad \Delta_i \in \mathrm{Conv}\left\{0, z_i^{(k-1)}\right\}, \qquad (b) \quad |z_i^{(k-1)} - \Delta_i| \leq \ell \vee \ell_+. \tag{69}$$

Since (a_{k-1}) is valid, (69.a) implies that

$$v_i^{(k)} = v_i^{(k-1)} + \Delta_i \in \left[v_i^{(k-1)} + \mathrm{Conv}\left\{0, w_i - v_i^{(k-1)}\right\}\right] \subseteq \mathrm{Conv}\{0, w_i\},$$

and (a_k) holds. Further, let I be the support of w^s. Relation (a_k) clearly implies that $|z_i^{(k)}| \leq |w_i|$, and we can write due to (69.b):

$$\begin{array}{rcl} \|w - v^{(k)}\|_1 & = & \sum\limits_{i \in I} |w_i - [v_i^{(k-1)} + \Delta_i]| + \sum\limits_{i \notin I} |z_i^{(k)}| \\ & \leq & \sum\limits_{i \in I} |z_i^{(k-1)} - \Delta_i| + \sum\limits_{i \notin I} |w_i| \leq s[\ell \vee \ell_+] + \mu = \alpha_k, \end{array}$$

which is (b_k). The induction step is justified.

It remains to show that (a_0, b_0) holds true. Since (a_0) is evident, all we need is to justify (b_0). Let

$$\alpha_* = \|w\|_1,$$

and let $u = Y^T y$. Same as above (cf. (68)), we have for all i:

$$|w_i - u_i| \leq \max\{\tau_-, \tau_+, \tau\}\alpha_* + \sigma\delta = \frac{\rho}{s}\alpha_* + \sigma\delta.$$

Then

$$\alpha_* = \sum_{i \in I} |w_i| + \sum_{i \notin I} |w_i| \leq \sum_{i \in I} [|u_i| + \frac{\rho}{s}\alpha_* + \sigma\delta] + \mu \leq \|u\|_{s,1} + \rho\alpha_* + s\sigma\delta + \mu.$$

Hence

$$\alpha_* \leq \alpha_0 = \frac{\|u\|_{s,1} + s\sigma\delta + \mu}{1 - \rho},$$

which implies (b_0). □

CHAPTER III

COMPRESSED SENSING SYNTHESIS PROBLEM

3.1 Overview

In this chapter we consider the *synthesis problem* of compressed sensing as follows:

> *Given s and an $M \times n$ matrix A, extract from it an $m \times n$ submatrix A_m, certified to be s-good, with m as small as possible.*

One can think, e.g., of a spatial or planar n-point grid $\mathcal{E}$ of possible locations of signal sources and an M-element grid $\mathcal{S}$ of possible locations of sensors. A sensor in a given location measures a known, depending on the location, linear form of the signals emitted at the nodes of $\mathcal{E}$, and the goal is to place a given number $m \ll M$ of sensors at the nodes of $\mathcal{S}$ in order to be able to recover the location of sources via the ℓ_1-minimization, under the condition that there are at most s sources. Since the exact verification of s-goodness is difficult, we will look for a submatrix of the original matrix A for which the s-goodness can be certified by the sufficient condition (10), introduced in [81]:

$$\exists Y \in \mathbb{R}^{m\times n} \text{ such that } \|I_n - Y^T A\|_\infty < \frac{1}{2s}$$

where $\|M\|_\infty = \max_{i,j} |M_{ij}|$ for a matrix M.

Suppose that along with A we know an $M \times n$ matrix Y_M which certifies that the "level of goodness" of A is at least s, that is, we have

$$\|I_n - Y_M^T A\|_\infty \le \mu < \frac{1}{2s}. \tag{70}$$

Then we can approach the synthesis problem as follows:

> Given $M \times n$ matrices Y_M and A and a tolerance $\epsilon > 0$, we want to extract from A m rows (the smaller is m, the better) to get an $m \times n$ matrix A_m which, along with properly chosen $Y_m \in \mathbb{R}^{m\times n}$, satisfies the relation $\|Y_M^T A - Y_m^T A_m\|_\infty \le \epsilon$.

Choosing $\epsilon < \frac{1}{2s} - \mu$ and invoking (70), we ensure that the output A_m of the above procedure is s-good. This simple observation motivates our interest to the problem of approximating a given matrix by a matrix of specified (low rank) in the *uniform norm*.

Note that in the existing literature on low rank approximation of matrices the emphasis is on efficient construction when the approximation error is measured in the Frobenius norm (for the Frobenius norm $\|A\|_F = \left(\sum_{i,j} A_{ij}^2\right)^{1/2}$). Though the Singular Value Decomposition (SVD) gives the best rank k approximation in terms of all the norms that are invariant under rotation (e.g., the Frobenius norm and the spectral norm), its computational cost may be prohibitive for applications involving large matrices. Recently, the properties of fast low rank approximations in the Frobenius norm based on the randomized sampling of rows (or columns) of the matrix (see, e.g., [50, 61]) or random sampling of a few individual entries (see [1] and references therein) have been studied extensively. Another randomized fast approximation based on the preprocessing by the Fast Fourier Transform or Fast Hadamard Transform has been studied in [107]. Yet we do not know explicit bounds available from the previous literature which concern numerically efficient low rank approximations in the uniform norm.

The only known to us result on low rank approximation of matrices in uniform norm is the one in [120]; it states then if $W = Y^T A \in \mathbb{R}^{m\times n}$ and the rows in Y, A are of Euclidean length at most D, then, for every k, W admits a k-rank approximation $W_k = Y_k^T A_k$ satisfying $\|W - W_k\|_\infty \leq O(1)D^2\sqrt{\ln(mn)/k}$, where Y_k and A_k are $k \times m$ and $k \times n$ matrices with rows that are linear combinations of those in Y, A, respectively. This result does not help in the synthesis problem, where we want the rows of A_k to be just rows of A, and not linear combinations of these rows.

The main result of this chapter is as follows. Let $W = Y^T A$, where Y and A are known $M \times n$ matrices. We consider the approximation $W_k = Y_k^T A_k$ of W such that the matrices Y_k and A_k of dimension $m_k \times n$, $m_k \leq k \leq M$, are composed of multiples of the rows of the matrices Y and A respectively[1]. We show that a fast (essentially, of numerical complexity

[1] Allowing rows of A_k to be *multiples* of rows of A in our context is the same as to require the rows of A_k to be among the rows of A – the corresponding factors can be moved from rows of A_k to those of Y_k.

$O(kMn^2)$) approximation W_k can be constructed which satisfies

$$\|W - W_k\|_\infty = O(1)L(Y,A)\sqrt{\frac{\ln(n)}{k}},$$

where $L(Y,A) = \sum_i \|y_i\|_\infty \|a_i\|_\infty$ and y_i^T, a_i^T denote the i-th rows of Y and A respectively. Note that for moderate values of $L(Y,A) = O(1)$ and $k < n/2$ this approximation is "quasi-optimal", as we know (cf., e.g., [81, Proposition 4.2]) that (for certain matrices W) the accuracy of such an approximation cannot be better than $O(k^{-1/2})$. Moreover, in Section 3.2.4, we show that when W is an $n \times n$ identity matrix, as in the case of compressed sensing synthesis problem, the above bound is unimprovable up to a logarithmic factor. See also Section 3.2.3 for a discussion of how large $L(Y,A)$ can be in the case of A being a Hadamard matrix. We propose two types of construction of fast approximations: we consider the randomized construction, for which the accuracy bounds above hold in expectation (or with significant probability). We also supply "derandomized" versions of the approximation algorithms which do not require random sampling of matrices and attain the same accuracy bounds as the randomized method.

3.2 Low Rank Approximation in Compressed Sensing

In this section, we suppose to be given s and an $M \times n$ matrix A and our objective is to extract from A a submatrix A_k which is composed of, at most, k rows of A, with as small k as possible, which is s-good. We assume that A admits a "goodness certificate" Y. Namely, we are given an $M \times n$ matrix Y such that

$$\mu := \|I_n - Y^T A\|_\infty < \frac{1}{2s}, \tag{71}$$

and we are looking for A_k and the corresponding Y_k such that $\|I_n - Y_k^T A_k\| < \frac{1}{2s}$.

3.2.1 Random Sampling Algorithm

The starting point of our developments is the following simple

Lemma 3.2.1 *Let for $\beta > 0$, let*

$$V_\beta(z) = \beta \ln\left(\sum_{i=1}^d \cosh\left(\frac{z_i}{\beta}\right)\right) - \beta \ln d : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+. \tag{72}$$

Then

(i) *we have* $\|z\|_\infty - \beta\ln(2d) \le V_\beta(z) \le \|z\|_\infty$;

(ii) *if* $\beta_1 \le \beta_2$ *then* $V_{\beta_1}(z) \ge V_{\beta_2}(z)$;

(iii) *function* V_β *is convex and continuously differentiable on* $\mathbb{R}^d$. *Further, its gradient* V'_β *is Lipschitz-continuous with the constant* β^{-1}:

$$\|V'_\beta(z_1) - V'_\beta(z_2)\|_1 \le \beta^{-1}\|z_1 - z_2\|_\infty, \tag{73}$$

and $\|V'_\beta(z)\|_1 \le 1$ *for all* $z \in \mathbb{R}^d$.

For proof, see Section 3.3.1.

Lemma 3.2.1 has the following immediate consequence:

Proposition 3.2.1 *Let* $\beta \ge \beta' > 0$ *(non-random) and let* ξ_1,...,ξ_k *be random vectors in* $\mathbb{R}^d$ *such that* $\mathbf{E}\{\xi_1\} = 0$ *and* $\mathbf{E}\{\xi_i|\xi_1,...,\xi_{i-1}\} = 0$ *a.s. for all* $i \in \{2,\dots,k\}$, *and* $\mathbf{E}\{\|\xi_i\|_\infty^2\} \le \sigma_i^2 < \infty$ *for all* $i \in \{1,\dots,k\}$, *and let* $S_k = \sum_{i=1}^k \xi_i$ *and* $S_0 = 0$. *Then for* $k \ge 1$

$$\mathbf{E}\{V_\beta(S_k)\} \le \mathbf{E}\{V_{\beta'}(S_{k-1})\} + \frac{\sigma_k^2}{2\beta}. \tag{74}$$

As a result,

$$\mathbf{E}\left\{\|S_k\|_\infty\right\} \le \sqrt{2\ln(2d)\sum_{i=1}^k \sigma_i^2}. \tag{75}$$

For proof, see Section 3.3.2.

The random sampling algorithm. Denoting y_i^T and a_i^T, $i = 1,...,M$, i-th rows of Y and A, respectively, let us set

$$\theta_i = \|y_i\|_\infty \|a_i\|_\infty, \quad L = \sum_i \theta_i, \quad \pi_i = \frac{\theta_i}{L}, \quad z_i = \frac{L}{\theta_i} y_i, \tag{76}$$

and let $W = Y^T A$. Observe that

$$\begin{array}{rcl} W & = & \sum_{i=1}^M \pi_i \left(z_i a_i^T\right), \\ \|z_i a_i^T\|_\infty & = & L, \quad 1 \le i \le M, \\ \sum_{i=1}^M \pi_i & = & 1, \quad \pi_i \ge 0, \quad 1 \le i \le M. \end{array} \tag{77}$$

Now let Ξ be random rank 1 matrix taking values $z_i a_i^T$ with probabilities π_i, and let $\Xi_1, \Xi_2, ...$ be a sample of independent realizations of Ξ. Consider the random matrix

$$W_k = \frac{1}{k} \sum_{\ell=1}^{k} \Xi_\ell.$$

Then W_k is, by construction, of the form $Y_k^T A_k$, where A_k is a random $m_k \times n$ submatrix of A with $m_k \leq k$.

As an immediate consequence of Proposition 3.2.1 we obtain the following statement:

Proposition 3.2.2 *One has*

$$\mathbf{E}\{\|W_k - W\|_\infty\} \leq 2Lk^{-1/2}\sqrt{2\ln(2n^2)}. \tag{78}$$

In particular, the probability of the event

$$\mathcal{E} = \{\Xi_1, ..., \Xi_k : \|W_k - W\|_\infty \leq 4Lk^{-1/2}\sqrt{2\ln(2n^2)}\}$$

is $\geq 1/2$*, and whenever this event takes place, we have in our disposal a matrix* Y_k *and a* $m_k \times n$ *submatrix* A_k *of* A *with* $m_k \leq k$ *such that*

$$\|I_n - Y_k^T A_k\|_\infty \leq \|I_n - W\|_\infty + \|W_k - W\|_\infty \leq \mu_k := \mu + 4Lk^{-1/2}\sqrt{2\ln(2n^2)}. \tag{79}$$

For proof, see Section 3.3.3.

Discussion. Proposition 3.2.2 suggests a certain approach to the synthesis problem. Indeed, according to this Proposition, picking at random k rows $a_{i_\ell}^T$, where $i_1, ..., i_k$ are sampled independently from the distribution π, we get with probability at least $1/2$ a random $m_k \times n$ matrix A_k, $m_k \leq k$, which is provably s-good with $s = O(1)(L\sqrt{\ln(n)/k} + \mu)^{-1}$. When $L = O(1)$, this is nearly as good as it could be, since the sufficient condition for s-goodness stated in (10) can justify s-goodness of an $m \times n$ sensing matrix with $n > O(1)m$ only when $s \leq O(1)\sqrt{m}$, see [81, Proposition 4.2].

3.2.2 Derandomization

Looking at the proof of Proposition 3.2.1, we see that the construction of A_k and Y_k can be derandomized. Indeed, (74) implies that

Whenever $S \in \mathbb{R}^{n\times n}$ and $\beta \geq \beta'$ there exists i such that

$$V_\beta(S + (z_i a_i^T - W)) \leq V_{\beta'}(S) + \frac{2L^2}{\beta}.$$

Specifically, the above bound is satisfied for every i such that

$$\langle V'_\beta(S), z_i a_i^T - W\rangle \leq 0,$$

and because $\pi_i \geq 0\ \forall i$ and $\sum_i \pi_i(z_i a_i^T - W) = 0$, the latter inequality is certainly satisfied for some i.

Now assume that given a sequence $\beta_0 \leq \beta_1 \leq ...$ of positive reals, we build a sequence of matrices S_i according to the following rules:

1. $S_0 = 0$;
2. $S_{k+1} = S_k + (v_k a_{\ell_k}^T - W)$ with $\ell_k \in \{1, ..., M\}$ and $v_k \in \mathbb{R}^n$ such that

$$V_{\beta_{k+1}}(S_{k+1}) \leq V_{\beta_k}(S_k) + \delta_{k+1},\ \ \delta_{k+1} \leq \frac{2L^2}{\beta_{k+1}}. \tag{80}$$

Then for every $k \geq 1$ the matrix $U_k = k^{-1}S_k$ is of the form $Y_k^T A_k - W$, where A_k is a $m_k \times n$ submatrix of A with $m_k \leq k$, and

$$\|S_k\|_\infty \leq \beta_k \ln(2n^2) + \sum_{\ell=1}^{k} \delta_\ell,$$

whence

$$\|Y_k^T A_k - I_n\|_\infty \leq \mu + k^{-1}\left(\beta_k \ln(2n^2) + \sum_{\ell=1}^{k} \delta_\ell\right).$$

In particular, for the choice $\beta_\ell = 2L\sqrt{\frac{\ell}{\ln(2n^2)}}$, $\ell = 1, 2, ...$, we obtain[2]

$$\|Y_k^T A_k - I_n\|_\infty \leq \mu + 4L\sqrt{\frac{\ln(2n^2)}{k}}$$

One can consider at least the following three (numerically efficient) policies for choosing v_k and ℓ_k satisfying (80); we order them according to their computational complexity.

[2] for a given k, setting $\beta_\ell = L\sqrt{\frac{2k}{\ln(2n^2)}}$, $1 \leq \ell \leq k$, the right hand side in the bound can be reduced to $\mu + 2L\sqrt{\frac{2\ln(2n^2)}{k}}$.

A. Given S_k, we test one by one the options $\ell_k = i$, $v_k = z_i$, $i = 1, ..., M$, until an option satisfying (80) is met (or test all the n options and choose the one which results in the smallest $V_{\beta_{k+1}}(S_{k+1})$). Note that accomplishing a step of this scheme requires $O(Mn^2)$ elementary operations.

A′. In this version of A, we test the options $\ell_k = i$, $v_k = z_i$ when picking i at random, as independent realizations of the random variable $\imath$ taking values $1, ..., M$ with probabilities π_i, until an option with $\langle V'_{\beta_{k+1}}(S_k), z_i a_i^T - W\rangle \le 0$ is met. Since $\mathbf{E}\left\{\langle V'_{\beta_{k+1}}(S_k), z_i a_i^T - W\rangle\right\} \le 0$, we may hope that this procedure will take essentially less steps than the ordered scan through the entire range $1, ..., M$ of values of i.

B. Given S_k we solve M one-dimensional convex optimization problems

$$t_i^* \in \operatorname*{Argmin}_{t\in\mathbb{R}_+} V_{\beta_{k+1}}(S_k + t z_i a_i^T - W),\ 1 \le i \le M, \tag{81}$$

then select the one, let its index be i_*, with the smallest value of $V_{\beta_{k+1}}(S_k + t_i^* z_i a_i^T - W)$, and put $v_k = t_{i_*}^* z_{i_*}$, $\ell_k = i_*$.

If the bisection algorithm is used to find t_i^*, solving the problem (81) for one i to the relative accuracy ϵ requires $O(n^2 \ln(1/\epsilon))$ elementary operations. The total numerical complexity of the step of the method is $O(Mn^2 \ln(1/\epsilon))$.

C. Given S_k, we solve M convex optimization problems

$$u_i^* \in \operatorname*{Argmin}_{u\in\mathbb{R}^n} V_{\beta_{k+1}}(S_k + u a_i^T - W),\ \ 1 \le i \le M, \tag{82}$$

t hen select the one, let its index be i_*, with the smallest value of $V_{\beta_{k+1}}(S_k + u_i^* a_i^T - W)$, and set $v_k = u_i^*$, $\ell_k = i_*$.

Note that due to the structure of V_β to solve (82) it suffices to find a solution to the system

$$\begin{array}{l} \sum_{\ell=1}^n \gamma_\ell \sinh(\alpha_{j\ell} + \gamma_\ell u_j) = 0, \\ \alpha_{j\ell} = \frac{[S_k]_{j\ell} - [W]_{j\ell}}{\beta_k},\ \ \gamma_\ell = \frac{[A]_{\ell i}}{\beta_k},\ \ 1 \le j, \ell \le n. \end{array} \tag{83}$$

Since the equations of the system (83) are independent, one can use bisection to find the component u_j of the solution.[3] Finding a solution of relative accuracy ϵ to each equation then requires $O(n\ln(1/\epsilon))$ arithmetical operations, and the total complexity of solving (82) becomes $O(Mn^2\ln(1/\epsilon))$.

Selecting Y and W. Note that the numerical schemes of this section should be initialized with matrices Y and $W = Y^TA$. We can do as follows:

1. We start with solving the problem
$$Y \in \mathop{\mathrm{Argmin}}_{Z=[z_1^T;...;z_M^T]\in\mathbb{R}^{M\times n}} \left\{\sum_{i=1}^M \|z_i\|_\infty \|a_i^T\|_\infty : \ \|I_n - Z^TA\|_\infty \le \mu\right\},$$
where μ is a certain fraction of $\frac{1}{2s}$. Assuming the problem is feasible for the chosen μ, we get in this way the "initial point" – the matrix $W = Y^TA$.

2. Then we apply the outlined procedure to find A_k and Y_k. At each step ℓ of this procedure, we get certain $m_\ell \times n$ submatrix A_ℓ of A and a matrix Y_ℓ. When $\|I_n - Y_\ell^TA_\ell\|_\infty$ becomes less than $\frac{1}{2s}$ we terminate. Alternatively, we can solve at each step ℓ an auxiliary problem $\min\limits_{U\in\mathbb{R}^{m_\ell\times n}} \|I_n - U^TA_\ell\|_\infty$ and terminate when the optimal value in this problem becomes less than $\frac{1}{2s}$.

3.2.3 Numerical Illustration

Here we report on preliminary numerical experiments with the synthesis problem as posed in the introduction. In our experiment, A is square, specifically, this is the Hadamard matrix H_{11} of order 2048.

Recall that the Hadamard matrix H_ν, $\nu = 0, 1, ...$ is a square matrix of order 2^ν given by the recurrence
$$H_0 = 1,\ H_{s+1} = \begin{bmatrix} H_s & H_s \\ H_s & -H_s \end{bmatrix},$$
whence H_ν is a symmetric matrix with entries ±1 and $H_\nu^TH_\nu = 2^\nu I_{2^\nu}$.

[3] Note that due to the convexity of the left-hand side of the equation in (83), even faster algorithm of Newton family can be used.

The goal of the experiment was to extract from $A = H_{11}$ an $m \times 2048$ submatrix A_m which satisfies the relation (cf. (10))

$$\mathrm{Opt}(A_m) := \min_{Y_m \in \mathbb{R}^{m \times n}} \|I_n - Y_m^T A_m\|_\infty < \frac{1}{2s}, \ n = 2048 \tag{84}$$

with $s = 10$; under this requirement, we would like to have m as small as possible. In compressed sensing terms, we are trying to solve the synthesis problem with $A = H_{11}$; in low rank approximation terms, we want to approximate I_{2048} in the uniform norm within accuracy < 0.05 by a rank m matrix of the form $Y_m^T A_m$, with the rows of A_m extracted from H_{11}. The advantages of the Hadamard matrix in our context is twofold:

1. The error bound (78) is proportional to the quantity L defined in (76). By the origin of this quantity, we clearly have $\|Y^T A\|_\infty = \|\sum_{i=1}^M y_i a_i^T\|_\infty \le L$, whence $L \ge 1 - \mu > 1 - \frac{1}{2s} \ge 1/2$ by (71). On the other hand, with $A = H_\nu$ being an Hadamard matrix, setting $Y = 2^{-\nu} H_\nu$, so that $Y^T A = I_{2^\nu}$, we ensure the validity of (71) with $\mu = 0$ and get $L = 1$, that is, μ is as small as it could be, and L is nearly as small as it could be.

2. Whenever A_m is a submatrix of H_ν, the optimization problem in the left hand side of (84) is easy to solve.

Item 2 deserves an explanation. Clearly, the optimization program in (84) reduces to the series of $n = 2048$ LP programs

$$\mathrm{Opt}_i(A_m) = \min_{y \in \mathbb{R}^m} \|e_i - A_m^T y\|_\infty, \ 1 \le i \le n, \tag{85}$$

where e_i is the standard basic orth in $\mathbb{R}^n$, and $\mathrm{Opt}(A_m) = \max_i \mathrm{Opt}_i(A_m)$. The point is given by

Proposition 3.2.3 *Suppose A_m is an $m \times n$ submatrix of the $n \times n$ Hadamard or Fourier matrix, then $\mathrm{Opt}_i(A_m)$ is independent of i, i.e., $\mathrm{Opt}(A_m) = \mathrm{Opt}_1(A_m)$.*

For proof of Proposition 3.2.3, see Section 3.3.4.

In the light of Proposition 3.2.3, checking the inequality in (84) requires solving a *single* LP program with m variables rather than solving n LO programs of the same size.

Table 9: Comparison of algorithms for compressed sensing synthesis problem

s	1	2	3	4	5	6	7	8	9	10
B	15	58	121	197	279	343	427	512	584	662
A	12	47	104	172	246	323	399	469	547	617

The experiment was organized as follows. As it was already mentioned, we used $\nu = 11$ (that is, $n = 2048$) and $s = 10$ (that is, the desired uniform norm of approximating I_{2048} by $Y_m^T A_m$ was 0.05). We compared two approximation policies:

- "Blind" approximation – we choose a random permutation $\sigma(\cdot)$ of the indices $1, ..., 2048$ and look at the submatrices A^k, $k = 1, 2, ...$ obtained by extracting from H_{11} rows with indices $\sigma(1), \sigma(2), ..., \sigma(k)$ until a submatrix satisfying (84) is met. This is a refinement of the Random sampling algorithm as applied to $A = H_{11}$ and $Y = 2^{-11}A$, which results in $W = I_{2048}$. The refinement is that instead of looking for approximation of $W = I_{2048}$ of the form $\frac{1}{k}\sum_{\ell=1}^{k} z_{i_\ell} a_{i_\ell}^T$, where $i_1, i_2, ...$ are independent realizations of random variable $\imath$ taking values $1, ..., \mu$ with equal probabilities (as prescribed by (76) in the case of $A = H_\nu$), we look for the best approximation of the form $Y_k^T A^k$, where A^k is the submatrix of A with the row indices $\sigma(1), ..., \sigma(k)$.

- "Active" approximation, which is obtained from algorithm $\mathbf{A}'$ by the same refinement as in the previous item.

In our experiments, we ran every policy 6 times. The results were as follows:

"Blind" policy $\mathcal{B}$: the rank of 0.05-approximation of $W = I_{2048}$ varied from 662 to 680.

"Active" policy $\mathcal{A}$: the rank of 0.05-approximation of W varied from 617 to 630.

Note that in both algorithms the resulting matrix A_m is built "row by row", and the certified levels of goodness of the intermediate matrices $A^1, A^2, ...$ are computed. In Table 9, we indicate, for the most successful (resulting in the smallest m) of the 6 runs of each algorithm, the smallest values of k for which A^k was certified to be s-good, $s = 1, 2, ..., 10$:

Finally, we remark that with A being the Hadamard matrix H_ν, the "no refinement" versions of our policies would terminate according to the criterion $\|I_n - \frac{1}{k}A_k^T A_k\|_\infty < \frac{1}{2s}$,

which, on a closest inspection, is nothing but a slightly spoiled version of the goodness test based on mutual incoherence [43][4]. In the experiments we are reporting, this criterion is essentially weaker that the one based on (84): for the best, over the 6 runs of the algorithms $\mathcal{A}$ and $\mathcal{B}$, 10-good submatrices A_m of H_{11} we got the test based on mutual incoherence certifies the levels of goodness as low as 5 (in the case of $\mathcal{B}$) and 7 (in the case of $\mathcal{A}$).

3.2.4 Lower Bound

We have seen that if $Y^TA = W \in \mathbb{R}^{m\times n}$, then the $\|\cdot\|_\infty$-error of the best in this norm approximation of W by a matrix of rank k by selecting rows from Y and A is *at most* $O(1)L(Y,A)\sqrt{\frac{\ln(n)}{k}}$. We intend to demonstrate that in general this bound is unimprovable, up to a logarithmic in m and n factor even when we are allowed to use any rank k matrix in the approximation. Specifically, the following result holds:

Proposition 3.2.4 *When $n \geq 2k$, the $\|\cdot\|_\infty$ error of any approximation of the unit matrix I_n by a matrix of rank k is at least*

$$\frac{1}{2\sqrt{k}}. \tag{86}$$

Proof [cf. [81, Proposition 4.2]] Let $\alpha(n,k)$ be the minimal $\|\cdot\|_\infty$ error of approximation of I_n by a matrix of rank $\leq k$; this function clearly is nondecreasing in n. Let ν be an integer such that $k < \nu \leq n$, and W be an $\nu\times\nu$ matrix of rank $\leq k$ such that $\|I_\nu - W\|_\infty = \alpha := \alpha(\nu,k)$. By variational characterization of singular values, at least $\nu - k$ singular values of $I_\nu - W$ are ≥ 1, whence $\mathrm{Tr}([I_\nu - W][I_\nu - W]^T) \geq \nu - k$. On the other hand, $\|I_\nu - W\|_\infty \leq \alpha$, whence $\mathrm{Tr}([I_\nu - W][I_\nu - W]^T) \leq \nu^2\alpha^2$. We conclude that $\alpha^2 \geq \frac{\nu-k}{\nu^2}$ for all ν with $k < \nu \leq n$, whence $\alpha^2 \geq \frac{1}{4k}$ when $n \geq 2k$. □

[4]The mutual incoherence test is as follows: given a $k\times n$ matrix $B = [b_1, ..., b_n]$ with nonzero columns, we compute the quantity $\mu(B) = \max\limits_{i\neq j} |b_i^Tb_j|/b_i^Tb_i$ and claim that B is s-good for all s such that $s < \frac{1+\mu(B)}{2\mu(B)}$. With the Hadamard A, the "no refinement" criterion for our scheme is nothing but $s < \frac{1}{2\mu(A^k)}$.

3.3 *Proofs of Chapter 3*

3.3.1 Proof of Lemma 3.2.1

Properties (i) and (ii) are immediate consequences of the definition of V_β given in (72). Observe that V_β is convex and continuously differentiable with

$$\left|\frac{d}{dt}\big|_{t=0} V_\beta(x+th)\right| = \left|\frac{\sum_{i=1}^d \sinh(x_i/\beta)h_i}{\sum_{i=1}^d \cosh(x_i/\beta)}\right| \le \|h\|_\infty \,\forall h,$$

whence $\|V'_\beta(x)\|_1 \le 1$ for $x \in \mathbb{R}^d$. Verification of (73) takes one line: V_β is twice continuously differentiable with

$$\frac{d^2}{dt^2}\big|_{t=0} V_\beta(x+th) = \beta^{-1}\frac{\sum_{i=1}^d \cosh(x_i/\beta)h_i^2}{\sum_{i=1}^d \cosh(x_i/\beta)} - \beta^{-1}\frac{\left(\sum_{i=1}^d \sinh(x_i/\beta)h_i\right)^2}{\left(\sum_{i=1}^d \cosh(x_i/\beta)\right)^2} \le \beta^{-1}\|h\|_\infty^2.$$

□

3.3.2 Proof of Proposition 3.2.1

Let $\beta \ge \beta'$. By applying items (ii) and (iii) of the lemma for $k \ge 1$ we get:

$$\begin{aligned} V_\beta(S_k) &\le V_\beta(S_{k-1}) + \langle V'_\beta(S_{k-1}),\, \xi_k\rangle + \tfrac{1}{2\beta}\|\xi_k\|_\infty^2 \\ &\le V_{\beta'}(S_{k-1}) + \langle V'_\beta(S_{k-1}),\, \xi_k\rangle + \tfrac{1}{2\beta}\|\xi_k\|_\infty^2 \end{aligned}$$

When taking the expectation (first conditional to $\xi_1, ..., \xi_{k-1}$), due to $\mathbf{E}\{\xi_k|\xi_1, ..., \xi_{k-1}\} = 0$ a.s. for $k \ge 2$ and then using $\mathbf{E}\{\langle V'_{\beta'}(S_0), \xi_1\rangle\} = 0$ (due to $\mathbf{E}\{\xi_1\} = 0$), we obtain for $k \ge 1$

$$\mathbf{E}\{V_\beta(S_k)\} \le \mathbf{E}\{V_{\beta'}(S_{k-1})\} + \frac{\mathbf{E}\{\|\xi_k\|_\infty^2\}}{2\beta} \le \mathbf{E}\{V_{\beta'}(S_{k-1})\} + \frac{\sigma_k^2}{2\beta},$$

which is (74). Now let us set $\beta' = \beta = \sqrt{\frac{\sum_{i=1}^k \sigma_i^2}{2\ln(2d)}}$. Since $V_\beta(0) = 0$ we conclude that

$$\mathbf{E}\{V_\beta(S_k)\} \le \sum_{i=1}^k \frac{\sigma_i^2}{2\beta}.$$

On the other hand, by item (i) of Lemma 3.2.1,

$$\mathbf{E}\{\|S_k\|_\infty\} \le \beta\ln(2d) + \mathbf{E}\{V_\beta(S_k)\} \le \beta\ln(2d) + \sum_{i=1}^k \frac{\sigma_i^2}{2\beta} \le \sqrt{2\ln(2d)\sum_{i=1}^k \sigma_i^2}$$

proving (75). □

3.3.3 Proof of Proposition 3.2.2

By (77) we have $\|z_i a_i^T\|_\infty = L$ for all i, and besides this, treating $\imath$ as random index distributed in $\{1, ..., M\}$ according to probability distribution $\pi = \{\pi_i\}_{i=1}^M$, we have $\mathbf{E}\{z_\imath a_\imath^T\} = W$. It follows that $\|\Xi_\ell - W\|_\infty \le 2L$ and $\mathbf{E}\{\Xi_\ell - W\} = 0$. If we denote $S_i = \sum_{\ell=1}^i (\Xi_\ell - W)$, when applying Proposition 3.2.1 we obtain

$$\mathbf{E}\{\|S_k\|_\infty\} \le 2L\sqrt{2k\ln(2n^2)},$$

and we arrive at (78). □

3.3.4 Proof of Proposition 3.2.3

We claim that if A_m is an $m \times 2^\nu$ submatrix of the Hadamard matrix H_ν of order $n = 2^\nu$, then the optimal values in all problems (85) are equal to each other. The explanation is a s follows. Let G be a finite abelian group of cardinality n. Recall that a *character* of G is a complex-valued function $\xi(g)$ such that $\xi(0) = 1$ and $\xi(g+h) = \xi(g)\xi(h)$ for all $g, h \in G$; from this definition it immediately follows that $|\xi(g)| \equiv 1$. The characters of a finite abelian group G form abelian group G_*, the multiplication being the pointwise multiplication of functions, and this group is isomorphic to G. The Fourier Transform matrix associated with G is the $n \times n$ matrix with rows indexed by $\xi \in G_*$, columns indexed by $g \in G$ and entries $\xi(g)$. For example, the usual DFT matrix of order n corresponds to the cyclic group $G = \mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$, while the Hadamard matrix H_ν is nothing but the Fourier Transform matrix associated with $G = [\mathbb{Z}_2]^\nu$ (in this case, all characters take values ± 1). For $g \in G$ let $e_g(h)$ stands for the function on G which is equal to 1 at $h = g$ and is equal to 0 at $h \ne g$. Given an m-element subset Q of G_*, consider the submatrix $A = [\xi(g)]_{\substack{\xi\in Q \\ g\in G}}$ of the Fourier Transform matrix, along with n optimization problems

$$\min_{y\in\mathbb{C}^m} \|\Re[e_g - A^T y]\|_\infty = \min_{y_\xi\in\mathbb{C}} \max_{h\in G} |\Re[e_g(h) - \sum_{\xi\in Q} y_\xi \xi(h)]| \qquad (P_g)$$

These problems clearly have equal optimal values, due to

$$\begin{aligned}
&\max_{h\in G} |\Re[e_g(h) - \textstyle\sum_{\xi\in Q} y_\xi \xi(h)]| = \max_{h\in G} |\Re[e_0(h-g) - \textstyle\sum_{\xi\in Q} [y_\xi \xi(g)]\xi(h-g)]| \\
&= \max_{f=h-g\in G} |\Re[e_0(f) - \textstyle\sum_{\xi\in Q} [y_\xi \xi(g)]\xi(f)|.
\end{aligned}$$

As applied to $G = \mathbb{Z}_2^\nu$, this observation implies that all quantities given by (85) are the same. $\square$

CHAPTER IV

RANDOMIZED ALGORITHMS FOR LARGE-SCALE OPTIMIZATION

4.1 Overview

This chapter is motivated by the desire to develop efficient *randomized* first-order methods for solving well-structured large-scale convex optimization problems. Our primary (but not the only) target is the ℓ_1-minimization problem

$$\mathrm{Opt}_p = \min_u \{\|u\|_1 : \|Au - b\|_p \le \delta\} \quad [A = [A_1, ..., A_n] \in \mathbb{R}^{m\times n}, m, n > 2], \tag{87}$$

where $p = \infty$ ("uniform fit") or $p = 2$ ("ℓ_2-fit"). We are interested in the large-scale case, where the sizes m, n of (possibly dense) matrix A are in the range of thousands/tens of thousands. Efficient solutions to the problems of this type are of paramount importance for sparsity-oriented signal processing, in particular, in compressed sensing (see [27, 25, 45] and references therein). To give an overview of our results, here is what our approach yields for (87):

Proposition 4.1.1 *Assume that (87) is feasible, δ is small enough, namely, $2m^{\frac{1}{p}}\delta \le \|b\|_p$. Given $\epsilon \in (0, \frac{1}{2}\mathrm{Opt}_p\|A\|_{1,p}]$,[1] let our goal be to find an ϵ-solution to (87), that is, a point x_ϵ satisfying*

$$\|x_\epsilon\|_1 \le \mathrm{Opt}_p \ \& \ \|Ax_\epsilon - b\|_p \le \delta + \epsilon.$$

Then, for every tolerance $\chi \in (0, 1/2]$, the outlined goal can be achieved with probability $\ge 1 - \chi$

(i) *in the case of $p = \infty$ (uniform fit) – in at most*

$$O(1)\left[\frac{\sqrt{\ln(m)\ln(n)}\|A\|_{1\to\infty}\mathrm{Opt}_\infty}{\epsilon}\ln\left(\frac{\sqrt{\ln(m)\ln(n)}\|A\|_{1\to\infty}\mathrm{Opt}_\infty}{\chi\epsilon}\right)\right]^2$$

[1] *Here and below $\|A\|_{1\to p} = \max_j \|A_j\|_p$ stands for the norm of the mapping $x \mapsto Ax$ induced by the norms $\|\cdot\|_1$ and $\|\cdot\|_p$ in the argument and the image spaces, respectively*

steps of a randomized algorithm, with computational effort per step reduced to extracting from A two columns and two rows, given their indexes, plus "computational overhead" of $O(1)(m+n)$ operations.

(ii) *in the case of $p=2$ (ℓ_2 fit) – in at most*

$$O(1)\left[\frac{\ln(mn)\kappa(A)\|A\|_{1\to 2}\mathrm{Opt}_2}{\epsilon}\ln\left(\frac{\ln(mn)\kappa(A)\|A\|_{1\to 2}\mathrm{Opt}_2}{\chi\epsilon}\right)\right]^2,\quad \kappa(A)=\frac{\sqrt{m}\|A\|_{1\to\infty}}{\|A\|_{1\to 2}},$$

steps of a randomized algorithm with the same as in (i) *computational effort per step.*

Furthermore, there exists a randomized preprocessing of the data $[A,b]$ of the problem (87) of computational cost not exceeding $O(1)mn\ln(m)$, which ensures with probability $\geq 1-\chi$ that $\kappa(A)\leq O(1)\sqrt{\ln(mn/\chi)}$.

Note that the best known so far complexity of finding ϵ-solution to a large-scale problem (87) by a *deterministic* algorithm is *at least* $O(1)\frac{\sqrt{\ln(m)\ln(n)}\|A\|_{1\to\infty}\mathrm{Opt}_\infty}{\epsilon}$ ($p=\infty$) or $\frac{\sqrt{\ln(n)}\|A\|_{1\to 2}\mathrm{Opt}_2}{\epsilon}$ ($p=2$) steps[2] with complexity of a step dominated by the necessity to perform $O(1)$ multiplications $x\mapsto Ax$, $y\mapsto A^Ty$. When A is dense, the resulting operations count is, up to logarithmic terms, of order of $N_{\mathrm{det}}=\frac{mn}{\nu}$, where $\nu=\frac{\epsilon}{\|A\|_{1\to p}\mathrm{Opt}_p}$ can be naturally interpreted as relative accuracy. For the randomized algorithms underlying Proposition 4.1.1, this count, again, up to logarithmic terms, is of order of $N_{\mathrm{rand}}=\frac{m+n}{\nu^2}$ (uniform fit) and $N_{\mathrm{rand}}=\frac{m+n}{\nu^2}+mn$ (ℓ_2 fit). We see that when $\nu\ll 1$ is fixed and m,n grow, the randomized algorithms eventually outperform the deterministic ones, becoming more significant as the problem size grows. Numerical results presented in Section 4.5 demonstrate that this acceleration is not a purely academic phenomenon and can be of real practical interest.

Our approach is based on saddle point reformulation of well-structured convex minimization problems and is applicable when the resulting saddle point problems are bilinear; in this respect, it goes back to the breakthrough paper of Nesterov [106]. The deterministic saddle point prototypes of the randomized algorithms we develop here were proposed in [100] and [101] and the prototypes of our randomization scheme were proposed in [102, Section 3.3] and [82]. In this chapter, we demonstrate that in the case of a bilinear saddle

[2]The indicated bounds are attainable, provided Opt_p is known in advance.

point problem, a better randomization is possible. The advantage of this new randomization over those prototypes lies in the immediate possibility to assess, in a computationally cheap fashion, the quality of the resulting approximate solutions. This possibility is instrumental when solving *parametric* bilinear saddle point problems. In particular many important applications including the problems of the form (87) reduce to the class of parametric bilinear saddle point problems which we introduce and study in Section 4.2.2. In the hindsight, one can recognize utilizing a particular case of this randomization technique leads to the sublinear time randomized algorithm for solving matrix games due to Grigoriadis and Khachiyan [66].

The main body of this chapter is organized as follows. In Section 4.2, we present a saddle-point-based framework for our developments together with a sample of interesting optimization problems fitting this framework. This sample includes, along with ℓ_1-minimization, the (semidefinite relaxation of the) problem of low-dimensional approximation to a collection of points in $\mathbb{R}^d$ and a specific version of the Support Vector Machine problem. Randomized algorithms for the problems fitting to our framework are developed and analyzed in Sections 4.3 and 4.4. Section 4.5 presents encouraging results of preliminary numerical experiments aimed at comparing the performance of the proposed randomized algorithm and a state-of-the-art deterministic algorithm as applied to large-scale ℓ_1-minimization problem. All proofs are relegated to the last section of this chapter.

4.2 Problems and Goals

We start with specifying and motivating two problems to be discussed in this chapter and our goals.

4.2.1 A Bilinear Saddle Point Problem

4.2.1.1 The problem

The first generic problem we are interested in is a Bilinear Saddle Point (BSP) problem

$$\begin{array}{rcl} \mathrm{SV} & = & \min\limits_{z_1\in Z_1}\max\limits_{z_2\in Z_2}\phi(z_1,z_2), \\ \phi(z_1,z_2) & = & v+\langle a_1,z_1\rangle+\langle a_2,z_2\rangle+\langle z_2,Bz_1\rangle: \quad Z[=Z_1\times Z_2]\to\mathbb{R}, \end{array} \qquad (\mathcal{S})$$

where Z_i are nonempty convex compact sets in Euclidean spaces E_i, $i = 1, 2$. Recall that $(\mathcal{S})$ gives rise to two dual to each other convex optimization programs

$$\begin{array}{rcll} \mathrm{Opt}(P) & = & \min\limits_{z_1 \in Z_1} \overline{\phi}(z_1) := \max\limits_{z_2 \in Z_2} \phi(z_1, z_2) & (P) \\ \mathrm{Opt}(D) & = & \max\limits_{z_2 \in Z_2} \underline{\phi}(z_2) := \min\limits_{z_1 \in Z_1} \phi(z_1, z_2) & (D) \end{array} \tag{88}$$

with $\mathrm{Opt}(P) = \mathrm{Opt}(D) = \mathrm{SV}$, and to the variational inequality : find $z_* \in Z := Z_1 \times Z_2$ such that

$$\langle F(z), z - z_* \rangle \geq 0 \ \text{ for all } z \in Z, \tag{89}$$

where $F : \ Z \mapsto E_1 \times E_2$, is an affine monotone operator given by

$$\begin{array}{rcl} F(z_1, z_2) & = & \left[F_1(z_2) = \dfrac{\partial \phi(z_1, z_2)}{\partial z_1}; F_2(z_1) = -\dfrac{\partial \phi(z_1, z_2)}{\partial z_2} \right] = a + \mathcal{A}[z_1; z_2], \\ & & a = [a_1; -a_2], \ \ \mathcal{A} = \left[\begin{array}{c|c} & B^* \\ \hline -B & \end{array} \right], \end{array}$$

(here B^* stands for the conjugate of B). Note that $\mathcal{A}$ is skew-symmetric: $\mathcal{A}^* = -\mathcal{A}$ and

$$\langle z, \mathcal{A} z \rangle = 0 \ \forall z \in E := E_1 \times E_2. \tag{90}$$

It is well known that the solutions to $(\mathcal{S})$ — the saddle points of ϕ on $Z_1 \times Z_2$ — are exactly the pairs $z = [z_1; z_2]$ comprised of optimal solutions to problems (P) and (D) in (88), same as are exactly the solutions to the variational inequality (89). We quantify the accuracy of candidate solutions $z = [z_1; z_2] \in Z$ to $(\mathcal{S})$ by the *saddle point residual*

$$\epsilon_{\mathrm{sad}}(z) = \overline{\phi}(z_1) - \underline{\phi}(z_2) = \underbrace{\left[\overline{\phi}(z_1) - \mathrm{Opt}(P) \right]}_{\geq 0} + \underbrace{\left[\mathrm{Opt}(D) - \underline{\phi}(z_2) \right]}_{\geq 0}. \tag{91}$$

4.2.1.2 Assumptions and goal

When speaking about a BSP problem $(\mathcal{S})$, our goal is to solve the problem within a given accuracy $\epsilon > 0$, that is, to find $z^\epsilon \in Z$ such that $\epsilon_{\mathrm{sad}}(z^\epsilon) \leq \epsilon$. Deterministic first order algorithms achieve this goal by working with the values of the associated operator F at the iterates z_t, $t = 1, 2, ...$, generated by the method. When Z is simple and the problem is large-scale, computing the values $F(z_t)$ is the "leading term" in the computational effort.

Our goal in this chapter is to replace relatively expensive (in the large-scale case) exact values $F(z_t)$ with their computationally cheap unbiased random estimates. Specifically, we assume that

[P] every point $z \in Z$ is associated with a probability distribution P_z such that

- P_z is supported on Z and $\mathbf{E}_{\zeta \sim P_z}\{\zeta\} = z$;
- Given z, we can sample from the distribution P_z.

Under these assumptions, in order to get an unbiased estimate of $F(z_t)$, it suffices to draw a $\zeta_t \sim P_{z_t}$ and to take $F(\zeta_t)$ as a desired estimate of $F(z_t)$. In order to make this approach meaningful, the computational price of generating ζ_t and subsequent computation of $F(\zeta_t)$ should be significantly less than the price of a straightforward computation of $F(z_t)$. This requirement guided us in the selection of applications to be considered below as well as in building the corresponding saddle point reformulations .

Note that the deterministic algorithms remain in the scope of our approach since we always have an option to define P_z as δ_z (the unit mass sitting at z).

4.2.1.3 Application example: low dimensional approximation

We consider the following problem (related to a dimension reduction problem in statistics, see, e.g., [37]): let $V = \{v_1, ..., v_N\}$ be a collection of unit vectors in $\mathbb{R}^n$, and $d < n$ be a positive integer. We want to find a linear subspace $E \subset \mathbb{R}^n$ of dimension d such that the *deviation* $\delta(V, E)$ of the collection from E — the maximal, over i, Euclidean distance between v_i and E — is as small as possible.

Letting Π^d be the family of all orthogonal projectors of $\mathbb{R}^n$ onto d-dimensional linear subspaces, the problem reads

$$\mathrm{Opt}_* = \max_{\Pi \in \Pi^d} \min_{1 \le i \le N} v_i^T \Pi v_i$$

and seems to be computationally intractable. It, however, admits the tractable relaxation

$$\mathrm{Opt} = \max_{Q \in \mathcal{P}^d} \min_{1 \le i \le N} v_i^T Q v_i, \quad \mathcal{P}^d = \{Q \in \mathbf{S}^n : \ 0 \preceq Q \preceq I, \ \mathrm{Tr}(Q) = d\}. \tag{92}$$

We refer to (92) as to the *problem of low dimensional approximation*. We clearly have $\mathrm{Opt}_* \le \mathrm{Opt} \le 1$, whence $\delta^2 := 1 - \mathrm{Opt} \le \delta_*^2 := 1 - \mathrm{Opt}_*$; note that δ_* is the deviation of V from the "ideal" d-dimensional space E_* underlying Opt_*. It is easily seen (see Lemma 4.6.1 of Section 4.6.1) that if Q_* is an optimal solution to the relaxation (92) and E is spanned by the d leading eigenvectors of Q_*, then $\delta(V, E) \le \sqrt{d+1}\delta_*$, that is, approximation (92) admits some quality guarantees.

