

INFERENCE AND ESTIMATION IN HIGH-DIMENSIONAL  
DATA ANALYSIS

A DISSERTATION  
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Adel Javanmard

July 2014

© 2014 by Adel Javanmard. All Rights Reserved.

Re-distributed by Stanford University under license with the author.



This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.

<http://creativecommons.org/licenses/by-nc/3.0/us/>

This dissertation is online at: <http://purl.stanford.edu/sj271jx5066>

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**Andrea Montanari, Primary Adviser**

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**Jonathan Taylor**

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**David Tse**

Approved for the Stanford University Committee on Graduate Studies.

**Patricia J. Gumport, Vice Provost for Graduate Education**

*This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.*

# Abstract

Modern technologies generate vast amounts of fine-grained data at an unprecedented speed. Nowadays, high-dimensional data, where the number of variables is much larger than the sample size, occur in many applications, such as healthcare, social networks, and recommendation systems, among others. The ubiquitous interest in these applications has spurred remarkable progress in the area of high-dimensional data analysis in terms of point estimation and computation. However, one of the fundamental inference task, namely quantifying uncertainty or assessing statistical significance, is still in its infancy for such models. In the first part of this dissertation, we present efficient procedures and corresponding theory for constructing classical uncertainty measures like confidence intervals and p-values for single regression coefficients in high-dimensional settings.

In the second part, we study the compressed sensing reconstruction problem, a well-known example of estimation in high-dimensional settings. We propose a new approach to this problem that is drastically different from the classical wisdom in this area. Our construction of the sensing matrix is inspired by the idea of spatial coupling in coding theory and similar ideas in statistical physics. For reconstruction, we use an approximate message passing algorithm. This is an iterative algorithm that takes advantage of the statistical properties of the problem to improve convergence rate. Finally, we prove that our method can effectively solve the reconstruction problem at (information-theoretically) optimal undersampling rate and show its robustness to measurement noise.

*To my parents, Elaheh and Morteza,  
my sister, Ghazal,  
and my brother, Milad*

# Acknowledgments

I was undecided on which area to work in when I arrived at Stanford five years ago. I found a lot of them interesting, and I was not sure if I would be successful and happy working in those fields. I am greatly indebted to my advisor, Andrea Montanari, and also Balaji Prabhakar for making this transition seamless and enjoyable. Their advice was simple: Broaden your interests, find what you love and do what you believe is great work. That is the only way you can be satisfied.

I would like to thank Andrea for his constant support and unwavering encouragement. He has always had his door open to me despite his busy schedule. From our very first discussions, it was clear to me that working with Andrea will be a very enriching and rewarding experience. His enthusiasm for research, wisdom and deep knowledge has been very inspirational for me. He nurtured my curiosity and helped me develop as an academic.

I am grateful to David Donoho, Benjamin Van Roy, and Balaji Prabhakar for fruitful collaborations and fun engagements. Writing papers with them and observing how they approach problems was greatly instructive to me. I would also like to thank Jonathan Taylor, David Tse, Mohsen Bayati and Tsachy Weissman for serving on my PhD defense committee, and for their advice and attention.

I have been fortunate to work closely with some great peers at Stanford. I thank my dear friends, Mohammadreza Alizadeh, Yash Deshpande and Mahdi Soltanolkotabi for their endless help through my career. I have benefited greatly from their feedback on my work as well as our countless discussions on technical and non-technical matters. I wish to thank my officemates who made my time at our office educative and enjoyable: Sewoong Oh, Raghunandan Keshavan, Morteza Ibrahimi, Yashodhan Kanoria, Hamid Hakim Javadi, and Sonia Bhaskar.

I would like to extend my gratitude to my wonderful collaborators outside Stanford, Sham Kakade, Daniel Hsu and Li Zhang for all their support and guidance during my two

internships at Microsoft. Working with them has been a tremendous learning experience.

There are many people who influenced me and helped me shape my mathematical skills in the first place. I am immensely indebted to Mr. Razizadeh and Mr. Moayedpoor, my math teachers at middle school and high school in Esfahan, Iran. Their support and passion made me confident in my abilities.

Special thanks to my invaluable friends at Stanford, who made these years very enjoyable; a part of my life that I will surely miss later: Tahereh, Mohammadreza, Shayan, Farnaz, Azar, Behnam, Zeinab, Mohsen, Hanieh, Hadi, Shima, Mohammad, and Alireza. I have learned and earned much from your support and maturity. Thank you.

And finally, yet most importantly, I am very grateful to my family. My brother, Milad, my sister, Ghazal, and my father, Morteza. I am very blessed to have you in my life. Thanks for believing in me, loving me unconditionally, and supporting me through all my endeavors. I lost my mom, Elaheh, when I was eighteen, a trauma. Mom, it is beyond words how grateful I am for your selfless love, your dedication to your family and their comfort. You raised me, taught me priceless lessons, and gave me all I could wish for. Still loved, still missed and very dear, I dedicate this dissertation to your memory and to the rest of our family.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgments</b>	<b>vi</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Structured estimation . . . . .	5
1.1.1 Some examples . . . . .	6
1.1.2 More background and related work . . . . .	8
1.2 Assigning statistical significance in high-dimensional problems . . . . .	11
1.2.1 Why is it important? . . . . .	14
1.2.2 Why is it hard? . . . . .	15
1.2.3 Contributions & Organization (Part I) . . . . .	16
1.2.4 Previously published material . . . . .	17
1.3 Optimal compressed sensing via spatial coupling and approximate message passing . . . . .	18
1.3.1 A toy example . . . . .	23
1.3.2 Organization (Part II) . . . . .	24
1.3.3 Previously published material . . . . .	25
<b>I Assigning Statistical Significance in High-Dimensional Problems</b>	<b>26</b>
<b>2 Confidence Intervals for High-Dimensional Models</b>	<b>27</b>
2.1 Preliminaries and notations . . . . .	28
2.2 The bias of the Lasso . . . . .	29

2.3	Compensating the bias of the Lasso . . . . .	31
2.3.1	A debiased estimator for $\theta_0$ . . . . .	31
2.3.2	Discussion: bias reduction . . . . .	36
2.3.3	Comparison with earlier results . . . . .	37
2.4	Confidence intervals . . . . .	39
2.4.1	Preliminary lemmas . . . . .	39
2.4.2	Generalization to simultaneous confidence intervals . . . . .	41
2.4.3	Non-Gaussian noise . . . . .	43
2.5	Comparison with Local Asymptotic Normality . . . . .	44
2.6	General regularized maximum likelihood . . . . .	44
<b>3</b>	<b>Hypothesis Testing in High-Dimensional Regression</b>	<b>47</b>
3.1	Minimax formulation . . . . .	49
3.2	Familywise error rate . . . . .	50
3.3	Minimax optimality of a test . . . . .	51
3.3.1	Upper bound on the minimax power . . . . .	52
3.3.2	Near optimality of $\hat{T}_{i,\mathbf{X}}(Y)$ . . . . .	55
3.4	Other proposals . . . . .	56
3.4.1	Multisample splitting . . . . .	56
3.4.2	Bias-corrected projection estimator . . . . .	57
<b>4</b>	<b>Hypothesis Testing under Optimal Sample Size</b>	<b>60</b>
4.1	Hypothesis testing for standard Gaussian designs . . . . .	61
4.1.1	Hypothesis testing procedure . . . . .	61
4.1.2	Asymptotic analysis . . . . .	62
4.1.3	Gaussian limit . . . . .	65
4.2	Hypothesis testing for nonstandard Gaussian designs . . . . .	65
4.2.1	Hypothesis testing procedure . . . . .	66
4.2.2	Asymptotic analysis . . . . .	66
4.2.3	Gaussian limit via the replica heuristics . . . . .	69
4.2.4	Covariance estimation . . . . .	71
4.3	Role of the factor $d$ . . . . .	72

<b>5</b>	<b>Numerical Validation</b>	<b>77</b>
5.1	Remark on optimization (5.0.1) . . . . .	77
5.2	Experiment 1 . . . . .	79
5.3	Experiment 2 . . . . .	82
5.4	Experiment 3 . . . . .	84
5.5	Experiment 4 . . . . .	90
5.6	Experiment 5 . . . . .	92
<b>6</b>	<b>Proof of Theorems in Part I</b>	<b>98</b>
6.1	Proof of Theorems in Chapter 2 . . . . .	98
6.1.1	Proof of Theorem 2.3.3 . . . . .	98
6.1.2	Proof of Theorem 2.3.4.(a) . . . . .	99
6.1.3	Proof of Theorem 2.3.4.(b) . . . . .	101
6.1.4	Proof of Theorem 2.2.1 . . . . .	102
6.1.5	Proof of Theorem 2.3.5 . . . . .	105
6.1.6	Proof of Corollary 2.3.7 . . . . .	105
6.1.7	Proof of Lemma 2.4.2 . . . . .	106
6.1.8	Proof of Theorem 2.4.6 . . . . .	108
6.2	Proof of Theorems in Chapter 3 . . . . .	109
6.2.1	Proof of Theorem 3.1.1 . . . . .	109
6.2.2	Proof of Theorem 3.2.1 . . . . .	110
6.2.3	Proof of Lemma 3.3.6 . . . . .	111
6.2.4	Proof of Lemma 3.3.7 . . . . .	112
6.2.5	Proof of Theorem 3.3.3 . . . . .	113
6.3	Proofs of Theorems in Chapter 4 . . . . .	114
6.3.1	Proof of Theorem 4.1.3 . . . . .	114
6.3.2	Proof of Theorem 4.2.3 . . . . .	116
6.3.3	Proof of Theorem 4.2.4 . . . . .	117
<b>II</b>	<b>Optimal Compressed Sensing via Spatial Coupling and Approximate Message Passing</b>	<b>118</b>
<b>7</b>	<b>Reconstruction at Optimal Rate</b>	<b>119</b>

7.1	Formal statement of the results . . . . .	119
7.1.1	Rényi information dimension . . . . .	123
7.1.2	MMSE dimension . . . . .	124
7.1.3	Main results . . . . .	125
7.2	Discussion . . . . .	127
7.3	Related work . . . . .	131
<b>8</b>	<b>Matrix and algorithm construction</b>	<b>133</b>
8.1	General matrix ensemble . . . . .	133
8.2	State evolution . . . . .	135
8.3	General algorithm definition . . . . .	137
8.4	Choices of parameters, and spatial coupling . . . . .	139
<b>9</b>	<b>Magic of Spatial Coupling</b>	<b>144</b>
9.1	How does spatial coupling work? . . . . .	144
9.2	Advantages of spatial coupling . . . . .	146
<b>10</b>	<b>Key Lemmas and Proof of the Main Theorems</b>	<b>148</b>
10.1	Numerical experiments . . . . .	151
10.1.1	Evolution of the AMP algorithm . . . . .	151
10.1.2	Phase diagram . . . . .	152
10.2	State evolution: an heuristic derivation . . . . .	154
<b>11</b>	<b>Analysis of state evolution: Proof of Lemma 10.0.2</b>	<b>157</b>
11.1	Outline . . . . .	157
11.2	Properties of the state evolution sequence . . . . .	158
11.3	Modified state evolution . . . . .	160
11.4	Continuum state evolution . . . . .	163
11.4.1	Free energy . . . . .	165
11.4.2	Analysis of the continuum state evolution . . . . .	167
11.4.3	Analysis of the continuum state evolution: robust reconstruction . . . . .	170
11.5	Proof of Lemma 10.0.2 . . . . .	172
<b>12</b>	<b>Gabor Transform &amp; Spatial Coupling</b>	<b>176</b>
12.1	Definitions . . . . .	177

12.2	Information theory model . . . . .	178
12.3	Contributions . . . . .	178
12.4	Sampling scheme . . . . .	179
12.4.1	Constructing the sensing matrix . . . . .	179
12.4.2	Algorithm . . . . .	181
12.5	Numerical simulations . . . . .	182
12.5.1	Evolution of the algorithm . . . . .	182
12.5.2	Phase diagram . . . . .	184
<b>A</b>	<b>Supplement to Chapter 2</b>	<b>188</b>
A.1	Proof of Lemma 2.4.1 . . . . .	188
A.2	Proof of Lemma 2.4.3 . . . . .	189
A.3	Proof of Lemma 6.1.3 . . . . .	190
<b>B</b>	<b>Supplement to Chapter 4</b>	<b>192</b>
B.1	Effective noise variance $\tau_0^2$ . . . . .	192
B.2	Tunned regularization parameter $\lambda$ . . . . .	193
B.3	Replica method calculation . . . . .	194
<b>C</b>	<b>Supplement to Chapter 7</b>	<b>205</b>
C.1	Dependence of the algorithm on the prior $p_X$ . . . . .	205
C.2	Lipschitz continuity of AMP . . . . .	208
<b>D</b>	<b>Supplement to Chapter 11</b>	<b>212</b>
D.1	Proof of Lemma 11.4.2 . . . . .	212
D.2	Proof of Proposition 11.4.9 . . . . .	215
D.3	Proof of Claim 11.4.11 . . . . .	215
D.4	Proof of Proposition 11.4.12 . . . . .	216
D.5	Proof of Proposition 11.4.13 . . . . .	218
D.6	Proof of Proposition 11.4.14 . . . . .	222
D.7	Proof of Claim 11.4.16 . . . . .	223
D.8	Proof of Proposition 11.4.17 . . . . .	224
	<b>Bibliography</b>	<b>226</b>

# List of Tables

1.3.1 Comparison between classical approach to compressed sensing and our approach. . . . .	23
5.2.1 Simulation results for the synthetic data described in Experiment 1. The results corresponds to 95% confidence intervals. . . . .	80
5.2.2 Simulation results for the synthetic data described in Experiment 1. The false positive rates (FP) and the true positive rates (TP) are computed at significance level $\alpha = 0.05$ . . . . .	82
5.4.1 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the asymptotic bound for SDL-TEST (cf. Theorem 4.1.3) in Experiment 3. The significance level is $\alpha = 0.05$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of $p = 1000$ , $n = 600$ , $s_0 = 50$ , $\gamma = 0.1$ . . . . .	88
5.4.2 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the asymptotic bound for SDL-TEST (cf. Theorem 4.1.3) in Experiment 3. The significance level is $\alpha = 0.025$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of $p = 1000$ , $n = 600$ , $s_0 = 50$ , $\gamma = 0.1$ . . . . .	89
5.5.1 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the lower bound for the statistical power of SDL-TEST (cf. Theorem 4.2.4) in Experiment 4. The significance level is $\alpha = 0.05$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of $p = 1000$ , $n = 600$ , $s_0 = 50$ , $\gamma = 0.1$ . . . . .	93

5.5.2 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the lower bound for the statistical power of SDL-TEST (cf. Theorem 4.2.4) in Experiment 4. The significance level is $\alpha = 0.025$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as $(1000, 600, 50, 0.1)$ denotes the values of $p = 1000$ , $n = 600$ , $s_0 = 50$ , $\gamma = 0.1$ . . . . .	94
5.6.1 Simulation results for the communities data set. . . . .	96
5.6.2 The relevant features (using the whole dataset) and the relevant features predicted by SDL-TEST and the Ridge-type projection estimator [16] for a random subsample of size $n = 84$ from the communities. The false positive predictions are in red. . . . .	97
7.2.1 Comparison between the minimax setup in [44] and the Bayesian setup considered in this dissertation. (cf. [44, Eq. (2.4)] for definition of $M(\varepsilon)$ ). . . .	131

# List of Figures

1.2.1	Frequency of the maximum absolute sample correlation between the first variable and the $i$ -th variables. This is an illustration of spurious correlation in high-dimensions. . . . .	12
1.2.2	Lefthand side: low dimensional projection of the high probability region given by (1.1.9) ( $q = 2$ ). The ball radius is $O(\sqrt{(s_0 \log p)/n})$ . Righthand side: confidence regions for low dimensional targets, $\theta_I$ , where $I$ is a subset of indices with constant size. The radius here is $O(1/\sqrt{n})$ . . . . .	13
1.3.3	Donoho-Tanner phase diagram for algorithm ( $P_1$ ) . . . . .	24
4.3.1	Empirical kurtosis of vector $v$ with and without normalization factor $d$ . In left panel $n = 3 s_0$ (with $\epsilon = 0.2$ , $\delta = 0.6$ ) and in the right panel $n = 30 s_0$ (with $\epsilon = 0.02$ , $\delta = 0.6$ ). . . . .	74
4.3.2	Histogram of $v$ for $n = 3 s_0$ ( $\epsilon = 0.2$ , $\delta = 0.6$ ) and $p = 3000$ . In left panel, factor $d$ is computed by Eq. (4.2.1) and in the right panel, $d = 1$ . . . . .	75
4.3.3	Histogram of $v$ for $n = 30 s_0$ ( $\epsilon = 0.02$ , $\delta = 0.6$ ) and $p = 3000$ . In left panel, factor $d$ is computed by Eq. (4.2.1) and in the right panel, $d = 1$ . . . . .	76
5.2.1	95% confidence intervals for one realization of configuration $(n, p, s_0, b) = (1000, 600, 10, 1)$ . For clarity, we plot the confidence intervals for only 100 of the 1000 parameters. The true parameters $\theta_{0,i}$ are in red and the coordinates of the debiased estimator $\hat{\theta}^u$ are in black. . . . .	81
5.2.2	Q-Q plot of $Z$ for one realization of configuration $(n, p, s_0, b) = (1000, 600, 10, 1)$ . . . . .	83
5.2.3	Empirical CDF of the computed $p$ -values (restricted to entries outside the support) for one realization of configuration $(n, p, s_0, b) = (1000, 600, 10, 1)$ . Clearly the plot confirms that the $p$ -values are distributed according to uniform distribution. . . . .	83

5.3.4 Empirical CDF of the computed $p$ -values for riboflavin example. Clearly the plot confirms that the $p$ -values are distributed according to uniform distribution. . . . .	84
5.4.5 Comparison between SDL-TEST ( Algorithm 4), Ridge-based regression [16] and the asymptotic bound for SDL-TEST (established in Theorem 4.1.3). Here, $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ . . . . .	85
5.4.6 Comparison between SDL-TEST , Ridge-based regression [16], and LDPE [149]. The curve corresponds to the asymptotic bound for SDL-TEST as established in Theorem 4.1.3. For the same values of type I error achieved by methods, SDL-TEST results in a higher statistical power. Here, $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ . . . . .	86
5.5.7 Numerical results for Experiment 4 and $p = 2000, n = 600, s_0 = 50, \gamma = 0.1$ .	91
5.5.8 Comparison between SDL-TEST , Ridge-based regression [16], and LDPE [149] in the setting of Experiment 4. For the same values of type I error achieved by methods, SDL-TEST results in a higher statistical power. Here, $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ . . . . .	92
5.6.9 Parameter vector $\theta_0$ for the communities data set. . . . .	95
5.6.10 Normalized histogram of $v_{S_0}$ (in red) and $v_{S_0^c}$ (in white) for the communities data set. . . . .	96
8.1.1 Construction of the spatially coupled measurement matrix $A$ as described in Section 8.1. The matrix is divided into blocks with size $M$ by $N$ . (Number of blocks in each row and each column are respectively $L_c$ and $L_r$ , hence $m = ML_r, n = NL_c$ ). The matrix elements $A_{ij}$ are chosen as $\mathbf{N}(0, \frac{1}{M}W_{\mathbf{g}(i),\mathbf{g}(j)})$ . In this figure, $W_{i,j}$ depends only on $ i - j $ and thus blocks on each diagonal have the same variance. . . . .	135
8.4.2 Matrix $W$ . The shaded region indicates the non zero entries in the lower part of the matrix. As shown (the lower part of ) the matrix $W$ is band diagonal. . . . .	140
9.1.1 Graph structure of a spatially coupled matrix. Variable nodes are shown as circle and check nodes are represented by square. . . . .	145
10.1.1 Profile $\phi_a(t)$ versus $a$ for several iteration numbers. . . . .	152
10.1.2 Comparison of $\text{MSE}_{\text{AMP}}$ and $\text{MSE}_{\text{SE}}$ across iteration. . . . .	153

10.1.3 Phase diagram for the spatially coupled sensing matrices and Bayes optimal AMP. . . . .	153
11.4.1 An illustration of function $\phi(x)$ and its perturbation $\phi_a(x)$ . . . . .	169
12.5.1 Profile $\phi_a(t)$ versus $a$ for several iteration numbers. . . . .	183
12.5.2 Comparison of $\text{MSE}_{\text{AMP}}$ and $\text{MSE}_{\text{SE}}$ across iteration. . . . .	184
12.5.3 Phase transition diagram for Scheme I, and $\varepsilon = 0.2$ . . . . .	185
12.5.4 Phase transition lines for Schemes I, II, III. . . . .	186

# Notational Conventions

Notation	Description
$\mathbb{R}$	real number
$\mathbb{R}^n$	vector space of $n$ -dimensional real valued vectors
$e_i$	vector with one at the $i$ -th position and zero everywhere else
$[p]$	$\{1, \dots, p\}$
$ \cdot $	if applied to a number, absolute value
$ \cdot $	if applied to a set, cardinality of the set
$\mathbb{I}$	indicator function
$(a)_+$	$a$ if $a > 0$ and zero otherwise
$\mathbb{P}(\cdot)$	probability of an event
$\mathbb{E}(\cdot)$	expected value of a random variable
$\mathbf{I}$	identity matrix in any dimension
$\phi(x)$	$e^{-x^2/2}/\sqrt{2\pi}$ , the Gaussian density
$\Phi(\cdot)$	$\int_{-\infty}^x e^{-u^2/2}/\sqrt{2\pi} du$ , the Gaussian distribution
$\ X\ _{\psi_1}$	sub-exponential norm of random variable (or vector) $X$
$\ X\ _{\psi_2}$	sub-gaussian norm of random variable (or vector) $X$
$v_i$	$i$ -th element of vector $v$
$\ v\ _p$	for a vector $v$ , $\ell_p$ norm defined as $(\sum_i  v_i ^p)^{1/p}$
$\ v\ _0$	$\ell_0$ norm of a vector. Number of nonzero elements in $v$ .
$\langle u, v \rangle$	$\sum_i u_i v_i$
$\text{supp}(v)$	for a vector $v$ , positions of nonzero entries in $v$
$v_I$	for vector $v$ , restriction of $v$ to indices in $I$

Notation	Description
$A_{ij}$	element $(i, j)$ of matrix $A$
$ A _\infty$	for a matrix $A$ , the maximum magnitude of the entries
$\ A\ _p$	for a matrix $A$ , $\ell_p$ operator norm
$\sigma_{\max}(A)$	maximum singular value of matrix $A$
$\sigma_{\min}(A)$	minimum singular value of matrix $A$
$A_J$	submatrix of $A$ with columns restricted to set $J$
$A_{I,J}$	submatrix of $A$ formed by rows in $I$ and columns in $J$
$A_{I,J}^{-1}$	shorthand for $(A^{-1})_{I,J}$
$A^\top$	transpose of matrix $A$
$f(n) = o(g(n))$	$f$ is dominated by $g$ asymptotically ( $\forall k > 0, \exists n_0$ , such that $\forall n > n_0,  f(n)  \leq k g(n) $ )
$f(n) = O(g(n))$	$f$ is bounded above by $g$ asymptotically ( $\exists k > 0, \exists n_0$ , such that $\forall n > n_0,  f(n)  \leq k g(n) $ )
$f(n) = \omega(g(n))$	$f$ dominates $g$ asymptotically ( $\forall k > 0, \exists n_0$ , such that $\forall n > n_0,  f(n)  \geq k g(n) $ )
$\bar{d}(p_X), \underline{d}(p_X)$	upper and lower Rényi information dimension of $p_X$
$\bar{D}(p_X), \underline{D}(p_X)$	upper and lower MMSE dimension of $p_X$

# Chapter 1

## Introduction

We are in the era of massive automated data collection, where we systematically obtain many measurements without knowing which ones are really relevant to the phenomena of interest. Microarrays and fMRI machines produce thousands of parallel datasets. Online social networks are constantly accumulating location, interaction and other information concerning hundreds of millions of users. Similar trend is now seen in healthcare systems, online advertising, and electronic commerce, among others.

This is a big break from traditional statistical theory in the following sense. In statistical data analysis, we have *samples* of a particular phenomena, and for each sample, we observe a vector of values measured on several *variables*. In traditional statistical methodology it was assumed that one has access to many samples and is dealing with a few relevant variables. For example, in studying a specific disease, doctor uses her domain expertise to measure just the right variables. However, the ubiquitous technological trend today is driving us towards the regime of more samples but even more so, to an extensively larger numbers of variables. Modern data sets are not only massive in sample size but also are remarkably feature-rich. As a concrete example, a typical electronic health record (EHR) database contains transcript records, lab results, medications, immunization status, medical images and a lot of other detailed information of patients, leading to a huge number of numerical variables (features). Moreover, one can construct new features by applying different functions to the current variables, or by considering higher order features, like  $k$ -tuples. Therefore, in principle one can construct an enormous set of features.

Variables (features) are commonly thought of as *dimensions* on which we are collecting information. In other words, the number of variables is regarded as the ambient dimension of

data. The focus of *high-dimensional* statistics is on the regime where the ambient dimension of data is of the same order or substantially larger than the sample size. The most useful statistical models in this context are over-parameterized: the number of parameters ( $p$ ) to estimate is far larger than the number of samples ( $n$ ).<sup>1</sup>

**Curses of dimensionality.** The expression “curse of dimensionality” is due to Bellman [10], where he used it to explain the difficulty of optimization and function evaluation on product space. Indeed, high dimensionality introduces both computational and statistical challenges.

- *Computational challenges:* Suppose that we have a function of  $d$  variables and we only know that it is Lipschitz. If we want to approximate this function over the unit hypercube  $[0, 1]^d$ , within uniform approximation error  $\epsilon$ , then we require  $\mathcal{O}(\epsilon^{-d})$  evaluations. A similar exponential explosion in computational complexity appears when we want to optimize such a function over a bounded region.
- *Statistical challenges:* Suppose that we are given  $n$  i.i.d. pairs  $(Y_1, X_1), (Y_2, X_2), \dots, (Y_n, X_n)$ , with vectors  $X_i \in \mathbb{R}^d$  and response variables  $Y_i$  given by

$$Y_i = f(X_i) + \text{noise}.$$

Further assume that we merely know  $f$  is a Lipschitz function and noise variables are i.i.d Gaussian with mean 0 and variance 1. Under these assumptions, we aim to estimate  $f$  from the observed samples.

We are interested in sample complexity for this task, namely how the accuracy of estimation depends on  $n$ . Let  $\mathcal{F}$  be the family of Lipschitz functions on  $[0, 1]^d$ . A standard argument in minimax decision theory [67] states that

$$\inf_{\hat{f}} \sup_{f \in \mathcal{F}} \mathbb{E}(\hat{f}(x) - f(x))^2 \geq Cn^{-2/(2+d)},$$

for some constant  $C$  that depends on the Lipschitz constant. Further, this lower bound is not asymptotic. Hence, in order to estimate  $f$  within an accuracy of  $\epsilon$ , we need  $\mathcal{O}(\epsilon^{-(2+d)/2})$  samples. The very slow rate of convergence in high dimensions is another aspect of the curse of dimensionality.

---

<sup>1</sup>One can think of associating one parameter to each measured variable.

Finally, over-parametrized models are prone to over-fitting in cases that the number of parameters are too large with respect to the size of training samples. Over-fitting implies poor generalization to correctly predict on new samples. Moreover, high-dimensionality brings noise accumulation and spurious correlations between response and unrelated features, which may lead to wrong statistical inference and false predictions.

**Blessings of dimensionality.** One of the main blessings of high-dimensionality is the phenomenon of “concentration measure”. Roughly speaking, having many “identical” dimensions allows one to “average” over them.

To be more specific, let  $S^{d-1}$  denote the surface of the unit sphere in  $\mathbb{R}^d$ , and let  $\mathbb{P}$  be the uniform measure over  $S^{d-1}$ . Then, for a function  $f : S^{d-1} \rightarrow \mathbb{R}$  that is  $L$ -Lipschitz, we have

$$\mathbb{P}\left(|f(x) - \mathbb{E}f(x)| > \epsilon\right) \leq 2e^{-d\epsilon^2/(2L^2)}.$$

The slogan is that *Lipschitz functions are nearly constant and the tails fall faster in higher dimensions*. Concentration of measure in high dimensions is the underlying tool in establishing many results in statistics and probability theory.

We refer to [35] for more discussions on curses and blessings of dimensionality.

## 1.1 Structured estimation

In the high dimensional models the number of parameters  $p$  is comparable to or larger than the sample size  $n$ , and therefore it is in general impossible to obtain consistent estimator procedures. Indeed, when  $p > n$ , the problem of parameter estimation is ill-posed. On the other hand, many such models enjoy various types of low-dimensional structures. Examples of such structures include sparsity, rank conditions, smoothness, symmetry, etc. A common tool in such settings is *regularization* that encourages the assumed structure. Regularization has played fundamental role in statistics and related mathematical areas. It was first introduced by Tikhonov [130] in connection with solving ill-posed integral equations. Since then, it has become a standard tool in statistics.

A widely applicable approach to estimation, in the context of high-dimensional models, is to solve a regularized optimization problem, which combines a loss function measuring fidelity of the model to the samples, with some regularization that promotes the underlying

structure. More precisely, let  $\mathcal{Z}^n = \{Z_1, Z_2, \dots, Z_n\}$  be a collection of samples drawn independently from some distribution, and let  $\mathcal{L}_n(\theta; \mathcal{Z}_n)$  be an empirical risk function defined as

$$\mathcal{L}(\theta; \mathcal{Z}_n) \equiv \frac{1}{n} \sum_{i=1}^n \mathcal{L}(\theta; Z_i).$$

Here  $\mathcal{L}(\theta; Z_i)$  measures the fit of parameter  $\theta$  to sample  $Z_i$ . For instance, in the regression setting, we have  $Z_i = (Y_i, X_i)$  with  $Y_i \in \mathbb{R}$  the response variable,  $X_i \in \mathbb{R}^p$  the covariate vector, and the least-squares loss  $\mathcal{L}(\theta; Z_i) = \frac{1}{2}(Y_i - \langle X_i, \theta \rangle)^2$ . A regularized M-estimator  $\hat{\theta}$  is constructed by minimizing a weighted combination of the empirical risk function with a convex regularizer  $\mathcal{R} : \mathbb{R}^p \rightarrow \mathbb{R}_+$ , that enforces a certain structure in the solution. Namely,

$$\hat{\theta} \in \operatorname{argmin}_{\theta \in \mathbb{R}^p} \left( \mathcal{L}_n(\theta; \mathcal{Z}_n) + \lambda_n \mathcal{R}(\theta) \right), \quad (1.1.1)$$

where  $\lambda_n > 0$  is a regularization parameter to be chosen. In case the right hand side has more than one minimizer, one of them can be selected arbitrarily for our purposes.

### 1.1.1 Some examples

We consider some classical examples of M-regularized estimators.

**Ridge regression estimator:** The simplest example of M-regularized estimators is the ridge regression estimate for linear models [65]. Given observations  $\mathcal{Z}_n = \{Z_1, Z_2, \dots, Z_n\}$ , with  $Z_i = (Y_i, X_i) \in \mathbb{R} \times \mathbb{R}^p$ , the Ridge regression estimator is defined by the following choice of loss function and regularizer:

$$\mathcal{L}_n(\theta; \mathcal{Z}_n) = \frac{1}{2n} \sum_{i=1}^n (Y_i - \langle \theta, X_i \rangle)^2, \quad \mathcal{R}(\theta) = \frac{1}{2} \|\theta\|_2^2. \quad (1.1.2)$$

**Lasso estimator:** In many applications only a relatively small subset of covariates are relevant to the response variable. Correspondingly, the parameter vector of interest is sparse. In these cases, a very successful estimator is Lasso which uses  $\ell_1$  norm regularizer to promote sparsity in the solution [30, 129]. For linear models, Lasso estimator is given by

$$\mathcal{L}_n(\theta; \mathcal{Z}_n) = \frac{1}{2n} \sum_{i=1}^n (Y_i - \langle \theta, X_i \rangle)^2, \quad \mathcal{R}(\theta) = \|\theta\|_1 = \sum_{i=1}^n |\theta_i|. \quad (1.1.3)$$

**Group-structured penalties:** In many applications, we are interested in finding important explanatory *factors* in predicting the response variable. Each explanatory factor may be represented by a group of dummy variables. As an example, consider predicting diabetes status of an individual based on her medical records. In this case, a factor might be a specific lab test and different levels of the test outcome can be represented through multiple variables. Since we are interested in finding the important factors (group of variables), we would like to impose sparsity at the group level, rather than individual variables. Various group-based regularizer have been studied to model such structured sparsity. Consider a collection of groups  $\mathcal{G} = \{G_1, G_2, \dots, G_k\}$ , such that  $G_i \subseteq [p]$ , and  $\cup_{i=1}^k G_i = [p]$ . Note that the groups may overlap. Moreover, for a vector  $\theta$ , let  $\theta_G$  denote the restriction of  $\theta$  to entries in  $G$ . Given a vector norm  $\|\cdot\|_{\#}$ , the regularizer is defined as follows:

$$\mathcal{R}(\theta) \equiv \sum_{G_i \in \mathcal{G}} \|\theta_{G_i}\|_{\#}. \quad (1.1.4)$$

Perhaps, the most common choice is  $\|\cdot\|_{\#} = \|\cdot\|_2$ , which is called group Lasso norm [80, 123, 147, 105]. The other choice studied by several researchers is  $\|\cdot\|_{\#} = \|\cdot\|_{\infty}$  [102, 132].

It is worth mentioning that when the groups are overlapping, the standard group Lasso has a property that may be undesirable. Let  $\hat{\theta}$  be regularized estimator (1.1.1), when the standard group norm (1.1.4) is used as regularizer. Further, let  $\hat{S} = \text{supp}(\hat{\theta})$ . Then, it can be shown that the complement  $\hat{S}^c$  of the support, i.e.,  $\hat{S}^c = \{i \in [p] : \hat{\theta}_i = 0\}$ , is always equal to the union of some of the groups. However, it is often natural to look for estimators  $\hat{\theta}$ , whose support (rather than its complement) is given by the union of some of the groups because these groups correspond to the important factors we are seeking. Jacob et al. [69] introduced a variant of the group lasso, known as *latent group lasso* to overcome this problem. It relies on the observation that for overlapping groups, a vector  $\theta$  has many possible group representations, where each representation is given by a set of vectors  $w_{G_i} \in \mathbb{R}^p$ , such that  $\sum_{G_i \in \mathcal{G}} w_{G_i} = \theta$  and  $\text{supp}(w_{G_i}) \subseteq G_i$ . Minimizing over all such representations gives the latent group lasso norm:

$$\mathcal{R}(\theta) \equiv \inf \left\{ \sum_{G_i \in \mathcal{G}} \|w_{G_i}\|_{\#} : \theta = \sum_{G_i \in \mathcal{G}} w_{G_i}, \text{supp}(w_{G_i}) \subseteq G_i \right\}.$$

Notice that when the groups are non-overlapping, the latent group norm reduced to the standard group norm by a simple use of triangle inequality. However, when the groups overlap, the solution  $\hat{\theta}$  (1.1.1) with latent group norm regularizer, is ensured to have its support equal to a union of a subset of groups [69].

**Low-rank matrix estimation:** There is a tremendous amount of work on estimating matrices with rank constraints. The rank constraints apply to many applications, including principle component analysis, clustering, matrix completion. A natural approach would be to enforce such constraints explicitly in the estimation procedure. However, the rank function is non-convex and in many cases, this approach leads to computationally infeasible schemes or resists a rigorous analysis because of the presence of local optima.

A natural surrogate for the rank function is the nuclear norm. Given a matrix  $\Theta \in \mathbb{R}^{n_1 \times n_2}$ , let  $\sigma_1, \sigma_2, \dots, \sigma_n$  be the singular values of  $\Theta$ , with  $n = \min(n_1, n_2)$ . Then,  $\text{rank}(\Theta)$  is merely the number of strictly positive singular values. The nuclear norm of  $\Theta$  is defined as  $\|\Theta\|_* = \sum_{i=1}^n \sigma_i$ . In other words, rank is the  $\ell_0$  norm of the vector of singular values, while the nuclear norm is its  $\ell_1$  norm. Analogous to the  $\ell_1$  norm as a relaxation of  $\ell_0$  norm, nuclear norm serves as a natural convex relaxation of the rank function. When nuclear norm is used as the regularizer in (1.1.1), it promotes low-rank solutions. The statistical and computational behavior of nuclear-norm regularized estimators has been well studied in various contexts [27, 26, 111, 59, 101].

### 1.1.2 More background and related work

High-dimensional regression has been the object of much theoretical investigation over the last few years. Here, we restrict ourselves to high-dimensional linear regression and the Lasso estimator (1.1.3). Before reviewing some of the obtained results, we need to establish some notations. Suppose that we are given  $n$  i.i.d. pairs  $(Y_1, X_1), (Y_2, X_2), \dots, (Y_n, X_n)$ , with  $X_i \in \mathbb{R}^p$  and response variables  $Y_i$  given by

$$Y_i = \langle \theta_0, X_i \rangle + W_i, \quad W_i \sim \mathbf{N}(0, \sigma^2). \quad (1.1.5)$$

Here  $\theta_0 \in \mathbb{R}^p$  is a vector of parameters to be learned. In matrix form, letting  $Y = (Y_1, \dots, Y_n)^\top$  and denoting  $\mathbf{X}$  the design matrix with rows  $X_1^\top, \dots, X_n^\top$ , we have

$$Y = \mathbf{X}\theta_0 + W, \quad W \sim \mathbf{N}(0, \sigma^2 \mathbf{I}_{n \times n}). \quad (1.1.6)$$

Recall Lasso estimator  $\hat{\theta} = \hat{\theta}^n(Y, \mathbf{X}; \lambda)$  defined as

$$\hat{\theta}^n(Y, \mathbf{X}; \lambda) \equiv \operatorname{argmin}_{\theta \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|_2^2 + \lambda \|\theta\|_1 \right\}. \quad (1.1.7)$$

We will omit the arguments  $Y, \mathbf{X}, \lambda$  and superscript  $n$ , when they are clear from the context. Further, let  $S = \operatorname{supp}(\theta_0)$ .

The focus has been so far on establishing order optimal guarantees on: (1) The prediction error  $\|\mathbf{X}(\hat{\theta} - \theta_0)\|_2$ , see e.g. [58]; (2) The estimation error, typically quantified through  $\|\hat{\theta} - \theta_0\|_q$ , with  $q \in [1, 2]$ , see e.g. [25, 14, 110]; (3) The model selection (or support recovery) properties typically by bounding  $\mathbb{P}\{\operatorname{supp}(\hat{\theta}) \neq S\}$ , see e.g. [95, 150, 140].

For prediction, there is no need to identify  $\theta_0$  since we are interested only in  $X_{\text{new}}^\top \theta_0$ . From this perspective, prediction is always an easier task than estimation of the parameter  $\theta_0$  or model selection. Roughly speaking, it is proved that with a proper choice for  $\lambda$  (of order  $\sigma \sqrt{(\log p)/n}$ ), one has the following ‘oracle inequality’ with high probability

$$\frac{1}{n} \|\mathbf{X}(\hat{\theta} - \theta_0)\|_2^2 \leq C_1 \sigma^2 \frac{s_0 \log p}{n}, \quad (1.1.8)$$

where  $C_1 > 0$  is a constant that depends on the Gram matrix  $\hat{\Sigma} = (\mathbf{X}^\top \mathbf{X}/n)$  [133].

For estimation guarantee, it is necessary to make specific assumptions on the design matrix  $\mathbf{X}$ , such as the restricted eigenvalue property of [14] or the compatibility condition of [133]. In particular, Bickel, Ritov and Tsybakov [14] show that, under such conditions, and for a suitable choice of  $\lambda$  (of order  $\sigma \sqrt{(\log p)/n}$ ), we have with high probability,

$$\|\hat{\theta} - \theta_0\|_q^q \leq C_2 s_0 \lambda^q, \quad (1.1.9)$$

for  $1 \leq q \leq 2$  and some constant  $C_2 > 0$  that depends on the Gram matrix  $\hat{\Sigma} = (\mathbf{X}^\top \mathbf{X}/n)$ .

For model selection guarantee, it was understood early on that even in the large-sample, low-dimensional limit  $n \rightarrow \infty$  at  $p$  constant,  $\operatorname{supp}(\hat{\theta}^n) \neq S$  unless the columns of  $\mathbf{X}$  with index in  $S$  are roughly orthogonal to the ones with index outside  $S$  [81]. This assumption

is formalized by the so-called *irrepresentability condition*, that can be stated in terms of the empirical covariance matrix  $\widehat{\Sigma} = (\mathbf{X}^\top \mathbf{X}/n)$ . Letting  $\widehat{\Sigma}_{A,B}$  be the submatrix  $(\widehat{\Sigma}_{i,j})_{i \in A, j \in B}$ , irrepresentability requires

$$\|\widehat{\Sigma}_{S^c, S} \widehat{\Sigma}_{S, S}^{-1} \text{sign}(\theta_{0, S})\|_\infty \leq 1 - \eta, \quad (1.1.10)$$

for some  $\eta > 0$  (here  $\text{sign}(u)_i = +1, 0, -1$  if, respectively,  $u_i > 0, = 0, < 0$ ). In an early breakthrough, Zhao and Yu [150] proved that, if this condition holds with  $\eta$  uniformly bounded away from 0, it guarantees correct model selection also in the high-dimensional regime  $p \gg n$ . Meinshausen and Bühlmann [95] independently established the same result for random Gaussian designs, with applications to learning Gaussian graphical models. These papers applied to very sparse models, requiring in particular  $s_0 = O(n^c)$ , for some  $c < 1$  with  $s_0 = \|\theta_0\|_0$  and parameter vectors with large coefficients. Namely, scaling the columns of  $X$  such that  $\widehat{\Sigma}_{i,i} \leq 1$ , for  $i \in [p]$ , they require  $\theta_{\min} \equiv \min_{i \in S} |\theta_{0,i}| \geq c' \sqrt{s_0/n}$ .

Wainwright [140] strengthened considerably these results by allowing for general scalings of  $s_0, p, n$  and proving that much smaller non-zero coefficients can be detected. Namely, he proved that for a broad class of empirical covariances it is only necessary that  $\theta_{\min} \geq c\sigma\sqrt{(\log p)/n}$ . This scaling of the minimum non-zero entry is optimal up to constants. Also, for specific classes of random Gaussian designs (including  $\mathbf{X}$  with i.i.d. standard Gaussian entries), the analysis of [140] provides tight bounds on the minimum sample size for correct model selection. Namely, there exists  $c_\ell, c_u > 0$  such that the Lasso fails with high probability if  $n < c_\ell s_0 \log p$  and succeeds with high probability if  $n \geq c_u s_0 \log p$ . Recently, [73] has introduced *generalized irrepresentability condition*, an assumption that is substantially weaker than irrepresentability. The authors prove that, under generalized irrepresentability condition, the Gauss-Lasso estimator correctly recovers the active set of variables.

A less ambitious goal than model selection is the task of variable screening, where one requires to find a subset  $\widehat{S} \subset [p]$  of the variables that contains  $S$ , with high probability, and  $|\widehat{S}|$  is much smaller than  $p$ . Therefore, variable screening allows for substantial dimensionality reduction, which is very useful when dealing with large data in practice. For variable screening, design matrix  $\mathbf{X}$  is required to have compatibility condition<sup>2</sup>. Further, the minimum nonzero parameter  $\theta_{\min}$  should be sufficiently large; similar assumption to the

<sup>2</sup>The compatibility condition will be explained in Section 2.1. It is weaker than the irrepresentability condition defined in (1.1.10) [133].

one required for model selection.

This dissertation consists in two parts. In Part I, we study the problem of assigning measures of confidence to single parameter estimates in high-dimensional models. This is an important problem of inference which has recently gained significance attention among researchers. In Part II, we study the reconstruction problem in compressed sensing, a well-known estimation problem in high-dimensional models, and present a practical scheme to achieve (information-theoretically) optimal undersampling rates for exact recovery. The proposed method is also shown to be robust with respect to measurement noise.

## 1.2 Assigning statistical significance in high-dimensional problems

To date, the bulk of high-dimensional statistical theory has focused on point estimation such as consistency for prediction, oracle inequalities and estimation of parameter vector, model selection, and variable screening. However, the fundamental problem of statistical significance is far less understood in the high-dimensional setting. Uncertainty assessment is particularly important when one seeks subtle statistical patterns in massive amount of data. In this case, any claimed pattern should be supported with some type of significance measure, which quantifies the confidence that the pattern is not a spurious correlation.

We consider a simple example to illustrate this point further. Let  $\mathbf{X} \in \mathbb{R}^{n \times p}$  be a design matrix with independent standard normal entries. For each configuration  $(n, p) = (100, 500), (100, 5000)$ , we generate 1000 realizations of  $\mathbf{X}$ , and for each realization compute  $\hat{r} = \max_{i \geq 2} |\widehat{\text{Corr}}(\mathbf{X}e_1, \mathbf{X}e_i)|$ , where  $\widehat{\text{Corr}}(\mathbf{X}e_1, \mathbf{X}e_i)$  denotes the sample correlation between the first and the  $i$ -th variables.<sup>3</sup> Figure 1.2.1 shows the empirical distribution of  $\hat{r}$ . As we observe, the maximum absolute sample correlation becomes higher as dimensionality increases. This example demonstrates that empirical correlation is not the right metric to claim statistically significant relationships between different variables when the design matrix is not orthogonal.

Following the traditional thinking in frequentist statistics, we treat the parameter vector  $\theta_0$  as a deterministic (unknown) object and the observations as random samples generated according to a model parametrized by  $\theta_0$ . We would like to estimate  $\theta_0$  based on the observed samples and accompany our point estimation with some measures of uncertainty.

---

<sup>3</sup>Recall that  $e_i$  is the vector with one at the  $i$ -th position and zero everywhere else.

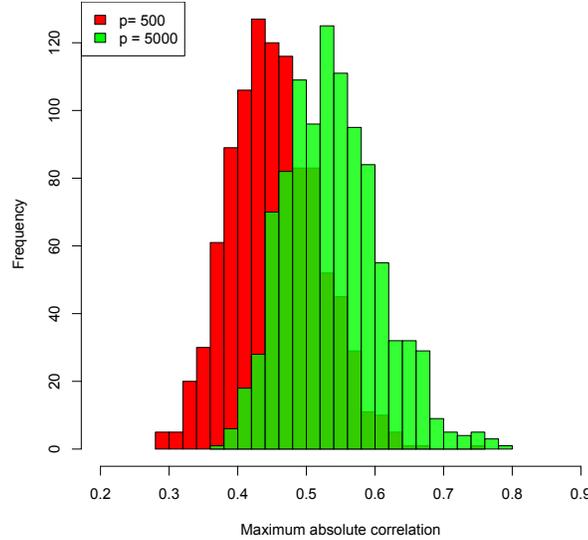


Figure 1.2.1: Frequency of the maximum absolute sample correlation between the first variable and the  $i$ -th variables. This is an illustration of spurious correlation in high-dimensions.

Two classical measures of uncertainty are confidence intervals and  $p$ -values, described below.

- *Confidence interval:* For each single parameter  $\theta_{0,i}$ ,  $i \in [p]$ , and a given significance level  $\alpha \in (0, 1)$ , we are interested in constructing intervals  $[\underline{\theta}_i, \bar{\theta}_i]$  such that

$$\mathbb{P}(\theta_{0,i} \in [\underline{\theta}_i, \bar{\theta}_i]) \geq 1 - \alpha.$$

- *Hypothesis testing and  $p$ -values:* For each  $i \in [p]$ , we are interested in testing whether variable  $i$  is significant in predicting the response variable. More specifically, we are interested in testing null hypotheses of the form

$$H_{0,i} : \theta_{0,i} = 0, \tag{1.2.1}$$

versus its alternative  $H_{A,i} : \theta_{0,i} \neq 0$ , for  $i \in [p]$ . Any hypothesis testing procedure faces two types of errors: false positives or type I errors (incorrectly rejecting  $H_{0,i}$ , while  $\theta_{0,i} = 0$ ), and false negatives or type II errors (failing to reject  $H_{0,i}$ , while  $\theta_{0,i} \neq 0$ ). The probabilities of these two types of errors will be denoted, respectively,

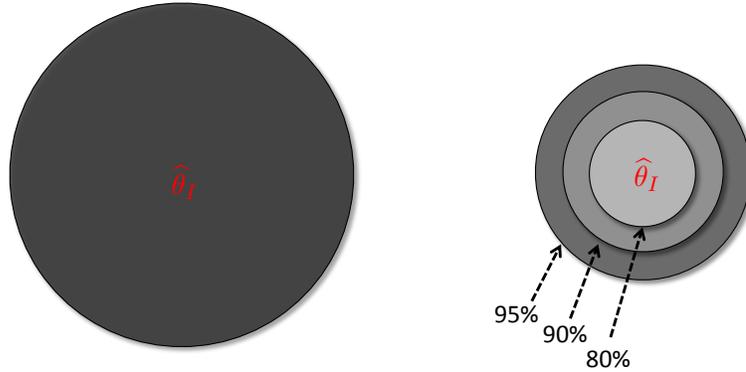


Figure 1.2.2: Lefthand side: low dimensional projection of the high probability region given by (1.1.9) ( $q = 2$ ). The ball radius is  $O(\sqrt{(s_0 \log p)/n})$ . Righthand side: confidence regions for low dimensional targets,  $\theta_I$ , where  $I$  is a subset of indices with constant size. The radius here is  $O(1/\sqrt{n})$ .

by  $\alpha$  and  $\beta$ . The quantity  $1 - \beta$  is also referred to as the power of the test, and  $\alpha$  as its significance level.

Central to any hypothesis testing procedure is the construction of  $p$ -value as a measure of statistical significance. The challenge in high-dimensional models is indeed the construction of  $p$ -values, which control type I error measure while having good power for detecting alternatives. We will discuss these challenges in Section 1.2.2.

It is instructive to see how the estimation error bound (1.1.9) compares to our goal of constructing confidence intervals. Note that the bound (1.1.9) is with high probability, which gives absolute (asymptotic) certainty for the intervals. However, we are interested in confidence intervals for single parameter  $\theta_{0,i}$  (low-dimensional targets), and if we use the bound (1.1.9), with  $q = 2$ , the resulting intervals will be of size  $O(\sqrt{(s_0 \log p)/n})$ . By contrast, for each single parameter  $\theta_{0,i}$  we construct  $(1 - \alpha)$  confidence interval of size  $O(1/\sqrt{n})$  which is much smaller. This is schematically illustrated in Figure 1.2.2.

Similarly, the results for model selection require the stringent irrepresentability condition to hold for the design matrix. Further, they assume the rather strong  $\theta_{\min}$  condition, saying that the nonzero parameters must be sufficiently large. These constraints are hard to be fulfilled in many real problems and indeed the  $\theta_{\min}$  condition cannot even be verified. Here, instead of making binary choices about the activeness of variables in the model, we are interested in developing procedures to quantify the statistical uncertainty that is intrinsic

in any such decision. Hypothesis testing provides a standard framework to address this type of problems. As we will see, to control type I error, there is no need to the  $\theta_{\min}$  condition. Further, we only assume compatibility condition on the design matrix which is weaker than the irrepresentability condition.

### 1.2.1 Why is it important?

We discuss the importance of uncertainty assessment from three different perspectives.

- *Scientific discoveries:* Consider a prostate data set that contains a few hundreds of samples in two classes, normal and tumor, along with expression levels of a few thousands of genes for each sample. Suppose that we are interested in finding the relevant genes in predicting prostate cancer. Clearly, this is a high-dimensional data set since the number of variables (gene expression levels) are much larger than the number of samples; a usual trend in genetics data analysis. As explained in the previous example (cf. [Figure 1.2.1](#)), empirical correlation is not the right indicator of relevance in this regime. If we make a claim about the importance of a gene on prostate tumor, we need to support our finding with a confidence measure like  $p$ -value. Otherwise, there is no evidence that our discovery is not just a spurious effect.
- *Decision making:* Uncertainty assessment is crucial whenever we intend to take actions on the basis of our statistical model of the data. For instance, in targeted online advertising, a typical inference task is to predict an individual's buying activity on the basis of her browsing history, location, position and relationships in a social network, and so on. Typically, these problems are high-dimensional due to the large number of attributes that are available for each individual. Existing methods allow to predict an expected behavior for each individual, but do not account for its intrinsic variability. Variability quantification is instead an important component of policy designs and decision making strategies.
- *Stopping rules in optimization:* M-estimators are constructed by optimizing a suitably regularized loss function, cf. [\(1.1.1\)](#). Solving such optimizations over millions of samples and billions of variables is computationally very challenging. Note that the classical polynomial time Interior Point methods (IPMs) are capable to solve convex programs within high accuracy at a low iteration count. However, the iteration cost of these methods scale nonlinearly with the problem's dimension (number of samples and

variables). As a result, these methods become impractical for very-large scale problems since a single iteration lasts forever! Motivated by the need for arbitrary-scale, decentralized solvers, the first order methods (FOMs) with computationally cheap iterations have been developed. Well-known FOMs include gradient descent, coordinate descent, Nesterov’s accelerated method [103, 104], Iterative Shrinkage Thresholding [9], and Alternating Direction Method of Multiplier (ADMM) [15]. For problems with favorable geometry, good FOMs exhibit dimension-independent convergence rate; however, they have only sublinear rate of convergence. As a result, these iterative algorithms should be run for a large number of iterations, and we need some type of stopping rule to know when to stop the algorithm if we desire to get within  $\epsilon$  accuracy of the solution.

Most of the stopping rules are based on the analysis of convergence rates of FOMs, saying that in order to be within  $\epsilon$  accuracy of  $\hat{\theta}$ , we need to run the method for  $n_\epsilon$ , or  $O(n_\epsilon)$  number of iterations. A subtle point to note is that here optimization serves as a tool to fulfill our inference goal, i.e, finding  $\theta_0$ . Hence, a holistic stopping rule must measure how far we are from the object of interest  $\theta_0$ , not  $\hat{\theta}$ . For instance, one choice would be based on the confidence intervals: At each iteration, construct a confidence interval for  $\theta_0$  as per the current estimate. Then stop the iterations when the change in the consecutive estimates is negligible with respect to the interval size.

### 1.2.2 Why is it hard?

In a nutshell, the main challenges are due to high-dimensionality. In the low-dimensional regime ( $p < n$ ), either exact distributional characterization of the estimators are available, or asymptotically exact ones can be derived from large sample theory [135]. On the other side, fitting high-dimensional statistical models often requires the use of non-linear parameter estimation procedures and in general, it is impossible to characterize distribution of such estimators.<sup>4</sup> Consequently, there is no commonly accepted procedure for constructing  $p$ -values in high-dimensional of statistics.

As we will discuss in Section 2.2, M-estimator  $\hat{\theta}$  is biased towards small  $\mathcal{R}(\theta)$ , and of course the bias vector is unknown. This is a major challenge in constructing confidence intervals and computing  $p$ -values based on M-estimators.

---

<sup>4</sup>In some special cases, such as design matrices with i.i.d. Gaussian entries, the distribution of Lasso estimator can be characterized. This will be discussed in details in Chapter 4.

Further, (limiting) distribution of an M-estimator is non-continuous. For example, the distribution of Lasso estimator is non-Gaussian with point mass at zero. Because of this, standard bootstrap or subsampling techniques do not give honest confidence regions or  $p$ -values.

### 1.2.3 Contributions & Organization (Part I)

We consider a general debiased estimator of the form  $\hat{\theta}^u = \hat{\theta}^n + (1/n)M\mathbf{X}^\top(Y - \mathbf{X}\hat{\theta}^n)$ , where  $\hat{\theta}^n \equiv \hat{\theta}(Y, \mathbf{X}; \lambda)$  denotes the Lasso estimator. We introduce a figure of merit of the pair  $M, \mathbf{X}$ , termed the generalized coherence parameter  $\mu_*(\mathbf{X}; M)$ . We show that, if the generalized coherence is small, then the debiasing procedure is effective in the sense that  $\text{Bias}(\hat{\theta}^u)$  is smaller than  $\text{Bias}(\hat{\theta}^n)$  in order of magnitude.<sup>5</sup>

In case of random designs, we show that the generalized coherence parameter can be made as small as  $\sqrt{(\log p)/n}$ , through a convex optimization procedure for computing  $M$  (cf. [Algorithm 1](#)). This results in a bound on the bias of  $\hat{\theta}^u$ : the largest entry of the bias is of order  $(s_0 \log p)/n$ . This must be compared with the standard deviation of  $\hat{\theta}_i^u$ , which is of order  $\sigma/\sqrt{n}$ . The conclusion is that, for  $n = \omega((s_0 \log p)^2)$ , the bias of  $\hat{\theta}^u$  is negligible.

**Distributional characterization.** We further characterize the (limiting) distribution of the debiased estimator  $\hat{\theta}^u$ , whence we derive confidence intervals and hypothesis testing procedures for low-dimensional marginals of  $\theta_0$ . The basic intuition is that  $\hat{\theta}^u$  is approximately Gaussian with mean  $\theta_0$ , and known covariance structure. Hence standard optimal tests can be applied.

Chapter 2 focuses on the debiasing approach, establishing its limiting distributional characterization, and deriving confidence intervals for low dimensional marginals of  $\theta_0$ . Section 2.4.3 extends our results to non-Gaussian noise using the central limit theorem for triangular arrays, and Section 2.6 provides a generalization of this approach to regularized maximum likelihood estimators.

In Chapter 3, we address the problem of hypothesis testing in high-dimensional regression models. Relying on the limiting distributional characterization of  $\hat{\theta}$ , we construct valid two-sided  $p$ -values for  $H_{0,i}$ . This controls the type I error of our proposed testing procedure (3.0.3).

---

<sup>5</sup>We refer to Section 2.1 for a formal definition of bias.

**Optimality.** In Chapter 3, we prove a general lower bound on the power of our testing procedure. Moreover, taking a minimax point of view, we prove a general upper bound on the minimax power of tests with a given significance level  $\alpha$ , in the case of Gaussian random designs with i.i.d. rows. By comparing these bounds, we conclude that the asymptotic efficiency of our approach is constant-optimal. Namely, it is lower bounded by a constant  $1/\eta_{\Sigma, s_0}$  which is bounded away from 0. Here  $\Sigma$  is the population level covariance matrix of the design, i.e.  $\Sigma = \mathbb{E}(X_1 X_1^\top)$ , and  $\eta_{\Sigma, s_0}$  is always upper bounded by the condition number of  $\Sigma$ . In particular,  $\eta_{I, s_0} = 1$ . Section 3.4 contains an overview of some of the most recent and related procedures for hypothesis testing in high-dimensional regression, namely multisample splitting [142, 96], Ridge-type projection estimator [16], and low dimensional projection estimator (LDPE), proposed by [149].

**Hypothesis testing under optimal sample size.** In Chapter 4, we focus on Gaussian random designs with i.i.d. rows  $X_i \sim \mathcal{N}(0, \Sigma)$ . In case of  $\Sigma = I$ , we build upon a rigorous characterization of the asymptotic distribution of the Lasso estimator and its debiased version, and propose a testing procedure under the optimal sample size, i.e.  $n = O(s_0 \log(p/s_0))$ . For general Gaussian designs, we show that a similar distributional characterization (termed ‘standard distributional limit’) can be derived from the replica heuristics in statistical physics. This derivation suggests near-optimality of the statistical power for a large class of Gaussian designs.

**Validation.** In Chapter 5, we validate our approach on both synthetic and real data, comparing it with other proposals. In the interest of reproducibility, an R implementation of our algorithm is available at <http://www.stanford.edu/~montanar/sslasso/>.

Proofs of theorems and technical lemmas in Part I are given in Chapter 6.

#### 1.2.4 Previously published material

The chapters of Part I are based on previous publications:

- Publication [75]: Adel Javanmard and Andrea Montanari. Confidence Intervals and Hypothesis Testing for High-Dimensional Regression. *To appear in Journal of Machine Learning Research*, 2014. (Short version [71] is published in *Advances in Neural Information Processing Systems*, pages 1187-1195, 2013.)

- Publication [72]: Adel Javanmard and Andrea Montanari. Hypothesis Testing in High-Dimensional Regression under the Gaussian Random Design Model: Asymptotic Theory. *To appear in IEEE Transaction on Information Theory*, 2014.

### 1.3 Optimal compressed sensing via spatial coupling and approximate message passing

Compressed sensing refers to a set of techniques that recover the signal accurately from undersampled measurements. In other words, instead of first measuring the signal and then compressing it, these techniques aim at measuring the signal in an already compressed form without missing any information.

Traditional sampling methods, such as Shannon-Nyquist-Whittaker, demand sampling rate proportional to the frequency bandwidth of the signal. By contrast, compressed sensing techniques require sampling rate proportional to the *information content* of the signal. Smaller sampling rate translates to faster and cheaper data collection and processing.

The theory of compressed sensing has three key ingredients: structure of the signal, sensing mechanism, and reconstruction algorithm. In the following, we briefly discuss these components.

**Structure of the signal:** Many real world signals enjoy some types of structure. For instance, the signal of interest belong to some known class, or it is generated according to some known distribution. A very common structure is *sparsity*, meaning that most of the information is concentrated in relatively few coordinates of the signal. More specifically, let  $x \in \mathbb{R}^n$  be an  $n$ -dimensional signal, and define its  $\ell_0$  norm as

$$\|\theta\|_0 = |\{i : \theta_i \neq 0\}|,$$

i.e., the number of non-zero coordinates of  $x$ . We refer to  $\epsilon \equiv \|\theta\|_0/n$  as the *sparsity level* of the signal as it represents the fraction of nonzero entries. Fortunately, many natural signals have a small sparsity level in some domain. Sparsity has long been used by compression algorithms to decrease storage costs; the goal of sparse recovery and compressed sensing is to exploit sparsity to decrease the required sampling rate for exact recovery.

**Sensing mechanism:** The simplest way of taking measurements of signal  $x \in \mathbb{R}^n$  is through a linear operator, namely the measurements are given by

$$y = Ax, \quad (1.3.1)$$

where  $A \in \mathbb{R}^{m \times n}$  is the measurement matrix. Of course, in most of the cases, the measurement are contaminated by noise and hence we consider a more general model

$$y = Ax + w, \quad (1.3.2)$$

with  $w$  representing the noise vector. We recall that the reconstruction problem in compressed sensing requires to reconstruct  $x$  from the measured vector  $y \in \mathbb{R}^m$ , and the measurement matrix  $A \in \mathbb{R}^{m \times n}$ .

Conventional wisdom in compressed sensing suggests that random isotropic vectors, with small coherence number, provide a suitable class of measurement vectors. To be more specific, in this case the sensing vectors are independently sampled from a population  $F$ , such that  $F$  obeys the isotropy property

$$\mathbb{E}aa^* = \mathbf{I}, \quad a \sim F.$$

Further, the coherence parameter  $\mu(F)$  is defined to be the smallest number such that

$$\max_{1 \leq i \leq n} |a_i|^2 \leq \mu(F)$$

holds deterministically or stochastically. It turns out the smaller  $\mu(F)$ , the fewer measurements are needed for accurate recovery. Some examples are the random matrices whose entries are drawn independently from Gaussian, Bernoulli, or any other sub-gaussian distribution. Another example is obtained by sampling, uniformly at random, from rows of the DFT matrix.

**Reconstruction algorithm:** Since  $m < n$ , the system of equation (1.3.1) does not, in general, admit a unique solution. Therefore the structure of the signal must be used as a side information in the recovery algorithm. This leads to nonlinear and more sophisticated schemes compared to the traditional sampling theory wherein the signals are reconstructed by applying simple linear operators.

One popular class of reconstruction schemes uses linear programming methods. Consider the noiseless model (1.3.1). Among the infinitely many solutions, we are interested in the sparsest one. This can be formulated as the following optimization problem:

$$\begin{aligned} & \underset{x}{\text{minimize}} && \|x\|_0 \\ & \text{subject to} && y = Ax \end{aligned} \tag{1.3.3}$$

However, this optimization is NP-complete and cannot be used in practice. Chen et al. [31] proposed the following convex optimization, called Basis Pursuit, as a convex relaxation of (1.3.3).

$$\begin{aligned} & \underset{x}{\text{minimize}} && \|x\|_1 \\ & \text{subject to} && y = Ax \end{aligned} \tag{P_1}$$

Clearly, this optimization can be cast as a linear programming problem. This method is shown to be remarkably successful in recovering the signal and an elegant theory has been developed for it [31, 30, 38, 23, 54].

There is a large corpus of research on recovering sparse signals from a few number of measurements [36, 23, 37]. It is shown that if only  $k$  entries of  $x$  are non-vanishing, then roughly  $m \gtrsim 2k \log(n/k)$  measurements are sufficient for  $A$  random, and reconstruction can be solved efficiently by convex programming. Deterministic sensing matrices achieve similar performance, provided they satisfy a suitable restricted isometry condition [28]. On top of this, reconstruction is robust with respect to additive noise in measurements [24, 44], namely, under the noisy model (1.3.2) with, say,  $w \in \mathbb{R}^m$  a random vector with i.i.d. components  $w_i \sim \mathbf{N}(0, \sigma^2)$ . In this context, the notions of ‘robustness’ or ‘stability’ refers to the existence of universal constants  $C$  such that the per-coordinate mean square error in reconstructing  $x$  from noisy observation  $y$  is upper bounded by  $C \sigma^2$ .

From an information-theoretic point of view it remains however unclear why we cannot achieve the same goal with far fewer than  $2k \log(n/k)$  measurements. Indeed, we can interpret Eq. (1.3.1) as describing an analog data compression process, with  $y$  a compressed version of  $x$ . From this point of view, we can encode all the information about  $x$  in a single real number  $y \in \mathbb{R}$  (i.e., use  $m = 1$ ), because the cardinality of  $\mathbb{R}$  is the same as the one of

$\mathbb{R}^n$ . Motivated by this puzzling remark, Wu and Verdú [143] introduced a Shannon-theoretic analogue of compressed sensing, wherein the vector  $x$  has i.i.d. components  $x_i \sim p_X$ . Crucially, the distribution  $p_X$  is available to, and may be used by the reconstruction algorithm. Under the mild assumptions that sensing is linear (as per (1.3.1)), and that the reconstruction mapping is Lipschitz continuous, they proved that compression is asymptotically lossless if and only if

$$m \geq n \bar{d}(p_X) + o(n). \quad (1.3.4)$$

Here  $\bar{d}(p_X)$  is the (upper) Rényi information dimension of the distribution  $p_X$ . We refer to Section 7.1 for a precise definition of this quantity. Suffices to say that, if  $p_X$  is  $\varepsilon$ -sparse (i.e., if it puts mass at most  $\varepsilon$  on nonzeros) then  $\bar{d}(p_X) \leq \varepsilon$ . Also, if  $p_X$  is the convex combination of a discrete part (sum of Dirac's delta) and an absolutely continuous part (with a density), then  $\bar{d}(p_X)$  is equal to the weight of the absolutely continuous part.

This result is quite striking. For instance, it implies that, for random  $k$ -sparse vectors,  $m \geq k + o(n)$  measurements are sufficient. Also, if the entries of  $x$  are random and take values in, say,  $\{-10, -9, \dots, -9, +10\}$ , then a sublinear number of measurements  $m = o(n)$ , is sufficient! At the same time, the result of Wu and Verdú presents two important limitations. First, it does not provide robustness guarantees of the type described above. Second and most importantly, it does not provide any computationally practical algorithm for reconstructing  $x$  from measurements  $y$ .

In an independent line of work, Krzakala et al. [82] developed an approach that leverages on the idea of *spatial coupling*. This idea was introduced in the compressed sensing literature by Kudekar and Pfister [84] (see [85] and Section 7.3 for a discussion of earlier work on this topic). Spatially coupled matrices are, roughly speaking, random sensing matrices with a band-diagonal structure. The analogy is, this time, with channel coding.<sup>6</sup> In this context, spatial coupling, in conjunction with message-passing decoding, allows to achieve Shannon capacity on memoryless communication channels. It is therefore natural to ask whether an approach based on spatial coupling can enable to sense random vectors  $x$  at an undersampling rate  $m/n$  close to the Rényi information dimension of the coordinates of  $x$ ,  $\bar{d}(p_X)$ . Indeed, the authors of [82] evaluate such a scheme numerically on a few classes of random vectors and demonstrate that it indeed achieves rates close to the fraction of

---

<sup>6</sup>Unlike [82], we follow here the terminology developed within coding theory.

non-zero entries. They also support this claim by insightful statistical physics arguments.

In Part II of this dissertation, we fill the gap between the above works, and present the following contributions:

**Construction.** We describe a construction for spatially coupled sensing matrices  $A$  that is somewhat broader than the one of [82] and give precise prescriptions for the asymptotic values of various parameters. We also use a somewhat different reconstruction algorithm from the one in [82], by building on the approximate message passing (AMP) approach of [42, 43]. AMP algorithms have the advantage of smaller memory complexity with respect to standard message passing, and of smaller computational complexity whenever fast multiplication procedures are available for  $A$ .

**Rigorous proof of convergence.** Our main contribution is a rigorous proof that the above approach indeed achieves the information-theoretic limits set out by Wu and Verdú [143]. Indeed, we prove that, for sequences of spatially coupled sensing matrices  $\{A(n)\}_{n \in \mathbb{N}}$ ,  $A(n) \in \mathbb{R}^{m(n) \times n}$  with asymptotic undersampling rate  $\delta = \lim_{n \rightarrow \infty} m(n)/n$ , AMP reconstruction is with high probability successful in recovering the signal  $x$ , provided  $\delta > \bar{d}(p_X)$ .

**Robustness to noise.** We prove that the present approach is robust<sup>7</sup> to noise in the following sense. For any signal distribution  $p_X$  and undersampling rate  $\delta$ , there exists a constant  $C$  such that the output  $\hat{x}(y)$  of the reconstruction algorithm achieves a mean square error per coordinate  $n^{-1} \mathbb{E}\{\|\hat{x}(y) - x\|_2^2\} \leq C \sigma^2$ . This result holds under the noisy measurement model (1.3.2) for a broad class of noise models for  $w$ , including i.i.d. noise coordinates  $w_i$  with  $\mathbb{E}\{w_i^2\} = \sigma^2 < \infty$ .

**Non-random signals.** Our proof does not apply uniquely to random signals  $x$  with i.i.d. components, but indeed to more general sequences of signals  $\{x(n)\}_{n \in \mathbb{N}}$ ,  $x(n) \in \mathbb{R}^n$  indexed by their dimension  $n$ . The conditions required are: (1) that the empirical distribution of the coordinates of  $x(n)$  converges (weakly) to  $p_X$ ; and (2) that  $\|x(n)\|_2^2$  converges to the second moment of the asymptotic law  $p_X$ .

There is a fundamental reason why this more general framework turns out to be equivalent to the random signal model. This can be traced back to the fact that, within

---

<sup>7</sup>This robustness bound holds for all  $\delta > \bar{D}(p_X)$ , where  $\bar{D}(p_X)$  is the upper MMSE dimension of  $p_X$ . (see Definition 7.1.4). It is worth noting that  $\bar{D}(p_X) = \bar{d}(p_X)$  for a broad class of distributions  $p_X$  including distributions without *singular continuous component*.

	Classical Compressed Sensing	Our Approach
<b>Structure</b>	Sparsity	Information dimension
<b>Rate</b>	$m = Ck \log(n/k)$	$m = \bar{d}(p_X) \cdot n$
<b>Measurements</b>	Random isotropic vectors	Spatially coupled matrices
<b>Reconstruction</b>	Convex optimization	Bayesian AMP
<b>Robustness</b>	$\text{MSE} \leq C\sigma^2$	$\text{MSE} \leq C(x)\sigma^2$

Table 1.3.1 Comparison between classical approach to compressed sensing and our approach.

our construction, the columns of the matrix  $A$  are probabilistically exchangeable. Hence any vector  $x(n)$  is equivalent to the one whose coordinates have been randomly permuted. The latter is in turn very close to the i.i.d. model. By the same token, the rows of  $A$  are exchangeable and hence the noise vector  $w$  does not need to be random either.

Interestingly, the present framework changes the notion of ‘structure’ that is relevant for reconstructing the signal  $x$ . Indeed, the focus is shifted from the *sparsity* of  $x$  to the *information dimension*  $\bar{d}(p_X)$ . In other words, the signal structure that facilitates recovery from a small number of linear measurements is the low-dimensional structure in an information theoretic sense, quantified by the information dimension of the signal. [Table 1.3.1](#) provides a comparison between pillars of traditional compressed sensing and principles of our approach. We refer to Part II for a detailed discussion on the salient features of our approach presented in [Table 1.3.1](#).

### 1.3.1 A toy example

The following example demonstrates the dramatic improvement achieved by our scheme. Consider a signal  $x \in \mathbb{R}^n$  whose coordinates are generated i.i.d. from the distribution  $p_X = 0.2\delta_0 + 0.3\delta_1 + 0.2\delta_{-1} + 0.2\delta_3 + 0.1\text{Uniform}(-2, 2)$ . Further suppose that we take  $m$  noiseless linear measurements of  $x$ . Classical scheme based on  $\ell_1$  minimization ( $P_1$ ) require  $m \geq 0.97n$  for exact recovery. More generally, Donoho and Tanner [34] showed that the reconstruction algorithm ( $P_1$ ) undergoes a phase transition: they characterized a curve  $\epsilon \rightarrow \delta_{\ell_1}(\epsilon)$  in the  $(\epsilon, \delta)$  plane such that the following happens for sensing matrices  $A \in \mathbb{R}^{m \times n}$

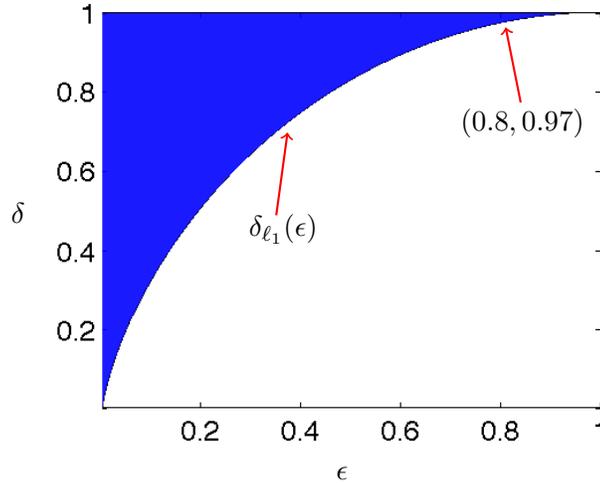


Figure 1.3.3: Donoho-Tanner phase diagram for algorithm  $(P_1)$

with i.i.d. Gaussian entries, in the large-system limit  $n, m \rightarrow \infty$ , with  $m/n = \delta$ . The reconstruction algorithm  $(P_1)$  correctly recovers the original signal, with high probability, provided  $\delta > \delta_{l_1}(\epsilon)$ , while for  $\delta < \delta_{l_1}(\epsilon)$  the algorithm fails with high probability.<sup>8</sup>

Donoho-Tanner phase diagram is depicted in [Figure 1.3.3](#). In our example,  $p_X(\{0\}) = 0.2$  and hence  $\epsilon = 0.8$ . Given that  $\delta_{l_1}(0.8) = 0.97$ , the reconstruction scheme  $(P_1)$  requires  $m = n\delta \geq 0.97n$  for exact recovery and this is indeed the fundamental limit for its performance.

On the other side,  $\bar{d}(p_X) = 0.1$  and therefore our method can recover signal  $x$  accurately from  $m = 0.1n$  spatially coupled measurements. Let us stress that our reconstruction approach requires to know the underlying distribution  $p_X$  and in this sense is not universal.

### 1.3.2 Organization (Part II)

In [Chapter 7](#), we state formally our results and discuss their implications and limitations, as well as relations with earlier work. [Chapter 8.3](#) provides a precise description of the matrix construction and the reconstruction algorithm. [Chapter 10](#) reduces the proof of our main results into two key lemmas. One of these lemmas is a (quite straightforward) generalization of the state evolution technique of [\[42, 7\]](#). The second lemma characterizes the behavior of

<sup>8</sup>A universality property is also conjectured for this phase transition phenomena, saying that if matrix  $A$  is sampled randomly from a “well-behaved” probability distribution, reconstruction  $(P_1)$  exhibits the same phase transition. Extensive numerical evidence is presented by [\[34\]](#) for a wide collection of ensembles, including partial Fourier, partial Hadamard, expander graphs, iid  $\pm 1$ .

the state evolution recursion, and is proved in Chapter 11. In Chapter 12, we study the problem of sampling a random signal with sparse support in frequency domain, and propose a sampling scheme inspired by the idea of spatial coupling. As we discuss one possible implementation of this idea is through Gabor transform. We show empirically that this scheme achieves correct reconstruction at information-theoretically optimal undersampling rate.

The proof of a number of intermediate technical steps is deferred to the appendices.

### 1.3.3 Previously published material

The chapters of Part II are based on previous publications:

- Publication [41]: David L. Donoho, Adel Javanmard and Andrea Montanari. Information-Theoretically Optimal Compressed Sensing via Spatial Coupling and Approximate Message Passing. *IEEE Transactions on Information Theory*, 59(11):7434-7464, 2013. (Short version [40] is published in *IEEE International Symposium on Information Theory (ISIT)*, pages 1231-1235, 2012.)
- Publication [70]: Adel Javanmard and Andrea Montanari. Subsampling at Information Theoretically Optimal Rates. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2431-2435, 2012.
- Publication [74]: Adel Javanmard and Andrea Montanari. State Evolution for General Approximate Message Passing Algorithms, with Applications to Spatial Coupling. *Journal of Information and Inference*, 2(2):115-144, 2013.

## Part I

# Assigning Statistical Significance in High-Dimensional Problems

## Chapter 2

# Confidence Intervals for High-Dimensional Models

We recall the linear regression model with  $n$  i.i.d. samples and  $p$  parameters, as described by (1.1.6), where the goal is to learn vector of parameters  $\theta_0$ .

In the classic setting,  $n \gg p$  and the estimation method of choice is ordinary least squares yielding  $\hat{\theta}^{\text{OLS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top Y$ . Least squares  $\hat{\theta}^{\text{OLS}}$  has an explicit formulation, and admits an exact distributional characterization. In particular  $\hat{\theta}^{\text{OLS}}$  is Gaussian with mean  $\theta_0$  and covariance  $\sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1}$ . This directly allows to construct confidence intervals for single parameters  $\theta_{0,i}$  based on the estimate  $\hat{\theta}^{\text{OLS}}$ .<sup>1</sup>

In the high-dimensional setting where  $p > n$ , the matrix  $(\mathbf{X}^\top \mathbf{X})$  is rank deficient and one has to resort to structured estimators. We discussed some remarkable properties of these estimator in the introduction. These properties, however, come at a price. Deriving an exact characterization for the distribution of structured estimators is not tractable in general, and hence there is no simple procedure to construct confidence intervals and  $p$ -values. A closely related property is that structured estimators are biased, an unavoidable property in high dimension, since a point estimate in  $p$ -dimension must be produced from data in lower dimension  $Y \in \mathbb{R}^n$ ,  $n < p$ .

In our presentation, we focus on Lasso estimator  $\hat{\theta}^n \equiv \hat{\theta}^n(Y, \mathbf{X}; \lambda)$ , as per (1.1.7) which promotes sparse reconstructions through an  $\ell_1$  penalty. Later in the chapter, we discuss the case of general structured estimators.

---

<sup>1</sup>For instance, letting  $Q \equiv (\mathbf{X}^\top \mathbf{X}/n)^{-1}$ ,  $[\hat{\theta}_i^{\text{OLS}} - 1.96\sigma\sqrt{Q_{ii}/n}, \hat{\theta}_i^{\text{OLS}} + 1.96\sigma\sqrt{Q_{ii}/n}]$  is a 95% confidence interval [141].