Now, (92) is nothing but the BSP problem:

$$1 - \mathrm{Opt} = \min_{Q \in \mathcal{P}^d} \max_{\lambda \in \Delta_N} \left[1 - \mathrm{Tr}\left(Q \sum\nolimits_{i=1}^N \lambda_i v_i v_i^T\right)\right], \quad \Delta_N = \left\{\lambda \in \mathbb{R}^N_+ : \sum\nolimits_i \lambda_i = 1\right\}. \tag{93}$$

In terms of $(\mathcal{S})$, E_1 is the space $\mathbf{S}^n$ of symmetric $n \times n$ matrices with Frobenius inner product, $Z_1 = \mathcal{P}^d \subset E_1$, $E_2 = \mathbb{R}^N$, $Z_2 = \Delta_N$. The associated operator F is

$$F(z_1, z_2) = F(Q, \lambda) = \big[\underbrace{-\sum\nolimits_{i=1}^N \lambda_i v_i v_i^T}_{F_1(z_2)}; \underbrace{[v_1^T Q v_1; \ ...; \ v_N^T Q v_N]}_{F_2(z_1)}\big]. \tag{94}$$

Assuming that v_i are dense, the arithmetic cost of computing the value of F at a given point is $O(n^2 N)$. Now let us specify the distributions P_z, $z = (Q, \lambda) \in Z = Z_1 \times Z_2$. In order to generate $\zeta \sim P_{(Q,\lambda)}$, we proceed as follows:

- Given $Q \in \mathcal{P}^d$, we build the eigenvalue decomposition $Q = U \mathrm{Diag}\{q\} U^T$. Note that $q \in \Delta_{n,d} = \{q \in \mathbb{R}^n : 0 \le q_i \le 1 \, \forall i, \sum_{i=1}^n q_i = d\}$. The extreme points of $\Delta_{n,d}$ are Boolean vectors with exactly d nonzero entries. There exists a simple algorithm (see Section 4.6.1) which, given as input a vector $q \in \Delta_{n,d}$, builds in $O(1)dn^2$ a.o. n extreme points q^j, $1 \le j \le n$, of $\Delta_{n,d}$ along with weights $\mu_j \ge 0$, $\sum_j \mu_j = 1$, such that $q = \sum_j \mu_j q^j$. We run this algorithm to build $\{q^j, \mu_j\}_{j=1}^n$, pick $\jmath \in \{1, ..., n\}$ at random, with $\mathrm{Prob}\{\jmath = j\} = \mu_j$, $j = 1, ..., n$, and set $\zeta_1^\jmath = U \mathrm{Diag}\{q^\jmath\} U^T$.
- Given $\lambda \in \Delta_N$, we pick $\imath \in \{1, ..., N\}$ at random, with $\mathrm{Prob}\{\imath = i\} = \lambda_i$, $1 \le i \le N$, and set $\zeta_2^\imath := e_\imath$, where e_i, $i = 1, ..., N$, are standard basic orths in $\mathbb{R}^N$.
- Finally, we set $\zeta = \zeta^{\imath\jmath} := [\zeta_1^\jmath; \zeta_2^\imath] \in \mathcal{P}^d \times \Delta_N$.

The family of distributions $P_{(Q,\lambda)}$ clearly satisfies [P]. The "setup costs" for sampling from $P_{(Q,\lambda)}$ reduce to those of 1) computing the eigenvalue decomposition of Q, 2) building

$q^1, ..., q^n$, $\mu_1, ..., \mu_n$ (this cost is $O(n^3 + dn^2)$ a.o.) and 3) computing the "cumulative distributions" $\{\mu^j = \sum_{s=1}^{j} \mu_s\}_{j=1}^{n}$ and $\{\lambda^i = \sum_{s=1}^{i} \lambda_s\}_{i=1}^{N}$ (what amounts to $O(n+N)$ a.o.). After the setup cost is paid, a sample $(\imath, \jmath)$ can be generated at the cost of just $O(\ln(n+N))$ a.o. Now let us look at the cost of computing $F(\zeta^{\imath\jmath})$ given $\imath, \jmath$. We have

$$F(\zeta^{\imath\jmath}) = \left[-v_\imath v_\imath^T; \{v_i^T U \mathrm{Diag}\{q^\jmath\} U^T v_i\}_{i=1}^{N}\right].$$

Since $q^\jmath$ has just d nonzero entries, all equal to 1, let the indices of the entries be $j_1, ..., j_d$, we have $v_i^T U \mathrm{Diag}\{q^\jmath\} U^T v_i = \sum_{\ell=1}^{d} (U_{j_\ell}^T v_i)^2$, where U_j is j^{th} column of U. We see that computing $F(\zeta^{\imath\jmath})$ costs $O(n^2 + dnN)$ a.o. Thus, the total cost (including that of the setup) of drawing a sample ζ from $P_{(Q,\lambda)}$ and computing $F(\zeta)$ is

$$O(n^3 + dn^2 + n^2 + dnN) = O(n^3 + dnN) \text{ a.o.}$$

When $d \ll n \ll N$, this cost is much smaller than the cost $O(n^2 N)$ of computing $F(z)$ at a "general position" point $z = (Q, \lambda) \in Z$.

4.2.2 A Generalized Bilinear Saddle Point Problem

4.2.2.1 The problem

Assume that we are given a *single-parameter family* of bilinear saddle point problems

$$\mathrm{SV}(\rho) = \min_{z_1 \in Z_1} \max_{z_2 \in Z_2} \phi^\rho(z_1, z_2) := \phi(z_1, z_2) + \rho\psi(z_1, z_2), \tag{95}$$

where $\rho \geq 0$ is a parameter and $\phi(z_1, z_2)$, $\psi(z_1, z_2)$ are bi-affine in z_1 and z_2. The *Generalized Bilinear Saddle Point* (GBSP) problem associated with this family is, by definition, the optimization program

$$\rho_* = \max\{\rho \geq 0 : \mathrm{SV}(\rho) \leq 0\} \tag{96}$$

A highly desirable property of a GBSP problem, relative to our approach, is the convexity of $\mathrm{SV}(\rho)$ as a function of $\rho \geq 0$. To ensure this property, we make from now on the following assumption on the structure of (95):

[A.1] *$Z_1 = Z_{11} \times Z_{12}$ is the direct product of two convex compact sets, and the bilinear functions $\phi(z_1, z_2)$, $\psi(z_1, z_2)$ in (95) are of the form*

$$\begin{array}{rcl} \phi(z_1 = [z_{11}; z_{12}], z_2) & = & \upsilon + \langle a_{11}, z_{11}\rangle + \langle b, z_2\rangle + \langle z_2, Bz_{11}\rangle, \\ \psi(z_1 = [z_{11}; z_{12}], z_2) & = & \chi + \langle a_{12}, z_{12}\rangle + \langle c, z_2\rangle + \langle z_2, Cz_{12}\rangle, \end{array} \tag{97}$$

that is, $\phi(z_1, z_2)$ and $\psi(z_1, z_2)$ as functions of z_1 depend each on its own "block" of z_1, and these blocks z_{11} and z_{12}, independently of each other, run through the respective convex compact sets Z_{11} and Z_{12}.

From now on, we denote by $F^\rho(z) = \Phi(z) + \rho\Psi(z)$ the affine monotone operator associated with ϕ^ρ according to (89).

Lemma 4.2.1 *In the case of* **A.1** *the function* $\mathrm{SV}(\rho)$ *given by* (95) *is convex in* $\rho \geq 0$.

From now on we assume, in addition to **A.1**, that

[A.2] *Function* $\mathrm{SV}(\rho)$ *given by* (95) *is nonpositive somewhere on* $\mathbb{R}_{++}$ *and tends to* $+\infty$ *as* $\rho \to +\infty$,

which implies solvability of (96) and positivity of ρ_*.

The goal. Given a GBSP problem (95) – (96) and a tolerance $\epsilon > 0$, our goal will be to find an ϵ-solution to the problem, that is, a pair ρ_ϵ, $z_1^\epsilon \in Z_1$ such that

$$\rho_\epsilon \geq \rho_* \text{ and } \max_{z_2 \in Z_2} \phi^{\rho_\epsilon}(z_1^\epsilon, z_2) \leq \rho_\epsilon \epsilon \tag{98}$$

We are about to point out several important application examples for GBSP problem.

4.2.2.2 Application example: ℓ_1-minimization with ℓ_p-fit

The problem of interest is

$$\mathrm{Opt} = \min_x \{\|x\|_1 : \|Ax - b\|_p \leq \delta\} \quad [A \in \mathbb{R}^{m \times n}]. \tag{99}$$

Different versions of this problem arise in sparsity-oriented signal processing and compressed sensing. Setting $x = \rho u$, $\|u\|_1 \leq 1$, we rewrite the problem equivalently as

$$\frac{1}{\mathrm{Opt}} = \rho_* = \max\left\{\rho : \min_{\|u\|_1 \leq 1} \|Au - \rho b\|_p - \rho\delta \leq 0\right\}, \tag{100}$$

or, which is the same as

$$\frac{1}{\mathrm{Opt}} = \rho_* = \max\left\{\rho : \Phi(\rho) = \min_{\|u\|_1 \leq 1,\, \|v\|_p \leq 1} \|Au - \rho b - \rho\delta v\|_\infty \leq 0\right\}.$$

This is nothing but the GBSP problem (95) with $SV(\rho) = \phi^\rho(z),\ z \in Z$, given by

$$\begin{aligned}\phi^\rho(z_1(=[z_{11}; z_{12}]), z_2) &= z_2^T J_m^T \left(AJ_n z_{11} - \rho[b + \delta z_{12}]\right), \\ Z_1 = \underbrace{\Delta_{2n}}_{Z_{11}} &\times \underbrace{\{z_{12} \in \mathbb{R}^m : \|z_{12}\|_p \le 1\}}_{Z_{12}},\ Z_2 = \Delta_{2m},\end{aligned} \tag{101}$$

where we denote $J_k = [I_k, -I_k]$, I_k being $k \times k$ identity matrix. This problem satisfies **[A.1]**; when $\|b\|_p > \delta$ (otherwise the optimal solution to (99) is $x = 0$), the problem satisfies **[A.2]** as well. The associated saddle value function is

$$\begin{aligned}\mathrm{SV}(\rho) &= \max_{z_2 \in \Delta_{2m}} \min_{z_{11} \in \Delta_{2n},\, z_{12} \in Z_{12}} \left[z_2^T J_m^T \left(AJ_n z_{11} - \rho[b - \delta z_{12}]\right)\right] \\ &= \max_{w = J_m z_2,\, z_2 \in \Delta_{2m}} \min_{u = J_n z_1,\, z_1 \in \Delta_{2n}} \min_{z_{12} \in Z_{12}} \left[w^T (Au - \rho[b + \delta z_{12}])\right] \\ &= \max_{\|w\|_1 \le 1} \min_{\|u\|_1 \le 1} \min_{\|v\|_p \le 1} \left[w^T (Au - \rho[b + \delta v])\right] = \Phi(\rho).\end{aligned}$$

Suppose that we are given an ε-solution ρ_ε, $z_1^\varepsilon = [z_{11}^\varepsilon; z_{12}^\varepsilon]$ to the problem (98), (101) with $\varepsilon = \epsilon m^{-\frac{1}{p}}$. When setting $x_\varepsilon = \rho_\varepsilon^{-1} J_n z_{11}^\varepsilon$ and $v_\varepsilon = z_{12}^\varepsilon$ we get an approximate solution to (99) such that

$$\|x_\epsilon\|_1 \le \mathrm{Opt}\ \&\ \|Ax_\epsilon - b\|_p \le \|\delta v_\varepsilon\|_p + \|Ax_\varepsilon - b - \delta v_\varepsilon\|_p \le \delta + \varepsilon m^{1/p} = \delta + \epsilon.$$

Finally, we associate with $z = [z_{11}; z_{12}; z_2] \in Z = Z_1 \times Z_2$ distribution P_z satisfying [P], namely, as follows. Note that for $z \in Z$, z_{11} and z_2 are vectors from the standard simplices and thus can be considered as probability distributions on the corresponding index sets $\{1, ..., 2n\}$, $\{1, ..., 2m\}$. To generate $\zeta = [\zeta_{11}; \zeta_{12}; \zeta_2] \sim P_z$, we draw at random index $\imath$ from the distribution z_{11} and make $[\zeta_{11}]_\imath = 1$ the only nonzero entry in ζ_{11}. ζ_2 is built similarly, with z_2 in the role of z_{11}, and ζ_{12} is nothing but z_{12}. It is immediately seen that it takes just $O(m + n)$ a.o. to generate a sample $\zeta \sim P_z$ and to compute the vector $F^\rho(\zeta)$.

It is worth to mention that in the important case $p = \infty$ the construction of the GBSP which corresponds to (99) can be substantially simplified. Indeed, one can see immediately that for $p = \infty$ (100) is equivalent to the GBSP problem on the direct product of just two unit ℓ_1-balls (since $\|Az_1 - b\|_\infty = \max_{\|z_2\|_1 \le 1} z_2^T (Az_1 - b)$). It is more convenient to pass from ℓ_1-balls to the standard simplexes, as it was done in the case of (101). The resulting GBSP problem is given by

$$\begin{aligned}\phi^\rho(z_1, z_2) &= z_2^T J_m^T AJ_n z_1 - \rho z_2^T J_m^T b - \rho\delta, \\ Z_1 = Z_{11} &= \Delta_{2n},\ Z_{12} = \{0\}, Z_2 = \Delta_{2m},\end{aligned} \tag{102}$$

and satisfies **[A.1]** and **[A.2]** when $\delta < \|b\|_\infty$.

4.2.2.3 Application example: ℓ_1 Support Vector Machine.

One of the "statistically solid" SVM models (see [36] and [115, Section 2.3.3]) is as follows. We are given a *training sample* — a matrix $X \in \mathbb{R}^{m\times n}$ with rows representing feature vectors, and a vector $y \in \mathbb{R}^m$ with entries ± 1 representing labels. Setting $R = \max_{i,j} |X_{ij}|$, $Y = \mathrm{Diag}\{y\}$ and $\mathbf{1} = [1; ...; 1] \in \mathbb{R}^m$, we want to solve the margin optimization problem

$$\mathrm{Opt} = \max_{w,b,\rho} \{\rho : \|w\|_1 \le 1, \|[\rho\mathbf{1} - Y[Xw + b\mathbf{1}]]_+\|_2 \le R\}, \tag{103}$$

where $[z]_+$ is the vector with coordinates $[z_i]_+ := \max[z_i, 0]$. We can convert this problem into a GBSP one as follows. Observe first that

$$\begin{array}{rcl} \mathrm{Opt} & = & \max\left\{\rho : \min\limits_{\|w\|_1\le 1,\, \|v\|_2\le 1} \max\limits_{1\le i\le m} [\rho\mathbf{1} - Y[Xw + b\mathbf{1}] - Rv]_i \le 0\right\} \\ & = & \max\left\{\rho : \min\limits_{\|w\|_1\le 1,\, \|v\|_2\le 1, b} \max\limits_{u\in\Delta_m} u^T[\rho\mathbf{1} - Y[Xw + b\mathbf{1}] - Rv] \le 0\right\}. \end{array}$$

Assuming that the entries of y contain both 1 and -1 and setting $\Delta_m^+ = \{u \in \Delta_m : y^T u = 0\}$, we have $\min\limits_b \max\limits_{u\in\Delta_m} u^T[\rho\mathbf{1} - Y[Xw + b\mathbf{1}] - Rv] = \max\limits_{u\in\Delta_m^+} u^T(\rho\mathbf{1} - YXw - Rv)$. Hence, when setting $w = J_n s$, we come to

$$\mathrm{Opt} = \max\left\{\rho : \min_{s\in\Delta_{2n}, \|v\|_2\le 1} \max_{u\in\Delta_m^+} u^T[\rho\mathbf{1} - YXJ_n s - Rv] \le 0\right\}.$$

We see that (103) is equivalent to the GBSP problem given by

$$\begin{array}{c} \phi^\rho(z_1 = [z_{11} = s; z_{12} = v], z_2 = u) = u^T[\rho\mathbf{1} - YXJ_n s - Rv], \\ Z_1 = \{[s; v] : s \in \Delta_{2n}, \|v\|_2 \le 1\}, Z_2 = \Delta_m^+. \end{array}$$

Note that this problem clearly satisfies $\mathbf{A.1-2}$. Besides this, $\sqrt{m}\max_i [x_i]_+ \ge \|[x]_+\|_2$, so that an ϵ-solution $(\rho_\epsilon, z_1^\epsilon = [s^\epsilon; v^\epsilon])$ to the GBSP problem induces the approximate solution $(\rho_\epsilon, w^\epsilon = J_n s^\epsilon, b^\epsilon)$ to (103) such that

$$\rho_\epsilon \ge \mathrm{Opt} \;\&\; \|[\rho_\epsilon\mathbf{1} - Y[Xw^\epsilon + b^\epsilon\mathbf{1}]]_+\|_2 \le R + \sqrt{m}\rho_\epsilon\epsilon,$$

whence $(\rho_\epsilon(1 - \sqrt{m}\epsilon), w^\epsilon, b^\epsilon)$ is a feasible solution to (103) with the value of the objective $\ge (1 - \sqrt{m}\epsilon)\mathrm{Opt}$.

Finally, we associate with $z = [z_{11} = s; z_{12} = v; z_2 = u] \in Z = Z_1 \times Z_2$ a distribution P_z on $Z = Z_1 \times Z_2$ defined as follows. To generate $\zeta = [\zeta_{11}; \zeta_{12}; \zeta_2] \sim P_z$, we pick at random $\imath \in \{1, ..., 2n\}$, with $\text{Prob}\{\imath = i\} = [z_{11}]_i$, $1 \leq i \leq 2n$, and set $\zeta_{11} = e_\imath$, e_i being the basic orths in $\mathbb{R}^{2n}$. We always set $\zeta_{12} = v$. To generate ζ_2, we act as follows. Let $I = \{i : y_i = 1\}$, $J = \{i : y_i = -1\}$, and let $p := \sum_{i \in I} u_i = \sum_{j \in J} u_j$ (recall that $\sum_i y_i u_i = 0$, that is, $\sum_{i \in I} u_i = \sum_{j \in J} u_j$). Note that $p \leq 1/2$ due to $\sum_{j=1}^m u_j \leq 1$. We first flip a coin with probability $1 - 2p$ to get head; if head appears, we set $\zeta_2 = 0$. If tail appears, we pick at random $\imath \in I$ with $\text{Prob}\{\imath = i\} = u_i/p$, $i \in I$, pick at random $\jmath \in J$ with $\text{Prob}\{\jmath = j\} = u_j/p$, $j \in J$, and set $\zeta_2 = \frac{1}{2}[e_\imath + e_j]$, e_i being the basic orths in $\mathbb{R}^m$. It is immediately seen that P_z satisfies [P], and that it takes just $O(m+n)$ a.o. to generate a sample $\zeta \sim P_z$ and to compute the vector $F^\rho(\zeta)$.

4.3 Solving Bilinear Saddle Point Problem

We are about to present two randomized first order methods for solving BSPs and hence will be utilized in solving GBSPs — the *Stochastic Approximation* (SA) and the *Stochastic Mirror Prox* (SMP) algorithms, which are the randomized versions of the methods proposed in [100] and [101] respectively. Both SA and SMP are directly applicable to a BSP problem, and this is the situation we are about to consider here; the GBSP case will be considered in Section 4.4.

4.3.1 The Setup

Both SA and SMP algorithms are aimed at solving a BSP problem $(\mathcal{S})$. The setup for these methods is given by

- a norm $\|\cdot\|$ on the Euclidean space E where the domain $Z = Z_1 \times Z_2$ of $(\mathcal{S})$ lives, along with the conjugate norm $\|\zeta\|_* = \max_{\|z\| \leq 1} \langle \zeta, z \rangle$;
- a *distance-generating function* (d.g.f.) $\omega(z)$ which is convex and continuous on Z, admits continuous on the set $Z^o = \{z \in Z : \partial\omega(z) \neq \emptyset\}$ selection $\omega'(z)$ of subgradient (here $\partial\omega(x)$ is a subdifferential of $\omega\big|_Z$ taken at z), and is strictly convex with modulus

1 w.r.t. $\|\cdot\|$:

$$\forall z', z'' \in Z^o : \langle \omega'(z') - \omega'(z''), z' - z'' \rangle \geq \|z' - z''\|^2.$$

We shall refer to the latter property as to *compatibility* of $\omega(\cdot)$ and $\|\cdot\|$.

A d.g.f. ω gives rise to several important for us entities:

1. *Bregman distance* $V_z(u) = \omega(u) - \omega(z) - \langle \omega'(z), u - z \rangle$, where $z \in Z^o$ and $u \in Z$;
2. *Prox-mapping* $\mathrm{Prox}_z(\xi) = \mathrm{argmin}_{w \in Z} \{\langle \xi, w \rangle + V_z(w)\} : E \to Z^o$; here $z \in Z^o$ is a "prox center;"
3. "ω-center" $z_\omega = \mathrm{argmin}_{z \in Z} \omega(z) \in Z^o$ of Z and the quantities

$$\Omega = \max_{z \in Z} V_{z_\omega}(z) \leq \max_{z \in Z} \omega(z) - \min_{z \in Z} \omega(z), \quad \Theta = \sqrt{2\Omega}. \tag{104}$$

In the sequel, we set

$$\mathcal{R} := \max_{z \in Z} \|z - z_\omega\| \leq \Theta, \tag{105}$$

where the concluding inequality follows from the fact that for every $z \in Z$ one has $\frac{1}{2}\|z - z_\omega\|^2 \leq V_{z_\omega}(z)$ by strong convexity of $\omega(\cdot)$. We also denote by $\mathcal{L}$ the $(\|\cdot\|, \|\cdot\|_*)$-Lipschitz constant of F:

$$\|F(z) - F(z')\|_* = \|\mathcal{A}(z - z')\|_* \leq \mathcal{L}\|z - z'\|, \quad \forall z, z'; \tag{106}$$

and set

$$M_* = \max_{z, z' \in Z} \|F(z) - F(z')\|_* \leq 2\mathcal{R}\mathcal{L} \leq 2\Theta\mathcal{L}, \tag{107}$$

$$F_* = \max_{z \in Z} \|F(z)\|_* \leq \|a\|_* + M_* \leq \|a\|_* + 2\Theta\mathcal{L}. \tag{108}$$

4.3.2 The SA and SMP Algorithms

Assume we have access to an "oracle" $\mathcal{O}$ which, at i-th call ($i = 1, 2, ...$), returns a vector $\xi_i \in E$ (this vector can be random with distribution depending on previous calls and, more

generally, on the history of our computational process before the call). This oracle gives rise to two conceptual algorithms:

$$\begin{array}{ll}(a): & z_1 = z_\omega; \{z_t, \xi_t\} \mapsto \{z_{t+1} = \mathrm{Prox}_{z_t}(\gamma_t \xi_t), \xi_{t+1}\},\ t = 1, 2, ... \\ (b): & z_1 = z_\omega; \{z_t, \xi_{2t-1}\} \mapsto \{w_t = \mathrm{Prox}_{z_t}(\gamma_t \xi_{2t-1}), \xi_{2t}\} \mapsto \{z_{t+1} = \mathrm{Prox}_{z_t}(\gamma_t \xi_{2t}), \xi_{2t+1}\}, \\ & t = 1, 2, ...\end{array} \tag{109}$$

here $\gamma_1, \gamma_2, ...$ are positive stepsizes defined in a non-anticipative fashion, that is, γ_t depends on oracle's answers obtained prior to step t (i.e., γ_t depends solely on $\xi_1, ..., \xi_{t-1}$ in the case of (a), and solely on $\xi_1, ..., \xi_{2t-2}$ in the case of (b)). We refer to (109.a,b) as the Stochastic Approximation (SA) and Stochastic Mirror Prox (SMP) schemes, respectively. We will consider two implementations of these schemes, the *basic* and the *advanced* ones.

4.3.2.1 Basic implementation

Recall that we have associated with $(\mathcal{S})$ the affine operator $F(z) : Z \to E$ given by (89), and with every point $z \in Z$ — a probability distribution P_z supported on Z satisfying $\mathbf{E}_{\zeta \sim P_z}\{\zeta\} = z$. Suppose that

- the stepsizes $\gamma_t > 0$ are chosen in a non-anticipating fashion such that $\gamma_1 \geq \gamma_2 \geq ...$;
- in SA: ζ_t is drawn at random from the distribution P_{z_t}, and $\xi_t = F(\zeta_t)$;
- in SMP: $\xi_{2t-1} = F(\eta_t)$ with η_t drawn at random from the distribution P_{z_t}, and $\xi_{2t} = F(\zeta_t)$ with ζ_t drawn at random from the distribution P_{w_t}.

The approximate solution generated by the short-step SA/SMP in course of $t = 1, 2, ...$ steps is

$$z^t = t^{-1} \sum_{\tau=1}^{t} \zeta_\tau. \tag{110}$$

4.3.2.2 Advanced implementation

In Advanced implementation of SA and SMP, same as in the Basic one, the stepsizes $\gamma_t > 0$ still are chosen in a non-anticipating fashion, but the restriction $\gamma_1 \geq \gamma_2 \geq ...$ is now lifted.

To explain how the oracle is built, observe that if $u \in Z$, then

$$\mathbf{E}_{\zeta\sim P_u}\{\langle F(\zeta), \zeta - u\rangle\} = 0$$

(recall that $F(z) = a + \mathcal{A}z$ with skew symmetric $\mathcal{A}$ and that $\mathbf{E}_{\zeta\sim P_u}\{\zeta\} = u$). It follows that given u and generating one by one independent samples $\eta^s \sim P_u$, $s = 1, 2, ...$, one with probability 1 eventually generates ζ such that

$$\langle F(\zeta), \zeta - u\rangle \leq 0. \tag{111}$$

At step t of SA, in order to define ξ_t, the oracle draws one by one samples $\eta^s \sim P_{z_t}$, $s = 1, 2, ...$, until a sample $\zeta_t := \eta^s$ satisfying (111) with $u = z_t$ is generated; when it happens, the oracle returns $\xi_t = F(\zeta_t)$. At a step t of SMP, the oracle is invoked twice, first to generate $\xi_{2t-1} = F(\eta_t)$, and then to generate $\xi_{2t} = F(\zeta_t)$. ξ_{2t-1} is generated exactly as in the basic implementation — by drawing a sample $\eta_t \sim P_{z_t}$ and returning $\xi_{2t-1} = F(\eta_t)$. To generate ξ_{2t}, the oracle draws one by one samples $\eta^s \sim P_{w_t}$, $s = 1, 2, ...$, until a sample $\zeta_t = \eta^s$ satisfying (111) with $u = w_t$ is generated; when it happens, the oracle returns $\xi_{2t} = F(\zeta_t)$.

Finally, in the advanced implementation we replace the rule (110) for generating approximate solutions with the rule

$$z^t = \frac{1}{\sum_{\tau=1}^t \gamma_\tau} \sum_{\tau=1}^t \gamma_\tau \zeta_\tau. \tag{112}$$

4.3.2.3 Quantifying quality of approximate solutions

Observe that by construction at a step τ both ζ_τ and $F(\zeta_\tau)$ become known. Recalling that F is affine, it follows that after t steps we have at our disposal both the approximate solution $z^t = [z_1^t; z_2^t]$ and the vector $F(z^t)$. As a result, with both Basic and Advanced implementations of both SA and SMP, after $t = 1, 2, ...$ steps we have at our disposal the quantities

$$\overline{\phi}(z_1^t) = \upsilon + \langle a_1, z_1^t\rangle + \max_{z_2\in Z_2} \langle z_2, -F_2(z_1^t)\rangle,\ \underline{\phi}(z_2^t) = \upsilon + \langle a_2, z_2^t\rangle + \min_{z_1\in Z_1} \langle z_1, F_1(z_2^t)\rangle \tag{113}$$

(see (89)) and consequently we know the residual $\epsilon_{\text{sad}}(z^t) = \overline{\phi}(z^t) - \underline{\phi}(z^t)$ of the current approximate solution z^t. As we shall see in Section 4.4, this feature of our algorithms

becomes instrumental when solving GBSP problems.[3] This is in sharp contrast with the prototypes of the SA and the SMP proposed, respectively, in [102, Section 3.3] and [82]. The approximate solutions z^t of those algorithms were computed according to the formula (112), but with z_τ [102] or w_τ [82] in the role of ζ_τ. As a result, in the prototype algorithms there is no computationally cheap way to quantify the quality of approximate solutions.

4.3.2.4 Efficiency estimates for basic implementation

The accuracy bounds for Basic SA and SMP algorithms are given by the following

Proposition 4.3.1 *Let the BSP problem $(\mathcal{S})$ be solved by the short-step SA or SMP algorithm with positive stepsizes $\gamma_1 \geq \gamma_2 \geq ...$ chosen in a non-anticipative fashion. Then*

(i) *For every $t \geq 1$, for both SA and SMP one has*

$$\epsilon_{sad}(z^t) \quad \leq \quad t^{-1}\left[\gamma_t^{-1}\Omega + R_t + S_t\right], \; R_t := \sum_{\tau=1}^{t} r_\tau, \; S_t := \sum_{\tau=1}^{t} s_\tau, \tag{114}$$

where

$$r_t \quad = \quad \begin{cases} \langle F(\zeta_t), \zeta_t - z_t\rangle & \text{in the case of SA,} \\ \langle F(\zeta_t), \zeta_t - w_t\rangle & \text{in the case of SMP,} \end{cases}$$

$$s_t \quad = \quad \begin{cases} \langle F(\zeta_t), z_t - z_{t+1}\rangle - \gamma_t^{-1} V_{z_t}(z_{t+1}), & \text{in the case of SA,} \\ \langle F(\zeta_t), w_t - z_{t+1}\rangle - \gamma_t^{-1} V_{z_t}(z_{t+1}), & \text{in the case of SMP.} \end{cases}$$

We have

$$s_t \leq \begin{cases} \frac{\gamma_t}{2}\|F(\zeta_t)\|_*^2, & \text{in the case of SA,} \\ \frac{\gamma_t}{2}\|F(\zeta_t) - F(\eta_t)\|_*^2 - \frac{1}{2\gamma_t}\|w_t - z_t\|^2, & \text{in the case of SMP,} \end{cases} \tag{115}$$

with

$$s_t \leq \begin{cases} \frac{\gamma_t}{2} F_*^2, & \text{in the case of SA,} \\ \frac{\gamma_t}{2} M_*^2, & \text{in the case of SMP.} \end{cases} \tag{116}$$

In particular, if the stepsizes $\gamma_t > 0$ satisfy $S_t \leq \Omega/\gamma_t$, $t = 1, 2, ...$, then

$$\epsilon_{sad}(z^t) \leq \frac{2\Omega}{t\gamma_t} + \frac{R_t}{t}. \tag{117}$$

[3] Of course, computing the quantities in (113) is not completely costless; note, however, that the cost of this computation is dominated by the cost of computing the prox-mapping(s) at a step and thus is a small fraction of the overall computational effort.

(ii) *Further,* $\mathbf{E}\{R_t\} = 0$*, and in the case of SMP, under additional assumption that*

$$\gamma_t \leq (\sqrt{3}\mathcal{L})^{-1}, \tag{118}$$

we have

$$s_t \leq \frac{3\gamma_t}{2}\left[\|\mathcal{A}(\zeta_t - w_t)\|_*^2 + \|\mathcal{A}(\eta_t - z_t)\|_*^2\right], \tag{119}$$

so that $\mathbf{E}\{s_t\} \leq 3\gamma_t\sigma^2$*, where*

$$\sigma^2 = \sup_{z\in Z} \mathbf{E}_{\zeta\sim P_z}\left\{\|\mathcal{A}(\zeta - z)\|_*^2\right\} \leq M_*^2. \tag{120}$$

In particular, if the stepsizes $\gamma_t > 0$ *satisfy* $\mathbf{E}\{S_t\} \leq \Omega/\gamma_t$ *for* $t = 1, 2, ...,$ *then*

$$\mathbf{E}\{\epsilon_{sad}(z^t)\} \leq \frac{2\Omega}{t\gamma_t}.$$

The bound of Proposition 4.3.1 allows to easily conceive stepsize policies. Let us start with *offline* policies, where γ_t are chosen in advance deterministic reals. If the number of steps N is fixed in advance, one can use constant stepsizes $\gamma_1 = ... = \gamma_N = \gamma$. In particular, when choosing

$$\gamma = \begin{cases} \frac{1}{F_*}\sqrt{\frac{2\Omega}{N}}, & \text{in the case of SA} \quad (a) \\ \min\left\{\frac{1}{\sigma}\sqrt{\frac{\Omega}{3N}}, \frac{1}{\sqrt{3}\mathcal{L}}\right\}, & \text{in the case of SMP} \quad (b) \end{cases} \tag{121}$$

(by (116), (108) this choice implies that $\mathbf{E}\{S_t\} \leq \Omega/\gamma_t$, $1 \leq t \leq N$), Proposition 4.3.1 implies the efficiency bound

$$\mathbf{E}\{\epsilon_{\rm sad}(z^N)\} \leq \begin{cases} F_*\sqrt{\frac{2\Omega}{N}}, & \text{in the case of SA} \quad (a) \\ \max\left\{2\sigma\sqrt{\frac{3\Omega}{N}}, \frac{2\sqrt{3}\Omega\mathcal{L}}{N}\right\}, & \text{in the case of SMP} \quad (b) \end{cases} \tag{122}$$

When the number of steps is not fixed in advance, one can use the decreasing stepsizes

$$\forall t \geq 1, \quad \gamma_t = \begin{cases} \frac{1}{F_*}\sqrt{\frac{\Omega}{t}}, & \text{in the case of SA,} \\ \min\left\{\frac{1}{\sigma}\sqrt{\frac{\Omega}{6t}}, \frac{1}{\sqrt{3}\mathcal{L}}\right\}, & \text{in the case of SMP,} \end{cases} \tag{123}$$

which result in the accuracy bound

$$\forall t \geq 1, \quad \mathbf{E}\{\epsilon_{\rm sad}(z^t)\} \leq \begin{cases} 2F_*\sqrt{\frac{\Omega}{t}}, & \text{in the case of SA} \quad (a) \\ \max\left\{2\sigma\sqrt{\frac{6\Omega}{t}}, \frac{2\sqrt{3}\Omega\mathcal{L}}{t}\right\}, & \text{in the case of SMP} \quad (b) \end{cases} \tag{124}$$

completely similar to (122).

4.3.2.5 *Online stepsize policies*

From theoretical viewpoint, the main advantage of the outlined versions of SA and SMP with the "theoretically optimal" offline stepsize policies (121) and (123) are the explicit (and in fact — the best known under circumstances) efficiency estimates (122), (124). While they may appear attractive also from the practical viewpoint because of their apparent simplicity, their use may present several disadvantages: the quantity σ involved in the stepsize computation may not be available at hand and should be evaluated. Besides this, these policies are offline and worst-case oriented; we would prefer more flexible on line adjustable stepsizes.

A natural way to adjust the stepsizes online would be to choose at each step $t \geq 1$ the largest $\gamma_t \leq \gamma_{t-1}$ ensuring the balance $\Omega/\gamma_t \geq S_t$, and thus the bound (117). This idea cannot be implemented "as is," since the stepsize policy should be non-anticipative, while s_t is not yet available when γ_t is computed. This difficulty can be easily circumvented by using instead of s_t its a priori upper bound, which is either $\frac{\gamma_t}{2}F_*$ for the SA algorithm or $\frac{\gamma_t}{2}M_*^2$ for the SMP, see (115). Specifically, consider the online policy of choosing γ_t, $t \geq 1$ as follows:

$$\Omega\gamma_t^{-2} = \begin{cases} 2\sum_{\tau=1}^{t-1}\gamma_\tau^{-1}[s_\tau]_+ + F_*^2 & \text{in the case of SA,} \\ 2\sum_{\tau=1}^{t-1}\gamma_\tau^{-1}[s_\tau]_+ + 8\Omega\mathcal{L}^2 & \text{in the case of SMP,} \end{cases} \tag{125}$$

where we set $\sum_{\tau=1}^{0}\gamma_\tau^{-1}[s_\tau]_+ = 0$. With this policy, one clearly has $\gamma_1 \geq \gamma_2 \geq ....$

Proposition 4.3.2 *Let positive stepsizes γ_t, $t = 1, 2, ...$ of the Basic SA/SMP implementation be chosen according to (125). Then the approximate solution z^t satisfies*

$$\epsilon_{sad}(z^t) \leq \frac{(1+\sqrt{2})\Omega}{t\gamma_t} + \frac{R_t}{t}. \tag{126}$$

As a consequence, we have

$$\epsilon_{sad}(z^t) \leq \begin{cases} \frac{(1+\sqrt{2})\sqrt{\Omega}}{t}\left(F_*^2 + \sum_{\tau=1}^{t-1}\|F(\zeta_\tau)\|_*^2\right)^{1/2} + \frac{R_t}{t}, & \text{in the case of SA} \quad (a) \\ \frac{(1+\sqrt{2})\sqrt{\Omega}}{t}\left(8\Omega\mathcal{L}^2 + \sum_{\tau=1}^{t-1}\varsigma_\tau\right)^{1/2} + \frac{R_t}{t} & \\ \qquad \leq \frac{7\Omega\mathcal{L}}{t} + \frac{R_t}{t} + \frac{(1+\sqrt{2})\sqrt{\Omega}}{t}\sqrt{\sum_{\tau=1}^{t-1}\varsigma_\tau}, & \text{in the case of SMP} \quad (b) \end{cases} \tag{127}$$

where

$$\varsigma_t = 3\left[\|F(\zeta_t) - F(w_t)\|_*^2 + \|F(\eta_t) - F(z_t)\|_*^2\right]. \tag{128}$$

Recalling that $\mathbf{E}\{R_t\} = 0$ and $\mathbf{E}\{\varsigma_t\} \le 6\sigma^2$ (see (120)), we arrive at

Corollary 4.3.1 *Under the premise of Proposition 4.3.2, for the SMP algorithm one has*

$$\mathbf{E}\{\epsilon_{sad}(z^t)\} \le \frac{7\Omega\mathcal{L}}{t} + \frac{6\sqrt{\Omega}\sigma}{\sqrt{t}}. \tag{129}$$

Note that the bounds (127.a) and (129) within an absolute constant factor coincide with the respective bounds in (124), that is, our online stepsizes policy (which, in contrast to (123), does not require knowledge of σ) is not worse that the "theoretically optimal" stepsize policies underlying (124).

4.3.2.6 Discussion

Since $F_* \ge \mathcal{R}\mathcal{L} \ge \sigma/2$ (cf. (105)), the SA efficiency estimate (124.a) is at most within an absolute constant factor better than the corresponding estimate for the SMP. Besides this, the SMP bound (124.b) says that when the noise level σ of the oracle is small enough (specifically, $\sigma^2 = O\left(\frac{\Omega\mathcal{L}^2}{N}\right)$), then $\mathbf{E}\{\epsilon_{\rm sad}(z^t)\} \le O(1)\frac{\Omega\mathcal{L}}{N}$, which, modulo expectation of the residual instead of the residual itself, coincides with the best known so far efficiency estimate of the deterministic first order algorithms solving bilinear saddle point problems. On the other hand, we do have a possibility to make σ small. The trivial way to do so is to use $P_z = \delta_z$, which results in $\sigma = 0$ and makes SMP a version of the Deterministic Mirror Prox algorithm (DMP) proposed in [101]. Another, more attractive, option to control σ is as follows. Given the family of distributions P_z supported on Z and such that $\mathbf{E}_{\zeta\sim P_z}\{\zeta\} = z$, and a positive integer k, we can convert P_z into the family of distributions $P_z^{(k)}$ with the same property as follows: in order to generate a random vector $\zeta \sim P_z^{(k)}$ and to compute $F(\zeta)$, we draw a k-element sample $\zeta^1, ..., \zeta^k$ from the distribution P_z, compute $F(\zeta^1), ..., F(\zeta^k)$ and then set $\zeta = \frac{1}{k}\sum_{i=1}^k \zeta^i$, so that

$$F(\zeta) = \frac{1}{k}\sum_{i=1}^k F(\zeta^i).$$

If, as in the examples of Section 4.2, drawing $\zeta^i \sim P_z$ and computing $F(z^i)$ is much cheaper than computing $F(z)$, the outlined procedure with a "reasonably large" value of k is still significantly cheaper than the direct computation of $F(z)$. At the same time, for "good enough" norms $\|\cdot\|_*$, passing from P_z to $P_z^{(k)}$ can significantly reduce the noise level σ. Specifically, given a norm $\|\cdot\|_*$ on a finite-dimensional Euclidean space E, one can associate with it its *regularity parameter* $\varkappa \geq 1$ (see Section 2.2, [77] for details) to ensure the following: whenever $k > 0$ is an integer and $\xi^1, ..., \xi^k$ are independent vectors from E with $\mathbf{E}\{\xi^i\} = 0$ and $\mathbf{E}\{\|\xi^i\|_*^2\} \leq \alpha_i^2$ and $\alpha = \max_i \alpha_i$, for $\xi = \frac{1}{k}\sum_{i=1}^k \xi_i$ it holds

$$\mathbf{E}\{\|\xi\|_*^2\} \leq \min\left[\frac{1}{k}, \frac{\varkappa}{k^2}\right] \sum_{i=1}^k \alpha_i^2 \leq \min\left[1, \frac{\varkappa}{k}\right] \alpha^2.$$

Suppose now that when running SMP we sample ζ_t, η_t from the distributions $P_z^{(k)}$ for some $k > 0$. It follows that if $\|\cdot\|_*$ is $\varkappa$-regular with certain $\varkappa$, then, passing from P_z to $P_z^+ = P_z^{(k)}$, we can reduce the "original" value of σ to the value $\sigma^+ = \min[1, \sqrt{\frac{\varkappa}{k}}]\sigma$. We shall see in a while that in the applications we have mentioned so far, $\varkappa$ is "small" — at most logarithmic in $\dim Z$. The bottom line is that there is a tradeoff between the computational cost of a call to a stochastic oracle and the noise level σ. Consequently, in the case of SMP, it is possible to tradeoff the computational effort per iteration and the iteration count to obtain an approximate solution of the desired expected quality, and we can use this tradeoff in order to save on the overall amount of computations. This option (which is the major advantage of SMP as compared to SA) is especially attractive when among the two components of our computational effort per iteration — one related to computing η_t, ζ_t, $F(\eta_t)$ $F(\zeta_t)$, and the other aimed at computing the prox mappings – the second component is essentially more significant than the first one. In such a situation, we basically can only gain by passing from P_z to $P_z^{(k)}$ with k chosen to balance the outlined two components of the computational effort.

4.3.2.7 Large deviations

In the above efficiency estimates, say, in (129), we upper-bounded the *expected* inaccuracy of approximate solutions z^t. In fact, one can get exponential upper bounds on probabilities

of large deviations for the inaccuracy of the approximate solution. Though we do not need such bounds to access the inaccuracy of solutions, they are still useful to provide *theoretical* guarantees for the complexity of our algorithms (cf. Theorem 4.4.1 in the next section).

For the sake of definiteness, when presenting large deviation results, we restrict ourselves to the SMP algorithm and the stepsize strategy (125). Note that one can easily derive a deviation bound from the bound (129) on the expectation of $\epsilon_{\rm sad}(z^t)$ in the previous section. Indeed, let us fix the number t of iterations, run the algorithm m times and select the best, in terms of $\epsilon_{\rm sad}(\cdot)$, of the resulting approximate solutions. The probability that for this solution $\epsilon_{\rm sad}(\cdot)$ is worse than, say, twice the right hand side of (129) is at most 2^{-m} and thus can be made negligibly small with quite moderate values of m.

We also have the following bound on the deviations of the algorithm without restarts:

Proposition 4.3.3 *Assume we are solving problem $(\mathcal{S})$ by Basic implementation of SMP where ζ_t, η_t are sampled from the distributions $P_z^{(k)}$, $k \geq 1$ being a parameter of the construction. Assume also that the norm $\|\cdot\|_*$ is $\varkappa$-regular, and the online stepsize policy (125) is used. Then there are absolute constants K_0, K_1 such that the approximate solution z^t satisfies for all $t \geq 1$ and λ, $\Lambda \geq 0$*

$$\text{Prob}\left\{\epsilon_{sad}(z^t) \geq K_0\left[\frac{\Theta^2\mathcal{L}}{t} + \frac{\varkappa_*(k,\Lambda)\Theta^2\mathcal{L}}{\sqrt{kt}} + \Theta(\|a\|_* + \Theta\mathcal{L})\sqrt{\frac{\lambda}{kt}}\right]\right\} \leq e^{-\Lambda t} + e^{-\lambda}, \quad (130)$$

where $\varkappa_*(k,\Lambda) = \sqrt{\min[k,\ (\varkappa+\Lambda)]}$. In particular, one has for all $\varepsilon > 0$:

$$\begin{array}{l}\text{Prob}\{\epsilon_{sad}(z^N) \geq \varepsilon\} \leq e^{-\Lambda N} + e^{-\lambda} \quad \text{for } N \geq N_\varepsilon, \text{ where}\\ N_\varepsilon = K_1\text{Ceil}\left(\max\left[\Theta^2\mathcal{L}\varepsilon^{-1},\ \frac{\varkappa_*^2(k,\Lambda)\Theta^4\mathcal{L}^2}{k\varepsilon^2},\ \frac{(\|a\|_*+\Theta\mathcal{L})^2\Theta^2\lambda}{k\varepsilon^2}\right]\right).\end{array} \quad (131)$$

4.3.3 Efficiency Estimates for Advanced Implementations of SA and SMP

The efficiency of Advanced implementations of SA and SMP stem from the following result (we use the notation from Section 4.3.1):

Proposition 4.3.4 *Let the BSP problem $(\mathcal{S})$ be solved by the advanced-step SA or SMP algorithms. Then for every $t \geq 1$, for both SA and SMP one has*

$$\epsilon_{sad}(z^t) \leq \Gamma_t^{-1}\left[\Omega + R_t + S_t\right] = \Gamma_t^{-1}\left[\Omega + \sum_{\tau=1}^t r_\tau + \sum_{\tau=1}^t s_\tau\right], \quad (132)$$

where

$$\begin{aligned}
\Gamma_t &= \sum_{\tau=1}^{t}\gamma_\tau,\\
r_t &= \begin{cases} \gamma_t\langle F(\zeta_t),\zeta_t-z_t\rangle & \text{in the case of SA}\\ \gamma_t\langle F(\zeta_t),\zeta_t-w_t\rangle & \text{in the case of SMP}\end{cases}\\
s_t &= \begin{cases} \left[\gamma_t\langle F(\zeta_t),z_t-z_{t+1}\rangle - V_{z_t}(z_{t+1})\right], & \text{in the case of SA}\\ \left[\gamma_t\langle F(\zeta_t),w_t-z_{t+1}\rangle - V_{z_t}(z_{t+1})\right], & \text{in the case of SMP}\end{cases}
\end{aligned}$$

with $r_t\le 0$ *and*

$$s_t\le\begin{cases}\frac{\gamma_t^2}{2}\|F(\zeta_t)\|_*^2\le\frac{\gamma_t^2}{2}F_*^2, & \text{in the case of SA}\\ \frac{\gamma_t^2}{2}\|F(\zeta_t)-F(\eta_t)\|_*^2-\frac{1}{2}\|w_t-z_t\|^2\le\frac{\gamma_t^2}{2}M_*^2, & \text{in the case of SMP.}\end{cases}\tag{133}$$

In order to extract from (132) explicit efficiency estimates, we need to specify a stepsize policy. In this respect, the advanced implementations offer more freedom than the basic ones, since now we should not ensure neither the martingale property of the random sums R_t, nor the monotonicity of the stepsizes. One option here is to use constant stepsize policy

$$\gamma_t=\sqrt{\frac{2\Omega}{N}}\cdot\begin{cases}\frac{1}{F_*}, & \text{in the case of SA}\\ \frac{1}{M_*}, & \text{in the case of SMP}\end{cases},\ 1\le t\le N.$$

As it is easily seen, with this policy, (132) results in efficiency estimate (cf. (124))

$$\forall t\ge 1,\ \ \mathbf{E}\left\{\epsilon_{\rm sad}(z^t)\right\}\le O(1)\begin{cases}F_*\sqrt{\frac{\Omega}{t}}, & \text{in the case of SA}\quad (a)\\ \mathcal{R}\mathcal{L}\sqrt{\frac{\Omega}{t}}, & \text{in the case of SMP}\quad (b)\end{cases}\tag{134}$$

Our preliminary experiments, however, suggest to equip the advanced implementations of SA and SMP with the online stepsize policy as follows. Let us set

$$\delta_t=\frac{\Theta^2}{t},\quad S_t^*=\sum_{\tau=1}^{t}\delta_\tau\quad[\le\Theta^2(1+\ln t)]\tag{135}$$

and let us choose γ_τ according to the "greedy" rule (the larger, the better) under the restriction that for all $t=1,2,...$ it holds

$$R_t+S_t\le S_t^*,\tag{$*_t$}$$

see (132). Specifically, assume that we have already carried out $t-1$ steps of the algorithm ensuring the relations $(*_\tau)$, $\tau\le t-1$, and are about to define γ_t in order to carry out step

t and to ensure $(*_t)$. At this time, we know $R_{t-1} \le 0$ and S_{t-1}, same as know for sure that whatever be our choice of $\gamma_t > 0$, we would have

$$R_t - R_{t-1} = r_t \le 0, \;\; S_t - S_{t-1} = s_t \le \theta\gamma_t^2, \;\; \theta = \begin{cases} \frac{F_*^2}{2}, & \text{in the case of SA} \\ \frac{M_*^2}{2} \le 2\mathcal{L}^2\mathcal{R}^2, & \text{in the case of SMP} \end{cases}$$

(see (133)). Thus, we can be sure that $S_t + R_t \le [S_{t-1} + R_{t-1}] + \theta\gamma_t^2$, meaning that when choosing

$$\gamma_t = \sqrt{[S_t^* - S_{t-1} - R_{t-1}]/\theta} \tag{136}$$

we guarantee the validity of $(*_t)$ and the inequality $\gamma_t \ge \sqrt{\delta_t/\theta}$. This observation combined with (132) and $(*_N)$ implies that

$$\begin{aligned} \forall N \ge 1: \quad \epsilon_{\text{sad}}(z^N) &\le \frac{\Theta^2/2 + R_N + S_N}{\sum_{\tau=1}^N \sqrt{\delta_\tau/\theta}} \le \frac{O(1)\Theta^2(1+\ln N)}{\sum_{\tau=1}^t \sqrt{\delta_\tau/\theta}} \\ &\le O(1)(1+\ln N) \cdot \begin{cases} \Theta F_* N^{-1/2}, & \text{in the case of SA,} \\ \Theta\mathcal{R}\mathcal{L}N^{-1/2}, & \text{in the case of SMP.} \end{cases} \end{aligned} \tag{137}$$

Observe that (137) is, within the logarithmic in N factor $O(1)(1+\ln N)$, the same as the bound (134). In fact, we could somehow reduce this logarithmic gap by modifying s_t^*, but we do not think this is necessary; we may hope (and the experiments to be reported in Section 4.5 fully support this hope) that "in reality" the rule (136) is much better than it is stated by the above worst-case analysis. The rationale behind this hope is that while we indeed are conservative when thinking how large could $S_t - S_{t-1}$ be, we account, to some extent, for the "past conservatism:" when $S_{t-1} + R_{t-1}$ is essentially less than S_{t-1}^*, γ_t as given by (136) is essentially larger than its lower bound used in the complexity analysis.

Finally, we remark that the major theoretical disadvantage of the efficiency estimate (137) as compared to (124) is much more serious than an extra log-factor. While with the basic implementation, in course of N steps the stochastic oracle is called $O(1)N$ times, the number of oracle calls in course of N steps of the advanced implementation is random and can be much larger than $O(1)N$; it is unclear why it should be $O(1)N$ even on average. Though for the time being we cannot support the empirical evidence by a solid theoretical complexity analysis, in our experiments the advanced implementation by far outperformed its basic counterpart.

4.3.4 The Favorable Geometry Case

We are about to present the "favorable geometry" case where we can point out the setup for SA/SMP which results in (nearly) *dimension-independent* efficiency estimates. Specifically, assume that

[G.1] The domain Z of $(\mathcal{S})$ is a *subset* of the direct product $Z^+ = B_1 \times .... \times B_{p+q}$ of $r = p + q$ "standard blocks" as follows:

- for $1 \le i \le p$, B_i is the unit Euclidean ball in $F_i = \mathbb{R}^{n_i}$;
- for $1 \le j \le q$, B_{p+j} is a subset of the space F_{p+j} of $n_{p+j} \times n_{p+j}$ ($n_{p+j} > 1$) symmetric block-diagonal matrices of a given block-diagonal structure and is the *spectahedron* of F_{p+j}, that is, the set of all positive semidefinite matrices from F_{p+j} with unit trace.

  In particular, B_{p+j} can be the standard simplex $\{x \in \mathbb{R}^k_+ : \sum_\ell x_\ell = 1\}$ (since the space of diagonal $k \times k$ matrices can be naturally identified with $\mathbb{R}^k$).

We equip $F_i = \mathbb{R}^{n_i}$, $i \le p$, with the standard Euclidean structure and the associated Euclidean norm $\|\cdot\|_{(i)}$, and F_{p+j} – with the Frobenius Euclidean structure and the trace-norm (the sum of singular values of a matrix) $\|\cdot\|_{(p+j)}$. In particular, the embedding space $E = F_1 \times ... \times F_r$ of Z^+ becomes equipped with the direct product of the indicated Euclidean structures. Note that the norm $\|\cdot\|_{(i,*)}$ conjugate to $\|\cdot\|_{(i)}$ is either the norm $\|\cdot\|_{(i)}$ itself (this is so when $i \le p$), or is the standard matrix norm (maximal singular value of a matrix) (this is so when $i > p$). We denote a vector form on E as $x = [x_1; ...; x_r]$, where x_ℓ is the F_ℓ-component of x.

G.2. The decomposition $Z = Z_1 \times Z_2 \subset E_1 \times E_2$ is compatible with the decomposition $Z = B_1 \times ... \times B_r$, that is, E_1 is the direct product of some of F_ℓ, $1 \le \ell \le p + q$, and E_2 is the direct product of the remaining F_ℓ. Besides this, we assume that Z intersects the relative interior of Z^+.

We refer to this case as to the one of *favorable geometry* and associate with this case the setup for SA and SMP as follows (cf. [101, Section 5]):

- The skew-symmetric linear mapping $\mathcal{A}$ (see (89)) can be written down as

$$\mathcal{A}[x_1;...;x_r] = [\sum_{j=1}^r A^{1j}x_j;...;\sum_{j=1}^r A^{rj}x_j],$$

where A^{ij} is a linear mapping from F_j to F_i and $[A^{ij}]^* = -A^{ji}$. We denote by L_{ij} an a priori upper bound on $L_{ij}^* := \max_{x_j}\left\{\|A^{ij}x_j\|_{(i,*)} : \|x_j\|_{(j)} \le 1\right\}$ such that $L_{ij} = L_{ji}$.[4]

- Further, we set

$$\omega_i(x_i) = \tfrac{1}{2}x_i^T x_i :\ B_i \to \mathbb{R},\ \Omega_i = \tfrac{1}{2},\ 1 \le i \le p$$
$$\omega_{p+j}(x_{p+j}) = 2\sum_{\ell=1}^{n_{p+j}} \lambda_\ell(x_{p+j})\ln(\lambda_\ell(x_{p+j})) : B_{p+j} \to \mathbb{R},\ \Omega_{p+j} = 2\ln(n_j),\ 1 \le j \le q$$

where $\lambda_\ell(u)$ are the eigenvalues of a symmetric matrix u taken with their multiplicities. It is known that $\omega_\ell(\cdot)$ is a d.g.f. for B_ℓ compatible with the norm $\|\cdot\|_\ell$, $1 \le \ell \le r$.

- Finally, we define the norm $\|\cdot\|$ on E and the d.g.f. $\omega(\cdot)$ for Z according to

$$\mu_\ell = \frac{1}{\Omega_\ell}\frac{\sum_{j=1}^r L_{\ell j}\sqrt{\Omega_\ell\Omega_j}}{\sum_{i,j=1}^r L_{ij}\sqrt{\Omega_i\Omega_j}},\ \|[x_1;...;x_r]\| = \sqrt{\sum_{\ell=1}^r \mu_\ell\|x_\ell\|_{(\ell)}^2},\ \omega(x) = \sum_{\ell=1}^r \mu_\ell\omega_\ell(x_\ell), \tag{138}$$

which results in

$$\Omega \le 1,\ \mathcal{R} \le \Theta \le \sqrt{2},\ \mathcal{L} = \sum_{i,j=1}^r L_{ij}\sqrt{\Omega_i\Omega_j}, \tag{139}$$

see [101, Section 5].

Remark 4.3.1 *From the results of [77] it follows that the norm* $\|\xi\|_* = \sqrt{\sum_{\ell=1}^r \mu_\ell^{-1}\|\xi_\ell\|_{(i,*)}^2}$ *is* $\varkappa$*-regular (see discussion in Section 4.3.2.4) with nearly dimension-independent* $\varkappa$*, namely,* $\varkappa = 3\max_{1\le j\le q}\ln(n_{p+j})$.

Note that the applications presented in Sections 4.2.2.2 and 4.2.2.3 are of favorable geometry; the same is true for the low dimension approximation problem of Section 4.2.1.3 after passing from the variable Q to the variable $R = d^{-1}Q$.

[4] The latter restriction is natural, since $L_{ij}^* = L_{ji}^*$ due to $[A^{ij}]^* = -A^{ji}$.

4.4 Solving the Generalized Bilinear Saddle Point Problem

Here we explain how a GBSP problem (95) – (96) can be reduced to a "small series" of BSP problems; the strategy to follow originates from [84]. From now on we assume, in addition to **A.1-2**, that we have an a priori upper bound $\bar{\rho}$ on the optimal value ρ_* of (96). For example, it is immediately seen that when finding an ϵ-solution to ℓ_1-minimization problem with ℓ_p-fit (Section 4.2.2.2) in the only nontrivial case $\|b\|_p > \delta$ relation (99) implies that

$$\bar{\rho} := \frac{\|A\|_{1\to p}}{\|b\|_p - \delta} \geq \rho_* := \frac{1}{\mathrm{Opt}}, \ \|A\|_{1\to p} = \max_j \|A_j\|_p, \tag{140}$$

where $A_1, ..., A_n$ are the columns of A. In particular, when finding an ϵ-solution to ℓ_1-minimization problem with the uniform fit in the only nontrivial case $\|b\|_\infty > \delta$ we have

$$\bar{\rho} := \frac{\|A\|_{1\to\infty}}{\|b\|_\infty - \delta} \geq \rho_* := \frac{1}{\mathrm{Opt}}, \ \|A\|_{1\to\infty} = \max_{i,j} |A_{ij}|; \tag{141}$$

.

For the sake of definiteness, we assume that we are in the Favorable Geometry case, and that the decomposition $Z = Z_{11} \times Z_{12} \times Z_2 \subset E$, see (97), is compatible with the decomposition $E = F_1 \times ... \times F_r$, that is, the embedding spaces of Z_{11}, Z_{12} and Z_2 are products of some of F_ℓ's. To save space, we restrict ourselves with the SMP algorithm; modifications in the case of SA are straightforward.

The algorithm solves the problem of interest (96) by applying to $\mathrm{SV}(\cdot)$ a Newton-type root finding routine, with (approximate) first order information on SV at a point ρ given by SMP as applied to the saddle point problem specifying $\mathrm{SV}(\rho)$. Specifically, the algorithm works stage by stage. At a stage s, we have at our disposal an upper bound ρ_s on ρ_* and a piecewise linear function $\ell_{s-1}(\rho)$ which underestimates $\mathrm{SV}(\cdot)$:

$$\mathrm{SV}(\rho) \geq \ell_{s-1}(\rho) \ \ \forall \rho \geq 0.$$

here $\rho_1 = \bar{\rho}$, $\ell_0 \equiv -\infty$. At a stage, we apply SMP to the BSP problem

$$\mathrm{SV}(\rho_s) = \min_{z_1 \in Z_1} \max_{z_2 \in Z_2} \phi^{\rho_s}(z_1, z_2) \tag{$\mathcal{S}_s$}$$

namely, act as follows.

A. We start stage s with building the setup for SMP as explained in Section 4.3.4. The affine operator associated with $(\mathcal{S}_s)$ is

$$\begin{array}{rcl} F^{\rho_s}(z_1=[z_{11};z_{12}],z_2) & = & \Phi(z_1,z_2)+\rho_s\Psi(z_1,z_2) \\ & = & [[a_{11}+B^*z_2;\rho_s(a_{12}+C^*z_2)]\,;-b-Bz_{11}-\rho_s(c+Cz_{12})]\,, \end{array}$$

see (95), (97). In matrix $\mathcal{A}=\mathcal{A}_s$ of the linear part of F^{ρ_s}, some blocks A^{ij} are independent of ρ_s, while the remaining blocks are proportional to ρ_s. Consequently, the Lipschitz constant of F^{ρ_s} as given by (139) is

$$\mathcal{L}=\mathcal{L}(\rho_s)=\mathcal{M}+\rho_s\mathcal{N},\ \mathcal{M},\mathcal{N}\geq 0. \tag{142}$$

An analogous decomposition holds for vector $a=a_s$:

$$\|a\|_*=\mu+\rho_s\nu,\ \mu,\nu\geq 0.$$

B. We apply to $(\mathcal{S}_s)$ either the basic, or the advanced implementation of the SMP. When running the basic SMP, we use the distributions $P_z^{(k)}$, see Section 4.3.2.4 (here $k\geq 1$ is a parameter of the construction) and use the online stepsize policy (125), where we set $\mathcal{L}=\mathcal{M}+\rho_s\mathcal{N}$ and $\Omega=1$ (see (139)). When $(\mathcal{S}_s)$ is solved by the advanced SMP, we use the online stepsize policy (135) – (136), with $\Theta=\sqrt{2}$ in (135).

B.1. Let $z^{ti}=[z_1^{ti},z_2^{ti}]$ be the approximate solution to $(\mathcal{S}_s)$ generated after t steps of stage s; recall that along with this solution, we have at our disposal the quantities

$$\begin{array}{rcl} \overline{\phi}^{ts} & = & \max_{z_2\in Z_2}\phi^{\rho_s}(z_1^{ts},z_2)=\upsilon+\langle a_{11},z_{11}^{ts}\rangle+\rho_s[\chi+\langle a_{12},z_{12}^{ts}\rangle] \\ & & \qquad +\min_{z_2\in Z_2}\langle z_2,b+\rho_s c+Bz_{11}^{ts}+\rho Cz_{12}^{ts}\rangle, \\ \underline{\phi}^{ts} & = & \min_{z_1\in Z_1}\phi^{\rho_s}(z_1,z_2^{ts})=\overbrace{\upsilon+\langle b,z_2^{ts}\rangle+\min_{z_{11}\in Z_{11}}\langle a_{11}+B^*z_2^{ts},z_{11}\rangle}^{p_{ts}} \\ & & \qquad +\rho_s\overbrace{\left[\kappa+\langle c,z_2^{ts}\rangle+\min_{z_{12}\in Z_{12}}\langle a_{12}+C^*z_2^{ts},z_{12}\rangle\right]}^{q_{ts}} \end{array} \tag{143}$$

(cf. (113) and see (95), (97)). We set

$$u^{ts}=\min_{\tau\leq t}\overline{\phi}^{\tau s},\ \ell^{ts}=\max_{\tau\leq t}\underline{\phi}^{\tau s},\ \ell_{ts}(\rho)=\max[\ell_{s-1}(\rho),\max_{1\leq\tau\leq t}[p_{\tau s}+q_{\tau s}\rho]].$$

Note that u^{ts} is a nonincreasing in t upper bound on $\mathrm{SV}(\rho_s)$, ℓ^{ts} is a nondecreasing in t lower bound on $\mathrm{SV}(\rho_s)$, and $\ell_{ts}(\rho)$ underestimates $\mathrm{SV}(\rho)$ for all $\rho\geq 0$. In addition, $\ell_{ts}(\rho_s)\geq\ell^{ts}$.

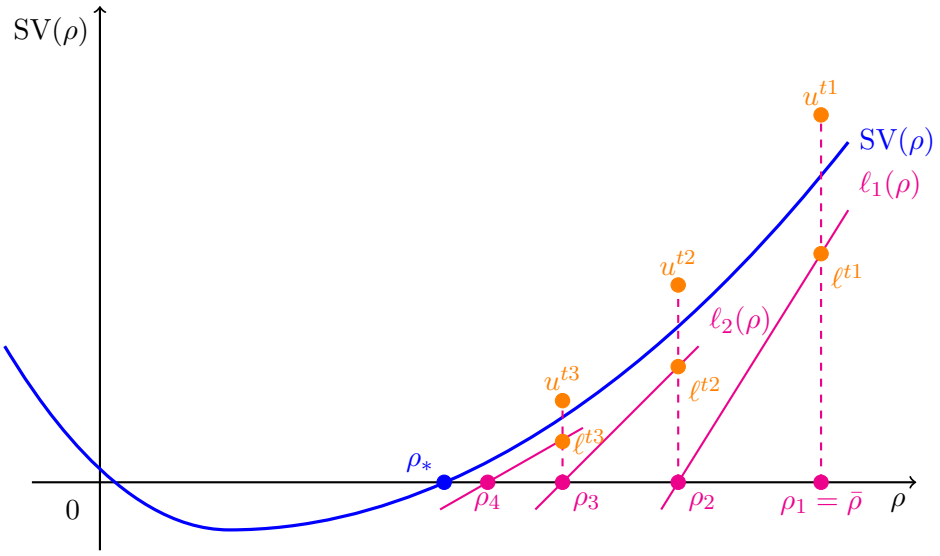

Figure 1: Illustration of the algorithm for solving GBSPP

Note also that after t steps we have at our disposal vectors $w_1^{ts} \in Z_1$, $w_2^{ts} \in Z_2$ such that

$$\max_{z_2 \in Z_2} \phi^{\rho_s}(w_1^{ts}, z_2) = u^{ts} \leq \overline{\phi}^{ts}, \min_{z_1 \in Z_1} \phi^{\rho_s}(z_1, w_2^{ts}) = \ell^{ts} \geq \underline{\phi}^{ts},$$

meaning that $w^{ts} = [w_1^{ts}; w_2^{ts}]$ is a feasible solution to $(\mathcal{S}_s)$ and $\epsilon_{\text{sad}}(w^{ts}) = u^{ts} - \ell^{ts} \leq \overline{\phi}^{ts} - \underline{\phi}^{ts} = \epsilon_{\text{sad}}(z^{ts})$.

B.2. We proceed with solving $(\mathcal{S}_s)$ until one of the following two situations occurs:

A) We get $u^{ts} \leq \epsilon\rho_s$. In this case we terminate with the claim that ρ_s, w_1^{ts} is the desired ϵ-solution to (95) – (96).

B) We get $\ell^{ts} \geq \frac{3}{4}u^{ts}$. When it happens, we set

$$\rho_{s+1} = \max\{\rho : \ell_{ts}(\rho) \leq 0\}, \ \ell_s(\cdot) \equiv \ell_{ts}(\cdot) \tag{144}$$

and pass to the stage $s+1$. An illustration of this algorithm is given in Figure 1.

Theorem 4.4.1 *When solving a Generalized Bilinear Saddle Point problem (95) – (96) by the outlined algorithm:*

(i) *The algorithm terminates in finite time with probability 1, and the resulting solution is an ϵ-solution, as defined in Section 4.2.2, to the GBSP problem in question;*

(ii) *The number of stages does not exceed the quantity $O(1)\ln\left(\frac{\|\phi\|_\infty + \bar{\rho}\|\psi\|_\infty}{\epsilon\rho_*} + 2\right)$, where $\|\phi\|_\infty = \max_{z \in Z} |\phi(z)|$, $\|\psi\|_\infty = \max_{z \in Z} |\psi(z)|$, see (95).*

(iii) *The (random) number N_s of steps at every stage s of the basic implementation satisfies for all $\varepsilon > 0$ the relation*

$$\mathrm{Prob}\{N_s \geq N(\epsilon)\} \leq e^{-\Lambda N(\epsilon)} + e^{\lambda},$$

where

$$N(\epsilon) = O(1)\mathrm{Ceil}\left[\frac{\mathcal{M}+\rho_*\mathcal{N}}{\epsilon\rho_*} + \frac{\varkappa_*(k,\Lambda)^2}{k}\left(\frac{\mathcal{M}+\rho_*\mathcal{N}}{\epsilon\rho_*}\right)^2 + \frac{\lambda}{k}\left(\frac{\mu+\rho_*\nu}{\epsilon\rho_*}\right)^2\right]. \tag{145}$$

The number of steps at every stage of the advanced implementation of the algorithm does not exceed

$$N_{\mathrm{adv}}(\epsilon) = O(1)\left[\frac{\mathcal{M}+\rho_*\mathcal{N}+2\epsilon\rho_*}{\epsilon\rho_*}\ln\left(\frac{\mathcal{M}+\rho_*\mathcal{N}+2\epsilon\rho_*}{\epsilon\rho_*}\right)\right]^2. \tag{146}$$

For proof, see Section 4.6.4.

In the case of ℓ_1-minimization problems with uniform- and ℓ_2-fits, Theorem 4.4.1 as applied to the basic implementation of SMP with $k = 1$, initialized according to (141), resp., (140), after completely straightforward computations implies the complexity bounds stated in Proposition 4.1.1. The preprocessing mentioned in item (ii) of Proposition is as follows: we choose an $m \times m$ orthogonal matrix U with moduli of entries not exceeding $O(1)/\sqrt{m}$ and such that multiplication of a vector by U takes $O(m \ln m)$ operations (e.g., U can be the matrix of the Cosine Transform). We then draw at random a ± 1 vector ξ from the uniform distribution on the vertices of the unit m-dimensional box and pass from the data $[A, b]$ to the data

$$[A' = U\mathrm{Diag}\{\xi\}A, \;\; b' = U\mathrm{Diag}\{\xi\}b],$$

thus obtaining an equivalent reformulation of the problem of interest. Note that this preprocessing costs $O(1)mn\ln(m)$ operations. We clearly have $\|A'\|_{1\to 2} = \|A\|_{1\to 2}$. Applying the Hoeffding inequality, it is immediately seen that with probability $\geq 1-\chi$ one has $\|A'\|_{1\to\infty} < O(1)\sqrt{\ln(mn/\chi)}m^{-1/2}\|A\|_{1\to 2}$, that is, $\Gamma(A') \leq O(1)\sqrt{\ln(mn/\chi)}$, as stated in Proposition 4.1.1.

4.5 Numerical Results

Below we report on a series of numerical experiments aimed at comparing the performances of the Stochastic Mirror Prox algorithm SMP (in its advanced implementation) and its prototype — Deterministic Mirror Prox algorithm (DMP) proposed in [101][5]. The algorithms were tested on the GBSP problems of ℓ_1-minimization with uniform and ℓ_2-fits, see Section 4.2.2.2. The MATLAB 7.10.0 implementation of the algorithms was executed on an eight-core machine with two quad-core Intel Xeon E5345 CPU@2.33GHz, 8 MB L2 cache per quad-core chip and 12GB FB-DIMM total RAM (the computations were running single-core and single-threaded).

Test problems we use are of the compressive sensing origin. Specifically, given the sizes m, n of a test problem, we picked at random an $m \times n$ matrix B with i.i.d. entries taking values ± 1 with probabilities 0.5, and a sparse (with Ceil$(\sqrt{m})$ nonzero entries) "true signal" x_* normalized to have $\|x_*\|_1 = 1$, thus giving rise to the test problem

$$\mathrm{Opt}_p = \min_x \{\|x\|_1 : \|Ax - y\|_p \le \delta\}, \; A = m^{-1/p}B, \; y = Ax_* + \xi \qquad (P_p)$$

where $p = \infty$ (uniform fit) or $p = 2$ (ℓ_2-fit). The "observation noise" ξ was chosen at random and then normalized to have $\|\xi\|_p = \delta$. Our goal is to solve (P_p) within accuracy ϵ, i.e., to find x_ϵ satisfying $\|x_\epsilon\|_1 \le \mathrm{Opt}_p$ and $\|Ax_\epsilon - y\|_p \le \delta + \epsilon$. In all our experiments, $\delta = 0.005$ and $\epsilon = 0.0025$ were used.

Implementation of the algorithms. The GBSP reformulations of problems (P_p) were solved by SMP (in advanced implementation) and DMP according to the scheme presented in Section 4.4. In the case $p = \infty$ of uniform fit, both SMP and DMP used the GBSP problem reformulation given by (102). In the case $p = 2$ of ℓ_2-fit, SMP used the GBSP reformulation (101), while DMP was applied to the GBSP problem stemming directly from (100) with $p = 2$, namely, given by

$$\phi^\rho(z_1, z_2) = z_2^T(AJ_n z_1 - \rho b) - \rho\delta, Z_1 = Z_{11} = \Delta_{2n}, \; Z_2 = \{\|z_2\|_2 \le 1\}. \qquad (147)$$

[5] DMP is nothing but SMP with precise information (i.e., P_z is the unit mass sitting at z) and on-line stepsize policy described in [101, Section 6].

The rationale here is that the GBSP given by (147) "by itself" is easier than the GBSP given by (101): an ϵ-solution to the latter problem induces straightforwardly an ϵ-solution to the former one, but not vice versa. As a compensation, the problem (101), in contrast to (147), is better suited for randomization[6]. The latter fact, which is crucial for SMP, is irrelevant for DMP, this is why we apply this algorithm to the GBSP given by (147). In order to make a fair comparison, when running SMP for ℓ_2-fit, we terminate the run based on the ℓ_2-residual of the solution.

When implementing SMP, we utilized the option, discussed in Section 4.3.2.4, of building an estimate $F(\zeta)$ of $F(z)$ by generating k samples $\zeta^\ell \sim P_z$, $\ell = 1, ..., k$, and setting $\zeta = \frac{1}{k}\sum_{\ell=1}^k \zeta^\ell$. The "multiplicity" k was set to 40 for small instances and 100 for large (those with at least 10^8 nonzeros in A) instances.

In our implementations, we have tested different policies for choosing the starting point at each stage and different choices of the distance generating function (d.g.f.) for the simplexes. Specifically, along with the entropy d.g.f. discussed in Section 4.3.4, we tested the power d.g.f. $\omega(x) = \frac{\mathrm{e}}{\kappa(1+\kappa)}\sum_{i=1}^n x_i^{1+\kappa} : \{x \in \mathbb{R}^n_+ : \sum_i x_i \le 1\} \to \mathbb{R}$, with $\kappa = \frac{1}{\ln(n)}$; the theoretical complexity bounds associated with this choice of d.-g.f. coincide, within absolute constant factors, with those for the entropy. The detailed results comparing the effects of these policies on the performance are provided in Tables 13 and 14 in Section 4.7. The best policies we ended up with are as follows:

— for SMP: entropy d.-g.f., restarts from the ω-center of Z ("C00E" implementation);

— for DMP, in the case of uniform fit: power d.-g.f., restarts from the convex combination of the best (with the smallest ϵ_{sad}) point found so far and the ω-center of Z, the weights being 0.75 and 0.25, respectively ("B75P" implementation);

— for DMP, in the case of ℓ_2-fit: power d.-g.f., restarts from the convex combination of the last search point of the previous stage and the ω-center of Z, the weights being 0.25 and

[6]Indeed, in the second problem all nontrivial matrix-vector multiplications required to compute $F^\rho(z)$ are multiplications of vectors from the ℓ_1-balls by A and A^T; since a vector from ℓ_1-ball is the expectation of an extremely sparse (just one nonzero entry) random vector taking values in the same ball, the required matrix-vector multiplications admit cheap randomized versions. In the first problem, some of the required matrix-vector multiplications involve vectors from the $\|\cdot\|_2$-ball, and such a vector typically cannot be represented as the expectation of a sparse random vector taking values in the ball.

0.75, respectively ("L25P" implementation).

The results, I. In order to avoid too time-consuming experimentation, we primarily dealt with "moderate size" test problems. These problems were split into four groups according to the total number of nonzeros in A ($2 \cdot 10^6$, $8 \cdot 10^6$, $32 \cdot 10^6$, $128 \cdot 10^6$). Every group was further split into two subgroups according to the ratio $n : m$ (8 and 2). For every one of the resulting pairs (m, n), we generated 5 instances of problem (P_2) and 5 instances of problem (P_∞) and solved them by DMP and SMP. Thus, the methods were compared on totally 70 problems split into 14 series of 5 experiments each, with common for all experiments of a series sizes m, n and the value of p. The results are presented in Tables 10 (uniform fit) and 11 (ℓ_2-fit). For every series of 5 experiments, we present the corresponding minimal, maximal and average values of several performance characteristics, specifically

- `CPU` — the CPU time (sec) of the entire computation
- `Calls` — the total number of computations of the values of F
- `FCalls` — the equivalent number of calls to the deterministic oracle for the randomized algorithm. This quantity is defined as follows. For DMP, computing a value of F at a point reduces to a pair of matrix-vector multiplications, one involving A and the other one involving A^T; the cost of this computation is $2mn$ operations. For SMP invoked with multiplicity k (see above), the computation of (an unbiased estimate of) $F(z)$ requires multiplying one vector with $\leq k$ nonzero entries by A, and another vector with $\leq k$ nonzero entries by A^T, the total cost of these two computations being $k(m+n)$ operations. Thus, the "deterministic equivalent" of the randomized computation of F used by SMP is $\frac{k(m+n)}{2mn}$. The quantity `FCalls` is the induced by this definition deterministic equivalent of all randomized computations of F in a run of the SMP.

The data in Tables 10, 11 and their summaries provided in Figures 2-7 suggest the following interpretations:

1. As the sizes of instances grow, the randomized algorithm eventually outperforms its deterministic counterpart in terms of the CPU time, and the corresponding "savings" grow with the size $m \times n$ of the instance, and for instances of a given size – grow as

the ratio n/m decreases. Both phenomena are quite natural: the larger is mn and the smaller is $n/m \geq 1$ for a given mn, the smaller is the deterministic equivalent $k\frac{m+n}{2mn}$ of a randomized computation of F.

2. Even for our "not too large" test problems, the savings stemming from randomization can be quite significant: for the 8000×16000 instances, SMP is, at average, nearly 4.6 times faster than the best version of DMP for problems with uniform fit and 2.1 times faster than DMP for problems with ℓ_2-fit.

   When interpreting the CPU time data one should keep in mind that oracle calls of DMP make use of very efficient MATLAB implementation of matrix-vector multiplication, while SMP relies upon much less efficient (with respect to, e.g., C language) implementation of long DO loops.

3. The advantages, if any, of SMP as compared to DMP are more significant in the case of uniform fit than in the case of ℓ_2-fit. This phenomenon is quite natural: as we have already explained, in the case of ℓ_2-fit the methods are applied to different GBSP reformulations of (P_2), and the reformulation DMP works with is easier than the one processed by SMP.

The results, II. In order to get impression of what happens when the matrix A in (P_p) is too large to be stored in RAM, we carried out two experiments where the goal was to solve the ℓ_1-minimization problem with uniform and with ℓ_2 fits and fully dense $(m = 32000) \times (n = 64000)$ matrix A given by a simple analytical expression. This expression allows to compute a column/a row of A with a given index in $O(m)$, respectively, $O(n)$ operations. Matrix $A = A_p$ was normalized to have $\|A\|_{1\to p} = 1$. While the sizes of A make it impossible to store the matrix in the RAM of the computer we used for the experiments, we still can multiply vectors by A and A^T by computing all necessary columns and rows, and thus can run DMP and SMP. In our related experiments, we generated at random a sparse (64 nonzeros) "true" signal $x_* \in \mathbb{R}^{64000}$ with $\|x_*\|_1 = 1$, computed $y = Ax + \xi$, ξ, $\|\xi\|_p = \delta = 0.005$, being observation noise, and ran DMP and SMP in order to find

an ϵ-solution x_ϵ, $\epsilon = 0.0025$, to the resulting problem (P_p); in particular, we should have $\|x_\epsilon\|_1 \leq \|x_*\|_1 = 1$ and $\|Ax_\epsilon - b\| \leq \delta + \epsilon = 0.0075$. In every experiment, each of the methods was allowed to run at most 7,200 sec. The results are as follows.

- In the allowed 7,200 sec, the deterministic algorithms on every one of the two test problems ($p = 2$ and $p = \infty$) was able to carry out just about 30 steps with the total of about 67 computations of $F(\cdot)$; this is by far not enough to get meaningful results, see Table 12. In contrast to this, the numbers of steps and randomized computations of F carried out by the randomized algorithm in the same 7,200 sec was in the range of tens of thousands, which was enough to fully achieve the required accuracy for both $p = \infty$ and $p = 2$.

- While the quality of approximation of x_* by the solution yielded by DMP is basically non-existing, the SMP produced fairy reasonable approximations of x_*, see Table 12 and Figure 8.

In our opinion, the preliminary numerical results we have reported suggest that "acceleration via randomization" possesses a significant practical potential when solving extremely large-scale convex programs of appropriate structure.

4.6 Proofs of Chapter 4

4.6.1 Low Dimensional Approximation

We use the notations of Section 4.2.1.3.

Lemma 4.6.1 *Let Q_* be an optimal solution to (92), $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ be the eigenvalues of Q_*, $e_1, ..., e_n$ be the corresponding eigenvectors of Q_*, and $E = \mathrm{Lin}(e_1, ..., e_d)$. Then for any $v \in V$, $\mathrm{dist}(v, E) \leq \delta_* \sqrt{d+1}$ (here $\mathrm{dist}(x, E)$ stands for the Euclidean distance from v to E).*

[7] Percents given in the table represent $\|\widehat{x} - x_*\|/\|x_*\|$ for the corresponding norms $\|\cdot\|$.

Table 10: Numerical Results for ℓ_1-minimization with $\|\cdot\|_\infty$-fit

Sizes		DMP Calls	DMP CPU	SMP Calls	SMP FCalls	SMP CPU	$\frac{\text{Calls,DMP}}{\text{FCalls,SMP}}$	$\frac{\text{CPU,DMP}}{\text{CPU,SMP}}$
500 x 4000	Mean (C00E)	2661.6	106.6	10511.0	236.5	57.2	11.89	1.98
	Min (C00E)	1683.0	50.0	8159.0	183.6	34.0	6.91	1.16
	Max (C00E)	4395.0	179.4	11783.0	265.1	83.4	23.94	4.14
	Mean (B25P)	1453.4	104.1				6.15	1.89
1000 x 2000	Mean (C00E)	1830.8	64.0	10568.8	158.5	42.9	11.69	1.54
	Min (C00E)	1344.0	41.0	8434.0	126.5	28.8	7.82	1.02
	Max (C00E)	2507.0	91.5	11576.0	173.6	70.4	15.83	2.02
	Mean (B25P)	1530.6	97.9				9.64	2.48
1000 x 8000	Mean (C00E)	2338.0	227.9	12406.6	139.6	113.2	16.68	1.99
	Min (C00E)	1453.0	119.4	11579.0	130.3	88.2	11.15	1.27
	Max (C00E)	2739.0	370.2	13895.0	156.3	168.9	18.99	2.39
	Mean (B25P)	1545.6	248.9				11.08	2.30
2000 x 8000	Mean (C00E)	2691.6	227.6	12922.8	96.9	74.5	27.93	3.10
	Min (C00E)	1132.0	97.7	10934.0	82.0	56.6	12.24	1.37
	Max (C00E)	3355.0	313.1	15632.0	117.2	88.8	35.46	4.25
	Mean (B25P)	1426.4	207.8				14.74	2.84
2000 x 16000	Mean (C00E)	2384.6	494.2	13174.8	74.1	184.9	32.30	2.68
	Min (C00E)	2288.0	486.3	11735.0	66.0	174.4	29.78	2.53
	Max (C00E)	2491.0	505.5	14729.0	82.9	195.3	34.66	2.84
	Mean (B25P)	1575.2	533.7				21.41	2.89
4000 x 8000	Mean (C00E)	2923.6	798.7	19750.2	74.1	228.4	39.42	3.30
	Min (C00E)	2032.0	407.6	17262.0	64.7	159.0	28.86	2.34
	Max (C00E)	3895.0	1539.7	22945.0	86.0	343.1	48.61	4.49
	Mean (B25P)	1554.6	576.2				21.12	2.63
4000 x 32000	Mean (C00E)	2482.8	2054.3	11973.2	84.2	515.8	29.47	3.98
	Min (C00E)	1826.0	1448.9	11331.0	79.7	499.9	22.39	2.90
	Max (C00E)	3479.0	2904.2	12715.0	89.4	525.0	42.65	5.70
	Mean (B25P)	1604.8	1736.3				19.19	3.36
8000 x 16000	Mean (C00E)	2680.4	2227.7	12474.6	58.5	375.0	45.78	5.92
	Min (C00E)	2297.0	1890.1	11493.0	53.9	341.9	41.12	5.44
	Max (C00E)	3177.0	2609.0	13759.0	64.5	408.8	49.26	6.48
	Mean (B25P)	1615.8	1752.7	12474.6	58.5	375.0	27.57	4.63

Table 11: Numerical Results for ℓ_1-minimization with $\|\cdot\|_2$-fit

Sizes		DMP Calls	DMP CPU	SMP Calls	SMP FCalls	SMP CPU	$\frac{\text{Calls,DMP}}{\text{FCalls,SMP}}$	$\frac{\text{CPU,DMP}}{\text{CPU,SMP}}$
500 x 4000	Mean (C00E)	579.8	21.0	4771.6	106.7	24.6	5.91	0.93
	Min (C00E)	410.0	14.5	3412.0	76.3	16.9	3.18	0.49
	Max (C00E)	722.0	40.3	6868.0	153.5	36.0	8.40	1.94
	Mean (L75P)	287.8	16.1				2.95	0.70
1000 x 2000	Mean (C00E)	553.0	19.0	3910.8	54.8	13.6	10.73	1.47
	Min (C00E)	463.0	9.1	3315.0	46.4	11.5	5.68	0.52
	Max (C00E)	664.0	30.1	5890.0	82.5	17.4	13.56	2.34
	Mean (L75P)	282.4	14.1				5.44	1.07
1000 x 8000	Mean (C00E)	617.0	56.6	5148.8	57.5	50.7	11.25	1.17
	Min (C00E)	486.0	34.7	3745.0	41.9	36.1	7.68	0.74
	Max (C00E)	794.0	87.1	6050.0	67.6	64.8	18.35	1.93
	Mean (L75P)	318.8	40.9				5.84	0.86
2000 x 8000	Mean (C00E)	634.8	39.8	5853.6	41.0	47.2	15.94	0.86
	Min (C00E)	487.0	30.0	3926.0	27.5	33.1	11.17	0.59
	Max (C00E)	796.0	51.0	6869.0	48.1	54.0	20.49	1.12
	Mean (L75P)	318.8	25.9				8.05	0.58
2000 x 16000	Mean (C00E)	531.8	150.7	5055.6	28.3	90.0	19.88	1.80
	Min (C00E)	438.0	108.3	3947.0	22.1	60.2	11.64	0.87
	Max (C00E)	608.0	180.3	6736.0	37.6	125.1	24.80	2.49
	Mean (L75P)	346.0	110.6				12.74	1.28
4000 x 8000	Mean (C00E)	675.2	138.5	6504.6	22.8	101.7	29.71	1.36
	Min (C00E)	531.0	99.1	5868.0	20.5	83.3	22.71	0.99
	Max (C00E)	810.0	193.6	7143.0	25.0	113.9	34.52	1.70
	Mean (L75P)	346.4	86.3				15.21	0.85
4000 x 32000	Mean (C00E)	672.2	486.0	5613.4	39.2	287.2	17.66	1.74
	Min (C00E)	506.0	382.5	3418.0	23.9	197.2	12.08	1.26
	Max (C00E)	817.0	579.1	6611.0	46.2	336.4	22.57	2.15
	Mean (L75P)	355.4	311.6				9.39	1.12
8000 x 16000	Mean (C00E)	592.4	591.4	5815.0	25.4	177.6	24.15	3.51
	Min (C00E)	509.0	472.4	3765.0	16.5	117.3	16.56	2.36
	Max (C00E)	696.0	798.1	7038.0	30.8	214.1	30.90	5.06
	Mean (L75P)	329.8	360.2				13.38	2.10

Table 12: Experiments with dense $32,000 \times 64,000$ matrices A

Method	p	Steps	Calls	FCalls	CPU (sec)	$\|A\widehat{x} - b\|_p$	$\|\widehat{x} - x_*\|_1$[7]	$\|\widehat{x} - x_*\|_2$	$\|\widehat{x} - x_*\|_\infty$
DMP (C00E)	∞	30	71	71	7564	0.16018	1.406 (141%)	0.143 (89%)	0.041 (79%)
DMP (B25P)	∞	31	67	67	7363	0.15975	1.361 (136%)	0.136 (85%)	0.035 (69%)
SMP (C00E)	∞	7501	22141	25.9	5352	0.00744	0.048 (5%)	0.005 (3%)	0.002 (4%)
DMP (C00E)	2	29	67	67	7471	0.03653	1.455 (146%)	0.135 (84%)	0.035 (68%)
DMP (L75P)	2	30	67	67	7536	0.02480	0.976 (98%)	0.093 (58%)	0.022 (42%)
SMP (C00E)	2	2602	7749	8.5	2350	0.00715	0.264 (26%)	0.021 (13%)	0.004 (7%)