## 2.1 Preliminaries and notations

In this section we introduce some basic definitions used throughout this part, starting with simple notations.

For a matrix  $A$  and set of indices  $I, J$ , we let  $A_{I,J}$  denote the submatrix formed by the rows in  $I$  and columns in  $J$ . Also,  $A_{I,\cdot}$  (resp.  $A_{\cdot,I}$ ) denotes the submatrix containing just the rows (reps. columns) in  $I$ . Likewise, for a vector  $v$ ,  $v_I$  is the restriction of  $v$  to indices in  $I$ . We use the shorthand  $A_{I,J}^{-1} = (A^{-1})_{I,J}$ . In particular,  $A_{i,i}^{-1} = (A^{-1})_{i,i}$ . The maximum and the minimum singular values of  $A$  are respectively denoted by  $\sigma_{\max}(A)$  and  $\sigma_{\min}(A)$ . We write  $\|v\|_p$  for the standard  $\ell_p$  norm of a vector  $v$ , i.e.,  $\|v\|_p = (\sum_i |v_i|^p)^{1/p}$ . and  $\|v\|_0$  for the number of nonzero entries of  $v$ . For a matrix  $A$ ,  $\|A\|_p$  is the  $\ell_p$  operator norm, and  $|A|_p$  is the elementwise  $\ell_p$  norm. For a vector  $v$ ,  $\text{supp}(v)$  represents the positions of nonzero entries of  $v$ . Throughout,  $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ , and  $\Phi(x) \equiv \int_{-\infty}^x \phi(t)dt$  respectively denote the PDF and the CDF of the standard normal distribution. Finally, *with high probability* (w.h.p) means with probability converging to one as  $n \rightarrow \infty$ .

We let  $\widehat{\Sigma} \equiv \mathbf{X}^\top \mathbf{X}/n$  be the sample covariance matrix. For  $p > n$ ,  $\widehat{\Sigma}$  is always singular. However, we may require  $\widehat{\Sigma}$  to be nonsingular for a restricted set of directions.

**Definition 2.1.1.** *Given a symmetric matrix  $\widehat{\Sigma} \in \mathbb{R}^{p \times p}$  and a set  $S \subseteq [p]$ , the corresponding compatibility constant is defined as*

$$\phi^2(\widehat{\Sigma}, S) \equiv \min_{\theta \in \mathbb{R}^p} \left\{ \frac{|S| \langle \theta, \widehat{\Sigma} \theta \rangle}{\|\theta_S\|_1^2} : \theta \in \mathbb{R}^p, \|\theta_{S^c}\|_1 \leq 3\|\theta_S\|_1 \right\}. \quad (2.1.1)$$

We say that  $\widehat{\Sigma} \in \mathbb{R}^{p \times p}$  satisfies the compatibility condition for the set  $S \subseteq [p]$ , with constant  $\phi_0$  if  $\phi(\widehat{\Sigma}, S) \geq \phi_0$ . We say that it holds for the design matrix  $\mathbf{X}$ , if it holds for  $\widehat{\Sigma} = \mathbf{X}^\top \mathbf{X}/n$ .

In the following, we shall drop the argument  $\widehat{\Sigma}$  if clear from the context. Note that a slightly more general definition is used normally [19, Section 6.13], where the condition  $\|\theta_{S^c}\|_1 \leq 3\|\theta_S\|_1$ , is replaced by  $\|\theta_{S^c}\|_1 \leq L\|\theta_S\|_1$ . The resulting constant  $\phi(\widehat{\Sigma}, S, L)$  depends on  $L$ . For the sake of simplicity, we restrict ourselves to the case  $L = 3$ .

**Definition 2.1.2.** *The sub-gaussian norm of a random variable  $X$ , denoted by  $\|X\|_{\psi_2}$ , is defined as*

$$\|X\|_{\psi_2} = \sup_{q \geq 1} q^{-1/2} (\mathbb{E}|X|^q)^{1/q}.$$

For a random vector  $X \in \mathbb{R}^n$ , its sub-gaussian norm is defined as

$$\|X\|_{\psi_2} = \sup_{x \in S^{n-1}} \|\langle X, x \rangle\|_{\psi_2},$$

where  $S^{n-1}$  denotes the unit sphere in  $\mathbb{R}^n$ .

**Definition 2.1.3.** The sub-exponential norm of a random variable  $X$ , denoted by  $\|X\|_{\psi_1}$ , is defined as

$$\|X\|_{\psi_1} = \sup_{q \geq 1} q^{-1} (\mathbb{E}|X|^q)^{1/q}.$$

For a random vector  $X \in \mathbb{R}^n$ , its sub-exponential norm is defined as

$$\|X\|_{\psi_1} = \sup_{x \in S^{n-1}} \|\langle X, x \rangle\|_{\psi_1},$$

where  $S^{n-1}$  denotes the unit sphere in  $\mathbb{R}^n$ .

We next formally define bias of an estimator.

**Definition 2.1.4.** Given an estimator  $\tilde{\theta}^n$  of the parameter vector  $\theta_0$ , we define its bias to be the vector

$$\text{Bias}(\tilde{\theta}^n) \equiv \mathbb{E}\{\tilde{\theta}^n - \theta_0 | \mathbf{X}\}. \quad (2.1.2)$$

Note that, if the design is random,  $\text{Bias}(\tilde{\theta}^n)$  is a measurable function of  $\mathbf{X}$ . If the design is deterministic,  $\text{Bias}(\tilde{\theta}^n)$  is a deterministic quantity as well, and the conditioning is superfluous.

Throughout, we denote by  $S \equiv \text{supp}(\theta_0)$  the support of  $\theta_0 \in \mathbb{R}^p$ , defined as

$$\text{supp}(\theta_0) \equiv \{i \in [p] : \theta_{0,i} \neq 0\},$$

where we use the notation  $[p] = \{1, \dots, p\}$ . We further let  $s_0 \equiv |S|$ .

## 2.2 The bias of the Lasso

Structured estimators are biased due to the regularization term. In particular, Lasso estimator is biased towards small  $\ell_1$  norm.

Let us begin with the simple case of  $\mathbf{X} = \mathbf{I}$ ,  $n = p$ . The observed samples are given by  $Y = \theta_0 + W$ . In this case, Lasso estimator is given by the soft thresholding function, i.e.,  $\hat{\theta}^n = \eta(Y; \lambda)$ , where  $\eta : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is defined as

$$\eta(y; \lambda) = \begin{cases} y - \lambda & \text{if } y > \lambda, \\ 0 & \text{if } -\lambda \leq y \leq \lambda, \\ y + \lambda & \text{otherwise.} \end{cases} \quad (2.2.1)$$

and for a vector  $u \in \mathbb{R}^n$ , we let  $\eta(u)$  denote the vector  $(\eta(u_1), \dots, \eta(u_n))$  obtained by applying  $\eta$  component-wise. By a simple algebraic manipulation, we have

$$\text{Bias}(\hat{\theta}_i^n) = -\lambda + (\lambda - \theta_{0,i})\Phi(\lambda - \theta_{0,i}) + (\lambda + \theta_{0,i})\Phi(-\lambda - \theta_{0,i}) + \phi(\lambda - \theta_{0,i}) - \phi(-\lambda - \theta_{0,i}).$$

Clearly, by choosing  $\theta_{0,i}$  large enough,  $\text{Bias}(\hat{\theta}_i^n)$  and therefore  $\|\text{Bias}(\hat{\theta}^n)\|_\infty$  will be larger than  $c\lambda$  for any fixed  $0 < c < 1$ . Following theorem lower bounds the bias of the Lasso for a broad class of design matrices.

**Theorem 2.2.1.** *Consider linear model (1.1.6) and suppose that  $\mathbf{X}$  has independent subgaussian rows, with mean zero and subgaussian norm  $\|X_1\|_{\psi_2} = \kappa$ , for some constant  $\kappa \in (0, \infty)$ . Let  $\hat{\theta}^n$  be the Lasso estimator, with  $\lambda = c\sigma\sqrt{(\log p)/n}$ . Then, there exist positive constants  $c_*, c_{**}$  such that if  $n \geq (3c_{**}s_0/c)^2 \log p$  and  $p \geq 13^{48/(c^2-48)}$ , then the following holds true: there exist a set of design matrices  $\mathcal{B}_n \subseteq \mathbb{R}^{n \times p}$ , and coefficient vectors  $\theta_0 \in \mathbb{R}^p$ ,  $\|\theta_0\|_0 \leq s_0$ , such that*

$$\mathbf{X} \in \mathcal{B}_n \Rightarrow \|\text{Bias}(\hat{\theta}^n)\|_\infty \geq \frac{\lambda}{3} = \frac{c\sigma}{3} \sqrt{\frac{\log p}{n}}, \quad (2.2.2)$$

$$\mathbb{P}(\mathcal{B}_n) \geq 1 - 4e^{-c_1 n} - 2p^{-3}, \quad (2.2.3)$$

where  $c_1 = 1/(4c_*\kappa^4)$ .

A formal proof of this theorem is given in Section 6.1.4.

Bias of the Lasso estimator is a major difficulty in achieving our goal, namely constructing confidence intervals for single parameters  $\theta_{0,i}$ . To overcome this challenge, we construct a debiased estimator  $\hat{\theta}^u$  from the Lasso solution, and characterize its limiting distribution. Armed with such characterization, we construct confidence intervals for the true parameters.

## 2.3 Compensating the bias of the Lasso

In this section we present our construction of a debiased estimator  $\widehat{\theta}^u$  and compare its bias with the bias of the Lasso estimator.

### 2.3.1 A debiased estimator for $\theta_0$

We begin with the classical case  $n > p$ . Here, the least squares estimator is given by  $\widehat{\theta}^{\text{OLS}} = (\widehat{\Sigma}^{-1}\mathbf{X}^\top Y/n)$ , where  $\widehat{\Sigma} \equiv (\mathbf{X}\mathbf{X}^\top/n)$  denotes the sample covariance. Further,  $\widehat{\theta}^{\text{OLS}}$  is unbiased estimator for  $\theta_0$ .

In the high-dimensional regime, however,  $\widehat{\Sigma}$  is not invertible since  $n < p$  and therefore  $\widehat{\theta}^{\text{OLS}}$  cannot even be defined. Instead, we define a similar estimator  $\tilde{\theta}$ , using a matrix  $M \in \mathbb{R}^{p \times p}$  in lieu of  $\widehat{\Sigma}^{-1}$ . (Take  $M$  to be an arbitrary matrix for now. We will discuss a specific choice of  $M$  later.)

Define

$$\tilde{\theta} \equiv \frac{1}{n}M\mathbf{X}^\top Y. \quad (2.3.1)$$

Plugging  $Y = \mathbf{X}\theta_0 + W$  in (2.3.1), we get

$$\begin{aligned} \tilde{\theta} &= \frac{1}{n}M\mathbf{X}^\top\mathbf{X}\theta_0 + \frac{1}{n}M\mathbf{X}^\top W \\ &= \theta_0 + \underbrace{(M\widehat{\Sigma} - \mathbf{I})\theta_0}_{\text{Bias}} + \underbrace{\frac{1}{n}M\mathbf{X}^\top W}_{\text{zero-mean noise}}. \end{aligned}$$

Note that  $\text{Bias}(\tilde{\theta})$  depends on the unknown vector of parameters  $\theta_0$ . In order to debias the estimator, we use Lasso solution  $\widehat{\theta}^n$  as an estimate of  $\theta_0$ . This leads us to a new estimator  $\widehat{\theta}^*$  defined as follows.

$$\begin{aligned} \widehat{\theta}^*(Y, \mathbf{X}; M, \lambda) &\equiv \tilde{\theta} - (M\widehat{\Sigma} - \mathbf{I})\widehat{\theta}^n(\lambda) \\ &= \frac{1}{n}M\mathbf{X}^\top Y - (M\widehat{\Sigma} - \mathbf{I})\widehat{\theta}^n(\lambda) \\ &= \widehat{\theta}^n(\lambda) + \frac{1}{n}M\mathbf{X}^\top(Y - \mathbf{X}\widehat{\theta}^n(\lambda)). \end{aligned} \quad (2.3.2)$$

For notational simplicity, we shall omit the arguments  $Y, \mathbf{X}, M, \lambda$  unless they are required for clarity.

### 2.3.1.1 How to choose $M$ ?

The quality of this debiasing procedure depends of course on the choice of  $M$ , as well as on the design  $\mathbf{X}$ . We characterize the pair  $(\mathbf{X}, M)$  by the following figure of merit.

**Definition 2.3.1.** *Given the pair  $\mathbf{X} \in \mathbb{R}^{n \times p}$  and  $M \in \mathbb{R}^{p \times p}$ , let  $\widehat{\Sigma} = \mathbf{X}^\top \mathbf{X}/n$  denote the associated sample covariance. Then, the generalized coherence parameter of  $\mathbf{X}, M$ , denoted by  $\mu_*(\mathbf{X}; M)$ , is*

$$\mu_*(\mathbf{X}; M) \equiv |M\widehat{\Sigma} - \mathbf{I}|_\infty. \quad (2.3.3)$$

Here and below  $|\cdot|_\infty$  denotes the entrywise  $\ell_\infty$  norm. The minimum (generalized) coherence of  $\mathbf{X}$  is  $\mu_{\min}(\mathbf{X}) = \min_{M \in \mathbb{R}^{p \times p}} \mu_*(\mathbf{X}; M)$ . We denote by  $M_{\min}(\mathbf{X})$  any minimizer of  $\mu_*(\mathbf{X}; M)$ .

Note that the minimum coherence can be computed efficiently since  $M \mapsto \mu_*(\mathbf{X}; M)$  is a convex function (even more, the optimization problem is a linear program).

The motivation for our terminology can be grasped by considering the following special case.

**Remark 2.3.2.** *Assume that the columns of  $\mathbf{X}$  are normalized to have  $\ell_2$  norm equal to  $\sqrt{n}$  (i.e.  $\|\mathbf{X}e_i\|_2 = \sqrt{n}$  for all  $i \in [p]$ ), and  $M = \mathbf{I}$ . Then  $(M\widehat{\Sigma} - \mathbf{I})_{i,i} = 0$ , and the maximum  $|M\widehat{\Sigma} - \mathbf{I}|_\infty = \max_{i \neq j} |(\widehat{\Sigma})_{ij}|$ . In other words  $\mu(\mathbf{X}; \mathbf{I})$  is the maximum normalized scalar product between distinct columns of  $\mathbf{X}$ :*

$$\mu_*(\mathbf{X}; \mathbf{I}) = \frac{1}{n} \max_{i \neq j} |\langle \mathbf{X}e_i, \mathbf{X}e_j \rangle|. \quad (2.3.4)$$

The quantity (2.3.4) is known as the *coherence parameter* of the matrix  $\mathbf{X}/\sqrt{n}$  and was first defined in the context of approximation theory by Mallat and Zhang [92], and by Donoho and Huo [39].

Assuming, for simplicity, that the columns of  $\mathbf{X}$  are normalized so that  $\|\mathbf{X}e_i\|_2 = \sqrt{n}$ , a small value of the coherence parameter  $\mu_*(\mathbf{X}; \mathbf{I})$  means that the columns of  $\mathbf{X}$  are roughly orthogonal. We emphasize however that  $\mu_*(\mathbf{X}; M)$  can be much smaller than its classical coherence parameter  $\mu_*(\mathbf{X}; \mathbf{I})$ . For instance,  $\mu_*(\mathbf{X}; \mathbf{I}) = 0$  if and only if  $\mathbf{X}/\sqrt{n}$  is an orthogonal matrix. On the other hand,  $\mu_{\min}(\mathbf{X}) = 0$  if and only if  $\mathbf{X}$  has rank  $p$ .<sup>2</sup>

<sup>2</sup>Of course this example requires  $n \geq p$ . It is the simplest example that illustrates the difference between coherence and generalized coherence, and it is not hard to find related examples with  $n < p$ .

The following theorem is a slight generalization of a result of [134] and its proof is given in Section 6.1.1. Let us emphasize that it applies to deterministic design matrices  $\mathbf{X}$ .

**Theorem 2.3.3.** *Let  $\mathbf{X} \in \mathbb{R}^{n \times p}$  be any (deterministic) design matrix, and  $\hat{\theta}^* = \hat{\theta}^*(Y, \mathbf{X}; M, \lambda)$  be a general debiased estimator as per Eq. (2.3.2). Then, setting  $Z = M\mathbf{X}^\top W/\sqrt{n}$ , we have*

$$\sqrt{n}(\hat{\theta}^* - \theta_0) = Z + \Delta, \quad Z \sim \mathbf{N}(0, \sigma^2 M \hat{\Sigma} M^\top), \quad \Delta = \sqrt{n}(M \hat{\Sigma} - \mathbf{I})(\theta_0 - \hat{\theta}^n). \quad (2.3.5)$$

Further, assume that  $\mathbf{X}$  satisfies the compatibility condition for the set  $S = \text{supp}(\theta_0)$ ,  $|S| \leq s_0$ , with constant  $\phi_0$ , and has generalized coherence parameter  $\mu_* = \mu_*(\mathbf{X}; M)$ , and let  $K \equiv \max_{i \in [p]} \hat{\Sigma}_{i,i}$ . Then, letting  $\lambda = \sigma \sqrt{(c^2 \log p)/n}$ , we have

$$\mathbb{P}\left(\|\Delta\|_\infty \geq \frac{4c\mu_*\sigma s_0}{\phi_0^2} \sqrt{\log p}\right) \leq 2p^{-c_0}, \quad c_0 = \frac{c^2}{32K} - 1. \quad (2.3.6)$$

Further, if  $M = M_{\min}(\mathbf{X})$  minimizes the convex cost function  $|M \hat{\Sigma} - \mathbf{I}|_\infty$ , then  $\mu_*$  can be replaced by  $\mu_{\min}(\mathbf{X})$  in Eq. (2.3.6).

The above theorem decomposes the estimation error  $(\hat{\theta}^* - \theta_0)$  into a zero mean Gaussian term  $Z/\sqrt{n}$  and a bias term  $\Delta/\sqrt{n}$  whose maximum entry is bounded as per Eq. (2.3.6). This estimate on  $\|\Delta\|_\infty$  depends on the design matrix through two constants: the compatibility constant  $\phi_0$  and the generalized coherence parameter  $\mu_*(\mathbf{X}; M)$ . The former is a well studied property of the design matrix [19, 133], and assuming  $\phi_0$  of order one is nearly necessary for the Lasso to achieve optimal estimation rate in high dimension. On the contrary, the definition of  $\mu_*(\mathbf{X}; M)$  is a new contribution of the present work.

The next theorem establishes that, for a natural probabilistic model of the design matrix  $\mathbf{X}$ , both  $\phi_0$  and  $\mu_*(\mathbf{X}; M)$  can be bounded with probability converging rapidly to one as  $n, p \rightarrow \infty$ .

**Theorem 2.3.4.** *Let  $\Sigma \in \mathbb{R}^{p \times p}$  be such that  $\sigma_{\min}(\Sigma) \geq C_{\min} > 0$ , and  $\sigma_{\max}(\Sigma) \leq C_{\max} < \infty$ , and  $\max_{i \in [p]} \Sigma_{ii} \leq 1$ . Assume  $\mathbf{X}\Sigma^{-1/2}$  to have independent subgaussian rows, with zero mean and subgaussian norm  $\|\Sigma^{-1/2} X_1\|_{\psi_2} = \kappa$ , for some constant  $\kappa \in (0, \infty)$ .*

- (a) *For  $\phi_0, s_0, K \in \mathbb{R}_{>0}$ , let  $\mathcal{E}_n = \mathcal{E}_n(\phi_0, s_0, K)$  be the event that the compatibility condition holds for  $\hat{\Sigma} = (\mathbf{X}^\top \mathbf{X}/n)$ , for all sets  $S \subseteq [p]$ ,  $|S| \leq s_0$  with constant  $\phi_0 > 0$ , and that*

$\max_{i \in [p]} \widehat{\Sigma}_{i,i} \leq K$ . Explicitly

$$\mathcal{E}_n(\phi_0, s_0, K) \equiv \left\{ \mathbf{X} \in \mathbb{R}^{n \times p} : \min_{S: |S| \leq s_0} \phi(\widehat{\Sigma}, S) \geq \phi_0, \max_{i \in [p]} \widehat{\Sigma}_{i,i} \leq K, \widehat{\Sigma} = (\mathbf{X}^\top \mathbf{X}/n) \right\}. \quad (2.3.7)$$

Then there exists  $c_* \leq 2000$  such that the following happens. If  $n \geq \nu_0 s_0 \log(p/s_0)$ ,  $\nu_0 \equiv 5 \times 10^4 c_* (C_{\max}/C_{\min})^2 \kappa^4$ ,  $\phi_0 = \sqrt{C_{\min}}/2$ , and  $K \geq 1 + 20\kappa^2 \sqrt{(\log p)/n}$ , then

$$\mathbb{P}(\mathbf{X} \in \mathcal{E}_n(\phi_0, s_0, K)) \geq 1 - 4e^{-c_1 n}, \quad c_1 \equiv \frac{1}{4c_* \kappa^4}. \quad (2.3.8)$$

(b) For  $a > 0$ , let  $\mathcal{G}_n = \mathcal{G}_n(a)$  be the following event:

$$\mathcal{G}_n(a) \equiv \left\{ \mathbf{X} \in \mathbb{R}^{n \times p} : \mu_{\min}(\mathbf{X}) < a \sqrt{\frac{\log p}{n}} \right\}. \quad (2.3.9)$$

Then, for  $n \geq a^2 C_{\min} \log p / (4e^2 C_{\max} \kappa^4)$

$$\mathbb{P}(\mathbf{X} \in \mathcal{G}_n(a)) \geq 1 - 2p^{-c_2}, \quad c_2 \equiv \frac{a^2 C_{\min}}{24e^2 \kappa^4 C_{\max}} - 2. \quad (2.3.10)$$

The proof of this theorem is given in Section 6.1.2 (for part (a)) and Section 6.1.3 (part (b)).

The proof that event  $\mathcal{E}_n$  holds with high probability relies crucially on a theorem by Rudelson and Zhou [116, Theorem 6]. Simplifying somewhat, the latter states that, if the restricted eigenvalue condition of [14] holds for the population covariance  $\Sigma$ , then it holds with high probability for the sample covariance  $\widehat{\Sigma}$ . (Recall that the restricted eigenvalue condition is implied by a lower bound on the minimum singular value<sup>3</sup>, and that it implies the compatibility condition [133].)

Motivated by Theorem 2.3.3, we propose a procedure for constructing matrix  $M$ . Our approach is to construct  $M$  by solving a convex program that aims at optimizing two objectives. On one hand, we try to control  $|M\widehat{\Sigma} - \mathbf{I}|_\infty$  which, as shown in Theorem 2.3.3, controls the non-Gaussianity and bias of the estimator. On the other, we minimize  $[M\widehat{\Sigma}M]_{i,i}$ , for each  $i \in [p]$ , which controls the variance of the  $i$ -th coordinate of the estimator. The procedure and the corresponding debiased estimator  $\widehat{\theta}^u$  are described in Algorithm 1. Note that

<sup>3</sup>Note, in particular, at the cost of further complicating the last statement, the condition  $\sigma_{\min}(\Sigma) = \Omega(1)$  can be further weakened.

---

**Algorithm 1** Unbiased estimator for  $\theta_0$  in high-dimensional linear regression models

---

**Input:** Measurement vector  $Y$ , design matrix  $\mathbf{X}$ , parameters  $\lambda, \mu$ .

**Output:** Unbiased estimator  $\hat{\theta}^u$ .

- 1: Let  $\hat{\theta}^n = \hat{\theta}^n(Y, \mathbf{X}; \lambda)$  be the Lasso estimator as per Eq. (1.1.7).
- 2: Set  $\hat{\Sigma} \equiv (\mathbf{X}^\top \mathbf{X})/n$ .
- 3: **for**  $i = 1, 2, \dots, p$  **do**
- 4:   Let  $m_i$  be a solution of the convex program:

$$\begin{aligned} & \text{minimize} && m^\top \hat{\Sigma} m \\ & \text{subject to} && \|\hat{\Sigma} m - e_i\|_\infty \leq \mu, \end{aligned} \tag{2.3.11}$$

where  $e_i \in \mathbb{R}^p$  is the vector with one at the  $i$ -th position and zero everywhere else.

- 5: Set  $M = (m_1, \dots, m_p)^\top$ . If any of the above problems is not feasible, then set  $M = \mathbf{I}_{p \times p}$ .
- 6: Define the estimator  $\hat{\theta}^u$  as follows:

$$\hat{\theta}^u = \hat{\theta}^n(\lambda) + \frac{1}{n} M \mathbf{X}^\top (Y - \mathbf{X} \hat{\theta}^n(\lambda)) \tag{2.3.12}$$


---

event  $\mathcal{G}_n$ , defined by (2.3.9), can be equivalently stated as the event that problem (2.3.11), for all  $i \in [p]$ , is feasible with  $\mu = a\sqrt{(\log p)/n}$ .

Finally, by putting together Theorem 2.3.3 and Theorem 2.3.4, we obtain the following conclusion. We refer to Section 6.1.5 for the proof of Theorem 2.3.5.

**Theorem 2.3.5.** *Consider the linear model (1.1.6) and let  $\hat{\theta}^u$  be defined as per Eq. (2.3.12) in Algorithm 1, with  $\mu = a\sqrt{(\log p)/n}$ . Then, setting  $Z = M \mathbf{X}^\top W / \sqrt{n}$ , we have*

$$\sqrt{n}(\hat{\theta}^u - \theta_0) = Z + \Delta, \quad Z | \mathbf{X} \sim \mathbf{N}(0, \sigma^2 M \hat{\Sigma} M^\top), \quad \Delta = \sqrt{n}(M \hat{\Sigma} - \mathbf{I})(\theta_0 - \hat{\theta}^n). \tag{2.3.13}$$

Further, under the assumptions of Theorem 2.3.4, and for  $n \geq \max(\nu_0 s_0 \log(p/s_0), \nu_1 \log p)$ ,  $\nu_1 = \max(1600\kappa^4, a^2/(4e^2\kappa^4))$ , and  $\lambda = \sigma\sqrt{(c^2 \log p)/n}$ , we have

$$\mathbb{P} \left\{ \|\Delta\|_\infty \geq \left( \frac{16ac\sigma}{C_{\min}} \right) \frac{s_0 \log p}{\sqrt{n}} \right\} \leq 4e^{-c_1 n} + 4p^{-\tilde{c}_0 \wedge c_2}. \tag{2.3.14}$$

where  $\tilde{c}_0 = (c^2/48) - 1$  and  $c_1, c_2$  are given by Eqs. (2.3.8) and (2.3.10).

Finally, the tail bound (2.3.14) holds for any choice of  $M$  that is only function of the design matrix  $\mathbf{X}$ , and satisfies the feasibility condition in Eq. (2.3.11), i.e.  $\|M \hat{\Sigma} - \mathbf{I}\|_\infty \leq \mu$ .

Assuming  $\sigma, C_{\min}$  of order one, the last theorem establishes that, for random designs,

the maximum size of the ‘bias term’  $\Delta_i$  over  $i \in [p]$  is:

$$\|\Delta\|_\infty = O\left(\frac{s_0 \log p}{\sqrt{n}}\right) \quad (2.3.15)$$

On the other hand, the ‘noise term’  $Z_i$  is roughly of order  $\sqrt{[M\widehat{\Sigma}M^\top]_{ii}}$ . Bounds on the variances  $[M\widehat{\Sigma}M^\top]_{ii}$  will be given in next Chapter (cf. Eq. (6.2.1) in the Proof of Theorem 3.1.1) showing that, if  $M$  is computed through Algorithm 1,  $[M\widehat{\Sigma}M^\top]_{ii}$  is of order one for a broad family of random designs. As a consequence  $|\Delta_i|$  is much smaller than  $|Z_i|$  whenever  $s_0 = o(\sqrt{n}/\log p)$ . We summarize these remarks below.

**Remark 2.3.6.** *Theorem 2.3.5 only requires that the support size satisfies  $s_0 = O(n/\log p)$ . If we further assume  $s_0 = o(\sqrt{n}/\log p)$ , then we have  $\|\Delta\|_\infty = o(1)$  with high probability. Hence,  $\widehat{\theta}^u$  is an asymptotically unbiased estimator for  $\theta_0$ .*

A more formal comparison of the bias of  $\widehat{\theta}^u$ , and of the one of the Lasso estimator  $\widehat{\theta}^n$  can be found in Section 2.3.2 below. Section 2.3.3 compares our approach with other related work.

As it can be seen from the statement of Theorem 2.3.3 and Theorem 2.3.4, the claim of Theorem 2.3.5 does not rely on the specific choice of the objective function in optimization problem (2.3.11) and only uses the constraint on  $\|\widehat{\Sigma}m - e_i\|_\infty$ . In particular it holds for any matrix  $M$  that is feasible. On the other hand, the specific objective function problem (2.3.11) minimizes the variance of the noise term,  $\text{Var}(Z_i)$ .

### 2.3.2 Discussion: bias reduction

Theorems 2.3.3 and 2.3.4 provide a quantitative framework to discuss in what sense the Lasso estimator  $\widehat{\theta}^n$  is asymptotically biased, while the debiased estimator  $\widehat{\theta}^u$  is asymptotically unbiased.

Invoking definition of bias (2.1.2), it follows from Eq. (2.3.5) that

$$\text{Bias}(\widehat{\theta}^u) = \frac{1}{\sqrt{n}} \mathbb{E}\{\Delta | \mathbf{X}\}. \quad (2.3.16)$$

Applying Theorem 2.3.5 with high probability,  $\|\Delta\|_\infty = O(s_0 \log p / \sqrt{n})$ . The next corollary states that this bound on  $\|\Delta\|_\infty$  translates into a bound on  $\text{Bias}(\widehat{\theta}^u)$  for all  $\mathbf{X}$  in a set that has probability rapidly converging to one as  $n, p$  get large.

**Corollary 2.3.7.** *Under the assumptions of Theorem 2.3.5, let  $c_1, c_2$  be defined as per equations (2.3.8) and (2.3.10). Then we have*

$$\mathbf{X} \in \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a) \Rightarrow \|\text{Bias}(\hat{\theta}^u)\|_{\infty} \leq \frac{160a}{C_{\min}} \cdot \frac{\sigma s_0 \log p}{n}, \quad (2.3.17)$$

$$\mathbb{P}\left(\mathbf{X} \in \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)\right) \geq 1 - 4e^{-c_1 n} - 2p^{-c_2}. \quad (2.3.18)$$

The proof of this corollary can be found in Section 6.1.6.

This result can be contrasted with a converse result for the Lasso estimator. Namely, as stated in Theorem 2.2.1, there are choices of the vector  $\theta_0$ , and of the design covariance  $\Sigma$ , such that  $\|\text{Bias}(\hat{\theta}^n)\|_{\infty}$  is of order  $\sigma\sqrt{(\log p)/n}$ . Comparing this bound with the bound on  $\|\text{Bias}(\hat{\theta}^u)\|_{\infty}$ , as established in Corollary 2.3.7, we conclude that if  $s_0$  is significantly smaller than  $\sqrt{n/\log p}$  (which is the main regime studied in this chapter), then  $\|\text{Bias}(\hat{\theta}^n)\|_{\infty} \gg \|\text{Bias}(\hat{\theta}^u)\|_{\infty}$ . Hence, our approach has effectively reduced the bias.

### 2.3.3 Comparison with earlier results

The idea of constructing a debiased estimator of the form  $\hat{\theta}^u = \hat{\theta}^n + (1/n)M\mathbf{X}^{\top}(Y - \mathbf{X}\hat{\theta}^n)$  was used by Javanmard and Montanari in [72], that suggested the choice  $M = c\Sigma^{-1}$ , with  $\Sigma = \mathbb{E}(X_1 X_1^{\top})$  the population covariance matrix and  $c$  a positive constant. We discuss this approach meticulously in Chapter 4. Van de Geer, Bühlmann, Ritov and Dezeure [134] used the same construction with  $M$  an estimate of  $\Sigma^{-1}$  which is appropriate for sparse inverse covariances. These authors prove semi-parametric optimality in a non-asymptotic setting, provided the sample size is at least  $n = \Omega((s_0 \log p)^2)$ .

The rationale for choosing  $M$  to be an estimate of  $\Sigma^{-1}$  is as follows.

**One step of Newton method.** Recall that Lasso estimator is given as the solution of the optimization problem below:

$$\hat{\theta}^n \equiv \underset{\theta \in \mathbb{R}^p}{\text{argmin}} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|^2 + \lambda \|\theta\|_1 \right\}.$$

Lasso solution  $\hat{\theta}^n$  is biased due to the regularization term. On the other hand, it is easy to see that

$$\underset{\theta \in \mathbb{R}^p}{\text{argmin}} \frac{1}{2n} \mathbb{E}(\|Y - \mathbf{X}\theta\|^2) = \theta_0, \quad (2.3.19)$$

where the expectation is taken with respect to the law of the samples  $\{(Y_i, X_i)\}_{i=1}^n$ . Therefore, a natural approach for debiasing is to take  $\hat{\theta}^n$  as the initial guess and perform one step of Newton method to move towards the minimizer of (2.3.19), i.e.,  $\theta_0$ . This leads to

$$\hat{\theta}^*(Y, \mathbf{X}; \Sigma^{-1}, \lambda) = \hat{\theta}^n + \frac{1}{n} \Sigma^{-1} \mathbf{X}^\top (Y - \mathbf{X} \hat{\theta}^n).$$

This justifies the choice of  $M = \Sigma^{-1}$ .

From a technical point of view, the presented approach in this chapter starts from a simple decomposition of the debiased estimator  $\hat{\theta}^u$  into a Gaussian part and an error term, already used in [134]. However, departing radically from earlier work, we realize that  $M$  need not be a good estimator of  $\Sigma^{-1}$  in order for the debiasing procedure to work. We instead set  $M$  as to minimize the error term and the variance of the Gaussian term. As a consequence of this choice, our approach applies to general covariance structures  $\Sigma$ . By contrast, earlier approaches applied only to sparse  $\Sigma$ , as in [72], or sparse  $\Sigma^{-1}$  as in [134]. The only assumptions we make on  $\Sigma$  are the standard compatibility conditions required for high-dimensional consistency [19].

In the case of linear statistical models considered here, the authors of [134] construct a debiased estimator of the form (2.3.2). However, instead of solving the optimization problem (2.3.11), they follow [149] and use the regression coefficients of the  $i$ -th column of  $\mathbf{X}$  on the other columns to construct the  $i$ -th row of  $M$ . These regression coefficients are computed, once again, using the Lasso (node-wise Lasso). It useful to spell out the most important differences between our contribution and the ones of [134]:

1. The case of fixed non-random designs is covered by [134, Theorem 2.1], which should be compared to our Theorem 2.3.3. While in our case the bias is controlled by the generalized coherence parameter, a similar role is played in [134] by the regularization parameters of the nodewise Lasso.
2. The case of random designs is covered by [134, Theorem 2.2, Theorem 2.4], which should be compared with our Theorem 2.3.5. In this case, the assumptions underlying our result are less restrictive. More precisely:
  - (a) [134, Theorem 2.2, Theorem 2.4] assume  $\mathbf{X}$  has i.i.d. rows, while we only assume the rows are independent.
  - (b) [134, Theorem 2.2, Theorem 2.4] assumes the rows of the inverse covariance

matrix  $\Sigma^{-1}$  are sparse. More precisely, letting  $s_j$  be the number of non-zero entries of the  $j$ -th row of  $\Sigma^{-1}$ , [134] assumes  $\max_{j \in [p]} s_j = o(n/\log p)$ , that is much smaller than  $p$ . We do not make any sparsity assumption on  $\Sigma^{-1}$ , and  $s_j$  can be as large as  $p$ . [134, Theorem 2.4] also considers a slightly different setting, where  $\mathbf{X}$  has bounded entries, under analogous sparsity assumptions.

It is currently unknown whether the sparsity assumption in [134] is required for that approach to work, or it is rather an artifact of the specific analysis. Indeed [134, Theorem 2.1] can in principle be used to weaken this condition.

In addition, our Theorem 2.3.5 provides the specific dependence on the maximum and minimum singular value of  $\widehat{\Sigma}$ .

Note that solving the convex problem (2.3.11) is not more burdensome than solving the nodewise Lasso as in [149, 134], This can be confirmed by checking that the dual of problem (2.3.11) is an  $\ell_1$ -regularized quadratic optimization problem. It has therefore the same complexity as the nodewise Lasso (but it is different from the nodewise Lasso). We refer to Section 5.1 for a thorough discussion on this.

## 2.4 Confidence intervals

A direct application of Theorem 2.3.5 is to derive confidence intervals for each single parameter  $\theta_{0,i}$  of a high-dimensional model. Throughout, we make the sparsity assumption  $s_0 = o(\sqrt{n}/\log p)$  and omit explicit constants that can be readily derived from Theorem 2.3.5.

### 2.4.1 Preliminary lemmas

As discussed above, the bias term  $\Delta$  is negligible with respect to the random term  $Z$  in the decomposition (2.3.13), provided the latter has variance of order one. Our first lemma establishes that this is indeed the case.

**Lemma 2.4.1.** *Let  $M = (m_1, \dots, m_p)^\top$  be the matrix with rows  $m_i^\top$  obtained by solving convex program (2.3.11) in Algorithm 1. Then for all  $i \in [p]$ ,*

$$[M\widehat{\Sigma}M^\top]_{i,i} \geq \frac{(1-\mu)^2}{\widehat{\Sigma}_{i,i}}.$$

Lemma 2.4.1 is proved in Appendix A.1.

Using this fact, we can then characterize the asymptotic distribution of the residuals  $(\hat{\theta}^u - \theta_{0,i})$ . Theorem 2.3.5 naturally suggests to consider the scaled residual  $\sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})/(\sigma[M\hat{\Sigma}M^\top]_{i,i}^{1/2})$ . In the next lemma we consider a slightly more general scaling, replacing  $\sigma$  by a consistent estimator  $\hat{\sigma}$ .

**Lemma 2.4.2.** *Consider a sequence of design matrices  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , with dimensions  $n \rightarrow \infty$ ,  $p = p(n) \rightarrow \infty$  satisfying the following assumptions, for constants  $C_{\min}, C_{\max}, \kappa \in (0, \infty)$  independent of  $n$ . For each  $n$ ,  $\Sigma \in \mathbb{R}^{p \times p}$  is such that  $\sigma_{\min}(\Sigma) \geq C_{\min} > 0$ , and  $\sigma_{\max}(\Sigma) \leq C_{\max} < \infty$ , and  $\max_{i \in [p]} \Sigma_{ii} \leq 1$ . Assume  $\mathbf{X}\Sigma^{-1/2}$  to have independent subgaussian rows, with zero mean and subgaussian norm  $\|\Sigma^{-1/2}X_1\|_{\psi_2} \leq \kappa$ ,*

*Consider the linear model (1.1.6) and let  $\hat{\theta}^u$  be defined as per Eq. (2.3.12) in Algorithm 1, with  $\mu = a\sqrt{(\log p)/n}$  and  $\lambda = \sigma\sqrt{(c^2 \log p)/n}$ , with  $a, c$  large enough constants. Finally, let  $\hat{\sigma} = \hat{\sigma}(y, \mathbf{X})$  be an estimator of the noise level satisfying, for any  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \sup_{\theta_0 \in \mathbb{R}^p; \|\theta_0\|_0 \leq s_0} \mathbb{P}\left(\left|\frac{\hat{\sigma}}{\sigma} - 1\right| \geq \epsilon\right) = 0. \quad (2.4.1)$$

If  $s_0 = o(\sqrt{n}/\log p)$  ( $s_0 \geq 1$ ), then, for all  $x \in \mathbb{R}$ , we have

$$\lim_{n \rightarrow \infty} \sup_{\theta_0 \in \mathbb{R}^p; \|\theta_0\|_0 \leq s_0} \left| \mathbb{P}\left\{ \frac{\sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})}{\hat{\sigma}[M\hat{\Sigma}M^\top]_{i,i}^{1/2}} \leq x \right\} - \Phi(x) \right| = 0. \quad (2.4.2)$$

The proof of this lemma can be found in Section 6.1.7. We also note that the dependence of  $a$  and  $c$  on  $C_{\min}, C_{\max}, \kappa$  can be easily reconstructed from Theorem 2.3.4.

The last lemma requires a consistent estimator of  $\sigma$ , in the sense of Eq. (2.4.1). Several proposals have been made to estimate the noise level in high-dimensional linear regression. A short list of references includes [50, 51, 122, 148, 124, 11, 49, 112, 32, 52, 5]. Consistency results have been proved for several of these estimators.

In order to demonstrate that the consistency criterion (2.4.1) can be achieved, we use the scaled Lasso [124] given by

$$\{\hat{\theta}^n(\tilde{\lambda}), \hat{\sigma}(\tilde{\lambda})\} \equiv \arg \min_{\theta \in \mathbb{R}^p, \sigma > 0} \left\{ \frac{1}{2\sigma n} \|Y - \mathbf{X}\theta\|_2^2 + \frac{\sigma}{2} + \tilde{\lambda} \|\theta\|_1 \right\}. \quad (2.4.3)$$

This is a joint convex optimization problem which provides an estimate of the noise level in addition to an estimate of  $\theta_0$ .

The following lemma uses the analysis of [124] to show that  $\hat{\sigma}$  satisfies the consistency criterion (2.4.1).

**Lemma 2.4.3.** *Under the assumptions of Lemma 2.4.2, let  $\hat{\sigma} = \hat{\sigma}(\tilde{\lambda})$  be the scaled Lasso estimator of the noise level, see (2.4.3), with  $\tilde{\lambda} = 10\sqrt{(2\log p)/n}$ . Then  $\hat{\sigma}$  satisfies (2.4.1).*

The proof of this lemma is fairly straightforward and can be found in Appendix A.2.

In view of Lemma 2.4.2, it is quite straightforward to construct asymptotically valid confidence intervals. Namely, for  $i \in [p]$  and significance level  $\alpha \in (0, 1)$ , we let

$$\begin{aligned} J_i(\alpha) &\equiv [\hat{\theta}_i^u - \delta(\alpha, n), \hat{\theta}_i^u + \delta(\alpha, n)], \\ \delta(\alpha, n) &\equiv \Phi^{-1}(1 - \alpha/2) \frac{\hat{\sigma}}{\sqrt{n}} [M\hat{\Sigma}M^\top]_{i,i}^{1/2}. \end{aligned} \tag{2.4.4}$$

**Theorem 2.4.4.** *Consider a sequence of design matrices  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , with dimensions  $n \rightarrow \infty$ ,  $p = p(n) \rightarrow \infty$  satisfying the assumptions of Lemma 2.4.2.*

*Consider the linear model (1.1.6) and let  $\hat{\theta}^u$  be defined as per Eq. (2.3.12) in Algorithm 1, with  $\mu = a\sqrt{(\log p)/n}$  and  $\lambda = \sigma\sqrt{(c^2 \log p)/n}$ , with  $a, c$  large enough constants. Finally, let  $\hat{\sigma} = \hat{\sigma}(y, \mathbf{X})$  a consistent estimator of the noise level in the sense of Eq. (2.4.1). Then the confidence interval  $J_i(\alpha)$  is asymptotically valid, namely*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\theta_{0,i} \in J_i(\alpha)) = 1 - \alpha. \tag{2.4.5}$$

*Proof:* The proof is an immediate consequence of Lemma 2.4.2 since

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(\theta_{0,i} \in J_i(\alpha)) &= \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{\sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})}{\hat{\sigma}[M\hat{\Sigma}M^\top]_{i,i}^{1/2}} \leq \Phi^{-1}(1 - \alpha/2) \right\} \\ &\quad - \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{\sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})}{\hat{\sigma}[M\hat{\Sigma}M^\top]_{i,i}^{1/2}} \leq -\Phi^{-1}(1 - \alpha/2) \right\} \\ &= 1 - \alpha. \end{aligned} \tag{2.4.6}$$

## 2.4.2 Generalization to simultaneous confidence intervals

In many situations, it is necessary to perform statistical inference on more than one of the parameters simultaneously. For instance, we might be interested in performing inference about  $\theta_{0,R} \equiv (\theta_{0,i})_{i \in R}$  for some set  $R \subseteq [p]$ .

The simplest generalization of our method is to the case in which  $|R|$  stays finite as  $n, p \rightarrow \infty$ . In this case we have the following generalization of Lemma 2.4.2. (The proof is the same as for Lemma 2.4.2, and hence we omit it.)

**Lemma 2.4.5.** *Under the assumptions of Lemma 2.4.2, define*

$$Q^{(n)} \equiv \frac{\widehat{\sigma}^2}{n} [M\widehat{\Sigma}M^\top]. \quad (2.4.7)$$

Let  $R = R(n)$  be a sequence of sets  $R(n) \subseteq [p]$ , with  $|R(n)| = k$  fixed as  $n, p \rightarrow \infty$ , and further assume  $s_0 = o(\sqrt{n}/\log p)$ , with  $s_0 \geq 1$ . Then, for all  $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ , we have

$$\lim_{n \rightarrow \infty} \sup_{\theta_0 \in \mathbb{R}^p; \|\theta_0\|_0 \leq s_0} \left| \mathbb{P} \left\{ (Q_{R,R}^{(n)})^{-1/2} (\widehat{\theta}_R^u - \theta_{0,R}) \leq x \right\} - \Phi_k(x) \right| = 0, \quad (2.4.8)$$

where  $(a_1, \dots, a_k) \leq (b_1, \dots, b_k)$  indicates that  $a_1 \leq b_1, \dots, a_k \leq b_k$ , and  $\Phi_k(x) = \Phi(x_1) \times \dots \times \Phi(x_k)$ .

This lemma allows to construct confidence regions for low-dimensional projections of  $\theta_0$ , much in the same way as we used Lemma 2.4.2 to compute confidence intervals for one-dimensional projections in Section 2.4.

Explicitly, let  $\mathcal{C}_{k,\alpha} \subseteq \mathbb{R}^k$  be any Borel set such that  $\int_{\mathcal{C}_{k,\alpha}} \phi_k(x) dx \geq 1 - \alpha$ , where

$$\phi_k(x) = \frac{1}{(2\pi)^{k/2}} \exp\left(-\frac{\|x\|^2}{2}\right),$$

is the  $k$ -dimensional Gaussian density. Then, for  $R \subseteq [p]$ , we define  $J_R(\alpha) \subseteq \mathbb{R}^k$  as follows

$$J_R(\alpha) \equiv \widehat{\theta}_R^u + (Q_{R,R}^{(n)})^{1/2} \mathcal{C}_{k,\alpha}. \quad (2.4.9)$$

Then Lemma 2.4.5 implies (under the assumptions stated there) that  $J_R(\alpha)$  is a valid confidence region

$$\lim_{n \rightarrow \infty} \mathbb{P}(\theta_{0,R} \in J_R(\alpha)) = 1 - \alpha. \quad (2.4.10)$$

### 2.4.3 Non-Gaussian noise

As can be seen from the proof of Theorem 2.3.5,  $Z = M\mathbf{X}^\top W/\sqrt{n}$ , and since the noise is Gaussian, i.e.,  $W \sim \mathbf{N}(0, \sigma^2\mathbf{I})$ , we have  $Z|\mathbf{X} \sim \mathbf{N}(0, \sigma^2 M\widehat{\Sigma}M^\top)$ . We claim that the distribution of the coordinates of  $Z$  is asymptotically Gaussian, even if  $W$  is non-Gaussian, provided the definition of  $M$  is modified slightly. As a consequence, the definition of confidence intervals in Corollary 2.4.4 remains valid in this broader setting.

In case of non-Gaussian noise, we write

$$\begin{aligned} \frac{\sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})}{\sigma[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} &= \frac{1}{\sqrt{n}} \frac{m_i^\top \mathbf{X}^\top W}{\sigma[m_i^\top \widehat{\Sigma} m_i]^{1/2}} + o(1) \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{m_i^\top X_j W_j}{\sigma[m_i^\top \widehat{\Sigma} m_i]^{1/2}} + o(1). \end{aligned}$$

Conditional on  $\mathbf{X}$ , the summands  $\xi_j = m_i^\top X_j W_j / (\sigma[m_i^\top \widehat{\Sigma} m_i]^{1/2})$  are independent and zero mean. Further,  $\sum_{j=1}^n \mathbb{E}(\xi_j^2 | \mathbf{X}) = 1$ . Therefore, if Lindeberg condition holds, namely for every  $\epsilon > 0$ , almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E}(\xi_j^2 \mathbb{I}_{\{|\xi_j| > \epsilon\sqrt{n}\}} | \mathbf{X}) = 0, \quad (2.4.11)$$

then  $\sum_{j=1}^n \xi_j / \sqrt{n} | \mathbf{X} \xrightarrow{d} \mathbf{N}(0, 1)$ , from which we can build the valid confidence intervals as in (2.4.4).

In order to ensure that the Lindeberg condition holds, we modify the optimization problem (2.4.12) as follows:

$$\begin{aligned} &\text{minimize} && m^\top \widehat{\Sigma} m \\ &\text{subject to} && \|\widehat{\Sigma} m - e_i\|_\infty \leq \mu \\ &&& \|\mathbf{X} m\|_\infty \leq n^\beta \quad \text{for arbitrary fixed } 1/4 < \beta < 1/2 \end{aligned} \quad (2.4.12)$$

Next theorem shows the validity of the proposed  $p$ -values in the non-Gaussian noise setting.

**Theorem 2.4.6.** *Suppose that the noise variables  $W_i$  are independent with  $\mathbb{E}(W_i) = 0$ ,  $\mathbb{E}(W_i^2) = \sigma^2$ , and  $\mathbb{E}(|W_i|^{2+a}) \leq C \sigma^{2+a}$  for some  $a > 0$ .*

*Let  $M = (m_1, \dots, m_p)^\top$  be the matrix with rows  $m_i^\top$  obtained by solving optimization*

problem (2.4.12). Then under the assumptions of Theorem 2.3.5, and for sparsity level  $s_0 = o(\sqrt{n}/\log p)$ , an asymptotic two-sided confidence interval for  $\theta_{0,i}$  with significance  $\alpha$  is given by  $I_i = [\hat{\theta}_i^u - \delta(\alpha, n), \hat{\theta}_i^u + \delta(\alpha, n)]$  where

$$\delta(\alpha, n) = \Phi^{-1}(1 - \alpha/2) \hat{\sigma} n^{-1/2} \sqrt{[M\hat{\Sigma}M^\top]_{i,i}}. \quad (2.4.13)$$

Theorem 2.4.6 is proved in Section 6.1.8.

## 2.5 Comparison with Local Asymptotic Normality

Our approach is based on an asymptotic distributional characterization of a debiased estimator, cf. Theorem 2.3.5. As shown, debiased estimator  $\hat{\theta}^u$  is asymptotically normal under proper scaling of sample size. This is analogous to what happens in classical statistics, where *local asymptotic normality* (LAN) can be used to characterize an estimator distribution, and hence derive test statistics [87, 135].

This analogy is only superficial, and the mathematical phenomenon underlying Theorem 2.3.5 is altogether different from the one in local asymptotic normality. We refer to [8] for a more complete understanding, and only mention a few points:

1. LAN theory holds in the low-dimensional limit, where the number of parameters  $p$  is much smaller than the number of samples  $n$ . Even more, the focus is on  $p$  fixed, and  $n \rightarrow \infty$ .

In contrast, the Gaussian limit in Theorem 2.3.5 holds with  $p$  proportional to  $n$ .

2. Indeed, in the present case, the Lasso estimator (which is of course a special case of M-estimator)  $\hat{\theta}$  is *not normal*. Only the debiased estimator  $\hat{\theta}^u$  is asymptotically normal. Further, while LAN theory holds quite generally in the classical asymptotics, the present theory is more sensitive to the properties of the design matrix  $\mathbf{X}$ .

## 2.6 General regularized maximum likelihood

In this section, we generalize our results beyond the linear regression model to regularized maximum likelihood. For univariate  $y$ , and vector  $x \in \mathbb{R}^q$ , we let  $\{f_\theta(y|x)\}_{\theta \in \mathbb{R}^p}$  be a family of conditional probability densities parameterized by  $\theta$ , that are absolutely continuous with

respect to a common measure  $\omega(dy)$ , and suppose that the gradient  $\nabla_{\theta} f_{\theta}(y|x)$  exists and is square integrable.

As in for linear regression, we assume that the data is given by  $n$  i.i.d. pairs  $(X_1, Y_1), \dots, (X_n, Y_n)$ , where conditional on  $X_i$ , the response variable  $Y_i$  is distributed as

$$Y_i \sim f_{\theta_0}(\cdot | X_i).$$

for some parameter vector  $\theta_0 \in \mathbb{R}^p$ . We consider the following regularized estimator:

$$\hat{\theta} \equiv \arg \min_{\theta \in \mathbb{R}^p} \{ \ell(\theta) + \lambda \mathcal{R}(\theta) \}, \quad (2.6.1)$$

with  $\ell(\theta) = \sum_{i=1}^n \ell_i(\theta)/n$ , and  $\ell_i(\theta) = -\log f_{\theta}(Y_i | X_i)$  the normalized negative log-likelihood.

We next generalize the definition of  $\hat{\Sigma}$ . Let  $\mathcal{I}_i(\theta)$  be the Fisher information of  $f_{\theta}(Y | X_i)$ , defined as

$$\mathcal{I}_i(\theta) \equiv \mathbb{E} \left[ \left( \nabla_{\theta} \log f_{\theta}(Y | X_i) \right) \left( \nabla_{\theta} \log f_{\theta}(Y | X_i) \right)^{\top} \middle| X_i \right] = -\mathbb{E} \left[ \left( \nabla_{\theta}^2 \log f_{\theta}(Y | X_i, \theta) \right) \middle| X_i \right],$$

where expectation is taken with respect to the law  $f_{\theta_0}(\cdot | X_i)$ , and  $\nabla_{\theta}^2$  denotes the Hessian operator. The second identity here holds under suitable regularity conditions [88].

We assume  $\mathbb{E}[\mathcal{I}_i(\theta)] \succ 0$ , with expectation taken with respect to the law of  $\{X_i\}_{i=1}^n$ , and define  $\hat{\Sigma} \in \mathbb{R}^{p \times p}$  as follows:

$$\hat{\Sigma} \equiv \frac{1}{n} \sum_{i=1}^n \mathcal{I}_i(\hat{\theta}). \quad (2.6.2)$$

Note that (in general)  $\hat{\Sigma}$  depends on  $\hat{\theta}$ . Finally, the debiased estimator  $\hat{\theta}^u$  is defined by

$$\hat{\theta}^u \equiv \hat{\theta} - M_* \nabla_{\theta} \ell(\hat{\theta}), \quad (2.6.3)$$

with  $M_*$  given again by the solution of the convex program (2.3.11), with the definition of  $\hat{\Sigma}$  provided here. Notice that this construction is analogous to the one in [134] (although the present setting is somewhat more general) with the crucial difference of the construction of  $M_*$ .

A formal justification can be given in complete analogy with the linear case. Here we only provide a simple heuristic.

By Taylor expansion of  $\ell(\widehat{\theta})$  around  $\theta_0$  we get  $\widehat{\theta}^u \approx \widehat{\theta} - M_* \nabla_{\theta} \ell(\theta_0) - M_* \nabla_{\theta}^2 \ell(\theta_0)(\widehat{\theta} - \theta_0)$ . Approximating  $\nabla_{\theta}^2 \ell(\theta_0) \approx \widehat{\Sigma}$  (which amounts to taking expectation with respect to the response variables  $Y_i$ ), we get  $\widehat{\theta}^u - \theta_0 \approx -M_* \nabla_{\theta} \ell(\theta_0) - (M_* \widehat{\Sigma} - \mathbf{I})(\widehat{\theta} - \theta_0)$ . Conditionally on  $\{X_i\}_{1 \leq i \leq n}$ , the first term has zero expectation and covariance  $M_* \widehat{\Sigma} M_*^{\top}$ . Further, by central limit theorem, its low-dimensional marginals are approximately Gaussian. The bias term  $-(M_* \widehat{\Sigma} - \mathbf{I})(\widehat{\theta} - \theta_0)$  can be bounded as in the linear regression case, building on the fact that  $M_*$  is chosen as to minimize  $|M \widehat{\Sigma} - \mathbf{I}|_{\infty}$ .

**Example ( $\ell_1$ -regularized logistic regression).** Under this model, the binary response  $Y_i \in \{0, 1\}$  is distributed as  $Y_i \sim f_{\theta_0}(\cdot | X_i)$  where

$$\begin{aligned} f_{\theta_0}(1|x) &= \frac{1}{1 + e^{-\langle x, \theta_0 \rangle}}, \\ f_{\theta_0}(0|x) &= \frac{1}{1 + e^{\langle x, \theta_0 \rangle}}. \end{aligned}$$

It is easy to see that in this case  $\mathcal{I}_i(\widehat{\theta}) = \widehat{q}_i(1 - \widehat{q}_i)X_i X_i^{\top}$ , with  $\widehat{q}_i = (1 + e^{-\langle \widehat{\theta}, X_i \rangle})^{-1}$ , and thus

$$\widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \widehat{q}_i(1 - \widehat{q}_i)X_i X_i^{\top}.$$

## Chapter 3

# Hypothesis Testing in High-Dimensional Regression

An important advantage of sparse linear regression models is that they provide parsimonious explanations of the data in terms of a small number of covariates. The easiest way to select the ‘active’ covariates is to choose the indexes  $i$  for which  $\hat{\theta}_i^n \neq 0$ . This approach however does not provide a measure of statistical significance for the finding that the coefficient is non-zero.

In this chapter, we focus on hypothesis testing in high-dimensional regression models. More specifically, considering the linear regression model (1.1.6), we are interested in testing null hypothesis of the form:

$$H_{0,i} : \theta_{0,i} = 0, \tag{3.0.1}$$

versus the alternative  $H_{A,i} : \theta_{0,i} \neq 0$ , for  $i \in [p] \equiv \{1, 2, \dots, p\}$ , and constructing  $p$ -values for these tests. A testing procedure faces two types of error: type I error and type II error.

- *Type I error*: It is the incorrect rejection of a true null hypothesis.
- *Type II error*: It is the failure to reject a false null hypothesis.

The quality of a test is measured in terms of the probabilities of type I and type II errors, respectively denoted by  $\alpha$  and  $\beta$ . The quantity  $\alpha$  is also referred to as significance level (false positive rate) of the test, and  $1 - \beta$  as its statistical power (true positive rate).

It is important to consider the tradeoff between statistical significance and power. Indeed any significance level  $\alpha$  can be achieved by randomly rejecting  $H_{0,i}$  with probability  $\alpha$ . This test achieves power  $1 - \beta = \alpha$ . Further note that, without further assumption, no nontrivial power can be achieved. In fact, choosing  $\theta_{0,i} \neq 0$  arbitrarily close to zero,  $H_{0,i}$  becomes indistinguishable from its alternative. We will therefore assume that, whenever  $\theta_{0,i} \neq 0$ , we have  $|\theta_{0,i}| > \gamma$  as well. The smallest value of  $\mu$  such that the power and significance reach some fixed non-trivial value (e.g.,  $\alpha = 0.05$  and  $1 - \beta \geq 0.9$ ) has a particularly compelling interpretation, and provides an answer to the following question:

What is the minimum magnitude of  $\theta_{0,i}$  to be able to distinguish it from the noise level, with a given degree of confidence?

Some intuition can be gained by considering special cases. For the case of orthogonal designs we have  $n = p$  and  $\mathbf{X}^T \mathbf{X} = n\mathbf{I}_{n \times n}$ . By an orthogonal transformation, we can restrict ourselves to  $\mathbf{X} = \sqrt{n}\mathbf{I}_{n \times n}$ , i.e.,  $y_i = \sqrt{n}\theta_{0,i} + w_i$ . Hence testing hypothesis  $H_{0,i}$  reduces to testing for the mean of a univariate Gaussian. It is easy to see that we can distinguish the  $i$ -th entry from noise only if its size is at least of order  $\sigma/\sqrt{n}$ . More precisely, for any  $\alpha \in (0, 1)$ ,  $\beta \in (0, \alpha)$ , we can achieve significance  $\alpha$  and power  $1 - \beta$  if and only if  $|\theta_{0,i}| \geq c(\alpha, \beta) \sigma/\sqrt{n}$  for some constant  $c(\alpha, \beta)$  [89, Section 3.9].

In this chapter, we propose a procedure for testing  $H_{0,i}$  versus its alternative  $H_{A,i}$ . we further prove that this test achieves a ‘nearly optimal’ power-significance trade-off in the case of Gaussian random designs with i.i.d. rows. Here, ‘optimality’ is in the minimax sense. More specifically, we first prove a general upper bound on the minimax power of tests with a given significance level  $\alpha$ , over the family of  $s_0$ -sparse parameter vectors. We next show that our procedure exhibits the same trade-off, provided that the sample size is increased by a factor  $\eta_{\Sigma, s_0}$ , or equivalently if the noise level  $\sigma$  is decreased by a factor  $\sqrt{\eta_{\Sigma, s_0}}$ . Here,  $\Sigma$  denotes the population level covariance of the Gaussian design, i.e.  $\Sigma = \mathbb{E}(X_1 X_1^T)$ , and  $\eta_{\Sigma, s_0}$  is always upper bounded by the condition number of  $\Sigma$ . In particular,  $\eta_{\mathbf{I}, s_0} = 1$ .

Motivated by the distributional characterization of the debiased estimator  $\hat{\theta}^u$ , provided in Theorem 2.3.5, we construct a  $p$ -value  $P_i$  for the test  $H_{0,i}$  as follows:

$$P_i = 2 \left( 1 - \Phi \left( \frac{\sqrt{n} |\hat{\theta}_i^u|}{\hat{\sigma}_{[M \hat{\Sigma} M^T]_{i,i}^{1/2}}} \right) \right). \tag{3.0.2}$$

The decision rule is then based on the  $p$ -value  $P_i$ :

$$\widehat{T}_{i,\mathbf{X}}(Y) = \begin{cases} 1 & \text{if } P_i \leq \alpha & (\text{reject } H_{0,i}), \\ 0 & \text{otherwise} & (\text{accept } H_{0,i}), \end{cases} \quad (3.0.3)$$

where  $\alpha$  is the fixed target type I error probability.

### 3.1 Minimax formulation

We consider the minimax criterion to measure the quality of a testing procedure. In order to define it formally, we first need to establish some notations.

A testing procedure for the family of hypotheses  $H_{0,i}$ , cf. Eq. (3.0.1), is given by a family of measurable functions  $\{T_{i,\mathbf{X}}\}_{i \in [p]}$ , with  $T_{i,\mathbf{X}} : \mathbb{R}^n \rightarrow \{0, 1\}$ . Here  $T_{i,\mathbf{X}}(Y) = 1$  has the interpretation that hypothesis  $H_{0,i}$  is rejected when the observation is  $Y \in \mathbb{R}^n$  and the design matrix is  $\mathbf{X}$ . We will hereafter drop the subscript  $\mathbf{X}$  whenever clear from the context.

As mentioned above, we will measure the quality of a test  $T_i$  in terms of its significance level  $\alpha$  and power  $1 - \beta$ . Adopting a minimax point of view, we require that these metrics are achieved uniformly over  $s_0$ -sparse vectors. Formally, for  $\gamma > 0$ , we let

$$\alpha_n(T_i) \equiv \sup \left\{ \mathbb{P}_{\theta_0}(T_{i,\mathbf{X}}(Y) = 1) : \theta_0 \in \mathbb{R}^p, \|\theta_0\|_0 \leq s_0(n), \theta_{0,i} = 0 \right\}. \quad (3.1.1)$$

$$\beta_n(T_i; \gamma) \equiv \sup \left\{ \mathbb{P}_{\theta_0}(T_{i,\mathbf{X}}(Y) = 0) : \theta_0 \in \mathbb{R}^p, \|\theta_0\|_0 \leq s_0(n), |\theta_{0,i}| \geq \gamma \right\}. \quad (3.1.2)$$

In words, for any  $s_0$ -sparse vector with  $\theta_i = 0$ , the probability of false alarm is upper bounded by  $\alpha_n(T_i)$ . On the other hand, if  $\theta$  is  $s_0$ -sparse with  $|\theta_i| \geq \gamma$ , the probability of misdetection is upper bounded by  $\beta_n(T_i; \mu)$ . Note that  $\mathbb{P}_{\theta}(\cdot)$  is the induced probability distribution on  $(Y, \mathbf{X})$  for random design  $\mathbf{X}$  and noise realization  $w$ , given the fixed parameter vector  $\theta$ . Throughout we will accept randomized testing procedures as well.<sup>1</sup>

Our next theorem establishes bounds on  $\alpha_n(\widehat{T}_i)$  and  $\beta_n(\widehat{T}_i; \gamma)$  for our decision rule (3.0.3).

**Theorem 3.1.1.** *Consider a sequence of design matrices  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , with dimensions  $n \rightarrow \infty$ ,  $p = p(n) \rightarrow \infty$  satisfying the assumptions of Lemma 2.4.2.*

*Consider the linear model (1.1.6) and let  $\widehat{\theta}^u$  be defined as per Eq. (2.3.12) in Algorithm 1, with  $\mu = a\sqrt{(\log p)/n}$  and  $\lambda = \sigma\sqrt{(c^2 \log p)/n}$ , with  $a, c$  large enough constants. Finally,*

<sup>1</sup>Formally, this corresponds to assuming  $T_i(Y) = T_i(Y; U)$  with  $U$  uniform in  $[0, 1]$  and independent of the other random variables.

let  $\hat{\sigma} = \hat{\sigma}(y, \mathbf{X})$  a consistent estimator of the noise level in the sense of Eq. (2.4.1), and  $\hat{T}$  be the test defined in Eq. (3.0.3).

Then the following holds true for any fixed sequence of integers  $i = i(n)$ :

$$\lim_{n \rightarrow \infty} \alpha_n(\hat{T}_i) \leq \alpha. \quad (3.1.3)$$

$$\liminf_{n \rightarrow \infty} \frac{1 - \beta_n(\hat{T}_i; \gamma)}{1 - \beta_{i,n}^*(\gamma)} \geq 1, \quad 1 - \beta_{i,n}^*(\gamma) \equiv G\left(\alpha, \frac{\sqrt{n} \gamma}{\sigma[\Sigma_{i,i}^{-1}]^{1/2}}\right), \quad (3.1.4)$$

where, for  $\alpha \in [0, 1]$  and  $u \in \mathbb{R}_+$ , the function  $G(\alpha, u)$  is defined as follows:

$$G(\alpha, u) = 2 - \Phi(\Phi^{-1}(1 - \alpha/2) + u) - \Phi(\Phi^{-1}(1 - \alpha/2) - u).$$

Theorem 3.1.1 is proved in Section 6.2.1. It is easy to see that, for any  $\alpha > 0$ ,  $u \mapsto G(\alpha, u)$  is continuous and monotone increasing. Moreover,  $G(\alpha, 0) = \alpha$  which is the trivial power obtained by randomly rejecting  $H_{0,i}$  with probability  $\alpha$ . As  $\gamma$  deviates from zero, we obtain nontrivial power. Notice that in order to achieve a specific power  $\beta > \alpha$ , our scheme requires  $\gamma \geq c_\beta(\sigma/\sqrt{n})$ , for some constant  $c_\beta$  that depends on  $\beta$ . This is because  $\Sigma_{i,i}^{-1} \leq \sigma_{\max}(\Sigma^{-1}) \leq (\sigma_{\min}(\Sigma))^{-1} = O(1)$ .

## 3.2 Familywise error rate

A common scenario in much of empirical research is to test multiple hypotheses simultaneously. Accounting for the multiplicity of individual tests can be achieved by controlling an appropriate error rate. The traditional or classical *familywise error rate* (FWER) is the probability of one or more false discoveries. Formally considering the family of hypotheses  $\{H_{0,i} : \theta_{0,i} = 0\}_{i \in [p]}$ , we want to propose  $T_{i,\mathbf{X}} : \mathbb{R}^n \rightarrow \{0, 1\}$ , for each  $i \in [p]$ ,  $\mathbf{X} \in \mathbb{R}^{n \times p}$  such that

$$\text{FWER}(T, n) \equiv \sup_{\theta_0 \in \mathbb{R}^p, \|\theta_0\|_0 \leq s_0} \mathbb{P}\left\{\exists i \in [p] : \theta_{0,i} = 0, T_{i,\mathbf{X}}(Y) = 1\right\}, \quad (3.2.1)$$

Here  $T = \{T_{i,\mathbf{X}}\}_{i \in [p]}$  represents the family of tests.

In order to achieve familywise error control, we adopt a standard trick based on Bonferroni inequality. Given  $p$ -values defined as per Eq. (3.0.2), we let

$$\widehat{T}_{i,\mathbf{X}}^{\text{F}}(Y) = \begin{cases} 1 & \text{if } P_i \leq \alpha/p & (\text{reject } H_{0,i}), \\ 0 & \text{otherwise} & (\text{accept } H_{0,i}). \end{cases} \quad (3.2.2)$$

Then we have the following error control guarantee.

**Theorem 3.2.1.** *Consider a sequence of design matrices  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , with dimensions  $n \rightarrow \infty$ ,  $p = p(n) \rightarrow \infty$  satisfying the assumptions of Lemma 2.4.2.*

*Consider the linear model (1.1.6) and let  $\widehat{\theta}^u$  be defined as per Eq. (2.3.12) in Algorithm 1, with  $\mu = a\sqrt{(\log p)/n}$  and  $\lambda = \sigma\sqrt{(c^2 \log p)/n}$ , with  $a, c$  large enough constants. Finally, let  $\widehat{\sigma} = \widehat{\sigma}(y, \mathbf{X})$  be a consistent estimator of the noise level in the sense of Eq. (2.4.1), and  $\widehat{T}$  be the test defined in Eq. (3.2.2). Then:*

$$\limsup_{n \rightarrow \infty} \text{FWER}(\widehat{T}^{\text{F}}, n) \leq \alpha. \quad (3.2.3)$$

The proof of this theorem is similar to the one of Lemma 2.4.2 and Theorem 3.1.1, and is deferred to Section 6.2.2.

### 3.3 Minimax optimality of a test

We show near optimality of testing procedure (3.0.3) for the case of random designs with i.i.d. Gaussian rows. To this end, we first develop an upper bound for the minimax power of tests with a given significance level  $\alpha$ , under Gaussian random design model. Near optimality of test (3.0.3) is then proved by comparing its statistical power, as derived in Theorem 3.1.1 with the developed upper bound.

**Definition 3.3.1.** *The minimax power for testing hypothesis  $H_{0,i}$  against the alternative  $|\theta_i| \geq \gamma$  is given by the function  $1 - \beta_{i,n}^{\text{opt}}(\cdot; \gamma) : [0, 1] \rightarrow [0, 1]$  where, for  $\alpha \in [0, 1]$*

$$1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \equiv \sup_{T_i} \left\{ 1 - \beta_n(T_i; \gamma) : \alpha_n(T_i) \leq \alpha \right\}, \quad (3.3.1)$$

where  $\alpha_n(T_i)$  and  $\beta_n(T_i; \gamma)$  are respectively given by (3.1.1) and (3.1.2).

Note that for standard Gaussian designs (and more generally for designs with exchangeable columns),  $\alpha_n(T_i)$ ,  $\beta_n(T_i; \gamma)$  do not depend on the index  $i \in [p]$ .

The following are straightforward yet useful properties.

**Remark 3.3.2.** *The optimal power  $\alpha \mapsto 1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma)$  is non-decreasing. Further, by using a test such that  $T_{i,\mathbf{X}}(Y) = 1$  with probability  $\alpha$  independently of  $Y$ ,  $\mathbf{X}$ , we conclude that  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \geq \alpha$ .*

*Proof.* To prove the first property, notice that, for any  $\alpha \leq \alpha'$  we have  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \leq 1 - \beta_{i,n}^{\text{opt}}(\alpha'; \gamma)$ . Indeed  $1 - \beta_{i,n}^{\text{opt}}(\alpha'; \gamma)$  is obtained by taking the supremum in Eq. (3.3.1) over a family of tests that includes those over which the supremum is taken for  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma)$ .

Next, a completely randomized test outputs  $T_{i,\mathbf{X}}(Y) = 1$  with probability  $\alpha$  independently of  $\mathbf{X}, Y$ . We then have  $\mathbb{P}_\theta(T_{i,\mathbf{X}}(Y) = 0) = 1 - \alpha$  for any  $\theta$ , whence  $\beta_n(T_i; \gamma) = 1 - \alpha$ . Since this test offers –by definition– the prescribed control on type I errors, we have, by Eq. (3.3.1),  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \geq 1 - \beta_n(T_i; \gamma) = \alpha$ .  $\square$

### 3.3.1 Upper bound on the minimax power

Our upper bound on the minimax power is stated in terms of the function  $G : [0, 1] \times \mathbb{R}_+ \rightarrow [0, 1]$ ,  $(\alpha, u) \mapsto G(\alpha, u)$ , defined as follows.

$$G(\alpha, u) \equiv 2 - \Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) + u\right) - \Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - u\right). \quad (3.3.2)$$

It is easy to check that, for any  $\alpha > 0$ ,  $u \mapsto G(\alpha, u)$  is continuous and monotone increasing. For  $u$  fixed  $\alpha \mapsto G(\alpha, u)$  is continuous and monotone increasing. Finally  $G(\alpha, 0) = \alpha$  and  $\lim_{u \rightarrow \infty} G(\alpha, u) = 1$ .

We then have the following upper bound on the optimal power of random Gaussian designs. We provide a proof outline in Section 3.3.1.1, and refer to Section 6.2.5 for the complete proof.

**Theorem 3.3.3.** *For  $i \in [p]$ , let  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma)$  be the minimax power of a Gaussian random design  $\mathbf{X}$  with covariance matrix  $\Sigma \in \mathbb{R}^{p \times p}$ , as per Definition 3.3.1. For  $S \subseteq [p] \setminus \{i\}$ , define*

$\Sigma_{i|S} \equiv \Sigma_{ii} - \Sigma_{i,S} \Sigma_{S,S}^{-1} \Sigma_{S,i} \in \mathbb{R}$ . Then, for any  $\ell \in \mathbb{R}$  and  $|S| < s_0$ ,

$$1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \leq G\left(\alpha, \frac{\gamma}{\sigma_{\text{eff}}(\ell)}\right) + F_{n-s_0+1}(n - s_0 + \ell), \quad (3.3.3)$$

$$\sigma_{\text{eff}}(\ell) \equiv \frac{\sigma}{\sqrt{\Sigma_{i|S}(n - s_0 + \ell)}},$$

where  $F_k(x) = \mathbb{P}(Z_k \geq x)$ , and  $Z_k$  is a chi-squared random variable with  $k$  degrees of freedom.

In other words, the statistical power is upper bounded by the one of testing the mean of a scalar Gaussian random variable, with effective noise variance  $\sigma_{\text{eff}}^2 \approx \sigma^2 / [\Sigma_{i|S}(n - s_0)]$ . (Note indeed that by concentration of a chi-squared random variable around their mean,  $\ell$  can be taken small as compared to  $n - s_0$ .)

The next corollary specializes the above result to the case of standard Gaussian designs. (The proof is immediate and hence we omit it.)

**Corollary 3.3.4.** *For  $i \in [p]$ , let  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma)$  be the minimax power of a standard Gaussian design  $\mathbf{X}$  with covariance matrix  $\Sigma = \mathbf{I}_{p \times p}$ , cf. Definition 3.3.1. Then, for any  $\xi \in [0, (3/2)\sqrt{n - s_0 + 1}]$  we have*

$$1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \leq G\left(\alpha, \frac{\gamma(\sqrt{n - s_0 + 1} + \xi)}{\sigma}\right) + e^{-\xi^2/8}. \quad (3.3.4)$$

It is instructive to look at the last result from a slightly different point of view. Given  $\alpha \in (0, 1)$  and  $1 - \beta \in (\alpha, 1)$ , how big does the entry  $\gamma$  need to be so that  $1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \geq 1 - \beta$ ? It follows from Corollary 3.3.4 that to achieve a pair  $(\alpha, \beta)$  as above we require  $\gamma \geq \gamma_{\text{UB}} = c\sigma/\sqrt{n}$  for some  $c = c(\alpha, \beta)$ .

As further clarified in the next section and in Section 6.2.3, Theorem 3.3.3 by an oracle-based argument. Namely, we upper bound the power of any hypothesis testing method, by the power of an oracle that knows, for each coordinates  $j \in [p] \setminus i$ , whether  $\theta_{0,j} \in \text{supp}(\theta_0)$  or not. In other words the procedure has access to  $\text{supp}(\theta_0) \setminus \{i\}$ . At first sight, this oracle appears exceedingly powerful, and hence the bound might be loose. Surprisingly, the bound turns out to be tight, at least in an asymptotic sense, as demonstrated in Section 4.1.

### 3.3.1.1 Proof outline

The proof of Theorem 3.3.3 is based on a simple reduction to the binary hypothesis testing problem. We first introduce the binary testing problem, in which the vector of coefficients  $\theta$  is chosen randomly according to one of two distributions.

**Definition 3.3.5.** Let  $Q_0$  be a probability distribution on  $\mathbb{R}^p$  supported on  $\mathcal{R}_0 \equiv \{\theta \in \mathbb{R}^p : \|\theta\|_0 \leq s_0, \theta_i = 0\}$ , and  $Q_1$  a probability distribution supported on  $\mathcal{R}_1 \equiv \{\theta \in \mathbb{R}^p : \|\theta\|_0 \leq s_0, |\theta_i| \geq \gamma\}$ . For fixed design matrix  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , and  $z \in \{0, 1\}$ , let  $\mathbb{P}_{Q,z,\mathbf{X}}$  denote the law of  $y$  as per model (1.3.2) when  $\theta_0$  is chosen randomly with  $\theta_0 \sim Q_z$ .

We denote by  $1 - \beta_{i,\mathbf{X}}^{\text{bin}}(\cdot; Q)$  the optimal power for the binary hypothesis testing problem  $\theta_0 \sim Q_0$  versus  $\theta_0 \sim Q_1$ , namely:

$$\beta_{i,\mathbf{X}}^{\text{bin}}(\alpha_{\mathbf{X}}; Q) \equiv \inf_T \left\{ \mathbb{P}_{Q,1,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 0) : \mathbb{P}_{Q,0,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 1) \leq \alpha_{\mathbf{X}} \right\}. \quad (3.3.5)$$

The reduction is stated in the next lemma.

**Lemma 3.3.6.** Let  $Q_0, Q_1$  be any two probability measures supported, respectively, on  $\mathcal{R}_0$  and  $\mathcal{R}_1$  as per Definition 3.3.5. Then, the minimax power for testing hypothesis  $H_{0,i}$  under the random design model, cf. Definition 3.3.1, is bounded as

$$\beta_{i,n}^{\text{opt}}(\alpha; \gamma) \geq \inf \left\{ \mathbb{E} \beta_{i,\mathbf{X}}^{\text{bin}}(\alpha_{\mathbf{X}}; Q) : \mathbb{E}(\alpha_{\mathbf{X}}) \leq \alpha \right\}. \quad (3.3.6)$$

Here expectation is taken with respect to the law of  $\mathbf{X}$  and the inf is over all measurable functions  $\mathbf{X} \mapsto \alpha_{\mathbf{X}}$ .

For the proof we refer to Section 6.2.3.

The binary hypothesis testing problem is characterized in the next lemma by reducing it to a simple regression problem. Let  $\tilde{x}_i = \mathbf{X}e_i$  denote the  $i$ -th column of design  $\mathbf{X}$ . For  $S \subseteq [p]$ , we denote by  $P_S$  the orthogonal projector on the linear space spanned by the columns  $\{\tilde{x}_i\}_{i \in S}$ . We also let  $P_S^\perp = I_{n \times n} - P_S$  be the projector on the orthogonal subspace.