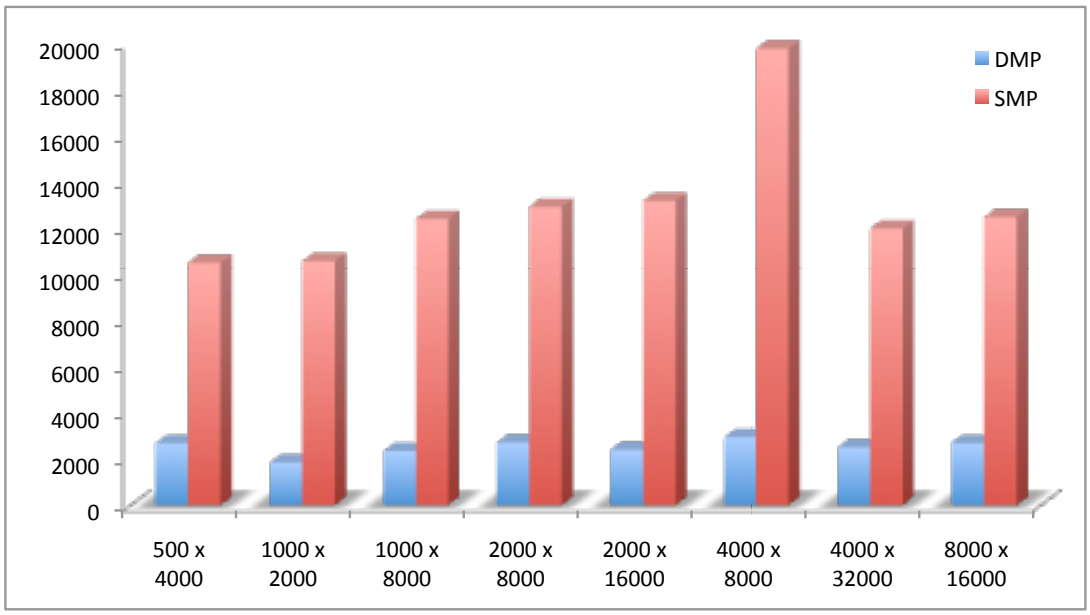

Figure 2: Number of oracle calls comparison for ℓ_∞-fit

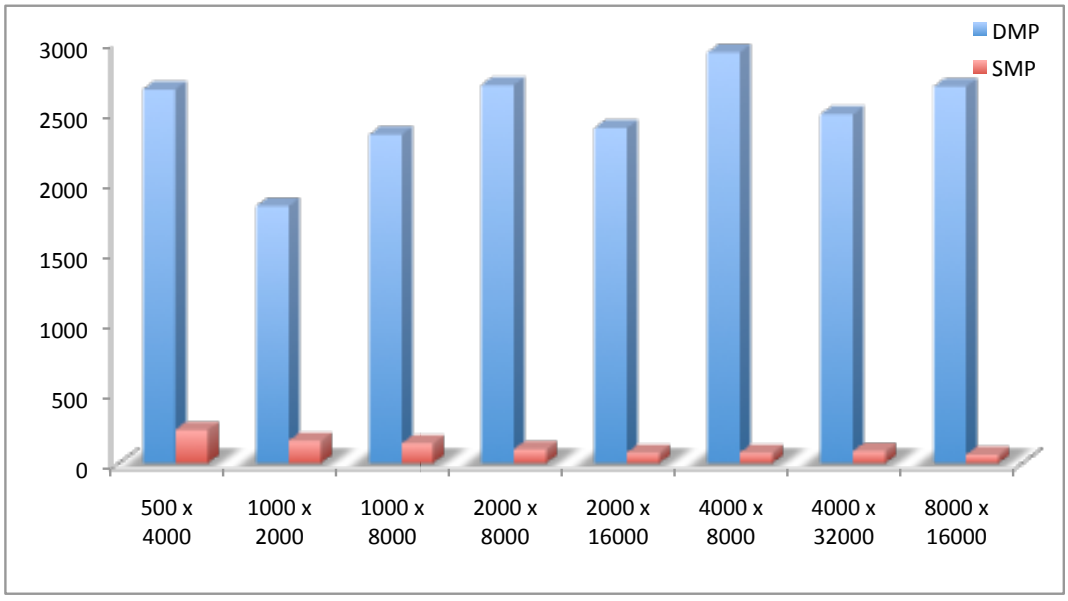

Figure 3: Number of equivalent oracle calls comparison for ℓ_∞-fit

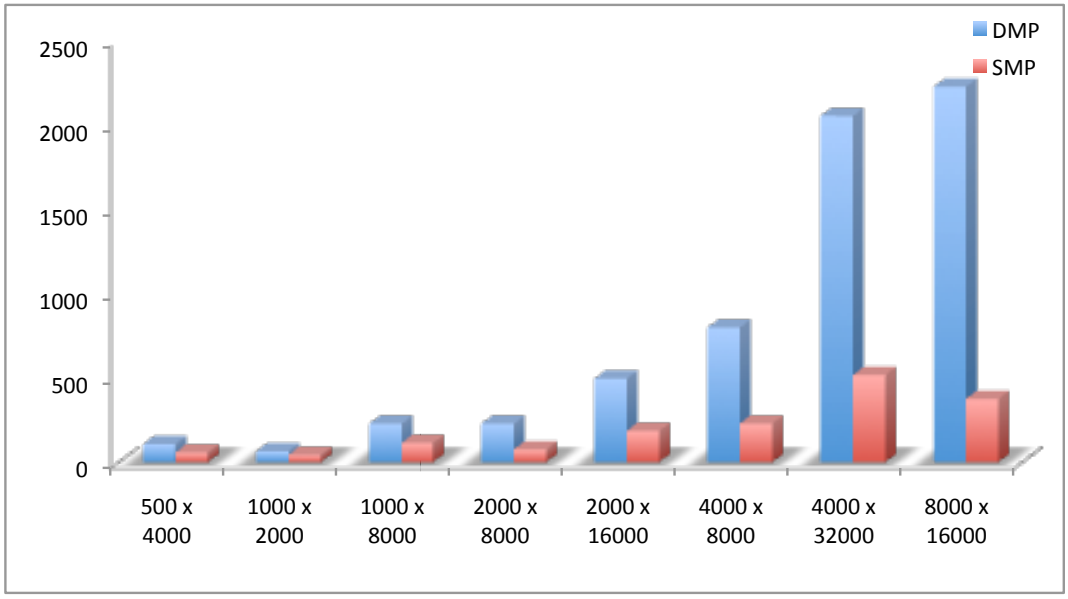

Figure 4: CPU time comparison for ℓ_∞-fit

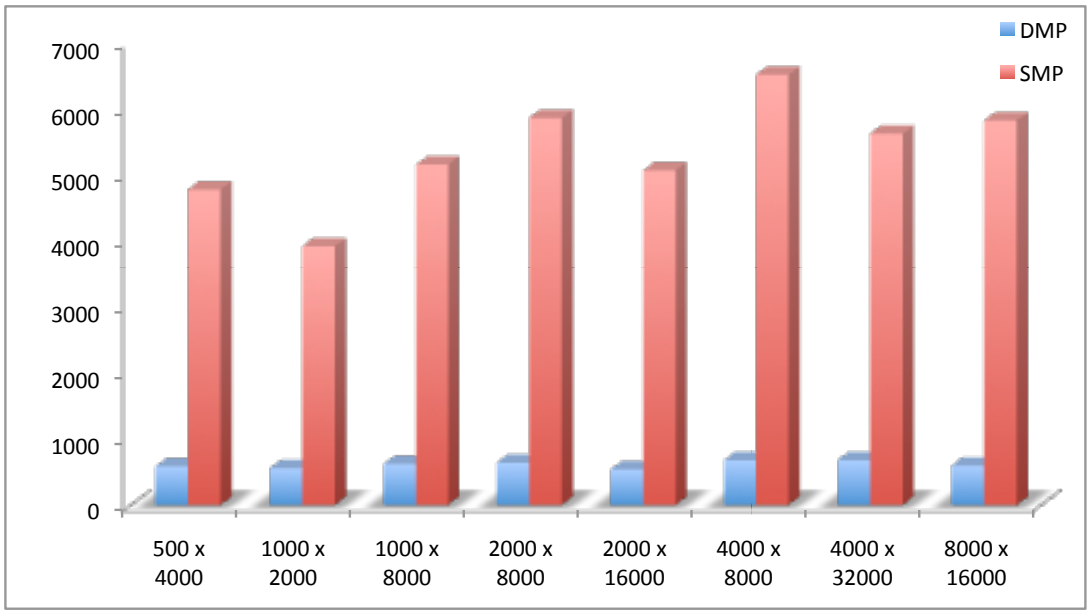

Figure 5: Number of oracle calls comparison for ℓ_2-fit

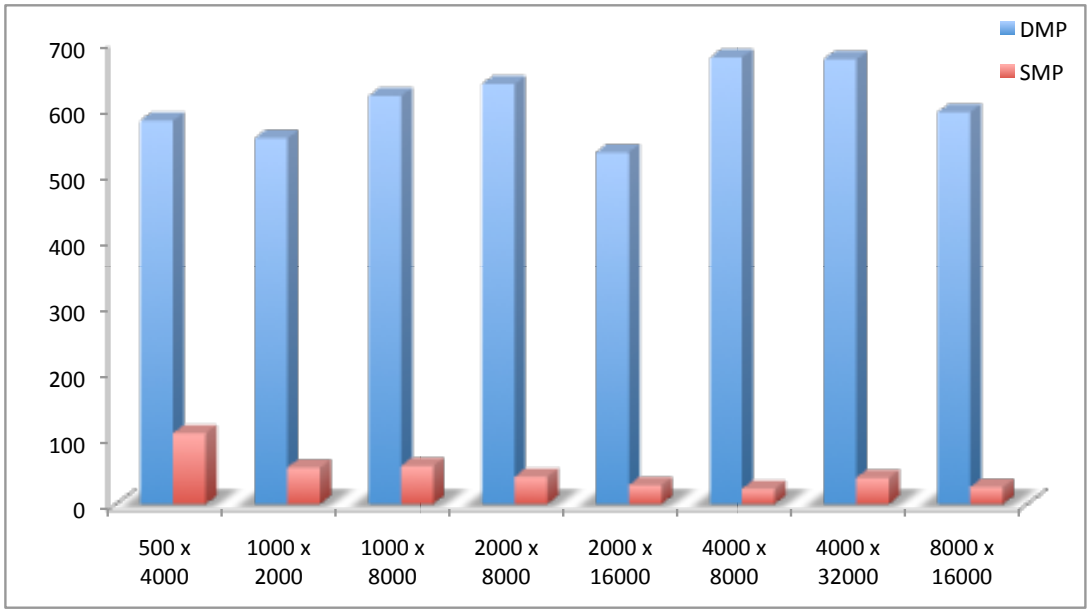

Figure 6: Number of equivalent oracle calls comparison for ℓ_2-fit

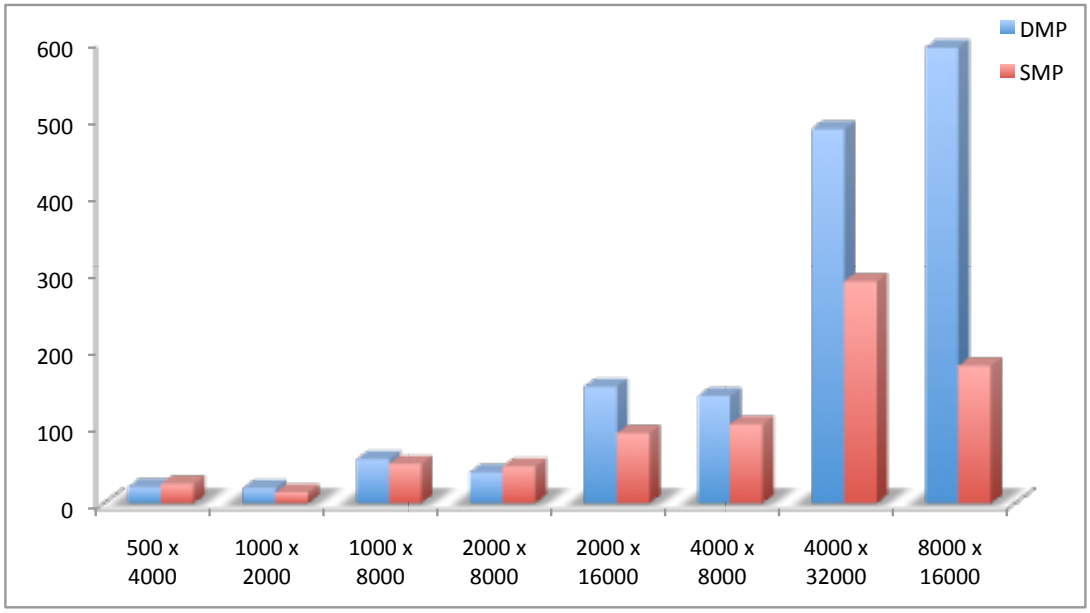

Figure 7: CPU time comparison for ℓ_2-fit

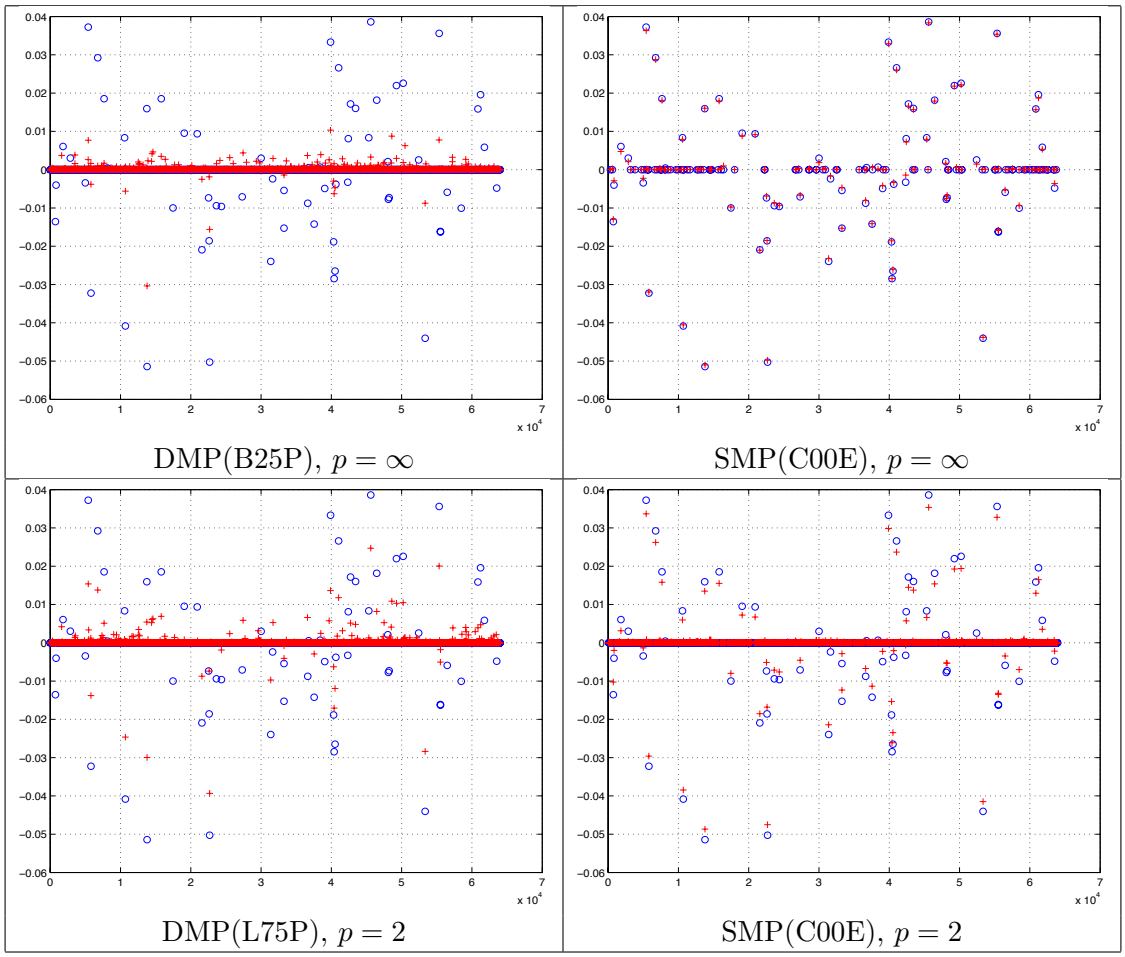

Figure 8: DMP-based (left) and SMP-based (right) recovery of sparse signals in the $32,000 \times 64,000$ experiment, entries vs. their indexes. Circles: x_*; crosses: recovery.

Proof. Note that $0 \le \lambda_j \le 1$ and $\sum_j \lambda_j = d$, so that $\lambda_j \le \frac{d}{d+1}$ when $j \ge d+1$. Denoting by $y_j(x)$ the coordinates of x in the eigenbasis $\{e_j\}$, we have

$$\begin{aligned}
&\mathrm{dist}^2(v,E) = 1 - \textstyle\sum_{j=1}^d y_j^2(v) \le 1 - \sum_{j=1}^d \lambda_j y_j^2(v) \text{ [since } \lambda_j \le 1]\\
&\le 1 - \textstyle\sum_{j=1}^n \lambda_j y_j^2(v) + \lambda_{d+1} \sum_{j=d+1}^n y_j^2(v) = 1 - v^T Q v + \lambda_{d+1}\mathrm{dist}^2(v,E)\\
&\Rightarrow \quad (1-\lambda_{d+1})\mathrm{dist}^2(v,E) \le 1 - v^T Q v \le 1 - \mathrm{Opt} \le 1 - \mathrm{Opt}_* = \delta_*^2\\
&\Rightarrow \quad \mathrm{dist}^2(v,E) \le (d+1)\delta_*^2 \text{ [since } \lambda_{d+1} \le \tfrac{d}{d+1}],
\end{aligned}$$

as claimed. □

Representing a vector from $\Delta_{n,d}$ as a convex combination of extreme points. The case of $d = n$ is trivial, thus, let $d < n$. Let

$$q \in \Delta_{n,d} = \left\{ q \in \mathbb{R}^n_+ : \ 0 \le q_i \le 1 \, \forall i, \ \sum_{i=1}^n q_i = d \right\}.$$

To represent q as a convex combination of n extreme points of $\Delta_{n,d}$ we act as follows:

- *Initialization:* We set $p^0 = [1; q]$, $\mu^0 = 1$. Note that $p^0 \in \Delta = \{p = [1; p_1; ...; p_n] \in \Delta_{n+1,d+1}\}$.

- *Step $t = 1, 2, ...$:* Given $p^{t-1} = [1; p_1^{t-1}; ...; p_n^{t-1}] \in \Delta$, we find the $d+1$ largest among the entries p_i^{t-1}, $i = 1, ..., n$, let their indexes be $i_1, ..., i_{d+1}$, where $p_{i_1}^{t-1} \ge p_{i_2}^{t-1} \ge ... \ge p_{i_{d+1}}^{t-1}$.

  a) It may happen that $p_{i_\ell}^{t-1} = 1$ for $1 \le \ell \le d$; since $p^{t-1} \in \Delta$, $r^t := p^{t-1}$ is a Boolean vector with exactly $d+1$ entries equal to 1, and $q^t = [p_1^{t-1}; ...; p_n^{t-1}]$ is an extreme point of $\Delta_{n,d}$. We set $\nu_t = 1$, $p^t = 0$ and terminate.

  b) When not all $p_{i_\ell}^{t-1}$, $1 \le \ell \le d$, are equal to 1, we set $\nu_t = \min[1 - p_{i_{d+1}}^{t-1}, p_{i_d}^{t-1}]$, define r^t as Boolean $(n+1)$-dimensional vector with $d+1$ entries equal to 1, the indexes of the entries being $0, i_1, ..., i_d$, set $p^t = [p^{t-1} - \nu_t r^t]/(1-\nu_t)$, $q^t = [r_1^t; ...; r_n^t]$ (note that q^t is an extreme point of $\Delta_{n,d}$) and pass to step $t+1$.

Observe that the algorithm is well defined. Indeed, $0 \le \nu_t \le 1$ by construction, and $\nu_t = 1$ if and only if $p_{i_{d+1}}^{t-1} = 0$ and $p_{i_d}^{t-1} = 1$, that is, when we terminate at step t according to a). Thus, p^t is well defined at every non-termination step t. Moreover, from b) it is immediately

seen that at such a step we have $p^t_0 = 1$, $0 \le p^t_i \le 1$ for all i and $\sum_{i=0}^n p^T_i = d+1$, that is, $p^t \in \Delta$ for all t for which p^t is well defined. Beside this, it is immediately seen that those entries in p^{t-1} which are zeros and ones remain zeros and ones in p^t as well, and that the total number of these entries increases at every step of the algorithm by at least 1. The latter observation implies that the algorithm terminates in at most n steps. Finally, by construction $p^{t-1} = (1-\nu_t)p^t + \nu_t r^t$, whence, denoting by $\bar{t}$ the termination step, p^0 is a convex combination of $r^1, ..., r^{\bar{t}}$ with coefficients μ_t readily given by $\nu_1, ..., \nu_{\bar{t}}$. Discarding in $r^1, ..., r^{\bar{t}}$ the entries with index 0, we get extreme points $q^1, ..., q^{\bar{t}}$ of $\Delta_{n,d}$ such that $q = \sum_{t=1}^{\bar{t}} \mu_t q^t$. Finally, the computational effort per step clearly does not exceed $O(1)dn$, that is, the total computational effort is at most $O(1)dn^2$.

4.6.2 Proof of Lemma 4.2.1

We have

$$\begin{aligned}
\mathrm{SV}(\rho) &= \max_{z_2\in Z_2}\min_{z_1\in Z_1}\phi^\rho(z_1,z_2)\\
&= \max_{z_2\in Z_2}\min_{z_{11}\in Z_{11},z_{12}\in Z_{12}}\Big[\upsilon+\rho\chi+\langle a_{11},z_{11}\rangle+\langle b,z_2\rangle+\langle z_2,Bz_{11}\rangle\\
&\qquad\qquad +\rho\left[\langle a_{12},z_{12}\rangle+\langle c,z_2\rangle+\langle z_2,Cz_{12}\rangle\right]\Big]\\
&= \max_{z_2\in Z_2}\Big[\upsilon+\rho\chi+\langle b,z_2\rangle+\rho\langle c,z_2\rangle+\min_{z_{11}\in Z_{11}}\big[\langle a_{11},z_{11}\rangle+\langle z_2,Bz_{11}\rangle\\
&\qquad\qquad +\rho\overbrace{\min_{z_{12}\in Z_{12}}\left[\langle a_{12}+C^*z_2,z_{12}\rangle\right]}^{g(z_2)}\big]\Big]\\
&= \max_{z_2\in Z_2}\big[\upsilon+\rho\chi+\langle b,z_2\rangle+\rho\langle c,z_2\rangle+\rho g(z_2)+\overbrace{\min_{z_{11}\in Z_{11}}\left[\langle a_{11},z_{11}\rangle+\langle z_2,Bz_{11}\rangle\right]}^{h(z_2)}\big]\\
&= \max_{z_2\in Z_2}\left[\upsilon+\langle b,z_2\rangle+h(z_2)+\rho\left[\chi+\langle c,z_2\rangle+g(z_2)\right]\right]
\end{aligned}$$

and thus $\mathrm{SV}(\rho)$ is the supremum of affine functions of ρ. □

4.6.3 Proofs for Section 4.3

We start with the following

Lemma 4.6.2 [cf. [101], Lemma 3.1.(*b*)] *Given $z \in Z^o$, $\gamma > 0$ and $\xi, \eta \in E$, let us set*

$$\begin{aligned}
w &= \mathrm{Prox}_z(\gamma\xi) = \mathrm{argmin}_{v\in Z}\left\{\langle\gamma\xi-\omega'(z),v\rangle+\omega(v)\right\},\\
z_+ &= \mathrm{Prox}_z(\gamma\eta) = \mathrm{argmin}_{v\in Z}\left\{\langle\gamma\eta-\omega'(z),v\rangle+\omega(v)\right\}.
\end{aligned}$$

Then $w, z_+ \in Z^o$, and for every $u \in Z$ one has

$$\begin{array}{lll}(a) & \gamma\langle\eta, w-u\rangle \le V_z(u) - V_{z_+}(u) + \gamma\langle\eta, w-z_+\rangle - V_z(z^+) \\ (b) & \le V_z(u) - V_{z_+}(u) + \gamma\langle\eta-\xi, w-z_+\rangle - V_z(w) - V_w(z_+) \\ (c) & \le V_z(u) - V_{z_+}(u) + \gamma\|\eta-\xi\|_*\|w-z_+\| - \frac{1}{2}\left[\|w-z\|^2 + \|w-z_+\|^2\right] \\ (d) & \le V_z(u) - V_{z_+}(u) + \frac{1}{2}\left[\gamma^2\|\eta-\xi\|_*^2 - \|w-z\|^2\right].\end{array} \tag{148}$$

Proof. The inclusions $w, z_+ \in Z^o$ are evident (a subgradient of $\omega(\cdot)$ at w, taken w.r.t. Z, is, e.g., $\omega'(z) - \gamma\xi$, and similarly for z_+). Now let $u \in Z$. z_+ is an optimal solution of certain explicit convex optimization problem; taking into account that $\omega'(\cdot)$ is continuous on Z^o, it is easily seen that the necessary optimality condition in this problem reads $\langle\gamma\eta + \omega'(z_+) - \omega'(z), u - z_+\rangle \ge 0$, whence $\gamma\langle\eta, w-u\rangle \le \gamma\langle\eta, w-z_+\rangle + \langle\omega'(z_+) - \omega'(z), u-z_+\rangle$, and the latter inequality, after rearranging terms in the right hand side, becomes (a). By similar reasons, $0 \le \langle\gamma\xi + \omega'(w) - \omega'(z), v-w\rangle$ for all $v \in Z$; setting $v = z_+$, summing up the resulting inequality with (a) and rearranging terms in the right hand side of what we get, we arrive at (b). (c) follows from (b) due to $V_a(b) \ge \frac{1}{2}\|a-b\|^2$ (recall that ω is strongly convex, modulus 1 w.r.t. $\|\cdot\|$, on Z). Finally, (d) follows from (c) due to $\mu\nu - \frac{1}{2}\mu^2 \le \frac{1}{2}\nu^2$. □

4.6.3.1 Proof of Proposition 4.3.1

Let us prove the bound (114). Consider first the case of SMP. Applying Lemma 4.6.2 to $z = z_\tau$, $\gamma = \gamma_\tau$, $\xi = F(\eta_\tau)$, $\eta = F(\zeta_\tau)$, which results in $w = w_\tau$ and $z_+ = z_{\tau+1}$, we get for all $u \in Z$:

$$\gamma_\tau\langle F(\zeta_\tau), w_\tau - u\rangle \le V_{z_\tau}(u) - V_{z_{\tau+1}}(u) + [\gamma_\tau\langle F(\zeta_\tau), w_\tau - z_{\tau+1}\rangle - V_{z_\tau}(z_{\tau+1})]$$

whence for all $u \in Z$

$$\begin{array}{rcl}\langle F(\zeta_\tau), \zeta_\tau - u\rangle & \le & \gamma_\tau^{-1}(V_{z_\tau}(u) - V_{z_{\tau+1}}(u)) + r_\tau + s_\tau, \\ s_\tau & = & \langle F(\zeta_\tau), w_\tau - z_{\tau+1}\rangle - \gamma_\tau^{-1}V_{z_\tau}(z_{\tau+1}) \\ & \le & \frac{1}{2}\left[\gamma_\tau\|F(\zeta_\tau) - F(\eta_\tau)\|_*^2 - \gamma_\tau^{-1}\|w_\tau - z_\tau\|^2\right], \quad (*) \\ r_\tau & = & \langle F(\zeta_\tau), \zeta_\tau - w_\tau\rangle.\end{array} \tag{149}$$

with $(*)$ given by (148). When summing up inequalities (149) over τ and taking into account that $\gamma_1 \geq \gamma_2 \geq ...$, $V_z(u) \geq 0$ and $V_{z_1}(u) = V_{z_\omega}(u) \leq \Omega$ by definition of Ω, we get

$$\sum_{\tau=1}^{t} \langle F(\zeta_\tau), \zeta_\tau - u \rangle \leq \gamma_t^{-1}\Omega + \sum_{\tau=1}^{t} [s_\tau + r_\tau]. \tag{150}$$

On the other hand,

$$\begin{aligned}
\sum_{\tau=1}^{t} \langle F(\zeta_\tau), \zeta_\tau - u \rangle &= \sum_{\tau=1}^{t} \langle a + \mathcal{A}\zeta_\tau, \zeta_\tau - u \rangle \\
&= t\langle a, z^t - u \rangle - \sum_{\tau=1}^{t} \langle \mathcal{A}\zeta_\tau, u \rangle \quad [\mathcal{A} \text{ is skew symmetric}] \\
&= t\left[\langle a, z^t - u \rangle - \langle \mathcal{A}z^t, u \rangle\right] = t\left[\langle a, z^t - u \rangle + \langle \mathcal{A}z^t, z^t - u \rangle\right] \\
&= t\langle F(z^t), z^t - u \rangle.
\end{aligned}$$

Thus, for all $u \in Z$ it holds

$$t\langle F(z^t), z^t - u \rangle \leq \Omega\gamma_t^{-1} + \sum_{\tau=1}^{t} [s_\tau + r_\tau] = \gamma_t^{-1}\Omega + S_t + R_t. \tag{151}$$

Setting $z^t = [z_1^t; z_2^t]$ and $u = [u_1; u_2]$, we get from (89) $\langle F(z^t), z^t - u \rangle = \phi(z_1^t, u_2) - \phi(u_1, z_2^t)$; the supremum of the latter quantity over $u \in Z$ is nothing that the saddle point residual $\epsilon_{\text{sad}}(z^t)$. Since the right hand side in (151) is independent of u, we arrive at the SMP-version of (114).

Now consider the case of SA. Applying Lemma 4.6.2 to $\gamma = \gamma_\tau$, $z = z_\tau$, $\xi = 0$, $\eta = F(\zeta_\tau)$, which results in $w = z_\tau$ and $z_+ = z_{\tau+1}$, and acting exactly as in the case of SMP, we arrive at the SA-version of (114).

Let us prove (ii). The conditional to the "past" (the answers of the oracle prior to the call for $\xi_{2\tau}$) distribution of ζ_τ is P_{w_τ}, which combines with the affinity of F and the facts that the linear part of F is skew symmetric and the expectation of P_z is z, to imply that

$$\begin{aligned}
\mathbf{E}\{\langle F(\zeta_\tau), \zeta_\tau - w_\tau \rangle\} &= \langle a, \mathbf{E}\{\zeta_\tau\} - w_\tau \rangle + \mathbf{E}\{\langle \mathcal{A}\zeta_\tau, \zeta_\tau - w_\tau \rangle\} = -\mathbf{E}\{\langle \mathcal{A}\zeta_\tau, w_\tau \rangle\} \\
&= \mathbf{E}\{\langle \mathcal{A}(w_\tau - \zeta_\tau), w_\tau \rangle\} = 0,
\end{aligned}$$

whence $\mathbf{E}\{R_t\} = 0$ for all t. By completely similar reasoning, $\mathbf{E}\{R_t\} = 0$ in the case of SA.

To complete the proof (ii), we need to prove (119). We have

$$\begin{aligned} s_t &\le \frac{\gamma_t}{2}\|F(\zeta_t)-F(\eta_t)\|_*^2 - \frac{1}{2\gamma_t}\|w_t-z_t\|^2 \ [\text{see (115)}] \\ &\le \frac{\gamma_t}{2}\left[\|F(w_t)-F(z_t)\|_* + \|F(\zeta_t)-F(w_t)\|_* + \|F(\eta_t)-F(z_t)\|_*\right]^2 - \frac{1}{2\gamma_t}\|w_t-z_t\|^2 \\ &\le \underbrace{\left[\frac{3\gamma_t}{2}\mathcal{L}^2 - \frac{1}{2\gamma_t}\right]}_{\le 0 \text{ by (118)}}\|w_t-z_t\|^2 + \frac{3\gamma_t}{2}\left[\|F(\zeta_t)-F(w_t)\|_*^2 + \|F(\eta_t)-F(z_t)\|_*^2\right]. \end{aligned}$$

It remains to note that $\|F(\zeta_t)-F(w_t)\|_*^2 + \|F(\eta_t)-F(z_t)\|_*^2 \le 2M_*^2$ since $\zeta_t, w_t, \eta_t, z_t \in Z$ and that the conditional, over the respective pasts, expectations of $\|F(\zeta_t)-F(w_t)\|_*^2$ and $\|F(\eta_t)-F(z_t)\|_*^2$ do not exceed σ^2. □

4.6.3.2 Proof of Proposition 4.3.2

We start with observing that (125) $\gamma_1 \ge \gamma_2 \ge ....$

1^0. Let us verify first that with the choice (125) of γ_τ, $\tau = 1, 2, ...$ we have for all $t = 1, 2, ...,$

$$\sqrt{2}\Omega\gamma_t^{-1} \ge S_t. \tag{152}$$

Indeed, for $t = 2, 3, ...$ we have (with $2S_0 = F_*^2$ in the case of SA and $2S_0 = 8\Omega\mathcal{L}^2$, $\ge M_*^2$ by (107), in the case of SMP)

$$\frac{\gamma_{t-1}^2}{\gamma_t^2} = \frac{\sum_{\tau=1}^{t-1} 2[s_\tau]_+/\gamma_\tau + 2S_0}{\sum_{\tau=1}^{t-2} 2[s_\tau]_+/\gamma_\tau + 2S_0} \le 1 + \frac{2[s_{t-1}]_+/\gamma_{t-1}}{2S_0} \le 2 \tag{153}$$

(recall that $2s_t/\gamma_\tau \le 2S_0$ by (116)). On the other hand

$$\gamma_t^{-2} - \gamma_{t-1}^{-2} = \frac{2[s_{t-1}]_+}{\Omega\gamma_{t-1}},$$

and

$$\gamma_t^{-1} - \gamma_{t-1}^{-1} \ge \tfrac{\gamma_t}{2}(\gamma_t^{-2} - \gamma_{t-1}^{-2}) = \tfrac{\gamma_t[s_{t-1}]_+}{\gamma_{t-1}\Omega} \ge \tfrac{[s_{t-1}]_+}{\sqrt{2}\Omega} \Rightarrow \sqrt{2}\Omega[\gamma_t^{-1} - \gamma_{t-1}^{-1}] \ge [s_{t-1}]_+$$

where the second inequality in this chain follows from $\gamma_{t-1} \le \sqrt{2}\gamma_t$ which is implied by (153). By summing up the resulting inequalities in the above chain, we get

$$\sqrt{2}\Omega\gamma_t^{-1} \ge \sum_{\tau=1}^{t-1} s_\tau + \sqrt{2}\Omega\gamma_1^{-1}. \tag{154}$$

In the case of SMP, we have $\gamma_1 = (2\sqrt{2}\mathcal{L})^{-1}$, whence $\sqrt{2}\Omega\gamma_1^{-1} = 4\Omega\mathcal{L} \geq \gamma_1 M_*^2 \geq \gamma_t M_*^2$ (see (107)), whence $\sqrt{2}\Omega\gamma_1^{-1} \geq s_t$ in view of (115), and (154) implies (152). In the case of SA, we have $\gamma_1 = \sqrt{\Omega}/F_*$, whence $\sqrt{2}\Omega\gamma_1^{-1} = \sqrt{2}\sqrt{\Omega}F_* \geq \gamma_1 F_*^2 \geq \gamma_t F_*^2$, whence $\sqrt{2}\Omega\gamma_1^{-1} \geq s_t$ by (115), and (152) again is given by (154).

2^0. Invoking (114), (152) implies (126). Now, by (115) in the case of SA we have $2[s_\tau]_+/\gamma_\tau \leq \|F(\zeta_\tau)\|_*^2$. In the case of SMP we have

$$
\begin{aligned}
&2[s_\tau]_+/\gamma_\tau \leq \|F(\zeta_\tau) - F(\eta_\tau)\|_*^2 - \gamma_\tau^{-2}\|w_\tau - z_\tau\|^2 \ \text{[see (115)]}\\
&\leq [\|F(\zeta_\tau) - F(w_\tau)\|_* + \|F(w_\tau) - F(z_\tau)\|_* + \|F(z_\tau) - F(\eta_\tau)\|_*]^2 - \gamma_1^{-2}\|w_\tau - z_\tau\|^2\\
&\leq 3\left[\|F(\zeta_\tau) - F(w_\tau)\|_*^2 + \|F(z_\tau) - F(\eta_\tau)\|_*^2\right] + \left[3\|F(w_\tau) - F(z_\tau)\|_*^2 - \gamma_1^{-2}\|w_\tau - z_\tau\|^2\right]\\
&\leq \varsigma_\tau := 3\left[\|F(\zeta_\tau) - F(w_\tau)\|_*^2 + \|F(z_\tau) - F(\eta_\tau)\|_*^2\right] \ \text{[by (106) due to } \gamma_1^{-1} = 2\sqrt{2}\mathcal{L}]
\end{aligned}
$$

Invoking (125), we get

$$
\gamma_t^{-1} \leq \Omega^{-1/2} \cdot \begin{cases} \left(F_*^2 + \sum_{\tau=1}^{t-1}\|F(\zeta_\tau)\|_*^2\right)^{1/2}, & \text{in the case of SA} \\ \left(8\Omega\mathcal{L}^2 + \sum_{\tau=1}^{t-1}\varsigma_\tau\right)^{1/2}, & \text{in the case of SMP} \end{cases} \tag{155}
$$

which combines with (126) to imply (127). □

4.6.3.3 Proof of Proposition 4.3.3

1^0. Let us denote

$$
\varphi_t = 8\Omega\mathcal{L}^2 + \sum_{\tau=1}^{t-1}\varsigma_\tau,
$$

where $\varsigma_t = 3\left[\|F(\zeta_t) - F(w_t)\|_*^2 + \|F(\eta_t) - F(z_t)\|_*^2\right]$ (cf. (128)). Let us show that under the premise of Proposition 4.3.3

$$
\forall \Lambda \geq 0: \quad \text{Prob}\left\{\varphi_t \geq O(1)\left[\Omega\mathcal{L}^2 + \frac{M_*^2 t}{k}\varkappa_*^2(k,\Lambda)\right]\right\} \leq \exp\{-\Lambda t\}, \tag{156}
$$

where $O(1)$ is an absolute constant factor. We use the following result (see, e.g., Theorem 2.1 (iii) of [77]): let $\xi^i, ..., \xi^k$ be k independent vectors from E with $\|\xi^i\|_* \leq \sigma$ and $\mathbf{E}\{\xi^i\} = 0$, where the norm $\|\cdot\|_*$ is $\varkappa$-regular, $\varkappa \geq 1$. Then for any $u \geq 0$

$$
\text{Prob}\left\{\|\sum_{i=1}^k \xi^i\|_* \geq \left[\sqrt{2\varkappa} + u\sqrt{2}\right]\sigma\sqrt{k}\right\} \leq \exp\{-u^2/2\}.
$$

When rewriting the above bound for $\xi^i = F(\zeta^i) - F(w)$ and $\xi^i = F(\eta^i) - F(z)$ and taking into account that $\|\xi^i\|_* \le M_*$ we obtain

$$\forall u \ge 0: \quad \mathrm{Prob}\left\{\|\sum_{i=1}^k \xi^i\|_*^2 \ge M_*^2 k(\sqrt{2\varkappa} + \sqrt{2}u)^2\right\} \le \exp\{-u^2/2\}.$$

So, if we denote Prob_t conditional probability over $\zeta_1, \eta_1, ..., \zeta_{t-1}, \eta_{t-1}$ being fixed, we get

$$\forall u \ge 0: \quad \mathrm{Prob}_t\left\{\varsigma_t \ge \frac{24M_*^2}{k}(\varkappa + u)\right\} \le 2\exp\{-u/2\},$$

where $\varsigma_t = 3\left[\|F(\zeta_t) - F(w_t)\|_*^2 + \|F(\eta_t) - F(z_t)\|_*^2\right]$ (cf. (128)). When setting $\nu_t = \frac{\varsigma_t k}{24M_*^2}$, we have for the conditional expectation $\mathbf{E}_t$ over $\zeta_1, \eta_1, ..., \zeta_{t-1}, \eta_{t-1}$ being fixed and $0 \le \alpha < 1$

$$\begin{aligned}\mathbf{E}_t\left\{\exp\{\frac{\alpha}{2}\nu_t\}\right\} &\le e^{\frac{\alpha\varkappa}{2}} + \frac{\alpha}{2}\int_\varkappa^\infty e^{\frac{\alpha u}{2}}\mathrm{Prob}_t\{\nu_t \ge u\}du \\ &\le e^{\frac{\alpha\varkappa}{2}} + \alpha\int_\varkappa^\infty \exp\{-\frac{(1-\alpha)u}{2}\}du = \frac{1+\alpha}{1-\alpha}\exp\{\frac{\alpha\varkappa}{2}\}\end{aligned}$$

When choosing $\alpha_* = \frac{\exp\{1\}-1}{\exp\{1\}+1}$ we get $\mathbf{E}_t\left\{\exp\{\frac{\alpha_*\nu_t}{2}\}\right\} \le \exp\{\frac{\alpha_*\varkappa}{2} + 1\}$, so that

$$\begin{aligned}\mathbf{E}\left\{\exp\{\sum_{\tau=1}^t \frac{\alpha_*\nu_\tau}{2}\}\right\} &= \mathbf{E}\left\{\mathbf{E}_t\left\{\exp\{\sum_{\tau=1}^{t-1}\frac{\alpha_*\nu_\tau}{2}\}\exp\{\frac{\alpha_*\nu_t}{2}\}\right\}\right\} \\ &= \mathbf{E}\left\{\exp\{\sum_{\tau=1}^{t-1}\frac{\alpha_*\nu_\tau}{2}\}\mathbf{E}_t\left\{\exp\{\frac{\alpha_*\nu_t}{2}\}\right\}\right\} \le \exp\{t(\frac{\alpha_*\varkappa}{2} + 1)\}\end{aligned}$$

Hence, when applying the Tchebychev inequality we find

$$\forall \Lambda \ge 0: \quad \mathrm{Prob}\left\{\sum_{\tau=1}^t \nu_\tau \ge t\left(\varkappa + \frac{2}{\alpha_*}(1+\Lambda)\right)\right\} \le \exp\{-\Lambda t\}.$$

When recalling that $\varsigma_t \le 6M_*^2$, we conclude that

$$\forall \Lambda \ge 0: \quad \mathrm{Prob}\left\{\sum_{\tau=1}^{t-1}\varsigma_\tau \ge \min\left[6M_*^2 t,\ \frac{24M_*^2 t}{k}\left(\varkappa + \frac{2}{\alpha_*}(1+\Lambda)\right)\right]\right\} \le \exp\{-\Lambda t\}.$$

Since $\varkappa \ge 1$, $\varkappa + \frac{2}{\alpha_*}(1+\Lambda) \le O(1)\varkappa_*^2(k,\Lambda)$, and we arrive at (156).

2^0. We have

$$\forall \lambda \ge 0: \quad \mathrm{Prob}\left\{\frac{R_t}{t} \ge O(1)F_*\sqrt{\frac{\Omega\lambda}{kt}}\right\} \le e^{-\lambda}. \tag{157}$$

Indeed, since $\mathcal{A}$ is skew-symmetric, i.e., $\langle \mathcal{A}z, z\rangle = 0$,

$$r_t \;=\; \langle F(\zeta_t),\, \zeta_t - w_t\rangle = \langle a + \mathcal{A}\zeta_t,\, \zeta_t - w_t\rangle = \langle a + \mathcal{A}w_t, \zeta_t - w_t\rangle = \langle F(w_t), \zeta_t - w_t\rangle.$$

We conclude that

$$\begin{aligned}\frac{R_t}{t} &= \frac{1}{t}\sum_{\tau=1}^{t} r_t = \frac{1}{t}\sum_{\tau=1}^{t}\langle F(w_\tau), \zeta_\tau - w_\tau\rangle = \frac{1}{t}\sum_{\tau=1}^{t}\left\langle F(w_\tau),\, \frac{1}{k}\sum_{i=1}^{k}\zeta_\tau^i - w_\tau\right\rangle \\ &= \frac{1}{tk}\sum_{\tau=1}^{t}\sum_{i=1}^{k}\langle F(w_\tau),\, \zeta_\tau^i - w_\tau\rangle = \frac{1}{tk}\sum_{\tau=1}^{t}\sum_{i=1}^{k}\xi_\tau^i,\end{aligned}$$

where $\xi_\tau^i := \langle F(w_\tau),\, \zeta_\tau^i - w_\tau\rangle$ is a scalar martingale-difference with $|\xi_\tau^i| \le 2\mathcal{R}F_* \le 2\Theta F_*$ (cf. (105)). Then by the Azuma-Hoeffding inequality [8],

$$\forall \lambda \ge 0: \quad \mathrm{Prob}\left\{\frac{R_t}{t} \ge 2\Theta F_*\sqrt{\frac{2\lambda}{kt}}\right\} \le e^{-\lambda},$$

which implies (157).

Now we are done – when substituting the bounds (156) and (157) into (127) we get

$$\mathrm{Prob}\left\{\epsilon_{\mathrm{sad}}(z^t) \ge O(1)\left[\frac{\Omega\mathcal{L}}{t} + M_*\varkappa_*(k,\Lambda)\sqrt{\frac{\Omega}{kt}} + \Theta F_*\sqrt{\frac{\lambda}{kt}}\right]\right\} \le e^{-\Lambda t} + e^{-\lambda},$$

which is (130) if we recall that $\Theta = \sqrt{2\Omega}$ and $F_* \le \|a\|_* + 2\Theta\mathcal{L}$ (cf. (108)). □

4.6.3.4 Proof of Proposition 4.3.4

This proof is completely similar to the one of Proposition 4.3.1 and is omitted.

4.6.4 Proof of Theorem 4.4.1

1^0. From the description of the method it follows that

$$\forall t, s \ge 1, \rho \ge 0: u^{ts} \ge \mathrm{SV}(\rho_s) \ge \ell^{ts},\ \ell_{ts}(\rho) \le \mathrm{SV}(\rho), \ell^{ts} \le \ell_{ts}(\rho_s). \tag{158}$$

Let us prove by induction in s that $\rho_* \le \rho_s \le \rho_1$. The base $s = 1$ is evident. Now let $\rho_* \le \rho_s \le \rho_1$, and let stage $s+1$ take place. When passing from stage s to stage $s+1$, we are in the case B) and thus have $u^{ts} > \epsilon\rho_s$, $\ell^{ts} \ge \frac{3}{4}u^{ts} > \frac{3}{4}\epsilon\rho_s$, whence, in view of (158),

$$\ell_s(\rho_s) = \ell_{ts}(\rho_s) \ge \ell^{ts} \ge \frac{3}{4}\max[\epsilon\rho_s, \mathrm{SV}(\rho_s)] \;\&\; \ell_s(\rho_s) > 0. \tag{159}$$

This combines with $\ell_{ts}(\rho_*) \le \mathrm{SV}(\rho_*) \le 0$ and convexity of $\ell_{ts}(\cdot)$ to imply that $\rho_* \le \rho_{s+1} < \rho_s$. Induction is complete.

Since $\rho_s \ge \rho_*$, u^{ts} is an upper bound on $\mathrm{SV}(\rho_s)$ and $u^{ts} \ge \overline{\phi}^{\rho_s}(w_1^{ts})$, we conclude that if the algorithm terminates at stage s, then the result ρ_s, w_1^{ts} is an ϵ-solution to the GBSP in question.

2^0. Let us prove (ii). The reasoning to follow goes back to [84]; we reproduce it here to make the chapter self-contained. Let s be such that the stage $s+1$ takes place, and let u_s be the last bound u^{ts} built at stage s. Observe that

$$\frac{3}{4}\epsilon\rho_s < \frac{3}{4}u_s \le \ell_s(\rho_s) \le \mathrm{SV}(\rho_s) \le u_s. \tag{160}$$

Since the convex function $\ell_s(\rho)$ is nonpositive at $\rho = \rho_{s+1}$ and is $\ge \frac{3}{4}u_s > 0$ at $\rho = \rho_s > \rho_{s+1}$, we have $g_s := \ell_s'(\rho_s) > 0$ and

$$\rho_s - \rho_{s+1} \ge \ell_s(\rho_s)/g_s \ge \frac{3}{4}u_s/g_s. \tag{161}$$

Now assume that $s > 1$ is such that the stage $s+1$ takes place. Applying (161) and (160) to $s-1$ in the role of s, we get $\rho_{s-1} - \rho_s \ge \frac{3}{4}u_{s-1}/g_{s-1}$ and $\frac{3}{4}u_s \le \ell_s(\rho_s)$, whence, by convexity of $\ell_s(\cdot)$ and in view of (158), $u_{s-1} \ge \mathrm{SV}(\rho_{s-1}) \ge \ell_s(\rho_{s-1}) \ge \ell_s(\rho_s) + g_s(\rho_{s-1} - \rho_s) \ge \frac{3}{4}u_s + g_s\frac{3}{4}\frac{u_{s-1}}{g_{s-1}}$, so that $\frac{4}{3}u_{s-1} \ge u_s + \frac{g_s u_{s-1}}{g_{s-1}}$, or $\frac{u_s}{u_{s-1}} + \frac{g_s}{g_{s-1}} \le \frac{4}{3}$, whence $\frac{u_s g_s}{u_{s-1}g_{s-1}} \le (1/4)(4/3)^2 = 4/9$. It follows that

$$\sqrt{u_s g_s} \le (2/3)^{s-1}\sqrt{u_1 g_1}. \tag{162}$$

We have $\ell_s(\rho_*) \le \mathrm{SV}(\rho_*) = 0$, $\ell_s(\rho_s) \ge \frac{3}{4}u_s$ (see (160)) and $\ell_s(\rho_s) - \ell_s(\rho_*) \le g_s(\rho_s - \rho_*)$ (convexity of $\ell_s(\cdot)$), whence $g_s \ge \frac{3}{4}u_s(\rho_s - \rho_*)^{-1} \ge \frac{3}{4\rho_1}u_s$, and (162) implies that

$$u_s \le (2/3)^{s-1}\sqrt{u_1 g_1}\sqrt{4\rho_1/3}. \tag{163}$$

Now, $g_1 = \ell_1'(\rho_1)$ and $\ell_1(\rho) \le \mathrm{SV}(\rho) \le \|\phi\|_\infty + \rho\|\psi\|_\infty$, whence $g_1 \le \|\psi\|_\infty$, and clearly $u_1 \le \|\phi\|_\infty + \rho_1\|\psi\|_\infty$. At the same time, $u_s > \epsilon\rho_s \ge \epsilon\rho_*$, so that (163) implies that $\epsilon\rho_* \le (2/3)^{s-1}[\|\phi\|_\infty + \rho_1\|\psi\|_\infty]$. The resulting upper bound on s implies (ii).

3^0. Let us prove (iii). From the description of the algorithm it follows that at every stage s before termination of the stage the residual of current approximate solutions w^{ts} is $\geq \frac{1}{4}\epsilon\rho_s$ (since $u^{ts} > \epsilon\rho_s$ and $\ell^{ts} < \frac{3}{4}u^{ts}$). In the case of short-step implementation we use the result of Proposition (4.3.3) with $\varepsilon = \epsilon\rho_s$. Let us denote $N_s(\epsilon)$ the corresponding value of N_ε as in (131). We conclude that the number N_s of steps at stage s is finite with probability 1 and satisfies $\mathrm{Prob}\{N_s > N_s(\epsilon)\} \leq \exp\{-\Lambda N_s(\epsilon)\} + \exp\{-\lambda\}$. As we have seen, $\rho_* \leq \rho_s$ for all s, and therefore $N_s \leq N(\epsilon)$ for all s, provided that the absolute constant $O(1)$ in (145) is properly chosen.

For the aggressive-step implementation, similar reasoning based on the bound (137) with $\mathcal{L} = \mathcal{M} + \rho_s\mathcal{N}$ justifies (146).

4^0. Combining (ii), (iii) and the concluding claim in item 1^0 above, we arrive at (i). □

4.7 Detailed Numerical Experiments of Chapter 4

Here we provide detailed results comparing the effects of different policies for choosing the starting point at each stage and different choices of the distance generating function (d.g.f.) for the simplexes on the performance of our algorithms. We encode various different policies with codes "$XNNY$" where

- X can be [C]enter of the domain, or a weighted combination of center point with the [B]est solution or [L]ast solution from the previous stages;
- NN gives the percentage for the convex combination weight used for the given starting point X and $100 - NN$ is the percentage for the convex combination weight used for the center of the domain, the possible values tested are 0.[00], 0.[25], 0.[75];
- Y denotes the distance generating function used for simplexes, with the [E]xponential, [P]ower d.g.f. options.