**Lemma 3.3.7.** Let  $\mathbf{X} \in \mathbb{R}^{n \times p}$  and  $i \in [p]$ . For  $S \subset [p] \setminus \{i\}$ ,  $\alpha \in [0, 1]$ , define

$$1 - \beta_{i,\mathbf{X}}^{\text{oracle}}(\alpha; S, \gamma) = G\left(\alpha, \frac{\gamma \|P_S^\perp \tilde{x}_i\|_2}{\sigma}\right). \quad (3.3.7)$$

If  $|S| < s_0$  then for any  $\xi > 0$  there exists distributions  $Q_0, Q_1$  as per Definition 3.3.5, depending on  $i, S, \gamma$  but not on  $\mathbf{X}$ , such that  $\beta_{i,\mathbf{X}}^{\text{bin}}(\alpha; Q) \geq \beta_{i,\mathbf{X}}^{\text{oracle}}(\alpha; S, \gamma) - \xi$ .

The proof of this Lemma is presented in Section 6.2.4.

The proof of Theorem 3.3.3 follows from Lemmas 3.3.6 and 3.3.7, cf. Section 6.2.5.

### 3.3.2 Near optimality of $\widehat{T}_{i,\mathbf{X}}(Y)$

Comparing the established upper bound in Theorem 3.3.3 with the result of Theorem 3.1.1, we get the following corollary.

**Corollary 3.3.8.** *Consider a Gaussian random design model that satisfies the conditions of Theorem 3.1.1, and for  $i \in [p]$ , let  $\widehat{T}_i$  be the testing procedure defined in Eq. (3.0.3), with  $\widehat{\theta}^u$  as in Algorithm 1. Further, let*

$$\eta_{\Sigma, s_0} \equiv \min_{i \in [p]; S} \left\{ \Sigma_{i|S} \Sigma_{ii}^{-1} : S \subseteq [p] \setminus \{i\}, |S| < s_0 \right\}. \quad (3.3.8)$$

Under the sparsity assumption  $s_0 = o(\sqrt{n}/\log p)$ , the following holds true. If  $\{T_i\}$  is any sequence of tests with  $\limsup_{n \rightarrow \infty} \alpha_n(T_i) \leq \alpha$ , then for any fixed sequence of integers  $i = i(n)$ , we have

$$\liminf_{n \rightarrow \infty} \frac{1 - \beta_n(\widehat{T}_i; \gamma)}{1 - \beta_{n/\eta_{\Sigma, s_0}}(T_i; \gamma)} \geq 1, \quad (3.3.9)$$

where  $\beta_n(T; \gamma)$  for a test  $T$  is defined by (3.1.2).

In other words, the asymptotic efficiency of the sequence tests  $\{\widehat{T}_i\}_{i \in [p]}$  is at least  $1/\eta_{\Sigma, s_0}$ .

Hence, our testing procedure has nearly optimal power in the following sense. It has power at least as large as the power of any other testing procedure, provided the latter is applied to a sample size decreased by a factor  $\eta_{\Sigma, s_0}$ .

Further, under the assumptions of Theorem 2.3.5, the factor  $\eta_{\Sigma, s_0}$  is a bounded constant. Indeed

$$\eta_{\Sigma, s_0} \leq \Sigma_{i,i}^{-1} \Sigma_{i,i} \leq \frac{\sigma_{\max}(\Sigma)}{\sigma_{\min}(\Sigma)} \leq \frac{C_{\max}}{C_{\min}}, \quad (3.3.10)$$

since  $\Sigma_{ii}^{-1} \leq (\sigma_{\min}(\Sigma))^{-1}$ , and  $\Sigma_{i|S} \leq \Sigma_{i,i} \leq \sigma_{\max}(\Sigma)$  due to  $\Sigma_{S,S} \succ 0$ .

Note that  $n$ ,  $\gamma$  and  $\sigma$  appears in our upper bound (3.3.3) in the combination  $\gamma\sqrt{n}/\sigma$ , which is the natural measure of the signal-to-noise ratio (where, for simplicity, we neglected  $s_0 = o(\sqrt{n}/\log p)$  with respect to  $n$ ). Hence, the above result can be restated as follows. The testing procedure (3.0.3) has power at least as large as the power of any other procedure, provided the latter is applied at a noise level augmented by a factor  $\sqrt{\eta_{\Sigma, s_0}}$ .

### 3.4 Other proposals

The method of Van de Geer, Bühlmann, Ritov and Dezeure [134], explained in Section 2.3.3 can also be used for constructing  $p$ -values and testing null hypotheses  $H_{0,i} : \theta_{0,i} = 0$  versus alternative  $H_{0,A} : \theta_{0,i} \neq 0$ . Here, we discuss two other proposals. Numerical comparison among the methods is given in Chapter 5.

#### 3.4.1 Multisample splitting

As we discussed in Section sec:WhyHard, while the problem of uncertainty assessment is challenging in high-dimensional models ( $n < p$ ), well-established methods are proposed for low-dimensional models ( $n > p$ ) using the large sample theory. The idea behind sample splitting method is based on a reduction from the high-dimensional domain to the low-dimensional domain.

More specifically, in this approach the sample set is split into two equal parts. On the first part, Lasso estimator is used to select the variables and then restricting to the chosen variables, statistical inference is performed based on the second half of the data. Note that the latter is a low-dimensional problem as the size of chosen variables is expected to be much smaller than the sample size.

The procedure is explained in Algorithm 2 aiming for controlling the familywise error rate (FWER).

One major assumption here is that the screening property holds for the first part, namely  $\widehat{S} \supseteq S$ . The whole idea is implicit in [142]. In practice, however, the screening property is not satisfied exactly. We refer to [18] for constructing valid  $p$ -values under weaker condition.

A nice property of Algorithm 2 is that the correction for multiplicity of testing involves only the factor  $|\widehat{S}|$ , whereas in classical Bonferroni the  $p$ -values are multiplied by  $p$  to adjust for multiple testing.

A major difficulty of the single sample splitting method is that it undergoes a  $p$ -value

---

**Algorithm 2** Single sample splitting for multiple testing of  $H_{0,i}$ , for  $i \in [p]$

---

**Input:** Measurement vector  $Y$ , design matrix  $\mathbf{X}$ .

**Output:** Adjusted  $p$ -values for controlling FWER.

- 1: Split sample set  $[n]$  into two equal subsets  $I_1$  and  $I_2$ .
- 2: Applying Lasso estimator on  $I_1$ , select the variables  $\widehat{S} \subseteq [p]$ . Choosing regularization parameter large enough, ensure that  $|\widehat{S}| \leq |I_2| = n/2$ .
- 3: Consider the sample set  $I_2$  restricted to the reduced set of variables  $\widehat{S}$ . Compute  $p$ -values  $P_i$  for  $H_{0,i}$ ,  $i \in \widehat{S}$ , using classical least squares (the t-test). This is well defined since  $|\widehat{S}| \leq |I_2|$ . For  $i \notin \widehat{S}$ , assign  $P_i = 1$ .
- 4: For multiple testing, adjust the  $p$ -values as

$$P_{\text{corr},i} = \min(P_i \cdot |\widehat{S}|, 1).$$


---

lottery phenomenon: Different splits leads to widely different corresponding  $p$ -values. To overcome this problem, Meinshausen et al. [96] proposed the multisample splitting procedure, summarized in [Algorithm 3](#)

The idea is to run the single sample splitting [Algorithm 2](#)  $B$  times, for a large  $B$ , leading to  $p$ -values  $P_{\text{corr},i}^{(1)}, \dots, P_{\text{corr},i}^{(B)}$ , for  $i \in [p]$ . These  $p$ -values are then aggregated to a single  $p$ -value  $P_i$ .

Note that  $p$ -values  $P_{\text{corr},i}^{(1)}, \dots, P_{\text{corr},i}^{(B)}$  are dependent since all the different splits are stem from the same full data set, and therefore the aggregation should be done carefully. The method of [96] aggregates the  $p$ -values through empirical  $\eta$ -quantiles  $Q(\eta)$ :

$$Q_i(\eta) \equiv \min \left\{ \text{empirical } \eta\text{-quantile } (P_{\text{corr},i}^{(b)}/\eta; b = 1, \dots, B), 1 \right\}. \quad (3.4.1)$$

### 3.4.2 Bias-corrected projection estimator

We begin with a general linear estimator for estimating parameters  $\theta_{0,i}$ . Namely, for each  $i \in [p]$ , consider a vector  $Z_i \in \mathbb{R}^n$  and a corresponding estimator

$$\widehat{\theta}_{0,i} = \frac{\langle Z_i, Y \rangle}{\langle Z_i, \widetilde{x}_i \rangle},$$

---

**Algorithm 3** Multisample splitting for multiple testing of  $H_{0,i}$ , for  $i \in [p]$

---

**Input:** Measurement vector  $Y$ , design matrix  $\mathbf{X}$ , parameters  $B$ ,  $0 < \eta_{\min} < 1$ .

**Output:** Adjusted  $p$ -values for controlling FWER.

- 1: Run the single sample splitting **Algorithm 2**  $B$  times. Let  $\{P_{\text{corr},i}^{(b)} : b = 1, \dots, B\}$  be the produced  $p$ -values.
- 2: Aggregate the  $p$ -values from the previous step, leading to a single  $p$ -value  $P_i$ , for  $i \in [p]$ , in the following way:

$$P_i = \min \left\{ (1 - \log(\eta_{\min})) \inf_{\eta \in (\eta_{\min}, 1)} Q_i(\eta), 1 \right\},$$

where  $Q_i(\eta)$  is given by (3.4.1).

---

where  $\tilde{x}_i = \mathbf{X}e_i$  denotes the  $i$ -th column of  $\mathbf{X}$ . Substituting for  $Y = \sum_{j=1}^p \tilde{x}_j \theta_{0,j} + W$  and using  $\mathbb{E}(W) = 0$ , we obtain

$$\mathbb{E}(\hat{\theta}_i) = \theta_{0,i} + \sum_{j \in [p], j \neq i} P_{ij} \theta_{0,j}, \quad P_{ij} = \frac{\langle Z_i, \tilde{x}_j \rangle}{\langle Z_i, \tilde{x}_j \rangle}.$$

Clearly, any linear estimator is biased in the high-dimensional regime ( $n < p$ ), since there is no vector  $Z_i \in \mathbb{R}^n$  that is orthogonal to all the  $(p-1)$  vectors  $\tilde{x}_j$ , for  $j \in [p]$ ,  $j \neq i$ .

The idea is to correct bias of such linear estimator using the Lasso. This leads to the following bias-corrected estimator:

$$\hat{\theta}_{\text{corr},i} = \hat{\theta}_i - \sum_{j \neq i} P_{ij} \hat{\theta}_{\text{Lasso},j}.$$

Concrete suggestions are proposed for choosing vectors  $Z_i$ . In [16], these vectors are chosen based on Ridge regression (*Ridge-type projection estimator*  $\hat{\theta}_{\text{corr-Ridge}}$ ). The authors in [149], use  $Z_i$  constructed based on Lasso regression (*Low-dimensional projection estimator*  $\hat{\theta}_{\text{corr-Lasso}}$ ).

It is shown that  $\hat{\theta}_{\text{corr-Ridge}}$  and  $\hat{\theta}_{\text{corr-Lasso}}$  are unbiased under sparsity scaling  $s_0 = o((n/\log p)^\xi)$  for some  $\xi \in (0, 1/2)$ , or equivalently under sample size  $n = \omega((\log p)s_0^{1/\xi})$ . More specifically,

$$\sqrt{n}(\hat{\theta}_{\text{corr},i} - \theta_{0,i})/\hat{\sigma}_i \rightarrow \text{N}(0, 1),$$

where  $\hat{\sigma}_i^2 = \hat{\sigma}^2 \omega_i$  with  $\hat{\sigma}$  a consistent estimate of  $\sigma$  (e.g., using the scaled Lasso (2.4.3)), and  $\omega_i$  is given explicitly based on the design  $\mathbf{X}$ . Confidence intervals and  $p$ -values are then

derived based on such distributional characterization.

When specialized to the case of standard Gaussian designs  $X_i \sim \mathbf{N}(0, \mathbf{I}_{p \times p})$ , these methods require  $|\theta_{0,i}| \geq c \max\{\sigma s_0 \log p/n, \sigma/\sqrt{n}\}$  to reject hypothesis  $H_{0,i}$  with a given degree of confidence (with  $c$  being a constant independent of the problem dimensions). By contrast, as discussed below Eq. (3.1.4), testing procedure (3.0.3) requires  $|\theta_{0,i}| \geq c_\beta(\sigma/\sqrt{n})$  to achieve a specific power  $\beta$ , with  $c_\beta$  depending upon power  $\beta$ . In other words, the Ridge-based and the low-dimensional projection estimators require the coefficient of interest to be larger by a diverging factor  $O(s_0 \log p/\sqrt{n})$  than what is required by testing procedure (3.0.3).

In Chapter 5, we also present extensive simulation results indicating that the proposed methods of [16] and [149] are overly conservative. Namely, they achieve smaller type I error than the prescribed level  $\alpha$  and this comes at the cost of a smaller statistical power than testing procedure (3.0.3).

## Chapter 4

# Hypothesis Testing under Optimal Sample Size

The methods presented in previous chapters for constructing confidence intervals and hypothesis testing require the sample size  $n$  that asymptotically dominates  $(s_0 \log p)^2$ . On the other hand, nearly optimal estimation via Lasso is possible for significantly smaller sample size, namely for  $n \geq Cs_0 \log p$ , for some constant  $C$  [25, 14]. This suggests the following natural question

*Is it possible to design a minimax optimal test for hypotheses  $H_{0,i}$ , for optimal sample size  $n = O(s_0 \log(p/s_0))$ ?*

In this chapter we provide an affirmative answer in case of random designs with i.i.d. Gaussian rows.

Our approach is based on a debiasing of the Lasso estimator, slightly different from the one presented in Chapter 2. In case of standard Gaussian designs, our analysis is based on rigorous characterization of the asymptotic distribution of the Lasso estimator and its debiased version. In case of non-standard Gaussian designs, i.e. i.i.d. rows  $X_i \sim \mathbf{N}(0, \Sigma)$ , we derive a similar distributional characterization, termed *standard distributional limit*. The analysis assumes  $\Sigma$  is known and is based on the powerful replica heuristics in statistical physics.

The Gaussian design assumption arises naturally in some important applications. Consider for instance the problem of learning a high-dimensional Gaussian graphical model from data. In this case we are given i.i.d. samples  $Z_1, Z_2, \dots, Z_n \sim \mathbf{N}(0, K^{-1})$ , with  $K$  a sparse

positive definite matrix whose non-zero entries encode the underlying graph structure. As first shown by Meinshausen and Bühlmann [95], the  $i$ -th row of  $K$  can be estimated by performing linear regression of the  $i$ -th entry of the samples  $Z_1, Z_2, \dots, Z_n$  onto the other entries [113]. This reduces the problem to a high-dimensional regression model under Gaussian designs. Standard Gaussian designs were also shown to provide useful insights for compressed sensing applications [37, 45, 46, 48].

## 4.1 Hypothesis testing for standard Gaussian designs

In Subsection 4.1.1 we describe our hypothesis testing procedure (that we refer to as `SDL-TEST`) in the case of standard Gaussian designs. In Subsection 4.1.2, we develop asymptotic bounds on the probability of type I and type II errors. The test is shown to nearly achieve the ideal tradeoff between significance level  $\alpha$  and power  $1 - \beta$ , using the upper bound stated in Section 3.3.

Our results are based on a characterization of the high-dimensional behavior of the Lasso estimator, developed in [8]. For the reader's convenience, and to provide further context, we recall this result in subsection 4.1.3.

### 4.1.1 Hypothesis testing procedure

Our `SDL-TEST` procedure for standard Gaussian designs is described in Algorithm 4.

The key is the construction of the *unbiased estimator*  $\hat{\theta}^u$  in step 3. The asymptotic analysis developed in [8] and in the next section establishes that  $\hat{\theta}^u$  is an *asymptotically unbiased estimator* of  $\theta_0$ , and the empirical distribution of  $\{\hat{\theta}_i^u - \theta_{0,i}\}_{i=1}^p$  is asymptotically normal with variance  $\tau^2$ . Further, the variance  $\tau^2$  can be consistently estimated using the residual vector  $r \equiv d(y - \mathbf{X}\hat{\theta})/\sqrt{n}$ . These results establish that (in a sense that will be made precise next) the regression model (1.3.2) is asymptotically equivalent to a simpler sequence model

$$\hat{\theta}^u = \theta_0 + \text{noise} \tag{4.1.1}$$

with `noise` having zero mean. In particular, under the null hypothesis  $H_{0,i}$ ,  $\hat{\theta}_i^u$  is asymptotically gaussian with mean 0 and variance  $\tau^2$ . This motivates rejecting the null if  $|\hat{\theta}_i^u| \geq \tau\Phi^{-1}(1 - \alpha/2)$ . Note that debiased estimator  $\hat{\theta}^u$  (4.1.3) is defined in a very similar

---

**Algorithm 4** SDL-TEST: Testing hypothesis  $H_{0,i}$  under standard Gaussian design.

---

**Input:** regularization parameter  $\lambda$ , significance level  $\alpha$

**Output:** p-values  $P_i$ , test statistics  $T_{i,\mathbf{X}}(Y)$

1: Let

$$\hat{\theta}(\lambda) = \operatorname{argmin}_{\theta \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|^2 + \lambda \|\theta\|_1 \right\}.$$

2: Let

$$\mathbf{d} = \left( 1 - \frac{1}{n} \|\hat{\theta}(\lambda)\|_0 \right)^{-1}, \quad \tau = \frac{1}{\Phi^{-1}(0.75)} \frac{\mathbf{d}}{\sqrt{n}} |(Y - \mathbf{X}\hat{\theta}(\lambda))|_{(n/2)}, \quad (4.1.2)$$

where for  $v \in \mathbb{R}^K$ ,  $|v|_\ell$  is the  $\ell$ -th largest entry in the vector  $(|v_1|, \dots, |v_n|)$ .

3: Let

$$\hat{\theta}^u = \hat{\theta}(\lambda) + \frac{\mathbf{d}}{n} \mathbf{X}^\top (Y - \mathbf{X}\hat{\theta}(\lambda)). \quad (4.1.3)$$

4: Assign the p-values  $P_i$  for the test  $H_{0,i}$  as follows.

$$P_i = 2 \left( 1 - \Phi \left( \left| \frac{\hat{\theta}_i^u}{\tau} \right| \right) \right).$$

5: The decision rule is then based on the p-values:

$$T_{i,\mathbf{X}}(Y) = \begin{cases} 1 & \text{if } P_i \leq \alpha & \text{(reject the null hypothesis } H_{0,i}), \\ 0 & \text{otherwise} & \text{(accept the null hypothesis).} \end{cases}$$


---

way to the debiased estimator (2.3.12), in Chapter 2, with  $M = \mathbf{d} \cdot \mathbf{I}$ . We refer to Section 4.3 for a thorough discussion on the role of this factor.

### 4.1.2 Asymptotic analysis

For given dimension  $p$ , an *instance* of the standard Gaussian design model is defined by the tuple  $(\theta_0, n, \sigma)$ , where  $\theta_0 \in \mathbb{R}^p$ ,  $n \in \mathbb{N}$ ,  $\sigma \in \mathbb{R}_+$ . We consider sequences of instances indexed by the problem dimension  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$ .

**Definition 4.1.1.** *The sequence of instances  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  indexed by  $p$  is said to be a converging sequence if  $n(p)/p \rightarrow \delta \in (0, \infty)$ ,  $\sigma(p)^2/n \rightarrow \sigma_0^2$ , and the empirical*

distribution of the entries  $\theta_0(p)$  converges weakly to a probability measure  $p_{\Theta_0}$  on  $\mathbb{R}$  with bounded second moment. Further  $p^{-1} \sum_{i \in [p]} \theta_{0,i}(p)^2 \rightarrow \mathbb{E}_{p_{\Theta_0}} \{\Theta_0^2\}$ .

Note that this definition assumes the coefficients  $\theta_{0,i}$  are of order one, while the noise is scaled as  $\sigma(p)^2 = \Theta(n)$ . Equivalently, we could have assumed  $\theta_{0,i} = \Theta(1/\sqrt{n})$  and  $\sigma^2(p) = \Theta(1)$ : the two settings only differ by a scaling of  $Y$ . We favor the first scaling as it simplifies somewhat the notation in the following.

As before, we will measure the quality of the proposed test in terms of its significance level (size)  $\alpha$  and power  $1 - \beta$ . Recall that  $\alpha$  and  $\beta$  respectively indicate the type I error (false positive) and type II error (false negative) rates. The following theorem establishes that the  $P_i$ 's are indeed valid p-values, i.e., allow to control type I errors. Throughout  $S_0(p) = \{i \in [p] : \theta_{0,i}(p) \neq 0\}$  is the support of  $\theta_0(p)$ .

**Theorem 4.1.2.** *Let  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  be a converging sequence of instances of the standard Gaussian design model. Assume  $\lim_{p \rightarrow \infty} |S_0(p)|/p = \mathbb{P}(\Theta_0 \neq 0)$ . Then, for  $i \in S_0^c(p)$ , we have*

$$\lim_{p \rightarrow \infty} \mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1) = \alpha. \quad (4.1.4)$$

A more general form of Theorem 4.1.2 (cf. Theorem 4.2.3) is proved in Section 6.3.2. We indeed prove the stronger claim that the following holds true almost surely

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0^c(p)|} \sum_{i \in S_0^c(p)} T_{i,\mathbf{X}}(Y) = \alpha. \quad (4.1.5)$$

The result of Theorem 4.1.2 follows then by taking the expectation of both sides of Eq. (4.1.5) and using bounded convergence theorem and exchangeability of the columns of  $\mathbf{X}$ .

Our next theorem proves a lower bound for the power of the proposed test. Here, we assume that the non-zero entries of  $\theta_0$  are lower bounded in magnitude. Otherwise, it would be impossible to distinguish arbitrarily small parameters  $\theta_{0,i}$  from  $\theta_{0,i} = 0$ . Notice that for the testing procedure (3.0.3) to achieve the near optimal minimax power, we only required  $|\theta_{0,i}| \geq \gamma$  where  $\theta_{0,i}$  is the parameter of interest. By contrast, here we are assuming that  $|\theta_{0,i}| \geq \gamma$  for all  $i \in S_0(p)$ .

**Theorem 4.1.3.** *There exists a (deterministic) choice of  $\lambda = \lambda(p_{\Theta_0}, \sigma, \varepsilon, \delta)$  such that the following happens.*

Let  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  be a converging sequence of instances under the standard Gaussian design model. Assume that  $|S_0(p)| \leq \varepsilon p$ , and for all  $i \in S_0(p)$ ,  $|\theta_{0,i}(p)| \geq \gamma$  with  $\gamma = \gamma_0 \sigma(p) / \sqrt{n(p)}$ . for  $i \in S_0(p)$ , we have

$$\lim_{p \rightarrow \infty} \mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1) \geq G\left(\alpha, \frac{\gamma_0}{\tau_*}\right), \quad (4.1.6)$$

where  $\tau_* = \tau_*(\sigma_0, \varepsilon, \delta)$  is defined as follows

$$\tau_*^2 = \begin{cases} \frac{1}{1 - M(\varepsilon)/\delta}, & \text{if } \delta > M(\varepsilon), \\ \infty, & \text{if } \delta \leq M(\varepsilon). \end{cases} \quad (4.1.7)$$

Here,  $M(\varepsilon)$  is given by the following parametric expression in terms of the parameter  $\kappa \in (0, \infty)$ :

$$\varepsilon = \frac{2(\phi(\kappa) - \kappa\Phi(-\kappa))}{\kappa + 2(\phi(\kappa) - \kappa\Phi(-\kappa))}, \quad M(\varepsilon) = \frac{2\phi(\kappa)}{\kappa + 2(\phi(\kappa) - \kappa\Phi(-\kappa))}. \quad (4.1.8)$$

In Appendix B.2, we also provide an explicit formula for the regularization parameter  $\lambda = \lambda(p_{\Theta_0}, \sigma, \varepsilon, \delta)$  that achieves this power. Theorem 4.1.3 is proved in Section 6.3.1. We indeed prove the stronger claim that the following holds true almost surely:

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} T_{i,\mathbf{X}}(Y) \geq G\left(\alpha, \frac{\gamma_0}{\tau_*}\right). \quad (4.1.9)$$

The result of Theorem 4.1.3 follows then by taking the expectation of both sides of Eq. (4.1.9) and using exchangeability of the columns of  $\mathbf{X}$ .

Again, it is convenient to rephrase Theorem 4.1.3 in terms of the minimum value of  $\gamma$  for which we can achieve statistical power  $1 - \beta \in (\alpha, 1)$  at significance level  $\alpha$ . It is known that  $M(\varepsilon) = 2\varepsilon \log(1/\varepsilon) (1 + O(\varepsilon))$  [44]. Hence, for  $n \geq 2s_0 \log(p/s_0) (1 + O(s_0/p))$ , we have  $\tau_*^2 = O(1)$ . Since  $\lim_{u \rightarrow \infty} G(\alpha, u) = 1$ , any pre-assigned statistical power can be achieved by taking  $\gamma \geq C(\varepsilon, \delta)\sigma/\sqrt{n}$ .

Let us finally comment on the choice of the regularization parameter  $\lambda$ . Theorem 4.1.2 holds irrespective of  $\lambda$ , as long as it is kept fixed in the asymptotic limit. In other words, control of type I errors is fairly insensitive to the regularization parameters. On the other hand, to achieve optimal minimax power, it is necessary to tune  $\lambda$  to the correct value. The tuned value of  $\lambda = \lambda(p_{\Theta_0}, \sigma, \varepsilon, \delta)$  for the standard Gaussian sequence model is provided in

Appendix B.2. Further, the factor  $\sigma$  (and hence the need to estimate the noise level) can be omitted if –instead of the Lasso– we use the scaled Lasso [124]. In Chapter 5, Section 5.4, we discuss another way of choosing  $\lambda$  that also avoid estimating the noise level.

### 4.1.3 Gaussian limit

Theorems 4.1.2 and 4.1.3 are based on an asymptotic distributional characterization of the Lasso estimator developed in [8]. We restate it here for the reader’s convenience.

**Theorem 4.1.4** ([8]). *Let  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  be a converging sequence of instances of the standard Gaussian design model. Denote by  $\hat{\theta} = \hat{\theta}(Y, \mathbf{X}, \lambda)$  the Lasso estimator given as per Eq. (1.1.7) and define  $\hat{\theta}^u \in \mathbb{R}^p$ ,  $r \in \mathbb{R}^n$  by letting*

$$\hat{\theta}^u \equiv \hat{\theta} + \frac{\mathbf{d}}{n} \mathbf{X}^\top (Y - \mathbf{X}\hat{\theta}), \quad r \equiv \frac{\mathbf{d}}{\sqrt{n}} (Y - \mathbf{X}\hat{\theta}), \quad (4.1.10)$$

with  $\mathbf{d} = (1 - \|\hat{\theta}\|_0/n)^{-1}$ .

Then, with probability one, the empirical distribution of  $\{(\theta_{0,i}, \hat{\theta}_i^u)\}_{i=1}^p$  converges weakly to the probability distribution of  $(\Theta_0, \Theta_0 + \tau_0 Z)$ , for some  $\tau_0 \in \mathbb{R}$ , where  $Z \sim \mathbf{N}(0, 1)$ , and  $\Theta_0 \sim p_{\Theta_0}$  is independent of  $Z$ . Furthermore, with probability one, the empirical distribution of  $\{r_i\}_{i=1}^p$  converges weakly to  $\mathbf{N}(0, \tau_0^2)$ .

Finally  $\tau_0 \in \mathbb{R}$  is defined by the unique solution of Eqs. (B.1.4) and (B.1.5) in Appendix B.1.

In particular, this result implies that the empirical distribution of  $\{\hat{\theta}_i^u - \theta_{0,i}\}_{i=1}^p$  is asymptotically normal with variance  $\tau_0^2$ . This naturally motivates the use of  $|\hat{\theta}_i^u|/\tau_0$  as a test statistics for hypothesis  $H_{0,i} : \theta_{0,i} = 0$ .

The definitions of  $\mathbf{d}$  and  $\tau$  in step 2 are also motivated by Theorem 4.1.4. In particular,  $\mathbf{d}(Y - \mathbf{X}\hat{\theta})/\sqrt{n}$  is asymptotically normal with variance  $\tau_0^2$ . This is used in step 2, where  $\tau$  is just the robust median absolute deviation (MAD) estimator (we choose this estimator since it is more resilient to outliers than the sample variance [66]).

## 4.2 Hypothesis testing for nonstandard Gaussian designs

In this section, we generalize our testing procedure to nonstandard Gaussian design models where the rows of the design matrix  $\mathbf{X}$  are drawn independently from distribution  $\mathbf{N}(0, \Sigma)$ .

We first describe the generalized SDL-TEST procedure in subsection 4.2.1 under the assumption that  $\Sigma$  is known. In subsection 4.2.2, we show that this generalization can be justified from a certain generalization of the Gaussian limit theorem 4.1.4 to nonstandard Gaussian designs. Establishing such a generalization of Theorem 4.1.4 appears extremely challenging. We nevertheless show that such a limit theorem follows from the replica method of statistical physics in section 4.2.3.

Finally, in Section 4.2.4 we propose Algorithm 6, a procedure for estimating the covariance  $\Sigma$ .

### 4.2.1 Hypothesis testing procedure

The hypothesis testing procedure SDL-TEST for general Gaussian designs is defined in Algorithm 5. The basic intuition of this generalization is that  $(\hat{\theta}_i^u - \hat{\theta}_{0,i})/(\tau[(\Sigma^{-1})_{ii}]^{1/2})$  is expected to be asymptotically  $\mathbf{N}(0, 1)$ , whence the definition of (two-sided) p-values  $P_i$  follows as in step 4. Parameters  $\mathbf{d}$  and  $\tau$  in step 2 are defined in the same manner to the standard Gaussian designs.

### 4.2.2 Asymptotic analysis

For given dimension  $p$ , an *instance* of the nonstandard Gaussian design model is defined by the tuple  $(\Sigma, \theta_0, n, \sigma)$ , where  $\Sigma \in \mathbb{R}^{p \times p}$ ,  $\Sigma \succ 0$ ,  $\theta_0 \in \mathbb{R}^p$ ,  $n \in \mathbb{N}$ ,  $\sigma \in \mathbb{R}_+$ . We are interested in the asymptotic properties of sequences of instances indexed by the problem dimension  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$ . Motivated by Proposition 4.1.4, we define a property of a sequence of instances that we refer to as *standard distributional limit*.

**Definition 4.2.1.** *A sequence of instances  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  indexed by  $p$  is said to have an (almost sure) standard distributional limit if there exist  $\tau, \mathbf{d} \in \mathbb{R}$  (with  $\mathbf{d}$  potentially random, and both  $\tau, \mathbf{d}$  potentially depending on  $p$ ), such that the following holds. Denote by  $\hat{\theta} = \hat{\theta}(y, \mathbf{X}, \lambda)$  the Lasso estimator given as per Eq. (1.1.7) and define  $\hat{\theta}^u \in \mathbb{R}^p$ ,  $r \in \mathbb{R}^n$  by letting*

$$\hat{\theta}^u \equiv \hat{\theta} + \frac{\mathbf{d}}{n} \Sigma^{-1} \mathbf{X}^\top (Y - \mathbf{X} \hat{\theta}), \quad r \equiv \frac{\mathbf{d}}{\sqrt{n}} (Y - \mathbf{X} \hat{\theta}). \quad (4.2.3)$$

Let  $v_i = (\theta_{0,i}, (\hat{\theta}_i^u - \theta_{0,i})/\tau, (\Sigma^{-1})_{ii})$ , for  $1 \leq i \leq p$ , and  $\nu^{(p)}$  be the empirical distribution of

---

**Algorithm 5** SDL-TEST: Testing hypothesis  $H_{0,i}$  under nonstandard Gaussian design.

---

**Input:** regularization parameter  $\lambda$ , significance level  $\alpha$ , covariance matrix  $\Sigma$

**Output:** p-values  $P_i$ , test statistics  $T_{i,\mathbf{X}}(Y)$

1: Let

$$\hat{\theta}(\lambda) = \operatorname{argmin}_{\theta \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|^2 + \lambda \|\theta\|_1 \right\}.$$

2: Let

$$\mathbf{d} = \left( 1 - \frac{1}{n} \|\hat{\theta}(\lambda)\|_0 \right)^{-1}, \quad \tau = \frac{1}{\Phi^{-1}(0.75)} \frac{\mathbf{d}}{\sqrt{n}} |(Y - \mathbf{X}\hat{\theta}(\lambda))|_{(n/2)}, \quad (4.2.1)$$

where for  $v \in \mathbb{R}^K$ ,  $|v|_\ell$  is the  $\ell$ -th largest entry in the vector  $(|v_1|, \dots, |v_n|)$ .

3: Let

$$\hat{\theta}^u = \hat{\theta}(\lambda) + \frac{\mathbf{d}}{n} \Sigma^{-1} \mathbf{X}^\top (Y - \mathbf{X}\hat{\theta}(\lambda)). \quad (4.2.2)$$

4: Assign the p-values  $P_i$  for the test  $H_{0,i}$  as follows.

$$P_i = 2 \left( 1 - \Phi \left( \left| \frac{\hat{\theta}_i^u}{\tau [(\Sigma^{-1})_{ii}]^{1/2}} \right| \right) \right).$$

5: The decision rule is then based on the p-values:

$$T_{i,\mathbf{X}}(Y) = \begin{cases} 1 & \text{if } P_i \leq \alpha & \text{(reject the null hypothesis } H_{0,i}), \\ 0 & \text{otherwise} & \text{(accept the null hypothesis).} \end{cases}$$


---

$\{v_i\}_{i=1}^p$  defined as

$$\nu^{(p)} = \frac{1}{p} \sum_{i=1}^p \delta_{v_i}, \quad (4.2.4)$$

where  $\delta_{v_i}$  denotes the Dirac delta function centered at  $v_i$ . Then, with probability one, the empirical distribution  $\nu^{(p)}$  converges weakly to a probability measure  $\nu$  on  $\mathbb{R}^3$  as  $p \rightarrow \infty$ . Here,  $\nu$  is the probability distribution of  $(\Theta_0, \Upsilon^{1/2}Z, \Upsilon)$ , where  $Z \sim \mathbf{N}(0, 1)$ , and  $\Theta_0$  and  $\Upsilon$  are random variables independent of  $Z$ . Furthermore, with probability one, the empirical distribution of  $\{r_i/\tau\}_{i=1}^n$  converges weakly to  $\mathbf{N}(0, 1)$ .

**Remark 4.2.2.** *This definition is non-empty by Theorem 4.1.4. Indeed, if  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  is converging as per Definition 4.1.1, and  $a > 0$  is a constant, then Theorem 4.1.4 states that  $\{(\Sigma(p) = a \mathbf{I}_{p \times p}, \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  has a standard distributional limit.*

Proving the standard distributional limit for general sequences  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  is an outstanding mathematical challenge. In Section 4.2.3 we discuss non-rigorous evidence towards its validity using replica method in statistical physics. The numerical simulations in Chapter 5 further support the usefulness of this notion.

We will next show that the SDL-TEST procedure is appropriate for any random design model for which the standard distributional limit holds. Our first theorem is a generalization of Theorem 4.1.2 to this setting.

**Theorem 4.2.3.** *Let  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  be a sequence of instances for which a standard distributional limit holds. Further assume  $\lim_{p \rightarrow \infty} |S_0(p)|/p = \mathbb{P}(\Theta_0 \neq 0)$ . Then,*

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0^c(p)|} \sum_{i \in S_0^c(p)} \mathbb{P}_{\theta_0(p)}(T_{i, \mathbf{X}}(Y) = 1) = \alpha. \quad (4.2.5)$$

The proof of Theorem 4.2.3 is deferred to Section 6.3.2. In the proof, we show the stronger result that the following holds true almost surely

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0^c(p)|} \sum_{i \in S_0^c(p)} T_{i, \mathbf{X}}(Y) = \alpha. \quad (4.2.6)$$

The result of Theorem 4.2.3 follows then by taking the expectation of both sides of Eq. (4.2.6) and using bounded convergence theorem.

The following theorem characterizes the power of SDL-TEST for general  $\Sigma$ , and under the assumption that a standard distributional limit holds.

**Theorem 4.2.4.** *Let  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  be a sequence of instances with standard distributional limit. Assume (without loss of generality)  $\sigma(p) = \sqrt{n(p)}$ , and further  $|\theta_{0,i}(p)|/[(\Sigma^{-1})_{ii}]^{1/2} \geq \gamma_0$  for all  $i \in S_0(p)$ , and  $\lim_{p \rightarrow \infty} |S_0(p)|/p = \mathbb{P}(\Theta_0 \neq 0) \in (0, 1)$ . Then,*

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{P}_{\theta_0(p)}(T_{i, \mathbf{X}}(Y) = 1) \geq G\left(\alpha, \frac{\gamma_0}{\tau}\right). \quad (4.2.7)$$

Theorem 4.2.4 is proved in Section 6.3.3. We indeed prove the stronger result that the following holds true almost surely

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} T_{i, \mathbf{x}}(Y) \geq G\left(\alpha, \frac{\gamma_0}{\tau}\right). \quad (4.2.8)$$

Note that in Theorem 4.1.3,  $\tau_*$  admits an explicit formula that, for a suitable choice of  $\lambda$ , leads to an analytical lower bound for the power. By contrast, in Theorem 4.2.4,  $\tau$  depends upon  $\lambda$  implicitly and can be estimated from the data as in step 3 of SDL-TEST procedure. The result of Theorem 4.2.4 holds for any value of  $\lambda$ .

### 4.2.3 Gaussian limit via the replica heuristics

As mentioned above, the standard distributional limit follows from Theorem 4.1.4 for  $\Sigma = \mathbf{I}_{p \times p}$ . Even in this simple case, the proof is rather challenging [8]. Partial generalization to non-gaussian designs and other convex problems appeared recently in [6] and [106], each requiring over 50 pages of proofs.

On the other hand, these and similar asymptotic results can be derived heuristically using the ‘replica method’ from statistical physics. In Appendix B.3, we use this approach to derive the following claim<sup>1</sup>.

**Replica Method Claim 4.2.5.** *Assume the sequence of instances  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  is such that, as  $p \rightarrow \infty$ : (i)  $n(p)/p \rightarrow \delta > 0$ ; (ii)  $\sigma(p)^2/n(p) \rightarrow \sigma_0^2 > 0$ ; (iii) The sequence of functions*

$$\mathfrak{E}^{(p)}(a, b) \equiv \frac{1}{p} \mathbb{E} \min_{\theta \in \mathbb{R}^p} \left\{ \frac{b}{2} \|\theta - \theta_0 - \sqrt{a} \Sigma^{-1/2} z\|_{\Sigma}^2 + \lambda \|\theta\|_1 \right\}, \quad (4.2.9)$$

with  $\|v\|_{\Sigma}^2 \equiv \langle v, \Sigma v \rangle$  and  $z \sim \mathbf{N}(0, \mathbf{I}_{p \times p})$  admits a differentiable limit  $\mathfrak{E}(a, b)$  on  $\mathbb{R}_+ \times \mathbb{R}_+$ , with  $\nabla \mathfrak{E}^{(p)}(a, b) \rightarrow \nabla \mathfrak{E}(a, b)$ . Then the sequence has a standard distributional limit. Further let

$$\eta_b(Y) \equiv \arg \min_{\theta \in \mathbb{R}^p} \left\{ \frac{b}{2} \|\theta - Y\|_{\Sigma}^2 + \lambda \|\theta\|_1 \right\}, \quad (4.2.10)$$

$$\mathbf{E}_1(a, b) \equiv \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \left\{ \|\eta_b(\theta_0 + \sqrt{a} \Sigma^{-1/2} z) - \theta_0\|_{\Sigma}^2 \right\}, \quad (4.2.11)$$

<sup>1</sup>In Appendix B.3 we derive indeed a more general result, where the  $\ell_1$  regularization is replaced by an arbitrary separable penalty.

where the limit exists by the above assumptions on the convergence of  $\mathfrak{E}^{(p)}(a, b)$ . Then, the parameters  $\tau$  and  $\mathbf{d}$  of the standard distributional limit are obtained by setting  $\mathbf{d} = (1 - \widehat{\theta}/n)^{-1}$  and solving the following with respect to  $\tau^2$ :

$$\tau^2 = \sigma_0^2 + \frac{1}{\delta} \mathbf{E}_1(\tau^2, 1/\mathbf{d}). \quad (4.2.12)$$

In other words, the replica method indicates that the standard distributional limit holds for a large class of non-diagonal covariance structures  $\Sigma$ . It is worth stressing that convergence assumption for the sequence  $\mathfrak{E}^{(p)}(a, b)$  is quite mild, and is satisfied by a large family of covariance matrices. For instance, it can be proved that it holds for block-diagonal matrices  $\Sigma$  as long as the blocks have bounded length and the blocks empirical distribution converges.

The replica method is a non-rigorous but highly sophisticated calculation procedure that has proved successful in a number of very difficult problems in probability theory and probabilistic combinatorics. Attempts to make the replica method rigorous have been pursued over the last 30 years by some world-leading mathematicians [126, 107, 60, 2]. This effort achieved spectacular successes, but so far does not provide tools to prove the above replica claim. In particular, the rigorous work mainly focuses on ‘i.i.d. randomness’, corresponding to the case covered by Theorem 4.1.4.

Over the last ten years, the replica method has been used to derive a number of fascinating results in information theory and communications theory, see e.g. [127, 63, 128, 21, 145]. More recently, several groups used it successfully in the analysis of high-dimensional sparse regression under standard Gaussian designs [109, 77, 61, 145, 125, 131, 76]. The rigorous analysis of ours and other groups [100, 8, 6, 106] subsequently confirmed these heuristic calculations in several cases.

There is a fundamental reason that makes establishing the standard distributional limit a challenging task. This requires in fact to characterize the distribution of the estimator (1.1.7) in a regime where the standard deviation of  $\widehat{\theta}_i$  is of the same order as its mean. Further,  $\widehat{\theta}_i$  does not converge to the true value  $\theta_{0,i}$ , hence making perturbative arguments ineffective.

#### 4.2.4 Covariance estimation

So far we assumed that the design covariance  $\Sigma$  is known. This setting is relevant for semi-supervised learning applications, where the data analyst has access to a large number  $N \gg p$  of ‘unlabeled examples’. These are i.i.d. feature vectors  $U_1, U_2, \dots, U_N$  with  $U_1 \sim \mathbf{N}(0, \Sigma)$  distributed as  $X_1$ , for which the response variable  $Y_i$  is not available. In this case  $\Sigma$  can be estimated accurately by  $N^{-1} \sum_{i=1}^n u_i u_i^\top$ . We refer to [29] for further background on such applications.

In other applications,  $\Sigma$  is unknown and no additional data is available. In this case we proceed as follows:

1. We estimate  $\Sigma$  from the design matrix  $X$  (equivalently, from the feature vectors  $X_1, X_2, \dots, X_n$ ). We let  $\hat{\Sigma}$  denote the resulting estimate.
2. We use  $\hat{\Sigma}$  instead of  $\Sigma$  in step 3 of our hypothesis testing procedure.

The problem of estimating covariance matrices in high-dimensional setting has attracted considerable attention in the past. Several estimation methods provide a consistent estimate  $\hat{\Sigma}$ , under suitable structural assumptions on  $\Sigma$ . For instance if  $\Sigma^{-1}$  is sparse, one can apply the graphical model method of [95], the regression approach of [113], or CLIME estimator [20], to name a few.

Since the covariance estimation problem is not the focus here, we will test the above approach using a very simple covariance estimation method. Namely, we assume that  $\Sigma$  is sparse and estimate it by thresholding the empirical covariance. A detailed description of this estimator is given in [Algorithm 6](#). We refer to [13] for a theoretical analysis of this type of methods. Note that the Lasso is unlikely to perform well if the columns of  $\mathbf{X}$  are highly correlated and hence the assumption of sparse  $\Sigma$  is very natural. On the other hand, we would like to emphasize that this covariance thresholding estimation is only one among many possible approaches.

In our numerical experiments in [Chapter 5](#), we use the estimated covariance returned by [Algorithm 6](#). As shown there, computed p-values appear to be fairly robust with respect to errors in the estimation of  $\Sigma$ . It would be interesting to develop a rigorous analysis of `SDL-TEST` that accounts for the covariance estimation error.

---

**Algorithm 6** Subroutine for estimating covariance  $\Sigma$

---

**Input:** Design matrix  $\mathbf{X}$

**Output:** Estimate  $\hat{\Sigma}$

- 1: Let  $C = (1/n)\mathbf{X}^\top \mathbf{X} \in \mathbb{R}^{p \times p}$ .
- 2: Let  $\sigma_1$  be the empirical variance of the entries in  $S$  and let  $\mathcal{A} = \{C_{ij} : |C_{ij}| \leq 3\sigma_1\}$ .
- 3: Let  $\sigma_2$  be the variance of entries in  $\mathcal{A}$ .
- 4: Construct  $\hat{C}$  as follows:

$$\hat{C}_{ij} = C_{ij} \mathbb{I}(|C_{ij}| \geq 3\sigma_2). \quad (4.2.13)$$

- 5: Denote by  $\zeta_1$  and  $\zeta_2$  the smallest and the smallest positive eigenvalues of  $\hat{C}$  respectively.
- 6: Set

$$\hat{\Sigma} = \hat{C} + (\zeta_2 - \zeta_1)\mathbf{I}. \quad (4.2.14)$$


---

### 4.3 Role of the factor $\mathbf{d}$

It is worth stressing one subtle, yet interesting, difference between debiasing methods of the Lasso stated in the previous chapter, with the one of the present chapter, i.e. (4.2.2). In all of these approaches, a debiased estimator is constructed of the form (2.3.2). However:

- The approach of [149, 134] sets  $M$  to be an estimate of  $\Sigma^{-1}$ . In the idealized situation where  $\Sigma$  is known, this construction reduces to setting  $M = \Sigma^{-1}$ .
- By contrast, prescription (4.2.2) amounts to setting  $M = \mathbf{d}\Sigma^{-1}$ , with  $\mathbf{d} = (1 - \|\hat{\theta}\|_0/n)^{-1}$ . In other words, we choose  $M$  as a *scaled version of the inverse covariance*.

The mathematical reason for the specific scaling factor is elucidated by the proof of Theorem 4.1.4 in [8]. Here we limit ourselves to illustrating through numerical simulations that this factor is indeed crucial to ensure the normality of  $(\hat{\theta}_i^u - \theta_{0,i})$  in the regime  $n = \Theta(s_0 \log(p/s_0))$ .

We consider a setup where the rows of the design matrix are generated independently from  $\mathbf{N}(0, \Sigma)$  with  $\Sigma_{jk}$  (for  $j \leq k$ ) given by

$$\Sigma_{jk} = \begin{cases} 1 & \text{if } k = j, \\ 0.1 & \text{if } k \in \{j+1, \dots, j+5\} \text{ or } k \in \{j+p-5, \dots, j+p-1\}, \\ 0 & \text{for all other } j \leq k. \end{cases} \quad (4.3.1)$$

Elements below the diagonal are given by the symmetry condition  $\Sigma_{kj} = \Sigma_{jk}$ . (Notice that this is a circulant matrix.)

We fix undersampling ratio  $\delta = n/p$  and sparsity level  $\epsilon = s_0/p$  and consider values  $p \in \{250, 500, 750, \dots, 3500\}$ . We also take active sets  $S_0$  with  $|S_0| = s_0$  chosen uniformly at random from the index set  $\{1, \dots, p\}$  and set  $\theta_{0,i} = 0.15$  for  $i \in S$ .

The goal is to illustrate the effect of the scaling factor  $\mathbf{d}$  on the empirical distribution of  $(\hat{\theta}_i^u - \theta_{0,i})$ , for large  $n, p, s_0$ . As we will see, the effect becomes more pronounced as the ratio  $n/s_0 = \delta/\epsilon$  (i.e. the number of samples per non-zero coefficient) becomes smaller. As above, we use  $\hat{\theta}^u$  for the unbiased estimator (4.2.2) (which amounts to Eq. (2.3.2) with  $M = \mathbf{d}\Sigma^{-1}$ ). We will use  $\hat{\theta}^{\mathbf{d}=1}$  for the ‘ideal’ unbiased estimator corresponding to the proposal of [149, 134] (which amounts to Eq. (2.3.2) with  $M = \Sigma^{-1}$ ).

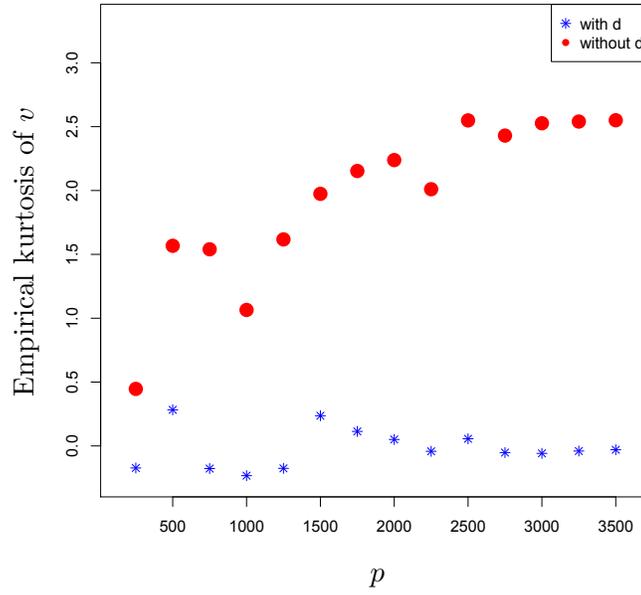
- **Small  $n/s_0$**  : We set  $\epsilon = 0.2$ ,  $\delta = 0.6$  (and thus  $n = 3s_0$ ). Let  $v = (v_i)_{i=1}^p$  with  $v_i \equiv (\hat{\theta}_i - \theta_{0,i})/(\tau[(\Sigma^{-1})_{ii}]^{1/2})$ . In Fig 4.3.1(a), the empirical kurtosis<sup>2</sup> of  $\{v_i\}_{i=1}^p$  is plotted for the two cases  $\hat{\theta}_i = \hat{\theta}_i^u$ , and  $\hat{\theta}_i = \hat{\theta}_i^{\mathbf{d}=1}$ . When using  $\hat{\theta}^u$ , the kurtosis is very small and data are consistent with the kurtosis vanishing as  $p \rightarrow \infty$ . This is suggestive of the fact that  $(\hat{\theta}_i^u - \theta_{0,i})/(\tau[(\Sigma^{-1})_{ii}]^{1/2})$  is asymptotically Gaussian, and hence satisfies a standard distributional limit. However, if we use  $\hat{\theta}^{\mathbf{d}=1}$ , the empirical kurtosis of  $v$  does not converge to zero.

In Fig. 4.3.2, we plot the histogram of  $v$  for  $p = 3000$  and using both  $\hat{\theta}^u$  and  $\hat{\theta}^{\mathbf{d}=1}$ . Again, the plots clearly demonstrate importance of  $\mathbf{d}$  in obtaining a Gaussian behavior.

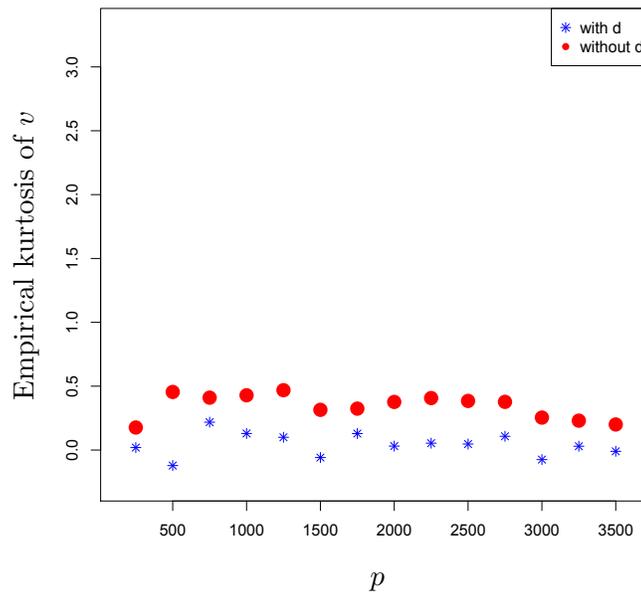
- **Large  $n/s_0$**  : We set  $\epsilon = 0.02$ ,  $\delta = 0.6$  (and thus  $n = 30s_0$ ). Figures 4.3.1(b) and 4.3.3 show similar plots for this case. As we see, the effect of  $\mathbf{d}$  becomes less noticeable here. The reason is that we expect  $\|\hat{\theta}\|_0/n = O(s_0/n)$ , and  $\mathbf{d} = (1 - \|\hat{\theta}\|_0/n)^{-1} = 1 + O(s_0/n) \approx 1$  for  $s_0$  much smaller than  $n$ .

---

<sup>2</sup>Recall that the empirical of sample kurtosis is defined as  $\kappa \equiv (m_4/m_2^2) - 3$  with  $m_\ell \equiv p^{-1} \sum_{i=1}^p (v_i - \bar{v})^\ell$  and  $\bar{v} \equiv p^{-1} \sum_{i=1}^p v_i$ .

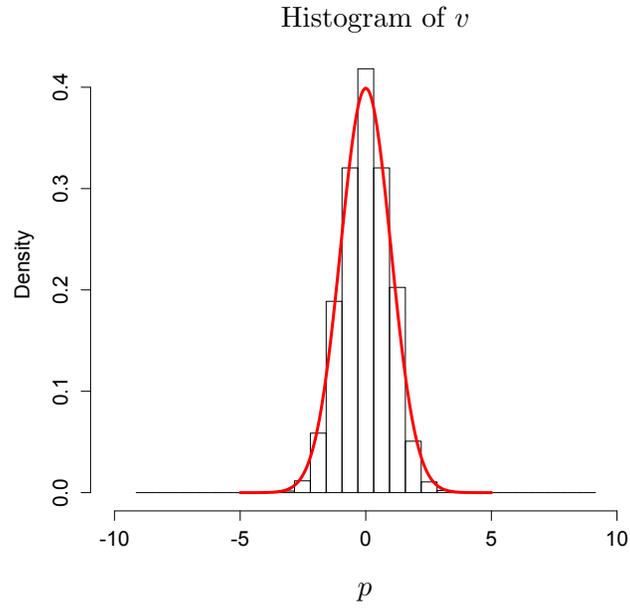


(a)  $\epsilon = 0.2$

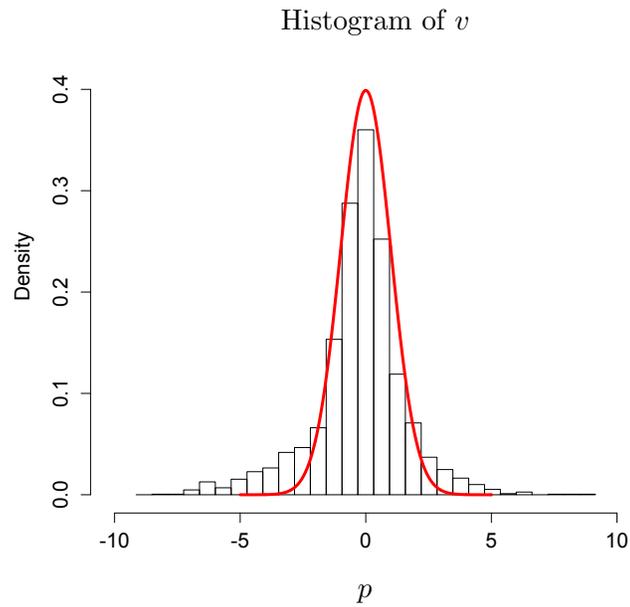


(b)  $\epsilon = 0.02$

Figure 4.3.1: Empirical kurtosis of vector  $v$  with and without normalization factor  $d$ . In left panel  $n = 3s_0$  (with  $\epsilon = 0.2$ ,  $\delta = 0.6$ ) and in the right panel  $n = 30s_0$  (with  $\epsilon = 0.02$ ,  $\delta = 0.6$ ).



(a) with factor  $d$



(b) without factor  $d$

Figure 4.3.2: Histogram of  $v$  for  $n = 3s_0$  ( $\epsilon = 0.2$ ,  $\delta = 0.6$ ) and  $p = 3000$ . In left panel, factor  $d$  is computed by Eq. (4.2.1) and in the right panel,  $d = 1$ .

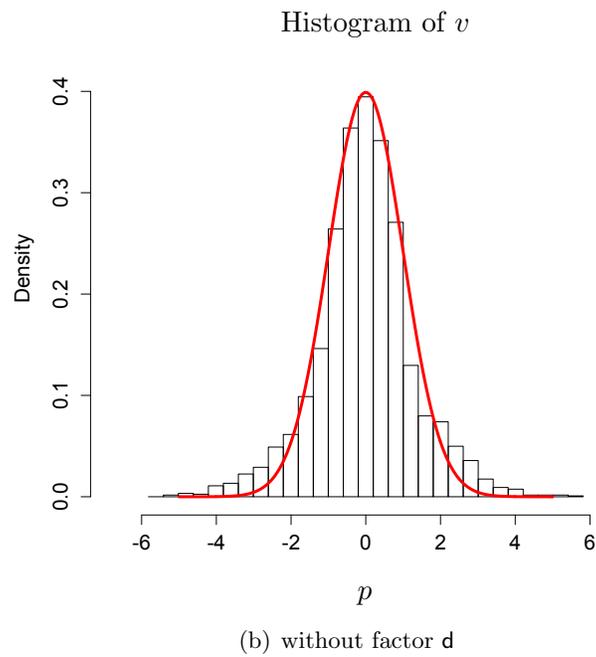
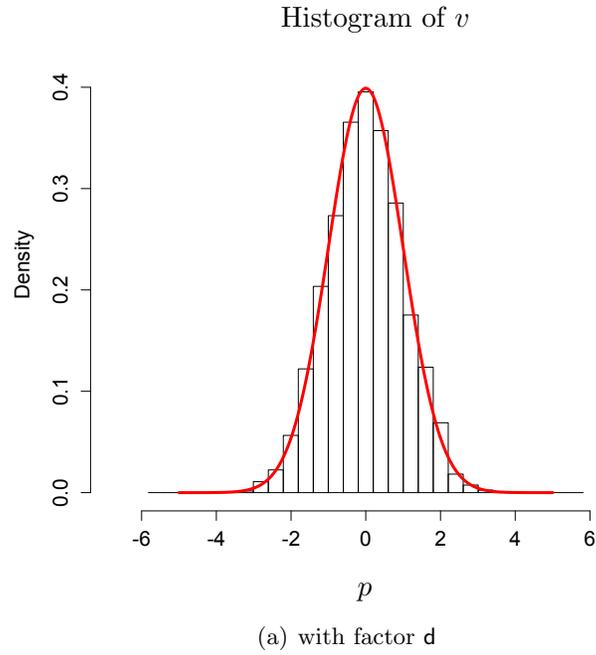


Figure 4.3.3: Histogram of  $v$  for  $n = 30 s_0$  ( $\epsilon = 0.02$ ,  $\delta = 0.6$ ) and  $p = 3000$ . In left panel, factor  $d$  is computed by Eq. (4.2.1) and in the right panel,  $d = 1$ .

## Chapter 5

# Numerical Validation

We compare performance of the debiasing approach [Algorithm 1](#) with the Ridge-type and the low-dimensional projection estimators (explained in [Section 3.4.2](#)), and the SDL-TEST test ([Algorithm 4](#) and [Algorithm 5](#)) on both synthetic and real data. We also validate the theoretical results about the type I error and power of testing procedure [\(3.0.3\)](#) ([Section 3.1](#)) and the results about the SDL-TEST ([Sections 4.1.2, 4.2.2](#)) by numerical simulations.

In the interest of reproducibility, an R implementation of [Algorithm 1](#) is available at <http://www.stanford.edu/~montanar/sslasso/>. Recall that in step 4 of [Algorithm 1](#) we need to solve the following optimization for  $i \in [p]$ :

$$\begin{aligned} & \text{minimize} && m^\top \widehat{\Sigma} m \\ & \text{subject to} && \|\widehat{\Sigma} m - e_i\|_\infty \leq \mu, \end{aligned} \tag{5.0.1}$$

These optimizations can be solved in parallel, since they are decoupled. Before proceeding to the numerical experiments, we will briefly remark on this optimization.

### 5.1 Remark on optimization [\(5.0.1\)](#)

We recast the problem as follows:

$$\begin{aligned} & \text{minimize} && m^\top \widehat{\Sigma} m \\ & \text{subject to} && \langle z, \widehat{\Sigma} m - e_i \rangle \leq \mu, \\ & && \|z\|_1 = 1. \end{aligned}$$

The Lagrangian is then given by

$$\mathcal{L}(\nu, m, z) = m^\top \widehat{\Sigma} m + \nu \left( \langle z, \widehat{\Sigma} m - e_i \rangle - \mu \right), \quad \|z\|_1 = 1.$$

Optimizing over  $m$ , we get

$$\frac{\partial \mathcal{L}}{\partial m} = 2\widehat{\Sigma} m_* + \nu \widehat{\Sigma} z = 0,$$

with  $m_*$  the optimizer. Therefore,  $m_* = -\nu z/2$ . The dual problem is then given by

$$\begin{aligned} & \text{maximize} && -\frac{\nu^2}{4} z^\top \widehat{\Sigma} z - \nu \langle z, e_i \rangle - \nu \mu \\ & \text{subject to} && \|z\|_1 = 1. \end{aligned}$$

We let  $\beta = -m_* = \nu z/2$ . Hence,  $\|\beta\|_1 = \nu/2$  since  $\|z\|_1 = 1$ . Rewriting the dual problem in terms of  $\beta$ , we arrive at:

$$\underset{\beta \in \mathbb{R}^p}{\text{minimize}} \quad \frac{1}{2} \beta^\top \widehat{\Sigma} \beta + \langle \beta, e_i \rangle + \mu \|\beta\|_1 \quad (5.1.1)$$

We use a line search to find the smallest value of  $\mu$  that makes (5.0.1) feasible.

For a fixed  $\mu$ , we solve the dual problem (5.1.1) by coordinate descent. Write (5.1.1) as

$$\frac{1}{2} \widehat{\Sigma}_{j,j} \beta_j^2 + \widehat{\Sigma}_{j,\sim j} \beta_{\sim j} \beta_j + \sum_{\ell, k \neq j} \widehat{\Sigma}_{\ell, k} \beta_\ell \beta_k + \beta_i + \mu \|\beta_{\sim j}\|_1 + \mu |\beta_j|,$$

where  $\widehat{\Sigma}_{j,\sim j}$  and  $\beta_{\sim j}$  respectively denote the  $j$ -th row of  $\widehat{\Sigma}$  with  $\widehat{\Sigma}_{j,j}$  removed, and the restriction of  $\beta$  to coordinates other than  $j$ . Minimizing w.r.t.  $\beta_j$ , we get

$$\beta_j + \frac{\widehat{\Sigma}_{j,\sim j} \beta_{\sim j} + \delta_{ij}}{\widehat{\Sigma}_{j,j}} + \left( \frac{\mu}{\widehat{\Sigma}_{j,j}} \right) \xi = 0,$$

where  $\delta_{ij} = \mathbb{I}(i = j)$  and  $\xi \in \partial |\beta_j|$ , with  $\partial |\cdot|$  denoting the subdifferential of  $|\cdot|$  at  $\beta_j$ .

It is easy to see that  $\beta_j$  is given by

$$\beta_j \longleftarrow \frac{1}{\widehat{\Sigma}_{j,j}} \eta(-\widehat{\Sigma}_{j,\sim j} \beta_{\sim j} - \delta_{ij}; \mu), \quad (5.1.2)$$

where  $\eta(\cdot)$  is the soft thresholding function defined by (2.2.1). The update (5.1.2) is repeated for  $j = 1, 2, \dots, p, 1, 2, \dots, p, \dots$  until convergence.

## 5.2 Experiment 1

In this experiment, we compare [Algorithm 1](#), multisample splitting (cf. [Section 3.4.1](#)), and the Ridge-type projection estimator (cf. [Section 3.4.2](#)) on synthetic data.

Consider linear model [\(1.3.2\)](#), where the rows of design matrix  $\mathbf{X}$  are fixed i.i.d. realizations from  $\mathbf{N}(0, \Sigma)$ , where  $\Sigma \in \mathbb{R}^{p \times p}$  is a circulant symmetric matrix with entries  $\Sigma_{jk}$  given as follows for  $j \leq k$ :

$$\Sigma_{jk} = \begin{cases} 1 & \text{if } k = j, \\ 0.1 & \text{if } k \in \{j + 1, \dots, j + 5\} \\ & \text{or } k \in \{j + p - 5, \dots, j + p - 1\}, \\ 0 & \text{for all other } j \leq k. \end{cases} \quad (5.2.1)$$

Regarding the regression coefficient, we consider a uniformly random support  $S \subseteq [p]$ , with  $|S| = s_0$  and let  $\theta_{0,i} = b$  for  $i \in S$  and  $\theta_{0,i} = 0$  otherwise. The measurement errors are  $W_i \sim \mathbf{N}(0, 1)$ , for  $i \in [n]$ . We consider several configurations of  $(n, p, s_0, b)$  and for each configuration report our results based on 20 independent realizations of the model with fixed design and fixed regression coefficients. In other words, we repeat experiments over 20 independent realization of the measurement errors.

We use the regularization parameter  $\lambda = 4\hat{\sigma}\sqrt{(2\log p)/n}$ , where  $\hat{\sigma}$  is given by the scaled Lasso [\(2.4.3\)](#) with  $\tilde{\lambda} = 10\sqrt{(2\log p)/n}$ . Furthermore, parameter  $\mu$  (cf. [Eq. \(2.3.11\)](#)) is set to  $\mu = 2\sqrt{(\log p)/n}$ . This choice of  $\mu$  is guided by [Theorem 2.3.4 \(b\)](#).

Throughout, we set the significance level  $\alpha = 0.05$ .

**Confidence intervals.** For each configuration, we consider 20 independent realizations of measurement noise and for each parameter  $\theta_{0,i}$ , we compute the average length of the corresponding confidence interval, denoted by  $\text{Avglength}(J_i(\alpha))$  where  $J_i(\alpha)$  is given by [equation \(2.4.4\)](#) and the average is taken over the realizations. We then define

$$\ell \equiv p^{-1} \sum_{i \in [p]} \text{Avglength}(J_i(\alpha)). \quad (5.2.2)$$

We also consider the average length of intervals for the active and inactive parameters, as

Measure \ Configuration	$\ell$	$\ell_S$	$\ell_{S^c}$	$\widehat{\text{Cov}}$	$\widehat{\text{Cov}}_S$	$\widehat{\text{Cov}}_{S^c}$
(1000, 600, 10, 0.5)	0.1870	0.1834	0.1870	0.9766	0.9600	0.9767
(1000, 600, 10, 0.25)	0.1757	0.1780	0.1757	0.9810	0.9000	0.9818
(1000, 600, 10, 0.1)	0.1809	0.1823	0.1809	0.9760	1	0.9757
(1000, 600, 30, 0.5)	0.2107	0.2108	0.2107	0.9780	0.9866	0.9777
(1000, 600, 30, 0.25)	0.1956	0.1961	0.1956	0.9660	0.9660	0.9659
(1000, 600, 30, 0.1)	0.2023	0.2043	0.2023	0.9720	0.9333	0.9732
(2000, 1500, 50, 0.5)	0.1383	0.1391	0.1383	0.9754	0.9800	0.9752
(2000, 1500, 50, 0.25)	0.1356	0.1363	0.1355	0.9720	0.9600	0.9723
(2000, 1500, 50, 0.1)	0.1361	0.1361	0.1361	0.9805	1	0.9800
(2000, 1500, 25, 0.5)	0.1233	0.1233	0.1233	0.9731	0.9680	0.9731
(2000, 1500, 25, 0.25)	0.1208	0.1208	0.1208	0.9735	1	0.9731
(2000, 1500, 25, 0.1)	0.1242	0.1237	0.1242	0.9670	0.9200	0.9676

Table 5.2.1 Simulation results for the synthetic data described in Experiment 1. The results corresponds to 95% confidence intervals.

follows:

$$\ell_S \equiv s_0^{-1} \sum_{i \in S} \text{Avglength}(J_i(\alpha)), \quad \ell_{S^c} \equiv (p - s_0)^{-1} \sum_{i \in S^c} \text{Avglength}(J_i(\alpha)). \quad (5.2.3)$$

Similarly, we consider average coverage for individual parameters. We define the following three metrics:

$$\widehat{\text{Cov}} \equiv p^{-1} \sum_{i \in [p]} \widehat{\mathbb{P}}[\theta_{0,i} \in J_i(\alpha)], \quad (5.2.4)$$

$$\widehat{\text{Cov}}_S \equiv s_0^{-1} \sum_{i \in S} \widehat{\mathbb{P}}[\theta_{0,i} \in J_i(\alpha)], \quad (5.2.5)$$

$$\widehat{\text{Cov}}_{S^c} \equiv (p - s_0)^{-1} \sum_{i \in S^c} \widehat{\mathbb{P}}[0 \in J_i(\alpha)], \quad (5.2.6)$$

where  $\widehat{\mathbb{P}}$  denotes the empirical probability computed based on the 20 realizations for each configuration. The results are reported in Table 5.2.1. In Fig. 5.2.1, we plot the constructed 95%-confidence intervals for one realization of configuration  $(n, p, s_0, b) = (1000, 600, 10, 1)$ . For sake of clarity, we plot the confidence intervals for only 100 of the 1000 parameters.

**False positive rates and statistical powers.** Table 5.2.2 summarizes the false positive rates and the statistical powers achieved by our proposed method, the multisample-splitting

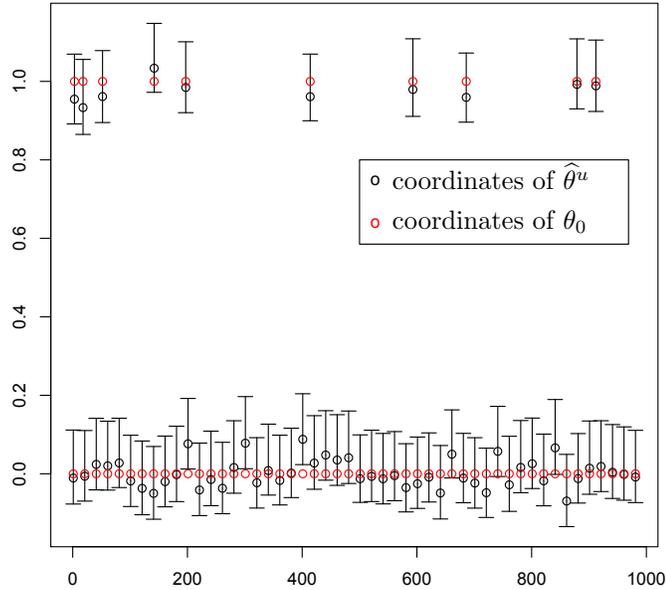


Figure 5.2.1: 95% confidence intervals for one realization of configuration  $(n, p, s_0, b) = (1000, 600, 10, 1)$ . For clarity, we plot the confidence intervals for only 100 of the 1000 parameters. The true parameters  $\theta_{0,i}$  are in red and the coordinates of the debiased estimator  $\hat{\theta}^u$  are in black.

method [96], and the Ridge-type projection estimator [16] for several configurations. The results are obtained by taking average over 20 independent realizations of measurement errors for each configuration. As we see the multisample-splitting achieves false positive rate 0 on all of the configurations considered here, making no type I error. However, the true positive rate is always smaller than that of our proposed method. By contrast, our method achieves false positive rate close to the pre-assigned significance level  $\alpha = 0.05$  and obtains much higher true positive rate. Similar to the multisample-splitting, the Ridge-type projection estimator is conservative and achieves false positive rate smaller than  $\alpha$ . This, however, comes at the cost of a smaller true positive rate than our method. It is worth noting that an ideal testing procedure should allow to control the level of statistical significance  $\alpha$ , and obtain the maximum true positive rate at that level.

Here, we used the R-package `hdi` to test multisample-splitting and the Ridge-type projection estimator.

Let  $Z = (z_i)_{i=1}^p$  denote the vector with  $z_i \equiv \sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})/\hat{\sigma}\sqrt{[M\hat{\Sigma}M^T]_{i,i}}$ . Fig. 5.2.2 shows the sample quantiles of  $Z$  versus the quantiles of the standard normal distribution for

Configuration	Our method		Multisample-splitting		Ridge-type projection estimator	
	FP	TP	FP	TP	FP	TP
(1000, 600, 10, 0.5)	0.0452	1	0	1	0.0284	0.8531
(1000, 600, 10, 0.25)	0.0393	1	0	0.4	0.02691	0.7506
(1000, 600, 10, 0.1)	0.0383	0.8	0	0	0.2638	0.6523
(1000, 600, 30, 0.5)	0.0433	1	0	1	0.0263	0.8700
(1000, 600, 30, 0.25)	0.0525	1	0	0.4	0.2844	0.8403
(1000, 600, 30, 0.1)	0.0402	0.7330	0	0	0.2238	0.6180
(2000, 1500, 50, 0.5)	0.0421	1	0	1	0.0301	0.9013
(2000, 1500, 50, 0.25)	0.0415	1	0	1	0.0292	0.8835
(2000, 1500, 50, 0.1)	0.0384	0.9400	0	0	0.02655	0.7603
(2000, 1500, 25, 0.5)	0.0509	1	0	1	0.0361	0.9101
(2000, 1500, 25, 0.25)	0.0481	1	0	1	0.3470	0.8904
(2000, 1500, 25, 0.1)	0.0551	1	0	0.16	0.0401	0.8203

Table 5.2.2 Simulation results for the synthetic data described in Experiment 1. The false positive rates (FP) and the true positive rates (TP) are computed at significance level  $\alpha = 0.05$ .

one realization of the configuration  $(n, p, s_0, b) = (1000, 600, 10, 1)$ . The scattered points are close to the line with unit slope and zero intercept. This confirms the result of Theorem 2.4.2 regarding the gaussianity of the entries  $z_i$ .

For the same problem, in Fig. 5.2.3 we plot the empirical CDF of the computed  $p$ -values restricted to the variables outside the support. Clearly, the  $p$ -values for these entries are uniformly distributed as expected.

### 5.3 Experiment 2

In this experiment, we compare Algorithm 1, multisample splitting (cf. Section 3.4.1), and the Ridge-type projection estimator (cf. Section 3.4.2) on real data.

We consider a high-throughput genomic data set concerning riboflavin (vitamin  $B_2$ ) production rate. This data set is made publicly available by [17] and contains  $n = 71$  samples and  $p = 4,088$  covariates corresponding to  $p = 4,088$  genes. For each sample, there is a real-valued response variable indicating the logarithm of the riboflavin production rate along with the logarithm of the expression level of the  $p = 4,088$  genes as the covariates.

Following [17], we model the riboflavin production rate as a linear model with  $p = 4,088$  covariates and  $n = 71$  samples, as in Eq. (1.1.6). We use the R package `glmnet` [56] to fit the Lasso estimator. Similar to the previous section, we use the regularization parameter

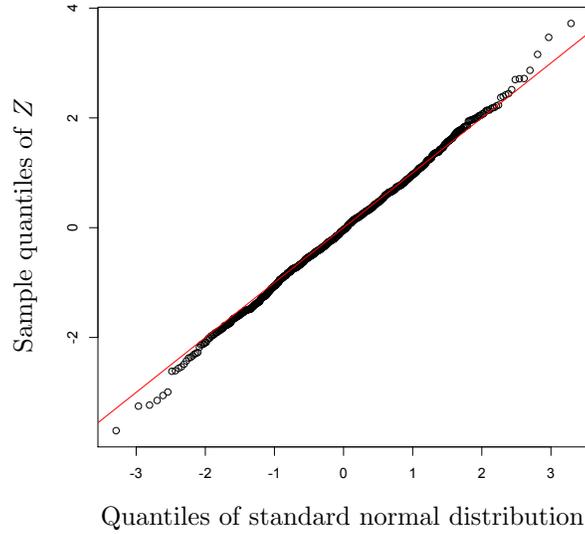


Figure 5.2.2: Q-Q plot of  $Z$  for one realization of configuration  $(n, p, s_0, b) = (1000, 600, 10, 1)$ .

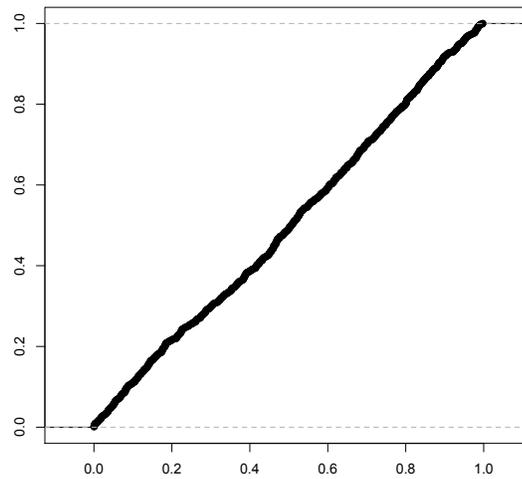


Figure 5.2.3: Empirical CDF of the computed  $p$ -values (restricted to entries outside the support) for one realization of configuration  $(n, p, s_0, b) = (1000, 600, 10, 1)$ . Clearly the plot confirms that the  $p$ -values are distributed according to uniform distribution.

$\lambda = 4\hat{\sigma}\sqrt{(2\log p)/n}$ , where  $\hat{\sigma}$  is given by the scaled Lasso as per equation (2.4.3) with  $\tilde{\lambda} = 10\sqrt{(2\log p)/n}$ . This leads to the choice  $\lambda = 0.036$ . The resulting model contains 30 genes (plus an intercept term) corresponding to the nonzero parameters of the lasso estimator.

We use Eq. (3.0.2) to construct  $p$ -values for different genes. Adjusting FWER to 5% significance level, we find two significant genes, namely genes YXLD-at and YXLE-at. By contrast, the multisample-splitting method proposed in [96] finds only the gene YXLD-at at the FWER-adjusted 5% significance level. Also the Ridge-type projection estimator, proposed in [16], returns no significance gene. (See [17] for further discussion on these methods.) This indicates that these methods are more conservative and produce typically larger  $p$ -values.

In Fig. 5.3.4 we plot the empirical CDF of the computed  $p$ -values for riboflavin example. Clearly the plot confirms that the  $p$ -values are distributed according to uniform distribution.

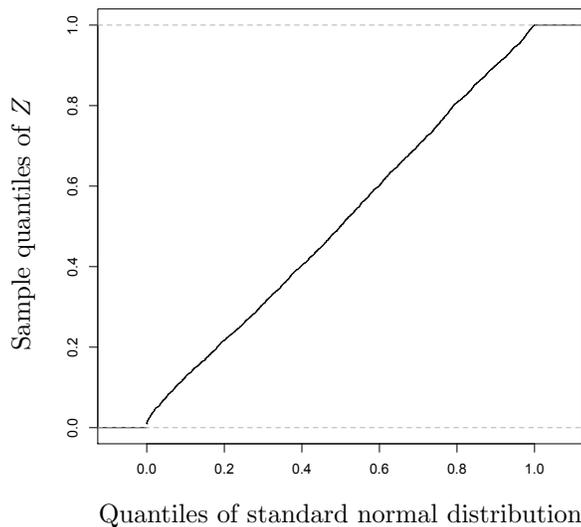


Figure 5.3.4: Empirical CDF of the computed  $p$ -values for riboflavin example. Clearly the plot confirms that the  $p$ -values are distributed according to uniform distribution.

## 5.4 Experiment 3

The goal of this experiment is to compare the SDL-TEST test (cf. Algorithm 4) with the Ridge-type projection estimator [16] and the low-dimensional projection estimator

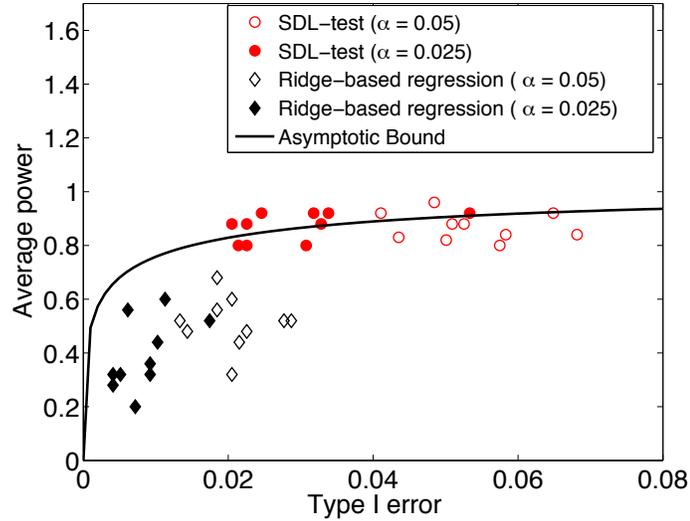


Figure 5.4.5: Comparison between SDL-TEST (Algorithm 4), Ridge-based regression [16] and the asymptotic bound for SDL-TEST (established in Theorem 4.1.3). Here,  $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ .

(LDPE) [149] (cf. Section Subsection 3.4.2) on synthetic data and standard Gaussian design. we further corroborate the theoretical results of Section 4.1.2 by simulation.

We generate synthetic data from the linear model (1.1.6) with  $W \sim N(0, I_{p \times p})$  and the following configurations.

*Design matrix:* For pairs of values  $(n, p) = \{(300, 1000), (600, 1000), (600, 2000)\}$ , the design matrix is generated from a realization of  $n$  i.i.d. rows  $X_i \sim N(0, I_{p \times p})$ .

*Regression parameters:* We consider active sets  $S_0$  with  $|S_0| = s_0 \in \{10, 20, 25, 50, 100\}$ , chosen uniformly at random from the index set  $\{1, \dots, p\}$ . We also consider two different strengths of active parameters  $\theta_{0,i} = \gamma$ , for  $i \in S_0$ , with  $\gamma \in \{0.1, 0.15\}$ .

We examine the performance of SDL-TEST (cf. Algorithm 4) at significance levels  $\alpha = 0.025, 0.05$ . The experiments are done using glmnet-package in R that fits the entire Lasso path for linear regression models. Let  $\varepsilon = s_0/p$  and  $\delta = n/p$ . We do not assume  $\varepsilon$  is known, but rather estimate it as  $\bar{\varepsilon} = 0.25 \delta / \log(2/\delta)$ . The value of  $\bar{\varepsilon}$  is half the maximum sparsity level  $\varepsilon$  for the given  $\delta$  such that the Lasso estimator can correctly recover the parameter vector if the measurements were noiseless [42, 8]. Provided it makes sense to use Lasso at all,  $\bar{\varepsilon}$  is thus a reasonable ballpark estimate.

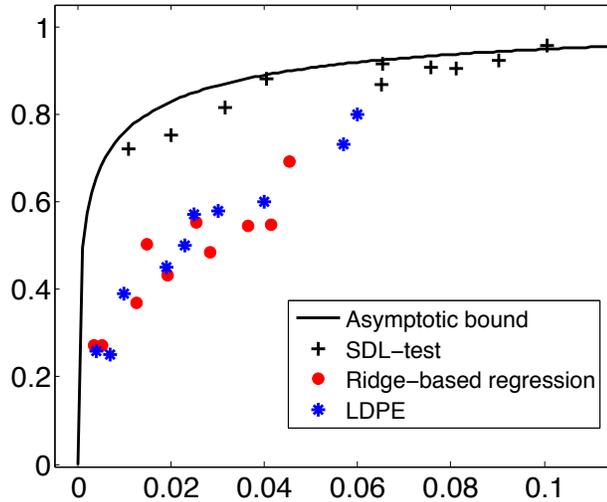


Figure 5.4.6: Comparison between SDL-TEST, Ridge-based regression [16], and LDPE [149]. The curve corresponds to the asymptotic bound for SDL-TEST as established in Theorem 4.1.3. For the same values of type I error achieved by methods, SDL-TEST results in a higher statistical power. Here,  $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ .

The regularization parameter  $\lambda$  is chosen as to satisfy

$$\lambda d = \kappa_* \tau \quad (5.4.1)$$

where  $\tau$  and  $d$  are determined in step 2 of the procedure. Here  $\kappa_* = \kappa_*(\bar{\varepsilon})$  is the minimax threshold value for estimation using soft thresholding in the Gaussian sequence model, see [44] and Remark B.2.1. Note that  $\tau$  and  $d$  in the equation above depend implicitly upon  $\lambda$ . Since `glmnet` returns the entire Lasso path, the value of  $\lambda$  solving the above equation can be computed by the bisection method.

As mentioned above, the control of type I error is fairly robust for a wide range of values of  $\lambda$ . However, the above is an educated guess based on the analysis of [42, 8]. We also tried the values of  $\lambda$  proposed for instance in [133, 16] on the basis of oracle inequalities.

Figure 5.4.5 shows the results of SDL-TEST and the method of [16] for parameter values  $p = 1000, n = 600, s_0 = 25, \gamma = 0.15$ , and significance levels  $\alpha \in \{0.025, 0.05\}$ . Each point in the plot corresponds to one realization of this configuration (there are a total of 10 realizations). We also depict the theoretical curve  $(\alpha, G(\alpha, \gamma_0/\tau_*))$ , predicted by Theorem 4.1.3. The empirical results are in good agreement with the asymptotic prediction.

We compare `SDL-TEST` with the Ridge-based regression method [16] and the low dimensional projection estimator (LDPE) [149]. Tables 5.4.1 and 5.4.2 summarize the results for a few configurations  $(p, n, s_0, \gamma)$ , and  $\alpha = 0.05, 0.025$ .

As demonstrated by these results, LDPE [149] and the Ridge-based regression [16] are both overly conservative. Namely, they achieve smaller type I error than the prescribed level  $\alpha$  and this comes at the cost of a smaller statistical power than our testing procedure. This is to be expected since the approach of [16] and [149] cover a broader class of design matrices  $\mathbf{X}$ , and are not tailored to random designs.

Note that being overly conservative is a drawback, when this comes at the expense of statistical power. The data analysts should be able to decide the level of statistical significance  $\alpha$ , and obtain optimal statistical power at that level.

The reader might wonder whether the loss in statistical power of methods in [16] and [149] is entirely due to the fact that these methods achieve a smaller number of false positives than requested. In Fig. 5.4.6, we run `SDL-TEST`, Ridge-based regression [16], and LDPE for  $\alpha \in \{0.01, 0.02, \dots, 0.1\}$  and for 10 realizations of the problem per each value of  $\alpha$ . We plot the average type I error and the average power of each method versus  $\alpha$ . As we see *even for the same empirical fraction of type I errors*, `SDL-TEST` results in a higher statistical power.

Method	Type I err (mean)	Type I err (std.)	Avg. power (mean)	Avg. power (std)
<b>SDL-test</b> (1000, 600, 50, 0.15)	0.06189	0.01663	0.83600	0.04300
<b>Ridge-based regression</b> (1000, 600, 50, 0.15)	<b>0.00989</b>	<b>0.00239</b>	<b>0.35000</b>	<b>0.07071</b>
<b>LDPE</b> (1000, 600, 50, 0.15)	<b>0.03925</b>	<b>0.00588</b>	<b>0.55302</b>	<b>0.07608</b>
<b>Asymptotic Bound</b> (1000, 600, 50, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.84721</b>	<b>NA</b>
<b>SDL-test</b> (1000, 600, 25, 0.15)	0.0572	0.0190	0.8840	0.0638
<b>Ridge-based regression</b> (1000, 600, 25, 0.15)	<b>0.0203</b>	<b>0.0052</b>	<b>0.3680</b>	<b>0.1144</b>
<b>LDPE</b> (1000, 600, 25, 0.15)	<b>0.04010</b>	<b>0.00917</b>	<b>0.62313</b>	<b>0.05408</b>
<b>Asymptotic Bound</b> (1000, 600, 25, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.9057</b>	<b>NA</b>
<b>SDL-test</b> (1000, 300, 50, 0.15)	0.05547	0.01554	0.45800	0.06957
<b>Ridge-based regression</b> (1000, 300, 50, 0.15)	<b>0.01084</b>	<b>0.00306</b>	<b>0.19200</b>	<b>0.04541</b>
<b>LDPE</b> (1000, 300, 50, 0.15)	<b>0.03022</b>	<b>0.00601</b>	<b>0.23008</b>	<b>0.08180</b>
<b>Asymptotic Bound</b> (1000, 300, 50, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.31224</b>	<b>NA</b>
<b>SDL-test</b> (1000, 300, 25, 0.15)	0.05149	0.01948	0.55600	0.11384
<b>Ridge-based regression</b> (1000, 300, 25, 0.15)	<b>0.00964</b>	<b>0.00436</b>	<b>0.32400</b>	<b>0.09324</b>
<b>LDPE</b> (1000, 300, 25, 0.15)	<b>0.04001</b>	<b>0.00531</b>	<b>0.34091</b>	<b>0.06408</b>
<b>Asymptotic Bound</b> (1000, 300, 25, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.51364</b>	<b>NA</b>
<b>SDL-test</b> (2000, 600, 100, 0.1)	0.05037	0.00874	0.44800	0.04940
<b>Ridge-based regression</b> (2000, 600, 100, 0.1)	<b>0.01232</b>	<b>0.00265</b>	<b>0.21900</b>	<b>0.03143</b>
<b>LDPE</b> (2000, 600, 100, 0.1)	<b>0.03012</b>	<b>0.00862</b>	<b>0.31003</b>	<b>0.06338</b>
<b>Asymptotic Bound</b> (2000, 600, 100, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.28324</b>	<b>NA</b>
<b>SDL-test</b> (2000, 600, 50, 0.1)	0.05769	0.00725	0.52800	0.08548
<b>Ridge-based regression</b> (2000, 600, 50, 0.1)	<b>0.01451</b>	<b>0.00303</b>	<b>0.27000</b>	<b>0.04137</b>
<b>LDPE</b> (2000, 600, 50, 0.1)	<b>0.03221</b>	<b>0.01001</b>	<b>0.35063</b>	<b>0.05848</b>
<b>Asymptotic Bound</b> (2000, 600, 50, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.46818</b>	<b>NA</b>
<b>SDL-test</b> (2000, 600, 20, 0.1)	0.05167	0.00814	0.58000	0.11595
<b>Ridge-based regression</b> (2000, 600, 20, 0.1)	<b>0.01879</b>	<b>0.00402</b>	<b>0.34500</b>	<b>0.09846</b>
<b>LDPE</b> (2000, 600, 20, 0.1)	<b>0.04021</b>	<b>0.00608</b>	<b>0.42048</b>	<b>0.08331</b>
<b>Asymptotic Bound</b> (2000, 600, 20, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.58879</b>	<b>NA</b>
<b>SDL-test</b> (2000, 600, 100, 0.15)	0.05368	0.01004	0.64500	0.05104
<b>Ridge-based regression</b> (2000, 600, 100, 0.15)	<b>0.00921</b>	<b>0.00197</b>	<b>0.30700</b>	<b>0.04877</b>
<b>LDPE</b> (2000, 600, 100, 0.15)	<b>0.02890</b>	<b>0.00493</b>	<b>0.58003</b>	<b>0.06338</b>
<b>Asymptotic Bound</b> (2000, 600, 100, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.54728</b>	<b>NA</b>
<b>SDL-test</b> (2000, 600, 20, 0.15)	0.04944	0.01142	0.89500	0.07619
<b>Ridge-based regression</b> (2000, 600, 20, 0.15)	<b>0.01763</b>	<b>0.00329</b>	<b>0.64000</b>	<b>0.08756</b>
<b>LDPE</b> (2000, 600, 20, 0.15)	<b>0.03554</b>	<b>0.005047</b>	<b>0.73560</b>	<b>0.04008</b>
<b>Asymptotic Bound</b> (2000, 600, 20, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.90608</b>	<b>NA</b>

Table 5.4.1 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the asymptotic bound for SDL-TEST (cf. Theorem 4.1.3) in Experiment 3. The significance level is  $\alpha = 0.05$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of  $p = 1000$ ,  $n = 600$ ,  $s_0 = 50$ ,  $\gamma = 0.1$ .

Method	Type I err (mean)	Type I err (std.)	Avg. power (mean)	Avg. power (std)
<b>SDL-test</b> (1000, 600, 50, 0.15)	0.02874	0.00546	0.75600	0.07706
<b>Ridge-based regression</b> (1000, 600, 50, 0.15)	<b>0.00379</b>	<b>0.00282</b>	<b>0.22800</b>	<b>0.06052</b>
<b>LDPE</b> (1000, 600, 100, 0.1)	0.01459	0.00605	0.41503	0.08482
<b>Asymptotic Bound</b> (1000, 600, 50, 0.15)	0.025	NA	0.77107	NA
<b>SDL-test</b> (1000, 600, 25, 0.15)	0.03262	0.00925	0.79200	0.04131
<b>Ridge-based regression</b> (1000, 600, 25, 0.15)	<b>0.00759</b>	<b>0.00223</b>	<b>0.28800</b>	<b>0.07729</b>
<b>LDPE</b> (1000, 600, 25, 0.15)	0.01032	0.00490	0.55032	0.07428
<b>Asymptotic Bound</b> (1000, 600, 25, 0.15)	0.025	NA	0.84912	NA
<b>SDL-test</b> (1000, 300, 50, 0.15)	0.02916	0.00924	0.36000	0.08380
<b>Ridge-based regression</b> (1000, 300, 50, 0.15)	<b>0.00400</b>	<b>0.00257</b>	<b>0.10800</b>	<b>0.05432</b>
<b>LDPE</b> (1000, 300, 50, 0.15)	0.01520	0.00652	0.25332	0.06285
<b>Asymptotic Bound</b> (1000, 300, 50, 0.15)	0.025	NA	0.22001	NA
<b>SDL-test</b> (1000, 300, 25, 0.15)	0.03005	0.00894	0.42400	0.08884
<b>Ridge-based regression</b> (1000, 300, 25, 0.15)	<b>0.00492</b>	<b>0.00226</b>	<b>0.21600</b>	<b>0.06310</b>
<b>LDPE</b> (1000, 300, 25, 0.15)	0.00881	0.00377	0.31305	0.05218
<b>Asymptotic Bound</b> (1000, 300, 25, 0.15)	0.025	NA	0.40207	NA
<b>SDL-test</b> (2000, 600, 100, 0.1)	0.03079	0.00663	0.33000	0.05033
<b>Ridge-based regression</b> (2000, 600, 100, 0.1)	<b>0.00484</b>	<b>0.00179</b>	<b>0.11200</b>	<b>0.03615</b>
<b>LDPE</b> (2000, 600, 100, 0.1)	0.01403	0.00970	0.24308	0.06041
<b>Asymptotic Bound</b> (2000, 600, 100, 0.1)	0.025	NA	0.19598	NA
<b>SDL-test</b> (2000, 600, 50, 0.1)	0.02585	0.00481	0.41200	0.06197
<b>Ridge-based regression</b> (2000, 600, 50, 0.1)	<b>0.00662</b>	<b>0.00098</b>	<b>0.20600</b>	<b>0.03406</b>
<b>LDPE</b> (2000, 600, 50, 0.1)	0.01601	0.00440	0.27031	0.03248
<b>Asymptotic Bound</b> (2000, 600, 50, 0.1)	0.025	NA	0.35865	NA
<b>SDL-test</b> (2000, 600, 20, 0.1)	0.02626	0.00510	0.47500	0.10607
<b>Ridge-based regression</b> (2000, 600, 20, 0.1)	<b>0.00838</b>	<b>0.00232</b>	<b>0.23500</b>	<b>0.08182</b>
<b>LDPE</b> (2000, 600, 20, 0.1)	0.02012	0.00628	0.34553	0.09848
<b>Asymptotic Bound</b> (2000, 600, 20, 0.1)	0.025	NA	0.47698	NA
<b>SDL-test</b> (2000, 600, 100, 0.15)	0.02484	0.00691	0.52700	0.09522
<b>Ridge-based regression</b> (2000, 600, 100, 0.15)	<b>0.00311</b>	<b>0.00154</b>	<b>0.22500</b>	<b>0.04007</b>
<b>LDPE</b> (2000, 600, 100, 0.15)	0.01482	0.00717	0.38405	0.03248
<b>Asymptotic Bound</b> (2000, 600, 100, 0.15)	0.025	NA	0.43511	NA
<b>SDL-test</b> (2000, 600, 20, 0.15)	0.03116	0.01304	0.81500	0.09443
<b>Ridge-based regression</b> (2000, 600, 20, 0.15)	<b>0.00727</b>	<b>0.00131</b>	<b>0.54500</b>	<b>0.09560</b>
<b>LDPE</b> (2000, 600, 20, 0.15)	0.01801	0.00399	0.68101	0.06255
<b>Asymptotic Bound</b> (2000, 600, 20, 0.15)	0.025	NA	0.84963	NA

Table 5.4.2 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the asymptotic bound for SDL-TEST (cf. Theorem 4.1.3) in Experiment 3. The significance level is  $\alpha = 0.025$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of  $p = 1000$ ,  $n = 600$ ,  $s_0 = 50$ ,  $\gamma = 0.1$

## 5.5 Experiment 4

In carrying out comparison between SDL-TEST, Ridge-type projection estimator [16] and the low-dimensional projection estimator (LDPE) [149] for correlated Gaussian designs, we consider the same setup as Experiment 3. The only difference is that the rows of the design matrix are independently  $X_i \sim \mathbf{N}(0, \Sigma)$ . We choose  $\Sigma \in \mathbb{R}^{p \times p}$  to be a the symmetric matrix with entries  $\Sigma_{jk}$  are defined as follows for  $j \leq k$

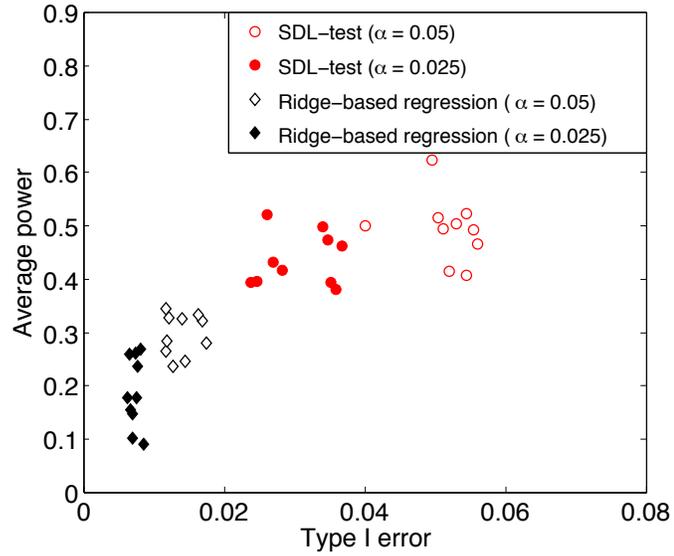
$$\Sigma_{jk} = \begin{cases} 1 & \text{if } k = j, \\ 0.1 & \text{if } k \in \{j + 1, \dots, j + 5\} \\ & \text{or } k \in \{j + p - 5, \dots, j + p - 1\}, \\ 0 & \text{for all other } j \leq k. \end{cases} \quad (5.5.1)$$

Elements below the diagonal are given by the symmetry condition  $\Sigma_{kj} = \Sigma_{jk}$ . (Notice that this is a circulant matrix.)

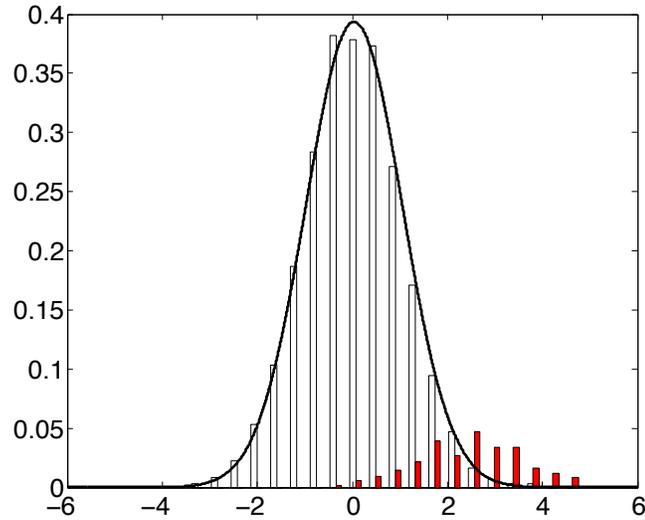
In Fig. 5.5.7(a), we compare SDL-TEST with the Ridge-based regression method proposed in [16]. While the type I errors of SDL-TEST are in good match with the chosen significance level  $\alpha$ , the method of [16] is conservative. As in the case of standard Gaussian designs, this results in significantly smaller type I errors than  $\alpha$  and smaller average power in return. Also, in Fig. 5.5.8, we run SDL-TEST, Ridge-based regression [16], and LDPE [149] for  $\alpha \in \{0.01, 0.02, \dots, 0.1\}$  and for 10 realizations of the problem per each value of  $\alpha$ . We plot the average type I error and the average power of each method versus  $\alpha$ . As we see, similar to the case of standard Gaussian designs, *even for the same empirical fraction of type I errors*, SDL-TEST results in a higher statistical power.

Tables 5.5.1 and 5.5.2 summarize the results of comparison among these methods for several configurations  $(p, n, s_0, \gamma)$ , and  $\alpha = 0.05, 0.025$ .

Let  $z = (z_i)_{i=1}^p$  denote the vector with entries  $z_i \equiv \hat{\theta}_i^u / (\tau[(\Sigma^{-1})_{ii}]^{1/2})$ . In Fig. 5.5.7(b) we plot the normalized histograms of  $z_{S_0}$  (in red) and  $z_{S_0^c}$  (in white), where  $z_{S_0}$  and  $z_{S_0^c}$  respectively denote the restrictions of  $z$  to the active set  $S_0$  and the inactive set  $S_0^c$ . The plot clearly exhibits the fact that  $z_{S_0^c}$  has (asymptotically) standard normal distribution and the histogram of  $z_{S_0}$  appears as a distinguishable bump. This is the core intuition in defining SDL-TEST.



(a) Comparison between SDL-TEST and Ridge-type projection estimator [16] (cf. Section 3.4.2).



(b) Normalized histograms of  $z_{S_0}$  (in red) and  $z_{S_0^c}$  (in white) for one realization.

Figure 5.5.7: Numerical results for Experiment 4 and  $p = 2000$ ,  $n = 600$ ,  $s_0 = 50$ ,  $\gamma = 0.1$ .

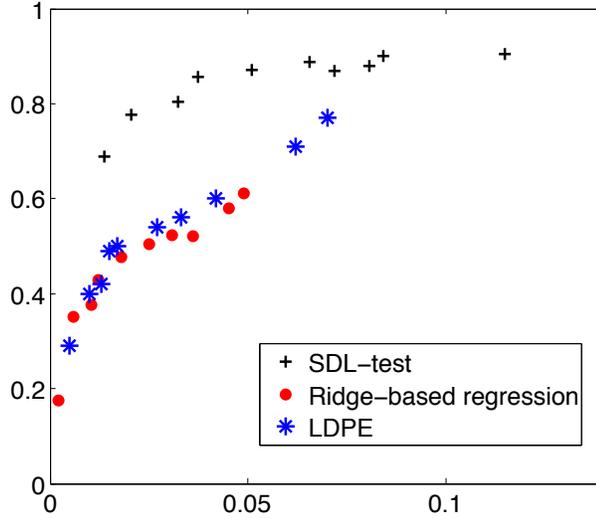


Figure 5.5.8: Comparison between SDL-TEST, Ridge-based regression [16], and LDPE [149] in the setting of Experiment 4. For the same values of type I error achieved by methods, SDL-TEST results in a higher statistical power. Here,  $p = 1000$ ,  $n = 600$ ,  $s_0 = 25$ ,  $\gamma = 0.15$ .

## 5.6 Experiment 5

In this experiment, we apply SDL-TEST and the Ridge-type projection estimator [16] on the UCI communities and crimes dataset [55]. This concerns the prediction of the rate of violent crime in different communities within US, based on other demographic attributes of the communities. The dataset consists of a response variable along with 122 predictive attributes for 1994 communities. Covariates are quantitative, including e.g., the fraction of urban population or the median family income. We consider a linear model as in (1.1.6) and hypotheses  $H_{0,i}$ . Rejection of  $H_{0,i}$  indicates that the  $i$ -th attribute is significant in predicting the response variable.

We perform the following preprocessing steps: (i) Each missing value is replaced by the mean of the non missing values of that attribute for other communities. (ii) We eliminate 16 attributes to make the ensemble of the attribute vectors linearly independent. Thus we obtain a design matrix  $\mathbf{X}_{\text{tot}} \in \mathbb{R}^{n_{\text{tot}} \times p}$  with  $n_{\text{tot}} = 1994$  and  $p = 106$ ; (iii) We normalize each column of the resulting design matrix to have mean zero and  $\ell_2$  norm equal to  $\sqrt{n_{\text{tot}}}$ .

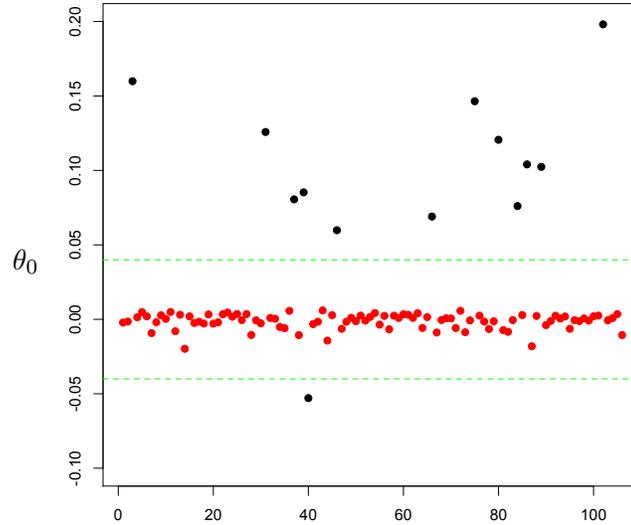
In order to evaluate various hypothesis testing procedures, we need to know the true significant variables. To this end, we let  $\theta_0 = (\mathbf{X}_{\text{tot}}^T \mathbf{X}_{\text{tot}})^{-1} \mathbf{X}_{\text{tot}}^T Y$  be the least-square estimator, using the whole data set. Figure 5.6.9 shows the the entries of  $\theta_0$ . Clearly, only

Method	Type I err (mean)	Type I err (std.)	Avg. power (mean)	Avg. power (std)
<b>SDL-test</b> (1000, 600, 50, 0.15)	0.05179	0.01262	0.81400	0.07604
<b>Ridge-based regression</b> (1000, 600, 50, 0.15)	<b>0.01095</b>	<b>0.00352</b>	<b>0.34000</b>	<b>0.05735</b>
<b>LDPE</b> (1000, 600, 50, 0.15)	<b>0.02653</b>	<b>0.00574</b>	<b>0.66800</b>	<b>0.07823</b>
<b>Lower bound</b> (1000, 600, 50, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.84013</b>	<b>0.03810</b>
<b>SDL-test</b> (1000, 600, 25, 0.15)	0.04937	0.01840	0.85600	0.06310
<b>Ridge-based regression</b> (1000, 600, 25, 0.15)	<b>0.01969</b>	<b>0.00358</b>	<b>0.46800</b>	<b>0.08011</b>
<b>LDPE</b> (1000, 600, 25, 0.15)	<b>0.01374</b>	<b>0.00709</b>	<b>0.63200</b>	<b>0.07155</b>
<b>Lower bound</b> (1000, 600, 25, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.86362</b>	<b>0.02227</b>
<b>SDL-test</b> (1000, 300, 50, 0.15)	0.05111	0.01947	0.43800	0.09402
<b>Ridge-based regression</b> (1000, 300, 50, 0.15)	<b>0.01011</b>	<b>0.00362</b>	<b>0.20200</b>	<b>0.05029</b>
<b>LDPE</b> (1000, 300, 50, 0.15)	<b>0.03621</b>	<b>0.00701</b>	<b>0.37600</b>	<b>0.07127</b>
<b>Lower bound</b> (1000, 300, 50, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.43435</b>	<b>0.03983</b>
<b>SDL-test</b> (1000, 300, 25, 0.15)	0.05262	0.01854	0.53600	0.08044
<b>Ridge-based regression</b> (1000, 300, 25, 0.15)	<b>0.01344</b>	<b>0.00258</b>	<b>0.33200</b>	<b>0.08230</b>
<b>LDPE</b> (1000, 300, 25, 0.15)	<b>0.01682</b>	<b>0.00352</b>	<b>0.36800</b>	<b>0.10354</b>
<b>Lower bound</b> (1000, 300, 25, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.50198</b>	<b>0.05738</b>
<b>SDL-test</b> (2000, 600, 100, 0.1)	0.05268	0.01105	0.43900	0.04383
<b>Ridge-based regression</b> (2000, 600, 100, 0.1)	<b>0.01205</b>	<b>0.00284</b>	<b>0.21200</b>	<b>0.04392</b>
<b>LDPE</b> (2000, 600, 100, 0.1)	<b>0.028102</b>	<b>0.00720</b>	<b>0.33419</b>	<b>0.04837</b>
<b>Lower bound</b> (2000, 600, 100, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.41398</b>	<b>0.03424</b>
<b>SDL-test</b> (2000, 600, 50, 0.1)	0.05856	0.00531	0.50800	0.05350
<b>Ridge-based regression</b> (2000, 600, 50, 0.1)	<b>0.01344</b>	<b>0.00225</b>	<b>0.26000</b>	<b>0.03771</b>
<b>LDPE</b> (2000, 600, 50, 0.1)	<b>0.03029</b>	<b>0.00602</b>	<b>0.37305</b>	<b>0.07281</b>
<b>Lower bound</b> (2000, 600, 50, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.49026</b>	<b>0.02625</b>
<b>SDL-test</b> (2000, 600, 20, 0.1)	0.04955	0.00824	0.57500	0.13385
<b>Ridge-based regression</b> (2000, 600, 20, 0.1)	<b>0.01672</b>	<b>0.00282</b>	<b>0.35500</b>	<b>0.08960</b>
<b>LDPE</b> (2000, 600, 20, 0.1)	<b>0.03099</b>	<b>0.00805</b>	<b>0.31350</b>	<b>0.04482</b>
<b>Lower bound</b> (2000, 600, 20, 0.1)	<b>0.05</b>	<b>NA</b>	<b>0.58947</b>	<b>0.04472</b>
<b>SDL-test</b> (2000, 600, 100, 0.15)	0.05284	0.00949	0.61600	0.06802
<b>Ridge-based regression</b> (2000, 600, 100, 0.15)	<b>0.00895</b>	<b>0.00272</b>	<b>0.31800</b>	<b>0.04131</b>
<b>LDPE</b> (2000, 600, 100, 0.15)	<b>0.01022</b>	<b>0.00570</b>	<b>0.35904</b>	<b>0.05205</b>
<b>Lower bound</b> (2000, 600, 100, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.64924</b>	<b>0.05312</b>
<b>SDL-test</b> (2000, 600, 20, 0.15)	0.05318	0.00871	0.85500	0.11891
<b>Ridge-based regression</b> (2000, 600, 20, 0.15)	<b>0.01838</b>	<b>0.00305</b>	<b>0.68000</b>	<b>0.12517</b>
<b>LDPE</b> (2000, 600, 20, 0.15)	<b>0.02512</b>	<b>0.00817</b>	<b>0.36434</b>	<b>0.05824</b>
<b>Lower bound</b> (2000, 600, 20, 0.15)	<b>0.05</b>	<b>NA</b>	<b>0.87988</b>	<b>0.03708</b>

Table 5.5.1 Comparison between **SDL-TEST**, Ridge-based regression [16], **LDPE** [149] and the lower bound for the statistical power of **SDL-TEST** (cf. Theorem 4.2.4) in Experiment 4. The significance level is  $\alpha = 0.05$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of  $p = 1000$ ,  $n = 600$ ,  $s_0 = 50$ ,  $\gamma = 0.1$

Method	Type I err (mean)	Type I err (std.)	Avg. power (mean)	Avg. power (std)
<b>SDL-test</b> (1000, 600, 50, 0.15)	0.02579	0.00967	0.71800	0.03824
<b>Ridge-based regression</b> (1000, 600, 50, 0.15)	<b>0.00326</b>	<b>0.00274</b>	<b>0.21000</b>	<b>0.05437</b>
<b>LDPE</b> (1000, 600, 50, 0.15)	0.01245	0.00391	0.64807	0.065020
<b>Lower bound</b> (1000, 600, 50, 0.15)	0.025	NA	0.75676	0.05937
<b>SDL-test</b> (1000, 600, 25, 0.15)	0.02462	0.00866	0.75600	0.12429
<b>Ridge-based regression</b> (1000, 600, 25, 0.15)	<b>0.01077</b>	<b>0.00346</b>	<b>0.30400</b>	<b>0.08262</b>
<b>LDPE</b> (1000, 600, 25, 0.15)	0.00931	0.00183	0.68503	0.17889
<b>Lower bound</b> (1000, 600, 25, 0.15)	0.025	NA	0.80044	0.05435
<b>SDL-test</b> (1000, 300, 50, 0.15)	0.02646	0.01473	0.39200	0.11478
<b>Ridge-based regression</b> (1000, 300, 50, 0.15)	<b>0.00368</b>	<b>0.00239</b>	<b>0.15000</b>	<b>0.04137</b>
<b>LDPE</b> (1000, 300, 50, 0.15)	0.01200	0.00425	0.28800	0.09654
<b>Lower bound</b> (1000, 300, 50, 0.15)	0.025	NA	0.36084	0.04315
<b>SDL-test</b> (1000, 300, 25, 0.15)	0.02400	0.00892	0.42400	0.09834
<b>Ridge-based regression</b> (1000, 300, 25, 0.15)	<b>0.00513</b>	<b>0.00118</b>	<b>0.18800</b>	<b>0.07786</b>
<b>LDPE</b> (1000, 300, 25, 0.15)	0.00492	0.00169	0.24500	0.07483
<b>Lower bound</b> (1000, 300, 25, 0.15)	0.025	NA	0.42709	0.03217
<b>SDL-test</b> (2000, 600, 100, 0.1)	0.03268	0.00607	0.32600	0.07412
<b>Ridge-based regression</b> (2000, 600, 100, 0.1)	<b>0.00432</b>	<b>0.00179</b>	<b>0.14100</b>	<b>0.05065</b>
<b>LDPE</b> (2000, 600, 100, 0.1)	0.01240	0.00572	0.20503	0.09280
<b>Lower bound</b> (2000, 600, 100, 0.1)	0.025	NA	0.32958	0.03179
<b>SDL-test</b> (2000, 600, 50, 0.1)	0.03108	0.00745	0.41800	0.04662
<b>Ridge-based regression</b> (2000, 600, 50, 0.1)	<b>0.00687</b>	<b>0.00170</b>	<b>0.18800</b>	<b>0.06680</b>
<b>LDPE</b> (2000, 600, 50, 0.1)	0.014005	0.00740	0.25331	0.04247
<b>Lower bound</b> (2000, 600, 50, 0.1)	0.025	NA	0.40404	0.06553
<b>SDL-test</b> (2000, 600, 20, 0.1)	0.02965	0.00844	0.38500	0.07091
<b>Ridge-based regression</b> (2000, 600, 20, 0.1)	<b>0.00864</b>	<b>0.00219</b>	<b>0.22500</b>	<b>0.07906</b>
<b>LDPE</b> (2000, 600, 20, 0.1)	0.01912	0.00837	0.31551	0.06288
<b>Lower bound</b> (2000, 600, 20, 0.1)	0.025	NA	0.47549	0.06233
<b>SDL-test</b> (2000, 600, 100, 0.15)	0.026737	0.009541	0.528000	0.062681
<b>Ridge-based regression</b> (2000, 600, 100, 0.15)	<b>0.002947</b>	<b>0.000867</b>	<b>0.236000</b>	<b>0.035653</b>
<b>LDPE</b> (2000, 600, 100, 0.15)	0.01012	0.00417	0.36503	0.05823
<b>Lower bound</b> (2000, 600, 100, 0.15)	0.025	NA	0.54512	0.05511
<b>SDL-test</b> (2000, 600, 20, 0.15)	0.03298	0.00771	0.79000	0.12202
<b>Ridge-based regression</b> (2000, 600, 20, 0.15)	<b>0.00732</b>	<b>0.00195</b>	<b>0.53500</b>	<b>0.07091</b>
<b>LDPE</b> (2000, 600, 20, 0.15)	0.01302	0.00711	0.60033	0.03441
<b>Lower bound</b> (2000, 600, 20, 0.15)	0.025	NA	0.81899	0.03012

Table 5.5.2 Comparison between SDL-TEST, Ridge-based regression [16], LDPE [149] and the lower bound for the statistical power of SDL-TEST (cf. Theorem 4.2.4) in Experiment 4. The significance level is  $\alpha = 0.025$ . The means and the standard deviations are obtained by testing over 10 realizations of the corresponding configuration. Here a quadruple such as (1000, 600, 50, 0.1) denotes the values of  $p = 1000$ ,  $n = 600$ ,  $s_0 = 50$ ,  $\gamma = 0.1$

Figure 5.6.9: Parameter vector  $\theta_0$  for the communities data set.

a few entries have non negligible values which correspond to the predictive attributes. In computing type I errors and powers, we take the elements in  $\theta_0$  with magnitude larger than 0.04 as active and the others as inactive.

In order to validate our approach in the high-dimensional  $p > n$  regime, we take random subsamples of the communities (hence subsamples of the rows of  $\mathbf{X}_{\text{tot}}$ ) of size  $n = 84$ . We compare SDL-TEST with the method of [16], over 20 realizations and significance levels  $\alpha = 0.01, 0.025, 0.05$ . The fraction of type I errors and statistical power is computed by comparing to  $\theta_0$ . Table 5.6.1 summarizes the results. As the reader can see, Ridge-type projection estimator is very conservative yielding to no type-I errors but a much smaller power than SDL-TEST.

In Table 5.6.2, we report the relevant features obtained from the whole dataset as described above, corresponding to the nonzero entries in  $\theta_0$ . We also report the features identified as relevant by SDL-TEST and those identified as relevant by Ridge-type projection estimator, from one random subsample of communities of size  $n = 84$ . Features description is available in [55].

Finally, in Fig. 5.6.10 we plot the normalized histograms of  $v_{S_0}$  (in red) and  $v_{S_0^c}$  (in white). Recall that  $v = (v_i)_{i=1}^p$  denotes the vector with  $v_i \equiv \hat{\theta}_i^u / (\tau[(\Sigma^{-1})_{ii}]^{1/2})$ . Further,  $v_{S_0}$  and  $v_{S_0^c}$  respectively denote the restrictions of  $v$  to the active set  $S_0$  and the inactive set  $S_0^c$ . This plot demonstrates that  $v_{S_0^c}$  has roughly standard normal distribution as expected

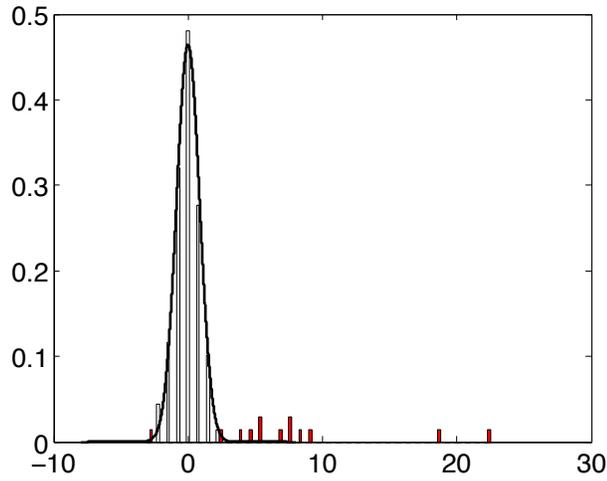


Figure 5.6.10: Normalized histogram of  $v_{S_0}$  (in red) and  $v_{S_0^c}$  (in white) for the communities data set.

by the developed theory in Chapter 4.

Method	Type I err (mean)	Avg. power (mean)
<b>SDL-test</b> ( $\alpha = 0.05$ )	0.0172043	0.4807692
<b>Ridge-based regression</b>	0	0.1423077
<b>SDL-test</b> ( $\alpha = 0.025$ )	0.01129032	0.4230769
<b>Ridge-based regression</b>	0	0.1269231
<b>SDL-test</b> ( $\alpha = 0.01$ )	0.008602151	0.3576923
<b>Ridge-based regression</b>	0	0.1076923

Table 5.6.1 Simulation results for the communities data set.

Relevant features		racePctHisp, PctTeen2Par, PctImmigRecent, PctImmigRec8, PctImmigRec10, PctNotSpeakEnglWell, OwnOccHiQuart, NumStreet, PctSameState85, LemasSwFTFieldPerPop, LemasTotReqPerPop, RacialMatchCommPol, PolicOperBudg
$\alpha = 0.01$	Relevant features (SDL-TEST)	racePctHisp, PctTeen2Par, PctImmigRecent, PctImmigRec8, PctImmigRec10, PctNotSpeakEnglWell, OwnOccHiQuart, NumStreet, PctSameState85, LemasSwFTFieldPerPop, LemasTotReqPerPop, RacialMatchCommPol, PolicOperBudg
	Relevant features (Ridge-based regression)	racePctHisp, PctSameState85
$\alpha = 0.025$	Relevant features (SDL-TEST)	racePctHisp, PctTeen2Par, PctImmigRecent, PctImmigRec8, PctImmigRec10, PctNotSpeakEnglWell, <b>PctHousOccup</b> , OwnOccHiQuart, NumStreet, PctSameState85, LemasSwFTFieldPerPop, LemasTotReqPerPop, RacialMatchCommPol, PolicOperBudg
	Relevant features (Ridge-based regression)	racePctHisp, PctSameState85
$\alpha = 0.05$	Relevant features (SDL-TEST)	racePctHisp, <b>PctUnemployed</b> , PctTeen2Par, PctImmigRecent, PctImmigRec8, PctImmigRec10, PctNotSpeakEnglWell, <b>PctHousOccup</b> , OwnOccHiQuart, NumStreet, PctSameState85, <b>LemasSwornFT</b> , LemasSwFTFieldPerPop, LemasTotReqPerPop, RacialMatchCommPol, <b>PctPolicWhite</b>
	Relevant features (Ridge-based regression)	racePctHisp, PctSameState85

Table 5.6.2 The relevant features (using the whole dataset) and the relevant features predicted by SDL-TEST and the Ridge-type projection estimator [16] for a random subsample of size  $n = 84$  from the communities. The false positive predictions are in red.

# Chapter 6

## Proof of Theorems in Part I

This chapter contains proofs of Theorems and technical lemmas stated in Chapter 2.

### 6.1 Proof of Theorems in Chapter 2

#### 6.1.1 Proof of Theorem 2.3.3

Substituting  $Y = \mathbf{X}\theta_0 + W$  in the definition (2.3.2), we get

$$\begin{aligned}\widehat{\theta}^* &= \widehat{\theta}^n + \frac{1}{n}M\mathbf{X}^\top\mathbf{X}(\theta_0 - \widehat{\theta}^n) + \frac{1}{n}M\mathbf{X}^\top W \\ &= \theta_0 + \frac{1}{\sqrt{n}}Z + \frac{1}{\sqrt{n}}\Delta,\end{aligned}\tag{6.1.1}$$

with  $Z, \Delta$  defined as per the theorem statement. Further  $Z$  is Gaussian with the stated covariance because it is a linear function of the Gaussian vector  $W \sim \mathbf{N}(0, \sigma^2 \mathbf{I}_{p \times p})$ .

We are left with the task of proving the bound (2.3.6) on  $\Delta$ . Note that by definition (2.3.1), we have

$$\|\Delta\|_\infty \leq \sqrt{n} |M\widehat{\Sigma} - \mathbf{I}|_\infty \|\widehat{\theta}^n - \theta_0\|_1 = \sqrt{n} \mu_* \|\widehat{\theta}^n - \theta_0\|_1.\tag{6.1.2}$$

By [19, Theorem 6.1, Lemma 6.2], we have, for any  $\lambda \geq 4\sigma\sqrt{2K\log(pe^{t^2/2})/n}$

$$\mathbb{P}\left(\|\widehat{\theta}^n - \theta_0\|_1 \geq \frac{4\lambda s_0}{\phi_0^2}\right) \leq 2e^{-t^2/2}.\tag{6.1.3}$$

(More precisely, we consider the trivial generalization of [19, Lemma 6.2] to the case

$(\mathbf{X}^T \mathbf{X}/n)_{ii} \leq K$ , instead of  $(\mathbf{X}^T \mathbf{X}/n)_{ii} = 1$  for all  $i \in [p]$ .)

Substituting Eq. (6.1.2) in the last bound, we get

$$\mathbb{P}\left(\|\Delta\|_\infty \geq \frac{4\lambda\mu_*s_0\sqrt{n}}{\phi_0^2}\right) \leq 2e^{-t^2/2}. \quad (6.1.4)$$

Finally, the claim follows by selecting  $t$  so that  $e^{t^2/2} = p^{c_0}$ .

### 6.1.2 Proof of Theorem 2.3.4.(a)

Note that the event  $\mathcal{E}_n$  requires two conditions. Hence, its complement is given by

$$\mathcal{E}_n(\phi_0, s_0, K)^c = \mathcal{B}_{1,n}(\phi_0, s_0) \cup \mathcal{B}_{2,n}(K), \quad (6.1.5)$$

$$\mathcal{B}_{1,n}(\phi_0, s_0) \equiv \left\{ \mathbf{X} \in \mathbb{R}^{n \times p} : \min_{S: |S| \leq s_0} \phi(\widehat{\Sigma}, S) < \phi_0, \widehat{\Sigma} = (\mathbf{X}^T \mathbf{X}/n) \right\}, \quad (6.1.6)$$

$$\mathcal{B}_{2,n}(K) \equiv \left\{ \mathbf{X} \in \mathbb{R}^{n \times p} : \max_{i \in [p]} \widehat{\Sigma}_{i,i} > K, \widehat{\Sigma} = (\mathbf{X}^T \mathbf{X}/n) \right\}. \quad (6.1.7)$$

We will bound separately the probability of  $\mathcal{B}_{1,n}$  and the probability of  $\mathcal{B}_{2,n}$ . The claim of Theorem 2.3.4.(a) follows by union bound.

#### 6.1.2.1 Controlling $\mathcal{B}_{1,n}(\phi_0, s_0)$

It is also useful to recall the notion of restricted eigenvalue, introduced by Bickel, Ritov and Tsybakov [14].

**Definition 6.1.1.** *Given a symmetric matrix  $Q \in \mathbb{R}^{p \times p}$  an integer  $s_0 \geq 1$ , and  $L > 0$ , the restricted eigenvalue of  $Q$  is defined as*

$$\phi_{\text{RE}}^2(Q, s_0, L) \equiv \min_{S \subseteq [p], |S| \leq s_0} \min_{\theta \in \mathbb{R}^p} \left\{ \frac{\langle \theta, Q \theta \rangle}{\|\theta_S\|_2^2} : \theta \in \mathbb{R}^p, \|\theta_{S^c}\|_1 \leq L \|\theta_S\|_1 \right\}. \quad (6.1.8)$$

[116] proves that, if the population covariance satisfies the restricted eigenvalue condition, then the sample covariance satisfies it as well, with high probability. More precisely, by [116, Theorem 6] we have

$$\mathbb{P}\left(\phi_{\text{RE}}(\widehat{\Sigma}, s_0, 3) \geq \frac{1}{2} \phi_{\text{RE}}(\Sigma, s_0, 9)\right) \geq 1 - 2e^{-n/(4c_*\kappa^4)}, \quad (6.1.9)$$

for some  $c_* \leq 2000$ ,  $m \equiv 6 \times 10^4 s_0 C_{\text{max}}^2 / \phi_{\text{RE}}^2(\Sigma, s_0, 9)$ , and every  $n \geq 4c_* m \kappa^4 \log(120ep/m)$ .

Note that  $\phi_{\text{RE}}(\Sigma, s_0, 9) \geq \sigma_{\min}(\Sigma)^{1/2} \geq \sqrt{C_{\min}}$ , and by Cauchy-Schwartz,

$$\min_{S:|S|\leq s_0} \phi(\widehat{\Sigma}, S) \geq \phi_{\text{RE}}(\widehat{\Sigma}, s_0, 3).$$

With the definitions in the statement (cf. Eq. (2.3.8)), we therefore have

$$\mathbb{P}\left(\min_{S:|S|\leq s_0} \phi(\widehat{\Sigma}, S) \geq \frac{1}{2}\sqrt{C_{\min}}\right) \geq 1 - 2e^{-c_1 n}. \quad (6.1.10)$$

Equivalently,  $\mathbb{P}(\mathcal{B}_{1,n}(\phi_0, s_0)) \leq 2e^{-c_1 n}$ .

### 6.1.2.2 Controlling $\mathcal{B}_{2,n}(K)$

By definition

$$\widehat{\Sigma}_{ii} - 1 = \frac{1}{n} \sum_{\ell=1}^n (\langle X_{\ell}, e_i \rangle^2 - 1) = \frac{1}{n} \sum_{\ell=1}^n u_{\ell}, \quad (6.1.11)$$

Note that  $u_{\ell}$  are independent centered random variables. Further, (recalling that, for any random variables  $U, V$ ,  $\|U + V\|_{\psi_1} \leq \|U\|_{\psi_1} + \|V\|_{\psi_1}$ , and  $\|U^2\|_{\psi_1} \leq 2\|U\|_{\psi_2}^2$ ) they are subexponential with subexponential norm

$$\begin{aligned} \|u_{\ell}\|_{\psi_1} &\leq 2\|\langle X_{\ell}, e_i \rangle^2\|_{\psi_1} \leq 4\|\langle X_{\ell}, e_i \rangle\|_{\psi_2}^2 \\ &\leq 4\|\langle \Sigma^{-1/2} X_{\ell}, \Sigma^{1/2} e_i \rangle\|_{\psi_1}^2 \\ &\leq 4\kappa^2 \|\Sigma^{1/2} e_i\|_2^2 = 4\kappa^2 \Sigma_{ii} = 4\kappa^2. \end{aligned}$$

By Bernstein-type inequality for centered subexponential random variables, cf. [136], we get

$$\mathbb{P}\left\{\frac{1}{n} \left| \sum_{\ell=1}^n u_{\ell} \right| \geq \epsilon\right\} \leq 2 \exp\left[-\frac{n}{6} \min\left(\left(\frac{\epsilon}{4e\kappa^2}\right)^2, \frac{\epsilon}{4e\kappa^2}\right)\right]. \quad (6.1.12)$$

Hence, for all  $\epsilon$  such that  $\epsilon/(e\kappa^2) \in [\sqrt{(48 \log p)/n}, 4]$ ,

$$\mathbb{P}\left(\max_{i \in [p]} \widehat{\Sigma}_{ii} \geq 1 + \epsilon\right) \leq 2p \exp\left(-\frac{n\epsilon^2}{24e^2\kappa^4}\right) \leq 2e^{-c_1 n}, \quad (6.1.13)$$

which implies  $\mathbb{P}(\mathbf{X} \in \mathcal{B}_{2,n}(K)) \leq 2e^{-c_1 n}$  for all  $K-1 \geq 20\kappa^2 \sqrt{(\log p)/n} \geq \sqrt{(48e^2\kappa^4 \log p)/n}$ .

### 6.1.3 Proof of Theorem 2.3.4.(b)

Obviously, we have

$$\mu_{\min}(\mathbf{X}) \leq |\Sigma^{-1}\widehat{\Sigma} - \mathbf{I}|, \quad (6.1.14)$$

and hence the statement follows immediately from the following estimate.

**Lemma 6.1.2.** *Consider a random design matrix  $\mathbf{X} \in \mathbb{R}^{p \times p}$ , with i.i.d. rows having mean zero and population covariance  $\Sigma$ . Assume that*

(i) *We have  $\sigma_{\min}(\Sigma) \geq C_{\min} > 0$ , and  $\sigma_{\max}(\Sigma) \leq C_{\max} < \infty$ .*

(ii) *The rows of  $X\Sigma^{-1/2}$  are sub-gaussian with  $\kappa = \|\Sigma^{-1/2}X_1\|_{\psi_2}$ .*

Let  $\widehat{\Sigma} = (\mathbf{X}^T\mathbf{X})/n$  be the empirical covariance. Then, for any constant  $C > 0$ , the following holds true.

$$\mathbb{P}\left\{|\Sigma^{-1}\widehat{\Sigma} - \mathbf{I}|_{\infty} \geq a\sqrt{\frac{\log p}{n}}\right\} \leq 2p^{-c_2}, \quad (6.1.15)$$

with  $c_2 = (a^2C_{\min})/(24e^2\kappa^4C_{\max}) - 2$ .

*Proof of Lemma 6.1.2.* The proof is based on Bernstein-type inequality for sub-exponential random variables [136]. Let  $\tilde{X}_{\ell} = \Sigma^{-1/2}X_{\ell}$ , for  $\ell \in [n]$ , and write

$$Z \equiv \Sigma^{-1}\widehat{\Sigma} - \mathbf{I} = \frac{1}{n} \sum_{\ell=1}^n \left\{ \Sigma^{-1}X_{\ell}X_{\ell}^T - \mathbf{I} \right\} = \frac{1}{n} \sum_{\ell=1}^n \left\{ \Sigma^{-1/2}\tilde{X}_{\ell}\tilde{X}_{\ell}^T\Sigma^{1/2} - \mathbf{I} \right\}.$$

Fix  $i, j \in [p]$ , and for  $\ell \in [n]$ , let  $v_{\ell}^{(ij)} = \langle \Sigma_{i,\cdot}^{-1/2}, \tilde{X}_{\ell} \rangle \langle \Sigma_{j,\cdot}^{1/2}, \tilde{X}_{\ell} \rangle - \delta_{i,j}$ , where  $\delta_{i,j} = \mathbf{1}_{\{i=j\}}$ . Notice that  $\mathbb{E}(v_{\ell}^{(ij)}) = 0$ , and the  $v_{\ell}^{(ij)}$  are independent for  $\ell \in [n]$ . Also,  $Z_{i,j} = (1/n) \sum_{\ell=1}^n v_{\ell}^{(ij)}$ . By [136, Remark 5.18], we have

$$\|v_{\ell}^{(ij)}\|_{\psi_1} \leq 2\|\langle \Sigma_{i,\cdot}^{-1/2}, \tilde{X}_{\ell} \rangle \langle \Sigma_{j,\cdot}^{1/2}, \tilde{X}_{\ell} \rangle\|_{\psi_1}.$$

Moreover, for any two random variables  $X$  and  $Y$ , we have

$$\begin{aligned} \|XY\|_{\psi_1} &= \sup_{p \geq 1} p^{-1} \mathbb{E}(|XY|^p)^{1/p} \\ &\leq \sup_{p \geq 1} p^{-1} \mathbb{E}(|X|^{2p})^{1/2p} \mathbb{E}(|Y|^{2p})^{1/2p} \\ &\leq 2 \left( \sup_{q \geq 2} q^{-1/2} \mathbb{E}(|X|^q)^{1/q} \right) \left( \sup_{q \geq 2} q^{-1/2} \mathbb{E}(|Y|^q)^{1/q} \right) \\ &\leq 2 \|X\|_{\psi_2} \|Y\|_{\psi_2}. \end{aligned}$$

Hence, by assumption (ii), we obtain

$$\begin{aligned} \|v_\ell^{(ij)}\|_{\psi_1} &\leq 2 \|\langle \Sigma_{i,\cdot}^{-1/2}, \tilde{X}_\ell \rangle\|_{\psi_2} \|\langle \Sigma_{j,\cdot}^{1/2}, \tilde{X}_\ell \rangle\|_{\psi_2} \\ &\leq 2 \|\Sigma_{i,\cdot}^{-1/2}\|_2 \|\Sigma_{j,\cdot}^{1/2}\|_2 \kappa^2 \leq 2 \sqrt{C_{\max}/C_{\min}} \kappa^2. \end{aligned}$$

Let  $\kappa' = 2\sqrt{C_{\max}/C_{\min}}\kappa^2$ . Applying Bernstein-type inequality for centered sub-exponential random variables [136], we get

$$\mathbb{P}\left\{ \frac{1}{n} \left| \sum_{\ell=1}^n v_\ell^{(ij)} \right| \geq \epsilon \right\} \leq 2 \exp \left[ -\frac{n}{6} \min \left( \left( \frac{\epsilon}{e\kappa'} \right)^2, \frac{\epsilon}{e\kappa'} \right) \right].$$

Choosing  $\epsilon = a\sqrt{(\log p)/n}$ , and assuming  $n \geq [a/(e\kappa')]^2 \log p$ , we arrive at

$$\mathbb{P}\left\{ \frac{1}{n} \left| \sum_{\ell=1}^n v_\ell^{(ij)} \right| \geq a\sqrt{\frac{\log p}{n}} \right\} \leq 2p^{-a^2/(6e^2\kappa'^2)}.$$

The result follows by union bounding over all possible pairs  $i, j \in [p]$ . □

#### 6.1.4 Proof of Theorem 2.2.1

The KKT condition for the Lasso estimator reads

$$\frac{1}{n} \mathbf{X}^\top (Y - \mathbf{X}\hat{\theta}^n) = \lambda v(\hat{\theta}^n),$$

with  $v(\hat{\theta}^n) \in \mathbb{R}^p$  being a vector in the subgradient of the  $\ell_1$  norm at  $\hat{\theta}^n$ . Adding  $\hat{\theta}^n - \theta_0$  to both sides, and taking expectation over the noise, we get

$$\text{Bias}(\hat{\theta}^*) = \text{Bias}(\hat{\theta}^n) + \lambda \mathbb{E}\{v(\hat{\theta}^n)|\mathbf{X}\}, \quad (6.1.16)$$

where  $\hat{\theta}^*$  is a debiased estimator of the general form (2.3.2), for  $M = \mathbf{I}$ .

In the following, we show that there exist a set of design matrices  $\mathcal{B}_n \subseteq \mathbb{R}^{n \times p}$  satisfying the probability bound (2.2.3) and a constant  $0 < c_{**} \leq 4800$ , such that for  $\mathbf{X} \in \mathcal{B}_n$  we have

$$\|\text{Bias}(\hat{\theta}^*)\|_\infty \leq \frac{c_{**}\sigma s_0 \log p}{n}, \quad (6.1.17)$$

$$\|\mathbb{E}\{v(\hat{\theta}^n)|\mathbf{X}\}\|_\infty \geq \frac{2}{3}. \quad (6.1.18)$$

Note that by applying triangle inequality to (6.1.16), we have

$$\|\text{Bias}(\hat{\theta}^n)\|_\infty \geq \lambda \|\mathbb{E}\{v(\hat{\theta}^n)|\mathbf{X}\}\|_\infty - \|\text{Bias}(\hat{\theta}^*)\|_\infty.$$

Using the bounds (6.1.17) and (6.1.18) in the above inequality, and invoking the assumption  $n \geq (3c_{**}\sigma s_0/c)^2 \log p$ , we obtain the desired result.

We begin with proving inequality (6.1.17). By Theorem 2.3.4.(a), we have

$$\mathbb{P}\left(\mathbf{X} \in \mathcal{E}_n(1/2, s_0, 3/2)\right) \geq 1 - 4e^{-c_1 n}. \quad (6.1.19)$$

Further, by Lemma 6.1.2, with  $\hat{\Sigma} \equiv \mathbf{X}^\top \mathbf{X}/n$ , we have

$$\mathbb{P}\left(\mu_*(\mathbf{X}; \mathbf{I}) \leq 30\sqrt{\frac{\log p}{n}}\right) \geq 1 - 2p^{-3}. \quad (6.1.20)$$

Hence, defining

$$\mathcal{B}_n \equiv \mathcal{E}_n(1/2, s_0, 3/2) \cap \left\{ \mathbf{X} \in \mathbb{R}^{n \times p} : \mu_*(\mathbf{X}; \mathbf{I}) \leq 30\sqrt{\frac{\log p}{n}} \right\} \quad (6.1.21)$$

we have the desired probability bound (2.2.3). Let  $\hat{\theta}^n = \hat{\theta}^n(Y, \mathbf{X}; \mathbf{I}, \lambda)$ . By Theorem 2.3.3, we have, for any  $\mathbf{X} \in \mathcal{B}_n$

$$\hat{\theta}^* = \theta_0 + \frac{1}{\sqrt{n}}Z + \frac{1}{\sqrt{n}}\Delta, \quad Z|\mathbf{X} \sim \mathbf{N}(0, \sigma^2 \hat{\Sigma}), \quad (6.1.22)$$

and further

$$\mathbb{P}\left\{\|\Delta\|_\infty \geq \frac{480c\sigma s_0 \log p}{\sqrt{n}} \mid \mathbf{X}\right\} \leq 2p^{1-(c^2/48)}. \quad (6.1.23)$$

Let  $L = (480\sigma s_0 \log p)/\sqrt{n}$ . Then,

$$\begin{aligned} \|\text{Bias}(\hat{\theta}^u)\|_\infty &\leq \frac{1}{\sqrt{n}} \mathbb{E}\{\|\Delta\|_\infty \mid \mathbf{X}\} \\ &= \frac{L}{\sqrt{n}} \int_0^\infty \mathbb{P}\left\{\|\Delta\|_\infty \geq Lc \mid \mathbf{X}\right\} dc \\ &\leq \frac{2L}{\sqrt{n}} \int_0^\infty \min(1, p^{1-(c^2/48)}) dc \leq \frac{10L}{\sqrt{n}}, \end{aligned}$$

where (6.1.23) is used in the second inequality. Therefore, for some constant  $c_{**} \leq 4800$ , we have

$$\|\text{Bias}(\hat{\theta}^*)\|_\infty \leq c_{**}\sigma \frac{s_0 \log p}{n}. \quad (6.1.24)$$

We next prove inequality (6.1.18). Note that  $v(\hat{\theta}^n)_i = 1$  whenever  $\hat{\theta}_i^n > 0$ , and  $|v(\hat{\theta}^n)_i| \leq 1$  for all coordinates  $i$ . Therefore, letting  $b_0 \equiv 480c\sigma(s_0 \log p)/n$  we have

$$\begin{aligned} 1 - \mathbb{E}\{v(\hat{\theta}^n)_i \mid \mathbf{X}\} &\leq 2\mathbb{P}\left(\hat{\theta}_i^n \leq 0 \mid \mathbf{X}\right) \leq 2\mathbb{P}\left(\hat{\theta}_i^u \leq \lambda \mid \mathbf{X}\right) \quad (6.1.25) \\ &\leq 2\mathbb{P}\left(\theta_{0,i} + \frac{1}{\sqrt{n}}Z_i + \frac{1}{\sqrt{n}}\Delta_i \leq \lambda \mid \mathbf{X}\right) \\ &\leq 2\mathbb{P}\left(\frac{1}{\sqrt{n}}Z_i \leq \lambda + b_0 - \theta_{0,i} \mid \mathbf{X}\right) + 2\mathbb{P}\left(\|\Delta\|_\infty > \sqrt{n}b_0\right) \\ &= 2\Phi\left((\lambda + b_0 - \theta_{0,i})\sqrt{n/(\sigma^2\hat{\Sigma}_{ii})}\right) + 4p^{1-(c^2/48)} \\ &\leq 2\Phi\left((\lambda + b_0 - \theta_{0,i})\sqrt{2n/(3\sigma^2)}\right) + 4p^{1-(c^2/48)} \quad (6.1.26) \end{aligned}$$

with  $\Phi(x)$  the standard normal distribution function. Here, we used the relation  $\hat{\theta}^u = \hat{\theta} + \lambda v(\hat{\theta})$  in Eq. (6.1.25). Further, Eq. (6.1.26) holds because  $\max_{i \in [p]} \hat{\Sigma}_{ii} \leq 3/2$  on  $\mathcal{B}_n$ . We then choose  $\theta_0$  so that  $\theta_{0,i} \geq b_0 + \lambda + \sqrt{30\sigma^2/n}$ , for  $i \in \text{supp}(\theta_0)$ . We therefore obtain

$$\mathbb{E}\{v(\hat{\theta}^n)_i \mid \mathbf{X}\} \geq 1 - 2\Phi(-\sqrt{20}) - 4p^{1-(c^2/48)} \geq \frac{2}{3}, \quad (6.1.27)$$

where in the last step we used the assumption  $p \geq 13^{48/(c^2-48)}$ .

This completes the proof of the theorem.

### 6.1.5 Proof of Theorem 2.3.5

Let

$$\Delta_0 \equiv \left( \frac{16ac\sigma}{C_{\min}} \right) \frac{s_0 \log p}{\sqrt{n}} \quad (6.1.28)$$

be a shorthand for the bound on  $\|\Delta\|_\infty$  appearing in Eq. (2.3.14). Then we have

$$\begin{aligned} \mathbb{P}(\|\Delta\|_\infty \geq \Delta_0) &\leq \mathbb{P}(\{\|\Delta\|_\infty \geq \Delta_0\} \cap \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)) \\ &\quad + \mathbb{P}(\mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2)) + \mathbb{P}(\mathcal{G}_n^c(a)) \\ &\leq \mathbb{P}(\{\|\Delta\|_\infty \geq \Delta_0\} \cap \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)) + 4e^{-c_1 n} + 2p^{-c_2}, \end{aligned}$$

where, in the first equation  $\mathcal{A}^c$  denotes the complement of event  $\mathcal{A}$  and the second inequality follows from Theorem 2.3.4. Notice, in particular, that the bound (2.3.8) can be applied for  $K = 3/2$  since, under the present assumptions  $20\kappa^2 \sqrt{(\log p)/n} \leq 1/2$ . Finally

$$\begin{aligned} &\mathbb{P}(\{\|\Delta\|_\infty \geq \Delta_0\} \cap \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)) \\ &\leq \sup_{\mathbf{X} \in \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)} \mathbb{P}(\|\Delta\|_\infty \geq \Delta_0 \mid \mathbf{X}) \leq 2p^{-\tilde{c}_0}. \end{aligned} \quad (6.1.29)$$

Here the last inequality follows from Theorem 2.3.3 applied per given  $\mathbf{X} \in \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)$  and hence using the bound (2.3.6) with  $\phi_0 = \sqrt{C_{\min}}/2$ ,  $K = 3/2$ ,  $\mu_* = a\sqrt{(\log p)/n}$ .

### 6.1.6 Proof of Corollary 2.3.7

By Theorem 2.3.3, for any  $\mathbf{X} \in \mathcal{E}_n(\sqrt{C_{\min}}/2, s_0, 3/2) \cap \mathcal{G}_n(a)$ , we have

$$\mathbb{P}\left\{\|\Delta\|_\infty \geq Lc \mid \mathbf{X}\right\} \leq 2p^{1-(c^2/48)}, \quad L \equiv \frac{16a\sigma}{C_{\min}} \frac{s_0 \log p}{\sqrt{n}}. \quad (6.1.30)$$

This is obtained by setting  $\phi_0 = \sqrt{C_{\min}}/2$ ,  $K = 3/2$ ,  $\mu_* = a\sqrt{(\log p)/n}$  in Eq. (2.3.6). Hence

$$\begin{aligned} \|\text{Bias}(\widehat{\theta}^u)\|_\infty &\leq \frac{1}{\sqrt{n}} \mathbb{E}\{\|\Delta\|_\infty | \mathbf{X}\} \\ &= \frac{L}{\sqrt{n}} \int_0^\infty \mathbb{P}\{\|\Delta\|_\infty \geq Lc | \mathbf{X}\} \, dc \\ &\leq \frac{2L}{\sqrt{n}} \int_0^\infty \min(1, p^{1-(c^2/48)}) \, dc \leq \frac{10L}{\sqrt{n}}, \end{aligned} \quad (6.1.31)$$

which coincides with Eq. (2.3.17). The probability estimate (2.3.18) simply follows from Theorem 2.3.4 using union bound.

### 6.1.7 Proof of Lemma 2.4.2

We will prove that, under the stated assumptions

$$\limsup_{n \rightarrow \infty} \sup_{\|\theta_0\|_0 \leq s_0} \mathbb{P}\left\{ \frac{\sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \leq x \right\} \leq \Phi(x). \quad (6.1.32)$$

A matching lower bound follows by a completely analogous argument.

Notice that by Eq. (2.3.13), we have

$$\frac{\sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})}{\sigma[M\widehat{\Sigma}M^\top]_{ii}^{1/2}} = \frac{e_i^\top M \mathbf{X}^\top W}{\sigma[M\widehat{\Sigma}M^\top]_{ii}^{1/2}} + \frac{\Delta_i}{\sigma[M\widehat{\Sigma}M^\top]_{ii}^{1/2}}. \quad (6.1.33)$$

Let  $V = \mathbf{X}M^\top e_i / (\sigma[M\widehat{\Sigma}M^\top]_{ii}^{1/2})$  and  $\widetilde{Z} \equiv V^\top W$ . We claim that  $\widetilde{Z} \sim \mathbf{N}(0, 1)$ . To see this, note that  $\|V\|_2 = 1$ , and  $V$  and  $W$  are independent. Hence,

$$\mathbb{P}(\widetilde{Z} \leq x) = \mathbb{E}\{\mathbb{P}(V^\top W \leq x | V)\} = \mathbb{E}\{\Phi(x) | V\} = \Phi(x), \quad (6.1.34)$$

which proves our claim. In order to prove Eq. (6.1.32), fix  $\epsilon > 0$  and write

$$\begin{aligned} \mathbb{P}\left(\frac{\sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \leq x\right) &= \mathbb{P}\left(\frac{\sigma}{\widehat{\sigma}}\widetilde{Z} + \frac{\Delta_i}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \leq x\right) \\ &\leq \mathbb{P}\left(\frac{\sigma}{\widehat{\sigma}}\widetilde{Z} \leq x + \epsilon\right) + \mathbb{P}\left(\frac{|\Delta_i|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \epsilon\right) \\ &\leq \mathbb{P}\left(\widetilde{Z} \leq x + 2\epsilon + \epsilon|x|\right) + \mathbb{P}\left(\frac{|\Delta_i|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \epsilon\right) \\ &\quad + \mathbb{P}\left(\left|\frac{\widehat{\sigma}}{\sigma} - 1\right| \geq \epsilon\right). \end{aligned}$$

By taking the limit and using the assumption (2.4.1), we obtain

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sup_{\|\theta_0\|_0 \leq s_0} \mathbb{P}\left(\frac{\sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \leq x\right) &\leq \\ &\Phi(x + 2\epsilon + \epsilon|x|) + \limsup_{n \rightarrow \infty} \sup_{\|\theta_0\|_0 \leq s_0} \mathbb{P}\left(\frac{|\Delta_i|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \epsilon\right). \end{aligned}$$

Since  $\epsilon > 0$  is arbitrary, it is therefore sufficient to show that the limit on the right hand side vanishes for any  $\epsilon > 0$ .

Note that  $[M\widehat{\Sigma}M^\top]_{i,i} \geq 1/(4\widehat{\Sigma}_{ii})$  for all  $n$  large enough, by Lemma 2.4.1, and since  $\mu = a\sqrt{(\log p)/n} \rightarrow 0$  as  $n, p \rightarrow \infty$ . We have therefore

$$\begin{aligned} \mathbb{P}\left(\frac{|\Delta_i|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \epsilon\right) &\leq \mathbb{P}\left(\frac{2}{\widehat{\sigma}}\widehat{\Sigma}_{ii}^{1/2}|\Delta_i| \geq \epsilon\right) \\ &\leq \mathbb{P}\left(\frac{5}{\sigma}|\Delta_i| \geq \epsilon\right) + \mathbb{P}\left(\frac{\widehat{\sigma}}{\sigma} \leq \frac{1}{2}\right) + \mathbb{P}(\widehat{\Sigma}_{ii} \geq \sqrt{2}). \end{aligned}$$

Note that  $\mathbb{P}((\widehat{\sigma}/\sigma) \leq 1/2) \rightarrow 0$  by assumption (2.4.1), and  $\mathbb{P}(\widehat{\Sigma}_{ii} \geq \sqrt{2}) \rightarrow 0$  by Theorem 2.3.4.(b). Hence

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sup_{\|\theta_0\|_0 \leq s_0} \mathbb{P}\left(\frac{|\Delta_i|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \epsilon\right) &\leq \limsup_{n \rightarrow \infty} \sup_{\|\theta_0\|_0 \leq s_0} \mathbb{P}\left(\|\Delta\|_\infty \geq \frac{\epsilon\sigma}{5}\right) \\ &\leq \limsup_{n \rightarrow \infty} (4e^{-c_1 n} + 4p^{-(\tilde{c}_0 \wedge c_2)}) = 0, \end{aligned}$$

where the last inequality follows from Eq. (2.3.14), recalling that  $s_0 = o(\sqrt{n}/\log p)$  and hence  $(16acs_0 \log p)/(C_{\min}\sqrt{n}) \leq \epsilon/5$  for all  $n$  large enough.

This completes the proof of Eq. (6.1.32). The matching lower bound follows by the same argument.

### 6.1.8 Proof of Theorem 2.4.6

Under the assumptions of Theorem 2.3.5 and assuming  $s_0 = o(\sqrt{n}/\log p)$ , we have

$$\sqrt{n}(\hat{\theta}^u - \theta_0) = \frac{1}{\sqrt{n}}M\mathbf{X}^\top W + \Delta$$

with  $\|\Delta\|_\infty = o(1)$ . Using Lemma 2.4.1, we have

$$\frac{\sqrt{n}(\hat{\theta}_i^u - \theta_{0,i})}{\sigma[M\hat{\Sigma}M^\top]_{i,i}^{1/2}} = Z_i + o(1), \quad \text{with } Z_i \equiv \frac{1}{\sqrt{n}} \frac{m_i^\top \mathbf{X}^\top W}{\sigma[m_i^\top \hat{\Sigma} m_i]^{1/2}}.$$

The following lemma characterizes the limiting distribution of  $Z_i|\mathbf{X}$  which implies the validity of the proposed confidence intervals.

**Lemma 6.1.3.** *Suppose that the noise variables  $W_i$  are independent with  $\mathbb{E}(W_i) = 0$ , and  $\mathbb{E}(W_i^2) = \sigma^2$ , and  $E(|W_i|^{2+a}) \leq C\sigma^{2+a}$  for some  $a > 0$ . Let  $M = (m_1, \dots, m_p)^\top$  be the matrix with rows  $m_i^\top$  obtained by solving optimization problem (2.4.12). For  $i \in [p]$ , define*

$$Z_i = \frac{1}{\sqrt{n}} \frac{m_i^\top \mathbf{X}^\top W}{\sigma[m_i^\top \hat{\Sigma} m_i]^{1/2}}.$$

*Under the assumptions of Theorem 2.3.5, for any sequence  $i = i(n) \in [p]$ , and any  $x \in \mathbb{R}$ , we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_i \leq x | \mathbf{X}) = \Phi(x).$$

Lemma 6.1.3 is proved in Appendix A.3.

## 6.2 Proof of Theorems in Chapter 3

### 6.2.1 Proof of Theorem 3.1.1

We begin with proving Eq. (3.1.3). Defining  $Z_i \equiv \sqrt{n}(\widehat{\theta}_i^u - \theta_{0,i})/(\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2})$ , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \alpha_n(\widehat{T}_i) &= \lim_{n \rightarrow \infty} \sup_{\theta_0} \left\{ \mathbb{P}(P_i \leq \alpha) : i \in [p], \|\theta_0\|_0 \leq s_0, \theta_{0,i} = 0 \right\} \\ &= \lim_{n \rightarrow \infty} \sup_{\theta_0} \left\{ \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \frac{\sqrt{n}|\widehat{\theta}_i^u|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}}\right) : i \in [p], \|\theta_0\|_0 \leq s_0, \theta_{0,i} = 0 \right\} \\ &= \lim_{n \rightarrow \infty} \sup_{\theta_0} \left\{ \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq |Z_i|\right) : i \in [p], \|\theta_0\|_0 \leq s_0 \right\} \leq \alpha, \end{aligned}$$

where the last inequality follows from Lemma 2.4.2.

We next prove Eq. (3.1.4). Recall that  $\Sigma_{\cdot,i}^{-1}$  is a feasible solution of (2.3.11), for  $1 \leq i \leq p$  with probability at least  $1 - 2p^{-c_2}$ , as per Lemma 6.1.2). On this event, letting  $m_i$  be the solution of the optimization problem (2.3.11), we have

$$\begin{aligned} m_i^\top \widehat{\Sigma} m_i &\leq \Sigma_{\cdot,i}^{-1} \widehat{\Sigma} \Sigma_{\cdot,i}^{-1} \\ &= (\Sigma_{\cdot,i}^{-1} \widehat{\Sigma} \Sigma_{\cdot,i}^{-1} - \Sigma_{ii}^{-1}) + \Sigma_{ii}^{-1} \\ &= \frac{1}{n} \sum_{j=1}^N (V_j^2 - \Sigma_{ii}^{-1}) + \Sigma_{ii}^{-1}, \end{aligned}$$

where  $V_j = \Sigma_{\cdot,i}^{-1} X_j$  are i.i.d. random variables with  $\mathbb{E}(V_j^2) = \Sigma_{ii}^{-1}$  and sub-gaussian norm

$$\|V_j\|_{\psi_2} \leq \|\Sigma_{\cdot,i}^{-1/2}\|_2 \|\Sigma^{-1/2} X_j\|_{\psi_2} \leq \kappa \sqrt{\Sigma_{ii}^{-1}}.$$

Letting  $U_j = V_j^2 - \Sigma_{ii}^{-1}$ , we have that  $U_j$  is zero mean and sub-exponential with  $\|U_j\|_{\psi_1} \leq 2\|V_j^2\|_{\psi_1} \leq 4\|V_j\|_{\psi_2}^2 \leq 4\kappa^2 \Sigma_{ii}^{-1} \leq 4\kappa^2 \sigma_{\min}(\Sigma)^{-1} \leq 4\kappa^2 C_{\min}^{-1} \equiv \kappa'$ . Hence, by applying Bernstein inequality (as, for instance, in the proof of Lemma 6.1.2), we have, for  $\epsilon \leq \epsilon \kappa'$ ,

$$\mathbb{P}\left(m_i^\top \widehat{\Sigma} m_i \geq \Sigma_{ii}^{-1} + \epsilon\right) \leq 2e^{-(n/6)(\epsilon/\epsilon \kappa')^2} + 2p^{-c_2}.$$

We can make  $c_2 \geq 2$  by a suitable choice of  $a$  and therefore, by Borel-Cantelli we have the

following almost surely

$$\limsup_{n \rightarrow \infty} [m_i^\top \widehat{\Sigma} m_i - \Sigma_{i,i}^{-1}] \leq 0. \quad (6.2.1)$$

Now we are ready to prove the lower bound for the power. Let  $z_* \equiv \Phi^{-1}(1 - \alpha/2)$ . Then,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1 - \beta_n(\widehat{T}_i; \gamma)}{1 - \beta_{i,n}^*(\gamma)} \\ &= \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} \inf_{\theta_0} \left\{ \mathbb{P}(P_i \leq \alpha) : \|\theta_0\|_0 \leq s_0, |\theta_{0,i}| \geq \gamma \right\} \\ &= \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} \inf_{\theta_0} \left\{ \mathbb{P}\left(z_* \leq \frac{\sqrt{n}|\widehat{\theta}_i^u|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}}\right) : \|\theta_0\|_0 \leq s_0, |\theta_{0,i}| \geq \gamma \right\} \\ &= \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} \inf_{\theta_0} \left\{ \mathbb{P}\left(z_* \leq \left|Z_i + \frac{\sqrt{n}\theta_{0,i}}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}}\right|\right) : \|\theta_0\|_0 \leq s_0, |\theta_{0,i}| \geq \gamma \right\} \\ &\stackrel{(a)}{\geq} \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} \inf_{\theta_0} \left\{ \mathbb{P}\left(z_* \leq \left|Z_i + \frac{\sqrt{n}\gamma}{\sigma[\Sigma_{i,i}^{-1}]^{1/2}}\right|\right) : \|\theta_0\|_0 \leq s_0 \right\} \\ &= \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} \left\{ 1 - \Phi\left(z_* - \frac{\sqrt{n}\gamma}{\sigma[\Sigma_{i,i}^{-1}]^{1/2}}\right) + \Phi\left(-z_* - \frac{\sqrt{n}\gamma}{\sigma[\Sigma_{i,i}^{-1}]^{1/2}}\right) \right\} \\ &= \liminf_{n \rightarrow \infty} \frac{1}{1 - \beta_{i,n}^*(\gamma)} G\left(\alpha, \frac{\sqrt{n}\gamma}{\sigma[\Sigma_{i,i}^{-1}]^{1/2}}\right) = 1. \end{aligned}$$

Here (a) follows from Eq. (6.2.1) and the fact  $|\theta_{0,i}| \geq \gamma$ .