Thus "B25P" implementation would mean that power d.-g.f. is used, and at each stage the algorithm restarts from the convex combination of the best (with the smallest ϵ_{sad}) point found so far and the ω-center of Z, the weights being 0.25 and 0.75, respectively.

The instances used are the same from Section 4.5 and for each given instance size, the numbers in the tables are the respective averages for performance of the selected policy on the corresponding instances.

Table 13: Detailed Results for ℓ_1-minimization with $\|\cdot\|_\infty$-fit

		DMP				**SMP**				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
500 × 4000	C00E	2661.6	106.6	912.2	4.0	10511.0	236.5	57.2	3505.2	4.8
	B25E	3157.4	124.9	1080.8	4.6	9235.6	207.8	52.1	3073.0	4.8
	B50E	3856.0	156.0	1317.6	4.8	10303.2	231.8	55.3	3421.0	4.8
	B75E	4495.4	180.8	1534.0	4.8	10524.4	236.8	55.3	3508.8	4.6
	B25P	1453.4	104.1	506.4	4.4	8607.8	193.7	106.4	2880.8	4.6
	B50P	1732.0	116.6	602.6	4.8	8708.4	195.9	107.1	2913.0	4.6
	B75P	1986.2	132.0	688.2	4.8	10162.2	228.6	115.6	3385.0	4.8
	L25P	1326.0	92.5	462.6	4.4	13666.4	307.5	166.3	4569.0	5.2
	L50P	1592.8	109.4	553.6	4.8	15013.2	337.8	179.6	4985.2	5.2
	L75P	1806.4	123.5	627.0	4.8	14762.2	332.1	181.4	4941.0	5.0
min		1326.0	92.5	462.6	4.0	8607.8	193.7	52.1	2880.8	4.6
max		4495.4	180.8	1534.0	4.8	15013.2	337.8	181.4	4985.2	5.2
1000 × 2000	C00E	1830.8	64.0	629.6	4.2	10568.8	158.5	42.9	3536.6	4.8
	B25E	2573.4	85.6	881.2	4.4	11136.4	167.1	38.1	3725.2	5.0
	B50E	3027.2	90.7	1038.4	4.8	12502.8	187.5	46.4	4185.4	5.0
	B75E	3759.2	126.0	1284.8	4.8	10881.6	163.2	41.9	3633.2	5.0
	B25P	1530.6	97.9	532.2	4.6	10709.2	160.6	113.3	3593.2	4.8
	B50P	1735.0	112.3	603.6	4.6	11285.6	169.3	117.9	3765.0	5.0
	B75P	1750.4	109.5	607.8	4.6	11339.4	170.1	117.3	3769.0	5.0
	L25P	1260.2	81.1	440.8	4.6	16031.8	240.5	172.6	5341.0	5.2
	L50P	1505.6	90.8	524.6	5.0	15116.4	226.7	148.7	5065.0	5.2
	L75P	1958.8	121.1	680.2	5.0	18181.4	272.7	198.5	6041.0	5.6
min		1260.2	64.0	440.8	4.2	10568.8	158.5	38.1	3536.6	4.8
max		3759.2	126.0	1284.8	5.0	18181.4	272.7	198.5	6041.0	5.6

Table 13: Detailed Results for ℓ_1-minimization with $\|\cdot\|_\infty$-fit (Continued)

		DMP				SMP				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
1000 × 8000	C00E	2338.0	227.9	801.6	4.0	12406.6	139.6	113.2	4136.6	5.0
	B25E	3212.2	300.9	1100.4	5.0	11324.2	127.4	122.7	3768.6	5.0
	B50E	3658.0	349.0	1252.2	5.0	14077.2	158.4	126.6	4700.8	5.0
	B75E	4721.4	455.8	1611.4	5.0	14301.2	160.9	137.2	4745.0	5.0
	B25P	1545.6	248.9	539.0	4.8	11355.4	127.7	266.0	3785.0	5.0
	B50P	2023.4	313.5	701.8	5.0	12386.2	139.3	297.4	4149.0	5.0
	B75P	2625.6	427.4	906.6	5.0	12935.0	145.5	299.0	4321.0	5.2
	L25P	1441.8	237.1	504.6	5.0	14922.2	167.9	343.0	4989.0	5.2
	L50P	1661.4	261.1	579.8	5.0	15944.6	179.4	378.7	5336.8	5.2
	L75P	2361.2	395.9	818.2	5.0	17466.4	196.5	425.7	5821.0	5.4
min		1441.8	227.9	504.6	4.0	11324.2	127.4	113.2	3768.6	5.0
max		4721.4	455.8	1611.4	5.0	17466.4	196.5	425.7	5821.0	5.4
2000 × 4000	C00E	2691.6	227.6	925.2	4.4	12922.8	96.9	74.5	4293.0	5.0
	B25E	2322.4	190.5	800.4	5.0	12303.8	92.3	79.4	4121.0	5.0
	B50E	2929.2	245.1	1006.0	5.0	13877.6	104.1	86.0	4613.0	5.2
	B75E	3839.6	346.6	1314.4	5.0	16366.0	122.7	109.2	5457.0	5.2
	B25P	1426.4	207.8	500.6	5.0	16294.8	122.2	334.5	5457.0	5.4
	B50P	1426.4	231.7	500.8	5.0	13462.2	101.0	271.7	4485.0	5.0
	B75P	1948.4	295.4	677.4	5.0	16149.8	121.1	311.0	5401.0	5.6
	L25P	1183.2	178.2	417.2	5.0	19280.8	144.6	367.2	6433.0	6.0
	L50P	1665.4	224.0	582.6	5.2	20075.8	150.6	390.5	6725.2	6.0
	L75P	2344.0	309.7	813.0	5.0	20953.8	157.2	413.9	6964.8	6.0
min		1183.2	178.2	417.2	4.4	12303.8	92.3	74.5	4121.0	5.0
max		3839.6	346.6	1314.4	5.2	20953.8	157.2	413.9	6964.8	6.0
2000 × 16000	C00E	2384.6	494.2	819.6	4.0	13174.8	74.1	184.9	4405.0	5.0
	B25E	3525.0	733.5	1205.6	4.8	15646.8	88.0	221.6	5229.0	5.0
	B50E	4303.2	912.2	1471.4	4.8	14498.6	81.6	207.5	4861.0	5.0
	B75E	5300.0	1130.6	1809.6	4.8	16908.6	95.1	238.9	5629.2	5.4
	B25P	1575.2	533.7	550.0	4.8	14623.4	82.3	599.3	4873.2	5.0
	B50P	1915.0	637.1	665.6	4.6	16554.0	93.1	671.9	5541.2	5.2
	B75P	2319.2	754.3	803.2	4.8	15903.6	89.5	639.2	5329.0	5.0
	L25P	1838.4	615.9	639.4	4.8	18717.0	105.3	773.4	6256.8	5.6
	L50P	2245.6	736.7	778.2	5.2	18882.8	106.2	775.1	6289.2	5.4
	L75P	2892.8	921.9	996.2	5.2	19391.2	109.1	796.5	6465.0	5.6
min		1575.2	494.2	550.0	4.0	13174.8	74.1	184.9	4405.0	5.0
max		5300.0	1130.6	1809.6	5.2	19391.2	109.1	796.5	6465.0	5.6

Table 13: Detailed Results for ℓ_1-minimization with $\|\cdot\|_\infty$-fit (Continued)

		DMP				SMP				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
4000 × 8000	C00E	2923.6	798.7	1005.0	4.4	19750.2	74.1	228.4	6589.0	5.4
	B25E	3386.8	795.1	1163.6	5.0	20047.2	75.2	215.8	6697.0	5.8
	B50E	4332.8	983.3	1482.4	5.2	18337.0	68.8	198.4	6101.0	5.4
	B75E	5197.6	1237.1	1778.0	5.2	19546.2	73.3	228.0	6529.0	5.8
	B25P	1554.6	576.2	544.4	5.0	19242.2	72.2	748.9	6425.0	5.6
	B50P	2068.8	704.1	720.2	5.2	20343.8	76.3	768.8	6781.0	5.8
	B75P	2684.8	882.7	928.4	5.4	21494.2	80.6	824.5	7145.0	6.0
	L25P	1823.4	641.4	635.0	5.0	20767.0	77.9	799.3	6937.0	6.0
	L50P	2433.8	766.7	843.2	5.4	22350.0	83.8	865.4	7465.2	6.4
	L75P	3130.0	987.1	1079.2	5.2	24185.8	90.7	934.4	8077.2	6.2
min		1554.6	576.2	544.4	4.4	18337.0	68.8	198.4	6101.0	5.4
max		5197.6	1237.1	1778.0	5.4	24185.8	90.7	934.4	8077.2	6.4

Table 14: Detailed Results for ℓ_1-minimization with $\|\cdot\|_2$-fit

		DMP				SMP				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
500 × 4000	C00E	579.8	21.0	206.2	5.0	4771.6	106.7	24.6	1589.0	4.4
	B25E	523.0	18.3	187.4	5.2	5186.0	115.9	27.7	1737.0	5.0
	B50E	527.6	19.3	189.2	5.2	5308.4	118.7	29.4	1781.0	5.0
	B75E	610.2	22.4	218.8	5.8	5204.4	116.3	28.9	1749.0	5.0
	B25P	539.4	30.8	194.6	4.8	5331.6	119.2	67.2	1781.0	5.0
	B50P	480.2	29.0	175.0	5.0	5215.0	116.6	61.7	1744.8	5.0
	B75P	421.2	24.7	156.2	5.4	5416.6	121.1	65.5	1797.0	5.0
	L25P	403.0	23.6	147.0	4.4	5453.8	121.9	64.3	1825.0	5.0
	L50P	334.8	19.7	125.2	5.0	5549.6	124.0	64.2	1857.0	5.0
	L75P	287.8	16.1	110.2	5.2	5938.4	132.7	66.9	1984.8	5.0
min		287.8	16.1	110.2	4.4	4771.6	106.7	24.6	1589.0	4.4
max		610.2	30.8	218.8	5.8	5938.4	132.7	67.2	1984.8	5.0

Table 14: Detailed Results for ℓ_1-minimization with $\|\cdot\|_2$-fit (Continued)

		DMP				**SMP**				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
1000 × 2000	C00E	553.0	19.0	198.0	5.0	3910.8	54.8	13.6	1296.2	4.2
	B25E	531.8	17.3	190.8	5.0	4880.0	68.3	17.2	1612.8	4.8
	B50E	534.0	15.5	192.4	5.2	5524.8	77.3	21.6	1849.0	5.0
	B75E	780.8	22.4	278.6	6.0	5294.8	74.1	21.6	1765.0	5.0
	B25P	465.0	17.7	170.4	4.8	5268.0	73.8	54.9	1757.0	5.0
	B50P	440.4	17.7	162.8	5.0	5256.6	73.6	52.6	1753.0	5.0
	B75P	431.4	15.9	159.4	5.0	5295.4	74.1	54.5	1772.8	5.0
	L25P	467.2	19.4	170.8	5.0	5863.2	82.1	59.4	1941.0	5.0
	L50P	374.4	17.6	139.6	5.0	5657.6	79.2	57.7	1885.0	5.0
	L75P	282.4	14.1	110.4	5.6	5923.6	82.9	61.7	1977.0	5.0
min		282.4	14.1	110.4	4.8	3910.8	54.8	13.6	1296.2	4.2
max		780.8	22.4	278.6	6.0	5923.6	82.9	61.7	1977.0	5.0
1000 × 8000	C00E	617.0	56.6	220.4	5.0	5148.8	57.5	50.7	1716.8	4.6
	B25E	493.4	41.4	178.8	5.2	5393.8	60.3	53.3	1797.0	5.0
	B50E	433.8	43.6	159.0	5.6	5552.4	62.1	55.0	1857.0	5.0
	B75E	597.4	61.5	216.6	6.0	5535.2	61.9	54.2	1825.0	5.0
	B25P	535.0	85.9	195.8	5.0	5279.4	59.0	134.6	1765.0	5.0
	B50P	471.2	74.9	173.8	5.0	5353.2	59.8	132.7	1789.0	5.0
	B75P	460.4	64.3	171.8	5.6	5425.8	60.6	135.0	1809.0	5.0
	L25P	438.4	68.4	162.6	5.0	5626.6	62.9	135.9	1881.0	5.0
	L50P	391.2	52.4	146.0	5.0	6238.4	69.7	155.1	2081.0	5.2
	L75P	318.8	40.9	124.0	5.8	5917.2	66.1	143.9	1957.0	5.0
min		318.8	40.9	124.0	5.0	5148.8	57.5	50.7	1716.8	4.6
max		617.0	85.9	220.4	6.0	6238.4	69.7	155.1	2081.0	5.2
2000 × 4000	C00E	634.8	39.8	227.2	5.0	5853.6	41.0	47.2	1952.8	4.8
	B25E	720.6	45.9	256.4	5.0	5653.4	39.6	50.3	1897.0	5.0
	B50E	617.8	39.3	222.6	5.4	5786.4	40.5	49.7	1937.0	5.0
	B75E	755.0	48.1	271.0	6.0	5946.2	41.6	50.4	1985.0	5.0
	B25P	427.8	37.0	159.8	5.0	5566.6	39.0	124.4	1845.0	5.0
	B50P	430.8	34.2	161.0	5.0	5524.8	38.7	121.2	1856.8	5.0
	B75P	479.0	37.8	178.8	5.6	5849.0	40.9	126.8	1941.0	5.0
	L25P	497.0	42.1	182.8	5.0	6241.2	43.7	138.6	2081.0	5.0
	L50P	383.6	32.0	144.0	5.0	6354.2	44.5	139.6	2105.0	5.0
	L75P	318.8	25.9	124.4	5.8	8832.8	61.8	201.3	2957.0	5.4
min		318.8	25.9	124.4	5.0	5524.8	38.7	47.2	1845.0	4.8
max		755.0	48.1	271.0	6.0	8832.8	61.8	201.3	2957.0	5.4

Table 14: Detailed Results for ℓ_1-minimization with $\|\cdot\|_2$-fit (Continued)

		DMP				**SMP**				
Size	Option	Calls	CPU	Steps	Stages	Calls	FCalls	CPU	Steps	Stages
2000×16000	C00E	531.8	150.7	192.6	5.0	5055.6	28.3	90.0	1689.0	4.4
	B25E	562.0	154.1	202.8	5.0	5802.4	32.4	99.6	1921.0	5.0
	B50E	623.8	161.7	224.2	5.0	6061.8	33.9	104.5	2017.0	5.0
	B75E	611.4	162.1	220.4	5.4	6049.2	33.8	108.2	2013.0	5.0
	B25P	500.4	175.8	184.4	4.8	5488.6	30.7	274.8	1841.0	5.0
	B50P	490.4	191.5	181.6	5.0	5820.4	32.5	281.5	1933.0	5.0
	B75P	478.6	156.5	179.6	5.6	5744.2	32.1	283.8	1917.0	5.0
	L25P	416.6	144.5	156.0	4.8	7620.0	42.6	374.2	2537.0	5.2
	L50P	402.8	132.1	151.6	5.0	9525.2	53.2	464.0	3185.0	5.4
	L75P	346.0	110.6	133.0	5.4	12614.4	70.5	629.5	4217.2	5.8
min		346.0	110.6	133.0	4.8	5055.6	28.3	90.0	1689.0	4.4
max		623.8	191.5	224.2	5.6	12614.4	70.5	629.5	4217.2	5.8
4000×8000	C00E	675.2	138.5	242.0	5.0	6504.6	22.8	101.7	2177.0	5.0
	B25E	685.8	141.7	245.6	5.0	5782.0	20.2	91.3	1929.0	5.0
	B50E	527.4	106.2	191.8	5.0	5894.4	20.6	96.2	1957.0	5.0
	B75E	870.4	182.0	312.0	6.0	5899.4	20.6	96.5	1973.0	5.0
	B25P	483.0	116.3	180.0	5.0	5439.2	19.0	247.4	1809.0	5.0
	B50P	453.2	112.8	169.6	5.0	5635.8	19.7	256.3	1885.0	5.0
	B75P	512.4	125.0	190.4	5.4	6126.6	21.4	274.0	2037.0	5.2
	L25P	411.2	104.3	155.0	5.0	12948.0	45.3	615.5	4285.0	6.0
	L50P	403.8	97.9	152.2	5.0	11588.2	40.6	532.2	3877.0	5.8
	L75P	346.4	86.3	135.0	5.8	12099.0	42.3	553.2	4021.0	5.8
min		346.4	86.3	135.0	5.0	5439.2	19.0	91.3	1809.0	5.0
max		870.4	182.0	312.0	6.0	12948.0	45.3	615.5	4285.0	6.0

CHAPTER V

EFFICIENTLY VERIFIABLE ACCURACY CERTIFICATES FOR NOISY RECOVERY

The problem we consider in this chapter is to estimate a linear transform $Bx \in \mathbb{R}^N$ of a vector $x \in \mathbb{R}^n$ from the observations

$$y = Ax + u + \sigma\xi. \tag{164}$$

Here A is a given $m \times n$ sensing matrix, B is a given $N \times n$ matrix, and $u + \sigma\xi$ is the observation error; in this error, u is an unknown *nuisance* known to belong to a given compact convex set $\mathcal{U} \subset \mathbb{R}^m$ symmetric w.r.t. the origin, $\sigma \geq 0$ is a known noise intensity, and ξ is random noise with known distribution P.

We assume that the space $\mathbb{R}^N$ where Bx lives is represented as $\mathbb{R}^N = \mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_K}$, so that a vector $w \in \mathbb{R}^N$ is a block vector: $w = [w[1]; ...; w[K]]$ with blocks $w[k] \in \mathbb{R}^{n_k}$, $1 \leq k \leq K$. In particular, $Bx = [B[1]x; ...; B[K]x]$ with $n_k \times n$ matrices $B[k]$, $1 \leq k \leq K$. While we do not assume that the vector x is sparse in the usual sense, we do assume that the linear transform Bx to be estimated is *block sparse*, meaning that at most a given number, s, of the blocks $B[k]x$, $1 \leq k \leq K$, are nonzero.

The recovery routines we intend to consider are based on *block-ℓ_1 minimization*, i.e., the estimate $\widehat{w}(y)$ of $w = Bx$ is $B\widehat{z}(y)$, where $\widehat{z}(y)$ is obtained by minimizing the norm $\sum_{k=1}^K \|B[k]z\|_{(k)}$ over signals $z \in \mathbb{R}^n$ with Az "fitting," in certain precise sense, the observations y. Above, $\|\cdot\|_{(k)}$ are given in advance norms on the spaces $\mathbb{R}^{n_k}$ where the blocks of Bx take their values.

In the sequel we refer to the given in advance collection $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ as the *representation structure*. Given such a structure and A, our ultimate goal is to understand how well one can recover the s-block-sparse transform Bx by appropriately implementing block-ℓ_1 minimization.

Related compressed sensing research. Our situation and goal form a straightforward extension of the usual sparse/block sparse compressed sensing framework. Indeed, the *standard representation structure* $B = I_n$, $n_k = 1$, $\|\cdot\|_{(k)} = |\cdot|$, $1 \le k \le K = n$, leads to the standard compressed sensing setting – recovering a sparse signal $x \in \mathbb{R}^n$ from its noisy observations (164) via ℓ_1-minimization. With the same $B = I_n$ and nontrivial block structure $\{n_k, \|\cdot\|_{(k)}\}_{k=1}^K$, we arrive at *block-sparsity* and related block-ℓ_1-minimization routines considered in numerous recent papers. Specifically, there is a number of applications where block-sparsity (with $B = I_n$) arises naturally, e.g., in multi-band signals, measurements of gene expression levels or in the estimation of multiple measurement vectors sharing a joint sparsity pattern (see [55] and references therein). In addition, in many studies (e.g., [54, 55]) it was shown that the block-sparsity model can be used in sampling signals that lie in a union of subspaces. Moreover, several methods of estimation and selection extending plain ℓ_1-minimization to block sparsity were proposed and investigated recently. Most of the related research focused so far on *block regularization schemes* — Lasso-type algorithms of the form

$$\widehat{x}(y) \in \operatorname*{Argmin}_{z=[z^1;...;z^K]\in\mathbb{R}^n=\mathbb{R}^{n_1}\times...\times\mathbb{R}^{n_K}} \left\{\|Az-y\|_2^2 + \lambda\|z\|_{\ell_1/\ell_q}\right\}, \quad \|z\|_{\ell_1/\ell_q} := \sum_{k=1}^K \|z^k\|_q.$$

In particular, there is a huge literature on plain Lasso ($n_k = 1, 1 \le k \le K = n$), see [23, 24, 75, 78, 87, 93, 94, 123] and references therein, there is a significant counterpart on group Lasso (arbitrary n_k, $q = 2$), see, e.g., [9, 14, 33, 51, 54, 55, 65, 70, 85, 92, 96, 108, 109, 121, 131], and references therein. Another avenue of research here [75, 86] deals with block-sparse analogies of the Dantzig selector originating from [29]. Most of the cited papers focus on bounding recovery errors in terms of magnitude of the observation noise and "s-concentration" of the true signal x (that is, its ℓ_1/ℓ_q distance from the space of signals with at most s nonzero blocks) or algorithms to solve block-ℓ_1 regularization problems. Typically, these results deal with the case of $q = 2$ and rely on natural block analogy ("Block RIP," see, e.g., [55]) of the celebrated Restricted Isometry Property (RIP) introduced by Candés and Tao [27, 26], or on block analogies [88] of the Restricted Eigenvalue Property introduced in [18].

Contributions of this chapter. The first (and by itself, minor) novelty in our problem setting is in the presence of matrix B; we are not aware of any preceding work handling the case of a "nontrivial" (i.e., different from the identity) B. Introducing this matrix seems to be natural and adds some useful flexibility (and costs nearly nothing, as far as the theoretical analysis is concerned). In addition to this, there are a number of important applications where a nontrivial B arises naturally. As a simple example consider the standard image reconstruction with Total Variation (TV) regularization, where one wants to recover an image x from noisy observations of its convolution with a given kernel. The rationale behind TV regularization stems from the fact that the (discretized) gradient field Bx of the image, not the image x itself, is (nearly) sparse. Note that this example leads to a block sparsity setting, since Bx is naturally split into 2-dimensional blocks representing gradients at the grid points, and TV is just the sum of ℓ_2 norms of these blocks. Another example is when x is the solution of a linear finite-difference equation with sparse right hand side ("evolution of a linear plant corrected from time to time by impulse control"); in this case, B is the matrix of the corresponding finite-difference operator.

We believe, however, that the major novelty in what follows is the emphasis on *verifiable* conditions on A and the representation structure which *guarantee* good recovery of transforms Bx from noisy observations of Ax, provided that the transforms are nearly s-sparse, and the observation noise is low. In this respect, this chapter continues the line of research started in [81, 76, 79], where ℓ_1-recovery of the usual sparse vectors was considered (in the first two papers – in the case of uncertain-but-bounded observation errors, and in the third – in the case of Gaussian observation noise). To give an impression of the approach, we present here a summary of our major results. To streamline this summary, we restrict ourselves for the time being with the case where (a) the random noise ξ in (164) is standard Gaussian: $\xi \sim \mathcal{N}(0, I)$, and (b) all the norms $\|\cdot\|_{(k)}$ are just $\|\cdot\|_r$-norms, with r common for all values of k. In this case, an (incomplete) summary of our (somehow simplified) constructions and results is as follows. Let s be a given positive integer — an a priori upper bound on the number of nonzero blocks $B[k]x$ in the transforms we intend to recover well, and $\epsilon \ll 1$ be the a given tolerance. We fix an $m \times n$ sensing matrix A and a representation

structure $(B, n_1, ..., n_K, \|\cdot\|_r, ..., \|\cdot\|_r)$.

A.1. Given s and $q \in [1, \infty]$ and the norm $\|\cdot\|$, we introduce a condition $\mathbf{Q}_{s,q}$ on an $m \times N$ *contrast matrix* H, specifically, the condition

$$\forall (x \in \mathbb{R}^n) : L_{s,q}(Bx) \leq s^{\frac{1}{q}} \|H^T Ax\| + \frac{s^{\frac{1}{q}-1}}{3} L_1(Bx)$$

where for $w = [w[1]; ...; w[K]] \in \mathbb{R}^N$ and $p \in [1, \infty]$, $L_p(w) = \|[\|w[1]\|_r; ...; \|w[K]\|_r]\|_p$ is the ℓ_p/ℓ_r norm of w, and $L_{s,p}(w)$ is the norm of w obtained as follows: we zero out all but the s largest in "magnitude" $\|w[k]\|_r$ blocks in w, and take the L_p-norm of the resulting s-block-sparse vector. For example, $L_{s,\infty}(w)$ is, independently of s, just the maximum of magnitudes $\|w[k]\|_r$ of blocks in w.

A.2. Given an $m \times N$ contrast matrix H, we introduce two recovery routines:

- *regular L_1-recovery* (cf. (block) Dantzig selector)

$$\widehat{x}_{\rm reg}(y) \in \operatorname*{Argmin}_{z \in \mathbb{R}^n} \left\{ L_1(Bz) : \|H^T(y - Az)\| \leq \nu(H) \right\},$$

  where $\nu(H) := \max\limits_{1 \leq j \leq N} \left[\max\limits_{u \in \mathcal{U}} u^T h^j + \sigma \mathrm{Erfinv}(\frac{\epsilon}{2N}) \|h^j\|_2 \right]$ and $\mathrm{Erfinv}(\delta)$ is the inverse error function[1], and,

- *penalized L_1-recovery* (cf. (block) Lasso)

$$\widehat{x}_{\rm pen}(y) \in \operatorname*{Argmin}_{z \in \mathbb{R}^n} \left[L_1(Bz) + 2s\|H^T(y - Az)\| \right].$$

Note that the regular L_1-recovery can be undefined; this happens when the corresponding optimization problem is infeasible. The penalized recovery always is well defined.

A.3. Our main related result is as follows (see Theorems 5.3.1, 5.3.2): *Let a contrast matrix H satisfy the condition $\mathbf{Q}_{s,q}$. Then there exists a set Ξ of realizations of ξ such that* $\mathrm{Prob}\{\xi \in \Xi\} \geq 1 - \epsilon$ *and for all $\xi \in \Xi$, $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$, $\widehat{x}_{\rm reg}(Ax + u + \xi)$ is well defined, and for both $\widehat{x} = \widehat{x}_{\rm reg}(Ax + u + \sigma\xi)$ and $\widehat{x} = \widehat{x}_{\rm pen}(Ax + u + \sigma\xi)$ one has*

$$\forall p \in [1, q] : L_p(B\widehat{x} - Bx) \leq O(1) s^{\frac{1}{p}} \left[\nu(H) + s^{-1} \upsilon_s(Bx) \right] \tag{165}$$

[1] i.e., $t = \mathrm{Erfinv}(\delta)$ means that $\frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-p^2/2} dp = \delta$.

where $O(1)$ is an absolute constant, and $\upsilon_s(w)$ is the "s-concentration of w," that is, the sum of magnitudes $\|w[k]\|_r$ of all *but the s largest in magnitude* blocks in w.

Note that for the case of the standard representation structure, the corresponding constructions and the result in **A** were developed in [79].

B.1. Similarly to the plain and block Restricted Isometry/Eigenvalue Properties, condition $\mathbf{Q}_{s,q}$ seems to be computationally intractable – given a candidate contrast matrix H, it is difficult to verify whether it satisfies $\mathbf{Q}_{s,q}$ or not; not speaking about designing the best — with the smallest $\nu(H)$ — contrast matrix satisfying this condition, if any exists. We, however, can point out a *verifiable sufficient condition* for H to satisfy $\mathbf{Q}_{s,q}$. Specifically, we demonstrate (Proposition 5.4.3) that *H definitely satisfies $\mathbf{Q}_{s,q}$, if there exists an $N \times N$ matrix V (which we treat as a $K \times K$ block matrix with $n_k \times n_\ell$ blocks $V^{k\ell}$) such that*

$$(a): \ B = VB + H^T A, \ \ \textit{and} \ \ (b): \ \|[\|V^{1\ell}\|_{r\to r}; \|V^{2\ell}\|_{r\to r}; ...; \|V^{K\ell}\|_{r\to r}]\|_{s,q} \le \frac{s^{\frac{1}{q}-1}}{3}, \quad (166)$$

where $\|V^{k\ell}\|_{r\to r} = \max_{u^\ell \in \mathbb{R}^{n_\ell}} \left\{\|V^{k\ell}u^\ell\|_r : \|u^\ell\|_r \le 1\right\}$, and $\|u\|_{s,p}$ is the norm on $\mathbb{R}^K$ defined as follows: we zero out all but the s largest in magnitude entries in vector u, and take the $\|\cdot\|_p$-norm of the resulting vector.

One can use this sufficient condition in order to build a "suboptimal" contrast matrix, specifically, by minimizing $\nu(H)$ over pairs (V, H) satisfying the system of convex constraints (166) (provided, of course, that this system of constraints is feasible). The resulting problem is computationally tractable, provided that the norms $\|\cdot\|_{r\to r}$ are efficiently computable, which indeed is the case when $r = 1$, or $r = 2$, or $r = \infty$.

B.2. In general, the verifiable (at least for $r \in \{1, 2, \infty\}$) sufficient condition for H to satisfy $\mathbf{Q}_{s,q}$ stated in **B.1** is *not* necessary, and the condition $\mathbf{Q}_{s,q}$ itself seems to be intractable. There exists, however, a notable exception – this is the case of $q = \infty$ and $r = \infty$. We show (Proposition 5.4.1) that here the verifiable sufficient condition stated in **B.1** is *necessary and sufficient* for H to satisfy $\mathbf{Q}_{s,\infty}$. Moreover, the latter condition is "fully computationally tractable," meaning that one can optimize efficiently the quantity $\nu(H)$ over the contrast matrices H satisfying $\mathbf{Q}_{s,\infty}$, thus ending up with an optimal, as far as the error bound (165) is concerned, recovery routines. Note that when $q = \infty$, the bound

(165) holds true in the largest possible range $1 \leq p \leq \infty$ of values of p.

Note that in the case of the standard representation structure, the sufficient condition in **B.1** reduces to the verifiable sufficient condition for the validity of ℓ_1-recovery established in [81]. It should further be mentioned that to the best of our knowledge, the only known so far *verifiable* sufficient condition for the validity of *block*-ℓ_1 recovery of block-sparse signals is the "mutual block-incoherence condition" [54] (slightly extended in [56]) dealing with the case of $B = I_n$ and $r = 2$; this is a block analogy of the usual mutual incoherence condition originating from [43]. We show in Section 5.4.3.1 that the mutual block-incoherence condition is covered by the case of $B = I_n$, $r = 2$ of condition **B.1**.

B.3. As the majority of good error bounds in compressed sensing, the error bound (165) expresses a quite intuitive fact, specifically, as follows: imagine that instead of implicit observations (164) of a transform $w = Bx$, we were observing this transform directly with random error Δw such that with probability $\geq 1 - \epsilon$ one has $L_\infty(\Delta w) \leq \nu(H)$. It is easily seen that in the latter case, *in the range $\upsilon_s(Bx) \leq s\nu(H)$ of s-concentrations of the transforms $w = Bx$, the best $(1-\epsilon)$-reliable bound on the $L_p(\cdot)$-norm of the recovery error of Bx coincides, within an absolute constant factor, with the right hand side of* (165). Thus, a natural interpretation of the error bound (165) is that *as far as recovery of transforms Bx with s-concentration $\upsilon_s(Bx) \leq s\nu(H)$ is concerned, everything is as if we were given their direct observations of Bx contaminated by noise of typical L_∞-magnitude $\leq \nu(H)$*. One of the main results presented in this chapter is that to some extent, the opposite also is true, provided that $r = \infty$ and thus the error bounds in (165) holds true in the entire range $1 \leq p \leq \infty$ of values of p. Specifically, we prove (see Proposition 5.4.2) the following. *Let all the norms $\|\cdot\|_{(k)}$ be the $\|\cdot\|_\infty$-norms, and let the observation error be present (that is, either $\sigma > 0$, or $\mathcal{U}$ contains a neighborhood of the origin). Let, further, for some integer S and positive ν there exist a routine $\widehat{w}(y) \equiv B\widehat{x}(y)$ for recovering Bx from observations* (164) *such that*

$$\forall(u \in \mathcal{U}, x \in \mathbb{R}^n : S^{-1}\upsilon_S(Bx) \leq \nu) :$$
$$\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I)}\{L_\infty(B[x - \widehat{x}(Ax + u + \sigma\xi)] \leq 8[\nu + S^{-1}\upsilon_S(Bx)]\} \geq 1 - \epsilon.$$

(cf. (165) *with* $p = \infty$*). Then for every integer* s, $1 \le s \le \frac{S}{51}$ *there exists an* $N \times N$ *contrast matrix* H *and a certificate* $V = [V^{k\ell}]_{k,\ell=1}^K \in \mathbb{R}^{N\times N}$ *such that* $B = VB + H^TA$, $\|V^{k\ell}\|_{\infty\to\infty} \le \frac{1}{3s}$, $1 \le k, \ell \le K$, *and* $\nu(H) \le \nu^+ := 17\nu\frac{Erfinv(\frac{\epsilon}{2N})}{Erfinv(\frac{\epsilon}{2})}$. *In other words, when* ϵ *is small, the condition* **B.1** *is satisfied by appropriate* s, H *such that* s *and* $\nu(H)$ *coincide, within absolute constant factors, with* S *and* ν*, respectively.*

The main body of this chapter is organized as follows. Section 5.1 contains detailed problem statement, Section 5.2 describes the condition $\mathbf{Q}_{s,q}(\kappa)$, Section 5.3 presents the regular and the penalized ℓ_1 recoveries and an analysis of their performance under condition $\mathbf{Q}_{s,q}(\kappa)$. Advanced properties of this condition are subject of Section 5.4, and its relations to the RIP from traditional compressed sensing are investigated in Section 5.5. In the concluding section, we present a Block Matching Pursuit "counterpart" of the regular and penalized recoveries.

All proofs are placed in the last section of this chapter.

5.1 Problem Statement

Notation. In the sequel, we deal with

- *signals* – vectors $x = [x_1; ...; x_n] \in \mathbb{R}^n$, and a $m \times n$ *sensing matrix* A;
- *representations of signals* – block vectors $w = [w[1]; ...; w[K]] \in \mathcal{W} := \mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_K}$, and the *representation matrix* $B = [B[1]; ...; B[K]]$, $B[k] \in \mathbb{R}^{n_k\times n}$; the representation of a signal $x \in \mathbb{R}^n$ is the block vector $w = Bx$ with the blocks $B[1]x$,...,$B[K]x$.

The dimension of $\mathcal{W}$ is denoted by N where $N = n_1 + ... + n_K$. The factors $\mathbb{R}^{n_k}$ of the space $\mathcal{W}$ of representations are equipped with the norms $\|\cdot\|_{(k)}$; the conjugate norms are denoted by $\|\cdot\|_{(k,*)}$. A vector $w = [w[1]; ...; w[K]]$ from $\mathcal{W}$ is called *s-sparse*, if the number of nonzero blocks $w[k] \in \mathbb{R}^{n_k}$ in w is at most s. We refer to the collection $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ as the *representation structure*.

For $w \in \mathcal{W}$, we call the number $\|w[k]\|_{(k)}$ as the *magnitude* of k-th block in w, and denote by w^s the representation vector obtained from w by zeroing out all but the s largest in magnitude blocks in w (with the ties resolved arbitrarily). For $w \in \mathcal{W}$ and $1 \le p \le \infty$,

we denote by $L_p(w)$ the $\|\cdot\|_p$-norm of the vector $[\|w[1]\|_{(1)};...;\|w[K]\|_{(K)}]$, so that $L_p(\cdot)$ is a norm on $\mathcal{W}$ with the conjugate norm $L_p^*(w) = \|[\|w[1]\|_{(1,*)};...;\|w[K]\|_{(K,*)}]\|_{p_*}$, $p_* = \frac{p}{p-1}$. Given positive integer $s \le K$, we set $L_{s,p}(w) = L_p(w^s)$; note that $L_{s,p}(\cdot)$ is a norm on $\mathcal{W}$.

We define the s-block concentration of a vector $w = Bx$ as the sum of magnitudes $\|w[k]\|_r$ of all but the s largest in magnitude blocks in w, and denote it by $\upsilon_s(w) := L_1(w) - L_{s,1}(w)$.

Problem of interest is as follows: given an observation

$$y = Ax + u + \sigma\xi, \tag{167}$$

of unknown signal $x \in \mathbb{R}^n$, we want to recover x and the representation Bx of x, knowing in advance that this representation is s-sparse, for some given s. In (167), the term $u + \sigma\xi$ is the observation error; in this error, u is an unknown *nuisance* known to belong to a given compact convex set $\mathcal{U} \subset \mathbb{R}^m$ symmetric w.r.t. the origin, $\sigma \ge 0$ is a known noise intensity, and ξ is random noise with known distribution P.

A recovery routine is a Borel function $\widehat{x}(y) : \mathbb{R}^m \to \mathbb{R}^n$ and we characterize the performance of such a routine by its risk

$$\begin{aligned}&\mathrm{Risk}_p^B(\widehat{x}(\cdot)|s,\sigma,\upsilon,\epsilon)\\ &= \inf\Big\{d : \mathrm{Prob}_{\xi\sim P}\left\{\xi : L_p\left(B[\widehat{x}(Ax+u+\sigma\xi)-x]\right) \le d \ \ \forall(u\in\mathcal{U}, x\in\mathbb{R}^n : \upsilon_s(Bx)\le\upsilon)\right\}\\ &\qquad\qquad \ge 1-\epsilon\Big\};\end{aligned} \tag{168}$$

here $0 \le \epsilon \le 1$ and $1 \le p \le \infty$. Thus, $\mathrm{Risk}_p^B(\widehat{x}(\cdot)|s,\sigma,\upsilon,\epsilon) \le d$ if and only if there exists a set $\Xi \in \mathbb{R}^m$ such that $P(\Xi) \ge 1-\epsilon$ and $L_p\left(B[\widehat{x}(Ax+u+\sigma\xi)-x]\right) \le d$ whenever $\xi \in \Xi$, $u \in \mathcal{U}$ and whenever $x \in \mathbb{R}^n$ is such that Bx can be approximated by s-sparse representation vector within accuracy υ (measured in $L_1(\cdot)$), i.e., $\upsilon_s(Bx) \le \upsilon$.

5.2 Condition $\mathbf{Q}_{s,q}(\kappa)$

Let a sensing matrix A and a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ be given, and let $s \le K$ be a positive integer, $q \in [1,\infty]$ and $\kappa > 0$. We say that a pair $(H, \|\cdot\|)$,

where $H \in \mathbb{R}^{m\times M}$ and $\|\cdot\|$ is a norm on $\mathbb{R}^M$, satisfies the condition $\mathbf{Q}_{s,q}(\kappa)$ associated with the matrices A, B and the representation structure, if

$$\forall x \in \mathbb{R}^n : \; L_{s,q}(Bx) \le s^{\frac{1}{q}}\|H^T Ax\| + \kappa s^{\frac{1}{q}-1} L_1(Bx). \tag{169}$$

The following is an evident observation

Observation 5.2.1 *Given A and a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$, let $(H, \|\cdot\|)$ satisfy $\mathbf{Q}_{s,q}(\kappa)$. Then $(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,q'}(\kappa')$ for all $q' \in (1, q)$ and $\kappa' \ge \kappa$. Besides this, if $s' \le s$ is a positive integer, $((s/s')^{\frac{1}{q}} H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s',q}((s'/s)^{1-\frac{1}{q}}\kappa)$.*

Whenever $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ is the standard representation structure, meaning that B is the identity matrix and $n_1 = ... = n_K = 1$ and $\|\cdot\|_{(k)} = |\cdot|$ for all k, the condition $\mathbf{Q}_{s,q}(\kappa)$ reduces to the condition $\mathbf{H}_{s,q}(\kappa)$ introduced in [79].

5.3 Recovery Routines

We are about to introduce two new recovery routines.

Regular L_1-recovery is

$$\widehat{x}_{\rm reg}(y) \in \operatorname*{Argmin}_u \left\{ L_1(Bu) : \|H^T(Au - y)\| \le \rho \right\}, \tag{170}$$

where $H \in \mathbb{R}^{m\times M}$, $\|\cdot\|$ and $\rho > 0$ are parameters of the construction.

Theorem 5.3.1 *Let s be a positive integer, $q \in [1, \infty]$ and $\kappa \in (0, 1/2)$. Let also $\epsilon \in (0, 1)$. Assume that the parameters H, $\|\cdot\|$, ρ of the regular L_1-recovery are such that*

A. *$(H, \|\cdot\|)$ satisfies the condition $\mathbf{Q}_{s,q}(\kappa)$ associated with matrices A, B;*

B. *There exists a set Ξ such that $P(\Xi) \ge 1 - \epsilon$ and*

$$\|H^T(u + \sigma\xi)\| \le \rho \; \forall (u \in \mathcal{U}, \xi \in \Xi). \tag{171}$$

Then for $1 \le p \le q$ and $\forall (\xi \in \Xi, u \in \mathcal{U}, x \in \mathbb{R}^n)$,

$$L_p(B[\widehat{x}_{\rm reg}(Ax + u + \sigma\xi) - x]) \le (4s)^{\frac{1}{p}} \frac{2\rho + s^{-1} L_1(Bx - [Bx]^s)}{1 - 2\kappa}. \tag{172}$$

Penalized L_1-recovery is

$$\widehat{x}_{\rm pen}(y) \in \operatorname*{Argmin}_u \left\{ L_1(Bu) + 2s\|H^T(Ax-y)\| \right\}, \tag{173}$$

where $H \in \mathbb{R}^{m\times M}$, $\|\cdot\|$ and a positive integer s are parameters of the construction.

Theorem 5.3.2 *Let s be a positive integer, $q \in [1,\infty]$ and $\kappa \in (0,1/2)$. Let also $\epsilon \in (0,1)$. Assume that the parameters H, $\|\cdot\|$, s of the penalized recovery and a $\rho \geq 0$ satisfy conditions* **A**, **B** *from Theorem 5.3.1. Then, similar to Theorem 5.3.1, we have*

$$L_p(B[\widehat{x}_{\rm pen}(Ax+u+\sigma\xi)-x]) \leq 2(2s)^{\frac{1}{p}} \frac{2\rho + s^{-1}L_1(Bx-[Bx]^s)}{1-2\kappa}, \tag{174}$$

for $1 \leq p \leq q$ and $\forall(\xi \in \Xi, u \in \mathcal{U}, x \in \mathbb{R}^n)$ cf. (172).

5.4 Properties of Condition $\mathbf{Q}_{s,\infty}(\kappa)$

In general, given a sensing matrix A and a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ...\|\cdot\|_{(K)})$, it seems to be difficult even to verify that a pair $(H, \|\cdot\|)$ satisfies condition $\mathbf{Q}_{s,q}(\kappa)$ associated with A, B, not speaking about synthesis of $(H, \|\cdot\|)$ satisfying this condition and resulting in the best possible error bounds (172), (174) for the regular and the penalized ℓ_1 recoveries. We are about to demonstrate that when all $\|\cdot\|_{(k)}$ are the uniform norms $\|\cdot\|_\infty$ and, in addition, $q = \infty$ (which, by Observation 5.2.1, corresponds to the strongest among the conditions $\mathbf{Q}_{s,q}(\kappa)$ associated with A and a given representation structure and ensures the validity of (172), (174) in the largest possible range $1 \leq p \leq \infty$ of values of p), the condition $\mathbf{Q}_{s,q}(\kappa)$ becomes "fully computationally tractable." We intend to demonstrate also that this condition $\mathbf{Q}_{s,\infty}(\kappa)$ is in fact necessary for the bounds of the form (172), (174) to be valid when $p = \infty$.

5.4.1 Condition $\mathbf{Q}_{s,\infty}(\kappa)$: Tractability

In the sequel, given $\pi, \theta \in [1,\infty]$ and a matrix M, we denote by $\|M\|_{\pi\to\theta}$ the norm of the linear operator $u \mapsto Mu$ induced by the norms $\|\cdot\|_\pi$ and $\|\cdot\|_\theta$ at the origin and the destination spaces:

$$\|M\|_{\pi\to\theta} = \max_{u:\|u\|_\pi \leq 1} \|Mu\|_\theta.$$

Consider the situation where $\|\cdot\|_{(k)}$, for all k, are the ℓ_∞ norms. We claim that in this case the condition $\mathbf{Q}_{s,\infty}(\kappa)$ becomes fully tractable. Specifically, we have the following

Proposition 5.4.1 *Let $\|\cdot\|_{(k)} = \|\cdot\|_\infty$ for all $k \le K$, and let a positive integer s and reals $\kappa > 0$, $\epsilon \in (0,1)$ be given.*

(i) *Assume that a triple $(H, \|\cdot\|, \rho)$, where $H \in \mathbb{R}^{M\times m}$, $\|\cdot\|$ is a norm on $\mathbb{R}^M$, and $\rho \ge 0$, is such that*

> (!) *$(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,\infty}(\kappa)$ and the set $\Xi = \{\xi : \|H^T[u + \sigma\xi]\| \le \rho \ \forall u \in \mathcal{U}\}$ is such that $P(\Xi) \ge 1 - \epsilon$.*

Then there exist $N = n_1 + ... + n_K$ vectors $h^1, ..., h^N$ in $\mathbb{R}^m$ and $N \times N$ matrix $V = [V^{k\ell}]_{k,\ell=1}^K$ (the blocks $V^{k\ell}$ are $n_k \times n_\ell$ matrices) such that

$$\begin{array}{ll} (a) & B = VB + [h^1, ..., h^N]^T A, \\ (b) & \|V^{k\ell}\|_{\infty\to\infty} \le s^{-1}\kappa \quad \forall k, \ell \le K, \\ (c) & P\left(\Xi^+ := \{\xi : \max\limits_{u\in\mathcal{U}} u^T h^i + \sigma|\xi^T h^i| \le \rho,\ 1 \le i \le N\}\right) \ge 1 - \epsilon. \end{array} \tag{175}$$

(ii) *Whenever vectors $h^1, ..., h^N \in \mathbb{R}^m$ and a matrix $V = [V^{k\ell}]_{k,\ell=1}^K$ with $n_k \times n_\ell$ blocks $V^{k\ell}$ satisfy (175), the $m \times N$ matrix $\widehat{H} = [h^1, ..., h^N]$, the norm $\|\cdot\|_\infty$ on $\mathbb{R}^N$ and ρ form a triple satisfying (!).*

Discussion. Let a sensing matrix $A \in \mathbb{R}^{m\times n}$ and a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ with all $\|\cdot\|_{(k)}$ being the ℓ_∞ norms be given, along with a positive integer s, an uncertainty set $\mathcal{U}$, a distribution P of ξ and quantities σ, ϵ. Theorems 5.3.1, 5.3.2 say that if a triple $(H, \|\cdot\|, \rho)$ is such that $(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,\infty}(\kappa)$ with $\kappa < 1/2$ and H, ρ are such that for the set

$$\Xi = \{\xi : \ \|H^T[u + \sigma\xi]\| \le \rho \ \forall u \in \mathcal{U}\}$$

it holds $P(\Xi) \ge 1 - \epsilon$, then for all $\upsilon \ge 0$, for the regular L_1-recovery associated with $(H, \|\cdot\|, \rho)$ and for the penalized L_1-recovery associated with $(H, \|\cdot\|, s)$ the following holds:

$$\mathrm{Risk}_p^B(\widehat{x}|s, \sigma, \upsilon, \epsilon) \le 2(2s)^{\frac{1}{p}} \frac{2\rho + s^{-1}\upsilon}{1 - 2\kappa}, \ 1 \le p \le \infty \tag{176}$$

for all x such that $v_s(Bx) \leq v$. Proposition 5.4.1 states that when applying this result, we loose nothing by restricting ourselves with triples $H = [h^1, ..., h^N] \in \mathbb{R}^{m\times N}$, $N = n_1 + ... + n_K$, $\|\cdot\| = \|\cdot\|_\infty$, $\rho \geq 0$ which can be augmented by an appropriately chosen matrix $N \times N$ matrix V to satisfy relations (175). In the rest of this discussion, it is assumed that we are speaking about triples $(H, \|\cdot\|, \rho)$ satisfying the just defined restrictions.

Now, as far as bounds (176) are concerned, they are completely determined by two parameters — κ (which should be $< 1/2$) and ρ; the smaller are these parameters, the better are the bounds. In what follows we address the issue of efficient synthesis of matrices H with "as good as possible" values of κ and ρ.

Observe, first, that $H = [h^1, ..., h^N]$ and κ should admit an extension by a matrix V to a solution of the system of convex constraints (175). In the case of $\sigma = 0$, the best choice of ρ, given H, is

$$\rho = \max_i \mu_{\mathcal{U}}(h^i), \text{ where } \mu_{\mathcal{U}}(h) = \max_{u\in\mathcal{U}} u^T h.$$

Consequently, in this case the "achievable pairs" ρ, κ form a computationally tractable convex set

$$G_s = \left\{(\kappa, \rho) : \begin{array}{l} \exists H = [h^1, ..., h^N] \in \mathbb{R}^{m\times N}, V = [V^{k\ell} \in \mathbb{R}^{n_k\times n_\ell}]_{k,\ell=1}^K : \\ \qquad B = VB + H^T A,\ \|V^{k\ell}\|_{\infty\to\infty} \leq \frac{\kappa}{s},\ \mu_{\mathcal{U}}(h^i) \leq \rho, 1 \leq i \leq N \end{array}\right\}.$$

When $\sigma > 0$, the situation is complicated by the necessity to maintain the validity of the restriction

$$P(\Xi^+) := P\left\{\xi : \mu_{\mathcal{U}}(h^i) + \sigma|\xi^T h^i| \leq \rho,\ 1 \leq i \leq N\right\} \geq 1 - \epsilon, \tag{177}$$

which is a chance constraint in variables $h^1, ..., h^N, \rho$ and as such can be "computationally intractable." Let us consider the "most standard" case of Gaussian zero mean noise ξ, that is, assume that $\xi = C\eta$ with $\eta \sim \mathcal{N}(0, I_m)$. Then (177) implies that

$$\rho \geq \max_i \left[\mu_{\mathcal{U}}(h^i) + \sigma \text{Erfinv}(\frac{\epsilon}{2})\|C^T h^i\|_2\right]$$

and is implied by

$$\rho \geq \max_i \left[\mu_{\mathcal{U}}(h^i) + \sigma \text{Erfinv}(\frac{\epsilon}{2N})\|C^T h^i\|_2\right],\ 1 \leq i \leq N.$$

Ignoring the "gap" between Erfinv$(\epsilon/2)$ and Erfinv$(\frac{\epsilon}{2N})$, we can safely model the restriction (177) by the system of convex constraints

$$\mu_{\mathcal{U}}(h^i) + \sigma \mathrm{Erfinv}(\frac{\epsilon}{2N})\|C^T h^i\|_2 \le \rho,\ 1 \le i \le N. \tag{178}$$

Thus, the set G_s of admissible κ, ρ can be safely approximated by the computationally tractable convex set

$$G_s^* = \left\{(\kappa,\rho) : \exists \left[\begin{array}{l} H = [h^1, ..., h^N] \in \mathbb{R}^{m\times N} \\ V = [V^{k\ell} \in \mathbb{R}^{n_k \times n_\ell}]_{k,\ell=1}^K \end{array}\right] : \left\{\begin{array}{l} B = BV + H^T A,\ \|V^{k\ell}\|_{\infty\to\infty} \le \frac{\kappa}{s}, 1 \le k,\ell \le K \\ \max\limits_{u\in\mathcal{U}} u^T h^i + \sigma \mathrm{Erfinv}\left(\frac{\epsilon}{2N}\right)\|C^T h^i\|_2 \le \rho, \\ \qquad\qquad\qquad\qquad 1 \le i \le N \end{array}\right\}. \tag{179}$$

5.4.2 Condition $\mathbf{Q}_{s,\infty}(\kappa)$: Necessity

In addition to the earlier assumption that for all k, $\|\cdot\|_{(k)}$ are the ℓ_∞ norms, we now assume ξ is a zero mean Gaussian noise: $\xi = C\eta$ with $\eta \sim \mathcal{N}(0, I_m)$. From the above discussion we know that if, for some $\kappa < 1/2$ and $\rho > 0$, there exist $H = [h^1, ..., h^N] \in \mathbb{R}^{m\times N}$ and $V = [V^{k\ell} \in \mathbb{R}^{n_k\times n_\ell}]_{k,\ell=1}^K$ satisfying (175), then regular and penalized ℓ_1 recoveries with appropriate choice of parameters ensure that

$$\mathrm{Prob}\{\|B[x - \widehat{x}(Ax + u + \sigma\xi)]\|_\infty \le 2\frac{2\rho + s^{-1}L_1(Bx - [Bx]^s)}{1 - 2\kappa}\} \ge 1 - \epsilon \tag{180}$$

for all $(x \in \mathbb{R}^n, u \in \mathcal{U})$.

We are about to demonstrate that this implication can be "nearly inverted:"

Proposition 5.4.2 *Let a sensing matrix A, a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$, the uncertainty set $\mathcal{U}$, and reals $\kappa > 0$, $\epsilon \in (0, 1)$ be given, with all $\|\cdot\|_{(k)}$ being ℓ_∞-norms. Assume also that the observation error "is present," specifically, that for every $r > 0$, the set $\{u + \sigma Ce : u \in \mathcal{U}, \|e\|_2 \le r\}$ contains a neighborhood of the origin.*

Given a positive integer S, assume that there exists a recovering routine $\widehat{x}$ satisfying an error bound of the form (180), specifically, the bound

$$\forall (x \in \mathbb{R}^n, u \in \mathcal{U}) : \mathrm{Prob}\{\|B[x - \widehat{x}(Ax + u + \sigma\xi)]\|_\infty \le \alpha + cS^{-1}L_1(Bx - [Bx]^S)\} \ge 1 - \epsilon. \tag{181}$$

where α and c are some positive constants. Then for every $s \le \frac{S}{6.3c}$ there exist $H = [h^1, ..., h^N] \in \mathbb{R}^{m\times N}$ and $V = [V^{k\ell} \in \mathbb{R}^{n_k\times n_\ell}]_{k,\ell=1}^K$ such that $h^i \in \mathbb{R}^m$, $v^i \in \mathbb{R}^k$, $1 \le i \le k$, satisfying

$$\begin{array}{ll}
(a) & B = VB + H^T A, \\
(b) & \|V^{k\ell}\|_{\infty\to\infty} \le \frac{1}{3s} \quad \forall k, \ell \le K, \\
(c) & \text{with } \rho := \max\limits_{1\le i\le N}\left[\max\limits_{u\in\mathcal{U}} u^T h^i + \sigma Erfinv(\frac{\epsilon}{2N})\|C^T h^i\|_2\right] \le 2.1\alpha\frac{Erfinv(\frac{\epsilon}{2N})}{Erfinv(\epsilon)} \\
 & \text{and } \xi = C\eta,\ \eta \sim \mathcal{N}(0, I_m), \\
 & \text{one has} \\
 & P\left(\Xi^+ := \{\xi : \max\limits_{u\in\mathcal{U}} u^T h^i + \sigma|\xi^T h^i| \le \rho,\ 1 \le i \le N\}\right) \ge 1 - \epsilon,
\end{array} \tag{182}$$

meaning that (see Proposition 5.4.1) $(H, \|\cdot\|_\infty)$ satisfies $\mathbf{Q}_{s,\infty}(1/3)$ for s "nearly as large as S," namely, $s \le \frac{S}{6.3c}$, and $H = [h^1, ..., h^k]$, ρ satisfy conditions (178) (and thus – condition **B** *from Theorem 5.3.1), with ρ being "nearly α', namely, $\rho \le 2.1\alpha\frac{Erfinv(\frac{\epsilon}{2N})}{Erfinv(\epsilon)}$.*

5.4.3 A Sufficient Condition for $\mathbf{Q}_{s,q}(\kappa)$

Proposition 5.4.3 *Let a sensing matrix A, a representation structure $(B, n_1, ..., n_K, \|\cdot\|_{(1)}, ..., \|\cdot\|_{(K)})$ be given. Let $N = n_1 + ... + n_K$, and let $N \times N$ matrix $V = [V^{k\ell}]_{k,\ell=1}^K$ ($V^{k\ell}$ are $n_k \times n_\ell$) and $m \times N$ matrix H satisfy the relation*

$$B = VB + H^T A. \tag{183}$$

Let

$$\nu^*_{s,q}(V) = \max_{1\le\ell\le K}\ \max_{w^\ell\in\mathbb{R}^{n_\ell}:\|w^\ell\|_{(\ell)}\le 1} L_{s,q}\left([V^{1\ell}w^\ell; ...; V^{K\ell}w^\ell]\right). \tag{184}$$

Then for all $s \le K$ and all $q \in [1, \infty]$, we have:

$$L_{s,q}(Bx) \le s^{\frac{1}{q}} L_\infty(H^T Ax) + \nu^*_{s,q}(V) L_1(Bx)\ \forall x. \tag{185}$$

In particular, whenever there is an upper bounding function $\nu_{s,q}(V)$ satisfying $\nu^*_{s,q}(V) \le \nu_{s,q}(V)$ for all $s \le K$, $q \in [1, \infty]$ and matrix V such that

$$\nu_{s,q}(V) \le s^{\frac{1}{q}-1}\kappa, \tag{186}$$

holds, then the pair $(H, L_\infty(\cdot))$ satisfies $\mathbf{Q}_{s,q}(\kappa)$.

Let $\mathrm{Col}_k(\Omega)$ *is* k*-th column of* Ω *and* $\|u\|_{s,q}$ *is the* ℓ_q*-norm of the vector obtained from* u *by zeroing all but the* s *largest entries in* u *and* $\|P\|_{(\ell)\to(k)}$ *is the norm of the linear mapping* $u \mapsto Pu : \mathbb{R}^{n_\ell} \to \mathbb{R}^{n_k}$ *induced by the norms* $\|\cdot\|_{(\ell)}$, $\|\cdot\|_{(k)}$ *on the argument and on the image spaces. Then*

$$\nu^*_{s,q}(V) \le \widehat{\nu}_{s,q}(V) = \max_{1\le k\le K} \|\mathrm{Col}_k[\Omega]\|_{s,q} \quad \textit{where } \Omega = [\|V^{k\ell}\|_{(\ell)\to(k)}]_{k,\ell=1}^K. \tag{187}$$

When all $\|\cdot\|_{(k)}$ are the ℓ_∞-norms and $q = \infty$, Proposition 5.4.3 recovers Proposition 5.4.1. In the general case, Proposition 5.4.3 suggests a way to synthesize matrices $H \in \mathbb{R}^{m\times N}$ which provably satisfy the condition $\mathbf{Q}_{s,q}(\kappa)$, along with a certificate V for this fact: H and V should satisfy a system of linear equations (183) and, in addition, V should satisfy (187) and (186), which is a system of convex constraints on V. Whenever these constraints are efficiently computable, we get a computationally tractable sufficient condition on H to satisfy $\mathbf{Q}_{s,q}(\kappa)$ – a condition which is expressed by an explicit system of efficiently computable convex constraints on H and additional matrix variable V. Now, efficient computability of the constraints (186) is the same as efficient computability of norms $\|\cdot\|_{(k)\to(\ell)}$. Assuming that $\|\cdot\|_{(k)} = \|\cdot\|_{\pi_k}$ for every k, the computability issue becomes the one of efficient computation of the norms $\|\cdot\|_{\pi_\ell\to\pi_k}$. The norm $\|\cdot\|_{\pi\to\theta}$ is known to be generically efficiently computable in just three cases:

1. $\theta = \infty$, where $\|M\|_{\pi\to\infty} = \|M^T\|_{1\to\frac{\pi}{\pi-1}} = \max_i \|\mathrm{Row}_i^T(M)\|_{\frac{\pi}{\pi-1}}$, where $\mathrm{Row}_i(M)$ is i-th row of M;
2. $\pi = 1$, where $\|M\|_{1\to\theta} = \max_j \|\mathrm{Col}_j(M)\|_\theta$;
3. $\pi = \theta = 2$, where $\|M\|_{2\to 2}$ is the usual spectral norm of a matrix M.

Assuming for the sake of simplicity that in our representation structure $\|\cdot\|_{(k)}$ are π-norms with common value of π, let us look at three "tractable cases" as specified by the above discussion – those of $\pi = \infty$, $\pi = 1$ and $\pi = 2$.

The case of $\pi = \infty$ was considered in full details in Section 5.4.1. In this case, we have

$$\nu_{s,q|\pi=\infty}(V) = \max_{1\le k,\ell\le K} \|V^{kl}\|_{\infty\to\infty}.$$

The case of $\pi = 1$. When all $\|\cdot\|_{(k)}$ are the ℓ_1 norms, the quantity $\nu^*_{s,q}(V)$ is easy to compute:

$$\nu_{s,q|\pi=1}(V) = \nu^*_{s,q}(V) = \max_{1\le k\le K} L_{s,q}(\mathrm{Col}_k(V)).$$

The bottom line is that *when $\|\cdot\|_{(k)}$ for all k are ℓ_1 norms, a verifiable sufficient condition of an $m \times N$ matrix H and the norm $L_\infty(\cdot)$ on $\mathbb{R}^N$ to satisfy $\mathbf{Q}_{s,q}(\kappa)$ is the existence of an $N \times N$ matrix V such that*

$$B = VB + H^T A \ \& \ \max_{1\le k\le K} L_{s,q}(\mathrm{Col}_k(V)) \le s^{-1}\kappa, \tag{188}$$

which is a system of efficiently computable convex constraints in variables V, H.

The case of $\pi = 2$. Now assume that all $\|\cdot\|_{(k)}$ are ℓ_2 norms, one has

$$\nu^*_{s,q}(V) = \max_{1\le\ell\le K} \max_{w\in\mathbb{R}^{n_\ell}:\|w\|_2\le 1} \|[\|V^{1\ell}w\|_2; \|V^{2\ell}w\|_2; ...; \|V^{K\ell}w\|_2]\|_{s,q}. \tag{189}$$

In order to convert (189) into a verifiable sufficient condition for H to satisfy $\mathbf{Q}_{s,q}(\kappa)$, we need an efficiently computable and convex in V upper bound on the quantity $\nu^*_{s,q}(V)$. To this end it suffices to find an efficiently computable upper bound on the function $\mu_{s,q}(U)$ of a block matrix $U = [U^1; ...; U^K]$ with $n_k \times q$ blocks U^k defined as follows:

$$\mu_{s,q}(U) = \max_{w\in\mathbb{R}^q:\|w\|_2\le 1} \|[\|U^1w\|_2; \|U^2w\|_2; ...; \|U^Kw\|_2]\|_{s,q}.$$

A trivial efficiently computable upper bound on $\mu_{s,q}(U)$ is $\|[\|U^1\|_{2\to2}; ...; \|U^K\|_{2\to2}]\|_{s,q}$; this bound brings us back to the function $\widehat{\nu}_{s,q}(V)$. Note that this bound is exact when $q = \infty$ (same as when $s = 1$, since $\mu_{1,q}(U) = \mu_{K,\infty}(U) = \max_k \|U^k\|_{2,2}$). A less trivial bound can be derived when $q = 1$.

The case of $\pi = 2$, $q = 1$. Let $\mathcal{Z} = \{z \in \mathcal{W} : \|z[k]\|_2 \le 1\, \forall k \le K, \sum_{k=1}^K \|z[k]\|_2 \le s\}$, so that for every vector $w \in \mathcal{W}$ we have $L_{s,1}(w) = \max_{z\in\mathcal{Z}} z^T w$. We have

$$\begin{aligned}
\mu^2_{s,1}(U) &= \max_{w:\|w\|_2\le 1} \|[\|U^1w\|_2; \|U^2w\|_2; ...; \|U^Kw\|_2]\|^2_{s,1} = \max_{w:\|w\|_2\le 1} \left[\max_{z\in\mathcal{Z}} z^TUw\right]^2 \\
&= \max_{w:\|w\|_2\le 1} \max_{z\in\mathcal{Z}} w^TU^Tzz^TUw = \max_{z\in\mathcal{Z}} \left[\max_{w:\|w\|_2\le 1} w^TU^Tzz^TUw\right] \\
&= \max_{z\in\mathcal{Z}} \mathrm{Tr}(U^Tzz^TU) \quad [\text{since } U^Tzz^TU \text{ is of rank 1}]
\end{aligned}$$

whence

$$\mu_{s,1}(U) = \max_{z \in \mathcal{Z}} \sqrt{\mathrm{Tr}(U^T z z^T U)}. \tag{190}$$

Now, when $z \in \mathcal{Z}$, the matrix $Z = zz^T$ is a block matrix: $Z = [Z^{k\ell}]_{k,\ell=1}^K$ with $n_k \times n_\ell$ blocks $Z^{k\ell} = z[k](z[\ell])^T$, and $\|z[k]\|_2 \le 1$ for all k, $\sum_k \|z[k]\|_2 \le s$. As a result, Z possesses the following properties:

$$\begin{array}{ll} (a) & Z = Z^T \succeq 0 \\ (b) & \exists t_k, 1 \le k \le K : \underbrace{\left\{ \begin{array}{l} \|Z^{k\ell}\|_* \le t_k, 1 \le k, \ell \le K \\ \sum_{\ell=1}^K \|Z^{k\ell}\|_* \le s t_k, \ 1 \le k \le K \\ t_k \le 1, \ 1 \le k \le K, \ \sum_{k=1}^K t_k \le s \end{array} \right.}_{(*)} \end{array} \tag{191}$$

where $\|\cdot\|_*$ is the nuclear norm (the sum of singular values) of a matrix.

Indeed, (a) is evident; to ensure $(*)$, it suffices to set $t_k = \|z[k]\|_2$, $1 \le k \le K$.

Now let $\mathcal{Z}^+$ be the set of all $N \times N$ matrices $Z = [Z^{k\ell}]_{k,\ell=1}^K$ satisfying (191), and let

$$\widehat{\mu}_{s,1}(U) = \sqrt{\max_{Z \in \mathcal{Z}^+} \mathrm{Tr}(U^T Z U)} \tag{192}$$

Observing that $\widehat{\mu}_{s,1}(U)$ is an upper bound on $\mu_{s,1}(U)$ (by (190) and due to $zz^T \in \mathcal{Z}^+$ when $z \in \mathcal{Z}$) and it is a convex efficiently computable function of U [2]; we have reached our goal – building an efficiently computable upper bound on $\mu_{s,1}(\cdot)$. It is easily seen that this bound is never worse than the simpler bound $\widetilde{\mu}_{s,1}(U) := \|[\|U^1\|_{2\to 2}; ...; \|U^K\|_{2\to 2}]\|_{s,1}$;[3] and numerical experiment shows that the ratio $\widetilde{\mu}/\widehat{\mu}$ can be quite significant. □

---

[2] Convexity follows from the fact that with $Z \succeq 0$, $\sqrt{\mathrm{Tr}(U^T Z U)}$ is the Frobenius norm of $\|Z^{1/2}U\|$, computability - from the fact that $\mathcal{Z}^+$ is a computationally tractable convex set, so that maximizing $\mathrm{Tr}(U^T Z U)$ over $\mathcal{Z}^+$ is a tractable task.

[3] Here is the justification: Let $Z \in \mathcal{Z}^+$ and t_ℓ be such that (191.b) takes place, and let $u_k = \|U^k\|_{2\to 2}$. Observe that

> (!) whenever $\tau \in \mathbb{R}^K$ and $\sigma \in \mathbb{R}$ are such that $|\tau_\ell| \le \sigma$, $1 \le \ell \le K$, and $\sum_\ell |\tau_\ell| \le s\sigma$, we have $|\tau^T w| \le \sigma \|w\|_{s,1}$ for all w.

We now have

$$\begin{array}{l} \mathrm{Tr}(U^T Z U) = \sum_{k,\ell=1}^K \mathrm{Tr}([U^k]^T Z^{k\ell} U^\ell) = \sum_{k,\ell=1}^k \mathrm{Tr}(Z^{k\ell} U^\ell [U^k]^T) \le \sum_{k,\ell=1}^K \|Z^{k\ell}\|_* \|U^k [U^\ell]^T\|_{2\to 2} \\ \le \sum_{k,\ell=1}^K \|Z^{k\ell}\|_* u_k u_\ell = \sum_{k=1}^K u_k \left[\sum_{\ell=1}^K \|Z_*^{k\ell}\| u_\ell\right] \le_{(a)} \sum_{k=1}^K u_k t_k \|u\|_{s,1} \le_{(b)} \|u\|_{s,1}^2 \end{array}$$

where $\le_{(a)}$ and $\le_{(b)}$ are given by (!) combined with (191.b).

5.4.3.1 *Proposition 5.4.3 and Mutual Block-Incoherence*

We have mentioned in Introduction section that, to the best of our knowledge, the only previously proposed verifiable sufficient condition for the validity of block-ℓ_1 recovery is the "mutual block incoherence condition" [54]; our local goal is to demonstrate that this condition is covered by Proposition 5.4.3.

The mutual block incoherence condition deals with the case where $B = I_n$, $n_k = d$, $\|\cdot\|_{(k)} = \|\cdot\|_2$, $1 \le k \le K$, and the columns of A are normalized to have Euclidean lengths equal to 1. The condition is as follows: let $A[\ell]$ be consecutive $m \times d$ submatrices of $m \times n$ matrix A. We set

$$\alpha = \max_{1\le\ell\le K} \max_{1\le j<j'\le d} |\mathrm{Col}_j^T(A[\ell])\mathrm{Col}_{j'}(A[\ell])|, \ \mu = \max_{1\le k<\ell\le K} \sigma_{\max}(A^T[k]A[\ell]) \tag{193}$$

where $\sigma_{\max}$ stands for the maximal singular value and $\mathrm{Col}_j(A)$ denotes the j-th column of the matrix A. [54] states that under the condition

$$s < \frac{1-(d-1)\alpha+\sigma}{2\sigma} \tag{194}$$

the block-ℓ_1/ℓ_2 recovery $\widehat{x}(y) \in \mathrm{Argmin}_z \left\{\sum_{\ell=1}^K \|z^\ell\|_2 : Az = y\right\}$ (z^ℓ are consecutive d-dimensional blocks in $z \in \mathbb{R}^n$) in the case of noiseless observations is exact on all s-block-sparse signals x. In the case of noisy observations, certain error bound for a (properly modified) block-ℓ_1/ℓ_2 recovery of nearly s-block-sparse signals is provided. We are about to demonstrate that these results of [54], are, essentially, covered by Proposition 5.4.3 and Theorems 5.3.1, 5.3.2 due to the following observation:

Proposition 5.4.4 *Given $m \times n$ sensing matrix A with unit Euclidean lengths of the columns and the representation structure $B = I_n$, $n_k = d$, $\|\cdot\|_{(k)} = \|\cdot\|_2$, $1 \le k \le K = n/d$, let $\alpha, \sigma > 0$ be defined according to (193), and let a positive integer s be such that (194) holds true (the latter clearly implies that $1-(d-1)\alpha \ge \sigma$). Then the matrices $B_\ell = A^T[\ell]A[\ell]$ are positive definite for all ℓ, and setting*

$$H = \theta[B_1^{-1}A[1], B_2^{-1}A[2], ..., B_K^{-1}A[K]], \quad \theta = \frac{1-(d-1)\alpha}{1-(d-1)\alpha+\sigma}$$

we get a contrast matrix which, along with the norm $L_\infty(\cdot)$, satisfies the condition $\mathbf{Q}_{s,\infty}(\kappa)$ with

$$\kappa = \frac{\sigma s}{1-(d-1)\alpha+\sigma} < 1/2.$$

The Euclidean lengths of the columns in H do not exceed $\frac{1}{1-(d-1)\rho}$.

5.4.3.2 *Sufficient Condition for $\mathbf{Q}_{s,q}(\kappa)$: Limits of Performance*

Consider the situation where all the norms $\|\cdot\|_{(k)}$ are $\|\cdot\|_\pi$, with $\pi \in \{1,2,\infty\}$. Here Proposition 5.4.3 offers a verifiable sufficient condition for a pair $(H \in \mathbb{R}^{m\times N}, L_\infty(\cdot))$ to satisfy the condition $\mathbf{Q}_{s,q}(\kappa)$. A natural question is, what are "limits of performance" of this sufficient condition, specifically, how large could be the values of s for which the condition can be satisfied by at least one contrast matrix. Here is a partial answer to this question:

Proposition 5.4.5 *Let A be an $m\times n$ sensing matrix with $m<n$, let $B=I_n$ and let $n_k=d$, $\|\cdot\|_{(k)}=\|\cdot\|_\pi$, $1\le k\le K=n/d$, with $\pi\in\{1,2,\infty\}$. Whenever an $m\times n$ matrix H and $n\times n$ matrix V satisfy the conditions*

$$I_n = V + H^T A \text{ and } \max_{1\le\ell\le K} \|[\|V^{1\ell}\|_{\pi\to\pi}; \|V^{2\ell}\|_{\pi\to\pi}; ...; \|V^{K\ell}\|_{\pi\to\pi}]\|_{s,q} \le s^{\frac{1}{q}-1}/2 \qquad (195)$$

(cf. (183), (187), and (186)), one has

$$s \le \frac{n}{2\sqrt{d(n-m)}} \qquad (196)$$

provided that either (a) $q\ge 2$, or (b) $q=1$ and A is "essentially non-square," namely, $m<3n/4$.

Discussion. Let the representation structure in question be the same as in Proposition 5.4.5, and let $m\times n$ sensing matrix A which is "sufficiently non-square," that is, $m\le\gamma n$ for some $\gamma<1$. Proposition 5.4.5 says that in this case, verifiable sufficient condition, stated by Proposition 5.4.3, for satisfiability of $\mathbf{Q}_{s,q}(\kappa)$ with $\kappa<1/2$ has rather restricted scope — it cannot certify the satisfiability of $\mathbf{Q}_{s,q}(\kappa)$, $\kappa\le 1/2$, when $s>\frac{\sqrt{n}}{2\sqrt{d(1-\gamma)}}$, at least in the case of $q\ge 2$; when $q=1$, the conclusion still holds true provided that $\gamma\le 3/4$). Note that in fact the condition $\mathbf{Q}_{s,q}(\kappa)$ can be satisfiable in a much larger range of values

of s; e.g., when the representation structure in question is the standard one, and A is a random Gaussian $m \times n$ matrix, the matrix $2A$ satisfies, with overwhelming probability as m, n grow, the condition $\mathbf{Q}_{s,2}(1/3)$ for s as large as $O(1)m/\sqrt{\ln(2n/m)}$, see Proposition 5.5.1 below. There is, however, an important case where the "limits of performance" of our *verifiable sufficient condition* for the satisfiability of $\mathbf{Q}_{s,q}(\kappa)$ implies severe restrictions on the range of values of s in which the "true" condition $Q_{s,q}(\kappa)$ is satisfiable – this is the case when $q = \infty$ and $\pi = \infty$. Combining Propositions 5.4.1 and 5.4.5, we conclude that *in the case of representation structure from Proposition 5.4.5 with $\pi = \infty$ and "sufficiently non-square" ($m \leq 3n/4$) $m \times n$ sensing matrix A, the associated condition $\mathbf{Q}_{s,\infty}(1/2)$ is unsatisfiable, provided that $s > \frac{\sqrt{n}}{2\sqrt{d(1-\gamma)}}$.* Invoking Proposition 5.4.2, we conclude that *with the representation structure in question, assuming $\xi \sim \mathcal{N}(0, I_m)$ and $\sigma > 0$, for every pair of constants $\alpha, C > 0$, the error bound*

$$\forall (x \in \mathbb{R}^n, u \in \mathcal{U}) : \text{Prob}\left\{\|\widehat{x}(Ax + u + \sigma\xi) - x\|_\infty \leq \alpha + Cs^{-1}L_1(x - x^s)\right\} \geq 1 - \epsilon$$

is not achievable for any estimate $\widehat{x}(\cdot)$, unless $s \leq O(1)C\sqrt{n/d}$. Informally speaking, at the (block) sparsity level $\sqrt{m}$ "something happens" – in particular, the nice picture "everything is as in the case of direct observations" outlined in item **B** of Introduction section ceases to exist, provided the approximation error is measured in the uniform norm.

5.5 RIP and Condition $\mathbf{Q}_{s,q}(\kappa)$

In this section, we restrict ourselves with the *standard* representation structure, meaning that $B = I_n$, $K = n$, $n_1 = ... = n_K = 1$ and all $\|\cdot\|_{(k)}$ are just the standard norms $|\cdot|$ on the real axis.

Recall that a sensing matrix $A \in \mathbb{R}^{m \times n}$ satisfies the Restricted Isometry Property RIP(δ, k) (here $\delta \geq 0$ and k is a positive integer) if for every $x \in \mathbb{R}^n$ with at most k nonzero entries one has

$$(1 - \delta)\|x\|_2^2 \leq x^T A^T A x \leq (1 + \delta)\|x\|_2^2 \tag{197}$$

Proposition 5.5.1 *Let $A \in \mathbb{R}^{m,n}$ satisfy* $\mathrm{RIP}(\delta, 2s)$ *for some $\delta < 1$ and positive integer s. Then*

(i) *The pair $\left(H = \frac{s^{-1/2}}{\sqrt{1-\delta}} I_m, \|\cdot\|_2\right)$ satisfies the condition $\mathbf{Q}_{s,2}\left(\frac{\delta}{1-\delta}\right)$ associated with A and the standard representation structure.*

(ii) *The pair $\left(H = \frac{1}{1-\delta} A, \|\cdot\|_\infty\right)$ satisfies the condition $\mathbf{Q}_{s,2}\left(\frac{\delta}{1-\delta}\right)$ associated with A and the standard representation structure.*

(iii) *Let $\widehat{s} = \mathrm{Floor}\left(\frac{(1-\delta)\sqrt{s}}{3\delta}\right)$. Then one can build efficiently a matrix $H = [h^1, ..., h^n] \in \mathbb{R}^{m\times n}$ such $\|h^i\|_2 \le \frac{1}{\sqrt{1-\delta}}$ for all i and the pair $(H, \|\cdot\|_\infty)$ satisfies the condition $\mathbf{Q}_{\widehat{s},\infty}(\frac{1}{3})$ associated with A and the standard representation structure.*

(iv) *Let m, s, n be such that $1 \le s \le m \le n$, $n \ge 1000$, and*

$$36s\sqrt{\ln(n)/m} \le 1, \tag{198}$$

and let A satisfy $\mathrm{RIP}(4\sqrt{\frac{s\ln(n)}{m}}, s)$. *Then $(H = \frac{3}{2}A, \|\cdot\|_\infty)$ satisfies $\mathbf{Q}_{s,\infty}(\frac{1}{3})$. Besides this, a Gaussian A (i.e., random $m \times n$ matrix with independent $\mathcal{N}(0, 1/m)$ entries) satisfies* $\mathrm{RIP}(4\sqrt{\frac{s\ln(n)}{m}}, s)$ *with probability at least $1 - 1/n$.*

Combining Theorems 5.3.1, 5.3.2 and Proposition 5.5.1, we arrive at the following conclusion:

Corollary 5.5.1 *Let $A = [A_1, ..., A_n] \in \mathbb{R}^{m\times n}$, $\delta < 1/3$ and positive integer s be such that A satisfies* $\mathrm{RIP}(\delta, 2s)$, *and let the representation structure be standard.*

(i) *Let $R > 0$ be such that for the set $\Xi = \{\xi : \|\sigma\xi + u\|_2 \le R \ \forall u \in \mathcal{U}\}$ one has $P(\Xi) \ge 1 - \epsilon$. Let $H = \frac{s^{-1/2}}{\sqrt{1-\delta}} I_m$, $\|\cdot\| = \|\cdot\|_2$ and*

$$\rho = \frac{s^{-1/2} R}{\sqrt{1-\delta}}.$$

Then for the regular ℓ_1-recovery associated with $H, \|\cdot\|_2, \rho$ and for the penalized ℓ_1-recovery associated with H, $\|\cdot\|_2$, s one has

$$\begin{array}{l} \forall (x \in \mathbb{R}^n, u \in \mathcal{U}, \xi \in \Xi): \\ \|\widehat{x}(Ax + u + \sigma\xi) - x\|_p \le 3s^{\frac{1}{p}} \frac{2\sqrt{1-\delta}s^{-1/2}R + (1-\delta)s^{-1}\|x - x^s\|_1}{1-3\delta}, \ 1 \le p \le 2. \end{array} \tag{199}$$

(ii) *Let*

$$\rho_i = \sigma \min\left\{d : \mathrm{Prob}_{\xi\sim P}\{(1-\delta)^{-1}|A_i^T\xi| > d\} \le \epsilon/n\right\},$$

so that with $\Xi = \{\xi : (1-\delta)^{-1}|A_i^T\xi| \le \rho_i, 1 \le i \le n\}$ *one has* $P(\Xi) \ge 1-\epsilon$. *Let also* $H = \frac{1}{1-\delta}A$ *and*

$$\rho = \max_i \left[(1-\delta)^{-1} \max_{u\in\mathcal{U}} u^T A_i + \rho_i\right].$$

Then for the regular ℓ_1*-recovery associated with* $H, \|\cdot\|_\infty, \rho$, *and the penalized* ℓ_1*-recovery associated with* $H, \|\cdot\|_\infty, s$, *and for all* $1 \le p \le 2$ *one has*

$$\forall(x \in \mathbb{R}^n, u \in \mathcal{U}, \xi \in \Xi) : \|\widehat{x}(Ax+u+\sigma\xi) - x\|_p \le 3s^{\frac{1}{p}} \frac{(1-\delta)[2\rho + s^{-1}\|x-x^s\|_1]}{1-3\delta}. \quad (200)$$

Discussion. **I.** Let $\mathcal{U} = \{u \in \mathbb{R}^m : \|u\|_2 \le r\}$ and $\sigma = 0$. In the situation of Corollary 5.5.1.i with $\epsilon = 0$ we get $R = r$, $\rho = s^{-1/2}r/\sqrt{1-\delta}$, the recoveries are

$$\begin{array}{rcl} \widehat{x}_{\rm reg}(y) & \in & {\rm Argmin}_u \{\|u\|_1 : \|Au-y\|_2 \le r\}, \\ \widehat{x}_{\rm pen}(y) & \in & {\rm Argmin}_u \left\{\|u\|_1 + 2\frac{s^{1/2}}{\sqrt{1-\delta}}\|Au-y\|_2\right\}, \end{array} \quad (201)$$

and the bound (199) reads

$$\|\widehat{x}(Ax+u) - x\|_p \le 3s^{\frac{1}{p}} \frac{2\sqrt{1-\delta}s^{-1/2}r + (1-\delta)s^{-1}\|x-x^s\|_1}{1-3\delta}, \quad (202)$$

$\forall(x \in \mathbb{R}^n, u : \|u\|_2 \le r)$ and $1 \le p \le 2$.

Note that this bound is not completely evident even in the case of direct observations $m = n$, $A = I_n$ (in this case, RIP$(0, 2s)$ holds true whenever $2s \le n$). As p grows from 1 to 2, the right hand side in the bound decreases from $3\frac{2\sqrt{s}\sqrt{1-\delta}r + (1-\delta)\|x-x^s\|_1}{1-3\delta}$ $(p=1)$ to $3\frac{2\sqrt{1-\delta}r + (1-\delta)s^{-1/2}\|x-x^s\|_1}{1-3\delta}$ $(p=2)$. We clearly have

$$\|\widehat{x}(Ax+u) - x\|_p \le \|\widehat{x}(Ax+u) - x\|_2 \le 3\frac{2\sqrt{1-\delta}r + (1-\delta)s^{-1/2}\|x-x^s\|_1}{1-3\delta}, \ 2 \le p \le \infty.$$

II. Let $\sigma \sim \mathcal{N}(0, I_m)$ and $\mathcal{U} = \{u : \|u\|_2 \le r\}$. In the situation of Corollary 5.5.1.ii with $\epsilon \in (0,1)$ we get

$$\rho = (1-\delta)^{-1}\bar{\rho}, \text{ where } \bar{\rho} = \max_i [r\|A_i\|_2 + \sigma{\rm Erfinv}(\epsilon/n)\|A_i\|_2] \le \sqrt{1+\delta}[r + \sigma{\rm Erfinv}(\epsilon/n)],$$

and the recoveries are

$$\begin{array}{rcl} \widehat{x}_{\rm reg}(y) & \in & {\rm Argmin}_u \left\{\|u\|_1 : \|A^T(Au-y)\|_\infty \le \bar{\rho}\right\}, \\ \widehat{x}_{\rm pen}(y) & \in & {\rm Argmin}_u \left\{\|u\|_1 + 2s(1-\delta)^{-1}\|A^T(Au-y)\|_\infty\right\}, \end{array} \quad (203)$$

and the bound (200) reads

$$\begin{array}{l}\forall(x\in\mathbb{R}^n, u\in\mathcal{U}, \xi\in\Xi:=\{\xi : |A_i^T\xi|\le \mathrm{Erfinv}(\epsilon/n),\ 1\le i\le n\}):\\ \|\widehat{x}(Ax+u+\sigma\xi)-x\|_p \le 3s^{\frac{1}{p}}\frac{2\sqrt{1+\delta}[r+\sigma\mathrm{Erfinv}(\epsilon/n)]+(1-\delta)s^{-1}\|x-x^s\|_1}{1-3\delta},\ 1\le p\le 2.\end{array} \tag{204}$$

5.5.1 ℓ_∞-error of Dantzig Selector.

We continue to consider the case of the standard representation structure. If $\ln(m) = O(\ln(n))$ and A is an $m\times n$ Gaussian matrix, Proposition 5.5.1.iv states that when $s\le O(1)\sqrt{m/\ln(m)}$ with an appropriately chosen $O(1)$, the probability for $(\frac{3}{2}A, \|\cdot\|_\infty)$ to satisfy the condition $\mathbf{Q}_{s,\infty}(1/3)$ approaches 1 as $n\to\infty$. Since for $\ln(m)=O(\ln(n))$, the Euclidean norms of columns in A are bounded by 1.1 with probability approaching 1 as $n\to\infty$, we conclude from our results on regular ℓ_1-recovery that

(!) *When $s\le O(1)\sqrt{m/\ln(m)}$, for n large and a typical realization of Gaussian A, the Dantzig selector*

$$y = Ax+\sigma\xi \mapsto \widehat{x}_{D,\epsilon}(y)\in \mathop{\mathrm{Argmin}}_u\left\{\|u\|_1 : \|A^T(Au-y)\|_\infty \le 6\sigma\sqrt{2\ln(n/\epsilon)}\right\}$$

under the standard Gaussian noise, $\xi\sim\mathcal{N}(0,I_m)$, satisfies

$$\mathrm{Prob}\{\xi : \|x-\widehat{x}_{D,\epsilon}(Ax+\sigma\xi)\|_\infty \le O(1)\sigma\sqrt{2\ln(n/\epsilon)}\}\ge 1-\epsilon \text{ for all } s\text{-sparse } x.$$

The question is, to which extent the bound $s\le O(1)\sqrt{m/\ln(m)}$ is important here. Specifically, let us pose the following question:

(?) *Consider Dantzig Selector recovery $\widehat{x}_{D,\epsilon}(\cdot)$, and let us fix ϵ, say, $\epsilon=0.01$. Given a constant $C>0$, how large, for n large, $\ln(m)=O(\ln(n))$ and a typical Gaussian $A\in\mathbb{R}^{m\times n}$, are those s for which the following holds*

$$\mathrm{Prob}\left\{\|\widehat{x}-x\|_\infty \le C\sigma\sqrt{2\ln(n/\epsilon)}\right\} > 1-\epsilon = 0.99 \tag{205}$$

for every s-sparse signal x?

The answer is given by the following

Proposition 5.5.2 *For given C, ϵ and for large m, n with $\ln(m) = O(1)\ln(n)$, a typical Gaussian $m \times n$ sensing matrix ensures (205) for every s-sparse x only when $s \leq O(1)\max[C, C^2]\sqrt{m\ln(m)}$.*

Thus, the restriction $s \leq O(1)\sqrt{m/\ln(m)}$ in (!) indeed is important: when s is by a logarithmic in m factor greater than this bound, the Dantzig selector associated with a typical Gaussian A stops to work properly in the ℓ_∞-norm.

5.6 Non-Euclidean Matching Pursuit Algorithm for Block Sparsity

The Matching Pursuit algorithm for sparse recovery is motivated by the desire to provide a reduced complexity alternative to the algorithms using ℓ_1-minimization. Several implementations of Matching Pursuit has been proposed in the compressed sensing literature including the ones for block-sparse recovery [10, 14, 54, 55]. In this section, we aim to show that for a given sparsity level s, whenever $\|\cdot\|_{(k)} = \|\cdot\|_\infty$ for all $k = 1, \ldots, K$, the verifiable condition $\mathbf{Q}_{s,\infty}$ can be used to design a specific version of the Matching Pursuit algorithm which we refer to as *Block Non-Euclidean Matching Pursuit (BNEMP) algorithm* for recovering vectors obeying a block sparsity structure. In this section, we will assume that the matrix B is invertible.

Suppose that we have in our disposal $\kappa > 0$ such that the condition $\mathbf{Q}_{s,\infty}$ is satisfied by some pair $(\hat{H}, \|\cdot\|)$. For the given $\epsilon \in (0, 1)$, let

$$\nu(\hat{H}) := \inf\left\{\rho : \ \mathrm{Prob}\{\xi : \|\hat{H}^T[u + \sigma\xi]\| \leq \rho \ \forall u \in \mathcal{U}\} \geq 1 - \epsilon\right\},$$

and by invoking Proposition 5.4.1, in this case, we can efficiently find $N = n_1 + ... + n_K$ vectors $h^1, ..., h^N$ in $\mathbb{R}^m$ and $N \times N$ matrix $V = [V^{k\ell}]_{k,\ell=1}^K$ (the blocks $V^{k\ell}$ are $n_k \times n_\ell$ matrices) such that

$$\begin{array}{ll} (a) & B = VB + [h^1, ..., h^N]^T A \\ (b) & \|V^{k\ell}\|_{\infty\to\infty} \leq \frac{\kappa}{s} \quad \forall k, \ell \leq K \\ (c) & \mathrm{Prob}\left(\Xi := \{\xi : \max\limits_{u\in\mathcal{U}} u^T h^i + \sigma|\xi^T h^i| \leq \nu(H),\ 1 \leq i \leq N\}\right) \geq 1 - \epsilon. \end{array} \tag{206}$$

In the remaining of this section, we will denote $\bar{\gamma} = s^{-1}\kappa$, $H = [h^1, ..., h^N]$ be the collection of vectors satisfying (206) and $\omega_*(\bar{\gamma}) = \nu(H)$. Let $\upsilon \geq 0$ be a given upper bound on the

"s-block tail" of the linear transform, Bx, to be recovered. Consider the following BNEMP algorithm:

Algorithm 2

1. *Initialization:* Set $v^{(0)} = 0$, $\alpha_0 = \frac{L_{s,1}(H^T y) + s\omega_*(\bar{\gamma}) + v}{1 - s\bar{\gamma}}$.

2. *Step k, $k = 1, 2, ...$:* Given $v^{(k-1)} \in \mathbb{R}^n$ and $\alpha_{k-1} \geq 0$, compute

   (a) $g = H^T(y - Av^{(k-1)})$ and vector $\Delta \in \mathbb{R}^n$ by setting

   $$\Delta_i = \mathrm{sign}(g_i)[|g_i| - \bar{\gamma}\alpha_{k-1} - \omega_*(\bar{\gamma})]_+, \quad 1 \leq i \leq n$$

   (here $[a]_+ = \max[0, a]$).

   (b) Set $v^{(k)} = v^{(k-1)} + B^{-1}\Delta$ and

   $$\alpha_k = 2s\bar{\gamma}\alpha_{k-1} + 2s\omega_*(\bar{\gamma}) + v. \tag{207}$$

   and loop to step $k + 1$.

3. The approximate solution found after k iterations is $v^{(k)}$.

Proposition 5.6.1 *Let $v \geq 0$ be given and assume that $s\bar{\gamma} < 1$ is such that (206) takes place. Then there exists a set $\Xi \subseteq \mathbb{R}^m$, $\mathrm{Prob}\{\xi \in \Xi\} \geq 1 - \epsilon$, of "good" realizations of ξ such that whenever $\xi \in \Xi$, for every $x \in \mathbb{R}^n$ satisfying $L_1(Bx - [Bx]^s) \leq v$ and every $u \in U$, the approximate solution $w^{(k)} := Bv^{(k)}$ for Bx and the value α_k after the k-th step of Algorithm 2 satisfy*

$$(a_k) \quad \text{for all } 1 \leq i \leq K \text{ and } 1 \leq j \leq n_i, \quad w^{(k)}[i]_j \in \mathrm{Conv}\{0; (Bx)[i]_j\},$$
$$(b_k) \quad L_1(Bx - w^{(k)}) \leq \alpha_k \text{ and } L_\infty(Bx - w^{(k+1)}) \leq 2\bar{\gamma}\alpha_k + 2\omega_*(\bar{\gamma}).$$

Note that if $\kappa < 1/2$, i.e., $2s\bar{\gamma} < 1$, then also $s\bar{\gamma} < 1$, so that Proposition 5.6.1 holds true. Furthermore, by (207) the sequence α_k converges exponentially fast to the limit $\alpha_\infty := \frac{2s\omega_*(\bar{\gamma}) + v}{1 - s\bar{\gamma}}$:

$$L_1(Bv^{(k)} - Bx) \leq \alpha_k = (2s\bar{\gamma})^k[\alpha_0 - \alpha_\infty] + \alpha_\infty.$$

Along with the second inequality of (b_k) this implies the bounds:

$$L_\infty(Bv^{(k)} - Bx) \le 2\bar\gamma\alpha_{k-1} + 2\omega_*(\bar\gamma) \le \frac{\alpha_k}{s},$$

and since $L_p(w) \le L_1(w)^{\frac{1}{p}} L_\infty(w)^{\frac{p-1}{p}}$ for $1 \le p \le \infty$, we have

$$L_p(Bv^{(k)} - Bx) \le s^{\frac{1-p}{p}}(2s\bar\gamma)^k[\alpha_0 - \alpha_\infty] + \alpha_\infty.$$

The bottom line here is as follows:

Corollary 5.6.1 *Let $\bar\gamma < 1/(2s)$ be such that $\mathbf{Q}_{s,\infty}$ takes place, so that we can find efficiently contrast matrices H, V satisfying (206). The recovery of Bx where $\upsilon_s(Bx) \le \upsilon$ with Algorithm 2 associated with H, V, one ensures that for every $t = 1, 2, \ldots$, the approximate solution $v^{(t)}$ found after t iterations satisfies*

$$\begin{aligned}&\mathrm{Risk}_p^B(v^{(t)}|s,\sigma,\upsilon,\epsilon)\\ &\quad\le s^{\frac{1}{p}}\left(\frac{2\omega_*(\bar\gamma)+s^{-1}v}{1-2s\bar\gamma} + (2s\bar\gamma)^t\left[\frac{s^{-1}(L_{s,1}(H^Ty)+v)+\omega_*(\bar\gamma)}{1-s\bar\gamma} - \frac{2\omega_*(\bar\gamma)+s^{-1}v}{1-2s\bar\gamma}\right]\right),\end{aligned}$$

for all $1 \le p \le \infty$ (cf. (176)).