### 6.2.2 Proof of Theorem 3.2.1

Let  $\mathcal{F}_{p,s_0} \equiv \{x \in \mathbb{R}^p : \|x\|_0 \leq s_0\}$ , and fix  $\epsilon \in (0, 1/10)$ . By definition,

$$\begin{aligned} \text{FWER}(\widehat{T}^F, n) &= \sup_{\theta_0 \in \mathcal{F}_{p,s_0}} \mathbb{P} \left\{ \exists i \in [p] \setminus \text{supp}(\theta_0), \text{ s.t. } \frac{\sqrt{n}|\widehat{\theta}_i^u - \theta_{0,i}|}{\widehat{\sigma}[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq \Phi^{-1}\left(1 - \frac{\alpha}{2p}\right) \right\} \\ &\leq \sup_{\theta_0 \in \mathcal{F}_{p,s_0}} \mathbb{P} \left\{ \exists i \in [p] \setminus \text{supp}(\theta_0), \text{ s.t. } \frac{\sqrt{n}|\widehat{\theta}_i^u - \theta_{0,i}|}{\sigma[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq (1 - \epsilon)\Phi^{-1}\left(1 - \frac{\alpha}{2p}\right) \right\} \\ &\quad + \sup_{\theta_0 \in \mathcal{F}_{p,s_0}} \mathbb{P} \left( \left| \frac{\widehat{\sigma}}{\sigma} - 1 \right| \geq \epsilon \right). \end{aligned}$$

Since the second term vanishes as  $n \rightarrow \infty$  by Eq. (2.4.1). Using Bonferroni inequality, letting  $z_\alpha(\epsilon) \equiv (1 - \epsilon)\Phi^{-1}(1 - \frac{\alpha}{2p})$ , we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \text{FWER}(\widehat{T}^F, n) &\leq \limsup_{n \rightarrow \infty} \sum_{i=1}^p \sup_{\theta_0 \in \mathcal{F}_{p,s_0}, \theta_{0,i}=0} \mathbb{P} \left\{ \frac{\sqrt{n} |\widehat{\theta}_i^u - \theta_{0,i}|}{\sigma[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \geq z_\alpha(\epsilon) \right\} \\ &= \limsup_{n \rightarrow \infty} \sum_{i=1}^p \sup_{\theta_0 \in \mathcal{F}_{p,s_0}, \theta_{0,i}=0} \mathbb{P} \left\{ \left| \widetilde{Z}_i + \frac{\Delta_i}{\sigma[M\widehat{\Sigma}M^\top]_{i,i}^{1/2}} \right| \geq z_\alpha(\epsilon) \right\}, \end{aligned}$$

where, by Theorem 2.3.5,  $\widetilde{Z}_i \sim \mathcal{N}(0, 1)$  and  $\Delta_i$  is given by Eq. (2.3.13). We then have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \text{FWER}(\widehat{T}^F, n) &\leq \limsup_{n \rightarrow \infty} \sum_{i=1}^p \mathbb{P}\{|\widetilde{Z}_i| \geq z_\alpha(\epsilon) - \epsilon\} \\ &\quad + \limsup_{n \rightarrow \infty} \sum_{i=1}^p \sup_{\theta_0 \in \mathcal{F}_{p,s_0}, \theta_{0,i}=0} \mathbb{P} \left\{ \|\Delta\|_\infty \geq \frac{\epsilon\sigma}{2\widehat{\Sigma}_{ii}^{1/2}} \right\} \\ &\leq 2p(1 - \Phi(z_\alpha(\epsilon) - \epsilon)) + \limsup_{n \rightarrow \infty} p \max_{i \in [p]} \mathbb{P}(\widehat{\Sigma}_{ii} \geq 2) \\ &\quad + \limsup_{n \rightarrow \infty} \sup_{\theta_0 \in \mathcal{F}_{p,s_0}, \theta_{0,i}=0} p \mathbb{P} \left\{ \|\Delta\|_\infty \geq \frac{\epsilon\sigma}{4} \right\}, \end{aligned} \quad (6.2.2)$$

where in the first inequality, we used  $[M\widehat{\Sigma}M^\top]_{i,i} \geq 1/(4\widehat{\Sigma}_{ii})$  for all  $n$  large enough, by Lemma 2.4.1, and since  $\mu = a\sqrt{(\log p)/n} \rightarrow 0$  as  $n, p \rightarrow \infty$ . Now, the second term in the right hand side of Eq. (6.2.2) vanishes by Theorem 2.3.4.(a), and the last term is zero by Theorem 2.3.5, since  $s_0 = o(\sqrt{n}/\log p)$ . Therefore

$$\limsup_{n \rightarrow \infty} \text{FWER}(\widehat{T}^F, n) \leq 2p(1 - \Phi(z_\alpha(\epsilon) - \epsilon)). \quad (6.2.3)$$

The claim follows by letting  $\epsilon \rightarrow 0$ .

### 6.2.3 Proof of Lemma 3.3.6

Fix  $\alpha \in [0, 1]$ ,  $\gamma > 0$ , and assume that the minimum error rate for type II errors in testing hypothesis  $H_{0,i}$  at significance level  $\alpha$  is  $\beta = \beta_{i,n}^{\text{opt}}(\alpha; \gamma)$ . Further fix  $\xi > 0$  arbitrarily small. By definition there exists a statistical test  $T_{i,\mathbf{X}} : \mathbb{R}^m \rightarrow \{0, 1\}$  such that  $\mathbb{P}_\theta(T_{i,\mathbf{X}}(Y) = 1) \leq \alpha$  for any  $\theta \in \mathcal{R}_0$  and  $\mathbb{P}_\theta(T_{i,\mathbf{X}}(Y) = 0) \leq \beta + \xi$  for any  $\theta \in \mathcal{R}_1$  (with  $\mathcal{R}_0, \mathcal{R}_1 \in \mathbb{R}^p$  defined as

in Definition 3.3.5). Equivalently:

$$\begin{aligned}\mathbb{E}\{\mathbb{P}_\theta(T_{i,\mathbf{X}}(Y) = 1|\mathbf{X})\} &\leq \alpha, & \text{for any } \theta \in \mathcal{R}_0, \\ \mathbb{E}\{\mathbb{P}_\theta(T_{i,\mathbf{X}}(Y) = 0|\mathbf{X})\} &\leq \beta + \xi, & \text{for any } \theta \in \mathcal{R}_1.\end{aligned}\tag{6.2.4}$$

We now take expectation of these inequalities with respect to  $\theta \sim Q_0$  (in the first case) and  $\theta \sim Q_1$  (in the second case) and we get, with the notation introduced in the Definition 3.3.5,

$$\begin{aligned}\mathbb{E}\{\mathbb{P}_{Q,0,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 1)\} &\leq \alpha, \\ \mathbb{E}\{\mathbb{P}_{Q,1,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 0)\} &\leq \beta + \xi.\end{aligned}$$

Call  $\alpha_{\mathbf{X}} \equiv \mathbb{P}_{Q,0,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 1)$ . By assumption, for any test  $T$ , we have  $\mathbb{P}_{Q,1,\mathbf{X}}(T_{i,\mathbf{X}}(Y) = 0) \geq \beta_{i,\mathbf{X}}^{\text{bin}}(\alpha_{\mathbf{X}}; Q)$  and therefore the last inequalities imply

$$\begin{aligned}\mathbb{E}\{\alpha_{\mathbf{X}}\} &\leq \alpha, \\ \mathbb{E}\{\beta_{i,\mathbf{X}}^{\text{bin}}(\alpha_{\mathbf{X}}; Q)\} &\leq \beta + \xi.\end{aligned}\tag{6.2.5}$$

The thesis follows since  $\xi > 0$  is arbitrary.

#### 6.2.4 Proof of Lemma 3.3.7

Fix  $\mathbf{X}$ ,  $\alpha$ ,  $i$ ,  $S$  as in the statement and assume, without loss of generality,  $P_{\tilde{S}}^\perp \tilde{x}_i \neq 0$ , and  $\text{rank}(\mathbf{X}_S) = |S| < n$ . We take  $Q_0 = \mathbf{N}(0, J)$  where  $J \in \mathbb{R}^{p \times p}$  is the diagonal matrix with  $J_{jj} = a$  if  $j \in S$  and  $J_{jj} = 0$  otherwise. Here  $a \in \mathbb{R}_+$  and will be chosen later. For the same covariance matrix  $J$ , we let  $Q_1 = \mathbf{N}(\gamma e_i, J)$  where  $e_i$  is the  $i$ -th element of the standard basis. Recalling that  $i \notin S$ , and  $|S| < s_0$ , the support of  $Q_0$  is in  $\mathcal{R}_0$  and the support of  $Q_1$  is in  $\mathcal{R}_1$ .

Under  $\mathbb{P}_{Q,0,\mathbf{X}}$  we have  $Y \sim \mathbf{N}(0, a \mathbf{X}_S \mathbf{X}_S^\top + \sigma^2 \mathbf{I})$ , and under  $\mathbb{P}_{Q,1,\mathbf{X}}$  we have  $Y \sim \mathbf{N}(\gamma \tilde{x}_i, a \mathbf{X}_S \mathbf{X}_S^\top + \sigma^2 \mathbf{I})$ . Hence the binary hypothesis testing problem under study reduces to the problem of testing a null hypothesis on the mean of a Gaussian random vector with known covariance against a simple alternative. It is well known that the most powerful test [89, Chapter 8] is obtained by comparing the ratio  $\mathbb{P}_{Q,0,\mathbf{X}}(Y)/\mathbb{P}_{Q,1,\mathbf{X}}(Y)$  with a threshold.

Equivalently, the most powerful test is of the form

$$T_{i,\mathbf{X}}(Y) = \mathbb{I}\left\{\langle \gamma \tilde{x}_i, (a\mathbf{X}_S\mathbf{X}_S^\top + \sigma^2\mathbf{I})^{-1}Y \rangle \geq c\right\}, \quad (6.2.6)$$

for some  $c \in \mathbb{R}$  that is to be chosen to achieve the desired significance level  $\alpha$ . Letting

$$\alpha \equiv 2\Phi\left(-\frac{c}{\gamma\|(a\mathbf{X}_S\mathbf{X}_S^\top + \sigma^2\mathbf{I})^{-1/2}\tilde{x}_i\|}\right), \quad (6.2.7)$$

it is a straightforward calculation to drive the power of this test as

$$G\left(\alpha, \gamma\|(a\mathbf{X}_S\mathbf{X}_S^\top + \sigma^2\mathbf{I})^{-1/2}\tilde{x}_i\|\right),$$

where the function  $G(\alpha, u)$  is defined as per Eq. (3.3.2). Next we show that the power of this test converges to  $1 - \beta_{i,\mathbf{X}}^{\text{oracle}}(\alpha; S, \gamma)$  as  $a \rightarrow \infty$ . Hence the claim is proved by taking  $a \geq a(\xi)$  for some  $a(\xi)$  large enough.

Write

$$\begin{aligned} (a\mathbf{X}_S\mathbf{X}_S^\top + \sigma^2\mathbf{I})^{-1/2} &= \frac{1}{\sigma} \left( \mathbf{I} + \frac{a}{\sigma^2} \mathbf{X}_S\mathbf{X}_S^\top \right)^{-1/2} \\ &= \frac{1}{\sigma} \left\{ \mathbf{I} - \mathbf{X}_S \left( \frac{\sigma^2}{a} \mathbf{I} + \mathbf{X}_S^\top \mathbf{X}_S \right)^{-1} \mathbf{X}_S^\top \right\}^{1/2}, \end{aligned} \quad (6.2.8)$$

where the second step follows from matrix inversion lemma. Clearly, as  $a \rightarrow \infty$ , the right hand side of the above equation converges to  $(1/\sigma)\mathbf{P}_S^\perp$ . Therefore, the power converges to  $1 - \beta_{i,\mathbf{X}}^{\text{oracle}}(\alpha; S, \gamma) = G(\alpha, \gamma\sigma^{-1}\|\mathbf{P}_S^\perp\tilde{x}_i\|)$ .

### 6.2.5 Proof of Theorem 3.3.3

Let  $u_{\mathbf{X}} \equiv \gamma\|\mathbf{P}_S^\perp\tilde{x}_i\|_2/\sigma$ . By Lemma 3.3.6 and 3.3.7, we have,

$$1 - \beta_{i,n}^{\text{opt}}(\alpha; \gamma) \leq \sup \left\{ \mathbb{E}G(\alpha_{\mathbf{X}}, u_{\mathbf{X}}) : \mathbb{E}(\alpha_{\mathbf{X}}) \leq \alpha \right\}, \quad (6.2.9)$$

with the sup taken over measurable functions  $\mathbf{X} \mapsto \alpha_{\mathbf{X}}$ , and  $G(\alpha, u)$  defined as per Eq. (3.3.2).

It is easy to check that  $\alpha \mapsto G(\alpha, u)$  is concave for any  $u \in \mathbb{R}_+$  and  $u \mapsto G(\alpha, u)$  is

non-decreasing for any  $\alpha \in [0, 1]$ . Further  $G$  takes values in  $[0, 1]$ . Hence

$$\begin{aligned} \mathbb{E}G(\alpha_{\mathbf{X}}, u_{\mathbf{X}}) &\leq \mathbb{E}\{G(\alpha_{\mathbf{X}}, u_{\mathbf{X}})\mathbb{I}(u \leq u_0)\} + \mathbb{P}(u_{\mathbf{X}} > u_0) \\ &\leq \mathbb{E}\{G(\alpha_{\mathbf{X}}, u_0)\} + \mathbb{P}(u_{\mathbf{X}} > u_0) \\ &\leq G(\mathbb{E}(\alpha_{\mathbf{X}}), u_0) + \mathbb{P}(u_{\mathbf{X}} > u_0) \\ &\leq G(\alpha, u_0) + \mathbb{P}(u_{\mathbf{X}} > u_0) \end{aligned} \tag{6.2.10}$$

Since  $\tilde{x}_i$  and  $\mathbf{X}_S$  are jointly Gaussian, we have

$$\tilde{x}_i = \Sigma_{i,S} \Sigma_{S,S}^{-1} \mathbf{X}_S + \Sigma_{i|S}^{1/2} z_i, \tag{6.2.11}$$

with  $z_i \sim \mathbf{N}(0, \mathbf{I}_{n \times n})$  independent of  $\mathbf{X}_S$ . It follows that

$$u_{\mathbf{X}} = (\gamma/\sigma) \Sigma_{i|S}^{1/2} \|\mathbf{P}_S^\perp z_i\|_2 \stackrel{d}{=} (\gamma/\sigma) \sqrt{\Sigma_{i|S} Z_{n-s_0+1}}, \tag{6.2.12}$$

with  $Z_{n-s_0+1}$  a chi-squared random variable with  $n-s_0+1$  degrees of freedom. The desired claim follows by taking  $u_0 = (\gamma/\sigma) \sqrt{\Sigma_{i|S}(n-s_0+\ell)}$ .

## 6.3 Proofs of Theorems in Chapter 4

### 6.3.1 Proof of Theorem 4.1.3

Since  $\{(\Sigma(p) = \mathbf{I}_{p \times p}, \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  has a standard distributional limit, the empirical distribution of  $\{(\theta_{0,i}, \hat{\theta}_i^u)\}_{i=1}^p$  converges weakly to  $(\Theta_0, \Theta_0 + \tau Z)$  (with probability one). By the portmanteau theorem, and the fact that  $\liminf_{p \rightarrow \infty} \sigma(p)/\sqrt{n(p)} = \sigma_0$ , we have

$$\mathbb{P}(0 < |\Theta_0| < \gamma_0 \sigma_0) \leq \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}\left(0 < \theta_{0,i} < \gamma_0 \frac{\sigma(p)}{\sqrt{n(p)}}\right) = 0. \tag{6.3.1}$$

In addition, since  $\gamma_0 \sigma_0/2$  is a continuity point of the distribution of  $\Theta_0$ , we have

$$\lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(|\theta_{0,i}| \geq \frac{\gamma_0 \sigma_0}{2}) = \mathbb{P}(|\Theta_0| \geq \frac{\gamma_0 \sigma_0}{2}). \tag{6.3.2}$$

Now, by Eq. (6.3.1),  $\mathbb{P}(|\Theta_0| \geq \gamma_0 \sigma_0 / 2) = \mathbb{P}(\Theta_0 \neq 0)$ . Further,  $\mathbb{I}(|\theta_{0,i}| \geq \gamma_0 \sigma_0 / 2) = \mathbb{I}(\theta_{0,i} \neq 0)$  for  $1 \leq i \leq p$ , as  $p \rightarrow \infty$ . Therefore, Eq. (6.3.2) yields

$$\lim_{p \rightarrow \infty} \frac{1}{p} |S_0(p)| = \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(\theta_{0,i} \neq 0) = \mathbb{P}(\Theta_0 \neq 0). \quad (6.3.3)$$

Hence,

$$\begin{aligned} \lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} T_{i,\mathbf{X}}(Y) &= \lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{I}(P_i \leq \alpha) \\ &= \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(P_i \leq \alpha, i \in S_0(p)) \\ &= \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \frac{|\hat{\theta}_i^u|}{\tau}, |\theta_{0,i}| \geq \frac{\gamma_0 \sigma(p)}{\sqrt{n(p)}}\right) \\ &\geq \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \left|\frac{\Theta_0}{\tau} + Z\right|, |\Theta_0| \geq \gamma_0 \sigma_0\right). \end{aligned}$$

Note that  $\tau$  depends on the distribution  $p_{\Theta_0}$ . Since  $|S_0(p)| \leq \varepsilon p$ , using Eq. (6.3.3), we have  $\mathbb{P}(\Theta_0 \neq 0) \leq \varepsilon$ , i.e.  $p_{\Theta_0}$  is  $\varepsilon$ -sparse. Let  $\tilde{\tau}$  denote the maximum  $\tau$  corresponding to densities in the family of  $\varepsilon$ -sparse densities. As shown in [42],  $\tilde{\tau} = \tau_* \sigma_0$ , where  $\tau_*$  is defined by Eqs. (4.1.7) and (4.1.8). Consequently,

$$\begin{aligned} \lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} T_{i,\mathbf{X}}(Y) &\geq \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \left|\frac{\gamma_0}{\tau_*} + Z\right|\right) \\ &= 1 - \mathbb{P}\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \frac{\gamma_0}{\tau_*} \leq Z \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \frac{\gamma_0}{\tau_*}\right) \\ &= 1 - \left\{\Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \gamma_0/\tau_*\right) - \Phi\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \gamma_0/\tau_*\right)\right\} \\ &= G\left(\alpha, \frac{\gamma_0}{\tau_*}\right). \end{aligned} \quad (6.3.4)$$

Now, we take the expectation of both sides of Eq. (6.3.4) with respect to the law of random design  $\mathbf{X}$  and random noise  $W$ . Changing the order of limit and expectation by applying dominated convergence theorem and using linearity of expectation, we obtain

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{E}_{\mathbf{X},W} \{T_{i,\mathbf{X}}(Y)\} \geq G\left(\alpha, \frac{\gamma_0}{\tau_*}\right). \quad (6.3.5)$$

Since  $T_{i,\mathbf{X}}(Y)$  takes values in  $\{0, 1\}$ , we have  $\mathbb{E}_{\mathbf{X},W}\{T_{i,\mathbf{X}}(Y)\} = \mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1)$ . The result follows by noting that the columns of  $\mathbf{X}$  are exchangeable and therefore  $\mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1)$  does not depend on  $i$ .

### 6.3.2 Proof of Theorem 4.2.3

Since the sequence  $\{\Sigma(p), \theta_0(p), n(p), \sigma(p)\}_{p \in \mathbb{N}}$  has a standard distributional limit, with probability one the empirical distribution of  $\{(\theta_{0,i}, \widehat{\theta}_i^u, (\Sigma^{-1})_{ii})\}_{i=1}^p$  converges weakly to the distribution of  $(\Theta_0, \Theta_0 + \tau\Upsilon^{1/2}Z, \Upsilon)$ . Therefore, with probability one, the empirical distribution of

$$\left\{ \frac{\widehat{\theta}_i^u - \theta_{0,i}}{\tau[(\Sigma^{-1})_{ii}]^{1/2}} \right\}_{i=1}^p$$

converges weakly to  $\mathbf{N}(0, 1)$ . Hence,

$$\begin{aligned} \lim_{p \rightarrow \infty} \frac{1}{|S_0^c(p)|} \sum_{i \in S_0^c(p)} T_{i,\mathbf{X}}(Y) &= \lim_{p \rightarrow \infty} \frac{1}{|S_0^c(p)|} \sum_{i \in S_0^c(p)} \mathbb{I}(P_i \leq \alpha) \\ &= \frac{1}{\mathbb{P}(\Theta_0 = 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(P_i \leq \alpha, i \in S_0^c(p)) \\ &= \frac{1}{\mathbb{P}(\Theta_0 = 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \frac{|\widehat{\theta}_i^u|}{\tau[(\Sigma^{-1})_{ii}]^{1/2}}, \theta_{0,i} = 0\right) \\ &= \frac{1}{\mathbb{P}(\Theta_0 = 0)} \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq |Z|, \Theta_0 = 0\right) \\ &= \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq |Z|\right) = \alpha. \end{aligned}$$

Applying the same argument as in the proof of Theorem 4.1.3, we obtain the following by taking the expectation of both sides of the above equation

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1) = \alpha. \quad (6.3.6)$$

In particular, for the standard Gaussian design (cf. Theorem 4.1.2), since the columns of  $\mathbf{X}$  are exchangeable we get  $\lim_{p \rightarrow \infty} \mathbb{P}_{\theta_0(p)}(T_{i,\mathbf{X}}(Y) = 1) = \alpha$  for all  $i \in S_0(p)$ .

### 6.3.3 Proof of Theorem 4.2.4

The proof of Theorem 4.2.4 proceeds along the same lines as the proof of Theorem 4.1.3. Since  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$  has a standard distributional limit, with probability one the empirical distribution of  $\{(\theta_{0,i}, \hat{\theta}_i^u, (\Sigma^{-1})_{ii})\}_{i=1}^p$  converges weakly to the distribution of  $(\Theta_0, \Theta_0 + \tau\Upsilon^{1/2}Z, \Upsilon)$ . Similar to Eq. (6.3.3), we have

$$\lim_{p \rightarrow \infty} \frac{1}{p} |S_0(p)| = \mathbb{P}(\Theta_0 \neq 0). \quad (6.3.7)$$

Also

$$\begin{aligned} & \lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} T_{i, \mathbf{X}}(Y) \\ &= \lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{I}(P_i \leq \alpha) \\ &= \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(P_i \leq \alpha, i \in S_0(p)) \\ &= \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \frac{|\hat{\theta}_i^u|}{\tau[(\Sigma^{-1})_{ii}]^{1/2}}, \frac{|\theta_{0,i}|}{[(\Sigma^{-1})_{ii}]^{1/2}} \geq \gamma_0\right) \\ &= \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \left|\frac{\Theta_0}{\tau\Upsilon^{1/2}} + Z\right|, \frac{|\Theta_0|}{\Upsilon^{1/2}} \geq \gamma_0\right) \\ &\geq \frac{1}{\mathbb{P}(\Theta_0 \neq 0)} \mathbb{P}\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \left|\frac{\gamma_0}{\tau} + Z\right|\right) \\ &= 1 - \left\{\Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \frac{\gamma_0}{\tau}\right) - \Phi\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \frac{\gamma_0}{\tau}\right)\right\} \\ &= G\left(\alpha, \frac{\gamma_0}{\tau}\right). \end{aligned} \quad (6.3.8)$$

Similar to the proof of Theorem 4.1.3, by taking the expectation of both sides of the above inequality we get

$$\lim_{p \rightarrow \infty} \frac{1}{|S_0(p)|} \sum_{i \in S_0(p)} \mathbb{P}_{\theta_0}(T_{i, \mathbf{X}}(Y) = 1) \geq G\left(\alpha, \frac{\gamma_0}{\tau}\right). \quad (6.3.9)$$

## Part II

# Optimal Compressed Sensing via Spatial Coupling and Approximate Message Passing

## Chapter 7

# Reconstruction at Optimal Rate

We consider noisy linear model wherein  $m$  linear measurements are taken of an unknown  $n$ -dimensional signal  $x \in \mathbb{R}^n$  as per the following

$$y = Ax + w, \tag{7.0.1}$$

where  $y \in \mathbb{R}^m$  denotes the vector of measurements,  $A \in \mathbb{R}^{m \times n}$  represents the measurement (sensing) matrix, and  $w \in \mathbb{R}^n$  is the measurement noise.

The reconstruction problem requires to reconstruct  $x$  given  $y$  and  $A$ . We let  $\delta = \lim_{n \rightarrow \infty} m/n$  be the asymptotic undersampling rate. In Part II of this dissertation, we present a construction of matrix  $A$  and a reconstruction algorithm that solves the reconstruction problem at information-theoretically optimal rate, namely  $\delta > \bar{d}(p_X)$  for exact reconstruction of  $x$  from noiseless measurements ( $w = 0$ ), and  $\delta > \bar{D}(p_X)$  for robust reconstruction under the noisy model. Here, it is assumed that the empirical distribution of the entries of  $x$  converges weakly to a probability measure  $p_X$  on the real line and  $\bar{d}(p_X)$  and  $\bar{D}(p_X)$  respectively denote the (upper) Rényi information dimension and the (upper) MMSE dimension of  $p_X$ , defined below. In the next section, we state formally our results and discuss their implications and limitations, as well as relations with earlier work.

### 7.1 Formal statement of the results

We consider the noisy model (1.3.2). An instance of the problem is therefore completely specified by the triple  $(x, w, A)$ . We will be interested in the asymptotic properties of

sequence of instances indexed by the problem dimensions  $\mathcal{S} = \{(x(n), w(n), A(n))\}_{n \in \mathbb{N}}$ . We recall a definition from [8]. (More precisely, [8] introduces the  $B = 1$  case of this definition.)

**Definition 7.1.1.** *The sequence of instances  $\mathcal{S} = \{x(n), w(n), A(n)\}_{n \in \mathbb{N}}$  indexed by  $n$  is said to be a  $B$ -converging sequence if  $x(n) \in \mathbb{R}^n$ ,  $w(n) \in \mathbb{R}^m$ ,  $A(n) \in \mathbb{R}^{m \times n}$  with  $m = m(n)$  is such that  $m/n \rightarrow \delta \in (0, \infty)$ , and in addition the following conditions hold<sup>1</sup>:*

- (a) *The empirical distribution of the entries of  $x(n)$  converges weakly to a probability measure  $p_X$  on  $\mathbb{R}$  with bounded second moment. Further  $n^{-1} \sum_{i=1}^n x_i(n)^2 \rightarrow \mathbb{E}\{X^2\}$ , where the expectation is taken with respect to  $p_X$ .*
- (b) *The empirical distribution of the entries of  $w(n)$  converges weakly to a probability measure  $p_W$  on  $\mathbb{R}$  with bounded second moment. Further  $m^{-1} \sum_{i=1}^m w_i(n)^2 \rightarrow \mathbb{E}\{W^2\} \equiv \sigma^2$ , where the expectation is taken with respect to  $p_W$ .*
- (c) *If  $\{e_i\}_{1 \leq i \leq n}$ ,  $e_i \in \mathbb{R}^n$  denotes the canonical basis, then  $\limsup_{n \rightarrow \infty} \max_{i \in [n]} \|A(n)e_i\|_2 \leq B$ ,  $\liminf_{n \rightarrow \infty} \min_{i \in [n]} \|A(n)e_i\|_2 \geq 1/B$ .*

We further say that  $\{(x(n), w(n))\}_{n \geq 0}$  is a converging sequence of instances, if they satisfy conditions (a) and (b). We say that  $\{A(n)\}_{n \geq 0}$  is a  $B$ -converging sequence of sensing matrices if they satisfy condition (c) above, and we call it a converging sequence if it is  $B$ -converging for some  $B$ . Similarly, we say  $\mathcal{S}$  is a converging sequence if it is  $B$ -converging for some  $B$ .

Finally, if the sequence  $\{(x(n), w(n), A(n))\}_{n \geq 0}$  is random, the above conditions are required to hold almost surely.

Notice that standard normalizations of the sensing matrix correspond to  $\|A(n)e_i\|_2^2 = 1$  (and hence  $B = 1$ ) or to  $\|A(n)e_i\|_2^2 = m(n)/n$ . The former corresponds to normalized columns and the latter corresponds to normalized rows. Since throughout we assume  $m(n)/n \rightarrow \delta \in (0, \infty)$ , these conventions only differ by a rescaling of the noise variance. In order to simplify the proofs, we allow ourselves somewhat more freedom by taking  $B$  a fixed constant.

---

<sup>1</sup>If  $(\mu_k)_{k \in \mathbb{N}}$  is a sequence of measures and  $\mu$  is another measure, all defined on  $\mathbb{R}$ , the weak convergence of  $\mu_k$  to  $\mu$  along with the convergence of their second moments to the second moment of  $\mu$  is equivalent to convergence in 2-Wasserstein distance [138]. Therefore, conditions (a)-(b) are equivalent to the following. The empirical distributions of the signal  $x(n)$  and the empirical distributions of noise  $w(n)$  converge in 2-Wasserstein distance.

Given a sensing matrix  $A$ , and a vector of measurements  $y$ , a reconstruction algorithm produces an estimate  $\hat{x}(A; y) \in \mathbb{R}^n$  of  $x$ . In our work we assume that the empirical distribution  $p_X$ , and the noise level  $\sigma^2$  are known to the estimator, and hence the mapping  $\hat{x} : (A, y) \mapsto \hat{x}(A; y)$  implicitly depends on  $p_X$  and  $\sigma^2$ . Since however  $p_X, \sigma^2$  are fixed throughout, we avoid the cumbersome notation  $\hat{x}(A, y, p_X, \sigma^2)$ .

Given a converging sequence of instances  $\mathcal{S} = \{x(n), w(n), A(n)\}_{n \in \mathbb{N}}$ , and an estimator  $\hat{x}$ , we define the asymptotic per-coordinate reconstruction mean square error as

$$\text{MSE}(\mathcal{S}; \hat{x}) = \limsup_{n \rightarrow \infty} \frac{1}{n} \|\hat{x}(A(n); y(n)) - x(n)\|^2. \quad (7.1.1)$$

Notice that the quantity on the right hand side depends on the matrix  $A(n)$ , which will be random, and on the signal and noise vectors  $x(n), w(n)$  which can themselves be random. Our results hold almost surely with respect to these random variables. In some applications it is more customary to take the expectation with respect to the noise and signal distribution, i.e., to consider the quantity

$$\overline{\text{MSE}}(\mathcal{S}; \hat{x}) = \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \|\hat{x}(A(n); y(n)) - x(n)\|^2. \quad (7.1.2)$$

It turns out that the almost sure bounds imply, in the present setting, bounds on the expected mean square error  $\overline{\text{MSE}}$ , as well.

In this part we study a specific low-complexity estimator, based on the AMP algorithm first proposed in [42]. AMP is an iterative algorithm derived from the theory of belief propagation in graphical models [99]. At each iteration  $t$ , it keeps track of an estimate  $x^t \in \mathbb{R}^n$  of the unknown signal  $x$ . This is used to compute residuals  $(y - Ax^t) \in \mathbb{R}^m$ . These correspond to the part of observations that is not explained by the current estimate  $x^t$ . The residuals are then processed through a matched filter operator (roughly speaking, this amounts to multiplying the residuals by the transpose of  $A$ ) and then applying a non-linear denoiser, to produce the new estimate  $x^{t+1}$ .

Formally, we start with an initial guess  $x_i^1 = \mathbb{E}\{X\}$  for all  $i \in [n]$  and proceed by

$$x^{t+1} = \eta_t(x^t + (Q^t \odot A)^* r^t), \quad (7.1.3)$$

$$r^t = y - Ax^t + \mathbf{b}^t \odot r^{t-1}. \quad (7.1.4)$$

The second equation corresponds to the computation of new residuals from the current

estimate. The memory term (also known as ‘Onsager term’ in statistical physics) plays a crucial role as emphasized in [42, 7, 6, 74]. The first equation describes matched filter, with multiplication by  $(Q_t \odot A)^*$ , followed by application of the denoiser  $\eta_t$ . Throughout  $\odot$  indicates Hadamard (entrywise) product and  $X^*$  denotes the transpose of matrix  $X$ .

For each  $t$ , the denoiser  $\eta_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a differentiable non-linear function that depends on the input distribution  $p_X$ . Further,  $\eta_t$  is separable<sup>2</sup>, namely, for a vector  $v \in \mathbb{R}^n$ , we have  $\eta_t(v) = (\eta_{1,t}(v_1), \dots, \eta_{n,t}(v_n))$ . The matrix  $Q^t \in \mathbb{R}^{m \times n}$  and the vector  $\mathbf{b}^t \in \mathbb{R}^m$  can be efficiently computed from the current state  $x^t$  of the algorithm, Further  $Q^t$  does not depend on the problem instance and hence can be precomputed. Both  $Q^t$  and  $\mathbf{b}^t$  are block-constants, i.e., they can be partitioned into blocks such that within each block all the entries have the same value. This property makes their evaluation, storage and manipulation particularly convenient.

We refer to the next section for explicit definitions of these quantities. A crucial element is the specific choice of  $\eta_{i,t}$ . The general guiding principle is that the argument  $y^t = x^t + (Q^t \odot A)^* r^t$  in Eq. (7.1.3) should be interpreted as a noisy version of the unknown signal  $x$ , i.e.,  $y^t = x + \text{noise}$ . The denoiser  $\eta_t$  must therefore be chosen as to minimize the mean square error at iteration  $(t + 1)$ . The papers [42, 33] take a minimax point of view, and hence study denoisers that achieve the smallest mean square error over the worst case signal  $x$  in a certain class. For instance, coordinate-wise soft thresholding is nearly minimax optimal over the class of sparse signals [33]. Here we instead assume that the prior  $p_X$  is known, and hence the choice of  $\eta_{i,t}$  is uniquely dictated by the objective of minimizing the mean square error at iteration  $t + 1$ . In other words  $\eta_{i,t}$  takes the form of a Bayes optimal estimator for the prior  $p_X$ . In order to stress this point, we will occasionally refer to this as the Bayes optimal AMP algorithm. As shown in Appendix C.2,  $x^t$  is (almost surely) a local Lipschitz continuous function of the observations  $y$ .

Finally notice that [43, 99] also derived AMP starting from a Bayesian graphical models point of view, with the signal  $x$  modeled as random with i.i.d. entries. The algorithm in Eqs. (7.1.3), (7.1.4) differs from the one in [43] in that the matched filter operation requires scaling  $A$  by the matrix  $Q^t$ . This is related to the fact that we will use a matrix  $A$  with independent but not identically distributed entries and, as a consequence, the accuracy of each entry  $x_i^t$  depends on the index  $i$  as well as on  $t$ .

We denote by  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2)$  the mean square error achieved by the Bayes optimal

---

<sup>2</sup>We refer to [33] for a study of non-separable denoisers in AMP algorithms.

AMP algorithm, where we made explicit the dependence on  $\sigma^2$ . Since the AMP estimate depends on the iteration number  $t$ , the definition of  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2)$  requires some care. The basic point is that we need to iterate the algorithm only for a constant number of iterations, as  $n$  gets large. Formally, we let

$$\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) \equiv \lim_{t \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \|x^t(A(n); y(n)) - x(n)\|^2. \quad (7.1.5)$$

As discussed above, limits will be shown to exist almost surely, when the instances  $(x(n), w(n), A(n))$  are random, and almost sure upper bounds on  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2)$  will be proved. (Indeed  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2)$  turns out to be deterministic.) On the other hand, one might be interested in the expected error

$$\overline{\text{MSE}}_{\text{AMP}}(\mathcal{S}; \sigma^2) \equiv \lim_{t \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}\{\|x^t(A(n); y(n)) - x(n)\|^2\}. \quad (7.1.6)$$

### 7.1.1 Rényi information dimension

We will tie the success of our compressed sensing scheme to the fundamental information-theoretic limit established in [143]. The latter is expressed in terms of the Rényi information dimension of the probability measure  $p_X$ .

**Definition 7.1.2.** *Let  $p_X$  be a probability measure over  $\mathbb{R}$ , and  $X \sim p_X$ . The upper and lower information dimension of  $p_X$  are defined as*

$$\bar{d}(p_X) = \limsup_{\ell \rightarrow \infty} \frac{H([X]_\ell)}{\log \ell}. \quad (7.1.7)$$

$$\underline{d}(p_X) = \liminf_{\ell \rightarrow \infty} \frac{H([X]_\ell)}{\log \ell}. \quad (7.1.8)$$

Here  $H(\cdot)$  denotes Shannon entropy and, for  $x \in \mathbb{R}$ ,  $[x]_\ell \equiv \lfloor \ell x \rfloor / \ell$ , and  $\lfloor x \rfloor \equiv \max\{k \in \mathbb{Z} : k \leq x\}$ . If the  $\limsup$  and  $\liminf$  coincide, then we let  $d(p_X) = \bar{d}(p_X) = \underline{d}(p_X)$ .

Whenever the limit of  $H([X]_\ell)/\log \ell$  exists and is finite<sup>3</sup>, the Rényi information dimension can also be characterized as follows. Write the binary expansion of  $X$ ,  $X = D_0.D_1D_2D_3\dots$  with  $D_i \in \{0, 1\}$  for  $i \geq 1$ . Then  $\bar{d}(p_X)$  is the entropy rate of the stochastic process  $\{D_1, D_2, D_3, \dots\}$ . It is also convenient to recall the following result from [114, 143].

<sup>3</sup>This condition can be replaced by  $H(\lfloor X \rfloor) < \infty$ . A sufficient condition is that  $\mathbb{E}[\log(1 + |X|)] < \infty$ , which is certainly satisfied if  $X$  has a finite variance [144].

**Proposition 7.1.3** ([114, 143]). *Let  $p_X$  be a probability measure over  $\mathbb{R}$ , and  $X \sim p_X$ . Assume  $H(\lfloor X \rfloor)$  to be finite. If  $p_X = (1 - \varepsilon)\nu_d + \varepsilon\tilde{\nu}$  with  $\nu_d$  a discrete distribution (i.e., with countable support), then  $\bar{d}(p_X) \leq \varepsilon$ . Further, if  $\tilde{\nu}$  has a density with respect to Lebesgue measure, then  $d(p_X) = \bar{d}(p_X) = \underline{d}(p_X) = \varepsilon$ . In particular, if  $\mathbb{P}\{X \neq 0\} \leq \varepsilon$  then  $\bar{d}(p_X) \leq \varepsilon$ .*

### 7.1.2 MMSE dimension

In order to present our result concerning the robust reconstruction, we need the definition of *MMSE dimension* of the probability measure  $p_X$ .

Given the signal distribution  $p_X$ , we let  $\text{mmse}(s)$  denote the minimum mean square error in estimating  $X \sim p_X$  from a noisy observation in gaussian noise, at signal-to-noise ratio  $s$ . Formally

$$\text{mmse}(s) \equiv \inf_{\eta: \mathbb{R} \rightarrow \mathbb{R}} \mathbb{E}\{[X - \eta(\sqrt{s}X + Z)]^2\}, \quad (7.1.9)$$

where  $Z \sim \mathcal{N}(0, 1)$ . Since the minimum mean square error estimator is just the conditional expectation, this is given by

$$\text{mmse}(s) = \mathbb{E}\{[X - \mathbb{E}[X|Y]]^2\}, \quad Y = \sqrt{s}X + Z. \quad (7.1.10)$$

Notice that  $\text{mmse}(s)$  is naturally well defined for  $s = \infty$ , with  $\text{mmse}(\infty) = 0$ . We will therefore interpret it as a function  $\text{mmse} : \bar{\mathbb{R}}_+ \rightarrow \bar{\mathbb{R}}_+$  where  $\bar{\mathbb{R}}_+ \equiv [0, \infty]$  is the completed non-negative real line.

We recall the inequality

$$0 \leq \text{mmse}(s) \leq \frac{1}{s}, \quad (7.1.11)$$

obtained by the estimator  $\eta(y) = y/\sqrt{s}$ . A finer characterization of the scaling of  $\text{mmse}(s)$  is provided by the following definition.

**Definition 7.1.4** ([144]). *The upper and lower MMSE dimension of the probability measure  $p_X$  over  $\mathbb{R}$  are defined as*

$$\bar{D}(p_X) = \limsup_{s \rightarrow \infty} s \cdot \text{mmse}(s), \quad (7.1.12)$$

$$\underline{D}(p_X) = \liminf_{s \rightarrow \infty} s \cdot \text{mmse}(s). \quad (7.1.13)$$

If the lim sup and lim inf coincide, then we let  $D(p_X) = \overline{D}(p_X) = \underline{D}(p_X)$ .

It is also convenient to recall the following result from [144].

**Proposition 7.1.5** ([144]). *If  $H(\lfloor X \rfloor) < \infty$ , then*

$$\underline{D}(p_X) \leq \underline{d}(p_X) \leq \overline{d}(p_X) \leq \overline{D}(p_X). \quad (7.1.14)$$

Hence, if  $D(p_X)$  exists, then  $d(p_X)$  exists and  $D(p_X) = d(p_X)$ . In particular, this is the case if  $p_X = (1 - \varepsilon)\nu_d + \varepsilon\tilde{\nu}$  with  $\nu_d$  a discrete distribution (i.e., with countable support), and  $\tilde{\nu}$  has a density with respect to Lebesgue measure.

### 7.1.3 Main results

We are now in position to state our main results. The first one states that for any under-sampling rate above Renyi information dimension  $\delta > \overline{d}(p_X)$ , we have  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) \rightarrow 0$  as  $\sigma^2 \rightarrow 0$  with, in particular,  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2 = 0) = 0$ .

**Theorem 7.1.6.** *Let  $p_X$  be a probability measure on the real line and assume*

$$\delta > \overline{d}(p_X). \quad (7.1.15)$$

Then there exists a random converging sequence of sensing matrices  $\{A(n)\}_{n \geq 0}$ ,  $A(n) \in \mathbb{R}^{m \times n}$ ,  $m(n)/n \rightarrow \delta$  (with distribution depending only on  $\delta$ ), for which the following holds. For any  $\varepsilon > 0$ , there exists  $\sigma_0 = \sigma_0(\varepsilon, \delta, p_X)$  such that for any converging sequence of instances  $\{(x(n), w(n))\}_{n \geq 0}$  with parameters  $(p_X, \sigma^2, \delta)$  and  $\sigma \in [0, \sigma_0]$ , we have, almost surely

$$\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) \leq \varepsilon. \quad (7.1.16)$$

Further, under the same assumptions, we have  $\overline{\text{MSE}}_{\text{AMP}}(\mathcal{S}; \sigma^2) \leq \varepsilon$ .

The second theorem characterizes the rate at which the mean square error goes to 0. In particular, we show that  $\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) = O(\sigma^2)$  provided  $\delta > \overline{D}(p_X)$ .

**Theorem 7.1.7.** *Let  $p_X$  be a probability measure on the real line and assume*

$$\delta > \overline{D}(p_X). \quad (7.1.17)$$

Then there exists a random converging sequence of sensing matrices  $\{A(n)\}_{n \geq 0}$ ,  $A(n) \in \mathbb{R}^{m \times n}$ ,  $m(n)/n \rightarrow \delta$  (with distribution depending only on  $\delta$ ) and a finite stability constant  $C = C(p_X, \delta)$ , such that the following is true. For any converging sequence of instances  $\{(x(n), w(n))\}_{n \geq 0}$  with parameters  $(p_X, \sigma^2, \delta)$ , we have, almost surely

$$\text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) \leq C \sigma^2. \quad (7.1.18)$$

Further, under the same assumptions, we have  $\overline{\text{MSE}}_{\text{AMP}}(\mathcal{S}; \sigma^2) \leq C \sigma^2$ .

Finally, the sensitivity to small noise is bounded as

$$\limsup_{\sigma \rightarrow 0} \frac{1}{\sigma^2} \text{MSE}_{\text{AMP}}(\mathcal{S}; \sigma^2) \leq \frac{4\delta - 2\overline{D}(p_X)}{\delta - \overline{D}(p_X)}. \quad (7.1.19)$$

The performance guarantees in Theorems 7.1.6 and 7.1.7 are achieved with special constructions of the sensing matrices  $A(n)$ . These are matrices with independent Gaussian entries with unequal variances (heteroscedastic entries), with a band diagonal structure. The motivation for this construction, and connection with coding theory is further discussed in Section 9.1, while formal definitions are given in Section 8.1 and 8.4.

Notice that, by Proposition 7.1.5,  $\overline{D}(p_X) \geq \overline{d}(p_X)$ , and  $\overline{D}(p_X) = \overline{d}(p_X)$  for a broad class of probability measures  $p_X$ , including all measures that do not have a singular continuous component (i.e., decomposes into a pure point mass component and an absolutely continuous component).

The noiseless model (1.3.1) is covered as a special case of Theorem 7.1.6 by taking  $\sigma^2 \downarrow 0$ . For the reader's convenience, we state the result explicitly as a corollary.

**Corollary 7.1.8.** *Let  $p_X$  be a probability measure on the real line. Then, for any  $\delta > \overline{d}(p_X)$  there exists a random converging sequence of sensing matrices  $\{A(n)\}_{n \geq 0}$ ,  $A(n) \in \mathbb{R}^{m \times n}$ ,  $m(n)/n \rightarrow \delta$  (with distribution depending only on  $\delta$ ) such that, for any sequence of vectors  $\{x(n)\}_{n \geq 0}$  whose empirical distribution converges to  $p_X$ , the Bayes optimal AMP asymptotically almost surely recovers  $x(n)$  from  $m(n)$  measurements  $y = A(n)x(n) \in \mathbb{R}^{m(n)}$ . (By ‘asymptotically almost surely’ we mean  $\text{MSE}_{\text{AMP}}(\mathcal{S}; 0) = 0$  almost surely, and  $\overline{\text{MSE}}_{\text{AMP}}(\mathcal{S}; 0) = 0$ .)*

Note that it would be interesting to prove a stronger guarantee in the noiseless case, namely  $\lim_{t \rightarrow \infty} x^t(A(n); y(n)) = x(n)$  with probability converging to 1 as  $n \rightarrow \infty$ . The present work does not lead to a proof of this statement.

## 7.2 Discussion

Theorem 7.1.6 and Corollary 7.1.8 are, in many ways, puzzling. It is instructive to spell out in detail a few specific examples, and discuss interesting features.

**Example 1 (Bernoulli-Gaussian signal).** Consider a Bernoulli-Gaussian distribution

$$p_X = (1 - \varepsilon) \delta_0 + \varepsilon \gamma_{\mu, \sigma} \quad (7.2.1)$$

where  $\gamma_{\mu, \sigma}(dx) = (2\pi\sigma^2)^{-1/2} \exp\{-(x-\mu)^2/(2\sigma^2)\}dx$  is the Gaussian measure with mean  $\mu$  and variance  $\sigma^2$ . This model has been studied numerically in a number of papers, including [4, 82]. By Proposition 7.1.3, we have  $\bar{d}(p_X) = \varepsilon$ , and by Proposition 7.1.5,  $\bar{D}(p_X) = \underline{D}(p_X) = \varepsilon$  as well.

Construct random signals  $x(n) \in \mathbb{R}^n$  by sampling i.i.d. coordinates  $x(n)_i \sim p_X$ . Glivenko-Cantelli's theorem implies that the empirical distribution of the coordinates of  $x(n)$  converges almost surely to  $p_X$ , hence we can apply Corollary 7.1.8 to recover  $x(n)$  from  $m(n) = n\varepsilon + o(n)$  spatially coupled measurements  $y(n) \in \mathbb{R}^{m(n)}$ . Notice that the number of non-zero entries in  $x(n)$  is, almost surely,  $k(n) = n\varepsilon + o(n)$ . Hence, we can restate the implication of Corollary 7.1.8 as follows. A sequence of vectors  $x(n)$  with Bernoulli-Gaussian distribution and  $k(n)$  nonzero entries can almost surely be recovered by  $m(n) = k(n) + o(n)$  spatially coupled measurements.

**Example 2 (Mixture signal with a point mass).** The above remarks generalize immediately to arbitrary mixture distributions of the form

$$p_X = (1 - \varepsilon) \delta_0 + \varepsilon q, \quad (7.2.2)$$

where  $q$  is a measure that is absolutely continuous with respect to Lebesgue measure, i.e.,  $q(dx) = f(x)dx$  for some measurable function  $f$ . Then, by Proposition 7.1.3, we have  $\bar{d}(p_X) = \varepsilon$ , and by Proposition 7.1.5,  $\bar{D}(p_X) = \underline{D}(p_X) = \varepsilon$  as well. Similar to Example 1, the number of non-zero entries in  $x(n)$  is, almost surely,  $k(n) = n\varepsilon + o(n)$ , and we can recover  $x(n)$  from  $m(n) = n\varepsilon + o(n)$  spatially coupled measurements. This can be recast as follows.

**Corollary 7.2.1.** *Let  $\{x(n)\}_{n \geq 0}$  be a sequence of vectors with i.i.d. components  $x(n)_i \sim p_X$*

where  $p_X$  is a mixture distribution as per Eq. (7.2.2). Denote by  $k(n)$  the number of nonzero entries in  $x(n)$ . Then, almost surely as  $n \rightarrow \infty$ , Bayes optimal AMP recovers the signal  $x(n)$  from  $m(n) = k(n) + o(n)$  spatially coupled measurements.

Under the regularity hypotheses of [143], no scheme can do substantially better, i.e., reconstruct  $x(n)$  from  $m(n)$  measurements if  $\limsup_{n \rightarrow \infty} m(n)/k(n) < 1$ .

One way to think about this result is the following. If an oracle gave us the support of  $x(n)$ , we would still need  $m(n) \geq k(n) - o(n)$  measurements to reconstruct the signal. Indeed, the entries in the support have distribution  $q$ , and  $\bar{d}(q) = 1$ . Corollary 7.1.8 implies that the measurements overhead for estimating the support of  $x(n)$  is sublinear,  $o(n)$ , even when the support is of order  $n$ .

It is sometimes informally argued that compressed sensing requires at least  $\Theta(k \log(n/k))$  for ‘information-theoretic reasons’, namely that specifying the support requires about  $nH(k/n) \approx k \log(n/k)$  bits. This argument is of course incomplete because it assumes that each measurement  $y_i$  is described by a bounded number of bits. Since it is folklore to say that sparse signal recovery requires  $m \geq C k \log(n/k)$  measurement, it is instructive to survey the results of this type and explain why they do not apply to the present setting. This elucidates further the implications of our results.

Specifically, [139, 1] prove information-theoretic lower bounds on the required number of measurements, under specific constructions for the random sensing matrix  $A$ . Further, these papers focus on the specific problem of exact support recovery. The paper [110] proves minimax bounds for reconstructing vectors belonging to  $\ell_p$ -balls. Notice that these bounds are usually proved by exhibiting a least favorable prior, which is close to a signal with i.i.d. coordinates. However, as the noise variance tends to zero, these bounds depend on the sensing matrix in a way that is difficult to quantify. In particular, they provide no explicit lower bound on the number of measurements required for exact recovery in the noiseless limit. Similar bounds were obtained for arbitrary measurement matrices in [22]. Again, these lower bounds vanish as noise tends to zero as soon as  $m(n) \geq k(n)$ .

A different line of work derives lower bounds from Gelfand’ width arguments [36, 79]. These lower bounds are only proved to be a necessary condition for a stronger reconstruction guarantee. Namely, these works require the vector of measurements  $y = Ax$  to enable recovery for *all*  $k$ -sparse vectors  $x \in \mathbb{R}^n$ . This corresponds to the ‘strong’ phase transition of [45, 37], and is also referred to as the ‘for all’ guarantee in the computer science literature [12].

The lower bound that comes closest to the present setting is the ‘randomized’ lower bound [3]. In this work the authors consider a fixed signal  $x$  and a random sensing matrix as in our setting. In other words they do not assume a standard minimax setting. However they require an  $\ell_1 - \ell_1$  error guarantee which is a stronger stability condition than what is achieved in Theorem 7.1.7, allowing for a more powerful noise process. Indeed the same paper also proves that recovery is possible from  $m(n) = O(k(n))$  measurements under stronger conditions.

**Example 3 (Discrete signal).** Let  $K$  be a fixed integer,  $a_1, \dots, a_K \in \mathbb{R}$ , and  $(p_1, p_2, \dots, p_K)$  be a collection of non-negative numbers that add up to one. Consider the probability distribution that puts mass  $p_i$  on each  $a_i$

$$p_X = \sum_{i=1}^K p_i \delta_{a_i}, \quad (7.2.3)$$

and let  $x(n)$  be a signal with i.i.d. coordinates  $x(n)_i \sim p_X$ . By Proposition 7.1.3, we have  $\bar{d}(p_X) = 0$ . As above, the empirical distribution of the coordinates of the vectors  $x(n)$  converges to  $p_X$ . By applying Corollary 7.1.8 we obtain the following

**Corollary 7.2.2.** *Let  $\{x(n)\}_{n \geq 0}$  be a sequence of vectors with i.i.d. components  $x(n)_i \sim p_X$  where  $p_X$  is a discrete distribution as per Eq. (7.2.3). Then, almost surely as  $n \rightarrow \infty$ , Bayes optimal AMP recovers the signal  $x(n)$  from  $m(n) = o(n)$  spatially coupled measurements.*

It is important to further discuss the last statement because the reader might be misled into too optimistic a conclusion. Consider any signal  $x \in \mathbb{R}^n$ . For practical purposes, this will be represented with finite precision, say as a vector of  $\ell$ -bit numbers. Hence, in practice, the distribution  $p_X$  is always discrete, with  $K = 2^\ell$  a fixed number dictated by the precision requirements. A sublinear number of measurements  $m(n) = o(n)$  will then be sufficient to achieve this precision.

On the other hand, Theorem 7.1.6 and Corollary 7.1.8 are asymptotic statements, and the convergence rate is not claimed to be uniform in  $p_X$ . In particular, the values of  $n$  at which it becomes accurate will likely increase with  $K$ .

**Example 4 (A discrete-continuous mixture).** Consider the probability distribution

$$p_X = \varepsilon_+ \delta_{+1} + \varepsilon_- \delta_{-1} + \varepsilon q, \quad (7.2.4)$$

where  $\varepsilon_+ + \varepsilon_- + \varepsilon = 1$  and the probability measure  $q$  has a density with respect to Lebesgue measure. Again, let  $x(n)$  be a vector with i.i.d. components  $x(n)_i \sim p_X$ . We can apply Corollary 7.1.8 to conclude that  $m(n) = n\varepsilon + o(n)$  spatially coupled measurements are sufficient. This should be contrasted with the case of sensing matrices with i.i.d. entries studied in [47] under convex reconstruction methods (namely solving the feasibility problem  $y = Ax$  under the constraint  $\|x\|_\infty \leq 1$ ). In this case  $m(n) = n(1+\varepsilon)/2 + o(n)$  measurements are necessary.

In the next section we describe the basic intuition behind the surprising phenomenon in Theorems 7.1.6 and 7.1.7, and why spatially coupled sensing matrices are so useful. We conclude by stressing once more the limitations of these results:

- The Bayes optimal AMP algorithm requires knowledge of the signal distribution  $p_X$ . Notice however that only a good approximation of  $p_X$  (call it  $p_{\tilde{X}}$ , and denote by  $\tilde{X}$  the corresponding random variable) is sufficient. Assume indeed that  $p_X$  and  $p_{\tilde{X}}$  can be coupled in such a way that  $\mathbb{E}\{(X - \tilde{X})^2\} \leq \tilde{\sigma}^2$ . Then

$$x = \tilde{x} + u \tag{7.2.5}$$

where  $\|u\|_2^2 \lesssim n\tilde{\sigma}^2$ . This is roughly equivalent to adding to the noise vector  $z$  further ‘noise’  $\tilde{z}$  with variance  $\tilde{\sigma}^2/\delta$ . By this argument the guarantee in Theorem 7.1.7 degrades gracefully as  $p_{\tilde{X}}$  gets different from  $p_X$ . Another argument that leads to the same conclusion consists in studying the evolution of the algorithm (7.1.3), (7.1.4) when  $\eta_t$  is matched to the incorrect prior, see Appendix C.1.

Finally, it was demonstrated numerically in [137, 82] that, in some cases, a good ‘proxy’ for  $p_X$  can be learned through an Expectation-Maximization-style iteration. A rigorous study of this approach goes beyond the scope of the present work.

- In particular, the present approach does not provide uniform guarantees over the class of, say, sparse signals characterized by  $p_X(\{0\}) \geq 1 - \varepsilon$ . In particular, both the phase transition location, cf. Eq. (7.1.15), and the robustness constant, cf. Eq. (7.1.18), depend on the distribution  $p_X$ . This should be contrasted with the minimax approach of [42, 44, 33] which provides uniform guarantees over sparse signals. See Table 7.2.1 for a comparison between the two schemes.
- As mentioned above, the guarantees in Theorems 7.1.6 and 7.1.7 are only asymptotic.

	Minimax setup [44]	Bayesian setup (present approach)
Sensing matrix	i.i.d. Gaussian entries (homoscedastic)	Spatially-coupled matrices (heteroscedastic)
Reconstruction algorithm	AMP with soft thresholding denoiser	Bayes optimal AMP
Compression rate	$\delta > M(\varepsilon)$	$\delta \geq \bar{d}(p_X)$
Sensitivity to noise	$\frac{M(\varepsilon)}{1-M(\varepsilon)/\delta}$	$\frac{4\delta-2\bar{D}(p_X)}{\delta-\bar{D}(p_X)}$

Table 7.2.1 Comparison between the minimax setup in [44] and the Bayesian setup considered in this dissertation. (cf. [44, Eq. (2.4)] for definition of  $M(\varepsilon)$ ).

It would be important to develop analogous non-asymptotic results.

- The stability bound (7.1.18) is non-uniform, in that the proportionality constant  $C$  depends on the signal distribution. It would be important to establish analogous bounds that are uniform over suitable classes of distributions. (We do not expect Eq. (7.1.18) to hold uniformly over *all* distributions.)

### 7.3 Related work

The most closely related earlier work was already discussed above.

More broadly, message passing algorithms for compressed sensing where the object of a number of studies, starting with [4]. As mentioned, we will focus on approximate message passing (AMP) as introduced in [42, 43]. As shown in [33] these algorithms can be used in conjunction with a rich class of denoisers  $\eta(\cdot)$ . A subset of these denoisers arise as posterior mean associated to a prior  $p_X$ . Several interesting examples were studied by Schniter and collaborators [117, 118, 120], and by Rangan and collaborators [108, 78].

Spatial coupling has been the object of growing interest within coding theory over the last few years. The first instance of spatially coupled code ensembles were the convolutional LDPC codes of Felström and Zigangirov [53]. While the excellent performances of such codes had been known for quite some time [121], the fundamental reason was not elucidated until recently [85] (see also [90]). In particular [85] proved, for communication over the binary erasure channel (BEC), that the thresholds of spatially coupled ensembles under message passing decoding coincide with the thresholds of the base LDPC code under MAP decoding. In particular, this implies that spatially coupled ensembles achieve capacity over the BEC.

The analogous statement for general memoryless symmetric channels was first elucidated in [83] and finally proved in [86]. The paper [64] discusses similar ideas in a number of graphical models.

The first application of spatial coupling ideas to compressed sensing is due to Kudekar and Pfister [84]. They consider a class of sparse spatially coupled sensing matrices, very similar to parity check matrices for spatially coupled LDPC codes. On the other hand, their proposed message passing algorithms do not make use of the signal distribution  $p_X$ , and do not fully exploit the potential of spatially coupled matrices. The message passing algorithm used here belongs to the general class introduced in [42]. The specific use of the minimum-mean square error denoiser was suggested in [43]. The same choice is made in [82], which also considers Gaussian matrices with heteroscedastic entries although the variance structure is somewhat less general.

Finally, let us mention that robust sparse recovery of  $k$ -sparse vectors from  $m = O(k \log \log(n/k))$  measurement is possible, using suitable ‘adaptive’ sensing schemes [68].

## Chapter 8

# Matrix and algorithm construction

In this chapter, we define an ensemble of random matrices, and the corresponding choices of  $Q^t$ ,  $\mathbf{b}^t$ ,  $\eta_t$  that achieve the reconstruction guarantees in Theorems 7.1.6 and 7.1.7. We proceed by first introducing a general ensemble of random matrices. Correspondingly, we define a deterministic recursion named state evolution, that plays a crucial role in the algorithm analysis. In Section 8.3, we define the algorithm parameters and construct specific choices of  $Q^t$ ,  $\mathbf{b}^t$ ,  $\eta_t$ . The last section also contains a restatement of Theorems 7.1.6 and 7.1.7, in which this construction is made explicit.

### 8.1 General matrix ensemble

The sensing matrix  $A$  will be constructed randomly, from an ensemble denoted by  $\mathcal{M}(W, M, N)$ . The ensemble depends on two integers  $M, N \in \mathbb{N}$ , and on a matrix with non-negative entries  $W \in \mathbb{R}_+^{\mathbf{R} \times \mathbf{C}}$ , whose rows and columns are indexed by the finite sets  $\mathbf{R}$ ,  $\mathbf{C}$  (respectively ‘rows’ and ‘columns’). The band-diagonal structure that is characteristic of spatial coupling is imposed by a suitable choice of the matrix  $W$ . In this section we define the ensemble for a general choice of  $W$ . In Section 8.4 we discuss a class of choices for  $W$  that corresponds to spatial coupling, and that yields Theorems 7.1.6 and 7.1.7.

In a nutshell, the sensing matrix  $A$  is obtained from  $W$  through a suitable ‘lifting’ procedure. Each entry  $W_{r,c}$  is replaced by an  $M \times N$  block with i.i.d. entries  $A_{ij} \sim \mathbf{N}(0, W_{r,c}/M)$ . Rows and columns of  $A$  are then re-ordered uniformly at random to ensure exchangeability. For the reader familiar with the application of spatial coupling to coding theory, it might be useful to notice the differences and analogies with graph liftings. In that

case, the ‘lifted’ matrix is obtained by replacing each edge in the base graph with a random permutation matrix.

Passing to the formal definition, we will assume that the matrix  $W$  is *roughly row-stochastic*, i.e.,

$$\frac{1}{2} \leq \sum_{c \in \mathbf{C}} W_{r,c} \leq 2, \quad \text{for all } r \in \mathbf{R}. \quad (8.1.1)$$

(This is a convenient simplification for ensuring correct normalization of  $A$ .) We will let  $|\mathbf{R}| \equiv L_r$  and  $|\mathbf{C}| \equiv L_c$  denote the matrix dimensions. The ensemble parameters are related to the sensing matrix dimensions by  $n = NL_c$  and  $m = ML_r$ .

In order to describe a random matrix  $A \sim \mathcal{M}(W, M, N)$  from this ensemble, partition the columns and row indices in, respectively,  $L_c$  and  $L_r$  groups of equal size. Explicitly

$$\begin{aligned} [n] &= \cup_{s \in \mathbf{C}} C(s), & |C(s)| &= N, \\ [m] &= \cup_{r \in \mathbf{R}} R(r), & |R(r)| &= M. \end{aligned}$$

Here and below we use  $[k]$  to denote the set of first  $k$  integers  $[k] \equiv \{1, 2, \dots, k\}$ . Further, if  $i \in R(r)$  or  $j \in C(s)$  we will write, respectively,  $r = \mathbf{g}(i)$  or  $s = \mathbf{g}(j)$ . In other words  $\mathbf{g}(\cdot)$  is the operator determining the group index of a given row or column.

With this notation we have the following concise definition of the ensemble.

**Definition 8.1.1.** *A random sensing matrix  $A$  is distributed according to the ensemble  $\mathcal{M}(W, M, N)$  (and we write  $A \sim \mathcal{M}(W, M, N)$ ) if the partition of rows and columns ( $[m] = \cup_{r \in \mathbf{R}} R(r)$  and  $[n] = \cup_{s \in \mathbf{C}} C(s)$ ) are uniformly random, and given this partitioning, the entries  $\{A_{ij}, i \in [m], j \in [n]\}$  are independent Gaussian random variables with <sup>1</sup>*

$$A_{ij} \sim \mathbf{N}\left(0, \frac{1}{M} W_{\mathbf{g}(i), \mathbf{g}(j)}\right). \quad (8.1.2)$$

We refer to Fig. 8.1.1 for an illustration. Note that the randomness of the partitioning of row and column indices is only used in the proof of Lemma 10.0.1 (cf. [74]), and hence this and other illustrations assume that the partitions are contiguous.

Within the applications of spatial coupling to LDPC codes, see [83, 64, 86], the spatially-coupled codes are constructed by ‘coupling’ or ‘chaining’ a sequence of sparse graphs. The

---

<sup>1</sup>As in many papers on compressed sensing, the matrix here has independent zero-mean Gaussian entries; however, unlike standard practice, here the entries are of widely different variances.

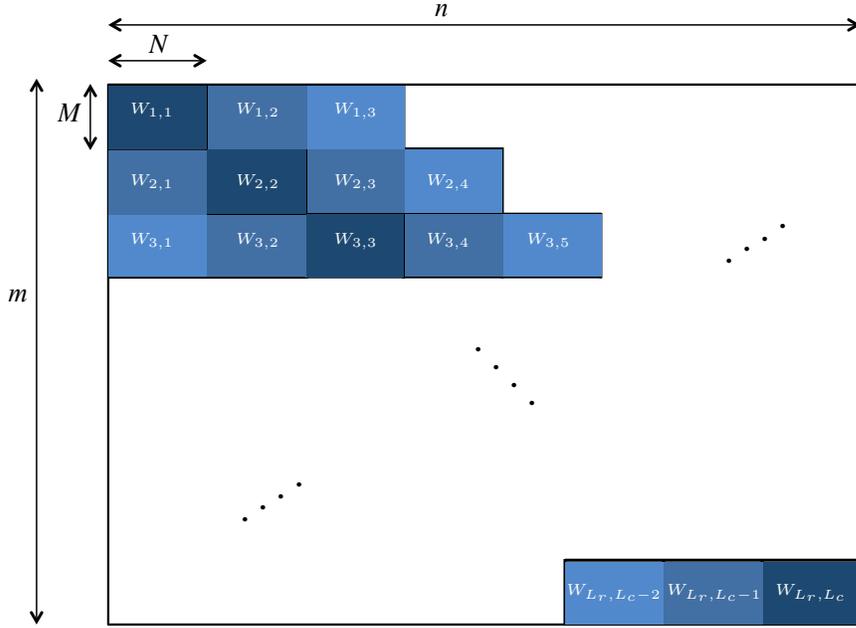


Figure 8.1.1: Construction of the spatially coupled measurement matrix  $A$  as described in Section 8.1. The matrix is divided into blocks with size  $M$  by  $N$ . (Number of blocks in each row and each column are respectively  $L_c$  and  $L_r$ , hence  $m = ML_r$ ,  $n = NL_c$ ). The matrix elements  $A_{ij}$  are chosen as  $N(0, \frac{1}{M}W_{g(i),g(j)})$ . In this figure,  $W_{i,j}$  depends only on  $|i - j|$  and thus blocks on each diagonal have the same variance.

indexes  $r \in R$ ,  $c \in C$  in the above construction correspond to the index of the graph along the chain in those constructions.

For proving Theorem 7.1.6 and Theorem 7.1.7 we will consider suitable sequences of ensembles  $\mathcal{M}(W, M, N)$  with undersampling ratio converging to  $\delta$ . While a complete description is given below, let us stress that we take the limit  $M, N \rightarrow \infty$  (with  $M = N\delta$ ) before the limit  $L_r, L_c \rightarrow \infty$ . Hence, the resulting matrix  $A$  is essentially dense: the fraction of non-zero entries per row vanishes only after the number of groups goes to  $\infty$ .

## 8.2 State evolution

State evolution allows an exact asymptotic analysis of AMP algorithms in the limit of a large number of dimensions. As indicated by the name, it bears close resemblance to the density evolution method in iterative coding theory [115]. Somewhat surprisingly, this analysis

approach is asymptotically exact despite the underlying factor graph being far from locally tree-like.

State evolution was first developed in [42] on the basis of heuristic arguments, and substantial numerical evidence. Subsequently, it was proved to hold for Gaussian sensing matrices with i.i.d. entries, and a broad class of iterative algorithms in [7]. These proofs were further generalized in [108], to cover ‘generalized’ AMP algorithms.

In the present case, state evolution takes the following form.<sup>2</sup>

**Definition 8.2.1.** *Given  $W \in \mathbb{R}_+^{\mathbb{R} \times \mathbb{C}}$  roughly row-stochastic, and  $\delta > 0$ , the corresponding state evolution maps  $\mathsf{T}'_W : \mathbb{R}_+^{\mathbb{R}} \rightarrow \mathbb{R}_+^{\mathbb{C}}$ ,  $\mathsf{T}''_W : \mathbb{R}_+^{\mathbb{C}} \rightarrow \mathbb{R}_+^{\mathbb{R}}$ , are defined as follows. For  $\phi = (\phi_a)_{a \in \mathbb{R}} \in \mathbb{R}_+^{\mathbb{R}}$ ,  $\psi = (\psi_i)_{i \in \mathbb{C}} \in \mathbb{R}_+^{\mathbb{C}}$ , we let:*

$$\mathsf{T}'_W(\phi)_i = \text{mmse}\left(\sum_{b \in \mathbb{R}} W_{b,i} \phi_b^{-1}\right), \quad (8.2.1)$$

$$\mathsf{T}''_W(\psi)_a = \sigma^2 + \frac{1}{\delta} \sum_{i \in \mathbb{C}} W_{a,i} \psi_i. \quad (8.2.2)$$

We finally define  $\mathsf{T}_W = \mathsf{T}'_W \circ \mathsf{T}''_W$ .

In the following, we shall omit the subscripts from  $\mathsf{T}_W$  whenever clear from the context.

**Definition 8.2.2.** *Given  $W \in \mathbb{R}_+^{L_r \times L_c}$  roughly row-stochastic, the corresponding state evolution sequence is the sequence of vectors  $\{\phi(t), \psi(t)\}_{t \geq 0}$ ,  $\phi(t) = (\phi_a(t))_{a \in \mathbb{R}} \in \mathbb{R}_+^{\mathbb{R}}$ ,  $\psi(t) = (\psi_i(t))_{i \in \mathbb{C}} \in \mathbb{R}_+^{\mathbb{C}}$ , defined recursively by  $\phi(t) = \mathsf{T}''_W(\psi(t))$ ,  $\psi(t+1) = \mathsf{T}'_W(\phi(t))$ , with initial condition*

$$\psi_i(0) = \infty \text{ for all } i \in \mathbb{C}. \quad (8.2.3)$$

Hence, for all  $t \geq 0$ ,

$$\begin{aligned} \phi_a(t) &= \sigma^2 + \frac{1}{\delta} \sum_{i \in \mathbb{C}} W_{a,i} \psi_i(t), \\ \psi_i(t+1) &= \text{mmse}\left(\sum_{b \in \mathbb{R}} W_{b,i} \phi_b(t)^{-1}\right). \end{aligned} \quad (8.2.4)$$

---

<sup>2</sup>In previous work, the state variable concerned a single scalar, representing the mean-squared error in the current reconstruction, averaged across all coordinates. In this work, the dimensionality of the state variable is much larger, because it contains  $\psi$ , an individualized MSE for each coordinate of the reconstruction and also  $\phi$ , a noise variance for the residuals  $r^t$  for each measurement coordinate.

The quantities  $\psi_i(t)$ ,  $\phi_a(t)$  correspond to the asymptotic MSE achieved by the AMP algorithm. More precisely,  $\psi_i(t)$  corresponds to the asymptotic mean square error  $\mathbb{E}\{(x_j^t - x_j)^2\}$  for  $j \in C(i)$ , as  $N \rightarrow \infty$ . Analogously,  $\phi_a(t)$  is the noise variance in residuals  $r_j^t$  corresponding to rows  $j \in R(a)$ . This correspondence is stated formally in Lemma 10.0.1 below. The state evolution (8.2.4) describes the evolution of these quantities. In particular, the linear operation in Eq. (7.1.4) corresponds to a sum of noise variances as per Eq. (8.2.2) and the application of denoisers  $\eta_t$  corresponds to a noise reduction as per Eq. (8.2.1).

As we will see, the definition of denoiser function  $\eta_t$  involves the state vector  $\phi(t)$ . (Notice that the state vectors  $\{\phi(t), \psi(t)\}_{t \geq 0}$  can be precomputed). Hence,  $\eta_t$  is ‘tuned’ according to the predicted reconstruction error at iteration  $t$ .

### 8.3 General algorithm definition

In order to fully define the AMP algorithm (7.1.3), (7.1.4), we need to provide constructions for the matrix  $Q^t$ , the nonlinearities  $\eta_t$ , and the vector  $\mathbf{b}^t$ . In doing this, we exploit the fact that the state evolution sequence  $\{\phi(t)\}_{t \geq 0}$  can be precomputed.

We define the matrix  $Q^t$  by

$$Q_{ij}^t \equiv \frac{\phi_{\mathbf{g}(i)}(t)^{-1}}{\sum_{k=1}^{L_r} W_{k,\mathbf{g}(j)} \phi_k(t)^{-1}}. \quad (8.3.1)$$

Notice that  $Q^t$  is block-constant: for any  $r, s \in [L]$ , the block  $Q_{R(r),C(s)}^t$  has all its entries equal.

As mentioned in Section ??, the function  $\eta_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is chosen to be separable, i.e., for  $v \in \mathbb{R}^N$ :

$$\eta_t(v) = (\eta_{t,1}(v_1), \eta_{t,2}(v_2), \dots, \eta_{t,N}(v_N)). \quad (8.3.2)$$

We take  $\eta_{t,i}$  to be a conditional expectation estimator for  $X \sim p_X$  in gaussian noise:

$$\eta_{t,i}(v_i) = \mathbb{E}\{X \mid X + s_{\mathbf{g}(i)}(t)^{-1/2} Z = v_i\}, \quad s_r(t) \equiv \sum_{u \in R} W_{u,r} \phi_u(t)^{-1}. \quad (8.3.3)$$

Notice that the function  $\eta_{t,i}(\cdot)$  depends on  $i$  only through the group index  $\mathbf{g}(i)$ , and in fact only parametrically through  $s_{\mathbf{g}(i)}(t)$ . It is also interesting to notice that the denoiser

$\eta_{t,i}(\cdot)$  does not have any tuning parameter to be optimized over. This was instead the case for the soft-thresholding AMP algorithm studied in [42] for which the threshold level had to be adjusted in a non-trivial manner to the sparsity level. This difference is due to the fact that the prior  $p_X$  is assumed to be known and hence the optimal denoiser is uniquely determined to be the posterior expectation as per Eq. (8.3.3).

Finally, in order to define the vector  $\mathbf{b}_i^t$ , let us introduce the quantity

$$\langle \eta'_t \rangle_u = \frac{1}{N} \sum_{i \in C(u)} \eta'_{t,i}(x_i^t + ((Q^t \odot A)^* r^t)_i). \quad (8.3.4)$$

Recalling that  $Q^t$  is block-constant, we define matrix  $\tilde{Q}^t \in \mathbb{R}^{L_r \times L_c}$  as  $\tilde{Q}_{r,u}^t = Q_{i,j}^t$ , with  $i \in R(r)$  and  $j \in C(u)$ . In words,  $\tilde{Q}^t$  contains one representative of each block. The vector  $\mathbf{b}^t$  is then defined by

$$\mathbf{b}_i^t \equiv \frac{1}{\delta} \sum_{u \in C} W_{\mathbf{g}(i),u} \tilde{Q}_{\mathbf{g}(i),u}^{t-1} \langle \eta'_{t-1} \rangle_u. \quad (8.3.5)$$

Again  $\mathbf{b}_i^t$  is block-constant: the vector  $\mathbf{b}_{C(u)}^t$  has all its entries equal.

This completes our definition of the AMP algorithm. Let us conclude with a few computational remarks:

1. The quantities  $\tilde{Q}^t$ ,  $\phi(t)$  can be precomputed efficiently iteration by iteration, because they are, respectively,  $L_r \times L_c$  and  $L_r$ -dimensional, and, as discussed further below,  $L_r, L_c$  are much smaller than  $m, n$ . The most complex part of this computation is implementing the iteration (8.2.4), which has complexity  $O((L_r + L_c)^3)$ , plus the complexity of evaluating the `mmse` function, which is a one-dimensional integral.
2. The vector  $\mathbf{b}^t$  is also block-constant, so can be efficiently computed using Eq. (8.3.5).
3. Instead of computing  $\phi(t)$  analytically by iteration (8.2.4),  $\phi(t)$  can also be estimated from data  $x^t, r^t$ . In particular, by generalizing the methods introduced in [42, 99], we get the estimator

$$\hat{\phi}_a(t) = \frac{1}{M} \|r_{R(a)}^t\|_2^2, \quad (8.3.6)$$

where  $r_{R(a)}^t = (r_j^t)_{j \in R(a)}$  is the restriction of  $r^t$  to the indices in  $R(a)$ . An alternative

more robust estimator (more resilient to outliers), would be

$$\widehat{\phi}_a(t)^{1/2} = \frac{1}{\Phi^{-1}(3/4)} |r_{R(a)}^t|_{(M/2)}, \quad (8.3.7)$$

where  $\Phi(z)$  is the Gaussian distribution function, and, for  $v \in \mathbb{R}^K$ ,  $|v|_{(\ell)}$  is the  $\ell$ -th largest entry in the vector  $(|v_1|, |v_2|, \dots, |v_K|)$ . (See, e.g., [66] for background in robust estimation.) The idea underlying both of the above estimators is that the components of  $r_{R(a)}^t$  are asymptotically i.i.d. with mean zero and variance  $\phi_a(t)$ .

## 8.4 Choices of parameters, and spatial coupling

In order to prove our main Theorem 7.1.6, we use a sensing matrix from the ensemble  $\mathcal{M}(W, M, N)$  for a suitable choice of the matrix  $W \in \mathbb{R}^{\mathbb{R} \times \mathbb{C}}$ . Our construction depends on parameters  $\rho \in \mathbb{R}_+$ ,  $L, L_0 \in \mathbb{N}$ , and on the ‘shape function’  $\mathcal{W}$ . As explained below,  $\rho$  will be taken to be small, and hence we will treat  $1/\rho$  as an integer to avoid rounding (which introduces in any case a negligible error).

Here and below  $\cong$  denotes identity between two sets up to a relabeling.

**Definition 8.4.1.** *A shape function is a function  $\mathcal{W} : \mathbb{R} \rightarrow \mathbb{R}_+$  continuously differentiable, with support in  $[-1, 1]$  and such that  $\int_{\mathbb{R}} \mathcal{W}(u) du = 1$ , and  $\mathcal{W}(-u) = \mathcal{W}(u)$ .*

We let  $\mathbb{C} \cong \{-2\rho^{-1}, \dots, 0, 1, \dots, L-1\}$ , so that  $L_{\mathbb{C}} = L + 2\rho^{-1}$ . Also let  $\mathbb{C}_0 = \{0, 1, \dots, L-1\}$ .

The rows are partitioned as follows:

$$\mathbb{R} = \mathbb{R}_0 \cup \left\{ \bigcup_{i=-2\rho^{-1}}^{-1} \mathbb{R}_i \right\},$$

where  $\mathbb{R}_0 \cong \{-\rho^{-1}, \dots, 0, 1, \dots, L-1 + \rho^{-1}\}$ , and  $\mathbb{R}_i = \{iL_0, \dots, (i+1)L_0 - 1\}$ , for  $i = -2\rho^{-1}, \dots, -1$ . Hence,  $|\mathbb{R}_i| = L_0$ , and  $L_{\mathbb{R}} = L_{\mathbb{C}} + 2\rho^{-1}L_0$ .

Finally, we take  $N$  so that  $n = NL_{\mathbb{C}}$ , and let  $M = N\delta$  so that  $m = ML_{\mathbb{R}} = N(L_{\mathbb{C}} + 2\rho^{-1}L_0)\delta$ . Notice that  $m/n = \delta(L_{\mathbb{C}} + 2\rho^{-1}L_0)/L_{\mathbb{C}}$ . Since we will take  $L_{\mathbb{C}}$  much larger than  $L_0/\rho$ , we in fact have  $m/n$  arbitrarily close to  $\delta$ .

Given these inputs, we construct the corresponding matrix  $W = W(L, L_0, \mathcal{W}, \rho)$  as follows.

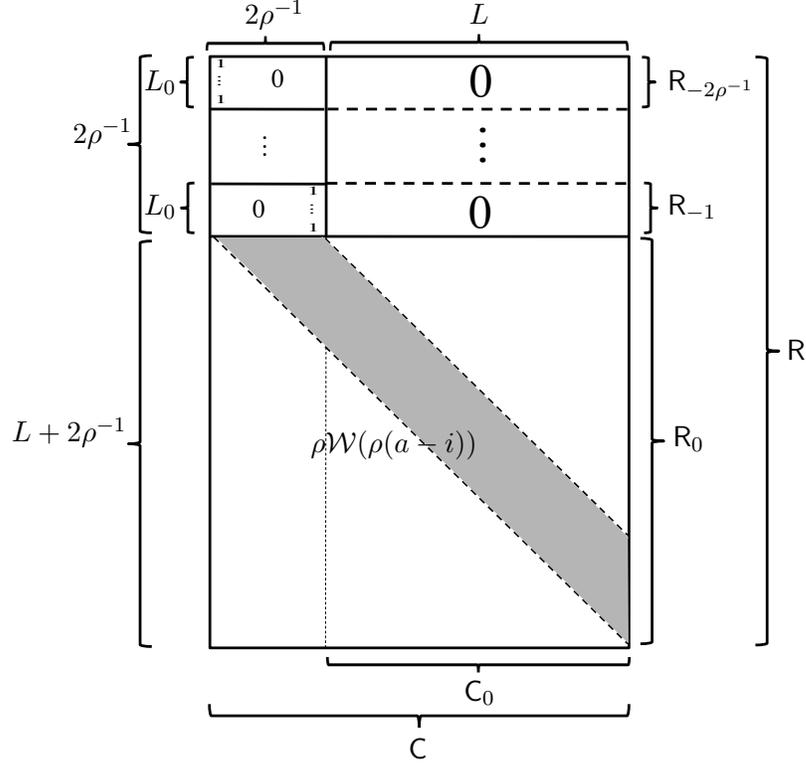


Figure 8.4.2: Matrix  $W$ . The shaded region indicates the non zero entries in the lower part of the matrix. As shown (the lower part of ) the matrix  $W$  is band diagonal.

1. For  $i \in \{-2\rho^{-1}, \dots, -1\}$ , and each  $a \in R_i$ , we let  $W_{a,i} = 1$ . Further,  $W_{a,j} = 0$  for all  $j \in C \setminus \{i\}$ .
2. For all  $a \in R_0 \cong \{-\rho^{-1}, \dots, 0, \dots, L - 1 + \rho^{-1}\}$ , we let

$$W_{a,i} = \rho\mathcal{W}(\rho(a-i)) \quad i \in \{-2\rho^{-1}, \dots, L-1\}. \quad (8.4.1)$$

The role of the rows in  $\left\{\bigcup_{i=-2\rho^{-1}}^{-1} R_i\right\}$  and the corresponding rows in  $A$  are to oversample the first few (namely the first  $2\rho^{-1}N$ ) coordinates of the signal as explained in Section 9.1. Furthermore, the restriction of  $W$  to the rows in  $R_0$  is band diagonal as  $\mathcal{W}$  is supported on  $[-1, 1]$ . See Fig. 8.4.2 for an illustration of the matrix  $W$ .

In the following we occasionally use the shorthand  $W_{a-i} \equiv \rho\mathcal{W}(\rho(a-i))$ . Note that  $W$

is roughly row-stochastic. Also, the restriction of  $W$  to the rows in  $\mathbf{R}_0$  is roughly column-stochastic. This follows from the fact that the function  $\mathcal{W}(\cdot)$  has continuous (and thus bounded) derivative on the compact interval  $[-1, 1]$ , and  $\int_{\mathbb{R}} \mathcal{W}(u) du = 1$ . Therefore, using the standard convergence of Riemann sums to Riemann integrals and the fact that  $\rho$  is small, we get the result.

We are now in position to restate Theorem 7.1.6 in a more explicit form.

**Theorem 8.4.2.** *Let  $p_X$  be a probability measure on the real line with  $\delta > \bar{d}(p_X)$ , and let  $\mathcal{W} : \mathbb{R} \rightarrow \mathbb{R}_+$  be a shape function. For any  $\varepsilon > 0$ , there exist  $L_0, L, \rho, t_0, \sigma_0^2 = \sigma_0(\varepsilon, \delta, p_X)^2$  such that  $L_0/(L\rho) \leq \varepsilon$ , and further the following holds true for  $W = W(L, L_0, \mathcal{W}, \rho)$ .*

*For  $N \geq 0$ , and  $A(n) \sim \mathcal{M}(W, M, N)$  with  $M = N\delta$ , and for all  $\sigma^2 \leq \sigma_0^2, t \geq t_0$ , we almost surely have*

$$\limsup_{N \rightarrow \infty} \frac{1}{n} \|x^t(A(n); y(n)) - x(n)\|^2 \leq \varepsilon. \quad (8.4.2)$$

*Further, under the same assumptions, we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{n} \mathbb{E}\{\|x^t(A(n); y(n)) - x(n)\|^2\} \leq \varepsilon. \quad (8.4.3)$$

In order to obtain a stronger form of robustness, as per Theorem 7.1.7, we slightly modify the sensing scheme. We construct the sensing matrix  $\tilde{A}$  from  $A$  by appending  $2\rho^{-1}L_0$  rows in the bottom.

$$\tilde{A} = \left( \begin{array}{c|c} A & \\ \hline 0 & I \end{array} \right), \quad (8.4.4)$$

where  $I$  is the identity matrix of dimensions  $2\rho^{-1}L_0$ . Note that this corresponds to increasing the number of measurements; however, the asymptotic undersampling rate remains  $\delta$ , provided that  $L_0/(L\rho) \rightarrow 0$ , as  $n \rightarrow \infty$ .

The reconstruction scheme is modified as follows. Let  $x_1$  be the vector obtained by restricting  $x$  to entries in  $\cup_i C(i)$ , where  $i \in \{-2\rho^{-1}, \dots, L - 2\rho^{-1} - 1\}$ . Also, let  $x_2$  be the vector obtained by restricting  $x$  to entries in  $\cup_i C(i)$ , where  $i \in \{L - 2\rho^{-1}, \dots, L - 1\}$ . Therefore,  $x = (x_1, x_2)^T$ . Analogously, let  $y = (y_1, y_2)^T$  where  $y_1$  is given by the restriction of  $y$  to  $\cup_{i \in \mathbf{R}} R(i)$  and  $y_2$  corresponds to the additional  $2\rho^{-1}L_0$  rows. Define  $w_1$  and  $w_2$  from

the noise vector  $w$ , analogously. Hence,

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \left( \begin{array}{c|c} A & \\ \hline 0 & I \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \quad (8.4.5)$$

Note that the sampling rate for vector  $x_2$  is one, i.e.,  $y_2$  and  $x_2$  are of the same length and are related to each other through the identity matrix  $I$ . Hence, we have a fairly good approximation of these entries. We use the AMP algorithm as described in the previous section to obtain an estimation of  $x_1$ . Formally, let  $x^t$  be the estimation at iteration  $t$  obtained by applying the AMP algorithm to the problem  $y_1 = Ax + w_1$ . The modified estimation is then  $\tilde{x}^t = (x_1^t, y_2)^T$ .

As we will see later, this modification in the sensing matrix and algorithm, while not necessary, simplifies some technical steps in the proof.

**Theorem 8.4.3.** *Let  $p_X$  be a probability measure on the real line with  $\delta > \overline{D}(p_X)$ , and let  $\mathcal{W} : \mathbb{R} \rightarrow \mathbb{R}_+$  be a shape function. There exist  $L_0, L, \rho, t_0$  and a finite stability constant  $C = C(p_X, \delta)$ , such that  $L_0/(L\rho) < \varepsilon$ , for any given  $\varepsilon > 0$ , and the following holds true for the modified reconstruction scheme.*

*For  $t \geq t_0$ , we almost surely have,*

$$\limsup_{N \rightarrow \infty} \frac{1}{n} \|\tilde{x}^t(\tilde{A}(n); y(n)) - x(n)\|^2 \leq C\sigma^2. \quad (8.4.6)$$

*Further, under the same assumptions, we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{n} \mathbb{E} \{ \|\tilde{x}^t(\tilde{A}(n); y(n)) - x(n)\|^2 \} \leq C\sigma^2. \quad (8.4.7)$$

*Finally, in the asymptotic case where  $\ell = L\rho \rightarrow \infty$ ,  $\rho \rightarrow 0$ ,  $L_0 \rightarrow \infty$ , we have*

$$\limsup_{\sigma \rightarrow 0} \frac{1}{\sigma^2} \left\{ \lim_{t \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{n} \|\tilde{x}^t(\tilde{A}(n); y(n)) - x(n)\|^2 \right\} \leq \frac{4\delta - 2\overline{D}(p_X)}{\delta - \overline{D}(p_X)}.$$

It is obvious that Theorems 8.4.2 and 8.4.3 respectively imply Theorems 7.1.6 and 7.1.7. We shall therefore focus on the proofs of Theorems 8.4.2 and 8.4.3 in the rest of this part.

It is worth noting that as per Theorem 8.4.3, the sensitivity constant at small noise depends on the signal distribution  $p_X$  only through its upper MMSE dimension  $\overline{D}(p_X)$ . In

particular, for signal distribution of the form (7.2.2), the robustness guarantee is independent of the  $q$  component.

Notice that the results of Theorems 8.4.2 and 8.4.3 only deal with a linear subsequence  $n = NL_c$  with  $N \rightarrow \infty$ . However, this is sufficient to prove the claim of Theorems 7.1.6 and 7.1.7. More specifically, suppose that  $n$  is not a multiple of  $L_c$ . Let  $n'$  be the smallest number greater than  $n$  which is divisible by  $L_c$ , i.e.,  $n' = \lceil n/L_c \rceil L_c$ , and let  $\hat{x} = (x, 0)^T \in \mathbb{R}^{n'}$  be obtained by padding  $x$  with zeros. Let  $\hat{x}^t$  denote the Bayes optimal AMP estimate of  $\hat{x}$  and  $x^t$  be the restriction of  $\hat{x}^t$  to the first  $n$  entries. We have  $(1/n)\|x^t - x\|^2 \leq (n'/n)(1/n')\|\hat{x}^t - \hat{x}\|^2$ . The result of Theorem 7.1.6 follows by applying Theorem 8.4.2 (for the sequence  $n = NL_c$ ,  $N \rightarrow \infty$ ), and noting that  $n'/n \leq (1 + L_c/n) \rightarrow 1$ , as  $N \rightarrow \infty$ . Similar comment applies to Theorems 8.4.3 and 7.1.7.

## Chapter 9

# Magic of Spatial Coupling

### 9.1 How does spatial coupling work?

Spatial coupling was developed in coding theory to construct capacity achieving LDPC codes [53, 121, 83, 64, 86]. The standard construction starts from the parity check matrix of an LDPC code that is sparse but unstructured apart from the degree sequence. A spatially coupled ensemble is then obtained by enforcing a band-diagonal structure, while keeping the degree sequence unchanged. Usually this is done by graph liftings, but the underlying principle is more general [64].

Following the above intuition, spatially coupled sensing matrices  $A$  are, roughly speaking, random band-diagonal matrices. The construction given below (as the one of [82]) uses matrices with independent zero-mean Gaussian entries, with non-identical variances (heteroscedastic entries). However, the simulations of [70] suggest that a much broader set of matrices display similar performances. As discussed in Section 8.1, the construction is analogous to graph liftings. We start by a matrix of variances  $W = (W_{r,c})$  and obtain the sensing matrix  $A$  by replacing each entry  $W_{r,c}$  by a block with i.i.d. Gaussian entries with variance proportional to  $W_{r,c}$ .

It is convenient to think of the graph structure that they induce on the reconstruction problem. Associate one node (a *variable node* in the language of factor graphs) to each coordinate  $i$  in the unknown signal  $x$ . Order these nodes on the real line  $\mathbb{R}$ , putting the  $i$ -th node at location  $i \in \mathbb{R}$ . Analogously, associate a node (a *factor node*) to each coordinate  $a$  in the measurement vector  $y$ , and place the node  $a$  at position  $a/\delta$  on the same line. Connect this node to all the variable nodes  $i$  such that  $A_{ai} \neq 0$ . If  $A$  is band diagonal, only

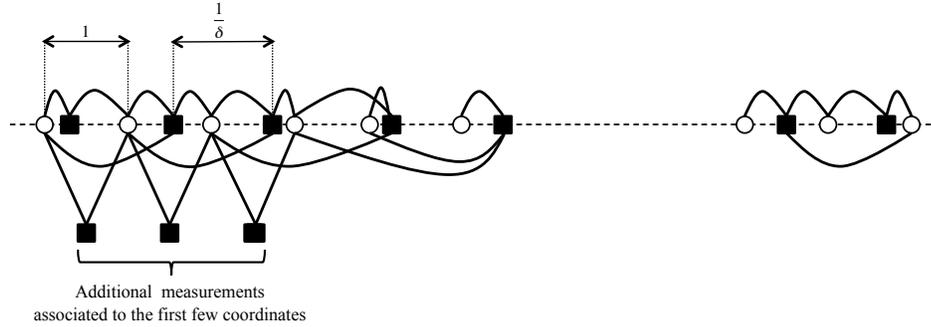


Figure 9.1.1: Graph structure of a spatially coupled matrix. Variable nodes are shown as circle and check nodes are represented by square.

nodes that are placed close enough will be connected by an edge. See Figure 9.1.1 for an illustration.

In a spatially coupled matrix, additional measurements are associated to the first few coordinates of  $x$ , say coordinates  $x_1, \dots, x_{n_0}$  with  $n_0$  much smaller than  $n$ . This has a negligible impact on the overall undersampling ratio as  $n/n_0 \rightarrow \infty$ . Although the overall undersampling remains  $\delta < 1$ , the coordinates  $x_1, \dots, x_{n_0}$  are oversampled. This ensures that these first coordinates are recovered correctly (up to a mean square error of order  $\sigma^2$ ). As the algorithm is iterated, the contribution of these first few coordinates is correctly subtracted from all the measurements, and hence we can effectively eliminate those nodes from the graph. In the resulting graph, the first few variables are effectively oversampled and hence the algorithm will reconstruct their values, up to a mean square error of order  $\sigma^2$ . As the process is iterated, variables are progressively reconstructed, proceeding from left to right along the node layout.

While the above explains the basic dynamics of AMP reconstruction algorithms under spatial coupling, a careful consideration reveals that this picture leaves open several challenging questions. In particular, why does the overall undersampling factor  $\delta$  have to exceed  $\bar{d}(p_X)$  for reconstruction to be successful? Our proof is based on a potential function argument. We will prove that there exists a potential function for the AMP algorithm, such that, when  $\delta > \bar{d}(p_X)$ , this function has its global minimum close to exact reconstruction. Further, we will prove that, unless this minimum is essentially achieved, AMP can always decrease the function. This technique is different from the one followed in [85] for the LDPC codes over the binary erasure channel, and we think it is of independent interest.

## 9.2 Advantages of spatial coupling

Within the construction proposed in this dissertation, spatially coupled sensing matrices have independent heteroscedastic entries (entries with different variances). In addition to this, we also oversample a few number of coordinates of the signal, namely the first  $2\rho^{-1}N$  coordinates. In this section we informally discuss the various components of this scheme.

It can be instructive to compare this construction with the case of homoscedastic Gaussian matrices (i.i.d. entries). For the reader familiar with coding theory, this comparison is analogous to the comparison between regular LDPC codes and spatially coupled regular LDPC codes. Regular LDPC codes have been known since Gallager [57, 94] to achieve the channel capacity, as the degree gets large, under maximum likelihood decoding. However their performances under practical (belief propagation) decoding is rather poor. When the code ensemble is modified via spatial coupling, the belief propagation performances improve to become asymptotically equivalent to the maximum likelihood performances. Hence spatially coupled LDPC codes achieve capacity under practical decoding schemes.

Similarly, standard (non-spatially coupled) sensing matrices achieve the information theoretic limit under computationally unpractical recovery schemes [143], but do not perform ideally under practical reconstruction algorithms. Consider for instance Bayes optimal AMP. Within the standard ensemble, the state evolution recursion reads

$$\begin{aligned}\phi(t) &= \sigma^2 + \frac{1}{\delta}\psi(t), \\ \psi(t+1) &= \text{mmse}(\phi(t)^{-1}).\end{aligned}\tag{9.2.1}$$

Note that  $\psi(t+1)$  is the minimum mean square error at signal-to-noise ratio  $\phi(t)^{-1}$ , i.e., treating the residual part as noise. Let  $\tilde{\delta}(p_X) \equiv \sup_{s \geq 0} s \cdot \text{mmse}(s) > \bar{d}(p_X)$ . It is immediate to see that the last recursion develops two (or possibly more) stable fixed points for  $\delta < \tilde{\delta}(p_X)$  and all  $\sigma^2$  small enough. The smallest fixed point, call it  $\phi_{\text{good}}$ , corresponds to correct reconstruction and is such that  $\phi_{\text{good}} = O(\sigma^2)$  as  $\sigma \rightarrow 0$ . The largest fixed point, call it  $\phi_{\text{bad}}$ , corresponds to incorrect reconstruction and is such that  $\phi_{\text{bad}} = \Theta(1)$  as  $\sigma \rightarrow 0$ . A study of the above recursion shows that  $\lim_{t \rightarrow \infty} \phi(t) = \phi_{\text{bad}}$ . State evolution converges to the ‘incorrect’ fixed point, hence predicting a large MSE for AMP.

On the contrary, for  $\bar{d}(p_X) < \delta < \tilde{\delta}(p_X)$  the recursion (9.2.1) converges (for appropriate choices of  $W$  as in the previous section) to the ‘ideal’ fixed point  $\lim_{t \rightarrow \infty} \phi_a(t) = \phi_{\text{good}}$  for all  $a$  (except possibly those near the boundaries). This is illustrated in Fig. 10.1.1. We also

refer to [64] for a survey of examples of the same phenomenon and to [82, 70] for further discussion in compressed sensing.

The above discussion also clarifies why the posterior expectation denoiser is useful. Spatially coupled sensing matrices do not yield better performances than the ones dictated by the best fixed point in the ‘standard’ recursion (9.2.1). In particular, replacing the Bayes optimal denoiser by another denoiser  $\eta_t$  amounts, roughly, to replacing  $\text{mmse}$  in Eq. (9.2.1) by the MSE of another denoiser, hence leading to worse performances.

In particular, if the posterior expectation denoiser is replaced by soft thresholding, the resulting state evolution recursion always has a unique stable fixed point for homoscedastic matrices [42]. This suggests that spatial coupling does not lead to any improvement for soft thresholding AMP and hence (via the correspondence of [8]) for Lasso or  $\ell_1$  reconstruction. This expectation is indeed confirmed numerically in [70].

## Chapter 10

# Key Lemmas and Proof of the Main Theorems

Our proof is based in a crucial way on state evolution. This effectively reduces the analysis of the algorithm (7.1.3), (7.1.4) to the analysis of the deterministic recursion (8.2.4).

**Lemma 10.0.1.** *Let  $W \in \mathbb{R}_+^{\mathbb{R} \times \mathbb{C}}$  be a roughly row-stochastic matrix (see Eq. (8.1.1)) and  $\phi(t)$ ,  $Q^t$ ,  $\mathbf{b}^t$  be defined as in Section 8.3. Let  $M = M(N)$  be such that  $M/N \rightarrow \delta$ , as  $N \rightarrow \infty$ . Define  $m = ML_r$ ,  $n = NL_c$ , and for each  $N \geq 1$ , let  $A(n) \sim \mathcal{M}(W, M, N)$ . Let  $\{(x(n), w(n))\}_{n \geq 0}$  be a converging sequence of instances with parameters  $(p_X, \sigma^2)$ . Then, for all  $t \geq 1$ , almost surely we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \|x_{C(i)}^t(A(n); y(n)) - x_{C(i)}\|_2^2 = \text{mmse} \left( \sum_{a \in \mathbb{R}} W_{a,i} \phi_a(t-1)^{-1} \right). \quad (10.0.1)$$

for all  $i \in \mathbb{C}$ .

This lemma is a straightforward generalization of [7]. We leave out the formal proof of this lemma since it does not require new ideas, but a significant amount of new notations. We instead refer to our publication [74] which covers an even more general setting. In the interest of self-containedness, and to develop useful intuition on state evolution, we present an heuristic derivation of the state evolution equations (8.2.4) in Section 10.2.

The next Lemma provides the needed analysis of the recursion (8.2.4).

**Lemma 10.0.2.** *Let  $\delta > 0$ , and  $p_X$  be a probability measure on the real line. Let  $\mathcal{W} : \mathbb{R} \rightarrow \mathbb{R}_+$  be a shape function.*

(a) If  $\delta > \bar{d}(p_X)$ , then for any  $\varepsilon > 0$ , there exist  $\sigma_0 = \sigma_0(\varepsilon, \delta, p_X), \rho, L_* > 0$ , such that for any  $\sigma^2 \in [0, \sigma_0^2], L_0 > 3/\delta$ , and  $L > L_*$ , the following holds for  $W = W(L, L_0, \mathcal{W}, \rho)$ :

$$\lim_{t \rightarrow \infty} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L+\rho^{-1}-1} \phi_a(t) \leq \varepsilon. \quad (10.0.2)$$

(b) If further  $\delta > \bar{D}(p_X)$ , then there exist  $\rho, L_* > 0$ , and a finite stability constant  $C = C(p_X, \delta)$ , such that for  $L_0 > 3/\delta$ , and  $L > L_*$ , the following holds for  $W = W(L, L_0, \mathcal{W}, \rho)$ .

$$\lim_{t \rightarrow \infty} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) \leq C\sigma^2. \quad (10.0.3)$$

Finally, in the asymptotic case where  $\ell = L\rho \rightarrow \infty, \rho \rightarrow 0, L_0 \rightarrow \infty$ , we have

$$\limsup_{\sigma \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{\sigma^2 L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) \leq \frac{3\delta - \bar{D}(p_X)}{\delta - \bar{D}(p_X)}. \quad (10.0.4)$$

The proof of this lemma is deferred to Section 11 and is indeed the technical core of our result.

Now, we have in place all we need to prove our main results.

*Proof (Theorem 8.4.2).* Recall that  $\mathbb{C} \cong \{-2\rho^{-1} \dots, L-1\}$ . Therefore,

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{n} \|x^t(A(n); y(n)) - x(n)\|^2 &\leq \frac{1}{L_c} \sum_{i \in \mathbb{C}} \limsup_{N \rightarrow \infty} \frac{1}{N} \|x_{C(i)}^t(A(n); y(n)) - x_{C(i)}(n)\|^2 \\ &\stackrel{(a)}{\leq} \frac{1}{L_c} \sum_{i=-2\rho^{-1}}^{L-1} \text{mmse} \left( \sum_{a \in \mathbb{R}} W_{a,i} \phi_a(t-1)^{-1} \right) \\ &\stackrel{(b)}{\leq} \frac{1}{L_c} \sum_{i=-2\rho^{-1}}^{L-1} \text{mmse} \left( \sum_{a \in \mathbb{R}_0} W_{a,i} \phi_a(t-1)^{-1} \right) \\ &\stackrel{(c)}{\leq} \frac{1}{L_c} \sum_{i=-2\rho^{-1}}^{L-1} \text{mmse} \left( \frac{1}{2} \phi_{i+\rho^{-1}}(t-1)^{-1} \right) \\ &\stackrel{(d)}{\leq} \frac{1}{L_c} \sum_{a=-\rho^{-1}}^{L+\rho^{-1}-1} 2\phi_a(t-1). \end{aligned} \quad (10.0.5)$$

Here, (a) follows from Lemma 10.0.1; (b) follows from the fact that mmse is non-increasing; (c) holds because of the following facts: (i)  $\phi_a(t)$  is nondecreasing in  $a$  for every  $t$  (see Lemma 11.3.4 below). (ii) Restriction of  $W$  to the rows in  $\mathbf{R}_0$  is roughly column-stochastic. (iii) mmse is non-increasing; (d) follows from the inequality  $\text{mmse}(s) \leq 1/s$ . The result is immediate due to Lemma 10.0.2, Part (a).

Now, we prove the claim regarding the expected error. Let  $f_n = \frac{1}{n} \|x^t(A(n); y(n)) - x(n)\|^2$ . Since  $\limsup_{n \rightarrow \infty} f_n \leq \varepsilon$ , there exists  $n_0$  such that  $f_n \leq 2\varepsilon$  for  $n \geq n_0$ . Applying reverse Fatou's lemma to the bounded sequence  $\{f_n\}_{n \geq n_0}$ , we have  $\limsup_{N \rightarrow \infty} \mathbb{E} f_n \leq \mathbb{E}[\limsup_{N \rightarrow \infty} f_n] \leq \varepsilon$ .  $\square$

*Proof (Theorem 8.4.3).* The proof proceeds in a similar manner to the proof of Theorem 8.4.2.

$$\begin{aligned}
& \limsup_{N \rightarrow \infty} \frac{1}{n} \|\tilde{x}^t(\tilde{A}(n); y(n)) - x(n)\|^2 \\
& \leq \frac{1}{L_c} \left\{ \sum_{i=-2\rho^{-1}}^{L-2\rho^{-1}-1} \limsup_{N \rightarrow \infty} \frac{1}{N} \|x_{C(i)}^t(A(n); y(n)) - x_{C(i)}(n)\|^2 + \lim_{N \rightarrow \infty} \frac{1}{N} \|w_2(n)\|^2 \right\} \\
& \leq \frac{1}{L_c} \left\{ \sum_{i=-2\rho^{-1}}^{L-2\rho^{-1}-1} \text{mmse} \left( \sum_{a \in \mathbf{R}} W_{a,i} \phi_a(t-1)^{-1} \right) + \lim_{N \rightarrow \infty} \frac{1}{N} \|w_2(n)\|^2 \right\} \\
& \leq \frac{1}{L_c} \left\{ \sum_{i=-2\rho^{-1}}^{L-2\rho^{-1}-1} \text{mmse} \left( \sum_{a \in \mathbf{R}_0} W_{a,i} \phi_a(t-1)^{-1} \right) + \lim_{N \rightarrow \infty} \frac{1}{N} \|w_2(n)\|^2 \right\} \tag{10.0.6} \\
& \leq \frac{1}{L_c} \left\{ \sum_{i=-2\rho^{-1}}^{L-2\rho^{-1}-1} \text{mmse} \left( \frac{1}{2} \phi_{i+\rho^{-1}}(t-1)^{-1} \right) + \lim_{N \rightarrow \infty} \frac{1}{N} \|w_2(n)\|^2 \right\} \\
& \leq \frac{1}{L_c} \left\{ \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} 2\phi_a(t-1) + \lim_{N \rightarrow \infty} \frac{1}{N} \|w_2(n)\|^2 \right\} \leq C \sigma^2,
\end{aligned}$$

where the last step follows from Part (b) in Lemma 10.0.2, and Part (b) in Definition 7.1.1.

The claim regarding the expected error follows by a similar argument to the one in the proof of Theorem 8.4.2.

Finally, in the asymptotic case, where  $\ell = L\rho \rightarrow \infty$ ,  $L_0 \rightarrow \infty$ ,  $\rho \rightarrow 0$ , we have  $\sum_{a \in \mathbf{R}_0} W_{a,i} = \sum_{a \in \mathbf{R}_0} \rho \mathcal{W}(\rho(a-i)) \rightarrow \int \mathcal{W}(u) du = 1$ , and the bound in Eq. (10.0.6) can be strengthened by replacing  $2\phi_a(t-1)$  with  $\phi_a(t-1)$ . Using Eq. (10.0.4) in Eq. (10.0.6), we obtain the desired result.  $\square$

## 10.1 Numerical experiments

We consider a Bernoulli-Gaussian distribution  $p_X = (1-\varepsilon)\delta_0 + \varepsilon\gamma_{0,1}$ . Recall that  $\gamma_{\mu,\sigma}(dx) = (2\pi\sigma^2)^{-1/2} \exp\{-(x-\mu)^2/(2\sigma^2)\}dx$ . We construct a random signal  $x(n) \in \mathbb{R}^n$  by sampling i.i.d. coordinates  $x(n)_i \sim p_X$ . We have  $\bar{d}(p_X) = \varepsilon$  by Proposition 7.1.3 and

$$\eta_{t,i}(v_i) = \frac{\varepsilon\gamma_{1+s_{\mathbf{g}(i)}^{-1}}(v_i)}{\varepsilon\gamma_{1+s_{\mathbf{g}(i)}^{-1}}(v_i) + (1-\varepsilon)\gamma_{s_{\mathbf{g}(i)}^{-1}}(v_i)} \cdot \frac{1}{1+s_{\mathbf{g}(i)}^{-1}}v_i. \quad (10.1.1)$$

In the experiments, we use  $\varepsilon = 0.1$ ,  $\sigma = 0.01$ ,  $\rho = 0.1$ ,  $M = 6$ ,  $N = 50$ ,  $L = 500$ ,  $L_0 = 5$ .

### 10.1.1 Evolution of the AMP algorithm

Our first set of experiments aims at illustrating the evolution of the profile  $\phi(t)$  defined by state evolution versus iteration  $t$ , and comparing the predicted errors by the state evolution with the empirical errors.

Figure 10.1.1 shows the evolution of profile  $\phi(t) \in \mathbb{R}^{L_r}$ , given by the state evolution recursion (8.2.4). As explained in Section 9.1, in the spatially coupled sensing matrix, additional measurements are associated to the first few coordinates of  $x$ , namely,  $2\rho^{-1}N = 1000$  first coordinates. This ensures that the values of these coordinates are recovered up to a mean square error of order  $\sigma^2$ . This is reflected in the figure as the profile  $\phi$  becomes of order  $\sigma^2$  on the first few entries after a few iterations (see  $t = 5$  in the figure). As the iteration proceeds, the contribution of these components is correctly subtracted from all the measurements, and essentially they are removed from the problem. Now, in the resulting problem the first few variables are effectively oversampled and the algorithm reconstructs their values up to a mean square error of  $\sigma^2$ . Correspondingly, the profile  $\phi$  falls to a value of order  $\sigma^2$  in the next few coordinates. As the process is iterated, all the variables are progressively reconstructed and the profile  $\phi$  follows a traveling wave with constant velocity. After a sufficient number of iterations ( $t = 800$  in the figure),  $\phi$  is uniformly of order  $\sigma^2$ .

Next, we numerically verify that the deterministic state evolution recursion predicts the performance of the AMP at each iteration. Define the empirical and the predicted mean

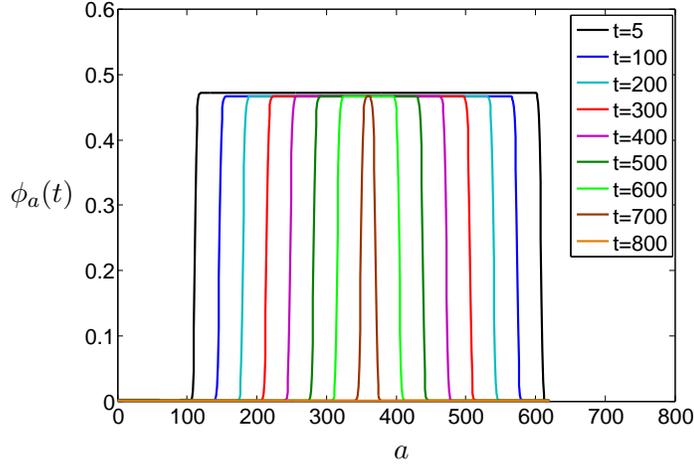


Figure 10.1.1: Profile  $\phi_a(t)$  versus  $a$  for several iteration numbers.

square errors respectively by

$$\text{MSE}_{\text{AMP}}(t) = \frac{1}{n} \|x^t(y) - x\|_2^2, \tag{10.1.2}$$

$$\text{MSE}_{\text{SE}}(t) = \frac{1}{L_c} \sum_{i \in \mathcal{C}} \text{mmse} \left( \sum_{a \in \mathcal{R}} W_{a,i} \phi_a^{-1}(t-1) \right). \tag{10.1.3}$$

The values of  $\text{MSE}_{\text{AMP}}(t)$  and  $\text{MSE}_{\text{SE}}(t)$  are depicted versus  $t$  in Fig. 10.1.2. (Values of  $\text{MSE}_{\text{AMP}}(t)$  and the error bars correspond to  $M = 30$  Monte Carlo instances). This verifies that the state evolution provides an iteration-by-iteration prediction of the AMP performance. We observe that  $\text{MSE}_{\text{AMP}}(t)$  (and  $\text{MSE}_{\text{SE}}(t)$ ) decreases linearly versus  $t$ .

### 10.1.2 Phase diagram

Consider a noiseless setting and let  $\mathcal{A}$  be a sensing matrix–reconstruction algorithm scheme. The curve  $\varepsilon \mapsto \delta_{\mathcal{A}}(\varepsilon)$  describes the sparsity–undersampling tradeoff of  $\mathcal{A}$  if the following happens in the large-system limit  $n, m \rightarrow \infty$ , with  $m/n = \delta$ . The scheme  $\mathcal{A}$  does (with high probability) correctly recover the original signal provided  $\delta > \delta_{\mathcal{A}}(\varepsilon)$ , while for  $\delta < \delta_{\mathcal{A}}(\varepsilon)$  the algorithm fails with high probability.

The goal of this section is to numerically compute the sparsity–undersampling tradeoff curve for the proposed scheme (spatially coupled sensing matrices and Bayes optimal AMP). We consider a set of sparsity parameters  $\varepsilon \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ , and for each value of  $\varepsilon$ , evaluate the empirical phase transition through a logit fit (we omit details, but follow

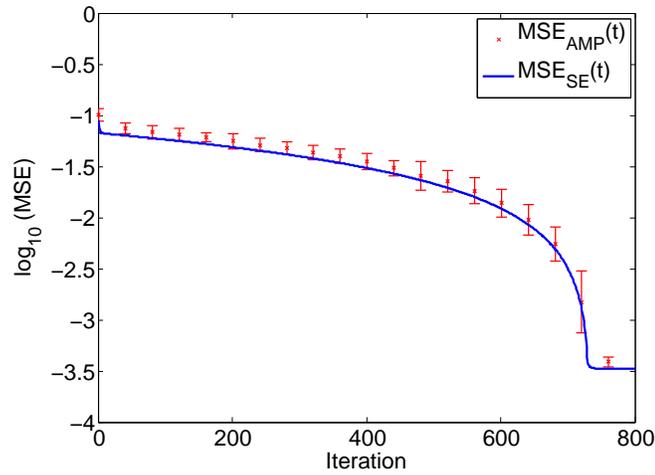


Figure 10.1.2: Comparison of  $\text{MSE}_{\text{AMP}}$  and  $\text{MSE}_{\text{SE}}$  across iteration.

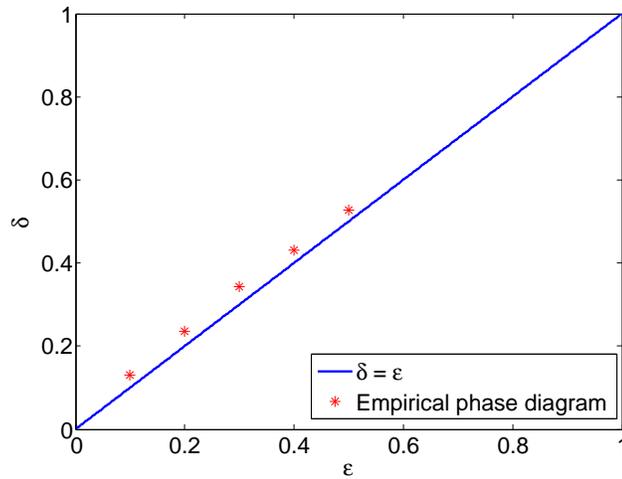


Figure 10.1.3: Phase diagram for the spatially coupled sensing matrices and Bayes optimal AMP.

the methodology described in [42]). As shown in Fig 10.1.3, the numerical results are consistent with the claim that this scheme achieves the information theoretic lower bound  $\delta > \bar{d}(p_X) = \varepsilon$ . (We indeed expect the gap to decrease further by taking larger values of  $L$ ).

## 10.2 State evolution: an heuristic derivation

This section presents an heuristic derivation of the state evolution equations (8.2.4). Our objective is to provide some basic intuition: a proof in a more general setting will appear in a separate publication [74]. An heuristic derivation similar to the present one, for the special cases of sensing matrices with i.i.d. entries was presented in [7].

Consider the recursion (7.1.3)-(7.1.4), and introduce the following modifications: (i) At each iteration, replace the random matrix  $A$  with a new independent copy  $A^t$ ; (ii) Replace the observation vector  $y$  with  $y^t = A^t x + w$ ; (iii) Eliminate the last term in the update equation for  $r^t$ . Then, we have the following update rules:

$$x^{t+1} = \eta_t(x^t + (Q^t \odot A^t)^* r^t), \quad (10.2.1)$$

$$r^t = y^t - A^t x^t, \quad (10.2.2)$$

where  $A^0, A^1, A^2, \dots$  are i.i.d. random matrices distributed according to the ensemble  $\mathcal{M}(W, M, N)$ , i.e.,

$$A_{ij}^t \sim \mathbf{N}\left(0, \frac{1}{M} W_{\mathbf{g}(i), \mathbf{g}(j)}\right). \quad (10.2.3)$$

Rewriting the recursion by eliminating  $r^t$ , we obtain:

$$\begin{aligned} x^{t+1} &= \eta_t((Q^t \odot A^t)^* y^t + (I - (Q^t \odot A^t)^* A^t) x^t) \\ &= \eta_t(x + (Q^t \odot A^t)^* w + B^t(x^t - x)), \end{aligned} \quad (10.2.4)$$

where  $B^t = I - (Q^t \odot A^t)^* A^t \in \mathbb{R}^{n \times n}$ . Note that the recursion (10.2.4) does not correspond to the AMP update rules defined per Eqs. (7.1.3) and (7.1.4). In particular, it does not correspond to any practical algorithm since the sensing matrix  $A$  is a fixed input to a reconstruction algorithm and is not resampled at each iteration. However, it is much easier to analyze, since  $A^t$  is independent of  $x^t$  and therefore the distribution of  $(Q^t \odot A^t)^* r^t$  can be easily characterized. Also, it is useful for presenting the intuition behind the AMP algorithm and to emphasize the role of the term  $\mathbf{b}^t \odot r^{t-1}$  in the update rule for  $r^t$ . As it emerges from the proof of [7], this term does asymptotically cancel dependencies across iterations.

By virtue of the central limit theorem, each entry of  $B^t$  is approximately normal. More

specifically,  $B_{ij}^t$  is approximately normal with mean zero and variance

$$\frac{1}{M} \sum_{r \in \mathbb{R}} W_{r, \mathbf{g}(i)} W_{r, \mathbf{g}(j)} (Q_{r, \mathbf{g}(i)}^t)^2,$$

for  $i, j \in [n]$ . Define  $\hat{\tau}_t(s) = \lim_{N \rightarrow \infty} \|x_{C(s)}^t - x_{C(s)}\|^2/N$ , for  $s \in \mathbb{C}$ . It is easy to show that distinct entries in  $B^t$  are approximately independent. Also,  $B^t$  is independent of  $\{B^s\}_{1 \leq s \leq t-1}$ , and in particular, of  $x^t - x$ . Hence,  $B^t(x^t - x)$  converges to a vector, say  $v$ , with i.i.d. normal entries, and for  $i \in [n]$ ,

$$\mathbb{E}\{v_i\} = 0, \quad \mathbb{E}\{v_i^2\} = \frac{N}{M} \sum_{u \in \mathbb{C}} \sum_{r \in \mathbb{R}} W_{r, \mathbf{g}(i)} W_{r, u} (Q_{r, \mathbf{g}(i)}^t)^2 \hat{\tau}_t(u). \quad (10.2.5)$$

Conditional on  $w$ ,  $(Q^t \odot A^t)^* w$  is a vector with i.i.d. zero-mean normal entries. Also, the variance of its  $i^{\text{th}}$  entry, for  $i \in [n]$ , is

$$\frac{1}{M} \sum_{r \in \mathbb{R}} W_{r, \mathbf{g}(i)} (Q_{r, \mathbf{g}(i)}^t)^2 \|w_{R(r)}\|^2, \quad (10.2.6)$$

which converges to  $\sum_{r \in \mathbb{R}} W_{r, \mathbf{g}(i)} (Q_{r, \mathbf{g}(i)}^t)^2 \sigma^2$ , by the law of large numbers. With slightly more work, it can be shown that these entries are approximately independent of the ones of  $B^t(x^t - x)$ .

Summarizing, the  $i^{\text{th}}$  entry of the vector in the argument of  $\eta_t$  in Eq. (10.2.4) converges to  $X + \tau_t(\mathbf{g}(i))^{1/2} Z$  with  $Z \sim \mathcal{N}(0, 1)$  independent of  $X$ , and

$$\tau_t(s) = \sum_{r \in \mathbb{R}} W_{r, s} (Q_{r, s}^t)^2 \left\{ \sigma^2 + \frac{1}{\delta} \sum_{u \in \mathbb{C}} W_{r, u} \hat{\tau}_t(u) \right\}, \quad (10.2.7)$$

for  $s \in \mathbb{C}$ . In addition, using Eq. (10.2.4) and invoking Eqs. (8.3.2), (8.3.3), each entry of  $x_{C(s)}^{t+1} - x_{C(s)}$  converges to  $\eta_{t,s}(X + \tau_t(s)^{1/2} Z) - X$ , for  $s \in \mathbb{C}$ . Therefore,

$$\begin{aligned} \hat{\tau}_{t+1}(s) &= \lim_{N \rightarrow \infty} \frac{1}{N} \|x_{C(s)}^{t+1} - x_{C(s)}\|^2 \\ &= \mathbb{E}\{[\eta_{t,s}(X + \tau_t(s)^{1/2} Z) - X]^2\} = \text{mmse}(\tau_t(s)^{-1}). \end{aligned} \quad (10.2.8)$$

Using Eqs. (10.2.7) and (10.2.8), we obtain:

$$\tau_{t+1}(s) = \sum_{r \in \mathbb{R}} W_{r,s} (Q_{r,s}^{t+1})^2 \left\{ \sigma^2 + \frac{1}{\delta} \sum_{u \in \mathbb{C}} W_{r,u} \text{mmse}(\tau_t(u)^{-1}) \right\}. \quad (10.2.9)$$

Applying the change of variable  $\tau_t(u)^{-1} = \sum_{b \in \mathbb{R}} W_{b,u} \phi_b(t)^{-1}$ , and substituting for  $Q_{r,s}^{t+1}$  from Eq. (8.3.1), we obtain the state evolution recursion, Eq. (8.2.4).

In conclusion, we showed that the state evolution recursion would hold if the matrix  $A$  was resampled independently from the ensemble  $\mathcal{M}(W, M, N)$ , at each iteration. However, in our proposed AMP algorithm, the matrix  $A$  is constant across iterations, and the above argument is not valid since  $x^t$  and  $A$  are dependent. The dependency between  $A$  and  $x^t$  cannot be neglected. Indeed, state evolution does not apply to the following naive iteration in which we dropped the memory term  $\mathbf{b}^t \odot r^{t-1}$ :

$$x^{t+1} = \eta_t(x^t + (Q^t \odot A)^* r^t), \quad (10.2.10)$$

$$r^t = y^t - Ax^t. \quad (10.2.11)$$

Indeed, the term  $\mathbf{b}^t \odot r^{t-1}$  leads to an asymptotic cancellation of the dependencies between  $A$  and  $x^t$  as proved in [7, 74].

## Chapter 11

# Analysis of state evolution: Proof of Lemma 10.0.2

This section is devoted to the analysis of the state evolution recursion for spatially coupled matrices  $A$ , hence proving Lemma 10.0.2.

In order to prove Lemma 10.0.2, we will construct a free energy functional  $E_{\mathcal{W}}(\phi)$  such that the fixed points of the state evolution are the stationary points of  $E_{\mathcal{W}}$ . We then assume by contradiction that the claim of the lemma does not hold, i.e.,  $\phi(t)$  converges to a fixed point  $\phi(\infty)$  with  $\phi_a(\infty) \gg \sigma^2$  for a significant fraction of the indices  $a$ . We then obtain a contradiction by describing an infinitesimal deformation of this fixed point (roughly speaking, a shift to the right) that decreases its free energy.

### 11.1 Outline

A more precise outline of the proof is given below:

- (i) We establish some useful properties of the state evolution sequence  $\{\phi(t), \psi(t)\}_{t \geq 0}$ . This includes a monotonicity property as well as a lower and an upper bound for the state vectors.
- (ii) We define a modified state evolution sequence, denoted by  $\{\phi^{\text{mod}}(t), \psi^{\text{mod}}(t)\}_{t \geq 0}$ . This sequence dominates the original state vectors (see Lemma 11.3.2) and hence it suffices to focus on the modified state evolution to get the desired result. As we will see the modified state evolution is more amenable to analysis.

- (iii) We next introduce continuum state evolution which serves as the continuous analog of the modified state evolution. (The continuum states are functions rather than vectors). The bounds on the continuum state evolution sequence lead to bounds on the modified state vectors.
- (iv) Analysis of the continuum state evolution incorporates the definition of a free energy functional defined on the space of non-negative measurable functions with bounded support. The energy is constructed in a way to ensure that the fixed points of the continuum state evolution are the stationary points of the free energy. Then, we show that if the undersampling rate is greater than the information dimension, the solution of the continuum state evolution can be made as small as  $O(\sigma^2)$ . If this were not the case, the (large) fixed point could be perturbed slightly in such a way that the free energy decreases to the first order. However, since the fixed point is a stationary point of the free energy, this leads to a contradiction.

## 11.2 Properties of the state evolution sequence

Throughout this section  $p_X$  is a given probability distribution over the real line, and  $X \sim p_X$ . Also, we will take  $\sigma > 0$ . The result for the noiseless model (Corollary 7.1.8) follows by letting  $\sigma \downarrow 0$ . Recall the inequality

$$\text{mmse}(s) \leq \min(\text{Var}(X), \frac{1}{s}). \quad (11.2.1)$$

**Definition 11.2.1.** For two vectors  $\phi, \tilde{\phi} \in \mathbb{R}^K$ , we write  $\phi \succeq \tilde{\phi}$  if all  $\phi_r \geq \tilde{\phi}_r$  for  $r \in \{1, \dots, K\}$ .

**Proposition 11.2.2.** For any  $W \in \mathbb{R}_+^{\mathbb{R} \times \mathbb{C}}$ , the maps  $\mathsf{T}'_W : \mathbb{R}_+^{\mathbb{R}} \rightarrow \mathbb{R}_+^{\mathbb{C}}$  and  $\mathsf{T}''_W : \mathbb{R}_+^{\mathbb{C}} \rightarrow \mathbb{R}_+^{\mathbb{R}}$ , as defined in Definition 8.2.1, are monotone; i.e., if  $\phi \succeq \tilde{\phi}$  then  $\mathsf{T}'_W(\phi) \succeq \mathsf{T}'_W(\tilde{\phi})$ , and if  $\psi \succeq \tilde{\psi}$  then  $\mathsf{T}''_W(\psi) \succeq \mathsf{T}''_W(\tilde{\psi})$ . Consequently,  $\mathsf{T}_W$  is also monotone.

*Proof.* It follows immediately from the fact that  $s \mapsto \text{mmse}(s)$  is a monotone decreasing function and the positivity of the matrix  $W$ .  $\square$

**Proposition 11.2.3.** The state evolution sequence  $\{\phi(t), \psi(t)\}_{t \geq 0}$  with initial condition  $\psi_i(0) = \infty$ , for  $i \in \mathbb{C}$ , is monotone decreasing, in the sense that  $\phi(0) \succeq \phi(1) \succeq \phi(2) \succeq \dots$  and  $\psi(0) \succeq \psi(1) \succeq \psi(2) \succeq \dots$ .

*Proof.* Since  $\psi_i(0) = \infty$  for all  $i$ , we have  $\psi(0) \succeq \psi(1)$ . The thesis follows from the monotonicity of the state evolution map.  $\square$

**Proposition 11.2.4.** *The state evolution sequence  $\{\phi(t), \psi(t)\}_{t \geq 0}$  is monotone increasing in  $\sigma^2$ . Namely, let  $0 \leq \sigma_1 \leq \sigma_2$  and  $\{\phi^{(1)}(t), \psi^{(1)}(t)\}_{t \geq 0}$ ,  $\{\phi^{(2)}(t), \psi^{(2)}(t)\}_{t \geq 0}$  be the state evolution sequences corresponding to setting, respectively,  $\sigma^2 = \sigma_1^2$  and  $\sigma^2 = \sigma_2^2$  in Eq. (8.2.4), with identical initial conditions. Then  $\phi^{(1)}(t) \preceq \phi^{(2)}(t)$ ,  $\psi^{(1)}(t) \preceq \psi^{(2)}(t)$  for all  $t$ .*

*Proof.* Follows immediately from Proposition 11.2.2 and the monotonicity of the one-step mapping (8.2.4).  $\square$

**Lemma 11.2.5.** *Assume  $\delta L_0 > 3$ . Then there exists  $t_0$  (depending only on  $p_X$ ), such that, for all  $t \geq t_0$  and all  $i \in \{-2\rho^{-1}, \dots, -1\}$ ,  $a \in \mathbf{R}_i$ , we have*

$$\psi_i(t) \leq \text{mmse}\left(\frac{L_0}{2\sigma^2}\right) \leq \frac{2\sigma^2}{L_0}, \quad (11.2.2)$$

$$\phi_a(t) \leq \sigma^2 + \frac{1}{\delta} \text{mmse}\left(\frac{L_0}{2\sigma^2}\right) \leq \left(1 + \frac{2}{\delta L_0}\right) \sigma^2. \quad (11.2.3)$$

*Proof.* Take  $i \in \{-2\rho^{-1}, \dots, -1\}$ . For  $a \in \mathbf{R}_i$ , we have  $\phi_a(t) = \sigma^2 + (1/\delta)\psi_i(t)$ . Further from  $\text{mmse}(s) \leq 1/s$ , we deduce that

$$\begin{aligned} \psi_i(t+1) &= \text{mmse}\left(\sum_{b \in \mathbf{R}} W_{b,i} \phi_b(t)^{-1}\right) \leq \left(\sum_{b \in \mathbf{R}} W_{b,i} \phi_b(t)^{-1}\right)^{-1} \\ &\leq \left(\sum_{a \in \mathbf{R}_i} W_{a,i} \phi_a(t)^{-1}\right)^{-1} = \left(L_0 \phi_a(t)^{-1}\right)^{-1} = \frac{\phi_a(t)}{L_0}. \end{aligned} \quad (11.2.4)$$

Here we used the facts that  $W_{a,i} = 1$ , for  $a \in \mathbf{R}_i$  and  $|\mathbf{R}_i| = L_0$ . Substituting in the earlier relation, we get  $\psi_i(t+1) \leq (1/L_0)(\sigma^2 + (1/\delta)\psi_i(t))$ . Recalling that  $\delta L_0 > 3$ , we have  $\psi_i(t) \leq 2\sigma^2/L_0$ , for all  $t$  sufficiently large. Now, using this in the equation for  $\phi_a(t)$ ,  $a \in \mathbf{R}_i$ , we obtain

$$\phi_a(t) = \sigma^2 + \frac{1}{\delta} \psi_i(t) \leq \left(1 + \frac{2}{\delta L_0}\right) \sigma^2. \quad (11.2.5)$$

We prove the other claims by repeatedly substituting in the previous bounds. In particular,

$$\begin{aligned}\psi_i(t) &= \text{mmse}\left(\sum_{b \in \mathbf{R}} W_{b,i} \phi_b(t-1)^{-1}\right) \leq \text{mmse}\left(\sum_{a \in \mathbf{R}_i} W_{a,i} \phi_a(t)^{-1}\right) \\ &= \text{mmse}(L_0 \phi_a(t)^{-1}) \leq \text{mmse}\left(\frac{L_0}{(1 + \frac{2}{\delta L_0}) \sigma^2}\right) \leq \text{mmse}\left(\frac{L_0}{2\sigma^2}\right),\end{aligned}\tag{11.2.6}$$

where we used Eq. (11.2.5) in the penultimate inequality. Finally,

$$\phi_a(t) \leq \sigma^2 + \frac{1}{\delta} \psi_i(t) \leq \sigma^2 + \frac{1}{\delta} \text{mmse}\left(\frac{L_0}{2\sigma^2}\right),\tag{11.2.7}$$

where the inequality follows from Eq. (11.2.6).  $\square$

Next we prove a lower bound on the state evolution sequence. Here and below  $\mathbf{C}_0 \equiv \mathbf{C} \setminus \{-2\rho^{-1}, \dots, -1\} \cong \{0, \dots, L-1\}$ . Also, recall that  $\mathbf{R}_0 \equiv \{-\rho^{-1}, \dots, 0, \dots, L-1 + \rho^{-1}\}$ . (See Fig. 8.4.2).

**Lemma 11.2.6.** *For any  $t \geq 0$ , and any  $i \in \mathbf{C}_0$ ,  $\psi_i(t) \geq \text{mmse}(2\sigma^{-2})$ . Further, for any  $a \in \mathbf{R}_0$  and any  $t \geq 0$  we have  $\phi_a(t) \geq \sigma^2 + (2\delta)^{-1} \text{mmse}(2\sigma^2)$ .*

*Proof.* Since  $\phi_a(t) \geq \sigma^2$  by definition, we have, for  $i \geq 0$ ,  $\psi_i(t) \geq \text{mmse}(\sigma^{-2} \sum_b W_{bi}) \geq \text{mmse}(2\sigma^{-2})$ , where we used the fact that the restriction of  $W$  to columns in  $\mathbf{C}_0$  is roughly column-stochastic. Plugging this into the expression for  $\phi_a$ , we get

$$\phi_a(t) \geq \sigma^2 + \frac{1}{\delta} \sum_{i \in \mathbf{C}} W_{a,i} \text{mmse}(2\sigma^{-2}) \geq \sigma^2 + \frac{1}{2\delta} \text{mmse}(2\sigma^{-2}).\tag{11.2.8}$$

$\square$

Notice that for  $L_{0,*} \geq 4$  and for all  $L_0 > L_{0,*}$ , the upper bound for  $\psi_i(t)$ ,  $i \in \{-2\rho^{-1}, \dots, -1\}$ , given in Lemma 11.2.5 is below the lower bound for  $\psi_i(t)$ , with  $i \in \mathbf{C}_0$ , given in Lemma 11.2.6; i.e., for all  $\sigma$ ,

$$\text{mmse}\left(\frac{L_0}{2\sigma^2}\right) \leq \text{mmse}\left(\frac{2}{\sigma^2}\right).\tag{11.2.9}$$

### 11.3 Modified state evolution

First of all, by Proposition 11.2.4 we can assume, without loss of generality  $\sigma > 0$ .

Motivated by the monotonicity properties of the state evolution sequence mentioned in Lemmas 11.2.5 and 11.2.6, we introduce a new state evolution recursion that dominates the original one and yet is more amenable to analysis. Namely, we define the modified state evolution maps  $F'_W : \mathbb{R}_+^{\mathbb{R}_0} \rightarrow \mathbb{R}_+^{\mathbb{C}_0}$ ,  $F''_W : \mathbb{R}_+^{\mathbb{C}_0} \rightarrow \mathbb{R}_+^{\mathbb{R}_0}$ . For  $\phi = (\phi_a)_{a \in \mathbb{R}_0} \in \mathbb{R}_+^{\mathbb{R}_0}$ ,  $\psi = (\psi_i)_{i \in \mathbb{C}_0} \in \mathbb{R}_+^{\mathbb{C}_0}$ , and for all  $i \in \mathbb{C}_0$ ,  $a \in \mathbb{R}_0$ , let:

$$F'_W(\phi)_i = \text{mmse}\left(\sum_{b \in \mathbb{R}_0} W_{b-i} \phi_b^{-1}\right), \quad (11.3.1)$$

$$F''_W(\psi)_a = \sigma^2 + \frac{1}{\delta} \sum_{i \in \mathbb{Z}} W_{a-i} \psi_i. \quad (11.3.2)$$

where, in the last equation we set by convention,  $\psi_i(t) = \text{mmse}(L_0/(2\sigma^2))$  for  $i \leq -1$ , and  $\psi_i = \infty$  for  $i \geq L$ , and recall the shorthand  $W_{a-i} \equiv \rho \mathcal{W}(\rho(a-i))$  introduced in Section 8.4. We also let  $F_W = F'_W \circ F''_W$ .

**Definition 11.3.1.** *The modified state evolution sequence is the sequence  $\{\phi(t), \psi(t)\}_{t \geq 0}$  with  $\phi(t) = F''_W(\psi(t))$  and  $\psi(t+1) = F'_W(\phi(t))$  for all  $t \geq 0$ , and  $\psi_i(0) = \infty$  for all  $i \in \mathbb{C}_0$ . We also adopt the convention that, for  $i \geq L$ ,  $\psi_i(t) = +\infty$  and for  $i \leq -1$ ,  $\psi_i(t) = \text{mmse}(L_0/(2\sigma^2))$ , for all  $t$ .*

Lemma 11.2.5 then implies the following.

**Lemma 11.3.2.** *Let  $\{\phi(t), \psi(t)\}_{t \geq 0}$  denote the state evolution sequence as per Definition 8.2.2, and  $\{\phi^{\text{mod}}(t), \psi^{\text{mod}}(t)\}_{t \geq 0}$  denote the modified state evolution sequence as per Definition 11.3.1. Then, there exists  $t_0$  (depending only on  $p_X$ ), such that, for all  $t \geq t_0$ ,  $\phi(t) \preceq \phi^{\text{mod}}(t - t_0)$  and  $\psi(t) \preceq \psi^{\text{mod}}(t - t_0)$ .*

*Proof.* Choose  $t_0 = t(L_0, \delta)$  as given by Lemma 11.2.5. We prove the claims by induction on  $t$ . For the induction basis ( $t = t_0$ ), we have from Lemma 11.2.5,  $\psi_i(t_0) \leq \text{mmse}(L_0/(2\sigma^2)) = \psi_i^{\text{mod}}(0)$ , for  $i \leq -1$ . Also, we have  $\psi_i^{\text{mod}}(0) = \infty \geq \psi_i(t_0)$ , for  $i \geq 0$ . Further,

$$\phi_a^{\text{mod}}(0) = F''_W(\psi^{\text{mod}}(0))_a \geq \mathsf{T}''_W(\psi^{\text{mod}}(0))_a \geq \mathsf{T}''_W(\psi(t_0))_a = \phi_a(t_0), \quad (11.3.3)$$

for  $a \in \mathbb{R}_0$ . Here, the last inequality follows from monotonicity of  $\mathsf{T}''_W$  (Proposition 11.2.2). Now, assume that the claim holds for  $t$ ; we prove it for  $t+1$ . For  $i \in \mathbb{C}_0$ , we have

$$\begin{aligned} \psi_i^{\text{mod}}(t+1-t_0) &= F'_W(\phi^{\text{mod}}(t-t_0))_i = \mathsf{T}'_W(\phi^{\text{mod}}(t-t_0))_i \\ &\geq \mathsf{T}'_W(\phi(t))_i = \psi_i(t+1), \end{aligned} \quad (11.3.4)$$

where the inequality follows from monotonicity of  $\mathsf{T}'_W$  (Proposition 11.2.2) and the induction hypothesis. In addition, for  $a \in \mathbb{R}_0$ ,

$$\begin{aligned} \phi_a^{\text{mod}}(t+1-t_0) &= \mathsf{F}''_W(\psi^{\text{mod}}(t+1-t_0))_a \geq \mathsf{T}''_W(\psi^{\text{mod}}(t+1-t_0))_a \\ &\geq \mathsf{T}''_W(\psi(t+1))_a = \phi_a(t+1). \end{aligned} \quad (11.3.5)$$

Here, the last inequality follows from monotonicity of  $\mathsf{T}''_W$  and Eq. (11.3.4).  $\square$

By Lemma 11.3.2, we can now focus on the modified state evolution sequence in order to prove Lemma 10.0.2. Notice that the mapping  $\mathsf{F}_W$  has a particularly simple description in terms of a shift-invariant state evolution mapping. Explicitly, define  $\mathsf{T}'_{W,\infty} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ ,  $\mathsf{T}''_{W,\infty} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ , by letting, for  $\phi, \psi \in \mathbb{R}^{\mathbb{Z}}$  and all  $i, a \in \mathbb{Z}$ :

$$\mathsf{T}'_{W,\infty}(\phi)_i = \text{mmse}\left(\sum_{b \in \mathbb{Z}} W_{b-i} \phi_b^{-1}\right), \quad (11.3.6)$$

$$\mathsf{T}''_{W,\infty}(\psi)_a = \sigma^2 + \frac{1}{\delta} \sum_{i \in \mathbb{Z}} W_{a-i} \psi_i. \quad (11.3.7)$$

Further, define the embedding  $\mathsf{H} : \mathbb{R}^{\mathbb{C}_0} \rightarrow \mathbb{R}^{\mathbb{Z}}$  by letting

$$(\mathsf{H}\psi)_i = \begin{cases} \text{mmse}(L_0/(2\sigma^2)) & \text{if } i < 0, \\ \psi_i & \text{if } 0 \leq i \leq L-1, \\ +\infty & \text{if } i \geq L, \end{cases} \quad (11.3.8)$$

And the restriction mapping  $\mathsf{H}'_{a,b} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{b-a+1}$  by  $\mathsf{H}'_{a,b}\psi = (\psi_a, \dots, \psi_b)$ .

**Lemma 11.3.3.** *With the above definitions,  $\mathsf{F}_W = \mathsf{H}'_{0,L-1} \circ \mathsf{T}_{W,\infty} \circ \mathsf{H}$ .*

*Proof.* Clearly, for any  $\psi = (\psi_i)_{i \in \mathbb{C}_0}$ , we have  $\mathsf{T}''_W \circ \mathsf{H}(\psi)_a = \mathsf{F}''_W \circ \mathsf{H}(\psi)_a$  for  $a \in \mathbb{R}_0$ , since the definition of the embedding  $\mathsf{H}$  is consistent with the convention adopted in defining the modified state evolution. Moreover, for  $i \in \mathbb{C}_0 \cong \{0, \dots, L-1\}$ , we have

$$\begin{aligned} \mathsf{T}'_{W,\infty}(\phi)_i &= \text{mmse}\left(\sum_{b \in \mathbb{Z}} W_{b-i} \phi_b^{-1}\right) = \text{mmse}\left(\sum_{-\rho^{-1} \leq b \leq L-1+\rho^{-1}} W_{b-i} \phi_b^{-1}\right) \\ &= \text{mmse}\left(\sum_{b \in \mathbb{R}_0} W_{b-i} \phi_b^{-1}\right) = \mathsf{F}'_W(\phi)_i. \end{aligned} \quad (11.3.9)$$

Hence,  $\mathsf{T}'_{W,\infty} \circ \mathsf{T}''_{W,\infty} \circ \mathsf{H}(\psi)_i = \mathsf{F}'_W \circ \mathsf{F}''_W \circ \mathsf{H}(\psi)_i$ , for  $i \in \mathbb{C}_0$ . Therefore,  $\mathsf{H}'_{0,L-1} \circ \mathsf{T}_{W,\infty} \circ \mathsf{H}(\psi) =$

$F_W \circ H(\psi)$ , for any  $\psi \in \mathbb{R}_+^{C_0}$ , which completes the proof.  $\square$

We will say that a vector  $\psi \in \mathbb{R}^K$  is *nondecreasing* if, for every  $1 \leq i < j \leq K$ ,  $\psi_i \leq \psi_j$ .

**Lemma 11.3.4.** *If  $\psi \in \mathbb{R}^{C_0}$  is nondecreasing, with  $\psi_i \geq \text{mmse}(L_0/(2\sigma^2))$  for all  $i$ , then  $F_W(\psi)$  is nondecreasing as well. In particular, if  $\{\phi(t), \psi(t)\}_{t \geq 0}$  is the modified state evolution sequence, then  $\phi(t)$  and  $\psi(t)$  are nondecreasing for all  $t$ .*

*Proof.* By Lemma 11.3.3, we know that  $F_W = H'_{0, L-1} \circ T_{W, \infty} \circ H$ . We first notice that, by the assumption  $\psi_i \geq \text{mmse}(L_0/(2\sigma^2))$ , we have that  $H(\psi)$  is nondecreasing.

Next, if  $\psi \in \mathbb{R}^{\mathbb{Z}}$  is nondecreasing,  $T_{W, \infty}(\psi)$  is nondecreasing as well. In fact, the mappings  $T'_{W, \infty}$  and  $T''_{W, \infty}$  both preserve the nondecreasing property, since both are shift invariant, and  $\text{mmse}(\cdot)$  is a decreasing function. Finally, the restriction of a nondecreasing vector is obviously nondecreasing.

This proves that  $F_W$  preserves the nondecreasing property. To conclude that  $\psi(t)$  is nondecreasing for all  $t$ , notice that the condition  $\psi_i(t) \geq \text{mmse}(L_0/(2\sigma^2))$  is satisfied at all  $t$  by Lemma 11.2.6 and condition (11.2.9). The claim for  $\psi(t)$  follows by induction.

Now, since  $F''_W$  preserves the nondecreasing property, we have  $\phi(t) = F''_W(\psi(t))$  is nondecreasing for all  $t$ , as well.  $\square$

## 11.4 Continuum state evolution

We start by defining the continuum state evolution mappings. For  $\Omega \subseteq \mathbb{R}$ , let  $\mathcal{M}(\Omega)$  be the space of non-negative measurable functions on  $\Omega$  (up to measure-zero redefinitions). Define  $\mathcal{F}'_{\mathcal{W}} : \mathcal{M}([-1, \ell + 1]) \rightarrow \mathcal{M}([0, \ell])$  and  $\mathcal{F}''_{\mathcal{W}} : \mathcal{M}([0, \ell]) \rightarrow \mathcal{M}([-1, \ell + 1])$  as follows. For  $\phi \in \mathcal{M}([-1, \ell + 1])$ ,  $\psi \in \mathcal{M}([0, \ell])$ , and for all  $x \in [0, \ell]$ ,  $y \in [-1, \ell + 1]$ , we let

$$\mathcal{F}'_{\mathcal{W}}(\phi)(x) = \text{mmse}\left(\int_{-1}^{\ell+1} \mathcal{W}(x-z)\phi(z)^{-1}dz\right), \quad (11.4.1)$$

$$\mathcal{F}''_{\mathcal{W}}(\psi)(y) = \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(y-x)\psi(x)dx, \quad (11.4.2)$$

where we adopt the convention that  $\psi(x) = \text{mmse}(L_0/(2\sigma^2))$  for  $x < 0$ , and  $\psi(x) = \infty$  for  $x > \ell$ .

**Definition 11.4.1.** *The continuum state evolution sequence is the sequence  $\{\phi(\cdot; t), \psi(\cdot; t)\}_{t \geq 0}$ , with  $\phi(t) = \mathcal{F}'_{\mathcal{W}}(\psi(t))$  and  $\psi(t+1) = \mathcal{F}'_{\mathcal{W}}(\phi(t))$  for all  $t \geq 0$ , and  $\psi(x; 0) = \infty$  for all*

$x \in [0, \ell]$ .

Recalling Eq. (11.2.1), we have  $\psi(x; t) = \mathcal{F}'_{\mathcal{W}}(\phi(t-1))(x) \leq \text{Var}(X)$ , for  $t \geq 1$ . Also,  $\phi(x; t) = \mathcal{F}''_{\mathcal{W}}(\psi(t))(x) \leq \sigma^2 + (1/\delta)\text{Var}(X)$ , for  $t \geq 1$ . Define,

$$\Phi_M = 1 + \frac{1}{\delta} \text{Var}(X). \quad (11.4.3)$$

Assuming  $\sigma < 1$ , we have  $\phi(x; t) < \Phi_M$ , for all  $t \geq 1$ .

The point of introducing continuum state evolution is that by construction of the matrix  $W$  and the continuity of  $\mathcal{W}$ , when  $\rho$  is small, one can approximate summation by integration and study the evolution of the continuum states which are represented by functions rather than vectors. This observation is formally stated in lemma below.

**Lemma 11.4.2.** *Let  $\{\phi(\cdot; t), \psi(\cdot; t)\}_{t \geq 0}$  be the continuum state evolution sequence and  $\{\phi(t), \psi(t)\}_{t \geq 0}$  be the modified discrete state evolution sequence, with parameters  $\rho$  and  $L = \ell/\rho$ . Then for any  $t \geq 0$*

$$\lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{i=0}^{L-1} |\psi_i(t) - \psi(\rho i; t)| = 0, \quad (11.4.4)$$

$$\lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi_a(t) - \phi(\rho a; t)| = 0. \quad (11.4.5)$$

Lemma 11.4.2 is proved in Appendix D.1.

**Corollary 11.4.3.** *The continuum state evolution sequence  $\{\phi(\cdot; t), \psi(\cdot; t)\}_{t \geq 0}$ , with initial condition  $\psi(x) = \text{mmse}(L_0/(2\sigma^2))$  for  $x < 0$ , and  $\psi(x) = \infty$  for  $x > \ell$ , is monotone decreasing, in the sense that  $\phi(x; 0) \geq \phi(x; 1) \geq \phi(x; 2) \geq \dots$  and  $\psi(x; 0) \geq \psi(x; 1) \geq \psi(x; 2) \geq \dots$ , for all  $x \in [0, \ell]$ .*

*Proof.* Follows immediately from Lemmas 11.2.3 and 11.4.2. □

**Corollary 11.4.4.** *Let  $\{\phi(\cdot; t), \psi(\cdot; t)\}_{t \geq 2}$  be the continuum state evolution sequence. Then for any  $t$ ,  $x \mapsto \psi(x; t)$  and  $x \mapsto \phi(x; t)$  are nondecreasing Lipschitz continuous functions.*

*Proof.* Nondecreasing property of functions  $x \mapsto \psi(x; t)$ , and  $x \mapsto \phi(x; t)$  follows immediately from Lemmas 11.3.4 and 11.4.2. Further, since  $\psi(x; t)$  is bounded for  $t \geq 1$ , and

$\mathcal{W}(\cdot)$  is Lipschitz continuous, recalling Eq. (11.4.2), the function  $x \mapsto \phi(x; t)$  is Lipschitz continuous as well, for  $t \geq 1$ . Similarly, since  $\sigma^2 < \phi(x; t) < \Phi_M$ , invoking Eq. (11.4.1), the function  $x \mapsto \psi(x; t)$  is Lipschitz continuous for  $t \geq 2$ .  $\square$

### 11.4.1 Free energy

A key role in our analysis is played by the free energy functional. In order to define the free energy, we first provide some preliminaries. Define the mutual information between  $X$  and a noisy observation of  $X$  at signal-to-noise ratio  $s$  by

$$I(s) \equiv I(X; \sqrt{s}X + Z), \quad (11.4.6)$$

with  $Z \sim \mathbf{N}(0, 1)$  independent of  $X \sim p_X$ . Recall the relation [62]

$$\frac{d}{ds} I(s) = \frac{1}{2} \text{mmse}(s). \quad (11.4.7)$$

Furthermore, the following identities relate the scaling law of mutual information under weak noise to Rényi information dimension [144].

**Proposition 11.4.5.** *Assume  $H(\lfloor X \rfloor) < \infty$ . Then*

$$\begin{aligned} \liminf_{s \rightarrow \infty} \frac{I(s)}{\frac{1}{2} \log s} &= \underline{d}(p_X), \\ \limsup_{s \rightarrow \infty} \frac{I(s)}{\frac{1}{2} \log s} &= \bar{d}(p_X). \end{aligned} \quad (11.4.8)$$

Now we are ready to define the free energy functional.

**Definition 11.4.6.** *Let  $\mathcal{W}(\cdot)$  be a shape function, and  $\sigma, \delta > 0$  be given. The corresponding free energy is the functional  $\mathbb{E}_{\mathcal{W}} : \mathcal{M}([-1, \ell + 1]) \rightarrow \overline{\mathbb{R}}$  defined as follows for  $\phi \in \mathcal{M}([-1, \ell + 1])$ :*

$$\mathbb{E}_{\mathcal{W}}(\phi) = \frac{\delta}{2} \int_{-1}^{\ell-1} \left\{ \frac{\varsigma^2(x)}{\phi(x)} + \log \phi(x) \right\} dx + \int_0^{\ell} I \left( \int \mathcal{W}(x-z) \phi(z)^{-1} dz \right) dx, \quad (11.4.9)$$

where

$$\varsigma^2(x) = \sigma^2 + \frac{1}{\delta} \left( \int_{y \leq 0} \mathcal{W}(y-x) dy \right) \text{mmse} \left( \frac{L_0}{2\sigma^2} \right). \quad (11.4.10)$$

The name ‘free energy’ is motivated by the connection with statistical physics, whereby  $E_{\mathcal{W}}(\phi)$  is the asymptotic log-partition function for the Gibbs-Boltzmann measure corresponding to the posterior distribution of  $x$  given  $y$ . (This connection is however immaterial for our proof and we will not explore it further, see for instance [82].)

Notice that this is where the Rényi information comes into the picture. The mutual information appears in the expression of the free energy and the mutual information is related to the Rényi information via Proposition 11.4.5.

Viewing  $E_{\mathcal{W}}$  as a function defined on the Banach space  $L_2([-1, \ell])$ , we will denote by  $\nabla E_{\mathcal{W}}(\phi)$  its Fréchet derivative at  $\phi$ . This will be identified, via standard duality, with a function in  $L_2([-1, \ell])$ . It is not hard to show that the Fréchet derivative exists on  $\{\phi : \phi(x) \geq \sigma^2\}$  and is such that

$$\begin{aligned} \nabla E_{\mathcal{W}}(\phi)(y) = & \\ & \frac{\delta}{2\phi^2(y)} \left\{ \phi(y) - \varsigma^2(y) - \frac{1}{\delta} \int_0^\ell \mathcal{W}(x-y) \text{mmse} \left( \int \mathcal{W}(x-z) \phi(z)^{-1} dz \right) dx \right\}, \end{aligned} \quad (11.4.11)$$

for  $-1 \leq y \leq \ell - 1$ . Note that the condition  $\phi(x) \geq \sigma^2$  is immediately satisfied by the state evolution sequence since, by Eq. (11.4.2),  $\mathcal{F}_{\mathcal{W}}''(\psi)(y) \geq \sigma^2$  for all  $y$  (because  $\mathcal{W}(y-x)$ ,  $\psi(x; t) \geq 0$ ); see also Definition 11.4.1.

The specific choice of the free energy in Eq. (11.4.9) ensures that the fixed points of the continuum state evolution are the stationary points of the free energy.

**Corollary 11.4.7.** *If  $\{\phi, \psi\}$  is the fixed point of the continuum state evolution, then  $\nabla E_{\mathcal{W}}(\phi)(y) = 0$ , for  $-1 \leq y \leq \ell - 1$ .*

*Proof.* We have  $\phi = \mathcal{F}_{\mathcal{W}}''(\psi)$  and  $\psi = \mathcal{F}_{\mathcal{W}}'(\phi)$ , whereby for  $-1 \leq y \leq \ell - 1$ ,

$$\begin{aligned} \phi(y) &= \sigma^2 + \frac{1}{\delta} \int \mathcal{W}(y-x) \psi(x) dx \\ &= \sigma^2 + \frac{1}{\delta} \left( \int_{x \leq 0} \mathcal{W}(y-x) dx \right) \text{mmse} \left( \frac{L_0}{2\sigma^2} \right) + \\ & \quad \frac{1}{\delta} \int_0^\ell \mathcal{W}(y-x) \text{mmse} \left( \int_{-1}^{\ell+1} \mathcal{W}(x-z) \phi(z)^{-1} dz \right) dx \\ &= \varsigma^2(y) + \frac{1}{\delta} \int_0^\ell \mathcal{W}(y-x) \text{mmse} \left( \int_{-1}^{\ell+1} \mathcal{W}(x-z) \phi(z)^{-1} dz \right) dx. \end{aligned} \quad (11.4.12)$$

The result follows immediately from Eq. (11.4.11). □

**Definition 11.4.8.** Define the potential function  $V : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  as follows.

$$V(\phi) = \frac{\delta}{2} \left( \frac{\sigma^2}{\phi} + \log \phi \right) + \mathfrak{l}(\phi^{-1}). \quad (11.4.13)$$

As we will see later, the analysis of the continuum state evolution involves a decomposition of the free energy functional into three terms and a careful treatment of each term separately. The definition of the potential function  $V$  is motivated by that decomposition.

Using Eq. (11.4.8), we have for  $\phi \rightarrow 0$ ,

$$\begin{aligned} V(\phi) &= \frac{\delta}{2} \left( \frac{\sigma^2}{\phi} + \log \phi \right) + \frac{1}{2} \bar{d}(p_X) \log(\phi^{-1})(1 + o(1)) \\ &= \frac{\delta \sigma^2}{2\phi} + \frac{1}{2} \left\{ \delta - \bar{d}(p_X)(1 + o(1)) \right\} \log(\phi). \end{aligned} \quad (11.4.14)$$

Define

$$\phi^* = \sigma^2 + \frac{1}{\delta} \text{mmse} \left( \frac{L_0}{2\sigma^2} \right). \quad (11.4.15)$$

Notice that  $\sigma^2 < \phi^* \leq (1 + 2/(\delta L_0))\sigma^2 < 2\sigma^2$ , given that  $\delta L_0 > 3$ . The following proposition upper bounds  $V(\phi^*)$  and its proof is deferred to Appendix D.2.