5.7 Proofs of Chapter 5

5.7.1 Proof of Theorem 5.3.1

Let us fix $x \in \mathbb{R}^n$, $u \in \mathcal{U}$ and $\xi \in \Xi$, and let us set $\eta = u + \sigma\xi$, $\widehat{x} = \widehat{x}_{\rm reg}(Ax + \eta)$. Let also $I \in \{1, ..., K\}$ be the set of indexes of the s largest in magnitude blocks in Bx, J be the complement of I in $\{1, ..., K\}$, and let for $w \in \mathcal{W}$, w_I and w_J be the vectors obtained from w by zeroing blocks $w[k]$ with indices $k \notin I$ and $k \notin J$, respectively, and keeping the remaining blocks intact. Finally, let $z = \widehat{x} - x$.

1^0. By **B** and due to $\xi \in \Xi$, $u \in \mathcal{U}$, we have

$$\|H^T([Ax + \eta] - Ax)\| \le \rho, \tag{208}$$

so that x is a feasible solution to the optimization problem specifying $\widehat{x}$, whence $L_1(B\widehat{x}) \leq \mathbf{L}_1(Bx)$. We therefore have

$$\begin{aligned} L_1([B\widehat{x}]_J) &= L_1(B\widehat{x}) - L_1([B\widehat{x}]_I) \leq L_1(Bx) - L_1([B\widehat{x}]_I) \\ &= L_1([Bx]_I) + L_1([Bx]_J) - L_1([B\widehat{x}]_I) \\ &\leq L_1([Bz]_I) + L_1([Bx]_J), \end{aligned}$$

whence

$$L_1([Bz]_J) \leq L_1([B\widehat{x}]_J) + L_1([Bx]_J) \leq L_1([Bz]_I) + 2L_1([Bx]_J).$$

It follows that

$$L_1(Bz) = L_1([Bz]_I) + L_1([Bz]_J) \leq 2L_1([Bz]_I) + 2L_1([Bx]_J). \tag{209}$$

Further, by definition of $\widehat{x}$ we have $\|H^T([Ax + u + \sigma\xi] - A\widehat{x})\| \leq \rho$, which combines with (208) to imply that

$$\|H^T A(\widehat{x} - x)\| \leq 2\rho. \tag{210}$$

2^0. Since $(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,q}(\kappa)$, it satisfies $\mathbf{Q}_{s,1}(\kappa)$ as well (Observation 5.2.1), that is,

$$L_{s,1}(Bz) \leq s\|H^T Az\| + \kappa L_1(Bz).$$

By (210), it follows that $L_{s,1}(Bz) \leq 2s\rho + \kappa L_1(Bz)$, which combines with the evident inequality $L_1([Bz]_I) \leq L_{s,1}(Bz)$ and with (209) to imply that

$$L_1([Bz]_I) \leq 2s\rho + \kappa L_1(Bz) \leq 2s\rho + 2\kappa L_1([Bz]_I) + 2\kappa L_1([Bx]_J),$$

whence

$$L_1([Bz]_I) \leq \frac{2s\rho + 2\kappa L_1([Bx]_J)}{1 - 2\kappa}.$$

Invoking (209), we conclude that

$$L_1(Bz) \leq \frac{4s\rho + 2L_1([Bx]_J)}{1 - 2\kappa}. \tag{211}$$

3^0. Since $(H, \|\cdot\|)$ satisfy $\mathbf{Q}_{s,q}(\kappa)$, we have

$$L_{s,q}(Bz) \leq s^{\frac{1}{q}}\|H^T Az\| + \kappa s^{\frac{1}{q}-1} L_1(Bz),$$

which combines with (211) and (210) to imply that

$$L_{s,q}(Bz) \le 2s^{\frac{1}{q}} \frac{\rho + \kappa s^{-1} L_1([Bx]_J)}{1-2\kappa}. \tag{212}$$

Let λ be the $(s+1)$-st largest among the magnitudes of blocks in Bz, and let $w = Bz - [Bz]^s$. We have $L_q(w) \le L_\infty(w)^{\frac{q-1}{q}} L_1(w)^{\frac{1}{q}} \le \lambda^{\frac{q-1}{q}} L_1(Bz)^{\frac{1}{q}} \le \lambda^{\frac{q-1}{q}} \left[\frac{4s\rho + 2L_1([Bx]_J)}{1-2\kappa}\right]^{\frac{1}{q}}$, where the concluding inequality is given by (211). Besides this, (212) implies that $\lambda \le 2\frac{\rho+\kappa s^{-1} L_1([Bx]_J)}{1-2\kappa} \le \frac{2\rho + s^{-1} L_1([Bx]_J)}{1-\kappa}$ (note that $\kappa < 1/2$), whence $L_q(w) \le (2s)^{\frac{1}{q}} \frac{2\rho + s^{-1} L_1([Bx]_J)}{1-2\kappa}$. Taking into account (212) and the fact that the supports of $[Bz]^s$ and w do not intersect, we get

$$\begin{aligned} L_q(Bz) &\le 2^{\frac{1}{q}} \max[L_q([Bz]^s), L_q(w)] = 2^{\frac{1}{q}} \max[L_{s,q}(Bz), L_q(w)] \\ &\le 2^{\frac{2}{q}} s^{\frac{1}{q}} \frac{2\rho + s^{-1} L_1([Bx]_J)}{1-2\kappa}. \end{aligned}$$

This relation combines with (211), Hölder inequality and the relation $\|[Bx]_J\| = \|Bx - [Bx]^s\|$ to imply (172). □

5.7.2 Proof of Theorem 5.3.2

Same as in the proof of Theorem 5.3.1, let us fix $x \in \mathbb{R}^n$, $u \in \mathcal{U}$ and $\xi \in \Xi$, and let us set $\eta = u + \sigma\xi$, $\widehat{x} = \widehat{x}_{\rm pen}(Ax + \eta)$. Let also $I \subset \{1, ..., K\}$ be the set of indices of the s largest in magnitude blocks in Bx, J be the complement of I in $\{1, ..., K\}$, and for $w \in \mathcal{W}$ let w_I, w_J be the vectors obtained from w by zeroing out all blocks with indexes not in I, respectively, not in J. Finally, let $z = \widehat{x} - x$.

1^0. We have

$$L_1(B\widehat{x}) + 2s\|H^T(A\widehat{x} - Ax - \eta)\| \le L_1(Bx) + 2s\|H^T\eta\|$$

and

$$\|H^T(A\widehat{x} - Ax - \eta)\| = \|H^T(Az - \eta)\| \ge \|H^T Az\| - \|H^T\eta\|,$$

whence

$$L_1(B\widehat{x}) + 2s\|H^T Az\| \le L_1(Bx) + 4s\|H^T\eta\| \le L_1(Bx) + 4s\rho, \tag{213}$$

where the concluding inequality follows from the fact that $u \in \mathcal{U}$, $\xi \in \Xi$ due to **B**. We have

$$\begin{aligned} L_1(B\widehat{x}) &= L_1(Bx + Bz) = L_1([Bx]_I + [Bz]_I) + L_1([Bx]_J + [Bz]_J) \\ &\ge L_1([Bx]_I) - L_1([Bz]_I) + L_1([Bz]_J) - L_1([Bx]_J), \end{aligned}$$

which combines with (213) to imply that

$$L_1([Bx]_I) - L_1([Bz]_I) + L_1([Bz]_J) - L_1([Bx]_J) + 2s\|H^T Az\| \le L_1(Bx) + 4s\rho,$$

or, which is the same,

$$L_1([Bz]_J) - L_1([Bz]_I) + 2s\|H^T Az\| \le 2L_1([Bx]_J) + 4s\rho. \tag{214}$$

Since $(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,q}(\kappa)$, $(H, \|\cdot\|)$ satisfies $\mathbf{Q}_{s,1}(\kappa)$ as well, whence

$$L_1([Bz]_I) \le L_{s,1}(Bz) \le s\|H^T Az\| + \kappa L_1(Bz),$$

or, which is the same,

$$(1-\kappa)L_1([Bz]_I) - \kappa L_1([Bz]_J) - s\|H^T Az\| \le 0. \tag{215}$$

Taking weighted sum of (214) and (215), the weights being 1, 2, respectively, we get

$$(1-2\kappa)\left[L_1([Bz]_I) + L_1([Bz]_J)\right] \le 2L_1([Bx]_J) + 4s\rho,$$

that is,

$$L_1(Bz) \le \frac{4s\rho + 2L_1([Bx]_J)}{1-2\kappa}, \tag{216}$$

exactly as in (211). Further, by (213) we have

$$2s\|H^T Az\| \le L_1(Bx) - L_1(B\widehat{x}) + 4s\rho \le L_1(Bz) + 4s\rho,$$

which combines with (216) to imply that

$$2s\|HA^T z\| \le \frac{4s\rho + 2L_1([Bx]_J)}{1-2\kappa} + 4s\rho = \frac{4s\rho(2-2\kappa) + 2L_1([Bx]_J)}{1-2\kappa}. \tag{217}$$

From $\mathbf{Q}_{s,q}(\kappa)$ it follows that

$$L_{s,q}(Bz) \le s^{\frac{1}{q}}\|H^T Az\| + \kappa s^{\frac{1}{q}-1}L_1(Bz),$$

which combines with (217) and (216) to imply that

$$\begin{aligned} L_{s,q}(Bz) &\le s^{\frac{1}{q}-1}\left[s\|H^T Az\| + \kappa L_1(Bz)\right] \\ &\le s^{\frac{1}{q}-1}\left[\frac{4s\rho(1-\kappa) + L_1([Bx]_J)}{1-2\kappa} + \frac{\kappa[4s\rho + 2L_1([Bx]_J)]}{1-2\kappa}\right] \\ &= s^{\frac{1}{q}}\frac{4\rho + 2s^{-1}L_1([Bx]_J)}{1-2\kappa}. \end{aligned}$$

It remains to repeat the reasoning following (212) in item 3^0 of the proof of Theorem 5.3.1. Specifically, denoting λ the $(s+1)$-st largest magnitude of entries in Bz, the above inequality results in

$$\lambda \le s^{-1/q} L_{s,q}(Bz) \le \frac{4\rho + 2s^{-1} L_1([Bx]_J)}{1-2\kappa}, \tag{218}$$

so that for the vector $w = Bz - [Bz]^s$ one has

$$\begin{array}{rcl} L_q(w) & \le & \lambda^{1-\frac{1}{q}} L_1(w)^{\frac{1}{q}} \le \lambda^{1-\frac{1}{q}} L_1(Bz)^{\frac{1}{q}} \\ & \le & \left[\frac{4\rho+2s^{-1}L_1([Bx]_J)}{1-2\kappa}\right]^{\frac{q-1}{q}} \left[\frac{4s\rho+2L_1([Bx]_J)}{1-2\kappa}\right]^{\frac{1}{q}} \ \text{[by (218) and (216)]} \\ & = & s^{\frac{1}{q}} \frac{4\rho+2s^{-1}L_1([Bx]_J)}{1-2\kappa}, \end{array}$$

whence, invoking (217) and taking into account that $[Bz]^s$ and w have non-intersecting supports,

$$L_q(Bz) \le 2^{\frac{1}{q}} \max[L_q([Bz]^s), L_q(w)] = 2^{\frac{1}{q}} \max[L_{s,q}(Bz), L_q(w)] \le 2(2s)^{\frac{1}{q}} \frac{2\rho + s^{-1} L_1([Bx]_J)}{1-2\kappa}.$$

This combines with (216) and Hölder inequality to imply (174). □

5.7.3 Proof of Proposition 5.4.1

(i): Let $H \in \mathbb{R}^{m\times M}$, $\|\cdot\|$, ρ satisfy (!). Then for every $k \le K$ and every $i \le n_i$, denoting by w_{ki} i-th entry in $w[k]$, $w \in \mathcal{W}$, we have

$$|[Bx]_{ki}| \le \|H^T Ax\| + s^{-1}\kappa L_1(Bx),$$

or, which is the same by homogeneity,

$$\min_x \left\{\|H^T Ax\| - [Bx]_{ki} : \ L_1(Bx) \le 1\right\} \ge -s^{-1}\kappa.$$

Equivalently the optimal value Opt_{ki} in the conic optimization problem

$$\mathrm{Opt}_{ki} = \min_{x,t} \left\{t - [e^{ki}]^T Bx : \ \|H^T Ax\| \le t, \ L_1(Bx) \le 1\right\},$$

where $e^{ki} \in \mathcal{W}$ is the vector with the only nonzero entry, equal to 1, placed at i-th position of the k-th block, is $\ge -s^{-1}\kappa$. Let $\|\cdot\|_*$ is the norm conjugate to $\|\cdot\|$. Since the problem clearly is strictly feasible, this is the same as to say that the dual problem

$$\max_{\mu\in\mathbb{R}, g\in\mathcal{W}, \eta\in\mathbb{R}^M} \left\{-\mu : \ A^T H\eta + B^T g = B^T e^{ki}, \|g[\ell]\|_1 \le \mu, 1 \le \ell \le K, \|\eta\|_* \le 1\right\}$$

has a feasible solution with the value of the objective $\geq -s^{-1}\kappa$. It follows that there exists $\eta = \eta^{ki}$ and $g = g^{ki}$ such that

$$\begin{array}{ll} (a) & B^T e^{ki} = A^T h^{ki} + B^T g^{ki}, \\ (b) & h^{ki} := H\eta^{ki}, \|\eta^{ki}\|_* \leq 1 \\ (c) & \|g^{ki}[\ell]\|_1 \leq s^{-1}\kappa,\ 1 \leq \ell \leq K. \end{array} \tag{219}$$

Denoting by h^i i-th column in the matrix $[h^{1,1}, ..., h^{1,n_1}, h^{2,1}, ..., h^{2,n_2}, ..., h^{K,1}, ..., h^{K,n_K}]$, defining $V^{k\ell}$ as the $n_k \times n_\ell$ matrix with the rows $(g^{ki}[\ell])^T$, $i = 1, ..., n_k$, and setting $V = [V^{k\ell}]_{k,\ell=1}^K$, (219.$a$,$c$) ensure the validity of (175a,b) (note that $\|M\|_{\infty\to\infty}$ is nothing but the maximum of $\|\cdot\|_1$-norms of the rows in M). Besides this, by (219.b) and the definition of Ξ (see (!)) we have

$$\begin{array}{rcl} \xi \in \Xi & \Rightarrow & \|H^T[u + \sigma\xi]\| \leq \rho \quad \forall u \in \mathcal{U} \\ & \Rightarrow_a & |[h^{ki}]^T[u + \sigma\xi]| \leq \rho \quad \forall u \in \mathcal{U} \\ & \Rightarrow_b & \max\limits_{u\in\mathcal{U}} u^T h^{ki} + \sigma|\xi^T h^{ki}| \leq \rho \end{array}$$

where the implication $\Rightarrow_a$ is due to the fact that $|[h^{ki}]^T\zeta| = |[\eta^{ki}]^T H^T\zeta| \leq \|H^T\zeta\|$ for all ζ because of $\|\eta^{ki}\|_* \leq 1$, and the implication $\Rightarrow_b$ is due to the fact that $\mathcal{U}$ is symmetric w.r.t. the origin. We conclude that $\Xi \subset \Xi^+$ and thus $P(\Xi^+) \geq P(\Xi) \geq 1 - \epsilon$, as required in (175.$c$). (i) is proved.

(ii): Let $\widehat{H} = [h^1, ..., h^N]$, $V = [V^{k\ell}]_{k,\ell=1}^K$, ρ satisfy (175). Then for every $x \in \mathbb{R}^n$ we have

$$w := Bx = VBx + \widehat{H}^T Ax = Vw + \underbrace{\widehat{H}^T Ax}_{v},$$

whence $w[k] = \sum_{\ell=1}^K V^{k\ell} w[\ell] + v[k]$, so that

$$\|w[k]\|_{(k)} = \|w[k]\|_\infty \leq \sum_{\ell=1}^K \|V^{k\ell}\|_{\infty\to\infty}\|w[\ell]\|_\infty + \|v[k]\|_\infty \leq s^{-1}\kappa L_1(w) + \|\widehat{H}^T Ax\|_\infty,$$

that is,

$$L_{s,\infty}(Bx) = \max_k \|w[k]\|_{(k)} \leq \|\widehat{H}^T Ax\|_\infty + s^{-1}\kappa L_1(Bx)$$

for all x, meaning that $(\widehat{H}, \|\cdot\|_\infty)$ satisfies $\mathbf{Q}_{s,\infty}(\kappa)$. Further, we have

$$\begin{array}{rcl} \Xi & := & \left\{\xi : \|\widehat{H}^T[u + \sigma\xi]\| \leq \rho\ \forall u \in \mathcal{U}\right\} = \left\{\xi : |[h^i]^T[u + \sigma\xi]| \leq \rho\ \forall u \in \mathcal{U},\ \forall i \leq N\right\} \\ & = & \left\{\xi : \max\limits_{u\in\mathcal{U}} u^T h^i + \sigma|[h^i]^T\xi| \leq \rho\ \forall i \leq N\right\} = \Xi^+, \end{array}$$

whence $P(\Xi) = P(\Xi^+) \geq 1-\epsilon$. Thus, $\widehat{H}$, $\|\cdot\|_\infty$, ρ satisfy (!). $\square$

5.7.4 Proof of Proposition 5.4.2

Notation. Let $1 \leq k \leq K$ and $1 \leq i \leq n_k$. For a vector $w \in \mathcal{W}$, we set $[w]_{ki}$ to be i-th coordinate in $w[k]$. For a vector $u \in \mathbb{R}^{n_k}$, we set $\|u\|_\infty^i = \max_{j \neq i} |u_j|$, with the convention that the latter maximum is 0 when $n_k = 1$. Further, let e^{ki} be the vector from $\mathcal{W}$ such that $[e^{ki}]_{\ell j} = 1$ when $\ell = k$ and $j = i$ and $[e^{ki}]_{\ell j} = 0$ for all remaining pairs ℓ, j. Finally, let $B = [B^1; ...; B^K]$ with $n_k \times n$ matrices B^k.

0^0. Let us fix k, i, $1 \leq k \leq K$, $1 \leq i \leq n_k$, and set

$$M = \tfrac{\alpha}{c} S,\ a = 1.1\alpha + S^{-1}cM = 2.1\alpha,$$
$$X_+^{ki} = \{x \in \mathbb{R}^n : [B^k x]_i = a, \|B^k x\|_\infty^i + \textstyle\sum_{\ell \neq k} \|B^\ell x\|_\infty \leq M\},\ X_-^{ki} = -X_+^{ki},$$
$$Y_+^{ki} = AX_+^{ki},\ Y_-^{ki} = AX_-^{ki} = -Y_+^{ki}$$

and let $\mathcal{V} = 2\mathcal{U} + 2\sigma\{C\eta : \|\eta\|_2 \leq \mathrm{Erfinv}(\epsilon)\}$.

0^0. It may happen that $X_\pm^{ki} = \emptyset$. This is exactly the same as to say that the optimal value in the strictly feasible conic optimization problem

$$\max_x \left\{ [e^{ki}]^T Bx :\ \|B^k x\|_\infty^i + \sum_{\ell \neq k} \|B^\ell x\|_\infty \leq M \right\}$$

is $< a$, meaning that the dual problem

$$\min_{v \in \mathcal{W}, t} \left\{ Mt :\ [v]_{ki} = 0, \sum_{\ell=1}^K [B^\ell]^T v[\ell] = B^T e^{ki}, \max_{1 \leq \ell \leq K} \|v[\ell]\|_1 \leq t \right\}$$

is $< a$, whence there exists $v^{ki} \in \mathcal{W}$ such that $[v^{ki}]_{ki} = 0$, $B^T v^{ki} = B^T e^{ki}$ and $M \max_\ell \|v^{ki}[\ell]\|_1 < a$, that is, $\max_\ell \|v^{ki}[\ell]\|_1 < a/M = 2.1cS^{-1}$. Thus, when $X_\pm^{ki}$ is empty, setting $h^{ki} = 0 \in \mathbb{R}^m$, we get vectors $h^{ki} \in \mathbb{R}^m$ and $v^{ki} \in \mathcal{W}$ such that there exists $v^{ki} \in \mathcal{W}$ such that

$$\begin{array}{ll} (a_{ki}) & B^T v^{ki} + Ah^{ki} = B^T e^{ki}, \\ (b_{ki}) & v_{ki}^{ki} = 0,\ \max\limits_{1 \leq \ell \leq K} \|v^{ki}[\ell]\|_1 \leq 2.1cS^{-1}, \\ (c_{ki}) & \max\limits_{u \in \mathcal{U}} u^T h^{ki} + \sigma \mathrm{Erfinv}(\epsilon) \|C^T h^{ki}\|_2 \leq 2.1\alpha \end{array} \tag{220}$$

1^0. Assume now that $X^{ki}_{\pm} \neq \emptyset$. Then $Y^{ki}_{\pm}$ are nonempty convex sets. We claim that whenever $0 < \theta < 1$, the convex compact set $\theta\mathcal{V}$ does not intersect the convex set $2Y^{ki}_{+}$. Indeed, if the opposite is true, there exists $v \in \mathcal{U}$ and e, $\|e\|_2 \leq \mathrm{Erfinv}(\epsilon)$, such that $\theta(v + \sigma Ce) = Az$ with $z \in X^{ki}_{+}$. Now consider two hypotheses on the distribution of a random vector $\zeta \in \mathbb{R}^m$: the first, H_+, states that $\zeta \sim P_+$, where P_+ is the distribution of $\theta\sigma Ce + \sigma C\eta$, $\eta \sim \mathcal{N}(0, I_m)$, and the second, H_-, states that $\zeta \sim P_-$, where P_- is the distribution of $-\theta\sigma Ce + \sigma C\eta$, $\eta \sim \mathcal{N}(0, I_m)$. Consider the following procedure for distinguishing between these two hypotheses: given ζ, we compute $\widehat{x}(\zeta)$ and accept H_+ when $[B\widehat{x}(\zeta)]_{ki} > 0$, otherwise we accept H_-. We claim that this procedure rejects the true hypothesis with probability $\leq \epsilon$. Indeed, applying (181) to $u = -\theta v$ and $x = z$, we get

$$\mathrm{Prob}_{\eta\sim\mathcal{N}(0,I_m)}\{\|B[z - \widehat{x}(Az - \theta v + \sigma C\eta)]\|_\infty \leq \alpha + cS^{-1}L_1(Bz - [Bz]^S)\} \geq 1 - \epsilon.$$

Since $Az = \theta v + \theta\sigma Ce$ and $L_1(Bz - [Bz]^S) \leq \sum_{\ell\neq k} \|B^\ell z\|_\infty \leq M$, we get $\alpha + cS^{-1}L_1(Bz - [Bz]^S) \leq \alpha + cS^{-1}M = 2\alpha$, while $[Bz]_{ki} = a = 2.1\alpha$; it follows that $\|B[z - \widehat{x}(Az - \theta v + \sigma C\eta)]\|_\infty \leq \alpha + cS^{-1}L_1(Bz - [Bz]^S)$ implies that $2.1\alpha - [B\widehat{x}(\theta\sigma Ce + \sigma C\eta)]_{ki} \leq 2\alpha$ and thus implies that $[B\widehat{x}(\theta\sigma Ce + \sigma C\eta)]_{ki} > 0$. We see that

$$\mathrm{Prob}_{\eta\sim\mathcal{N}(0,I_m)}\{[B\widehat{x}(\theta\sigma Ce + \sigma C\eta)]_{ki} > 0\} \geq 1 - \epsilon,$$

that is, our rule for distinguishing between H_+ and H_- rejects H_+ when this hypothesis is true with probability $\leq \epsilon$. Similarly, applying (181) to $u = \theta v$ and $x = -z$, we get

$$\mathrm{Prob}_{\eta\sim\mathcal{N}(0,I_m)}\{\|B[-z - \widehat{x}(-Az + \theta v + \sigma C\eta)]\|_\infty \leq \alpha + cS^{-1}L_1(Bz - [Bz]^S)\} \geq 1 - \epsilon.$$

Since $-Az = -\theta v - \theta\sigma Ce$, we, same as above, conclude that $\|B[-z - \widehat{x}(-Az + \theta v + \sigma C\eta)]\|_\infty \leq \alpha + cS^{-1}L_1(Bz - [Bz]^S)_1$ implies that $[B\widehat{x}(-\theta\sigma Ce + \sigma C\eta)]_{ki} < 0$, and thus

$$\mathrm{Prob}_{\eta\sim\mathcal{N}(0,I_m)}\{[B\widehat{x}(-\theta\sigma Ce + \sigma C\eta)]_{ki} < 0\} \geq 1 - \epsilon$$

that is, the probability to reject H_- when the hypothesis is true is $\leq \epsilon$. On the other hand, to distinguish between the hypotheses $H_\pm$ via observation ζ distributed according to the respective distribution P_+/P_- is the same as to distinguish between the distributions $\mathcal{N}(-\theta e, I_m)$ and $\mathcal{N}(\theta e, I_m)$; to do it with probabilities $\leq \epsilon$ to reject the true distribution is

possible only when $\|\theta e\|_2 \geq \mathrm{Erfinv}(\epsilon)$, which is not the case due to $\|e\|_2 \leq \mathrm{Erfinv}(\epsilon)$ and $0 < \theta < 1$. The resulting contradiction demonstrates that $\theta\mathcal{V}$ does not intersect $2Y_+^{ki}$.

2^0. Since $\theta\mathcal{V}$ does not intersect $2Y_+^{ki}$ when $\theta < 1$, the sets $\mathcal{V}$ and $2Y_+^{ki}$ can be separated by a linear form, which can be normalized to be ≥ 2 on $2Y_+^{ki}$ and ≤ 2 on $\mathcal{V}$ (recall that $0 \in \mathrm{int}\mathcal{V}$). In other words, there exists $g = g^{ki} \in \mathbb{R}^m$ such that $\max\limits_{v\in\mathcal{V}} g^T v \leq 2$ and $\inf\limits_{y\in 2Y_+^{ki}} g^T y \geq 2$. Recalling the origin of $\mathcal{V}$, the first relation amounts to

$$\max_{u\in\mathcal{U}} u^T g + \sigma\mathrm{Erfinv}(\epsilon)\|C^T g\|_2 \leq 1, \tag{221}$$

while the relation $g^T y \geq 2$ for all $y \in 2Y_+^{ki} = 2AX_+^{ki}$ amounts to $f^T x \geq 1$ for all $x \in X_+^{ki}$, where $f = A^T g$. Recalling the definition of X_+^{ki}, it follows that

$$\min_x \left\{ f^T x : (e^{ki})^T Bx = a, \|B^k x\|_\infty^i + \sum_{\ell\neq k} \|B^\ell x\|_\infty \leq M \right\} \geq 1.$$

Passing to the dual problem, the latter inequality results in

$$\exists(t \in \mathbb{R}, y \in \mathcal{W}) : f = B^T y,\ ay_{ki} - Mt \geq 1,\ \sum_{j\neq i} |y_{kj}| \leq t,\ \max_{\ell\neq k} \|y[\ell]\|_1 \leq t. \tag{222}$$

For the above t, y we have $0 \leq t \leq (ay_{ki} - 1)/M$, so that $y_{ki} > 0$; setting

$$[v^{ki}]_{\ell j} = \begin{cases} 0, & \ell = k, j = i \\ -[y]_{\ell j}/y_{ki}, & \text{otherwise} \end{cases}, \quad \text{and} \quad h^{ki} = y_{ki}^{-1} g,$$

(222) combines with $f = A^T g$ to imply that $B^T e^{ki} = B^T v^{ki} + Ah^{ki}$. Besides this, by construction we have $[v^{ki}]_{ki} = 0$. Further, by (222) we have $\|v^{ki}[\ell]\|_1 \leq t/y_{ki} \leq a/M = 2.1cS^{-1}$, so that v^{ki}, h^{ki} satisfy (220.(a_{ki}),(b_{ki})). Besides this, by (222) we have $0 < 1/y_{ki} \leq a$, which combines with (221) to imply (220.(c_{ki})).

4^0. The bottom line is that for every $k, 1 \leq k \leq K$, and every i, $1 \leq i \leq n_k$, there exist vectors $h^{ki} \in \mathbb{R}^m$ and $v^{ki} \in \mathcal{V}$ satisfying (220). Setting

$$H = [h^{1,1}, ..., h^{1,n_1}, h^{2,1}, ..., h^{2,n_2}, ..., h^{K,1}, ..., h^{K,n_K}],$$
$$V = [(v^{1,1})^T; ...; (v^{1,n_1})^T; (v^{2,1})^T; ...; (v^{2,n_2})^T; ...; (v^{K,1})^T; ...; (v^{K,n_K})^T],$$

we get (182) as an immediate consequence of (220) and the relation $s \leq \frac{S}{6.3c}$. □

5.7.5 Proof of Proposition 5.4.4

The diagonal entries in B_ℓ are equal to 1, and the moduli of the off-diagonal entries are $\le \alpha$. This implies, quite straightforwardly, that the minimal eigenvalue of B_ℓ is $\ge 1-(d-1)\alpha$, and in our situation the latter quantity is $\ge \sigma > 0$ (otherwise the right hand side in (194) were ≤ 0, which is not the case. Setting $V = I - H^T A$, it is immediately seen that the minimal eigenvalue of B is $\ge 1-(d-1)\alpha > 0$, so that the matrices B_ℓ indeed are positive definite, H is well defined, and the Euclidean lengths of columns in H do not exceed $1/(1-(d-1)\alpha$. Now let us set $V = I - H^T A$. The $d \times d$ blocks $V^{k\ell}$ in V are as follows: when $k = \ell$, we have $V^{k\ell} = (1-\theta)I_d$; since clearly $\theta < 1$, the $\|\cdot\|_{2\to 2}$ norms of these blocks are equal to $1-\theta$. The $\|\cdot\|_{2\to 2}$ norms of the off-diagonal blocks $V^{k\ell} = \theta B_k^{-1} A^T[k]A[\ell]$ clearly do not exceed $\theta\sigma/(1-(d-1)\alpha)$ by the definition of σ and since, as we have already seen, $\|B_k^{-1}\|_{2\to 2} \le 1/(1-(d-1)\alpha)$. It remains to note that $1-\theta = \frac{\theta\sigma}{1-(d-1)\alpha}$, that is, the $\|\cdot\|_{2\to 2}$ norms of all blocks $V^{k\ell}$ do not exceed $\beta := 1-\theta = \frac{\sigma}{1-(d-1)\alpha+\sigma}$. Setting $\kappa = \beta s$, we get $\kappa < 1/2$ by (194). By construction, $q = \infty$, H, V and $\|V^{k\ell}\|_{2\to 2} \le s^{-1}\kappa$ satisfy (183), (184), whence, by Proposition 5.4.3, $(H, L_\infty(\cdot))$ satisfies $\mathbf{Q}_{s,\infty}(\kappa)$. □

5.7.6 Proof of Proposition 5.4.5

Let H, V satisfy (195). We have $V = I_n - H^T A$, and the rank of $H^T A$ is $\le m$; therefore at least $n-m$ singular values of V are ≥ 1, and therefore the squared Frobenius norm $\|V\|_F^2$ of V is at least $n-m$. On the other hand, let us upper-bound this quantity as follows. For $\pi = \{1, 2, \infty\}$, it is immediately seen that, for every $d \times d$ block $V^{k\ell}$ in V we have $\|V^{k\ell}\|_F \le \sqrt{d}\|V^{k\ell}\|_{\pi\to\pi} := \sqrt{d}\Omega_{k\ell}$. The columns C_ℓ of the $K \times K$ matrix $\Omega = [\Omega_{k\ell}]_{k,\ell=1}^K$ satisfy $\|C_\ell\|_{s,k} \le \alpha := s^{\frac{1}{q}-1}/2$ by (195). Now, it is immediately seen that for every K-dimensional vector f one has $\|f\|_2^2 \le \max\left[\frac{K}{s^{2/q}}, 1\right]\|f\|_{s,q}^2$. We now have

$$\begin{aligned} n-m &\le \|V\|_F^2 = \sum_{\ell=1}^K\sum_{k=1}^K \|V^{k\ell}\|_F^2 \le d\sum_{\ell=1}^K\sum_{k=1}^K \|V^{k\ell}\|_{\pi\to\pi}^2 = d\sum_{\ell=1}^K \|C_\ell\|_2^2 \\ &\le d\sum_{\ell=1}^K \max\left[\frac{K}{s^{2/q}}, 1\right]\|C_\ell\|_{s,q}^2 \le \max\left[\frac{K}{s^{2/q}}, 1\right] dK\alpha^2 \\ &= \frac{1}{4}n\max\left[d^{-1}ns^{-2}, s^{\frac{2}{q}-2}\right]. \end{aligned}$$

The resulting inequality, as it is immediately seen, implies the desired conclusion. □

5.7.7 Proof of Proposition 5.4.3

Let $V^\ell = [V^{1\ell}; ...; V^{K\ell}]$, $1 \le \ell \le K$. Given $x \in \mathbb{R}^n$ and setting $w = Bx$, and using the relation (183), we have

$$w = Bx = [VB + H^T A]x = Vw + H^T Ax,$$

whence,

$$\begin{aligned}
L_{s,q}(w) &= L_{s,q}(Vw + H^T Ax) \\
&\le L_{s,q}(Vw) + L_{s,q}(H^T Ax) \\
&= L_{s,q}\left(\sum_{\ell=1}^K V^\ell w[\ell]\right) + s^{\frac{1}{q}} L_\infty(H^T Ax) \\
&\le \sum_{\ell=1}^K \|[\|V^{1\ell} w[\ell]\|_{(1)}; ...; \|V^{K\ell} w[\ell]\|_{(K)}]\|_{s,q} + s^{\frac{1}{q}} L_\infty(H^T Ax) \\
&\le \sum_{\ell=1}^K \nu^*_{s,q}(V)\|w[\ell]\|_{(\ell)} + s^{\frac{1}{q}} L_\infty(H^T Ax) \\
&= \nu^*_{s,q}(V) L_1(w) + s^{\frac{1}{q}} L_\infty(H^T Ax)
\end{aligned}$$

proving (185).

To verify (187), note that for every k and every ℓ, we have

$$0 \le \|V^{k\ell} w[\ell]\|_{(k)} \le \|w[\ell]\|_{(\ell)} \max_{w \in \mathbb{R}^{n_l}: \|w\|_{(\ell)} \le 1} \|V^{k\ell} w\|_{(k)} = \|w[\ell]\|_{(\ell)} \|V^{k\ell}\|_{(\ell)\to(k)} = \|w[\ell]\|_{(\ell)} \Omega_{k\ell}.$$

Since for any two nonnegative vectors, a, b satisfying $a_i \le b_i \ \forall i$, we have $\|a\|_{s,q} \le \|b\|_{s,q}$, we get

$$\|[\|[V^{1\ell} w[\ell]\|_{(1)}; ...; \|V^{K\ell} w[\ell]\|_{(K)}]\|_{s,q} \le \|w[\ell]\|_{(\ell)} \|\mathrm{Col}_\ell[\Omega]\|_{s,q}.$$

By taking the maximum of both sides first with respect to $w[\ell]$ subject to the constraint that $\|w[\ell]\|_{(\ell)} \le 1$, and then over $1 \le \ell \le k$, we arrive at (187) implying that $\nu^*_{s,q}(V) \le \widehat{\nu}_{s,q}(V)$. □

5.7.8 Proof of Proposition 5.5.1

Let $x \in \mathbb{R}^n$, and let $x^1, ..., x^q$ be obtained from x by the following construction: x^1 is obtained from x by zeroing all but the s largest in magnitude entries; x^2 is obtained by the same procedure applied to $x - x^1$, x^3 – by the same procedure applied to $x - x^1 - x^2$, and so on; the process is terminated at the first step q when it happens that $x = x^1 + ... + x^q$. Note for $j \geq 2$ we have $\|x^j\|_\infty \leq s^{-1}\|x^{j-1}\|_1$ and $\|x^j\|_1 \leq \|x^{j-1}\|_1$, whence also $\|x^j\|_2 \leq \sqrt{\|x_j\|_\infty \|x_j\|_1} \leq s^{-1/2}\|x^{j-1}\|_1$. Recall that if A is RIP$(\delta, 2s)$, then for every two s-sparse vectors u, v with non-overlapping support we have

$$|u^T A^T A u| \leq \delta \|u\|_2 \|v\|_2. \qquad (*)$$

(i): We have

$$\begin{aligned}
&\|Ax^1\|_2\|Ax\|_2 \geq [x^1]^T A^T Ax = \|Ax^1\|_2^2 - \textstyle\sum_{j=2}^q [x^1]^T A^T Ax^j \\
&\geq \|Ax^1\|_2^2 - \delta \textstyle\sum_{j=2}^t \|x^1\|_2 \|x^j\|_2 \text{ [by } (*)] \\
&\geq \|Ax^1\|_2^2 - \delta s^{-1/2}\|x^1\|_2 \textstyle\sum_{j=2}^q \|x^{j-1}\|_1 \geq \|Ax^1\|_2^2 - \delta s^{-1/2}\|x^1\|_2\|x\|_1 \\
\Rightarrow\ & \|Ax^1\|_2^2 \leq \|Ax^1\|_2\|Ax\|_2 + \delta s^{-1/2}\|x^1\|_2\|x\|_1 \\
\Rightarrow\ & \|x^1\|_2 = \tfrac{\|x^1\|_2}{\|Ax^1\|_2^2}\|Ax^1\|_2^2 \leq \tfrac{\|x^1\|_2}{\|Ax^1\|_2}\|Ax\|_2 + \delta s^{-1/2}\left(\tfrac{\|x^1\|_2}{\|Ax^1\|_2}\right)^2 \|x\|_1 \\
\Rightarrow\ & \|x\|_{s,2} = \|x^1\|_2 \leq \tfrac{1}{\sqrt{1-\delta}}\|Ax\|_2 + \tfrac{\delta s^{-1/2}}{1-\delta}\|x\|_1 \text{ [by RIP}(\delta, 2s)]
\end{aligned}$$

and we see that the pair $\left(H = \frac{s^{-1/2}}{\sqrt{1-\delta}} I_m, \|\cdot\|_2\right)$ satisfies $\mathbf{Q}_{s,2}(\frac{\delta}{1-\delta})$, as claimed in (i).

(ii): We have

$$\begin{aligned}
&\|x^1\|_1\|A^T Ax\|_\infty \geq [x^1]^T A^T Ax = \|Ax^1\|_2^2 - \textstyle\sum_{j=2}^q [x^1]^T A^T Ax^j \\
&\geq \|Ax^1\|_2^2 - \delta s^{-1/2}\|x^1\|_2\|x\|_1 \text{ [exactly as above]} \\
\Rightarrow\ & \|Ax^1\|_2^2 \leq \|x^1\|_1\|A^T Ax\|_\infty + \delta s^{-1/2}\|x^1\|_2\|x\|_1 \\
\Rightarrow\ & (1-\delta)\|x^1\|_2^2 \leq \|x^1\|_1\|A^T Ax\|_\infty + \delta s^{-1/2}\|x^1\|_2\|x\|_1 \text{ [by RIP}(\delta, 2s)] \\
& \leq s^{1/2}\|x^1\|_2\|A^T Ax\|_\infty + \delta s^{-1/2}\|x^1\|_2\|x\|_1 \\
\Rightarrow\ & \|x\|_{s,2} = \|x^1\|_2 \leq \tfrac{s^{1/2}}{1-\delta}\|A^T Ax\|_\infty + \tfrac{\delta}{1-\delta} s^{-1/2}\|x\|_1,
\end{aligned}$$

and we see that the pair $\left(H = \frac{1}{1-\delta} A, \|\cdot\|_\infty\right)$ satisfies the condition $\mathbf{Q}_{s,2}\left(\frac{\delta}{1-\delta}\right)$, as required in (ii).

(iii): By (i) we have

$$\forall x \in \mathbb{R}^n : \|x\|_{s,2} \leq \frac{1}{\sqrt{1-\delta}}\|Ax\|_2 + \underbrace{\frac{\delta}{1-\delta}s^{-1/2}}_{\omega}\|x\|_1,$$

whence for every i it holds

$$\forall x : e_i^T x \leq \frac{1}{\sqrt{1-\delta}}\|Ax\|_2 + \omega\|x\|_1 \ \forall x,$$

that is,

$$\mathrm{Opt}(P) = \min_{x,t}\left\{t - e_i^T x : t \geq \frac{1}{\sqrt{1-\delta}}\|Ax\|_2, \|x\|_1 \leq 1\right\} \geq -\omega.$$

The conic optimization problem in the right hand side of the latter relation clearly is strictly feasible and bounded, whence by Conic Duality Theorem the dual problem

$$\mathrm{Opt}(D) = \max_{g,f,s}\left\{-s : -\frac{1}{\sqrt{1-\delta}}A^T g - f = e_i,\ \|g\|_2 \leq 1, \|f\|_\infty \leq s\right\}$$

is solvable with optimal value $\mathrm{Opt}(D) = \mathrm{Opt}(P) \geq -\omega$. Thus, there exist vectors g^i and $f = f^i$ such that

$$-\frac{1}{\sqrt{1-\delta}}A^T g^i - f^i = e_i,\ \|g^i\|_2 \leq 1, \|f^i\|_\infty \leq \omega.$$

Setting $h^i = -\frac{1}{\sqrt{1-\delta}}g^i$ and $H = [h^1, ..., h^n]$, we get $\|h^i\|_2 \leq \frac{1}{\sqrt{1-\delta}}$ and $\|A^T h^i - e_i\|_\infty \leq \omega$, whence for every $x \in \mathbb{R}^n$ it holds $|x_i| \leq |x^T A^T h^i| + \omega\|x\|_1$, so that

$$\|x\|_\infty \leq \|H^T Ax\|_\infty + \omega\|x\|_1 = \frac{\delta}{1-\delta}s^{-1/2}\|x\|_1 \leq \|H^T Ax\|_\infty + \frac{1}{3\widehat{s}}\|x\|_1. \qquad \square$$

(iv): We start with proving the fact about Gaussian matrices mentioned in (iv).

Lemma 5.7.1 *Let $1 \leq s \leq m \leq n$ be such that $n \geq 1000$ and $\delta := 4\sqrt{\frac{s\ln(n)}{m}} \leq 1/3$, and let A be a Gaussian $m \times n$ random matrix (i.e., random $m \times n$ matrix with independent $\mathcal{N}(0, 1/m)$ entries). Then*

$$\mathrm{Prob}\{A \text{ does not satisfy } \mathrm{RIP}(\delta, s)\} \leq 1/n. \tag{223}$$

Proof. 1^0. Let I be an s-element subset of $\{1, ..., n\}$, let x be a $\|\cdot\|_2$-unit vector supported on I, let $\theta = 0.005$, $\Delta = (1-4\theta)\delta$, and let $\mathcal{A}_x = \{A \in \mathbb{R}^{m\times n} : 1 - \Delta \leq \|Ax\|_2^2 \leq 1 + \Delta\}$. Then

$$\mathrm{Prob}\{A \notin \mathcal{A}_x\} \leq 2\exp\{-m\kappa\Delta^2\},\ \kappa = 3(1 - 3\ln(4/3))/2. \tag{224}$$

Indeed, $\zeta = \sqrt{m}Ax$ is a vector with independent $\mathcal{N}(0,1)$ entries. Therefore , setting

$$f(\gamma) = \ln\left(\frac{1}{\sqrt{2\pi}}\int \exp\{\gamma s^2\}\exp\{-s^2/2\}ds\right) = -\frac{1}{2}\ln(1-2\gamma),$$

for $-1/2 < \gamma < 1/2$, and applying Bernstein bounding scheme, we get

$$\begin{array}{rcl}
\forall \gamma \in [0,1/2): & & \\
\ln(p_+) & := & \ln\left(\mathrm{Prob}\{\|Ax\|_2^2 > (1+\Delta)\}\right) = \ln\left(\mathrm{Prob}\{\|\zeta\|_2^2 \geq m(1+\Delta)\}\right) \\
& \leq & mf(\gamma) - \gamma m(1+\Delta) = -\frac{m}{2}\ln(1-2\gamma) - \gamma m(1+\Delta) \\
\Rightarrow \ln(p_+) & \leq & \inf\limits_{0\leq\gamma<1/2}\left[-\frac{m}{2}\ln(1-2\gamma) - \gamma m(1+\Delta)\right] \\
& = & m\left[\frac{1}{2}\ln(1+\Delta) - \Delta/2\right] \leq -m\kappa\Delta^2, \\
\forall \gamma \in (-1/2,0): & & \\
\ln(p_-) & := & \ln\left(\mathrm{Prob}\{\|Ax\|_2^2 < (1-\Delta)\}\right) = \ln\left(\mathrm{Prob}\{\|\zeta\|_2^2 \leq m(1-\Delta)\}\right) \\
& \leq & mf(\gamma) - \gamma m(1-\Delta) = -\frac{m}{2}\ln(1-2\gamma) - \gamma m(1-\Delta) \\
\Rightarrow \ln(p_-) & \leq & \inf\limits_{-1/2<\gamma\leq 0}\left[-\frac{m}{2}\ln(1-2\gamma) - \gamma m(1-\Delta)\right] \\
& = & m\left[\frac{1}{2}\ln(1-\Delta) + \Delta/2\right] \leq -m\kappa\Delta^2,
\end{array}$$

and (224) follows; note that we have used the evident fact $\frac{1}{2}\ln(1+s) - \frac{1}{2}s \leq -\kappa s^2$, $-1/3 < s \leq 1/3$.

2^0. Let I be a subset of $\{1,...,n\}$ of cardinality s, let $S = \{u \in \mathbb{R}^s : \|u\|_2 = 1\}$, let $\epsilon = \mu\delta$, and let $\mathcal{S}$ be a minimal ϵ-net, w.r.t. $\|\cdot\|_2$, in S, and N be the cardinality of $\mathcal{S}$. Finally, let A_I be the submatrix of A comprised of columns with indices in I, and let $B = A_I^T A_I$. We claim that

$$\left\{(1-\Delta) \leq u^T Bu \leq (1+\Delta)\,\forall u \in \mathcal{S}\right\} \Rightarrow \left\{(1-\delta) \leq u^T Bu \leq (1+\delta)\,\forall u \in S\right\}. \qquad (225)$$

Indeed, let the premise in (225) hold true. Let β be the spectral norm of the positive semidefinite symmetric matrix B, and $\bar{u} \in S$ be such that $u^T Bu = \beta$. There exists $v \in \mathcal{S}$ such that $\|u - v\|_2 \leq \epsilon$, whence $\beta = u^T Bu \leq 2\beta\epsilon + v^T Bv \leq 2\beta\epsilon + 1 + \Delta$ (since the quadratic form $z^T Bz$ is Lipschitz continuous, with constant 2β w.r.t. $\|\cdot\|_2$, on S), whence $\beta \leq \frac{1+\Delta}{1-2\epsilon} \leq 1+\delta$, where the concluding inequality is given by a straightforward computation.

Thus, $z^TBz \le 1+\delta$ for all $z \in S$. By similar reasons, if $u \in S$, and if $v \in \mathcal{S}$ is such that $\|u-v\|_2 \le \epsilon$, we have $u^TBu \ge v^TBv - 2\beta\epsilon \ge 1-\Delta-2(1+\delta)\epsilon = 1-\delta$. Thus, the conclusion in (225) holds true.