**Proposition 11.4.9.** *There exists  $\sigma_2 > 0$ , such that, for  $\sigma \in (0, \sigma_2]$ , we have*

$$V(\phi^*) \leq \frac{\delta}{2} + \frac{\delta - \bar{d}(p_X)}{4} \log(2\sigma^2). \quad (11.4.16)$$

Now, we write the energy functional in terms of the potential function.

$$\mathbf{E}_{\mathcal{W}}(\phi) = \int_{-1}^{\ell-1} V(\phi(x)) \, dx + \frac{\delta}{2} \int_{-1}^{\ell-1} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \, dx + \tilde{\mathbf{E}}_{\mathcal{W}}(\phi), \quad (11.4.17)$$

with,

$$\tilde{\mathbf{E}}_{\mathcal{W}}(\phi) = \int_0^\ell \left\{ \mathfrak{l}(\mathcal{W} * \phi(y)^{-1}) - \mathfrak{l}(\phi(y-1)^{-1}) \right\} dy. \quad (11.4.18)$$

## 11.4.2 Analysis of the continuum state evolution

Now we are ready to study the fixed points of the continuum state evolution.

**Lemma 11.4.10.** *Let  $\delta > 0$ , and  $p_X$  be a probability measure on the real line with  $\delta > \bar{d}(p_X)$ . For any  $\kappa > 0$ , there exist  $\ell_0, \sigma_0^2 = \sigma_0(\kappa, \delta, p_X)^2$ , such that, for any  $\ell > \ell_0$  and  $\sigma \in (0, \sigma_0]$ , and any fixed point of continuum state evolution,  $\{\phi, \psi\}$ , with  $\psi$  and  $\phi$  nondecreasing Lipschitz functions and  $\psi(x) \geq \text{mmse}(L_0/(2\sigma^2))$ , the following holds.*

$$\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx \leq \kappa\ell. \quad (11.4.19)$$

*Proof.* The claim is trivial for  $\kappa \geq \Phi_M$ , since  $\phi(x) \leq \Phi_M$ . Fix  $\kappa < \Phi_M$ , and choose  $\sigma_1$ , such that  $\phi^* < \kappa/2$ , for  $\sigma \in (0, \sigma_1]$ . Since  $\phi$  is a fixed point of continuum state evolution, we have  $\nabla E_{\mathcal{W}}(\phi) = 0$ , on the interval  $[-1, \ell - 1]$  by Corollary 11.4.7. Now, assume that  $\int_{-1}^{\ell-1} |\phi(x) - \phi^*| > \kappa\ell$ . We introduce an infinitesimal perturbation of  $\phi$  that decreases the energy in the first order; this contradicts the fact  $\nabla E_{\mathcal{W}}(\phi) = 0$  on the interval  $[-1, \ell - 1]$ .

**Claim 11.4.11.** *For each fixed point of continuum state evolution that satisfies the hypothesis of Lemma 11.4.10, the following holds. For any  $K > 0$ , there exists  $\ell_0$ , such that, for  $\ell > \ell_0$  there exist  $x_1 < x_2 \in [0, \ell - 1)$ , with  $x_2 - x_1 = K$  and  $\kappa/2 + \phi^* < \phi(x)$ , for  $x \in [x_1, x_2]$ .*

Claim 11.4.11 is proved in Appendix D.3.

Fix  $K > 2$  and let  $x_0 = (x_1 + x_2)/2$ . Thus,  $x_0 \geq 1$ . For  $a \in (0, 1]$ , define

$$\phi_a(x) = \begin{cases} \phi(x), & \text{for } x_2 \leq x, \\ \phi\left(\frac{x_2 - x_0}{x_2 - x_0 - a} x - \frac{ax_2}{x_2 - x_0 - a}\right), & \text{for } x \in [x_0 + a, x_2), \\ \phi(x - a), & \text{for } x \in [-1 + a, x_0 + a), \\ \phi^*, & \text{for } x \in [-1, -1 + a). \end{cases} \quad (11.4.20)$$

See Fig. 11.4.1 for an illustration. (Note that from Eq. (11.4.2),  $\phi(-1) = \phi^*$ ). In the following, we bound the difference of the free energies of functions  $\phi$  and  $\phi_a$ .

**Proposition 11.4.12.** *For each fixed point of continuum state evolution, satisfying the hypothesis of Lemma 11.4.10, there exists a constant  $C(K)$ , such that*

$$\int_{-1}^{\ell-1} \left\{ \frac{\zeta^2(x) - \sigma^2}{\phi_a(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx \leq C(K)a.$$

We refer to Appendix D.4 for the proof of Proposition 11.4.12.

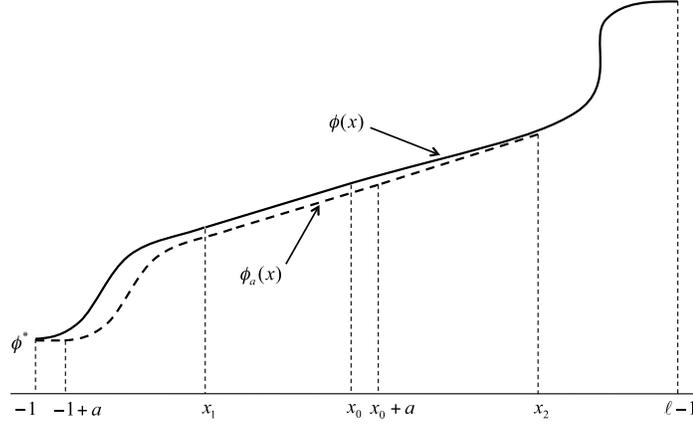


Figure 11.4.1: An illustration of function  $\phi(x)$  and its perturbation  $\phi_a(x)$ .

**Proposition 11.4.13.** *For each fixed point of continuum state evolution, satisfying the hypothesis of Lemma 11.4.10, there exists a constant  $C(\kappa, K)$ , such that,*

$$\tilde{E}_{\mathcal{W}}(\phi_a) - \tilde{E}_{\mathcal{W}}(\phi) \leq C(\kappa, K)a.$$

Proof of Proposition 11.4.13 is deferred to Appendix D.5.

Using Eq. (11.4.17) and Proposition 11.4.13, we have

$$E_{\mathcal{W}}(\phi_a) - E_{\mathcal{W}}(\phi) \leq \int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx + C(\kappa, K)a, \quad (11.4.21)$$

where the constants  $(\delta/2)C(K)$  and  $C(\kappa, K)$  are absorbed in  $C(\kappa, K)$ .

We proceed by proving the following proposition. Its proof is deferred to Appendix D.6.

**Proposition 11.4.14.** *For any  $C = C(\kappa, K)$ , there exists  $\sigma_0$ , such that for  $\sigma \in (0, \sigma_0]$  the following holds.*

$$\int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx < -2C(\kappa, K)a. \quad (11.4.22)$$

Fix  $C(\kappa, K) > 0$ . As a result of Eq. (11.4.21) and Proposition 11.4.14,

$$\begin{aligned} \mathbf{E}_{\mathcal{W}}(\phi_a) - \mathbf{E}_{\mathcal{W}}(\phi) &< \int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx + C(\kappa, K)a \\ &\leq -C(\kappa, K)a. \end{aligned} \quad (11.4.23)$$

Since  $\phi$  is a Lipschitz function by assumption, it is easy to see that  $\|\phi_a - \phi\|_2 \leq C a$ , for some constant  $C$ . By Taylor expansion of the free energy functional around function  $\phi$ , we have

$$\begin{aligned} \langle \nabla \mathbf{E}_{\mathcal{W}}(\phi), \phi_a - \phi \rangle &= \mathbf{E}_{\mathcal{W}}(\phi_a) - \mathbf{E}_{\mathcal{W}}(\phi) + o(\|\phi_a - \phi\|_2) \\ &\leq -C(\kappa, K)a + o(a). \end{aligned} \quad (11.4.24)$$

However, since  $\{\phi, \psi\}$  is a fixed point of the continuum state evolution, we have  $\nabla \mathbf{E}_{\mathcal{W}}(\phi) = 0$  on the interval  $[-1, \ell - 1]$  (cf. Corollary 11.4.7). Also,  $\phi_a - \phi$  is zero out of  $[-1, \ell - 1]$ . Therefore,  $\langle \nabla \mathbf{E}_{\mathcal{W}}(\phi), \phi_a - \phi \rangle = 0$ , which leads to a contradiction in Eq (11.4.24). This implies that our first assumption  $\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx > \kappa \ell$  is false. The result follows.  $\square$

### 11.4.3 Analysis of the continuum state evolution: robust reconstruction

Next lemma pertains to the robust reconstruction of the signal. Prior to stating the lemma, we need to establish some definitions. Due to technical reasons in the proof, we consider an alternative decomposition of  $\mathbf{E}_{\mathcal{W}}(\phi)$  to Eq. (11.4.17).

Define the potential function  $V_{\text{rob}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  as follows.

$$V_{\text{rob}}(\phi) = \frac{\delta}{2} \left( \frac{\sigma^2}{\phi} + \log \phi \right), \quad (11.4.25)$$

and decompose the Energy functional as:

$$\mathbf{E}_{\mathcal{W}}(\phi) = \int_{-1}^{\ell-1} V_{\text{rob}}(\phi(x)) dx + \frac{\delta}{2} \int_{-1}^{\ell-1} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx + \tilde{\mathbf{E}}_{\mathcal{W}, \text{rob}}(\phi), \quad (11.4.26)$$

with,

$$\tilde{\mathbf{E}}_{\mathcal{W}, \text{rob}}(\phi) = \int_0^\ell \mathbf{l}(\mathcal{W} * \phi(y)^{-1}) dy. \quad (11.4.27)$$

**Lemma 11.4.15.** *Let  $\delta > 0$ , and  $p_X$  be a probability measure on the real line with  $\delta >$*

$\bar{D}(p_X)$ . For any  $0 < \alpha < 1$ , there exist  $\ell_0 = \ell_0(\alpha)$ ,  $\sigma_0^2 = \sigma_0(p_X, \delta, \alpha)^2$ , such that, for any  $\ell > \ell_0$  and  $\sigma \in (0, \sigma_0]$ , and for any fixed point of continuum state evolution,  $\{\phi, \psi\}$ , with  $\psi$  and  $\phi$  nondecreasing Lipschitz functions and  $\psi(x) \geq \text{mmse}(L_0/(2\sigma^2))$ , the following holds.

$$\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx \leq C\sigma^2\ell, \quad (11.4.28)$$

with  $C = \frac{2\delta}{(1-\alpha)(\delta-\bar{D}(p_X))}$ .

*Proof.* Suppose  $\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx > C\sigma^2\ell$ , for the given  $C$ . Similar to the proof of Lemma 11.4.10, we obtain an infinitesimal perturbation of  $\phi$  that decreases the free energy in the first order, contradicting the fact  $\nabla E_{\mathcal{W}}(\phi) = 0$  on the interval  $[-1, \ell - 1]$ .

By definition of upper MMSE dimension (Eq. (7.1.12)), for any  $\varepsilon > 0$ , there exists  $\phi_1$ , such that, for  $\phi \in [0, \phi_1]$ ,

$$\text{mmse}(\phi^{-1}) \leq (\bar{D}(p_X) + \varepsilon)\phi. \quad (11.4.29)$$

Henceforth, fix  $\varepsilon$  and  $\phi_1$ .

**Claim 11.4.16.** For each fixed point of continuum state evolution that satisfies the hypothesis of Lemma 11.4.15, the following holds. For any  $K > 0$ ,  $0 < \alpha < 1$ , there exist  $\ell_0 = \ell_0(\alpha)$  and  $\sigma_0 = \sigma_0(\varepsilon, \alpha, p_X, \delta)$ , such that for  $\ell > \ell_0$  and  $\sigma \in (0, \sigma_0]$ , there exist  $x_1 < x_2 \in [0, \ell - 1]$ , with  $x_2 - x_1 = K$  and  $C\sigma^2(1 - \alpha) \leq \phi(x) \leq \phi_1$ , for  $x \in [x_1, x_2]$ .

Claim 11.4.16 is proved in Appendix D.7. For positive values of  $a$ , define

$$\phi_a(x) = \begin{cases} \phi(x), & \text{for } x \leq x_1, x_2 \leq x, \\ (1 - a)\phi(x) & \text{for } x \in (x_1, x_2). \end{cases} \quad (11.4.30)$$

Our aim is to show that  $E_{\mathcal{W}}(\phi_a) - E_{\mathcal{W}}(\phi) \leq -c a$ , for some constant  $c > 0$ .

Invoking Eq. (11.4.17), we have

$$\begin{aligned} E_{\mathcal{W}}(\phi_a) - E_{\mathcal{W}}(\phi) &= \int_{-1}^{\ell-1} \{V_{\text{rob}}\phi_a(x) - V_{\text{rob}}(\phi(x))\} dx \\ &\quad + \frac{\delta}{2} \int_{-1}^{\ell-1} (\varsigma^2(x) - \sigma^2) \left( \frac{1}{\phi_a(x)} - \frac{1}{\phi(x)} \right) dx + \tilde{E}_{\mathcal{W}, \text{rob}}(\phi_a) - \tilde{E}_{\mathcal{W}, \text{rob}}(\phi). \end{aligned} \quad (11.4.31)$$

The following proposition bounds each term on the right hand side separately.

**Proposition 11.4.17.** *For the function  $\phi(x)$  and its perturbation  $\phi_a(x)$ , we have*

$$\int_{-1}^{\ell-1} \{V_{\text{rob}}(\phi_a(x)) - V_{\text{rob}}(\phi(x))\} dx \leq \frac{\delta}{2} K \log(1-a) + K \frac{\delta a}{2C(1-\alpha)(1-a)}, \quad (11.4.32)$$

$$\int_{-1}^{\ell-1} (\varsigma^2(x) - \sigma^2) \left( \frac{1}{\phi_a(x)} - \frac{1}{\phi(x)} \right) dx \leq K \frac{a}{C(1-\alpha)(1-a)}, \quad (11.4.33)$$

$$\tilde{\mathbb{E}}_{\mathcal{W}, \text{rob}}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W}, \text{rob}}(\phi) \leq -\frac{\overline{D}(p_X) + \varepsilon}{2} (K+2) \log(1-a). \quad (11.4.34)$$

We refer to Appendix D.8 for the proof of Proposition 11.4.17.

Combining the bounds given by Proposition 11.4.17, we obtain

$$\begin{aligned} \mathbb{E}_{\mathcal{W}}(\phi_a) - \mathbb{E}_{\mathcal{W}}(\phi) &\leq \\ &\frac{K}{2} \log(1-a) \left\{ \delta - (\overline{D}(p_X) + \varepsilon) \left(1 + \frac{2}{K}\right) \right\} + K \frac{\delta a}{C(1-\alpha)(1-a)}. \end{aligned} \quad (11.4.35)$$

Since  $\delta > \overline{D}(p_X)$  by our assumption, and  $C = \frac{2\delta}{(1-\alpha)(\delta - \overline{D}(p_X))}$ , there exist  $\varepsilon, a$  small enough and  $K$  large enough, such that

$$c = \delta - (\overline{D}(p_X) + \varepsilon) \left(1 + \frac{2}{K}\right) - \frac{2\delta}{C(1-\alpha)(1-a)} > 0.$$

Using Eq. (11.4.35), we get

$$\mathbb{E}_{\mathcal{W}}(\phi_a) - \mathbb{E}_{\mathcal{W}}(\phi) \leq -\frac{cK}{2} a. \quad (11.4.36)$$

By an argument analogous to the one in the proof of Lemma 11.4.10, this is in contradiction with  $\nabla \mathbb{E}_{\mathcal{W}}(\phi) = 0$ . The result follows.  $\square$

## 11.5 Proof of Lemma 10.0.2

By Lemma 11.3.2,  $\phi_a(t) \leq \phi_a^{\text{mod}}(t-t_0)$ , for  $a \in \mathbb{R}_0 \cong \{\rho^{-1}, \dots, L-1+\rho^{-1}\}$  and  $t \geq t_1(L_0, \delta)$ . Therefore, we only need to prove the claim for the modified state evolution. The idea of the proof is as follows. In the previous section, we analyzed the continuum state evolution and showed that at the fixed point, the function  $\phi(x)$  is close to the constant  $\phi^*$ . Also, in Lemma 11.4.2, we proved that the modified state evolution is essentially approximated by

the continuum state evolution as  $\rho \rightarrow 0$ . Combining these results implies the thesis.

*Proof (Part(a)).* By monotonicity of continuum state evolution (cf. Corollary 11.4.3),  $\lim_{t \rightarrow \infty} \phi(x; t) = \phi(x)$  exists. Further, by continuity of state evolution recursions,  $\phi(x)$  is a fixed point. Finally,  $\phi(x)$  is a nondecreasing Lipschitz function (cf. Corollary 11.4.4). Using Lemma 11.4.10 in conjunction with the Dominated Convergence theorem, we have, for any  $\varepsilon > 0$

$$\lim_{t \rightarrow \infty} \frac{1}{\ell} \int_{-1}^{\ell-1} |\phi(x; t) - \phi^*| dx \leq \frac{\varepsilon}{4}, \quad (11.5.1)$$

for  $\sigma \in (0, \sigma_0^2]$  and  $\ell > \ell_0$ . Therefore, there exists  $t_2 > 0$  such that  $\frac{1}{\ell} \int_{-1}^{\ell-1} |\phi(x; t_2) - \phi^*| dx \leq \varepsilon/2$ . Moreover, for any  $t \geq 0$ ,

$$\begin{aligned} \frac{1}{\ell} \int_{-1}^{\ell-1} |\phi(x; t) - \phi^*| dx &= \lim_{\rho \rightarrow 0} \frac{\rho}{\ell} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi(\rho a; t) - \phi^*| \\ &= \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi(\rho a; t) - \phi^*|. \end{aligned} \quad (11.5.2)$$

By triangle inequality, for any  $t \geq 0$ ,

$$\begin{aligned} \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi_a(t) - \phi^*| &\leq \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi_a(t) - \phi(\rho a; t)| + \\ &\quad \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} |\phi(\rho a; t) - \phi^*| \\ &= \frac{1}{\ell} \int_{-1}^{\ell-1} |\phi(x; t) - \phi^*| dx, \end{aligned} \quad (11.5.3)$$

where the last step follows from Lemma 11.4.2 and Eq. (11.5.2).

Since the sequence  $\{\phi(t)\}$  is monotone decreasing in  $t$ , we have

$$\begin{aligned}
 \lim_{\rho \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) &\leq \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t_2) \\
 &\leq \lim_{\rho \rightarrow 0} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} (|\phi_a(t_2) - \phi^*| + \phi^*) \\
 &\leq \frac{1}{\ell} \int_{-1}^{\ell-1} |\phi(x; t_2) - \phi^*| dx + \phi^* \\
 &\leq \frac{\varepsilon}{2} + \phi^*.
 \end{aligned} \tag{11.5.4}$$

Finally,

$$\begin{aligned}
 \lim_{t \rightarrow \infty} \sum_{a=-\rho^{-1}}^{L+\rho^{-1}-1} \phi_a(t) &\leq \frac{2\rho^{-1}}{L} \Phi_M + \frac{\varepsilon}{2} + \phi^* \\
 &\leq \frac{2\rho^{-1}}{L_*} \Phi_M + \frac{\varepsilon}{2} + 2\sigma_0.
 \end{aligned} \tag{11.5.5}$$

Clearly, by choosing  $L_*$  large enough and  $\sigma_0$  sufficiently small, we can ensure that the right hand side of Eq. (11.5.5) is less than  $\varepsilon$ .  $\square$

*Proof (Part(b)).* Consider the following two cases.

- $\sigma \leq \sigma_0$ : In this case, proceeding along the same lines as the proof of Part (a), and using Lemma 11.4.15 in lieu of Lemma 11.4.10, we have

$$\lim_{t \rightarrow \infty} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) \leq C\sigma^2 + \phi^* \leq \left( \frac{2\delta}{(1-\alpha)(\delta - \overline{D}(p_X))} + 1 + \frac{2}{\delta L_0} \right) \sigma^2. \tag{11.5.6}$$

- $\sigma > \sigma_0$ : Since  $\phi_a(t) \leq \sigma^2 + (1/\delta)\text{Var}(X)$  for any  $t > 0$ , we have

$$\lim_{t \rightarrow \infty} \frac{1}{L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) \leq \sigma^2 + \frac{1}{\delta} \text{Var}(X). \tag{11.5.7}$$

Choosing

$$C = \max \left\{ \frac{2\delta}{(1-\alpha)(\delta - \overline{D}(p_X))} + 1 + \frac{2}{\delta L_0}, 1 + \frac{\text{Var}(X)}{\delta \sigma_0^2} \right\},$$

proves the claim in both cases.

Finally, in the asymptotic case where  $\ell = L\rho \rightarrow \infty$ ,  $\rho \rightarrow 0$ ,  $L_0 \rightarrow \infty$ , we have  $\alpha \rightarrow 0$  and using Eq. (11.5.6), we get

$$\limsup_{\sigma \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{\sigma^2 L} \sum_{a=-\rho^{-1}}^{L-\rho^{-1}-1} \phi_a(t) \leq \frac{3\delta - \overline{D}(p_X)}{\delta - \overline{D}(p_X)}.$$

□

## Chapter 12

# Gabor Transform & Spatial Coupling

As discussed in PartII, using spatial coupling and (approximate) message passing, our approach allows successful compressed sensing recovery from a number of measurements that achieves the information-theoretic limit. However, this approach presents in fact several unrealistic features. In particular, the entries of  $A$  are independent Gaussian entries with zero mean and suitably chosen variances. It is obviously difficult to implement such a measurement matrix through a physical sampling mechanism. This chapter aims at showing that the spatial coupling phenomenon is significantly more robust and general than suggested by the constructions of [82, 41].

We study the problem of sampling a random signal with sparse support in frequency domain. Shannon famously considered a scheme that *instantaneously* samples the signal at equispaced times. He proved that the signal can be reconstructed as long as the sampling rate exceeds twice the bandwidth (Nyquist rate). Candès, Romberg, Tao introduced a scheme that acquires *instantaneous* samples of the signal at random times. They proved that the signal can be uniquely and efficiently reconstructed, provided the sampling rate exceeds the frequency support of the signal, times logarithmic factors.

In this chapter, we consider a probabilistic model for the signal, and a sampling scheme inspired by the idea of spatial coupling in coding theory. Namely, we propose to acquire *non-instantaneous* samples at random times. Mathematically, this is implemented by acquiring a small random subset of Gabor coefficients. We show empirically that this scheme achieves correct reconstruction as soon as the sampling rate exceeds the frequency support of the

signal, thus reaching the information theoretic limit.

## 12.1 Definitions

For the sake of simplicity, we consider a discrete-time model (analogous to the one of [23]) and denote signals in time domain as  $x \in \mathbb{C}^n$ ,  $x = (x(t))_{1 \leq t \leq n} = (x(1), \dots, x(n))^T$ . Their discrete Fourier transform is denoted by  $\hat{x} \in \mathbb{C}^n$ ,  $\hat{x} = (\hat{x}(\omega))_{\omega \in \Omega_n}$ , where  $\Omega_n = \{\omega = 2\pi k/n : k \in \{0, 1, \dots, n-1\}\}$ . The Fourier transform  $\hat{x} = (\mathbf{F}x)$  is given by

$$\hat{x}(\omega) = \langle b_\omega, x \rangle = \sum_{t=1}^n \overline{b_\omega(t)} x(t), \quad b_\omega(t) \equiv \frac{1}{\sqrt{n}} e^{i\omega t}. \quad (12.1.1)$$

Here  $\langle \cdot, \cdot \rangle$  denotes the standard scalar product on  $\mathbb{C}^n$ . Also, for a complex variable  $z$ ,  $\bar{z}$  is the complex conjugate of  $z$ . Notice that  $(b_\omega)_{\omega \in \Omega_n}$  is an orthonormal basis of  $\mathbb{C}^n$ . This implies Parseval's identity  $\langle \hat{x}_1, \hat{x}_2 \rangle = \langle x_1, x_2 \rangle$ . In addition, the inverse transform is given by

$$x(t) = \sum_{\omega \in \Omega_n} \hat{x}(\omega) b_\omega(t) = \frac{1}{\sqrt{n}} \sum_{\omega \in \Omega_n} \hat{x}(\omega) e^{i\omega t}. \quad (12.1.2)$$

We will denote by  $T_n = \{1, \dots, n\}$  the time domain, and will consider signals that are sparse in the Fourier domain.

A sampling mechanism is defined by a measurement matrix  $A \in \mathbb{R}^{m \times n}$ . Measurement vector  $y = (y(1), \dots, y(m))^T \in \mathbb{R}^m$  is given by

$$y = Ax + w \equiv y_0 + w, \quad (12.1.3)$$

where  $w$  is a noise vector with variance  $\sigma^2$ , and  $y_0$  is the vector of ideal (noiseless) measurements. In other words,  $y(i) = \langle a_i, x \rangle$  where we let  $a_1^*, \dots, a_m^*$  be the rows of  $A$ . *Instantaneous* sampling corresponds to vectors  $a_i$  that are canonical base vectors.

Measurements can also be given in terms of the Fourier transform of the signal:

$$y = A_{\mathbf{F}} \hat{x} + w, \quad A_{\mathbf{F}} = A\mathbf{F}^*. \quad (12.1.4)$$

The rows of  $A_{\mathbf{F}}$  are denoted by  $\hat{a}_1^*, \dots, \hat{a}_m^*$ , and obviously  $\hat{a}_i = \mathbf{F}a_i$ . Here and below, for a matrix  $M$ ,  $M^*$  is the hermitian adjoint of  $M$ , i.e.  $M_{ij}^* = \overline{M_{ji}}$ .

## 12.2 Information theory model

In [23], Candès, Romberg, Tao studied a randomized scheme that samples the signal instantaneously at uniformly random times. Mathematically, this corresponds to choosing the measurement vectors  $a_i$  to be a random subset of the canonical basis in  $\mathbb{C}^n$ . They proved that, with high probability, these measurements allow to reconstruct  $x$  uniquely and efficiently, provided  $m \geq C|S|\log n$ , where  $S = \{\omega \in \Omega : \hat{x}(\omega) \neq 0\}$  is the frequency support of the signal.

In this chapter, we consider a probabilistic model for the signal  $\hat{x}$ , namely we assume that the components  $\hat{x}(\omega)$ ,  $\omega \in \Omega$  are i.i.d. with  $\mathbb{P}\{\hat{x}(\omega) \neq 0\} \leq \varepsilon$  and  $\mathbb{E}\{|\hat{x}(\omega)|^2\} \leq C < \infty$ . The distribution of  $\hat{x}(\omega)$  is assumed to be known. Indeed, information theoretic thinking has led to impressive progress in digital communication, as demonstrated by the development of modern iterative codes [115]. More broadly, probabilistic models can lead to better understanding of limits and assumptions in relevant applications to digital communication and sampling theory.

## 12.3 Contributions

The goal of present chapter is to show that the spatial coupling phenomenon is –in the present context– significantly more general than suggested by the constructions of [82, 41]. Unfortunately, a rigorous analysis of message passing algorithms is beyond reach for sensing matrices with dependent or deterministic entries. We thus introduce an ensemble of sensing matrices, and show numerically that, under approximate message passing (AMP) reconstruction, they allow recovery at undersampling rates close to the information dimension. Similar simulations were already presented by Krzakala et al. [82] in the case of matrices with independent entries.

Our matrix ensemble can be thought of as a modification of the one in [23] for implementing spatial coupling. As mentioned above, [23] suggests to sample the signal pointwise (instantaneously) in time. In the Fourier domain (in which the signal is sparse) this corresponds to taking measurements that probe all frequencies with the same weight. In other words,  $A_F$  is not band-diagonal as required in spatial coupling. Our solution is to ‘smear out’ the samples: instead of measuring  $x(t_*)$ , we modulate the signal with a wave of frequency  $\omega_*$ , and integrate it over a window of size  $W^{-1}$  around  $t_*$ . In Fourier space, this

corresponds to integrating over frequencies within a window  $W$  around  $\omega_*$ . Each measurement corresponds to a different time-frequency pair  $(t_*, \omega_*)$ . While there are many possible implementations of this idea, the Gabor transform offers an analytically tractable avenue. Our method can be thought of as a subsampling of a discretized Gabor transform of the signal.

In [93], Gabor frames have also been used to exploit the sparsity of signals in time and enable sampling multipulse signals at sub-Nyquist rates.

## 12.4 Sampling scheme

### 12.4.1 Constructing the sensing matrix

The sensing matrix  $A$  is drawn from a random ensemble denoted by  $\mathcal{M}(n, m_1, L, \ell, \xi, \delta)$ . Here  $n, m_1, L, \ell$  are integers and  $\xi, \delta \in (0, 1)$ . The rows of  $A$  are partitioned as follows:

$$\mathbf{R} = \left\{ \bigcup_{k=1}^{m_1} \mathbf{R}_k \right\} \cup \mathbf{R}_0, \quad (12.4.1)$$

where  $|\mathbf{R}_k| = L$ , and  $|\mathbf{R}_0| = \lfloor n\delta \rfloor$ . Hence,  $m = m_1 L + \lfloor n\delta \rfloor$ . Notice that  $m/n = (m_1 L + \lfloor n\delta \rfloor)/n$ . Since we will take  $n$  much larger than  $m_1 L$ , the undersampling ratio  $m/n$  will be arbitrary close to  $\delta$ . Indeed, with an abuse of language, we will refer to  $\delta$  as the undersampling ratio.

We construct the sensing matrix  $A$  as follows:

1. For each  $k \in \{1, \dots, m_1\}$ , and each  $r \in \mathbf{R}_k$ ,  $a_r = b_{2\pi k/n}$ .
2. The rows  $\{a_r\}_{r \in \mathbf{R}_0}$  are defined as

$$a_r(t) = a(t; t_r, \omega_r), \quad (12.4.2)$$

where  $\{t_r\}_{r \in \mathbf{R}_0}$  are independent and uniformly random in  $T_n$ , and  $\{\omega_r\}_{r \in \mathbf{R}_0}$  are equispaced in  $\Omega_n$ . Finally, for  $t_* \in T_n$ , and  $\omega_* \in \Omega_n$ , we define

$$a(t; t_*, \omega_*) = \frac{1}{C_\ell} e^{i\omega_* t} P_{\xi, \ell}(t_*, t), \quad C_\ell = \left\{ \sum_{t \in T_n} P_{\xi, \ell}(t_*, t)^2 \right\}^{1/2}.$$

Here  $P_{\xi, \ell}(t_*, t)$  is the probability that a random walk on the circle with  $n$  sites  $\{1, \dots, n\}$  starting at time 0 at site  $t_*$  is found at time  $\ell$  at site  $t$ . The random

walk is lazy, i.e. it stays on the same position with probability  $1 - \xi \in (0, 1)$  and moves with probability  $\xi$  choosing either of the adjacent sites with equal probability.

Notice that the probabilities  $P_{\xi, \ell}(t_*, t)$  satisfy the recursion

$$\begin{aligned} P_{\xi, \ell+1}(t_*, t) &= (1 - \xi) P_{\xi, \ell}(t_*, t) + \frac{\xi}{2} P_{\xi, \ell}(t_* - 1, t) \\ &\quad + \frac{\xi}{2} P_{\xi, \ell}(t_* + 1, t), \quad P_{\xi, 0}(t_*, t) = \mathbb{I}(t = t_*), \end{aligned} \quad (12.4.3)$$

where sums on  $T_n$  are understood to be performed modulo  $n$ . We can think of  $P_{\xi, \ell}$  as a discretization of a Gaussian kernel. Indeed, for  $1 \ll \ell \ll n^2$  we have, by the local central limit theorem,

$$P_{\xi, \ell}(t_*, t) \approx \frac{1}{(2\pi\xi\ell)^{1/2}} \exp\left\{-\frac{(t - t_*)^2}{2\xi\ell}\right\}. \quad (12.4.4)$$

and hence  $C_\ell \approx (4\pi\xi\ell)^{-1/4}$ .

The above completely define the sensing process. For the signal reconstruction we will use AMP in the Fourier domain, i.e. we will try to reconstruct  $\hat{x}$  from  $y = A_F \hat{x} + w$ . It is therefore convenient to give explicit expressions for the measurement matrix in this domain.

1. For each  $k \in \{1, \dots, m_1\}$ , and each  $r \in \mathbf{R}_k$ , we have  $\hat{a}_r = e_k$ , where  $e_k \in \mathbb{R}^n$  refers to the  $k^{\text{th}}$  standard basis element, e.g.,  $e_1 = (1, 0, 0, \dots, 0)$ . These rows are used to sense the extreme of the spectrum frequencies.
2. For  $r \in \mathbf{R}_0$ , we have  $\hat{a}_r(\omega) = \hat{a}(\omega; t_*, \omega_*)$ , where

$$\hat{a}(\omega; t_*, \omega_*) = \frac{1}{C_\ell \sqrt{n}} e^{-i(\omega - \omega_*)t_*} (1 - \xi + \xi \cos(\omega - \omega_*))^\ell.$$

Again, to get some insight, we consider the asymptotic behavior for  $1 \ll \ell \ll n^2$ . It is easy to check that  $\hat{a}$  is significantly different from 0 only if  $\omega - \omega_* = O(\ell^{-1/2})$  and

$$\hat{a}(\omega; t_*, \omega_*) \approx \frac{1}{C_\ell \sqrt{n}} \exp\left\{-i(\omega - \omega_*)t_* - \frac{1}{2}\xi\ell(\omega - \omega_*)^2\right\}.$$

Hence the measurement  $y_i$  depends on the signal Fourier transform only within a window of size  $W = O(\ell^{-1/2})$ , with  $1/n \ll W \ll 1$ . As claimed in the introduction, we recognize that the rows of  $A$  are indeed (discretized) Gabor filters. Also it is easy to check that  $A_F$  is roughly band-diagonal with width  $W$ .

### 12.4.2 Algorithm

We use a generalization of the AMP algorithm for spatially-coupled sensing matrices (cf. Chapter 8) to the complex setting. Assume that the empirical law of the entries of  $\hat{x}^{(n)}$  converges weakly to a limit  $p_{\hat{X}}$ , with bounded second moment. The algorithm proceeds by the following iteration (initialized with  $\hat{x}_i^1 = \mathbb{E}\{\hat{X}\}$  for all  $i \in [n]$ ). For  $\hat{x}^t \in \mathbb{C}^n$ ,  $r^t \in \mathbb{C}^m$ ,

$$\begin{aligned}\hat{x}^{t+1} &= \eta_t(\hat{x}^t + (Q^t \odot A_F)^* r^t), \\ r^t &= y - A_F \hat{x}^t + \mathbf{b}^t \odot r^{t-1} + \mathbf{d}^t \odot \bar{r}^{t-1}.\end{aligned}\tag{12.4.5}$$

Here  $\eta_t(v) = (\eta_{t,1}(v_1), \dots, \eta_{t,n}(v_n))$ , where  $\eta_{t,i} : \mathbb{C} \rightarrow \mathbb{C}$  is a scalar denoiser. Similar to Chapter 8, we assume that the prior  $p_{\hat{X}}$  is known and use the posterior expectation denoiser

$$\eta_{t,i}(v_i) = \mathbb{E}\{\hat{X} | \hat{X} + s_i^{-1/2} Z = v_i\}, \quad s_i = \sum_{a \in [m]} W_{ai} \phi_a(t)^{-1},$$

where  $\hat{X} \sim p_{\hat{X}}$  and  $Z \sim \mathbf{N}_{\mathbb{C}}(0, 1)$  is a standard complex normal random variable, independent of  $\hat{X}$ . Also,  $\bar{r}^t$  is the complex conjugate of  $r^t$  and  $\odot$  indicates Hadamard (entrywise) product. The matrix  $Q^t \in \mathbb{R}^{m \times n}$ , and the vector  $\mathbf{b}^t \in \mathbb{R}^m$  are given by

$$Q_{ai}^t = \frac{\phi_a(t)^{-1}}{\sum_{b \in [m]} W_{bi} \phi_b(t)^{-1}},\tag{12.4.6}$$

$$\mathbf{b}_a^t = \sum_{i \in [n]} Q_{ai}^{t-1} W_{ai} \partial \eta_{t-1,i},\tag{12.4.7}$$

$$\mathbf{d}_a^t = \sum_{i \in [n]} Q_{ai}^{t-1} (A_F)_{ai}^2 \bar{\partial} \eta_{t-1,i},\tag{12.4.8}$$

where  $W_{ai} \equiv |(A_F)_{ai}|^2$  and  $\partial \eta_{t,i} \equiv \partial \eta_{t,i}(\hat{x}_i^t + ((Q^t \odot A_F)^* r^t)_i)$ ,  $\bar{\partial} \eta_{t,i} \equiv \bar{\partial} \eta_{t,i}(\hat{x}_i^t + ((Q^t \odot A_F)^* r^t)_i)$ . Throughout,  $\eta_{t,i}(v)$  is viewed as a function of  $v$ ,  $\bar{v}$ , and  $v$ ,  $\bar{v}$  are taken as independent variables in the sense that  $\partial \bar{v} / \partial v = 0$ . Then,  $\partial \eta_{t,i}$  and  $\bar{\partial} \eta_{t,i}$  respectively denote the partial derivative of  $\eta_{t,i}$  with respect to  $v$  and  $\bar{v}$ . Also, derivative is understood here on the complex domain. (These are the principles of Wirtinger's calculus for the complex functions [119]). Finally, the sequence  $\{\phi(t)\}_{t \geq 0}$  is determined by the following state evolution recursion.

$$\phi_a(t+1) = \sigma^2 + \sum_{i \in [n]} W_{ai} \text{mmse} \left( \sum_{b \in [m]} W_{bi} \phi_b(t)^{-1} \right).\tag{12.4.9}$$

Here  $\text{mmse}(\cdot)$  is defined as follows. If  $\widehat{X} \sim p_{\widehat{X}}$  and  $Y = \widehat{X} + s^{-1/2}Z$  for  $Z \sim \mathbf{N}_{\mathbb{C}}(0, 1)$  independent of  $\widehat{X}$ , then

$$\text{mmse}(s) \equiv \frac{1}{2} \mathbb{E}\{|\widehat{X} - \mathbb{E}[\widehat{X}|Y]|^2\}. \quad (12.4.10)$$

## 12.5 Numerical simulations

We consider a Bernoulli-Gaussian distribution  $p_{\widehat{X}} = (1-\varepsilon)\delta_0 + \varepsilon\gamma_{\mathbb{C}}$ , where  $\gamma_{\mathbb{C}}$  is the standard complex gaussian measure and  $\delta_0$  is the delta function at 0. We construct a random signal  $(\widehat{x}(\omega))_{\omega \in \Omega_n}$  by sampling i.i.d. coordinates  $\widehat{x}(\omega) \sim p_{\widehat{X}}$ . We have  $\bar{d}(p_{\widehat{X}}) = \varepsilon$  [143] and

$$\eta_{t,i}(v_i) = \frac{\varepsilon\gamma_{1+s_i^{-1}}(v_i)}{\varepsilon\gamma_{1+s_i^{-1}}(v_i) + (1-\varepsilon)\gamma_{s_i^{-1}}(v_i)} \cdot \frac{1}{1+s_i^{-1}}v_i, \quad (12.5.1)$$

where  $\gamma_{\sigma^2}(z) = 1/(\pi\sigma^2) \exp\{-z\bar{z}/\sigma^2\}$  is the density function of the complex normal distribution with mean zero and variance  $\sigma^2$ .

### 12.5.1 Evolution of the algorithm

Our first set of experiments aims at illustrating the spatial coupling phenomenon and checking the predictions of state evolution. In these experiments we use  $\varepsilon = 0.1$ ,  $\sigma = 0.001$ ,  $\delta = 0.15$ ,  $n = 5000$ ,  $\ell = 800$ ,  $m_1 = 20$ ,  $L = 3$ , and  $\xi = 0.5$ .

State evolution yields an iteration-by-iteration prediction of the AMP performance in the limit of a large number of dimensions. State evolution can be proved rigorously for sensing matrices with independent entries [7, 6]. We also refer to [74] for a derivation which directly studies the case of spatially-coupled matrices. We expect however the prediction to be robust and will check it through numerical simulations for the current sensing matrix  $A_{\mathbb{F}}$ . In particular, state evolution predicts that

$$\mathbb{E}\{|\widehat{x}_i^t(y) - \widehat{x}_i|^2\} \approx \text{mmse}\left(\sum_{a \in \mathbb{R}} W_{a,i} \phi_a^{-1}(t-1)\right). \quad (12.5.2)$$

Figure 12.5.1 shows the evolution of profile  $\phi(t) \in \mathbb{R}^m$ , given by the state evolution recursion (12.4.9). This clearly demonstrates the spatial coupling phenomenon. In our sampling scheme, additional measurements are associated to the first few coordinates of  $\widehat{x}$ , namely,  $\widehat{x}_1, \dots, \widehat{x}_{m_1}$ . This has negligible effect on the undersampling rate ratio because

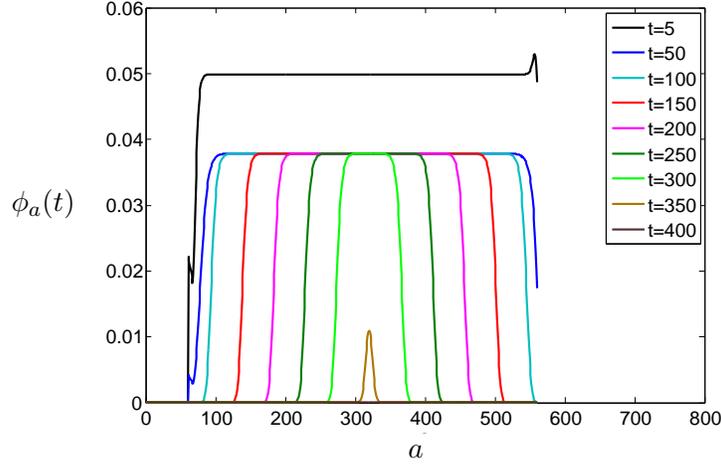


Figure 12.5.1: Profile  $\phi_a(t)$  versus  $a$  for several iteration numbers.

$m_1 L/n \rightarrow 0$ . However, the Fourier components  $\hat{x}_1, \dots, \hat{x}_{m_1}$  are oversampled. This leads to a correct reconstruction of these entries (up to a mean square error of order  $\sigma^2$ ). This is reflected by the fact that  $\phi$  becomes of order  $\sigma^2$  on the first few entries after a few iterations (see  $t = 5$  in the figure). As the iteration proceeds, the contribution of these components is correctly subtracted from all the measurements, and essentially they are removed from the problem. Now, in the resulting problem the first few variables are effectively oversampled and the algorithm reconstructs their values up to a mean square error of  $\sigma^2$ . Correspondingly, the profile  $\phi$  falls to a value of order  $\sigma^2$  in the next few coordinates. As the process is iterated, all the variables are progressively reconstructed and the profile  $\phi$  follows a traveling wave with constant velocity. After a sufficient number of iterations ( $t = 400$  in the figure),  $\phi$  is uniformly of order  $\sigma^2$ .

In order to check the prediction of state evolution, we compare the empirical and the predicted mean square errors

$$\text{MSE}_{\text{AMP}} = \frac{1}{n} \|\hat{x}^t(y) - \hat{x}\|_2^2, \quad (12.5.3)$$

$$\text{MSE}_{\text{SE}} = \frac{1}{n} \sum_{i=1}^n \text{mmse} \left( \sum_{a \in \mathcal{R}} W_{a,i} \phi_a^{-1}(t-1) \right). \quad (12.5.4)$$

The values of  $\text{MSE}_{\text{AMP}}$  and  $\text{MSE}_{\text{SE}}$  versus iteration are depicted in Fig. 12.5.2. (Values of  $\text{MSE}_{\text{AMP}}$  and the bar errors correspond to  $M = 30$  Monte Carlo instances). This verifies that the state evolution provides an iteration-by iteration prediction of AMP performance.

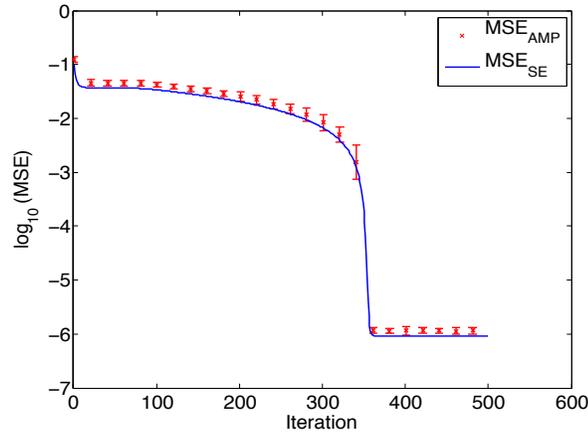


Figure 12.5.2: Comparison of  $\text{MSE}_{\text{AMP}}$  and  $\text{MSE}_{\text{SE}}$  across iteration.

We observe that  $\text{MSE}_{\text{AMP}}$  (and  $\text{MSE}_{\text{SE}}$ ) decreases linearly versus iteration.

## 12.5.2 Phase diagram

In this section, we consider the noiseless compressed sensing setting, and reconstruction through different algorithms and sensing matrix ensembles.

Let  $\mathcal{A}$  be a sensing matrix–reconstruction algorithm scheme. The curve  $\varepsilon \mapsto \delta_{\mathcal{A}}(\varepsilon)$  describes the sparsity–undersampling tradeoff of  $\mathcal{A}$  if the following happens in the large-system limit  $n, m \rightarrow \infty$ , with  $m/n = \delta$ . The scheme  $\mathcal{A}$  does (with high probability) correctly recover the original signal provided  $\delta > \delta_{\mathcal{A}}(\varepsilon)$ , while for  $\delta < \delta_{\mathcal{A}}(\varepsilon)$  the algorithm fails with high probability. We will consider three schemes. For each of them, we consider a set of sparsity parameters  $\varepsilon \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ , and for each value of  $\varepsilon$ , evaluate the empirical phase transition through a logit fit (we omit details, but follow the methodology described in [42]).

### 12.5.2.1 Scheme I

We construct the sensing matrix as described in Section 12.4.1 and for reconstruction, we use the algorithm described in Section 12.4.2. An illustration of the phase transition phenomenon is provided in Fig. 12.5.3. This corresponds to  $\varepsilon = 0.2$  and an estimated phase transition location  $\delta = 0.23$ .

As it is shown in Fig. 12.5.4, our results are consistent with the claim that this scheme

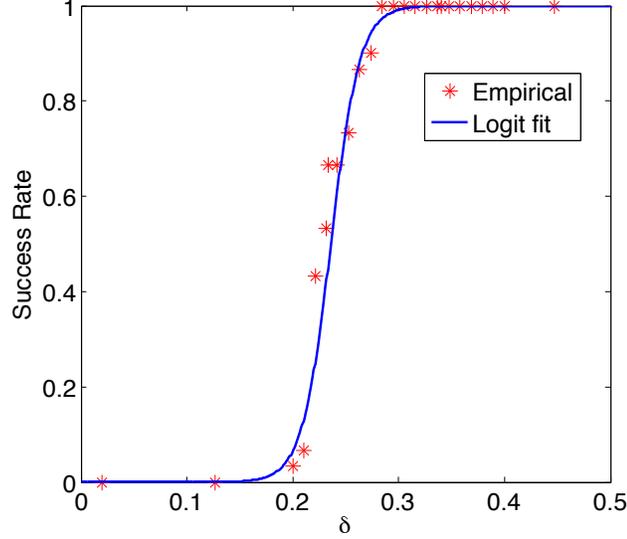


Figure 12.5.3: Phase transition diagram for Scheme I, and  $\varepsilon = 0.2$ .

achieves successful reconstruction at rates close to the information theoretic lower bound  $\delta > \bar{d}(p_{\hat{X}}) = \varepsilon$ . (We indeed expect the gap to decrease further by taking larger values of  $\ell, n$ .)

### 12.5.2.2 Scheme II

The sensing matrix  $A_F$  is obtained by choosing  $m$  rows of the Fourier matrix  $F$  at random. In time domain, this corresponds to sampling at  $m$  random time instants as in [23]. Reconstruction is done via AMP algorithm with posterior expectation as the denoiser  $\eta$ . More specifically, through the following iterative procedure.

$$\begin{aligned} \hat{x}^{t+1} &= \eta_t(\hat{x}^t + A^* r^t), \\ r^t &= y - A\hat{x}^t + \frac{1}{\delta} r^{t-1} \langle \partial \eta_{t-1} \rangle + \frac{1}{\delta} \bar{r}^{t-1} \langle \bar{\partial} \eta_{t-1} \rangle. \end{aligned} \quad (12.5.5)$$

Here  $\eta_t(v) = (\eta_t(v_1), \dots, \eta_t(v_n))$ , where  $\eta_t(v_i) = \mathbb{E}\{\hat{X} | \hat{X} + \phi_t^{1/2} Z = v_i\}$  and  $Z \sim N_{\mathbb{C}}(0, 1)$ . Also  $\partial \eta_{t,i} \equiv \partial \eta_t(\hat{x}_i^t + (A^* r^t)_i)$ ,  $\bar{\partial} \eta_{t,i} \equiv \bar{\partial} \eta_t(\hat{x}_i^t + (A^* r^t)_i)$  and for a vector  $u \in \mathbb{R}^n$ ,  $\langle u \rangle = n^{-1} \sum_{i=1}^n u_i$ .

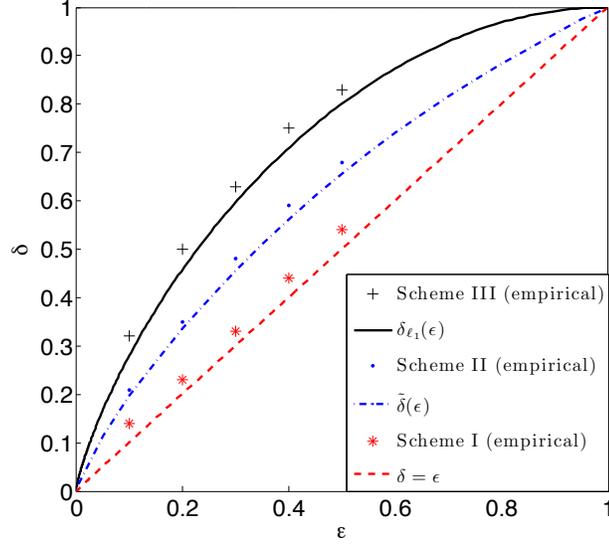


Figure 12.5.4: Phase transition lines for Schemes I, II, III.

The sequence  $\phi_t$  is determined by state evolution

$$\phi_{t+1} = \frac{1}{\delta} \text{mmse}(\phi_t^{-1}), \quad \phi_0 = \text{Var}(\widehat{X})/\delta. \quad (12.5.6)$$

When  $A$  has independent entries  $A_{ij} \sim \mathcal{N}(0, 1/m)$ , state evolution (12.5.6) predicts the performance of the algorithm (12.5.5) [7]. Therefore, the algorithm successfully recovers the original signal with high probability, provided

$$\delta > \tilde{\delta}(\epsilon) = \sup_{s \geq 0} s \cdot \text{mmse}(s). \quad (12.5.7)$$

As shown in Fig. 12.5.4, the empirical phase transition for scheme II is very close to the prediction  $\tilde{\delta}(\epsilon)$ . Note that schemes I, II both use posterior expectation denoising. However, as observed in [82], spatially-coupled matrices in scheme I significantly improve the performances.

### 12.5.2.3 Scheme III

We use the spatially-coupled sensing matrix described in Section 12.4.1, and an AMP algorithm with soft-thresholding denoiser

$$\eta_{ST}(z; \theta) = \left(1 - \frac{\theta}{|z|}\right)_+ z. \quad (12.5.8)$$

The algorithm is defined as in Eq. (12.4.5), except that the soft-thresholding denoiser is used in lieu of the posterior expectation. Formally, let  $\eta_t(v) = (\eta_{t,1}(v_1), \dots, \eta_{t,n}(v_n))$  with

$$\eta_{t,i}(v_i) = \eta_{ST}(v_i, \alpha^*(\varepsilon) s_i^{-1/2}), \quad s_i = \sum_{a \in [m]} W_{ai} \phi_a(t)^{-1}, \quad (12.5.9)$$

and the sequence of profiles  $\{\phi(t)\}_{t \geq 0}$  is given by the following recursion.

$$\phi_a(t+1) = \sum_{i \in [n]} W_{ai} \mathbb{E}\{|\eta_{t,i}(\widehat{X} + s_i^{-1/2} Z; \alpha^* s_i^{-1/2}) - \widehat{X}|^2\}.$$

Finally  $\alpha^* = \alpha^*(\varepsilon)$  is tuned to optimize the phase transition boundary.

This is in fact a generalization of the complex AMP (CAMP) algorithm that was developed in [91] for unstructured matrices. CAMP strives to solve the standard convex relaxation

$$\begin{aligned} & \text{minimize} && \|\widehat{x}\|_1 \\ & \text{subject to} && A_{\mathbb{F}} \widehat{x} = y, \end{aligned}$$

where we recall  $\|\widehat{x}\|_1 \equiv \sum_{\omega \in \Omega_n} |\widehat{x}(\omega)|$ . For a given  $\varepsilon$ , we denote by  $\delta_{\ell_1}(\varepsilon)$  the phase transition location for  $\ell_1$  minimization, when sensing matrices with i.i.d. entries are used. This coincides with the one of CAMP with optimally tuned  $\alpha = \alpha^*(\varepsilon)$  [146, 91].

The empirical phase transition of Scheme III is shown in Fig. 12.5.4. The results are consistent with the claim that the phase boundary coincides with  $\delta_{\ell_1}$ . In other words, spatially-coupled sensing matrix does not improve the performances under  $\ell_1$  reconstruction (or under AMP with soft-thresholding denoiser). This agrees with earlier findings by Krzakala et al. for Gaussian matrices ([82], and private communications). This can be inferred from the the state evolution map. For AMP with posterior expectation denoiser, and for  $\varepsilon < \delta < \tilde{\delta}(\varepsilon)$ , the state evolution map has two stable fixed points; one of order  $\sigma^2$ , and one much larger. Spatial coupling makes the algorithm converge to the ‘right’ fixed point. However, the state evolution map corresponding to the soft-thresholding denoiser is concave and has only one stable fixed point, much larger than  $\sigma^2$ . Therefore, spatial coupling is not helpful in this setting.

# Appendix A

## Supplement to Chapter 2

### A.1 Proof of Lemma 2.4.1

Let  $C_i(\mu)$  be the optimal value of the optimization problem (2.3.11). We claim that

$$C_i(\mu) \geq \frac{(1-\mu)^2}{\widehat{\Sigma}_{ii}}. \quad (\text{A.1.1})$$

To prove this claim notice that the constraint implies (by considering its  $i$ -th component):

$$1 - \langle e_i, \widehat{\Sigma}m \rangle \leq \mu.$$

Therefore if  $\tilde{m}$  is feasible and  $c \geq 0$ , then

$$\langle \tilde{m}, \widehat{\Sigma}\tilde{m} \rangle \geq \langle \tilde{m}, \widehat{\Sigma}\tilde{m} \rangle + c(1-\mu) - c\langle e_i, \widehat{\Sigma}\tilde{m} \rangle \geq \min_m \left\{ \langle m, \widehat{\Sigma}m \rangle + c(1-\mu) - c\langle e_i, \widehat{\Sigma}m \rangle \right\}.$$

Minimizing over all feasible  $\tilde{m}$  gives

$$C_i(\mu) \geq \min_m \left\{ \langle m, \widehat{\Sigma}m \rangle + c(1-\mu) - c\langle e_i, \widehat{\Sigma}m \rangle \right\}. \quad (\text{A.1.2})$$

The minimum over  $m$  is achieved at  $m = ce_i/2$ . Plugging in for  $m$ , we get

$$C_i(\mu) \geq c(1-\mu) - \frac{c^2}{4}\widehat{\Sigma}_{ii} \quad (\text{A.1.3})$$

Optimizing this bound over  $c$ , we obtain the claim (A.1.1), with the optimal choice being  $c = 2(1 - \mu)/\widehat{\Sigma}_{ii}$ .

## A.2 Proof of Lemma 2.4.3

Let  $\mathcal{E}_n = \mathcal{E}_n(\phi_0, s_0, K)$  be the event defined as per Theorem 2.3.4.(a). In particular, we take  $\phi_0 = \sqrt{C_{\min}}/2$ , and  $K \geq 1 + 20\kappa^2\sqrt{(\log p)/n}$ .<sup>1</sup> Further note that we can assume without loss of generality  $n \geq \nu_0 s_0 \log(p/s_0)$ , since  $s_0 = o(\sqrt{n}/\log p)$ . Fixing  $\epsilon > 0$ , we have

$$\begin{aligned} \mathbb{P}\left(\left|\frac{\widehat{\sigma}}{\sigma} - 1\right| \geq \epsilon\right) &\leq \sup_{\mathbf{X} \in \mathcal{E}_n} \mathbb{P}\left(\left|\frac{\widehat{\sigma}}{\sigma} - 1\right| \geq \epsilon \mid \mathbf{X}\right) + \mathbb{P}(\mathbf{X} \notin \mathcal{E}_n) \\ &\leq \sup_{\mathbf{X} \in \mathcal{E}_n} \mathbb{P}\left(\left|\frac{\widehat{\sigma}}{\sigma} - 1\right| \geq \epsilon \mid \mathbf{X}\right) + 4e^{-c_1 n}, \end{aligned}$$

where  $c_1 > 0$  is a constant defined as per Theorem 2.3.4.(a).

We are therefore left with the task of bounding the first term in the last expression above, uniformly over  $\theta_0 \in \mathbb{R}^p$ ,  $\|\theta_0\|_0 \leq s_0$ . For  $\mathbf{X} \in \mathcal{E}_n$ , we apply the result of [124, Theorem 1]. More precisely, using the notations of [124], with  $\lambda_0 = \widetilde{\lambda}$ ,  $\xi = 3$ ,  $T = \text{supp}(\theta_0)$ ,  $\kappa(\xi, T) \geq \phi_0$ , we have  $\eta_*(\widetilde{\lambda}, \xi) \leq 4s_0\widetilde{\lambda}^2/\phi_0^2$ . Further, let  $\sigma^*$  be the oracle estimator of  $\sigma$  introduced there. If  $\|\mathbf{X}^\top W/(n\sigma^*)\|_\infty \leq \widetilde{\lambda}/4$ , using Eq. (13) in [124], we obtain

$$\left|\frac{\widehat{\sigma}}{\sigma^*} - 1\right| \leq \frac{2\sqrt{s_0}\widetilde{\lambda}}{\sigma^*\phi_0} \leq \frac{\epsilon}{2}, \quad (\text{A.2.1})$$

where the last inequality follows for all  $n$  large enough since  $s_0 = o(\sqrt{n}/\log p)$ .

Hence

$$\sup_{\mathbf{X} \in \mathcal{E}_n} \mathbb{P}\left(\left|\frac{\widehat{\sigma}}{\sigma} - 1\right| \geq \epsilon \mid \mathbf{X}\right) \leq \sup_{\mathbf{X} \in \mathcal{E}_n} \mathbb{P}\left(\|\mathbf{X}^\top W/n\|_\infty > \widetilde{\lambda}/4 \mid \mathbf{X}\right) + \sup_{\mathbf{X} \in \mathcal{E}_n} \mathbb{P}\left(\left|\frac{\sigma^*}{\sigma} - 1\right| \geq \frac{\epsilon}{10} \mid \mathbf{X}\right), \quad (\text{A.2.2})$$

where we note that the right hand side is independent of  $\theta_0$ . The first term vanishes as  $n \rightarrow \infty$  by a standard tail bound on the supremum of  $p$  Gaussian random variables. The second term also vanishes because it is controlled by the tail of a chi-squared random variable (see [124]).

<sup>1</sup> For instance  $K = 1.1$  will work for all  $n$  large enough since  $(s_0 \log p)^2/n \rightarrow 0$ , with  $s_0 \geq 1$ , by assumption.

### A.3 Proof of Lemma 6.1.3

Write

$$Z_i = \frac{1}{\sqrt{n}} \sum_{j=1}^n \xi_j \quad \text{with} \quad \xi_j \equiv \frac{m_i^\top X_j W_j}{\sigma [m_i^\top \widehat{\Sigma} m_i]^{1/2}}.$$

Conditional on  $\mathbf{X}$ , the summands  $\xi_j$  are zero mean and independent. Further,  $\sum_{j=1}^n \mathbb{E}(\xi_j^2 | \mathbf{X}) = n$ . We next prove the Lindenberg condition as per Eq. (2.4.11). Let  $c_n \equiv (m_i^\top \widehat{\Sigma} m_i)^{1/2}$ . By Lemma 2.4.1, we have  $\liminf_{n \rightarrow \infty} c_n \geq c_\infty > 0$ , almost surely. If all the optimization problems in (2.4.12) are feasible, then  $|\xi_j| \leq c_n^{-1} \|\mathbf{X} m_i\|_\infty \|W\|_\infty / \sigma \leq c_n^{-1} n^\beta (\|W\|_\infty / \sigma)$ . Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E} \left( \xi_j^2 \mathbb{I}_{\{|\xi_j| > \epsilon \sqrt{n}\}} | \mathbf{X} \right) &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E} \left( \xi_j^2 \mathbb{I}_{\{\|W\|_\infty / \sigma > \epsilon c_n n^{1/2-\beta}\}} | \mathbf{X} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{m_i^\top X_j X_j^\top m_i}{m_i^\top \widehat{\Sigma} m_i} \mathbb{E} \left( \widetilde{W}_j^2 \mathbb{I}_{\{\|\widetilde{W}\|_\infty > \epsilon c_\infty n^{1/2-\beta}\}} \right) \\ &\leq \lim_{n \rightarrow \infty} \mathbb{E} \left( \widetilde{W}_1^2 \mathbb{I}_{\{\|\widetilde{W}_1\| > \epsilon c_\infty n^{1/2-\beta}\}} \right) \\ &\leq c'(\epsilon) \lim_{n \rightarrow \infty} n^{-a(1/2-\beta)} \mathbb{E} \{ |\widetilde{W}_1|^{2+a} \} = 0. \end{aligned}$$

where  $\widetilde{W}_j = W_j / \sigma$  and the last limit follows since  $\beta < 1/2$  and  $a > 0$ .

Using Lindenberg central limit theorem, we obtain  $Z_i | \mathbf{X}$  converges weakly to standard normal distribution, and hence,  $\mathbf{X}$ -almost surely

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_i \leq x | \mathbf{X}) = \Phi(x).$$

What remains is to show that with high probability all the  $p$  optimization problems in (2.4.12) are feasible. In particular, we show that  $\Sigma_{i,\cdot}^{-1}$  is a feasible solution to the  $i$ -th optimization problem, for  $i \in [p]$ . By Lemma 6.1.2,  $|\Sigma^{-1} \widehat{\Sigma} - \mathbf{I}|_\infty \leq \mu$ , with high probability. Moreover,

$$\begin{aligned} \sup_{j \in [p]} \|\Sigma_{i,\cdot}^{-1} X_j\|_{\psi_2} &= \sup_{j \in [p]} \|\Sigma_{i,\cdot}^{-1/2} \Sigma^{-1/2} X_j\|_{\psi_2} \\ &= \|\Sigma_{i,\cdot}^{-1/2}\|_2 \sup_{j \in [p]} \|\Sigma^{-1/2} X_j\|_{\psi_2} \\ &= [\Sigma_{i,i}^{-1}]^{1/2} \sup_{j \in [p]} \|\Sigma^{-1/2} X_j\|_{\psi_2} = O(1). \end{aligned}$$

Using tail bound for sub-gaussian variables  $\Sigma_{i,\cdot}^{-1} X_j$  and union bounding over  $j \in [n]$ , we get

$$\mathbb{P}(\|\mathbf{X}\Sigma_{\cdot,i}^{-1}\|_\infty > n^\beta) \leq ne^{-cn^{2\beta}},$$

for some constant  $c > 0$ . Note that  $s_0 = o(\sqrt{n}/\log p)$  and  $\beta > 1/4$  imply  $p = e^{o(n^{2\beta})}$ . Hence, almost surely,  $\Sigma_{i,\cdot}^{-1}$  is a feasible solution to optimization problem (2.4.12), for all  $i \in [p]$ .

# Appendix B

## Supplement to Chapter 4

### B.1 Effective noise variance $\tau_0^2$

As stated in Theorem 4.1.4 the unbiased estimator  $\hat{\theta}^u$  can be regarded –asymptotically– as a noisy version of  $\theta_0$  with noise variance  $\tau_0^2$ . An explicit formula for  $\tau_0$  is given in [8]. For the reader’s convenience, we explain it here using our notations.

Denote by  $\eta : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  the soft thresholding function

$$\eta(x; a) = \begin{cases} x - a & \text{if } x > a, \\ 0 & \text{if } -a \leq x \leq a \\ x + a & \text{otherwise.} \end{cases} \quad (\text{B.1.1})$$

Further define function  $F : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$  as

$$F(\tau^2, a) = \sigma^2 + \frac{1}{\delta} \mathbb{E}\{[\eta(\Theta_0 + \tau Z; a) - \Theta_0]^2\}, \quad (\text{B.1.2})$$

where  $\Theta_0$  and  $Z$  are defined as in Theorem 4.1.4. Let  $\kappa_{\min} = \kappa_{\min}(\delta)$  be the unique non-negative solution of the equation

$$(1 + \kappa^2)\Phi(-\kappa) - \kappa\phi(\kappa) = \frac{\delta}{2}. \quad (\text{B.1.3})$$

The effective noise variance  $\tau_0^2$  is obtained by solving the following two equations for  $\kappa$  and

$\tau$ , restricted to the interval  $\kappa \in (\kappa_{\min}, \infty)$ :

$$\tau^2 = F(\tau^2, \kappa\tau), \quad (\text{B.1.4})$$

$$\lambda = \kappa\tau \left[ 1 - \frac{1}{\delta} \mathbb{P}(|\Theta_0 + \tau Z| \geq \kappa\tau) \right]. \quad (\text{B.1.5})$$

Existence and uniqueness of  $\tau_0$  is proved in [8, Proposition 1.3].

## B.2 Tuned regularization parameter $\lambda$

In previous appendix, we provided the value of  $\tau_0$  for a given regularization parameter  $\lambda$ . In this appendix, we discuss the tuned value for  $\lambda$  to achieve the power stated in Theorem 4.1.3.

Let  $\mathcal{F}_\varepsilon \equiv \{p_{\Theta_0} : p_{\Theta_0}(\{0\}) \geq 1 - \varepsilon\}$  be the family of  $\varepsilon$ -sparse distributions. Also denote by  $M(\varepsilon, \kappa)$  the minimax risk of soft thresholding denoiser (at threshold value  $\kappa$ ) over  $\mathcal{F}_\varepsilon$ , i.e.,

$$M(\varepsilon, \kappa) = \sup_{p_{\Theta_0} \in \mathcal{F}_\varepsilon} \mathbb{E}\{\eta(\Theta_0 + Z; \kappa) - \Theta_0\}^2. \quad (\text{B.2.1})$$

The function  $M$  can be computed explicitly by evaluating the mean square error on the worst case  $\varepsilon$ -sparse distribution. A simple calculation gives

$$M(\varepsilon, \kappa) = \varepsilon(1 + \kappa^2) + (1 - \varepsilon)[2(1 + \kappa^2)\Phi(-\kappa) - 2\kappa\phi(\kappa)]. \quad (\text{B.2.2})$$

Further, let

$$\kappa_*(\varepsilon) \equiv \arg \min_{\kappa \in \mathbb{R}_+} M(\varepsilon, \kappa). \quad (\text{B.2.3})$$

In words,  $\kappa_*(\varepsilon)$  is the minimax optimal value of threshold  $\kappa$  over  $\mathcal{F}_\varepsilon$ . The value of  $\lambda$  for Theorem 4.1.3 is then obtained by solving Eq. (B.1.4) for  $\tau$  with  $\kappa = \kappa_*(\varepsilon)$ , and then substituting  $\kappa_*$  and  $\tau$  in Eq. (B.1.5) to get  $\lambda = \lambda(p_{\Theta_0}, \sigma, \varepsilon, \delta)$ .

**Remark B.2.1.** *The theory of [8, 44] implies that in the standard Gaussian setting and for a converging sequence of instances  $\{(\theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$ , Eq. (B.1.5) is equivalent to the following:*

$$\lambda d = \kappa\tau, \quad (\text{B.2.4})$$

where the normalization factor  $\mathbf{d}$  is given by Eq. (4.1.2).

### B.3 Replica method calculation

In this section we outline the replica calculation leading to the Claim 4.2.5. Indeed we consider an even more general setting, where the  $\ell_1$  regularization is replaced by an arbitrary separable penalty. Namely, instead of the Lasso, we consider regularized least squares estimators of the form

$$\widehat{\theta}(Y, \mathbf{X}) = \arg \min_{\theta \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|^2 + J(\theta) \right\}, \quad (\text{B.3.1})$$

with  $J(\theta)$  being a convex separable penalty function; namely for a vector  $\theta \in \mathbb{R}^p$ , we have  $J(\theta) = J_1(\theta_1) + \dots + J_p(\theta_p)$ , with  $J_\ell : \mathbb{R} \rightarrow \mathbb{R}$  a convex function. Important instances from this ensemble of estimators are Ridge-regression ( $J(\theta) = \lambda \|\theta\|^2/2$ ), and the Lasso ( $J(\theta) = \lambda \|\theta\|_1$ ). The Replica Claim 4.2.5 is generalized to the present setting replacing  $\lambda \|\theta\|_1$  by  $J(\theta)$ . The only required modification concerns the definition of the factor  $\mathbf{d}$ . We let  $\mathbf{d}$  be the unique positive solution of the following equation

$$1 = \frac{1}{\mathbf{d}} + \frac{1}{n} \text{Trace} \left\{ (\mathbf{I} + \mathbf{d} \Sigma^{-1/2} \nabla^2 J(\widehat{\theta}) \Sigma^{-1/2})^{-1} \right\}, \quad (\text{B.3.2})$$

where  $\nabla^2 J(\widehat{\theta})$  denotes the Hessian, which is diagonal since  $J$  is separable. If  $J$  is non-differentiable, then we formally set  $[\nabla^2 J(\widehat{\theta})]_{ii} = \infty$  for all the coordinates  $i$  such that  $J$  is non-differentiable at  $\widehat{\theta}_i$ . It can be checked that this definition is well posed and yields the previous choice for  $J(\theta) = \lambda \|\theta\|_1$ .

We pass next to establishing the claim. We limit ourselves to the main steps, since analogous calculations can be found in several earlier works [127, 63, 125]. For a general introduction to the method and its motivation we refer to [98, 97]. Also, for the sake of simplicity, we shall focus on characterizing the asymptotic distribution of  $\widehat{\theta}^u$ , cf. Eq. (4.2.3). The distribution of  $r$  is derived by the same approach.

Fix a sequence of instances  $\{(\Sigma(p), \theta_0(p), n(p), \sigma(p))\}_{p \in \mathbb{N}}$ . For the sake of simplicity, we assume  $\sigma(p)^2 = n(p)\sigma_0^2$  and  $n(p) = p\delta$  (the slightly more general case  $\sigma(p)^2 = n(p)[\sigma_0^2 + o(1)]$  and  $n(p) = p[\delta + o(1)]$  does not require any change to the derivation given here, but is more cumbersome notationally). Fix  $\widetilde{g} : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  a continuous function convex in its first argument, and let  $g(u, y, z) \equiv \max_{x \in \mathbb{R}} [ux - \widetilde{g}(x, y, z)]$  be its Lagrange dual. The

replica calculation aims at estimating the following moment generating function (*partition function*)

$$\begin{aligned} \mathcal{Z}_p(\beta, s) \equiv \int \exp \left\{ -\frac{\beta}{2n} \|Y - \mathbf{X}\theta\|_2^2 - \beta J(\theta) - \beta s \sum_{i=1}^p [g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) - u_i \hat{\theta}_i^u] \right. \\ \left. - \frac{\beta}{2n} (s\tilde{\mathbf{d}})^2 \|\mathbf{X}\Sigma^{-1}u\|_2^2 \right\} d\theta du. \end{aligned} \quad (\text{B.3.3})$$

Here  $(Y_i, X_i)$  are i.i.d. pairs distributed as per model (1.1.5) and  $\hat{\theta}^u = \theta + (\tilde{\mathbf{d}}/n) \Sigma^{-1} \mathbf{X}^\top (Y - \mathbf{X}\theta)$  with  $\tilde{\mathbf{d}} \in \mathbb{R}$  to be defined below. Further,  $g : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function strictly convex in its first argument. Finally,  $s \in \mathbb{R}_+$  and  $\beta > 0$  is a ‘temperature’ parameter not to be confused with the type II error rate as used in the main text. We will eventually show that the appropriate choice of  $\tilde{\mathbf{d}}$  is given by Eq. (B.3.2).

Within the replica method, it is assumed that the limits  $p \rightarrow \infty$ ,  $\beta \rightarrow \infty$  exist almost surely for the quantity  $(p\beta)^{-1} \log \mathcal{Z}_p(\beta, s)$ , and that the order of the limits can be exchanged. We therefore define

$$\mathfrak{F}(s) \equiv - \lim_{\beta \rightarrow \infty} \lim_{p \rightarrow \infty} \frac{1}{p\beta} \log \mathcal{Z}_p(\beta, s) \quad (\text{B.3.4})$$

$$\equiv - \lim_{p \rightarrow \infty} \lim_{\beta \rightarrow \infty} \frac{1}{p\beta} \log \mathcal{Z}_p(\beta, s). \quad (\text{B.3.5})$$

In other words  $\mathfrak{F}(s)$  is the exponential growth rate of  $\mathcal{Z}_p(\beta, s)$ . It is also assumed that  $p^{-1} \log \mathcal{Z}_p(\beta, s)$  concentrates tightly around its expectation so that  $\mathfrak{F}(s)$  can in fact be evaluated by computing

$$\mathfrak{F}(s) = - \lim_{\beta \rightarrow \infty} \lim_{p \rightarrow \infty} \frac{1}{p\beta} \mathbb{E} \log \mathcal{Z}_p(\beta, s), \quad (\text{B.3.6})$$

where expectation is being taken with respect to the distribution of  $(Y_1, X_1), \dots, (Y_n, X_n)$ . Notice that, by Eq. (B.3.5) and using Laplace method in the integral (B.3.3), we have

$$\begin{aligned} \mathfrak{F}(s) = \\ \lim_{p \rightarrow \infty} \frac{1}{p} \min_{\theta, u \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|Y - \mathbf{X}\theta\|_2^2 + J(\theta) + s \sum_{i=1}^p [g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) - u_i \hat{\theta}_i^u] + \frac{1}{2n} (s\tilde{\mathbf{d}})^2 \|\mathbf{X}\Sigma^{-1}u\|_2^2 \right\}. \end{aligned} \quad (\text{B.3.7})$$

Finally we assume that the derivative of  $\mathfrak{F}(s)$  as  $s \rightarrow 0$  can be obtained by differentiating

inside the limit. This condition holds, for instance, if the cost function is strongly convex at  $s = 0$ . We get

$$\frac{d\mathfrak{F}}{ds}(s=0) = \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \min_{u_i \in \mathbb{R}} [g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) - u_i \hat{\theta}_i^u] \quad (\text{B.3.8})$$

where  $\hat{\theta}^u = \hat{\theta} + (\tilde{d}/n) \Sigma^{-1} \mathbf{X}^\top (Y - \mathbf{X}\hat{\theta})$  and  $\hat{\theta}$  is the minimizer of the regularized least squares as per Eq. (B.3.1). Since, by duality  $\tilde{g}(x, y, z) \equiv \max_{u \in \mathbb{R}} [ux - g(u, y, z)]$ , we get

$$\frac{d\mathfrak{F}}{ds}(s=0) = - \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \tilde{g}(\hat{\theta}_i^u, \theta_{0,i}, (\Sigma^{-1})_{ii}). \quad (\text{B.3.9})$$

Hence, by computing  $\mathfrak{F}(s)$  using Eq. (B.3.6) for a complete set of functions  $\tilde{g}$ , we get access to the corresponding limit quantities (B.3.9) and hence, via standard weak convergence arguments, to the joint empirical distribution of the triple  $(\hat{\theta}_i^u, \theta_{0,i}, (\Sigma^{-1})_{ii})$ , cf. Eq. (4.2.4).

In order to carry out the calculation of  $\mathfrak{F}(s)$ , we begin by rewriting the partition function (B.3.3) in a more convenient form. Using the definition of  $\hat{\theta}^u$  and after a simple manipulation

$$\begin{aligned} \mathcal{Z}_p(\beta, s) = & \int \exp \left\{ -\frac{\beta}{2n} \|Y - \mathbf{X}(\theta + s\tilde{d}\Sigma^{-1}u)\|_2^2 - \beta J(\theta) + \beta s \langle u, \theta \rangle - \beta s \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) \right\} d\theta du. \end{aligned} \quad (\text{B.3.10})$$

Define the measure  $\nu(d\theta)$  over  $\theta \in \mathbb{R}^p$  as follows

$$\nu(d\theta) = \int \exp \left\{ -\beta J(\theta - s\tilde{d}\Sigma^{-1}u) + \beta s \langle \theta - s\tilde{d}\Sigma^{-1}u, u \rangle - \beta s \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) \right\} du. \quad (\text{B.3.11})$$

Using this definition and with the change of variable  $\theta' = \theta + s\tilde{d}\Sigma^{-1}u$ , we can rewrite

Eq. (B.3.10) as

$$\begin{aligned}
\mathcal{Z}_p(\beta, s) &\equiv \int \exp \left\{ -\frac{\beta}{2n} \|Y - \mathbf{X}\theta\|_2^2 \right\} \nu(d\theta) \\
&= \int \exp \left\{ i\sqrt{\frac{\beta}{n}} \langle z, Y - \mathbf{X}\theta \rangle \right\} \nu(d\theta) \gamma_n(dz) \\
&= \int \exp \left\{ i\sqrt{\frac{\beta}{n}} \langle W, z \rangle + i\sqrt{\frac{\beta}{n}} \langle z, \mathbf{X}(\theta_0 - \theta) \rangle \right\} \nu(d\theta) \gamma(dz), \tag{B.3.12}
\end{aligned}$$

where  $\gamma_n(dz)$  denotes the standard Gaussian measure on  $\mathbb{R}^n$ :

$$\gamma_n(dz) \equiv (2\pi)^{-\frac{n}{2}} \exp \left( -\frac{\|z\|_2^2}{2} \right) dz.$$

The replica method aims at computing the expected log-partition function, cf. Eq. (B.3.6) using the identity

$$\mathbb{E} \log \mathcal{Z}_p(\beta, s) = \frac{d}{dk} \Big|_{k=0} \log \mathbb{E} \{ \mathcal{Z}_p(\beta, s)^k \}. \tag{B.3.13}$$

This formula would require computing fractional moments of  $\mathcal{Z}_p$  as  $k \rightarrow 0$ . The replica method consists in a prescription that allows to compute a formal expression for the  $k$  integer, and then extrapolate it as  $k \rightarrow 0$ . Crucially, the limit  $k \rightarrow 0$  is inverted with the one  $p \rightarrow \infty$ :

$$\lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \log \mathcal{Z}_p(\beta, s) = \frac{d}{dk} \Big|_{k=0} \lim_{p \rightarrow \infty} \frac{1}{p} \log \mathbb{E} \{ \mathcal{Z}_p(\beta, s)^k \}. \tag{B.3.14}$$

In order to represent  $\mathcal{Z}_p(\beta, s)^k$ , we use the identity

$$\left( \int f(x) \rho(dx) \right)^k = \int f(x^1) f(x^2) \cdots f(x^k) \rho(dx^1) \cdots \rho(dx^k). \tag{B.3.15}$$

In order to apply this formula to Eq. (B.3.12), we let, with a slight abuse of notation,  $\nu^k(d\theta) \equiv \nu(d\theta^1) \times \nu(d\theta^2) \times \cdots \times \nu(d\theta^k)$  be a measure over  $(\mathbb{R}^p)^k$ , with  $\theta^1, \dots, \theta^k \in \mathbb{R}^p$ . Analogously  $\gamma_n^k(dz) \equiv \gamma_n(dz^1) \times \gamma_n(dz^2) \times \cdots \times \gamma_n(dz^k)$ , with  $z^1, \dots, z^k \in \mathbb{R}^n$ . With these

notations, we have

$$\mathbb{E}\{\mathcal{Z}_p(\beta, s)^k\} = \int \mathbb{E} \exp \left\{ i \sqrt{\frac{\beta}{n}} \langle W, \sum_{a=1}^k z^a \rangle + i \sqrt{\frac{\beta}{n}} \langle \mathbf{X}, \sum_{a=1}^k z^a (\theta_0 - \theta^a)^\top \rangle \right\} \nu^k(d\theta) \gamma_n^k(dz). \quad (\text{B.3.16})$$

In the above expression  $\mathbb{E}$  denotes expectation with respect to the noise vector  $W$ , and the design matrix  $\mathbf{X}$ . Further, we used  $\langle \cdot, \cdot \rangle$  to denote matrix scalar product as well:  $\langle A, B \rangle \equiv \text{Trace}(A^\top B)$ .