3^0 We can straightforwardly build an ϵ-net $\mathcal{S}'$ in S in such a way that the $\|\cdot\|_2$-distance between every two distinct points of the net is $> \epsilon$, so that the balls $B_v = \{z \in \mathbb{R}^s : \|z-v\|_2 \le \epsilon/2\}$ with $v \in \mathcal{S}$ are mutually disjoint. Since the union of these balls belongs to $B = \{z \in \mathbb{R}^s : \|z\|_2 \le 1+\epsilon/2\}$, we get $\mathrm{Card}(\mathcal{S}')(\epsilon/2)^s \le (1+\epsilon/2)^s$, that is, $N \le \mathrm{Card}(\mathcal{S}') \le (1+2/\epsilon)^s = (1+2/(\mu\delta))^s$. Invoking (224), we see that the probability of violating the premise in (225) for a given I does not exceed $\exp\{-m\kappa\delta^2 + \ln(2) + s\ln(1+2/(\mu\delta))\}$. The number of M of s-element subsets of $\{1,...,n\}$ does not exceed $\exp\{s\ln(n)\}$, and we conclude that the probability for A be such that $x^TA^TAx \notin [1-\delta, 1+\delta]$ for a $\|\cdot\|_2$-unit vector x with at most s nonzero entries does not exceed

$$p = \exp\{-m\kappa\delta^2 + \ln(2) + s\ln(n(1+2/(\mu\delta)))\}.$$

It is immediately seen that with the just defined δ and in the range $1 \le s \le m \le n$, $n \ge 1000$, we get $p \le 1/n$. □

Let $A \in \mathbb{R}^{m\times n}$ satisfy RIP(δ, s). Let x^1 be the vector obtained from x by zeroing all but the s largest in magnitude entries in x, x^2 is obtained by the same procedure from $x - x^1$, x^3 obtained by the same procedure from $x - x^1 - x^2$, and so on, until a vector x^q with at most s nonzero entries is built. Let, further, $h \in \mathbb{R}^n$ be supported on the support of x^1. We have

$$\begin{array}{rcl} j \ge 2 & \Rightarrow & \|x^j\|_\infty \le s^{-1}\|x^{j-1}\|_1 \\ & \Rightarrow & \|x^j\|_2 \le \|x^j\|_\infty^{1/2}\|x^j\|_1^{1/2} \le \|x^j\|_\infty^{1/2}\|x^{j-1}\|_1^{1/2} \le s^{-1/2}\|x^{j-1}\|_1. \end{array}$$

Moreover,

$$\begin{aligned}\|h\|_1\|A^TAx\|_\infty &\geq h^TA^TAx \geq h^TA^TAx^1 - \sum_{j=2}^{q} h^TA^TAx^j \\ &\geq h^TA^TAx - \delta\sum_{j=2}^{q}\|h\|_2\|x^j\|_2 \\ &\geq h^TA^TAx^1 - \delta s^{-1/2}\sum_{j=1}^{q-1}\|h\|_2\|x^j\|_1 \\ &\geq h^TA^TAx^1 - \delta s^{-1/2}\|h\|_2\|x\|_1\end{aligned}$$

Now let I be the support of x^1, A_I be the submatrix of A comprised of columns with indexes from I, and $B = A_I^TA_I$. Since h is supported on I, we have $h^TA^TAx^1 = h^TBx^1$. Now let $i_* \in I$ be the index of the largest in magnitude entry in x^1, and h be i_*-th basic orth times sign(x_{i_*}). Observing that the spectral norm of the difference D of B and the unit $s \times s$ matrix does not exceed δ, we have $h^TBx^1 \geq h^Tx^1 - h^TDx^1 = h^Tx^1 - \|Dx^1\|_2 = |x_{i_*}| - \delta\|x^1\|_2 \geq |x_{i_*}| - \delta\sqrt{s}|x_{i_*}|$, we get

$$\begin{aligned}\delta\sqrt{s} < 1 \quad &\Rightarrow \quad \|x\|_\infty(1-\delta\sqrt{s}) \leq h^TA^TAx^1 \leq \|h\|_1\|A^TAx\|_\infty + \delta s^{-1/2}\|h\|_2\|x\|_1 \\ &\Rightarrow \quad \|x\|_\infty \leq \frac{1}{1-\delta\sqrt{s}}\|A^TAx\|_\infty + \frac{\delta s^{-1/2}}{1-\delta\sqrt{s}}\|x\|_1.\end{aligned}$$

Now let $\delta = 4\sqrt{\frac{s\ln(n)}{m}}$; recall that when $n \geq 1000$ and $4\sqrt{\frac{s\ln(n)}{m}} \leq 1/3$, the probability for a Gaussian $m \times n$ matrix to satisfy RIP(δ, s) is at least $1 - 1/n$. When A is RIP$(4\sqrt{\frac{s\ln(n)}{m}}, s)$ and $\delta\sqrt{s} \leq 1/3$, the above computation shows that

$$\forall x : \|x\|_\infty \leq \frac{3}{2}\|A^TAx\|_\infty + \frac{3}{2}\delta s^{-1/2}\|x\|_1 = \frac{3}{2}\|A^TAx\|_\infty + 12\sqrt{\ln(n)/m}\|x\|_1,$$

so that in the case of (198) $(\frac{3}{2}A, \|\cdot\|_\infty)$ satisfies $\mathbf{Q}_{s,q}(1/3)$. □

5.7.9 Proof of Proposition 5.5.2

By homogeneity reasons, it suffices to consider the case of $\sigma = 1$, which we assume from now on.

Assume that (205) takes place for every s-sparse x, and let $\rho = \sqrt{2\ln(n/\epsilon)}$.

0^0. For a typical Gaussian $A \in \mathbb{R}^{m\times n}$ with $\ln(m) = O(1)\ln(n)$ and m, n large, we have

$$(a) \quad 0.99 \le \|A_i\|_2^2 \le 1.01,\ 1 \le i \le n,$$
$$(b) \quad \mu(A) := \max_{1\le i<j\le n} |A_i^T A_j| \le O(1)\sqrt{\ln(m)/m}.$$

From now on we assume that these relations hold true.

1^0. Let I be a subset of $\{1, ..., n\}$ including 1 and of cardinality $\le s$, and let $X^I = \{x \in \mathbb{R}^n : x_1 \ge 2C\rho, x_i = 0\,\forall i \notin I\}$ and $Y^I = A^T A X^I$. We claim that

$$\min_{y\in Y^I} \|y\|_\infty > \rho.$$

Indeed, otherwise we could find s-sparse signal $\bar{x}$ with $\bar{x}_1 \ge 2C\rho$ such that $\|A^T A\bar{x}\|_\infty \le \rho$, meaning that

$$\|A^T(A\bar{x} + \xi)\| \le 2\rho$$

for a typical realization of ξ. For such a ξ, we clearly have $\widehat{x} = 0$, whence $\|\widehat{x} - \bar{x}\|_\infty > C\rho$ for a typical ξ, which is impossible, since by (205) for a typical ξ the opposite inequality takes place.

Thus, the convex set Y^I is at $\|\cdot\|_\infty$ distance from the origin at least ρ, meaning that there exists $y^I \in \mathbb{R}^m$ with $\|y^I\|_1 = 1$ such that $[y^I]^T A^T A x \ge \rho$ for all $x \in X^I$, or, which is the same,

$$\langle \sum_{i=1}^n y_i^I A_i, A_\ell \rangle = \begin{cases} 0, & 2 \le \ell \in I \\ c_I \ge (2C)^{-1}, & \ell = 1 \end{cases},$$

Setting $\lambda_i^I = y_i^I / c_I$, we conclude that there exists vector $\lambda^I \in \mathbb{R}^n$ such that

$$\begin{aligned} &(a) \quad \|\lambda^I\|_1 \le 2C \\ &(b) \quad \langle A\lambda^I, A_1\rangle = 1 \\ &(c) \quad \langle A\lambda^I, A_\ell\rangle = 0,\ \ell \in I\backslash\{1\} \end{aligned} \tag{226}$$

2^0. Let I be as above. When $\ell = 1$, we have

$$1 = \langle A\lambda^I, A_1\rangle = \lambda_1^I \|A_1\|_2^2 + \sum_{i>1} \lambda_i^I A_i^T A_1$$

and $|\sum_{i>1} \lambda_i^I A_i^T A_1| \le \|\lambda\|_1 \mu(A)$, whence

$$|1 - \lambda_1^I \|A_1\|_2^2| \le 2C\mu(A),$$

whence $0.98 \le \lambda_1^I \le 1.02$, provided m is large enough. Similarly, for every $\ell \in I\backslash\{1\}$ we have

$$0 = \langle A\lambda^I, A_\ell\rangle = \lambda_\ell^I \|A_\ell\|_2^2 + \sum_{i\neq\ell} \lambda_i^I A_i^T A_\ell$$

and $|\sum_{i\neq\ell} \lambda_i^I A_i^T A_\ell| \le \|\lambda^I\|_1 \mu(A)$. Thus, we can assume that

$$\begin{array}{ll} (a) & 0.98 \le \lambda_1^I \le 1.02, \\ (b) & 1 \neq \ell \in I \Rightarrow |\lambda_\ell^I| \le O(1)C\sqrt{\ln(m)/m}. \end{array} \tag{227}$$

3^0. Let $k = s/2$ be integer, and let $I^* = \{1, ..., k\}$. Let

$$\Lambda = \{\lambda \in \mathbb{R}^n : \|\lambda\|_1 \le 2C, \langle A\lambda, A_\ell\rangle = \left\{ \begin{array}{ll} 1, & \ell = 1 \\ 0, & 2 \le \ell \le k \end{array} \right\}$$

so that Λ is a closed convex set in $\mathbb{R}^n$ which is nonempty (indeed, by (226) we have $\lambda^{I^*} \in \Lambda$). For $\lambda \in \Lambda$, let λ^- be the vector obtained from λ by zeroing the first k coordinates, and let

$$\gamma = \min_{\lambda\in\Lambda} \|\lambda^-\|_\infty.$$

We claim that

$$\gamma \le O(1)C\left[1/s + \sqrt{\ln(m)/m}\right] \tag{228}$$

Indeed, assuming $\gamma > 0$, there exists $e \in \mathbb{R}^n$, $\|e\|_1 = 1$, with $e_1 = e_2 = ... = e_k = 0$, such that $e^T\lambda \ge \gamma$ for every $\lambda \in \Lambda$. Let $i_1, ..., i_k$ be the indexes of the k largest in magnitude entries in e. Setting $I = \{1, 2, ..., k\} \cup \{i_1, ..., i_k\}$ and invoking (226) and (227), we have $\lambda^I \in \Lambda$ and therefore we should have $\gamma \le e^T[\lambda^I]^- \le \sum_{\ell=1}^k e_{i_\ell}[\lambda^I]_{i_\ell} + \|\lambda^I\|_1 \max_{i\notin\{i_1,...,i_k\}} |e_i|$. Invoking (227) and taking into account that $\|e\|_1 = 1$, we get $\sum_{\ell=1}^k e_{i_\ell}[\lambda^I]_{i_\ell} \le O(1)C\sqrt{\ln(m)/m}$, while $\|\lambda^I\|_1 \max_{i\notin\{i_1,...,i_k\}} |e_i| \le 2C/k$ due to $\lambda^I \in \Lambda$ and $\|e\|_1 = 1$, and (228) follows.

From (228) it follows that there exists $\bar\lambda \in \Lambda$ such that

$$|\bar\lambda_i| \le O(1)C\left[1/s + \sqrt{\ln(m)/m}\right], \ i > k$$

Besides this, from $\bar\lambda \in \Lambda$ it follows that $\|\bar\lambda\|_1 \le 2C$ and

$$\langle A\bar\lambda, A_i\rangle = \left\{ \begin{array}{ll} 1, & i = 1 \\ 0, & 2 \le i \le k \end{array} \right.$$

whence, applying the reasoning which led us to (227) to $\bar\lambda$ in the role of λ^I and I^* in the role of I,

$$0.98 \le \bar\lambda_1 \le 1.02, \ \ |\bar\lambda_i| \le O(1)C\sqrt{\ln(m)/m},\ 2 \le i \le k,$$

provided m is large enough. Thus,

$$\begin{array}{ll} (a) & \|\bar\lambda\|_1 \le 2C \\ (b) & 0.98 \le \bar\lambda_1 \le 1.02 \\ (c) & |\bar\lambda_i| \le O(1)C\left[1/s + \sqrt{\ln(m)/m}\right],\ 2 \le i \le n \\ (d) & \langle A\bar\lambda, A_i\rangle = 0, 2 \le i \le k. \end{array} \tag{229}$$

4^0. Let f be the orthogonal projection of A_1 onto the linear span L of $A_2, ..., A_k$. Since A is a Gaussian matrix, the typical Euclidean norm of f is $\ge O(1)\sqrt{(k-1)/m} = O(1)\sqrt{s/m}$. Now consider the vector

$$g = \sum_{i=2}^{\infty} \bar\lambda_i A_i = A\bar\lambda - \bar\lambda_1 A_1.$$

Note that by (229.d) the orthogonal projection of $A\bar\lambda$ onto L is zero. It follows that the orthogonal projection of g onto L is $-\bar\lambda_1 f$, whence

$$g^T g \ge \bar\lambda_1^2 f^T f \ge O(1)s/m.$$

On the other hand, $g = \sum_i \lambda_i A_i$ with $\|\lambda\|_1 \le 2C$ and $\|\lambda\|_\infty \le \delta := O(1)C[1/s + \sqrt{\ln(m)/m}]$, whence

$$\begin{array}{rcl} g^T g & = & \sum_{i,j} \lambda_i\lambda_i A_i^T A_j \le \sum_i \lambda_i^2 \|A_i\|_2^2 + \sum_{i\ne j} |\lambda_i||\lambda_j|\mu(A) \\ & \le & 1.01\|\lambda\|_1\|\lambda\|_\infty + \|\lambda\|_1^2\mu(A) \le O(1)C^2\left[[1/s + \sqrt{\ln(m)/m}] + \sqrt{\ln(m)/m}\right], \end{array}$$

and we arrive at the inequality

$$s/m \le O(1)C^2[1/s + \sqrt{\ln(m)/m}],$$

whence

$$s \le O(1)\max[C, C^2]\sqrt{\ln(m)m}.$$ □

5.7.10 Proof of Proposition 5.6.1

The proof below follows the lines of the proofs of Proposition 7 of [76] and Proposition 10 of [79]. Given $\epsilon \in (0,1)$, let $\Xi = \{\xi : \max_{u\in\mathcal{U}} h_i^T u + \sigma|h_i^T\xi| \le \nu(H),\ 1 \le i \le N\}$, so that $\mathrm{Prob}\{\xi \in \Xi\} \ge 1-\epsilon$. Let us fix $\xi \in \Xi$, $u \in U$ and $x \in \mathbb{R}^n$ such that $L_1(Bx - [Bx]^s) \le v$. For $\eta = y - Ax = u + \sigma\xi$, using the definition of Ξ and the fact that $\mathcal{U}$ is a symmetric set, we have

$$\begin{array}{rcl}\|H^T\eta\|_\infty &=& \max\limits_i |h_i^T u + \sigma h_i^T\xi| \le \max\limits_i\{|h_i^T u| + \sigma|h_i^T\xi|\}\\ &\le& \max\limits_i \left\{\max\limits_{u\in\mathcal{U}} h_i^T u + \sigma|h_i^T\xi|\right\} \le \nu(H) = \omega_*(\bar\gamma).\end{array}$$

We will proceed by induction. First, let us show that (a_{k-1}, b_{k-1}) implies (a_k, b_k). Thus, assume that (a_{k-1}, b_{k-1}) holds true. Let $z^{(k-1)} = x - v^{(k-1)}$. By (a_{k-1}), $Bz^{(k-1)}$ is supported on the support of Bx. Note that

$$\begin{array}{rcl}Bz^{(k-1)} - g &=& Bx - Bv^{(k-1)} - H^T(y - Av^{(k-1)}) = (B - H^TA)(x - v^{(k-1)}) - H^T\eta\\ &=& VBz^{(k-1)} - H^T\eta,\end{array}$$

where the last equality follows from $B = VB + H^TA$.

$$\begin{array}{rcl}\|(Bz^{(k-1)} - g)[i]\|_\infty &=& \|(VBz^{(k-1)} - H^T\eta)[i]\|_\infty\\ &\le& \|(VBz^{(k-1)})[i]\|_\infty + \|H^T\eta[i]\|_\infty\\ &\le& \|\sum\limits_\ell V^{i\ell}(Bz^{(k-1)})[\ell]\|_\infty + \omega_*(\bar\gamma)\\ &\le& \sum\limits_\ell \|V^{i\ell}\|_{\infty\to\infty}\|(Bz^{(k-1)})[\ell]\|_\infty + \omega_*(\bar\gamma)\\ &\le& \bar\gamma L_1(Bz^{(k-1)}) + \omega_*(\bar\gamma),\end{array}$$

consequently,

$$\|(Bz^{(k-1)} - g)[i]\|_\infty \le \bar\gamma\alpha_{k-1} + \omega_*(\bar\gamma) := \gamma. \tag{230}$$

We conclude that for any $1 \le i \le K$ and any $1 \le j \le n_i$, the interval $S[i]_j = [g[i]_j - \gamma, g[i]_j + \gamma]$ of the width $\ell = 2\bar\gamma\alpha_{k-1} + 2\omega_*(\bar\gamma)$, covers $(Bz^{(k-1)})[i]_j$ and the closest to 0

point of $S[i]_j$ is

$$\begin{array}{rcll}\widetilde{\Delta}[i]_j & = & [g[i]_j - \gamma]_+ & g[i]_j \geq 0, \\ \widetilde{\Delta}[i]_j & = & -[|g[i]_j| - \gamma]_+ & g[i]_j < 0,\end{array}$$

that is, $\widetilde{\Delta}[i]_j = \Delta[i]_j$ for all $1 \leq i \leq K$ and $1 \leq j \leq n_i$. Since the segment $S[i]_j$ covers $(Bz^{(k-1)})[i]_j$ and $\Delta[i]_j$ is the closest to 0 point in $S[i]_j$, while the width of $S[i]_j$ is at most ℓ, we clearly have

$$(a) \quad \Delta[i]_j \in \mathrm{Conv}\left\{0, (Bz^{(k-1)})[i]_j\right\}, \qquad (b) \quad L_\infty(Bz^{(k-1)} - \Delta) \leq \ell. \tag{231}$$

Since (a_{k-1}) is valid, (231.a) implies that

$$(Bv^{(k)})[i]_j = (Bv^{(k-1)} + \Delta)[i]_j \in \underbrace{\left[(Bv^{(k-1)})[i]_j + \mathrm{Conv}\left\{0, (Bx - Bv^{(k-1)})[i]_j\right\}\right]}_{\subseteq \mathrm{Conv}\{0, (Bx)[i]_j\}},$$

and (a_k) holds. Further, let $I \subset \{1, ..., K\}$ be the set of indices of the s largest in magnitude blocks in Bx and $\bar{I} = \{1, \ldots, K\} \setminus I$. Relation (a_k) clearly implies that $\|(Bz^{(k)})[i]\|_\infty \leq \|(Bx)[i]\|_\infty$, and we can write due to (231.b):

$$\begin{array}{rcl} L_1(Bx - Bv^{(k)}) & = & L_1([Bx - Bv^{(k-1)} - \Delta]_I) + L_1([Bz^{(k)}]_{\bar{I}}) \\ & \leq & \sum_{i \in I} \|(Bz^{(k-1)} - \Delta)[i]\|_\infty + \sum_{i \notin I} \|(Bx)[i]\|_\infty \leq s\ell + v = \alpha_k. \end{array}$$

Since by (231.b)

$$L_\infty(Bx - Bv^{(k)}) = L_\infty(Bx - Bv^{(k-1)} - \Delta) \leq \ell = 2\bar{\gamma}\alpha_{k-1} + 2\omega_*(\bar{\gamma}),$$

we conclude that (b_k) is satisfied. The induction step is justified.

It remains to show that (a_0, b_0) holds true. Since (a_0) is evident, all we need is to justify (b_0). Let

$$\alpha_* = L_1(Bx),$$

and let $g = H^T y$. Same as above (cf. (230)), we have for all i:

$$L_\infty(Bx - g) \leq 2\bar{\gamma}\alpha_* + 2\omega_*(\bar{\gamma}).$$

Then

$$\begin{aligned}\alpha_* &= L_1(Bx) = L_1([Bx]_I) + L_1(Bx - [Bx]^s) \\ &\leq \sum_{i\in I}[\|g[i]\|_\infty + \bar{\gamma}\alpha_* + \omega_*(\bar{\gamma})] + v \leq L_{s,1}(g) + s\bar{\gamma}\alpha_* + s\omega_*(\bar{\gamma}) + v.\end{aligned}$$

Hence

$$\alpha_* \leq \alpha_0 = \frac{L_{s,1}(g) + s\omega_*(\bar{\gamma}) + v}{1 - s\bar{\gamma}},$$

which implies (b_0). □

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this thesis we have investigated new techniques for both analysis and design of tractable relaxations and efficient algorithms with good performance guarantees in large-scale sparsity-oriented optimization. In terms of developing tractable relaxations, we have

- developed a unifying framework for building tractable relaxations for disjoint bilinear programs, based on linear and semidefinite programming;
- investigated the benefits of using additional information given in the form of sign restrictions for sparse ℓ_1-recovery, and presented necessary and sufficient as well as verifiable sufficient conditions which generalize their previous counter parts from the literature and analyzed their limits of performance;
- demonstrated that our verifiable sufficient conditions can be utilized in the efficient design of a measurement matrix with performance guarantees on the quality of recovery;
- investigated conditions and proposed new recovery methods for a more general sparse estimation problem –estimating a signal from its undersampled observations corrupted with nuisance and stochastic noise under the assumption that a known linear transform of the signal admits a good block-sparse approximation in a given block representation structure. We have shown that all our previous results on goodness of ℓ_1-recovery can be extended to this setting.

Motivated by the specific convex optimization problems arising in compressed sensing, we have introduced and investigated the notion of a generalized bilinear saddle point problem (GBSP). Specifically:

- We have shown that many interesting classes of problems from compressed sensing recovery and machine learning can be cast as GBSPs.

- We have suggested efficient first-order methods for solving GBSPs equipped with deterministic and stochastic oracles. These methods admit explicit theoretical efficiency estimates which for numerous applications, including those in compressed sensing and machine learning, are the best known so far.

- Through numerical experiments, we have shown that our algorithms with stochastic oracles besides achieving sublinear time behavior, exhibit excellent computational performance.

- We have investigated the effect of using previous information in our algorithms through testing different continuation schemas paired with new distance generating functions.

Compressed sensing is an emerging field with active research particularly because of its great promise for acquiring and processing massive data efficiently and accurately. Yet, its solid theoretical foundation is still evolving. Optimization techniques will play a key role in the development of this foundation. A few topics that are worth mentioning for future studies can be categorized as follows:

Efficient Algorithms for Block-Sparse Recovery: In Chapter 5, we have suggested new algorithms with provable accuracy certificates for a more general sparse estimation problem with stochastic noise. One of the important features of this problem is that it assumes the signals are *block-sparse with respect to a given representation structure* and specifically considers the problem of estimating a linear transform $Bx \in \mathbb{R}^N$ of a vector $x \in \mathbb{R}^n$, where Bx is assumed to be block sparse, from the observations

$$y = Ax + u + \sigma\xi.$$

Consequently in all of the resulting recovery procedures, a norm of this linear transform Bx, e.g., $L_1(Bx)$, appears in the objective. When the representation matrix, B is identity, these recovery procedures more or less reduces to the ℓ_1 regularization problems and we arrive back at the setting of Chapter 4. On the other hand, handling the general representation matrix B is a nontrivial and important task. There are a number of important applications

where a nontrivial B arises naturally such as standard image reconstruction with Total Variation regularization or finding the solution of a linear finite-difference equation with sparse right hand side. However the algorithms in Chapter 4 are explicitly making use of the ℓ_1 objective function and similar transformations for norms involving nontrivial B will no longer lead to "easy to solve" subproblems utilized in the state of the art first-order methods (or at least we don't know it yet). Therefore it will be an interesting task to develop efficient first-order methods for solving these estimation problems.

Verifiable Sufficient Conditions for Nuclear Norm Minimization: Minimizing the rank of a matrix subject to constraints is a challenging problem that arises in many applications in machine learning, control theory, and discrete geometry. This class of optimization problems, known as rank minimization, is NP-hard, and for most practical problems there are no efficient algorithms that yield exact solutions. A popular heuristic replaces minimizing the rank function of a matrix with minimizing its nuclear norm –the sum of the singular values. This practical approach has been shown to provide the optimal low rank solution in a variety of scenarios. The necessary and sufficient condition that characterizing when nuclear norm minimization finds the minimum rank matrix subject to linear matrix inequalities have been established in [114]. Despite the fact that the probabilistic performance bounds on the rank as a function of the matrix dimension and the number of constraints, for which the nuclear norm minimization succeeds with overwhelming probability are provided in the literature (see [52, 113, 114]), there is no tractable way of verifying these conditions. It will be interesting to provide verifiable sufficient conditions in this setting. The main difficulty in extending the previous work and ideas from ℓ_1-recovery to this setting lies in the fact that in the space of matrices, the unit ball in the nuclear norm is not a polyhedral set, i.e., it has infinitely many extreme points, as opposed to the ℓ_1-ball in $\mathbb{R}^n$.

Parallel Implementation of Efficient Algorithms for Compressed Sensing and Their Applications: The first-order methods presented in Chapter 4 provide excellent theoretical (they exhibit sublinear time behavior) and practical performance, yet they don't

take advantage of the emerging parallel and/or distributed computing architectures. Designing and implementing algorithms with parallel/distributed computing in mind to achieve both superb practical performance (thus satisfying the increasing demand for handling larger instances faster) and optimal rate of convergence, and conducting more computational studies for large-scale first-order methods in these settings will be very rewarding.

REFERENCES

[1] Achlioptas, D. and McSherry, F., "Fast computation of low rank matrix approximations," *Journal of the ACM*, vol. 54, pp. 1–19, 2007.

[2] Agarwal, A., Negahban, S. N., and Wainwright, M. J., "Fast global convergence of gradient methods for high-dimensional statistical recovery," tech. rep., 2011. http://arxiv.org/abs/1104.4824.

[3] Al-Khayyal, F. and Falk, J., "Jointly constrained biconvex programming," *Math. Oper. Res.*, vol. 8, pp. 273–286, 1983.

[4] Andersen, E. D. and Andersen, K. D., "The MOSEK optimization tools manual." http://www.mosek.com/fileadmin/products/6_0/tools/doc/pdf/tools.pdf.

[5] Aybat, N. S. and Iyengar, G., "A unified approach for minimizing composite norms," tech. rep., 2009. http://arxiv.org/abs/1005.4733.

[6] Aybat, N. S. and Iyengar, G., "A first-order augmented lagrangian method for compressed sensing," tech. rep., 2010. http://arxiv.org/abs/1005.5582.

[7] Aybat, N. S. and Iyengar, G., "A first-order smoothed penalty method for compressed sensing," *SIAM J. Optim.*, vol. 21, no. 1, pp. 287–313, 2011.

[8] Azuma, K., "Weighted sums of certain dependent random variables," *Tökuku Math. J.*, vol. 19, pp. 357–367, 1967.

[9] Bach, F., "Consistency of the group lasso and multiple kernel learning," *J. Mach. Learn. Res.*, vol. 9, pp. 1179–1225, 2008.

[10] Baraniuk, R. G., Cevher, V., Duarte, M. F., and Hegde, C., "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[11] Baraniuk, R., Davenport, M., Duarte, M., and Hegdek, C., *Introduction to Compressive Sensing*. Connexions e-textbook, 2011.

[12] Becker, S., Bobin, J., and Candès, E. J., "Nesta: a fast and accurate first-order method for sparse recovery," *SIAM J. on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[13] Becker, S., Candès, E. J., and Grant, M., "Templates for convex cone problems with applications to sparse signal recovery," tech. rep., 2010. http://arxiv.org/abs/1009.2065.

[14] Ben-Haim, Z. and Eldar, Y. C., "Near-oracle performance of greedy block-sparse estimation techniques from noisy measurements," tech. rep., 2010. http://arxiv.org/abs/1009.0906.

[15] Berinde, R., Gilbert, A. C., Indyk, P., Karloff, H. J., and Strauss, M. J., "Combining geometry and combinatorics: A unified approach to sparse signal recovery," 2008.

[16] Berinde, R. and Indyk, P., "Sequential sparse matching pursuit," in *Proceedings of the* 47^{th} *Annual Allerton Conference on Communication, Control, and Computing*, pp. 36–43, 2009.

[17] Bickel, P. J., "Discussion: the Dantzig selector," *Annals of Stat.*, vol. 35, no. 6, pp. 2352–2357, 2007.

[18] Bickel, P. J., Ritov, Y., and Tsybakov, A. B., "Simultaneous analysis of Lasso and Dantzig selector," *Annals of Stat.*, vol. 37, no. 4, pp. 1705–1732, 2008.

[19] Blumensath, T. and Davies, M., "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[20] Bourgain, J., Dilworth, S. J., Ford, K., Konyagin, S., and Kutzarova, D., "Explicit constructions of rip matrices and related problems," tech. rep., 2010. http://arxiv.org/abs/1008.4535.

[21] Bruckstein, A. M., Donoho, D., and Elad, M., "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[22] Bruckstein, A. M., Elad, M., and Zibulevsky, M., "A non-negative and sparse enough solution of an underdetermined linear system of equations is unique," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.

[23] Bühlmann, P. and van de Geer, S., "On the conditions used to prove oracle results for the Lasso," *Electron. J. Statist.*, vol. 3, pp. 1360–1392, 2009.

[24] Bunea, F., Tsybakov, A. B., and Wegkamp, M. H., "Sparsity oracle inequalities for the Lasso," *Electron. J. Statist.*, vol. 1, pp. 169–194, 2007.

[25] Candès, E. J., "Compressive sampling," in *Proceedings of International Congress of Mathematicians (ICM)*, pp. 1437–1452, 2006.

[26] Candès, E. J., "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus de l'Acad. des Sci.*, vol. Serie I, no. 346, pp. 589–592, 2008.

[27] Candès, E. J. and Tao, T., "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203–4215, 2006.

[28] Candès, E., Romberg, J., and Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, 2006.

[29] Candès, E. and Tao, T., "The Dantzig selector: Statistical estimation when p is much larger than n," *Annals of Stat.*, vol. 35, no. 6, pp. 2313–2351, 2007.

[30] Carathéodory, C., "Ueber den variabilitaetsbereich der fourierschen konstanten von positiven harmonischen funktionen," *Rend. Circ. Mat. Palermo*, vol. 32, pp. 193–217, 1911.

[31] Chandar, V., "A negative result concerning explicit matrices with the restricted isometry property," tech. rep., 2008. http://dsp.rice.edu/files/cs/Venkat_CS.pdf.

[32] Chen, S. S., Donoho, D. L., and Saunders, M. A., "Atomic decomposition by basis pursuit," *SIAM J. of Scientific Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[33] Chesneau, C. and Hebiri, M., "Some theoretical results on the grouped variables Lasso," *Mathematical Methods of Statistics*, vol. 27, no. 4, pp. 317–326, 2008.

[34] Cohen, A., Dahmen, W., and Devore, R., "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc*, vol. 22, pp. 211–231, 2009.

[35] CPLEX, "IBM ILOG," 12.2. http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/.

[36] Cristiani, N. and Shawe-Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, NY: Cambridge University Press, 2000.

[37] Dalalyan, A. S., Juditsky, A. B., and Spokoiny, V., "A new algorithm for estimating the effective dimension-reduction subspace," *J. Mach. Learn. Res.*, vol. 9, pp. 1647–1678, 2008.

[38] d'Aspremont, A. and Ghaoui, L. E., "Testing the nullspace property using semidefinite programming," *Math. Progr.*, vol. 127, no. 1, pp. 123–144, 2010. Special issue on machine learning.

[39] Davenport, M., Duarte, M., Eldar, Y., and Kutyniok, G., *Introduction to Compressed Sensing*. Cambridge, NY: Cambridge University Press, 2011.

[40] de Castro, Y., "Error prediction and model selection via unbalanced expander graphs," tech. rep., 2011. http://arxiv.org/abs/1010.2457v4.

[41] Devore, R. A., "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4–6, pp. 918–925, 2007.

[42] Donoho, D. L. and Elad, M., "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Natl. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, 2003.

[43] Donoho, D. L., Elad, M., and Temlyakov, V. N., "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, pp. 6–18, 2006.

[44] Donoho, D. L., Elad, M., and Temlyakov, V. N., "On Lebesgue-type inequalities for greedy approximation," *J. Approx. Theory*, vol. 147, no. 2, pp. 185–195, 2007.

[45] Donoho, D. L. and Huo, X., "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[46] Donoho, D. L. and Tanner, J., "Neighborliness of randomly-projected simplices in high dimensions," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 27, pp. 9452–9457, 2005.

[47] Donoho, D. L. and Tanner, J., "Sparse nonnegative solutions of underdetermined linear equations by linear programming," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 27, pp. 9446–9451, 2005.

[48] Donoho, D. L., Tsaig, Y., Drori, I., and luc Starck, J., "Sparse solution of underdetermined linear equations by stagewise Orthogonal Matching Pursuit (StOMP)," tech. rep., 2006. http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.5221.

[49] Dorfman, R., "The detection of defective members of large populations," *Annals of Stat.*, vol. 14, pp. 436–440, 1943.

[50] Drineas, P., Kannan, R., and Mahoney, M. W., "Fast Monte Carlo algorithms for matrices II: Computing a low-rank approximation to a matrix," *SIAM J. Comput.*, vol. 36, pp. 158–183, 2006.

[51] Duarte, M., Bajwa, W., and Calderbank, R., "The performance of group Lasso for linear regression of grouped variables," tech. rep., 2011. http://www.duke.edu/~wb40/pubs/sampta11_tr.pdf.

[52] Dvijotham, K. and Fazel, M., "A nullspace analysis of the nuclear norm heuristic for rank minimization," in *Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 3586–3589, 2010.

[53] Elad, M., "Optimized projections for compressed sensing," *IEEE Trans. on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, 2007.

[54] Eldar, Y. C., Kuppinger, P., and Bölcskei, H., "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[55] Eldar, Y. C. and Mishali, M., "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[56] Elhamifar, E. and Vidal, R., "Structured sparse recovery via convex optimization," tech. rep., 2011. http://arxiv.org/pdf/1104.0654.

[57] Farias, V., Jagabathula, S., and Shah, D., "A non-parametric approach to modeling choice with limited data," tech. rep., 2010. http://arxiv.org/pdf/0910.0063.

[58] Fornasier, M., *Theoretical Foundations and Numerical methods for sparse recovery*. Radon Series Comp. Appl. Math., Germany: De Gruyter, 2010.

[59] Fornasier, M. and Rauhut, H., "Compressive Sensing," in *Handbook of Mathematical Methods in Imaging* (Scherzer, O., ed.), pp. 187–228, Springer, 2011.

[60] Fornasier, M. and Rauhut, H., "Iterative thresholding algorithms," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 2, pp. 187–208, 2008.

[61] Frieze, A., Kannan, R., and Vempala, S., "Fast Monte-Carlo algorithms for finding low-rank approximations," *Journal of the ACM*, vol. 51, pp. 1025–1041, 2004.

[62] Gilbert, A. C. and Indyk, P., "Sparse recovery using sparse matrices," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 937 – 947, 2010.

[63] Goldfarb, D., Ma, S., and Scheinberg, K., "Fast alternating linearization methods for minimizing the sum of two convex functions," tech. rep., 2009. http://arxiv.org/abs/0912.4571.

[64] Goldfarb, D., Ma, S., and Scheinberg, K., "Fast multiple splitting algorithms for convex optimization," tech. rep., 2009. http://arxiv.org/abs/0912.4570.

[65] Gribonval, R. and Nielsen, R., "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, pp. 3320–3325, 2003.

[66] Grigoriadis, M. D. and Khachiyan, L. G., "A sublinear-time randomized approximation algorithm for matrix games," *Oper. Res. Lett.*, vol. 18, pp. 53–58, 1995.

[67] Gurobi Optimization, "Reference Manual," 4.0. http://www.gurobi.com/doc/40/refman/.

[68] Guruswami, V., Umans, C., and Vadhan, S., "Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes," *J. ACM*, vol. 56, no. 4, pp. 20–34, 2009.

[69] Herman, M. and Strohmer, T., "High-resolution radar via compressed sensing," *IEEE Trans. on Signal Processing*, vol. 57, no. 6, 2007.

[70] Huang, J. and Zhang, T., "The benefit of group sparsity," *Annals of Stat.*, vol. 38, no. 4, pp. 1978–2004, 2010.

[71] Huang, S. and Zhu, J., "Recovery of sparse signals using OMP and its variants: convergence analysis based on RIP," *Inverse Problems*, vol. 27, no. 3, p. 035003, 2011.

[72] Indyk, P., "Explicit constructions for compressed sensing of sparse signals," in *Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 30–33, 2008.

[73] Jafarpour, S., Cevher, V., and Shapire, R., "A game theoretic approach to expander-based compressive sensing," tech. rep., 2011. http:infoscience.epfl.ch/record/163465/files/isit2011.pdf.

[74] Jafarpour, S., Xu, W., Hassibi, B., and Calderbank, R., "Efficient and robust compressed sensing using optimized expander graphs," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4299–4308, 2009.

[75] James, G. M., Radchenko, P., and Lv, J., "Dasso: connections between the Dantzig selector and Lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 71, no. 1, pp. 127–142, 2009.

[76] Juditsky, A. B., Kilinç-Karzan, F., and Nemirovski, A. S., "Verifiable conditions of ℓ_1 recovery for sparse signals with sign restrictions," *Math. Progr.*, vol. 127, no. 1, pp. 89–122, 2010. http://www.optimization-online.org/DB_HTML/2009/03/2272.html.

[77] Juditsky, A. B. and Nemirovski, A. S., "Large deviations of vector-valued martingales in 2-smooth normed spaces," tech. rep., 2008. http://www2.isye.gatech.edu/~nemirovs/LargeDevSubmitted.pdf.

[78] Juditsky, A. B. and Nemirovski, A. S., "Nonparametric estimation by convex programming," *Annals of Stat.*, vol. 37, no. 5a, pp. 2278–2300, 2009.

[79] Juditsky, A. B. and Nemirovski, A. S., "Accuracy guarantees for ℓ_1-recovery," tech. rep., 2010. http://www.optimization-online.org/DB_HTML/2010/10/2778.html.

[80] Juditsky, A. B. and Nemirovski, A. S., "First order methods for large-scale convex optimization," in *Optimization for Machine Learning* (Sra, S., Nowozin, S., and Wright, S. J., eds.), Cambridge, Massachusetts: The MIT Press, 2010.

[81] Juditsky, A. B. and Nemirovski, A. S., "On verifiable sufficient conditions for sparse signal recovery via ℓ_1 minimization," *Math. Progr.*, vol. 127, no. 1, pp. 57–88, 2010. Special issue on machine learning.

[82] Juditsky, A. B., Nemirovski, A. S., and Tauvel, C., "Solving variational inequalities with stochastic mirror prox algorithm," tech. rep., 2008. http://www2.isye.gatech.edu/~nemirovs/SMP_240408.pdf.

[83] Lan, G., *Convex Optimization Under Inexact First-order Information*. PhD thesis, Georgia Institute of Technology, 2009. http://www.ise.ufl.edu/glan/papers/guanghui_lan_200908_phd.pdf.

[84] Lemarechal, C., Nemirovski, A. S., and Nesterov, Y., "New variants of bundle methods," *Math. Progr.*, vol. 69, no. 1, pp. 111–148, 1995.

[85] Liu, H. and Zhang, J., "Estimation consistency of the group Lasso and its applications," *Journal of Machine Learning Research - Proceedings Track*, vol. 5, pp. 376–383, 2009.

[86] Liu, H., Zhang, J., Jiang, X., and Liu, J., "The group Dantzig selector," *Journal of Machine Learning Research - Proceedings Track*, vol. 9, pp. 461–468, 2010.

[87] Lounici, K., "Sup-norm convergence rate and sign concentration property of Lasso and Dantzig estimators," *Electron. J. Statist.*, vol. 2, pp. 90–102, 2008.

[88] Lounici, K., Pontil, M., Tsybakov, A., and van de Geer, S., "Oracle inequalities and optimal inference under group sparsity," tech. rep., 2010. http://arxiv.org/pdf/1007.1771.

[89] Lustig, M., Donoho, D., and Pauly, J. M., "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 56, no. 6, pp. 1182–1195, 2007.

[90] Mallat, S. and Zhang, Z., "Matching pursuit with time-frequency dictionaries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[91] McCormick, G. P., "Computability of global solutions to factorable nonconvex programs: Part I Convex underestimating problems," *Math. Progr.*, vol. 10, no. 1, pp. 147–175, 1976.

[92] Meier, L., van de Geer, S., and Bühlmann, P., "The group Lasso for logistic regression," *J. Roy. Statist. Soc. Ser. B, Methodological*, vol. 70, no. 1, pp. 53–71, 2008.

[93] Meinshausen, N., Rocha, G., and Yu, B., "Discussion: A tale of three cousins: Lasso, l2boosting and Dantzig," *Annals of Stat.*, vol. 35, pp. 2373–2384, 2007.

[94] Meinshausen, N. and Yu, B., "Lasso-type recovery of sparse representations for high-dimensional data," *Annals of Stat.*, vol. 37, pp. 246–270, 2009.

[95] Milenkovic, O., Baraniuk, R., and Simunic-Rosing, T., "Compressed sensing meets bioinformatics: a new DNA microarray architecture," in *Information Theory and Applications Workshop*, 2007.

[96] Nardi, Y. and Rinaldo, A., "On the asymptotic properties of the group Lasso estimator for linear models," *Electron. J. Statist.*, vol. 2, pp. 605–633, 2008.

[97] Natarajan, B., "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, pp. 227–234, 1995.

[98] Needell, D. and Tropp, J., "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2008.

[99] Needell, D. and Vershynin, R., "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of Computational Mathematics*, vol. 9, no. 3, pp. 317–334, 2009.

[100] Nemirovski, A. S., "Efficient methods for large-scale convex problems," *Ekonomika i Matematicheskie Metody (in Russian)*, vol. 15, 1979.

[101] Nemirovski, A. S., "Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems," *SIAM J. Optim.*, vol. 15, pp. 229–251, 2004.

[102] Nemirovski, A. S., Juditsky, A. B., Lan, G., and Shapiro, A., "Stochastic approximation approach to stochastic programming," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1574–1609, 2009.

[103] Nemirovski, A. S. and Yudin, D., *Problem complexity and method efficiency in Optimization*. Somerset, NJ: John Wiley & Sons, 1983.

[104] Nesterov, Y., "A method for solving a convex programming problem with rate of convergence $O(1/k^2)$," *Soviet Math. Dokl.*, vol. 27, no. 2, pp. 372–376, 1983.

[105] Nesterov, Y., *Introductory Lectures on Convex Optimization: A Basic Course*. Norwell, MA: Springer, 2003.

[106] Nesterov, Y., "Smooth minimization of non-smooth functions," *Math. Progr.*, vol. 103, pp. 127–152, 2005.

[107] Nguyen, N. H., Do, T. T., and Tran, T. D., "A fast and efficient algorithm for low-rank approximation of a matrix," in *Proceedings of the* 41^{st} *annual ACM symposium on Theory of computing (STOC)*, pp. 215–224, 2009.

[108] Obozinski, G., Wainwright, M. J., and Jordan, M. I., "Support union recovery in high-dimensional multivariate regression," *Annals of Stat.*, vol. 39, no. 1, pp. 1–47, 2011.

[109] Parvaresh, F., Vikalo, H., Misra, S., and Hassibi, B., "Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays," *IEEE J. Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 275–285, 2008.

[110] Parvaresh, F. and Vardy, A., "Correcting errors beyond the Guruswami-Sudan radius in polynomial time," in *Proceedings of the* 46^{th} *Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 285–294, 2005.

[111] Pascal Koiran, A. Z., "On the certification of the restricted isometry property," tech. rep., 2011. http://arxiv.org/abs/1103.4984.

[112] Raginsky, M., Jafarpour, S., Harmany, Z. T., Marcia, R. F., Willett, R. M., and Calderbank, R., "Performance bounds for expander-based compressed sensing in poisson noise," tech. rep., 2010. http://arxiv.org/abs/1007.2377.

[113] Recht, B., Fazel, M., and Parrilo, P. A., "Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[114] Recht, B., Xu, W., and Hassibi, B., "Null space conditions and thresholds for rank minimization," *Math. Progr.*, vol. 127, pp. 175–211, 2011.

[115] Rubinstein, E., "Support vector machines via advanced optimization techniques," Master's thesis, Technion, 2005. http://www2.isye.gatech.edu/~nemirovs/Eitan.pdf.

[116] Rudin, L., Osher, S., and Fatemi, E., "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[117] Santosa, F. and Symes, W., "Linear inversion of band-limited reflection seismograms," *SIAM J. Sci. Statist. Comput.*, vol. 7, no. 4, pp. 1307–1330, 1986.

[118] Sherali, H. and Adams, W., *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Boston, MA: Kluwer Academic Publishers, 1999.

[119] Sipser, M. and Spielman, D. A., "Expander codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1710–1722, 1996.

[120] Srebro, N. and Shraibman, A., "Rank, trace-norm and max-norm," in *Proceedings of the* 18^{th} *Annual Conference on Learning Theory*, pp. 545–560, 2005.

[121] Stojnic, M., Parvaresh, F., and Hassibi, B., "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. on Signal Processing*, vol. 57, no. 8, pp. 3075–3085, 2009.

[122] Taylor, H., Banks, S., and McCoy, J., "Deconvolution with the ℓ_1-norm," *Geophys. J. Internat.*, vol. 44, no. 1, pp. 39–52, 1979.

[123] Tibshirani, R., "Regression shrinkage and selection via the Lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[124] Tropp, J., "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[125] Tropp, J. and Gilbert, A., "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[126] Unser, M., "Sampling–50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, 2000.

[127] Vasanawala, S. S., Alley, M. T., Hargreaves, B. A., Barth, R. A., Pauly, J. M., and Lustig, M., "Improved pediatric MR Imaging with compressed sensing," *Radiology*, vol. 256, no. 2, pp. 607–616, 2010.

[128] Wagner, G., Schmieder, P., Stern, A., and Hoch, J., "Application of nonlinear sampling schemes to cosy-type spectra," *J. Biomolecular NMR*, vol. 3, no. 5, pp. 569–576, 1993.

[129] Wakin, M., Laska, J., Duarte, M., Baron, D., Sarvotham, S., Takhar, D., Kelly, K., and Baraniuk, R., "An architecture for compressive imaging," in *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pp. 1273–1276, 2006.

[130] Yang, J. and Zhang, Y., "Alternating direction algorithms for ℓ_1-problems in compressive sensing," tech. rep., 2009. http://arxiv.org/abs/0912.1185.

[131] Yuan, M. and Lin, Y., "Model selection and estimation in regression with grouped variables," *J. Roy. Stat. Soc. Ser. B*, vol. 68, no. 1, pp. 49–67, 2006.

[132] Zhang, Y., "A simple proof for recoverability of ℓ_1-minimization: go over or under?," tech. rep., 2005. http://www.caam.rice.edu/~yzhang/reports/tr0509.pdf.

[133] Zhang, Y., "A simple proof for recoverability of ℓ_1-minimization (II): the nonnegative case," tech. rep., 2005. http://www.caam.rice.edu/~yzhang/reports/tr0510.pdf.

[134] Zhang, Y., "Theory of compressive sensing via ℓ_1 minimization: a non-rip analysis and extensions," tech. rep., 2008. http://www.caam.rice.edu/~yzhang/reports/tr0811.pdf.