At this point we can take the expectation with respect to  $W$ ,  $\mathbf{X}$ . We use the fact that, for any  $M \in \mathbb{R}^{n \times p}$ ,  $u \in \mathbb{R}^n$

$$\begin{aligned} \mathbb{E}\{\exp(i\langle W, u \rangle)\} &= \exp\left\{-\frac{1}{2}n\sigma_0^2 \|u\|_2^2\right\}, \\ \mathbb{E}\{\exp(i\langle M, \mathbf{X} \rangle)\} &= \exp\left\{-\frac{1}{2}\langle M, M\Sigma \rangle\right\}, \end{aligned} \quad (\text{B.3.17})$$

Using these identities in Eq. (B.3.16), we obtain

$$\begin{aligned} \mathbb{E}\{\mathcal{Z}_p^k\} &= \\ &\int \exp\left\{-\frac{1}{2}\beta\sigma_0^2 \sum_{a=1}^k \|z^a\|_2^2 - \frac{\beta}{2n} \sum_{a,b=1}^k \langle z^a, z^b \rangle \langle (\theta^a - \theta_0), \Sigma(\theta^b - \theta_0) \rangle\right\} \nu^k(d\theta) \gamma_n^k(dz). \end{aligned} \quad (\text{B.3.18})$$

We next use the identity

$$e^{-xy} = \frac{1}{2\pi i} \int_{(-i\infty, i\infty)} \int_{(-\infty, \infty)} e^{-\zeta q + \zeta x - qy} d\zeta dq, \quad (\text{B.3.19})$$

where the integral is over  $\zeta \in (-i\infty, i\infty)$  (imaginary axis) and  $q \in (-\infty, \infty)$ . We apply this identity to Eq. (B.3.18), and introduce integration variables  $Q \equiv (Q_{ab})_{1 \leq a, b \leq k}$  and

$\Lambda \equiv (\Lambda_{ab})_{1 \leq a, b \leq k}$ . Letting  $dQ \equiv \prod_{a,b} dQ_{ab}$  and  $d\Lambda \equiv \prod_{a,b} d\Lambda_{ab}$

$$\mathbb{E}\{\mathcal{Z}_p^k\} = \left(\frac{\beta n}{4\pi i}\right)^{k^2} \int \exp\left\{-p\mathcal{S}_k(Q, \Lambda)\right\} dQ d\Lambda, \quad (\text{B.3.20})$$

$$\mathcal{S}_k(Q, \Lambda) = \frac{\beta\delta}{2} \sum_{a,b=1}^k \Lambda_{ab} Q_{ab} - \frac{1}{p} \log \xi(\Lambda) - \delta \log \widehat{\xi}(Q), \quad (\text{B.3.21})$$

$$\xi(\Lambda) \equiv \int \exp\left\{\frac{\beta}{2} \sum_{a,b=1}^k \Lambda_{ab} \langle (\theta^a - \theta_0), \Sigma(\theta^b - \theta_0) \rangle\right\} \nu^k(d\theta), \quad (\text{B.3.22})$$

$$\widehat{\xi}(Q) \equiv \int \exp\left\{-\frac{\beta}{2} \sum_{a,b=1}^k (\sigma_0^2 \mathbf{I} + Q)_{a,b} z_1^a z_1^b\right\} \gamma_1^k(dz_1). \quad (\text{B.3.23})$$

Notice that above we used the fact that, after introducing  $Q, \Lambda$ , the integral over  $(z^1, \dots, z^k) \in (\mathbb{R}^n)^k$  factors into  $n$  integrals over  $(\mathbb{R})^k$  with measure  $\gamma_1^k(dz_1)$ .

We next use the saddle point method in Eq. (B.3.20) to obtain

$$-\lim_{p \rightarrow \infty} \frac{1}{p} \log \mathbb{E}\{\mathcal{Z}_p^k\} = \mathcal{S}_k(Q^*, \Lambda^*), \quad (\text{B.3.24})$$

where  $Q^*, \Lambda^*$  is the saddle-point location. The replica method provides a hierarchy of ansatz for this saddle-point. The first level of this hierarchy is the so-called *replica symmetric* ansatz postulating that  $Q^*, \Lambda^*$  ought to be invariant under permutations of the row/column indices. This is motivated by the fact that  $\mathcal{S}_k(Q, \Lambda)$  is indeed left unchanged by such change of variables. This is equivalent to postulating that

$$Q_{ab}^* = \begin{cases} q_1 & \text{if } a = b, \\ q_0 & \text{otherwise,} \end{cases}, \quad \Lambda_{ab}^* = \begin{cases} \beta\zeta_1 & \text{if } a = b, \\ \beta\zeta_0 & \text{otherwise,} \end{cases} \quad (\text{B.3.25})$$

where the factor  $\beta$  is for future convenience. Given that the partition function, cf. Eq. (B.3.3) is the integral of a log-concave function, it is expected that the replica-symmetric ansatz yields in fact the correct result [98, 97].

The next step consists in substituting the above expressions for  $Q^*, \Lambda^*$  in  $\mathcal{S}_k(\cdot, \cdot)$  and then taking the limit  $k \rightarrow 0$ . We will consider separately each term of  $\mathcal{S}_k(Q, \Lambda)$ , cf. Eq. (B.3.21).

Let us begin with the first term

$$\sum_{a,b=1}^k \Lambda_{ab}^* Q_{ab}^* = k \beta \zeta_1 q_1 + k(k-1) \beta \zeta_0 q_0. \quad (\text{B.3.26})$$

Hence

$$\lim_{k \rightarrow 0} \frac{\beta \delta}{2k} \sum_{a,b=1}^k \Lambda_{ab}^* Q_{ab}^* = \frac{\beta^2 \delta}{2} (\zeta_1 q_1 - \zeta_0 q_0). \quad (\text{B.3.27})$$

Let us consider  $\widehat{\xi}(Q^*)$ . We have

$$\log \widehat{\xi}(Q^*) = -\frac{1}{2} \log \text{Det}(\mathbf{I} + \beta \sigma^2 \mathbf{I} + \beta Q^*) \quad (\text{B.3.28})$$

$$= -\frac{k-1}{2} \log(1 + \beta(q_1 - q_0)) - \frac{1}{2} \log(1 + \beta(q_1 - q_0) + \beta k(\sigma^2 + q_0)). \quad (\text{B.3.29})$$

In the limit  $k \rightarrow 0$  we thus obtain

$$\lim_{k \rightarrow 0} \frac{1}{k} (-\delta) \log \widehat{\xi}(Q^*) = \frac{\delta}{2} \log(1 + \beta(q_1 - q_0)) + \frac{\delta}{2} \frac{\beta(\sigma^2 + q_0)}{1 + \beta(q_1 - q_0)}. \quad (\text{B.3.30})$$

Finally, introducing the notation  $\|v\|_{\Sigma}^2 \equiv \langle v, \Sigma v \rangle$ , we have

$$\begin{aligned} \xi(\Lambda^*) &\equiv \int \exp \left\{ \frac{\beta^2}{2} (\zeta_1 - \zeta_0) \sum_{a=1}^k \|\theta^a - \theta_0\|_{\Sigma}^2 + \frac{\beta^2 \zeta_0}{2} \sum_{a,b=1}^k \langle (\theta^a - \theta_0), \Sigma (\theta^b - \theta_0) \rangle \right\} \nu^k(d\theta), \\ &= \mathbb{E} \int \exp \left\{ \frac{\beta^2}{2} (\zeta_1 - \zeta_0) \sum_{a=1}^k \|\theta^a - \theta_0\|_{\Sigma}^2 + \beta \sqrt{\zeta_0} \sum_{a=1}^k \langle z, \Sigma^{1/2} (\theta^a - \theta_0) \rangle \right\} \nu^k(d\theta), \end{aligned} \quad (\text{B.3.31})$$

where expectation is with respect to  $z \sim \mathbf{N}(0, \mathbf{I}_{p \times p})$ . Notice that, given  $z \in \mathbb{R}^p$ , the integrals over  $\theta^1, \theta^2, \dots, \theta^k$  factorize, whence

$$\xi(\Lambda^*) = \mathbb{E} \left\{ \left[ \int \exp \left\{ \frac{\beta^2}{2} (\zeta_1 - \zeta_0) \|\theta - \theta_0\|_{\Sigma}^2 + \beta \sqrt{\zeta_0} \langle z, \Sigma^{1/2} (\theta - \theta_0) \rangle \right\} \nu(d\theta) \right]^k \right\}. \quad (\text{B.3.32})$$

Therefore

$$\begin{aligned} \lim_{k \rightarrow 0} \frac{(-1)}{pk} \log \xi(\Lambda^*) = \\ - \frac{1}{p} \mathbb{E} \left\{ \log \left[ \int \exp \left\{ \frac{\beta^2}{2} (\zeta_1 - \zeta_0) \|\theta - \theta_0\|_{\Sigma}^2 + \beta \sqrt{\zeta_0} \langle z, \Sigma^{1/2}(\theta - \theta_0) \rangle \right\} \nu(d\theta) \right] \right\}. \end{aligned} \quad (\text{B.3.33})$$

Putting Eqs. (B.3.27), (B.3.30), and (B.3.33) together we obtain

$$\begin{aligned} - \lim_{p \rightarrow \infty} \frac{1}{p\beta} \mathbb{E} \log \mathcal{Z}_p &= \lim_{k \rightarrow 0} \frac{1}{k\beta} \mathcal{S}_k(Q^*, \Lambda^*) \\ &= \frac{\beta\delta}{2} (\zeta_1 q_1 - \zeta_0 q_0) + \frac{\delta}{2\beta} \log(1 + \beta(q_1 - q_0)) + \frac{\delta}{2} \frac{\sigma^2 + q_0}{1 + \beta(q_1 - q_0)} \\ &\quad - \lim_{p \rightarrow \infty} \frac{1}{p\beta} \mathbb{E} \left\{ \log \left[ \int \exp \left\{ \frac{\beta^2}{2} (\zeta_1 - \zeta_0) \|\theta - \theta_0\|_{\Sigma}^2 \right. \right. \right. \\ &\quad \left. \left. \left. + \beta \sqrt{\zeta_0} \langle z, \Sigma^{1/2}(\theta - \theta_0) \rangle \right\} \nu(d\theta) \right] \right\}. \end{aligned} \quad (\text{B.3.34})$$

We can next take the limit  $\beta \rightarrow \infty$ . In doing this, one has to be careful with respect to the behavior of the saddle point parameters  $q_0, q_1, \zeta_0, \zeta_1$ . A careful analysis (omitted here) shows that  $q_0, q_1$  have the same limit, denoted here by  $q_0$ , and  $\zeta_0, \zeta_1$  have the same limit, denoted by  $\zeta_0$ . Moreover  $q_1 - q_0 = (q/\beta) + o(\beta^{-1})$  and  $\zeta_1 - \zeta_0 = (-\zeta/\beta) + o(\beta^{-1})$ . Substituting in the above expression, and using Eq. (B.3.6), we get

$$\begin{aligned} \mathfrak{F}(s) &= \frac{\delta}{2} (\zeta_0 q - \zeta q_0) + \frac{\delta}{2} \frac{q_0 + \sigma^2}{1 + q} \\ &\quad + \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \min_{\theta \in \mathbb{R}^p} \left\{ \frac{\zeta}{2} \|\theta - \theta_0\|_{\Sigma}^2 - \sqrt{\zeta_0} \langle z, \Sigma^{1/2}(\theta - \theta_0) \rangle + \tilde{J}(\theta; s) \right\}, \end{aligned} \quad (\text{B.3.35})$$

$$\tilde{J}(\theta; s) = \min_{u \in \mathbb{R}^p} \left\{ J(\theta - s\tilde{d}\Sigma^{-1}u) - s \langle \theta - s\tilde{d}\Sigma^{-1}u, u \rangle + s \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) \right\}. \quad (\text{B.3.36})$$

After the change of variable  $\theta - s\mathbf{d}\Sigma^{-1}u \rightarrow \theta$ , this reads

$$\begin{aligned}\mathfrak{F}(s) &= \frac{\delta}{2}(\zeta_0 q - \zeta q_0) + \frac{\delta}{2} \frac{q_0 + \sigma_0^2}{1+q} - \frac{\zeta_0}{2\zeta} \\ &\quad + \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \min_{\theta, u \in \mathbb{R}^p} \left\{ \frac{\zeta}{2} \left\| \theta - \theta_0 - \frac{\sqrt{\zeta_0}}{\zeta} \Sigma^{-1/2} z + s\mathbf{d}\Sigma^{-1}u \right\|_{\Sigma}^2 + \tilde{J}(\theta, u; s) \right\}, \\ \tilde{J}(\theta, u; s) &= J(\theta) - s\langle \theta, u \rangle + s \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}).\end{aligned}$$

Finally, we must set  $\zeta, \zeta_0$  and  $q, q_0$  to their saddle point values. We start by using the stationarity conditions with respect to  $q, q_0$ :

$$\frac{\partial \mathfrak{F}}{\partial q}(s) = \frac{\delta}{2} \zeta_0 - \frac{\delta}{2} \frac{q_0 + \sigma_0^2}{(1+q)^2}, \quad (\text{B.3.37})$$

$$\frac{\partial \mathfrak{F}}{\partial q_0}(s) = -\frac{\delta}{2} \zeta + \frac{\delta}{2} \frac{1}{1+q}. \quad (\text{B.3.38})$$

We use these to eliminate  $q$  and  $q_0$ . Renaming  $\zeta_0 = \zeta^2 \tau^2$ , we get our final expression for  $\mathfrak{F}(s)$ :

$$\begin{aligned}\mathfrak{F}(s) &= -\frac{1}{2}(1-\delta)\zeta\tau^2 - \frac{\delta}{2}\zeta^2\tau^2 + \frac{\delta}{2}\sigma_0^2\zeta \\ &\quad + \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \min_{\theta, u \in \mathbb{R}^p} \left\{ \frac{\zeta}{2} \left\| \theta - \theta_0 - \tau \Sigma^{-1/2} z + s\mathbf{d}\Sigma^{-1}u \right\|_{\Sigma}^2 + \tilde{J}(\theta, u; s) \right\},\end{aligned} \quad (\text{B.3.39})$$

with

$$\tilde{J}(\theta, u; s) = J(\theta) - s\langle \theta, u \rangle + s \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}). \quad (\text{B.3.40})$$

Here it is understood that  $\zeta$  and  $\tau^2$  are to be set to their saddle point values.

We are interested in the derivative of  $\mathfrak{F}(s)$  with respect to  $s$ , cf. Eq. (B.3.9). Consider first the case  $s = 0$ . Using the assumption  $\mathfrak{E}^{(p)}(a, b) \rightarrow \mathfrak{E}(a, b)$ , cf. Eq. (4.2.9), we get

$$\mathfrak{F}(s=0) = -\frac{1}{2}(1-\delta)\zeta\tau^2 - \frac{\delta}{2}\zeta^2\tau^2 + \frac{\delta}{2}\sigma_0^2\zeta + \mathfrak{E}(\tau^2, \zeta). \quad (\text{B.3.41})$$

The values of  $\zeta$ ,  $\tau^2$  are obtained by setting to zero the partial derivatives

$$\frac{\partial \mathfrak{F}}{\partial \zeta}(s=0) = -\frac{1}{2}(1-\delta)\tau^2 - \delta\zeta\tau^2 + \frac{\delta}{2}\sigma_0^2 + \frac{\partial \mathfrak{E}}{\partial \zeta}(\tau^2, \zeta), \quad (\text{B.3.42})$$

$$\frac{\partial \mathfrak{F}}{\partial \tau^2}(s=0) = -\frac{1}{2}(1-\delta)\zeta - \frac{\delta}{2}\zeta^2 + \frac{\partial \mathfrak{E}}{\partial \tau^2}(\tau^2, \zeta), \quad (\text{B.3.43})$$

Define, as in the statement of the Replica Claim

$$\mathbf{E}_1(a, b) \equiv \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \left\{ \left\| \eta_b(\theta_0 + \sqrt{a}\Sigma^{-1/2}z) - \theta_0 \right\|_{\Sigma}^2 \right\}, \quad (\text{B.3.44})$$

$$\begin{aligned} \mathbf{E}_2(a, b) &\equiv \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \left\{ \text{div} \eta_b(\theta_0 + \sqrt{a}\Sigma^{-1/2}z) \right\} \\ &= \lim_{p \rightarrow \infty} \frac{1}{p\tau} \mathbb{E} \left\{ \langle \eta_b(\theta_0 + \sqrt{a}\Sigma^{-1/2}z), \Sigma^{1/2}z \rangle \right\}, \end{aligned} \quad (\text{B.3.45})$$

where the last identity follows by integration by parts. These limits exist by the assumption that  $\nabla \mathfrak{E}^{(p)}(a, b) \rightarrow \nabla \mathfrak{E}(a, b)$ . In particular

$$\frac{\partial \mathfrak{E}}{\partial \zeta}(\tau^2, \zeta) = \frac{1}{2} \mathbf{E}_1(\tau^2, \zeta) - \tau^2 \mathbf{E}_2(\tau^2, \zeta) + \frac{1}{2} \tau^2, \quad (\text{B.3.46})$$

$$\frac{\partial \mathfrak{E}}{\partial \tau^2}(\tau^2, \zeta) = -\frac{\zeta}{2} \mathbf{E}_2(\tau^2, \zeta) + \frac{1}{2} \zeta. \quad (\text{B.3.47})$$

Substituting these expressions in Eqs. (B.3.42), (B.3.43), and simplifying, we conclude that the derivatives vanish if and only if  $\zeta, \tau^2$  satisfy the following equations

$$\tau^2 = \sigma_0^2 + \frac{1}{\delta} \mathbf{E}_1(\tau^2, \zeta), \quad (\text{B.3.48})$$

$$\zeta = 1 - \frac{1}{\delta} \mathbf{E}_2(\tau^2, \zeta). \quad (\text{B.3.49})$$

The solution of these equations is expected to be unique for  $J$  convex and  $\sigma_0^2 > 0$ .

Next consider the derivative of  $\mathfrak{F}(s)$  with respect to  $s$ , which is our main object of interest, cf. Eq. (B.3.9). By differentiating Eq. (B.3.39) and inverting the order of derivative and limit, we get

$$\frac{d\mathfrak{F}}{ds}(s=0) = \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \min_{u \in \mathbb{R}^p} \left\{ \zeta \tilde{d} \langle u, \hat{\theta} - \theta_0 - \tau \Sigma^{-1/2}z \rangle - \langle \hat{\theta}, u \rangle + \sum_{i=1}^p g(u_i, \theta_{0,i}, (\Sigma^{-1})_{ii}) \right\},$$

where  $\hat{\theta}$  is the minimizer at  $s=0$ , i.e.,  $\hat{\theta} = \eta_{\zeta}(\theta_0 + \tau \Sigma^{-1/2}z)$ , and  $\zeta, \tau^2$  solve Eqs. (B.3.48),

(B.3.49). At this point we choose  $\tilde{\mathbf{d}} = 1/\zeta$ . Minimizing over  $u$  (recall that  $\tilde{g}(x, y, z) = \max_{u \in \mathbb{R}}[ux - g(u, y, z)]$ ), we get

$$\frac{d\mathfrak{F}}{ds}(s=0) = - \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \tilde{g}(\theta_{0,i} + \tau(\Sigma^{-1/2}z)_i, \theta_{0,i}, (\Sigma^{-1})_{ii}). \quad (\text{B.3.50})$$

Comparing with Eq. (B.3.9), this proves the claim that the standard distributional limit does indeed hold.

Notice that  $\tau^2$  is given by Eq. (B.3.48) that, for  $\mathbf{d} = 1/\zeta$  coincides with the claimed Eq. (4.2.12). Finally consider the scale parameter  $\mathbf{d} = \mathbf{d}(p)$  defined by Eq. (B.3.2). We claim that

$$\lim_{p \rightarrow \infty} \mathbf{d}(p) = \tilde{\mathbf{d}} = \frac{1}{\zeta}. \quad (\text{B.3.51})$$

Consider, for the sake of simplicity, the case that  $J$  is differentiable and strictly convex (the general case can be obtained as a limit). Then the minimum condition of the proximal operator (4.2.10) reads

$$\theta = \eta_b(Y) \Leftrightarrow b\Sigma(Y - \theta) = \nabla J(\theta). \quad (\text{B.3.52})$$

Differentiating with respect to  $\theta$ , and denoting by  $D\eta_b$  the Jacobian of  $\eta_b$ , we get  $D\eta_b(Y) = (\mathbf{I} + b^{-1}\Sigma^{-1}\nabla^2 J(\theta))^{-1}$  and hence

$$\mathbb{E}_2(a, b) = \lim_{p \rightarrow \infty} \frac{1}{p} \mathbb{E} \text{Trace} \left\{ (1 + b^{-1}\Sigma^{-1/2}\nabla^2 J(\hat{\theta})\Sigma^{-1/2})^{-1} \right\}, \quad (\text{B.3.53})$$

$$\hat{\theta} \equiv \eta_b(\theta_0 + \sqrt{a}\Sigma^{-1/2}z). \quad (\text{B.3.54})$$

Hence, combining Eqs. (B.3.49) and (B.3.53) implies that  $\tilde{\mathbf{d}} = \zeta^{-1}$  satisfies

$$1 = \frac{1}{\tilde{\mathbf{d}}} + \lim_{p \rightarrow \infty} \frac{1}{n} \mathbb{E} \text{Trace} \left\{ (1 + \tilde{\mathbf{d}}\Sigma^{-1/2}\nabla^2 J(\hat{\theta})\Sigma^{-1/2})^{-1} \right\}, \quad (\text{B.3.55})$$

$$\hat{\theta} \equiv \eta_{1/\tilde{\mathbf{d}}}(\theta_0 + \tau\Sigma^{-1/2}z). \quad (\text{B.3.56})$$

The claim (B.3.51) follows by comparing this with Eq. (B.3.2), and noting that, by the above  $\hat{\theta}$  is indeed asymptotically distributed as the estimator (B.3.1).

# Appendix C

## Supplement to Chapter 7

### C.1 Dependence of the algorithm on the prior $p_X$

In this appendix we briefly discuss the impact of a wrong estimation of the prior  $p_X$  on the AMP algorithm. Namely, suppose that instead of the true prior  $p_X$ , we have an approximation of  $p_X$  denoted by  $p_{\tilde{X}}$ . The only change in the algorithm is in the posterior expectation denoiser. That is to say, the denoiser  $\eta$  in Eq. (7.1.3) will be replaced by a new denoiser  $\tilde{\eta}$ . We will quantify the discrepancy between  $p_X$  and  $p_{\tilde{X}}$  through their Kolmogorov-Smirnov distance  $D_{\text{KS}}(p_X, p_{\tilde{X}})$ . Denoting by  $F_X(z) = p_X((-\infty, z])$  and  $F_{\tilde{X}}(z) = p_{\tilde{X}}((-\infty, z])$  the corresponding distribution functions, we have

$$D_{\text{KS}}(p_X, p_{\tilde{X}}) = \sup_{z \in \mathbb{R}} |F_X(z) - F_{\tilde{X}}(z)|.$$

The next lemma establishes a bound on the pointwise distance between  $\eta$  and  $\tilde{\eta}$  in terms of  $D_{\text{KS}}(p_X, p_{\tilde{X}})$ .

Note that state evolution (8.2.4) applies also to the algorithm with the mismatched denoiser, provided the  $\text{mmse}(\cdot)$  function is replaced by the mean square error for the non-optimal denoiser  $\tilde{\eta}$ . Hence the bound on  $|\eta(y) - \tilde{\eta}(y)|$  given below can be translated into a bound on the performance of AMP with the mismatched prior. A full study of this issue goes beyond the scope of this dissertation and will be the object of a forthcoming publication.

For the sake of simplicity we shall assume that  $p_X, p_{\tilde{X}}$  have bounded supports. The general case requires a more careful consideration.

**Lemma C.1.1.** *Let  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  be the Bayes optimal estimator for estimating  $X \sim p_X$  in Gaussian noise  $\eta(y) = \mathbb{E}(X|X + Z = y)$ , with  $Z \sim \mathbf{N}(0, 1)$ . Define denoiser  $\tilde{\eta}$  similarly, with respect to  $p_{\tilde{X}}$ . Assume that  $p_X$  is supported in  $[-M, M]$ . Then for any  $p_{\tilde{X}}$  supported in  $[-M, M]$ , we have*

$$|\eta(y) - \tilde{\eta}(y)| \leq \frac{M(15 + 10M|y|)}{\mathbb{E}\{e^{-X^2/2}\}} D_{\text{KS}}(p_X, p_{\tilde{X}}) e^{2M|y|}.$$

*Proof.* Throughout the proof we let  $\Delta \equiv D_{\text{KS}}(p_X, p_{\tilde{X}})$ , and  $\Delta_1 \equiv \mathbb{E}\{e^{-X^2/2}\}$ .

Let  $\gamma(z) = \exp(-z^2/2)/\sqrt{2\pi}$  be the Gaussian density. We then have  $\eta(y) = \mathbb{E}\{X\gamma(X - y)\}/\mathbb{E}\{\gamma(X - y)\}$ . Let  $p_W$  be the probability measure with Radon-Nikodym derivative with respect to  $p_X$  given by

$$\frac{dp_W}{dp_X}(x) = \frac{e^{-x^2/2}}{\mathbb{E}\{e^{-X^2/2}\}}.$$

We define  $p_{\tilde{W}}$  analogously from the measure  $p_{\tilde{X}}$  and let  $W, \tilde{W}$  be two random variables with law  $p_W$  and  $p_{\tilde{W}}$ , respectively. We then have

$$\eta(y) = \frac{\mathbb{E}\{W e^{yW}\}}{\mathbb{E}\{e^{yW}\}}. \quad (\text{C.1.1})$$

Letting  $F_W, F_{\tilde{W}}$  denote the corresponding distribution functions, we have

$$F_W(x) = \int_{-\infty}^x dp_W(w) = \frac{\int_{-\infty}^x e^{-z^2/2} dp_X(z)}{\mathbb{E}\{e^{-X^2/2}\}} = \frac{e^{-x^2/2} F_X(x) + \int_{-\infty}^x z e^{-z^2/2} F_X(z) dz}{\int_{-\infty}^{\infty} z e^{-z^2/2} F_X(z) dz}.$$

Letting  $N_W(x)$  be the numerator in this expression, we have

$$|N_W(x) - N_{\tilde{W}}(x)| \leq |F_X(x) - F_{\tilde{X}}(x)| + \int_{-\infty}^x |z| e^{-z^2/2} |F_X(z) - F_{\tilde{X}}(z)| dz \leq 3\Delta.$$

Proceeding analogously for the denominator, we have

$$|\mathbb{E}\{e^{-X^2/2}\} - \mathbb{E}\{e^{-\tilde{X}^2/2}\}| \leq \int_{-\infty}^{\infty} |z| e^{-z^2/2} |F_X(z) - F_{\tilde{X}}(z)| dz \leq 2\Delta.$$

Combining these bounds, we obtain

$$\begin{aligned}
|F_W(x) - F_{\widetilde{W}}(x)| &= \left| \frac{N_W(x)}{\mathbb{E}\{e^{-X^2/2}\}} - \frac{N_{\widetilde{W}}(x)}{\mathbb{E}\{e^{-\widetilde{X}^2/2}\}} \right| \\
&\leq \left| \frac{N_W(x) - N_{\widetilde{W}}(x)}{\mathbb{E}\{e^{-X^2/2}\}} \right| + \left| N_{\widetilde{W}}(x) \left( \frac{1}{\mathbb{E}\{e^{-X^2/2}\}} - \frac{1}{\mathbb{E}\{e^{-\widetilde{X}^2/2}\}} \right) \right| \\
&= \frac{|N_W(x) - N_{\widetilde{W}}(x)|}{\mathbb{E}\{e^{-X^2/2}\}} + F_{\widetilde{W}}(x) \frac{|\mathbb{E}\{e^{-X^2/2}\} - \mathbb{E}\{e^{-\widetilde{X}^2/2}\}|}{\mathbb{E}\{e^{-X^2/2}\}} \\
&\leq \frac{3\Delta}{\Delta_1} + \frac{2\Delta}{\Delta_1} = \frac{5\Delta}{\Delta_1}.
\end{aligned} \tag{C.1.2}$$

Since, the above inequality holds for any  $x \in \mathbb{R}$ , we get

$$D_{\text{KS}}(p_W, p_{\widetilde{W}}) \leq \frac{5\Delta}{\Delta_1}. \tag{C.1.3}$$

Consider now Eq. (C.1.1). We have

$$\begin{aligned}
|\mathbb{E}\{e^{yW}\} - \mathbb{E}\{e^{y\widetilde{W}}\}| &= |y| \int e^{yx} |F_W(x) - F_{\widetilde{W}}(x)| dx \\
&\leq |y| D_{\text{KS}}(p_W, p_{\widetilde{W}}) \int_{-M}^M e^{yx} dx \leq e^{M|y|} D_{\text{KS}}(p_W, p_{\widetilde{W}}).
\end{aligned}$$

We proceed analogously for the numerator, namely,

$$\begin{aligned}
|\mathbb{E}\{W e^{yW}\} - \mathbb{E}\{\widetilde{W} e^{y\widetilde{W}}\}| &= \int (1 + |yx|) e^{yx} |F_W(x) - F_{\widetilde{W}}(x)| dx \\
&\leq D_{\text{KS}}(p_W, p_{\widetilde{W}}) \int_{-M}^M (1 + |yx|) e^{yx} dx \leq 2M(1 + M|y|) e^{M|y|} D_{\text{KS}}(p_W, p_{\widetilde{W}}).
\end{aligned}$$

Combining these bounds and proceeding along similar lines to Eq. (C.1.2), we obtain

$$|\tilde{\eta}(y) - \eta(y)| \leq \frac{2M(1 + M|y|) + \tilde{\eta}(y)}{\mathbb{E}\{e^{yW}\}} e^{M|y|} D_{\text{KS}}(p_W, p_{\widetilde{W}}). \tag{C.1.4}$$

Note that  $\tilde{\eta}(y) \in [-M, M]$  since  $p_{\widetilde{X}}$  is supported on  $[-M, M]$ , and thus  $|\tilde{\eta}(y)| \leq M$ . Also,  $p_W$  is supported on  $[-M, M]$  since it is absolutely continuous with respect to  $p_X$  and  $p_X$  is supported on  $[-M, M]$ . Therefore,  $\mathbb{E}\{e^{yW}\} \geq e^{-M|y|}$ . Using these bounds in Eq. (C.1.4),

we obtain

$$|\tilde{\eta}(y) - \eta(y)| \leq M(3 + 2M|y|) e^{2M|y|} D_{\text{KS}}(p_W, p_{\tilde{W}}). \quad (\text{C.1.5})$$

The result follows by plugging in the bound given by Eq. (C.1.3).  $\square$

## C.2 Lipschitz continuity of AMP

Let  $x^t$  be the Bayes optimal AMP estimation at iteration  $t$  as given by Eqs. (7.1.3), (7.1.4). We show that for each fixed iteration number  $t$ , the mapping  $y \rightarrow x^t(y)$  is locally Lipschitz continuous.

**Lemma C.2.1.** *For any  $R, B > 0$ ,  $t \in \mathbb{N}$ , there exists  $L = L(R, B; t) < \infty$  such that for any  $y, \tilde{y} \in \mathbb{R}^m$  with  $\|y\|, \|\tilde{y}\| \leq R$ , and any matrix  $A$  with  $\|A\|_2 \leq B$  we have*

$$\|x^t(y) - x^t(\tilde{y})\| \leq L \|y - \tilde{y}\|. \quad (\text{C.2.1})$$

Note that in the statement we assume  $\|A\|_2$  to be finite. This happens as long as the entries of  $A$  are bounded and hence almost surely within our setting.

Also, we assume  $\|y\|, \|\tilde{y}\| \leq R$  for some fixed  $R$ . In other words, we prove that the algorithm is locally Lipschitz. We can obtain an algorithm that is globally Lipschitz by defining  $x^t(y)$  via the AMP iteration for  $\|y\| \leq R$ , and by an arbitrary bounded Lipschitz extension for  $\|y\| \geq R$ . Notice that  $\|y\| \leq B\|x\| + \|w\|$ , and, by the law of large numbers,  $\|x\|^2 \leq (\mathbb{E}\{X^2\} + \epsilon)n$ ,  $\|w\|^2 \leq (\sigma^2 + \epsilon)m$  with probability converging to 1. Hence, the globally Lipschitz modification of AMP achieves the same performance as the original AMP, almost surely. (Note that  $R$  can depend on  $n$ ).

*Proof (Lemma C.2.1).* Suppose that we have two measurement vectors  $y$  and  $\tilde{y}$ . Note that the state evolution is completely characterized in terms of prior  $p_X$  and noise variance  $\sigma^2$ , and can be precomputed (independent of measurement vector).

Let  $(x^t, r^t)$  correspond to the AMP with measurement vector  $y$  and  $(\tilde{x}^t, \tilde{r}^t)$  correspond to the AMP with measurement vector  $\tilde{y}$ . (To clarify, note that  $x^t \equiv x^t(y)$  and  $\tilde{x}^t \equiv x^t(\tilde{y})$ ). Further define

$$\xi_t = \max(\|x^t - \tilde{x}^t\|, \|r^t - \tilde{r}^t\|, \|y - \tilde{y}\|).$$

We show that

$$\xi_t \leq C_t(1 + \|y\|) \xi_{t-1}, \quad (\text{C.2.2})$$

for a constant  $C_t$ . This establishes the claim since

$$\|x^t - \tilde{x}^t\| \leq \xi_t \leq C_t C_{t-1} \dots C_2 (1 + \|y\|)^{t-1} \xi_1 = C_t C_{t-1} \dots C_2 (1 + \|y\|)^{t-1} \|y - \tilde{y}\|,$$

where the last step holds since  $x_i^1 = \tilde{x}_i^1 = \mathbb{E}\{X\}$  and  $r^1 - \tilde{r}^1 = y - \tilde{y}$ .

In order to prove Eq. (C.2.2), we need to prove the following two claims.

**Claim C.2.2.** *For any fixed iteration number  $t$ , there exists a constant  $C_t$ , such that*

$$\|r^t\| \leq C_t \max(\|x^1\|, \|y\|).$$

*Proof (Claim C.2.2).* Define  $\lambda_t = \max(\|x^{t+1}\|, \|r^t\|, \|y\|)$ . Then,

$$\|r^t\| \leq \|y\| + \|A\|_2 \|x^t\| + \|b^t\|_\infty \|r^{t-1}\|.$$

Note that  $A$  has bounded operator by assumption. Also, the posterior mean  $\eta$  is a smooth function with bounded derivative. Therefore, recalling the definition of  $b^t$ ,

$$b^t \equiv \frac{1}{\delta} \sum_{u \in \mathcal{C}} W_{g(i),u} \tilde{Q}_{g(i),u}^{t-1} \langle \eta'_{t-1} \rangle_u,$$

we have  $\|b^t\|_\infty \leq C_{1,t}$  for some constant  $C_{1,t}$ . Hence,  $\|r^t\| \leq C_{2,t} \lambda_{t-1}$ . Moreover,

$$\|x^{t+1}\| = \|\eta_t(x^t + (Q^t \odot A)^* r^t)\| \leq C(\|x^t\| + \|Q^t \odot A\|_2 \|r^t\|) \leq C_{3,t} \max(\|x^t\|, \|r^t\|),$$

for some constant  $C_{3,t}$ . In the first inequality, we used the fact that  $\eta$  is Lipschitz continuous.

Therefore,  $\lambda_t \leq C'_t \lambda_{t-1}$ , where  $C'_t = \max(1, C_{2,t}, C_{3,t}, C_{2,t} C_{3,t})$ , and

$$\|r^t\| \leq \lambda_t \leq C'_t \dots C'_1 \lambda_0 \leq C'_t \dots C'_1 \max(\|x^1\|, \|y\|),$$

with  $x_i^1 = \mathbb{E}\{X\}$ , for  $i \in [n]$ . □

**Claim C.2.3.** *For any fixed iteration number  $t$ , there exists a constant  $C_t$ , such that*

$$\|\mathbf{b}^t \odot r^{t-1} - \tilde{\mathbf{b}}^t \odot \tilde{r}^{t-1}\| \leq C_t(1 + \|y\|) \max(\|x^{t-1} - \tilde{x}^{t-1}\|, \|r^{t-1} - \tilde{r}^{t-1}\|).$$

*Proof (Claim C.2.3).* Using triangle inequality, we have

$$\|\mathbf{b}^t \odot r^{t-1} - \tilde{\mathbf{b}}^t \odot \tilde{r}^{t-1}\| \leq \|(\mathbf{b}^t - \tilde{\mathbf{b}}^t) \odot r^{t-1}\| + \|\tilde{\mathbf{b}}^t \odot (r^{t-1} - \tilde{r}^{t-1})\|. \quad (\text{C.2.3})$$

Since  $\eta'$  is Lipschitz continuous, we have

$$\|\mathbf{b}^t - \tilde{\mathbf{b}}^t\| \leq C_{1,t}(\|x^{t-1} - \tilde{x}^{t-1}\| + \|r^{t-1} - \tilde{r}^{t-1}\|),$$

for some constant  $C_{1,t}$ . Also, as discussed in the proof of Claim C.2.2, the Onsager terms  $\mathbf{b}^t$  are uniformly bounded. Applying these bounds to the right hand side of Eq. (C.2.3), we obtain

$$\begin{aligned} \|\mathbf{b}^t \odot r^{t-1} - \tilde{\mathbf{b}}^t \odot \tilde{r}^{t-1}\| &\leq C_{1,t}(\|x^{t-1} - \tilde{x}^{t-1}\| + \|r^{t-1} - \tilde{r}^{t-1}\|) \|r^{t-1}\| + C_{2,t} \|r^{t-1} - \tilde{r}^{t-1}\| \\ &\leq C_t(1 + \|y\|) \max(\|x^{t-1} - \tilde{x}^{t-1}\|, \|r^{t-1} - \tilde{r}^{t-1}\|), \end{aligned}$$

for some constants  $C_{1,t}, C_{2,t}, C_t$ . The last inequality here follows from the bound given in Claim C.2.2.  $\square$

Now, we are ready to prove Eq. (C.2.2). We write

$$\begin{aligned} \|x^t - \tilde{x}^t\| &= \|\eta_{t-1}(x^{t-1} + (Q^{t-1} \odot A)^* r^{t-1}) - \eta_{t-1}(\tilde{x}^{t-1} + (Q^{t-1} \odot A)^* \tilde{r}^{t-1})\| \\ &\leq C(\|x^{t-1} - \tilde{x}^{t-1}\| + \|Q^{t-1} \odot A\|_2 \|r^{t-1} - \tilde{r}^{t-1}\|) \\ &\leq C_{1,t} \max(\|x^{t-1} - \tilde{x}^{t-1}\|, \|r^{t-1} - \tilde{r}^{t-1}\|, \|y - \tilde{y}\|) = C_{1,t} \xi_{t-1}, \end{aligned} \quad (\text{C.2.4})$$

for some constant  $C_{1,t}$ . Furthermore,

$$\begin{aligned} \|r^t - \tilde{r}^t\| &\leq \|y - \tilde{y}\| + \|A\|_2 \|x^t - \tilde{x}^t\| + \|\mathbf{b}^t \odot r^{t-1} - \tilde{\mathbf{b}}^t \odot \tilde{r}^{t-1}\| \\ &\leq \|y - \tilde{y}\| + \|A\|_2 C_{1,t} \xi_{t-1} + C'_t(1 + \|y\|) \max(\|x^{t-1} - \tilde{x}^{t-1}\|, \|r^{t-1} - \tilde{r}^{t-1}\|) \\ &\leq C_{2,t}(1 + \|y\|) \xi_{t-1}, \end{aligned} \quad (\text{C.2.5})$$

for some constant  $C_{2,t}$  and using Eq. (C.2.4) and Claim C.2.3 in deriving the second inequality. Combining Eqs. (C.2.4) and (C.2.5), we obtain

$$\xi_t \leq \max(1, C_{1,t}, C_{2,t}) (1 + \|y\|) \xi_{t-1}.$$

□

## Appendix D

# Supplement to Chapter 11

### D.1 Proof of Lemma 11.4.2

We prove the first claim, Eq. (11.4.4). The second one follows by a similar argument. The proof uses induction on  $t$ . It is a simple exercise to show that the induction basis ( $t = 1$ ) holds (the calculation follows the same lines as the induction step). Assuming the claim for  $t$ , we write, for  $i \in \{0, 1, \dots, L - 1\}$

$$\begin{aligned}
 |\psi_i(t+1) - \psi(\rho i; t+1)| &= \left| \text{mmse} \left( \sum_{b \in \mathbb{R}_0} W_{b-i} [\sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi_j(t)]^{-1} \right) \right. \\
 &\quad \left. - \text{mmse} \left( \int_{-1}^{\ell+1} \mathcal{W}(z - \rho i) [\sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z - y) \psi(y; t) dy]^{-1} dz \right) \right| \\
 &\leq \left| \text{mmse} \left( \sum_{b \in \mathbb{R}_0} W_{b-i} [\sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi_j(t)]^{-1} \right) \right. \\
 &\quad \left. - \text{mmse} \left( \sum_{b \in \mathbb{R}_0} W_{b-i} [\sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi(\rho j; t)]^{-1} \right) \right| \tag{D.1.1} \\
 &\quad + \left| \text{mmse} \left( \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b - i)) [\sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} \rho \mathcal{W}(\rho(b - j)) \psi(\rho j; t)]^{-1} \right) \right. \\
 &\quad \left. - \text{mmse} \left( \int_{-1}^{\ell+1} \mathcal{W}(z - \rho i) [\sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z - y) \psi(y; t) dy]^{-1} dz \right) \right|.
 \end{aligned}$$

Now, we bound the two terms on the right hand side separately. Note that the arguments of  $\text{mmse}(\cdot)$  in the above terms are at most  $2/\sigma^2$ . Since  $\text{mmse}$  has a continuous derivative, there exists a constant  $C$  such that  $|\frac{d}{ds} \text{mmse}(s)| \leq C$ , for  $s \in [0, 2/\sigma^2]$ . Then, considering

the first term in the upper bound (D.1.1), we have

$$\begin{aligned}
& \left| \text{mmse} \left( \sum_{b \in \mathbf{R}_0} W_{b-i} \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi_j(t) \right]^{-1} \right) - \text{mmse} \left( \sum_{b \in \mathbf{R}_0} W_{b-i} \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi(\rho j; t) \right]^{-1} \right) \right| \\
& \leq C \left| \sum_{b \in \mathbf{R}_0} W_{b-i} \left( \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi_j(t) \right]^{-1} - \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} W_{b-j} \psi(\rho j; t) \right]^{-1} \right) \right| \\
& \leq \frac{C}{\sigma^4} \sum_{b \in \mathbf{R}_0} W_{b-i} \frac{1}{\delta} \left| \sum_{j=-\infty}^{L-1} W_{b-j} (\psi(\rho j; t) - \psi_j(t)) \right| \\
& \leq \frac{C}{\delta \sigma^4} \sum_{b \in \mathbf{R}_0} W_{b-i} \sum_{j=-\infty}^{L-1} W_{b-j} |\psi(\rho j; t) - \psi_j(t)| \\
& = \frac{C}{\delta \sigma^4} \sum_{j=0}^{L-1} \left( \sum_{b \in \mathbf{R}_0} W_{b-i} W_{b-j} \right) |\psi(\rho j; t) - \psi_j(t)| \\
& \leq \frac{C}{\delta \sigma^4} \left( \sum_{i \in \mathbb{Z}} W_i^2 \right) \sum_{j=0}^{L-1} |\psi(\rho j; t) - \psi_j(t)| \\
& \leq \frac{C' \rho}{\delta \sigma^4} \sum_{j=0}^{L-1} |\psi(\rho j; t) - \psi_j(t)|.
\end{aligned} \tag{D.1.2}$$

Here we used  $\sum_{i \in \mathbb{Z}} W_i^2 = \sum_{i \in \mathbb{Z}} \rho^2 \mathcal{W}(\rho i)^2 \leq C \sum_{|i| \leq \rho^{-1}} \rho^2 \leq C \rho$  (where the first inequality follows from the fact that  $\mathcal{W}$  is bounded).

To bound the second term in Eq. (D.1.1), note that

$$\begin{aligned}
& \left| \text{mmse} \left( \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} \rho \mathcal{W}(\rho(b-j)) \psi(\rho j; t) \right]^{-1} \right) \right. \\
& \quad \left. - \text{mmse} \left( \int_{-1}^{\ell+1} \mathcal{W}(z - \rho i) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z-y) \psi(y; t) dy \right]^{-1} dz \right) \right| \\
& \leq C \left| \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} \rho \mathcal{W}(\rho(b-j)) \psi(\rho j; t) \right]^{-1} \right. \\
& \quad \left. - \int_{-1}^{\ell+1} \mathcal{W}(z - \rho i) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z-y) \psi(y; t) dy \right]^{-1} dz \right| \\
& \leq C \left| \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left[ \sigma^2 + \frac{1}{\delta} \sum_{j \in \mathbb{Z}} \rho \mathcal{W}(\rho(b-j)) \psi(\rho j; t) \right]^{-1} \right. \\
& \quad \left. - \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(\rho b - y) \psi(y; t) dy \right]^{-1} \right| \tag{D.1.3} \\
& + C \left| \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(\rho b - y) \psi(y; t) dy \right]^{-1} dz \right. \\
& \quad \left. - \int_{-1}^{\ell+1} \mathcal{W}(z - \rho i) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z-y) \psi(y; t) dy \right]^{-1} dz \right| \\
& \leq \frac{C}{\delta \sigma^4} \sum_{b \in \mathbb{R}_0} \rho \mathcal{W}(\rho(b-i)) \left| \sum_{j \in \mathbb{Z}} \rho F_1(\rho b; \rho j) - \int_{\mathbb{R}} F_1(\rho b; y) dy \right| \\
& + C \left| \sum_{b \in \mathbb{R}_0} \rho F_2(\rho b) - \int_{-1}^{\ell+1} F_2(z) dz \right|
\end{aligned}$$

where  $F_1(x; y) = \mathcal{W}(x-y)\psi(y; t)$  and  $F_2(z) = \mathcal{W}(z - \rho i) \left[ \sigma^2 + \frac{1}{\delta} \int_{\mathbb{R}} \mathcal{W}(z-y)\psi(y; t) dy \right]^{-1}$ . Since the functions  $\mathcal{W}(\cdot)$  and  $\psi(\cdot)$  have continuous (and thus bounded) derivative on compact interval  $[0, \ell]$ , the same is true for  $F_1$  and  $F_2$ . Using the standard convergence of Riemann sums to Riemann integrals, right hand side of Eq. (D.1.3) can be bounded by  $C_3 \rho / \delta \sigma^4$ , for some constant  $C_3$ . Let  $\epsilon_i(t) = |\psi_i(t) - \psi(\rho i; t)|$ . Combining Eqs. (D.1.2) and (D.1.3), we get

$$\epsilon_i(t+1) \leq \frac{\rho}{\delta \sigma^4} \left( C' \sum_{j=0}^{L-1} \epsilon_j(t) + C_3 \right). \tag{D.1.4}$$

Therefore,

$$\frac{1}{L} \sum_{i=0}^{L-1} \epsilon_i(t+1) \leq \frac{\ell}{\delta\sigma^4} \left( \frac{C'}{L} \sum_{j=0}^{L-1} \epsilon_j(t) \right) + \frac{C_3\rho}{\delta\sigma^4}. \quad (\text{D.1.5})$$

The claims follows from the induction hypothesis.

## D.2 Proof of Proposition 11.4.9

By Eq. (11.4.8), for any  $\varepsilon > 0$ , there exists  $\phi_0$ , such that for  $0 \leq \phi \leq \phi_0$ ,

$$l(\phi^{-1}) \leq \frac{\bar{d}(p_X) + \varepsilon}{2} \log(\phi^{-1}). \quad (\text{D.2.1})$$

Therefore,

$$V(\phi) \leq \frac{\delta\sigma^2}{2\phi} + \frac{\delta - \bar{d}(p_X) - \varepsilon}{2} \log \phi. \quad (\text{D.2.2})$$

Now let  $\varepsilon = (\delta - \bar{d}(p_X))/2$  and  $\sigma_2 = \sqrt{\phi_0/2}$ . Hence, for  $\sigma \in (0, \sigma_2]$ , we get  $\phi^* < 2\sigma^2 \leq \phi_0$ . Plugging in  $\phi^*$  for  $\phi$  in the above equation, we get

$$\begin{aligned} V(\phi^*) &\leq \frac{\delta\sigma^2}{2\phi^*} + \frac{\delta - \bar{d}(p_X)}{4} \log \phi^* \\ &< \frac{\delta}{2} + \frac{\delta - \bar{d}(p_X)}{4} \log(2\sigma^2). \end{aligned} \quad (\text{D.2.3})$$

## D.3 Proof of Claim 11.4.11

Recall that  $\kappa < \Phi_M$  and  $\phi(x)$  is nondecreasing. Let

$$0 < \theta = \frac{\Phi_M - \kappa}{\Phi_M - \frac{\kappa}{2}} < 1.$$

We show that  $\phi(\theta\ell - 1) \geq \kappa/2 + \phi^*$ . If this is not true, using the nondecreasing property of  $\phi(x)$ , we obtain

$$\begin{aligned} \int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx &= \int_{-1}^{\theta\ell-1} |\phi(x) - \phi^*| dx + \int_{\theta\ell-1}^{\ell-1} |\phi(x) - \phi^*| dx \\ &< \frac{\kappa}{2}\theta\ell + \Phi_M(1-\theta)\ell \\ &= \kappa\ell, \end{aligned} \tag{D.3.1}$$

contradicting our assumption. Therefore,  $\phi(x) \geq \kappa/2 + \phi^*$ , for  $\theta\ell - 1 \leq x \leq \ell - 1$ . For given  $K$ , choose  $\ell_0 = K/(1 - \theta)$ . Hence, for  $\ell > \ell_0$ , interval  $[\theta\ell - 1, \ell - 1)$  has length at least  $K$ . The result follows.

## D.4 Proof of Proposition 11.4.12

We first establish some properties of function  $\zeta^2(x)$ .

**Remark D.4.1.** *The function  $\zeta^2(x)$  as defined in Eq. (11.4.10), is non increasing in  $x$ . Also,  $\zeta^2(x) = \sigma^2 + (1/\delta) \text{mmse}(L_0/(2\sigma^2))$ , for  $x \leq -1$  and  $\zeta^2(x) = \sigma^2$ , for  $x \geq 1$ . For  $\delta L_0 > 3$ , we have  $\sigma^2 \leq \zeta^2(x) < 2\sigma^2$ .*

**Remark D.4.2.** *The function  $\zeta^2(x)/\sigma^2$  is Lipschitz continuous. More specifically, there exists a constant  $C$ , such that,  $|\zeta^2(\alpha_1) - \zeta^2(\alpha_2)| < C\sigma^2|\alpha_2 - \alpha_1|$ , for any two values  $\alpha_1, \alpha_2$ . Further, if  $L_0\delta > 3$  we can take  $C < 1$ .*

The proof of Remarks D.4.1 and D.4.2 are immediate from Eq. (11.4.10).

To prove the proposition, we split the integral over the intervals  $[-1, -1 + a)$ ,  $[-1 + a, x_0 + a)$ ,  $[x_0 + a, x_2)$ ,  $[x_2, \ell - 1)$ , and bound each one separately. Firstly, note that

$$\int_{x_2}^{\ell-1} \left\{ \frac{\zeta^2(x) - \sigma^2}{\phi_a(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx = 0, \tag{D.4.1}$$

since  $\phi_a(x)$  and  $\phi(x)$  are identical for  $x \geq x_2$ .

Secondly, let  $\alpha = (x_2 - x_0)/(x_2 - x_0 - a)$ , and  $\beta = (ax_2)/(x_2 - x_0 - a)$ . Then,

$$\begin{aligned}
& \int_{x_0+a}^{x_2} \left\{ \frac{\zeta^2(x) - \sigma^2}{\phi_a(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx \\
&= \int_{x_0}^{x_2} \frac{\zeta^2\left(\frac{x+\beta}{\alpha}\right) - \sigma^2}{\phi(x)} \frac{dx}{\alpha} - \int_{x_0+a}^{x_2} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx \\
&= \int_{x_0}^{x_2} \left\{ \frac{1}{\alpha} \frac{\zeta^2\left(\frac{x+\beta}{\alpha}\right) - \sigma^2}{\phi(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx + \int_{x_0}^{x_0+a} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx \\
&\stackrel{(a)}{\leq} \frac{1}{\sigma^2} \int_{x_0}^{x_2} \left| \frac{1}{\alpha} \zeta^2\left(\frac{x+\beta}{\alpha}\right) - \zeta^2(x) \right| dx + \left(1 - \frac{1}{\alpha}\right) \int_{x_0}^{x_2} \frac{\sigma^2}{\phi(x)} dx + \int_{x_0}^{x_0+a} \frac{\sigma^2}{\phi(x)} dx \quad (\text{D.4.2}) \\
&\leq \frac{1}{\sigma^2} \int_{x_0}^{x_2} \left(1 - \frac{1}{\alpha}\right) \zeta^2\left(\frac{x+\beta}{\alpha}\right) dx + \frac{1}{\sigma^2} \int_{x_0}^{x_2} \left| \zeta^2\left(\frac{x+\beta}{\alpha}\right) - \zeta^2(x) \right| dx + \frac{K}{2} \left(1 - \frac{1}{\alpha}\right) + a \\
&\leq \left(1 - \frac{1}{\alpha}\right) K + \frac{1}{\sigma^2} \int_{x_0}^{x_2} \left| \zeta^2\left(\frac{x+\beta}{\alpha}\right) - \zeta^2(x) \right| dx + \frac{K}{2} \left(1 - \frac{1}{\alpha}\right) + a \\
&\stackrel{(b)}{\leq} \left(1 - \frac{1}{\alpha}\right) K + CK^2 \left(1 - \frac{1}{\alpha}\right) + CKa + \frac{K}{2} \left(1 - \frac{1}{\alpha}\right) + a \\
&\leq C(K)a,
\end{aligned}$$

where (a) follows from the fact  $\sigma^2 \leq \phi(x)$  and Remark D.4.1; (b) follows from Remark D.4.2.

Thirdly, recall that  $\phi_a(x) = \phi(x - a)$ , for  $x \in [-1 + a, x_0 + a]$ . Therefore,

$$\begin{aligned}
& \int_{-1+a}^{x_0+a} \left\{ \frac{\zeta^2(x) - \sigma^2}{\phi_a(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx \\
&= \int_{-1}^{x_0} \frac{\zeta^2(x+a) - \sigma^2}{\phi(x)} dx - \int_{-1+a}^{x_0+a} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx \\
&= \int_{-1}^{x_0} \frac{\zeta^2(x+a) - \zeta^2(x)}{\phi(x)} dx - \int_{x_0}^{x_0+a} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx + \int_{-1}^{-1+a} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} dx \quad (\text{D.4.3}) \\
&\leq 0 + 0 + \int_{-1}^{-1+a} \frac{\sigma^2}{\phi(x)} dx \\
&\leq a,
\end{aligned}$$

where the first inequality follows from Remark D.4.1 and the second follows from  $\phi(x) \geq \sigma^2$ .

Finally, using the facts  $\sigma^2 \leq \zeta^2(x) \leq 2\sigma^2$ , and  $\sigma^2 \leq \phi(x)$ , we have

$$\int_{-1}^{-1+a} \left\{ \frac{\zeta^2(x) - \sigma^2}{\phi_a(x)} - \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \right\} dx \leq a. \quad (\text{D.4.4})$$

Combining Eqs. (D.4.1), (D.4.2), (D.4.3), and (D.4.4) implies the desired result.

## D.5 Proof of Proposition 11.4.13

*Proof.* Let  $\tilde{\mathbb{E}}_{\mathcal{W}}(\phi_a) = \tilde{\mathbb{E}}_{\mathcal{W},1}(\phi_a) + \tilde{\mathbb{E}}_{\mathcal{W},2}(\phi_a) + \tilde{\mathbb{E}}_{\mathcal{W},3}(\phi_a)$ , where

$$\begin{aligned}\tilde{\mathbb{E}}_{\mathcal{W},1}(\phi_a) &= \int_{x_0+a}^{\ell-1} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy, \\ \tilde{\mathbb{E}}_{\mathcal{W},2}(\phi_a) &= \int_a^{x_0+a} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy, \\ \tilde{\mathbb{E}}_{\mathcal{W},3}(\phi_a) &= \int_0^a \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy.\end{aligned}\tag{D.5.1}$$

Also let  $\tilde{\mathbb{E}}_{\mathcal{W}}(\phi) = \tilde{\mathbb{E}}_{\mathcal{W},1}(\phi) + \tilde{\mathbb{E}}_{\mathcal{W},2,3}(\phi)$ , where

$$\begin{aligned}\tilde{\mathbb{E}}_{\mathcal{W},1}(\phi) &= \int_{x_0+a}^{\ell-1} \{\mathbb{I}(\mathcal{W} * \phi(y)^{-1}) - \mathbb{I}(\phi(y-1)^{-1})\} dy, \\ \tilde{\mathbb{E}}_{\mathcal{W},2,3}(\phi) &= \int_0^{x_0+a} \{\mathbb{I}(\mathcal{W} * \phi(y)^{-1}) - \mathbb{I}(\phi(y-1)^{-1})\} dy.\end{aligned}\tag{D.5.2}$$

The following remark is used several times in the proof.

**Remark D.5.1.** For any two values  $0 \leq \alpha_1 < \alpha_2$ ,

$$\mathbb{I}(\alpha_2) - \mathbb{I}(\alpha_1) = \int_{\alpha_1}^{\alpha_2} \frac{1}{2} \text{mmse}(z) dz \leq \int_{\alpha_1}^{\alpha_2} \frac{1}{2z} dz = \frac{1}{2} \log\left(\frac{\alpha_2}{\alpha_1}\right) \leq \frac{1}{2} \left(\frac{\alpha_2}{\alpha_1} - 1\right).\tag{D.5.3}$$

• Bounding  $\tilde{\mathbb{E}}_{\mathcal{W},1}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W},1}(\phi)$ .

Notice that the functions  $\phi(x) = \phi_a(x)$ , for  $x_2 \leq x$ . Also  $\kappa/2 < \phi_a(x) \leq \phi(x) \leq \Phi_M$ , for  $x_1 < x < x_2$ . Let  $\alpha = (x_2 - x_1)/(x_2 - x_1 - a)$ , and  $\beta = (ax_2)/(x_2 - x_1 - a)$ . Then,

$\phi_a(x) = \phi(\alpha x - \beta)$  for  $x \in [x_0 + a, x_2]$ . Hence,

$$\begin{aligned}
& \tilde{\mathbb{E}}_{\mathcal{W},1}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W},1}(\phi) \\
&= \int_{x_0+a}^{x_2+1} \mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\mathcal{W} * \phi(y)^{-1}) \, dy + \int_{x_0+a}^{x_2+1} \mathbb{I}(\phi(y-1)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1}) \, dy \\
&\leq \frac{1}{2} \int_{x_0+a}^{x_2+1} \frac{1}{\mathcal{W} * \phi(y)^{-1}} (\mathcal{W} * \phi_a(y)^{-1} - \mathcal{W} * \phi(y)^{-1}) \, dy \\
&\leq \frac{\Phi_M}{2} \int_{x_0+a}^{x_2+1} \left( \int_{x_0+a-1}^{x_2} \mathcal{W}(y-z) \phi_a(z)^{-1} \, dz - \int_{x_0+a-1}^{x_2} \mathcal{W}(y-z) \phi(z)^{-1} \, dz \right) dy \\
&= \frac{\Phi_M}{2} \int_{x_0+a}^{x_2+1} \left( \int_{x_0+a}^{x_2} \mathcal{W}(y-z) \phi(\alpha z - \beta)^{-1} \, dz + \int_{x_0+a-1}^{x_0+a} \mathcal{W}(y-z) \phi(z-a)^{-1} \, dz \right. \\
&\quad \left. - \int_{x_0+a-1}^{x_2} \mathcal{W}(y-z) \phi(z)^{-1} \, dz \right) dy \\
&\leq \frac{\Phi_M}{2} \int_{x_0+a}^{x_2+1} \left\{ \int_{x_0}^{x_2} \left( \frac{1}{\alpha} \mathcal{W}\left(y - \frac{z+\beta}{\alpha}\right) - \mathcal{W}(y-z) \right) \phi(z)^{-1} \, dz \right. \\
&\quad \left. + \int_{x_0-1}^{x_0} \left( \mathcal{W}(y-z-a) - \mathcal{W}(y-z) \right) \phi(z)^{-1} \, dz \right. \\
&\quad \left. + \int_{x_0-1}^{x_0+a-1} \mathcal{W}(y-z) \phi(z)^{-1} \, dz \right\} dy \\
&\leq \frac{\Phi_M}{2} \int_{x_0+a}^{x_2+1} \left\{ \int_{x_0}^{x_2} \left( \mathcal{W}\left(y - \frac{z+\beta}{\alpha}\right) - \mathcal{W}(y-z) \right) \phi(z)^{-1} \, dz \right. \\
&\quad \left. + \int_{x_0-1}^{x_0} \left( \mathcal{W}(y-z-a) - \mathcal{W}(y-z) \right) \phi(z)^{-1} \, dz \right. \\
&\quad \left. + \int_{x_0-1}^{x_0+a-1} \mathcal{W}(y-z) \phi(z)^{-1} \, dz \right\} dy \\
&\leq C_1 \left(1 - \frac{1}{\alpha}\right) + C_2 \frac{\beta}{\alpha} + C_3 a \leq C_4 a. \tag{D.5.4}
\end{aligned}$$

Here  $C_1, C_2, C_3, C_4$  are some constants that depend only on  $K$  and  $\kappa$ . The last step follows from the facts that  $\mathcal{W}(\cdot)$  is a bounded Lipschitz function and  $\phi(z)^{-1} \leq 2/\kappa$  for  $z \in [x_1, x_2]$ . Also, note that in the first inequality,  $\mathbb{I}(\phi(y-1)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1}) \leq 0$ , since  $\phi(y-1)^{-1} \leq \phi_a(y-1)^{-1}$ , and  $\mathbb{I}(\cdot)$  is nondecreasing.

- Bounding  $\tilde{\mathbb{E}}_{\mathcal{W},2}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W},2,3}(\phi)$ .

We have

$$\begin{aligned} \tilde{E}_{\mathcal{W},2}(\phi_a) &= \int_{x_0+a-1}^{x_0+a} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy \\ &\quad + \int_a^{x_0+a-1} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy. \end{aligned} \tag{D.5.5}$$

We treat each term separately. For the first term,

$$\begin{aligned} &\int_{x_0+a-1}^{x_0+a} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy \\ &= \int_{x_0+a-1}^{x_0+a} \left\{ \mathbb{I} \left( \int_{x_0+a}^{x_0+a+1} \mathcal{W}(y-z)\phi_a(z)^{-1} dz + \int_{x_0+a-2}^{x_0+a} \mathcal{W}(y-z)\phi_a(z)^{-1} dz \right) - \mathbb{I}(\phi_a(y-1)^{-1}) \right\} dy \\ &= \int_{x_0+a-1}^{x_0+a} \left\{ \mathbb{I} \left( \int_{x_0}^{x_0+\alpha} \mathcal{W}(y - \frac{z+\beta}{\alpha})\phi(z)^{-1} \frac{dz}{\alpha} + \int_{x_0-2}^{x_0} \mathcal{W}(y-a-z)\phi(z)^{-1} dz \right) \right. \\ &\quad \left. - \int_{x_0-1}^{x_0} \mathbb{I}(\phi(y-1)^{-1}) dy \right\} \\ &= \int_{x_0-1}^{x_0} \left\{ \mathbb{I} \left( \int_{x_0}^{x_0+\alpha} \mathcal{W}(y+a - \frac{z+\beta}{\alpha})\phi(z)^{-1} \frac{dz}{\alpha} + \int_{x_0-2}^{x_0} \mathcal{W}(y-z)\phi(z)^{-1} dz \right) \right. \\ &\quad \left. - \int_{x_0-1}^{x_0} \mathbb{I}(\phi(y-1)^{-1}) dy \right\} \\ &\leq C_5 a + \int_{x_0-1}^{x_0} \left\{ \mathbb{I} \left( \int_{x_0-2}^{x_0+1} \mathcal{W}(y-z)\phi(z)^{-1} dz \right) - \mathbb{I}(\phi(y-1)^{-1}) \right\} dy \\ &= C_5 a + \int_{x_0-1}^{x_0} \left\{ \mathbb{I}(\mathcal{W} * \phi(y)^{-1}) - \mathbb{I}(\phi(y-1)^{-1}) \right\} dy, \end{aligned} \tag{D.5.6}$$

where the last inequality is an application of remark D.5.1. More specifically,

$$\begin{aligned}
& \mathbb{I} \left( \int_{x_0}^{x_0+\alpha} \mathcal{W}(y+a - \frac{z+\beta}{\alpha}) \phi(z)^{-1} \frac{dz}{\alpha} + \int_{x_0-2}^{x_0} \mathcal{W}(y-z) \phi(z)^{-1} dz \right) \\
& \quad - \mathbb{I} \left( \int_{x_0-2}^{x_0+1} \mathcal{W}(y-z) \phi(z)^{-1} dz \right) \\
& \leq \frac{\Phi_M}{2} \left( \int_{x_0}^{x_0+\alpha} \mathcal{W}(y+a - \frac{z+\beta}{\alpha}) \phi(z)^{-1} \frac{dz}{\alpha} - \int_{x_0}^{x_0+1} \mathcal{W}(y-z) \phi(z)^{-1} dz \right) \\
& \leq \frac{\Phi_M}{2} \int_{x_0+1}^{x_0+\alpha} \mathcal{W}(y+a - \frac{z+\beta}{\alpha}) \phi(z)^{-1} dz \\
& \quad + \frac{\Phi_M}{2} \int_{x_0}^{x_0+1} \left( \mathcal{W}(y+a - \frac{z+\beta}{\alpha}) - \mathcal{W}(y-z) \right) \phi(z)^{-1} dz \\
& \leq C'_1 \left(1 - \frac{1}{\alpha}\right) + C'_2 \frac{\beta}{\alpha} + C'_3 a \leq C_5 a,
\end{aligned}$$

where  $C'_1, C'_2, C'_3, C_5$  are constants that depend only on  $\kappa$ . Here, the penultimate inequality follows from  $\alpha > 1$ , and the last one follows from the fact that  $\mathcal{W}(\cdot)$  is a bounded Lipschitz function and that  $\phi(z)^{-1} \leq 2/\kappa$ , for  $z \in [x_1, x_2]$ .

To bound the second term on the right hand side of Eq. (D.5.6), notice that  $\phi_a(z) = \phi(z-a)$ , for  $z \in [-1+a, x_0+a)$ , whereby

$$\begin{aligned}
& \int_a^{x_0+a-1} \{\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) - \mathbb{I}(\phi_a(y-1)^{-1})\} dy \\
& = \int_0^{x_0-1} \{\mathbb{I}(\mathcal{W} * \phi(y)^{-1}) - \mathbb{I}(\phi(y-1)^{-1})\} dy.
\end{aligned} \tag{D.5.7}$$

Now, using Eqs. (D.5.2), (D.5.5) and (D.5.7), we obtain

$$\begin{aligned}
\tilde{\mathbb{E}}_{\mathcal{W},2}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W},2,3}(\phi) & \leq C_5 a - \int_{x_0}^{x_0+a} \{\mathbb{I}(\mathcal{W} * \phi(y)^{-1}) - \mathbb{I}(\phi(y-1)^{-1})\} dy \\
& \leq C_5 a + \int_{x_0}^{x_0+a} \log \left( \frac{\phi(y-1)^{-1}}{\mathcal{W} * \phi(y)^{-1}} \right) dy \\
& \leq C_5 a + a \log \left( \frac{\Phi_M}{\kappa} \right) = C_6 a,
\end{aligned} \tag{D.5.8}$$

where  $C_6$  is a constant that depends only on  $\kappa$ .

- Bounding  $\tilde{\mathbb{E}}_{\mathcal{W},3}(\phi_a)$ .

Notice that  $\phi_a(y) \geq \sigma^2$ . Therefore,  $\mathbb{I}(\mathcal{W} * \phi_a(y)^{-1}) \leq \mathbb{I}(\sigma^{-2})$ , since  $\mathbb{I}(\cdot)$  is nondecreasing.

Recall that  $\phi_a(y) = \phi^* < 2\sigma^2$ , for  $y \in [-1, -1 + a)$ . Consequently,

$$\tilde{\mathbb{E}}_{\mathcal{W},3}(\phi_a) \leq \int_0^a \{I(\sigma^{-2}) - I(\phi^{*-1})\} dy \leq \frac{a}{2} \log\left(\frac{\phi^*}{\sigma^2}\right) < \frac{a}{2} \log 2, \quad (\text{D.5.9})$$

where the first inequality follows from Remark D.5.1.

Finally, we are in position to prove the proposition. Using Eqs. (D.5.4), (D.5.8) and (D.5.9), we get

$$\tilde{\mathbb{E}}_{\mathcal{W}}(\phi_a) - \tilde{\mathbb{E}}_{\mathcal{W}}(\phi) \leq C_4 a + C_6 a + \frac{a}{2} \log 2 = C(\kappa, K) a. \quad (\text{D.5.10})$$

□

## D.6 Proof of Proposition 11.4.14

We have

$$\begin{aligned} \int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx &= \int_{x_2}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx \\ &+ \left( \int_{x_0+a}^{x_2} V(\phi_a(x)) dx - \int_{x_0}^{x_2} V(\phi(x)) dx \right) \\ &+ \left( \int_{-1+a}^{x_0+a} V(\phi_a(x)) dx - \int_{-1}^{x_0} V(\phi(x)) dx \right) \\ &+ \int_{-1}^{-1+a} V(\phi_a(x)) dx. \end{aligned} \quad (\text{D.6.1})$$

Notice that the first and the third terms on the right hand side are zero. Also,

$$\begin{aligned} \int_{x_0+a}^{x_2} V(\phi_a(x)) dx - \int_{x_0}^{x_2} V(\phi(x)) dx &= -\frac{a}{x_2 - x_0} \int_{x_0}^{x_2} V(\phi(x)) dx, \\ \int_{-1}^{-1+a} V(\phi_a(x)) dx &= aV(\phi^*), \end{aligned} \quad (\text{D.6.2})$$

where the second equation holds because  $\phi_a(x) = \phi^*$  for  $x \in [-1, -1 + a)$  in view of Eq. (11.4.20).

Substituting Eq. (D.6.2) in Eq. (D.6.1), we get

$$\int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx = \frac{a}{x_2 - x_0} \int_{x_0}^{x_2} \{V(\phi^*) - V(\phi(x))\} dx. \quad (\text{D.6.3})$$

Now we upper bound the right hand side of Eq. (D.6.3).

By Proposition 11.4.9, we have

$$V(\phi^*) \leq \frac{\delta}{2} + \frac{\delta - \bar{d}(p_X)}{4} \log(2\sigma^2), \quad (\text{D.6.4})$$

for  $\sigma \in (0, \sigma_2]$ . Also, since  $\phi(x) > \kappa/2$  for  $x \in [x_0, x_2]$ , we have  $V(\phi(x)) \geq (\delta/2) \log \phi > (\delta/2) \log(\kappa/2)$ . Therefore,

$$\begin{aligned} \frac{1}{2} \int_{-1}^{\ell-1} \{V(\phi_a(x)) - V(\phi(x))\} dx &= \frac{a}{2(x_2 - x_0)} \int_{x_0}^{x_2} \{V(\phi^*) - V(\phi(x))\} dx \\ &< \frac{a}{2} \left[ \frac{\delta}{2} + \frac{\delta - \bar{d}(p_X)}{4} \log(2\sigma^2) - \frac{\delta}{2} \log\left(\frac{\kappa}{2}\right) \right]. \end{aligned} \quad (\text{D.6.5})$$

It is now obvious that by choosing  $\sigma_0 > 0$  small enough, we can ensure that for values  $\sigma \in (0, \sigma_0]$ ,

$$\frac{a}{2} \left[ \frac{\delta}{2} + \frac{\delta - \bar{d}(p_X)}{4} \log(2\sigma^2) - \frac{\delta}{2} \log\left(\frac{\kappa}{2}\right) \right] < -2C(\kappa, K)a. \quad (\text{D.6.6})$$

(Notice that the right hand side of Eq. (D.6.6) does not depend on  $\sigma$ ).

## D.7 Proof of Claim 11.4.16

Similar to the proof of Claim 11.4.11, the assumption  $\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx > C\sigma^2\ell$  implies  $\phi(\theta\ell - 1) > C\sigma^2(1 - \alpha)$ , where

$$0 < \theta = \frac{\Phi_M - C\sigma^2}{\Phi_M - C\sigma^2(1 - \alpha)} < 1.$$

Choose  $\sigma$  small enough such that  $\phi^* < \phi_1$ . Let  $\kappa = (\phi_1 - \phi^*)(1 - \theta)/2$ . Applying Lemma 11.4.10, there exists  $\ell_0$ , and  $\sigma_0$ , such that,  $\int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx \leq \kappa\ell$ , for  $\ell > \ell_0$  and  $\sigma \in (0, \sigma_0]$ . We claim that  $\phi(\mu\ell - 1) < \phi_1$ , with

$$\mu = 1 - \frac{\kappa}{\phi_1 - \phi^*} = \frac{1 + \theta}{2}.$$

Otherwise, by monotonicity of  $\phi(x)$ ,

$$(\phi_1 - \phi^*)(1 - \mu)\ell \leq \int_{\mu\ell-1}^{\ell-1} |\phi(x) - \phi^*| dx < \int_{-1}^{\ell-1} |\phi(x) - \phi^*| dx \leq \kappa\ell. \quad (\text{D.7.1})$$

Plugging in for  $\mu$  yields a contradiction.

Therefore,  $C\sigma^2(1 - \alpha) < \phi(x) < \phi_1$ , for  $x \in [\theta\ell - 1, \mu\ell - 1]$ , and  $(\mu - \theta)\ell = (1 - \theta)\ell/2$ . Choosing  $\ell > \max\{\ell_0, 2K/(1 - \theta)\}$  gives the result.

## D.8 Proof of Proposition 11.4.17

To prove Eq. (11.4.32), we write

$$\begin{aligned} \int_{-1}^{\ell-1} \{V_{\text{rob}}(\phi_a(x)) - V_{\text{rob}}(\phi(x))\} dx &= - \int_{x_1}^{x_2} \int_{\phi_a(x)}^{\phi(x)} V'(s) ds dx \\ &\leq - \int_{x_1}^{x_2} \int_{\phi_a(x)}^{\phi(x)} \frac{\delta}{2s^2} (s - \sigma^2) ds dx \\ &= - \frac{\delta}{2} \int_{x_1}^{x_2} \left\{ \log \left( \frac{\phi(x)}{\phi_a(x)} \right) + \frac{\sigma^2}{\phi(x)} - \frac{\sigma^2}{\phi_a(x)} \right\} dx \\ &\leq \frac{\delta}{2} K \log(1 - a) + K \frac{\delta a}{2C(1 - \alpha)(1 - a)}, \end{aligned} \quad (\text{D.8.1})$$

where the second inequality follows from the fact  $C\sigma^2/2 < \phi(x)$ , for  $x \in [x_1, x_2]$ .

Next, we pass to prove Eq. (11.4.33).

$$\begin{aligned} \int_{-1}^{\ell-1} (\zeta^2(x) - \sigma^2) \left( \frac{1}{\phi_a(x)} - \frac{1}{\phi(x)} \right) dx &= \int_{x_1}^{x_2} \frac{\zeta^2(x) - \sigma^2}{\phi(x)} \left( \frac{1}{1 - a} - 1 \right) \\ &\leq \frac{a}{1 - a} \int_{x_1}^{x_2} \frac{\sigma^2}{\phi(x)} dx \leq K \frac{a}{C(1 - \alpha)(1 - a)}, \end{aligned} \quad (\text{D.8.2})$$

where the first inequality follows from Remark D.4.1.

Finally, we have

$$\begin{aligned}
\tilde{E}_{\mathcal{W},\text{rob}}(\phi_a) - \tilde{E}_{\mathcal{W},\text{rob}}(\phi) &= \int_0^\ell \{I(\mathcal{W} * \phi_a(y)^{-1}) - I(\mathcal{W} * \phi(y)^{-1})\} dy \\
&= \int_0^\ell \int_{\mathcal{W} * \phi(y)^{-1}}^{\mathcal{W} * \phi_a(y)^{-1}} \frac{1}{2} \text{mmse}(s) ds dy \\
&\leq \frac{\overline{D}(p_X) + \varepsilon}{2} \int_0^\ell \int_{\mathcal{W} * \phi(y)^{-1}}^{\mathcal{W} * \phi_a(y)^{-1}} s^{-1} ds dy \\
&\leq \frac{\overline{D}(p_X) + \varepsilon}{2} \int_0^\ell \log \left( \frac{\mathcal{W} * \phi_a(y)^{-1}}{\mathcal{W} * \phi(y)^{-1}} \right) dy \\
&\leq -\frac{\overline{D}(p_X) + \varepsilon}{2} (K + 2) \log(1 - a),
\end{aligned} \tag{D.8.3}$$

where the first inequality follows from Eq. (11.4.29) and Claim 11.4.16.

# Bibliography

- [1] S. Aeron, V. Saligrama, and M. Zhao. Information theoretic bounds for compressed sensing. *IEEE Trans. on Inform. Theory*, 56:5111 – 5130, 2010. [128](#)
- [2] M. Aizenman, R. Sims, and S. L. Starr. Extended variational principle for the sherrington-kirkpatrick spin-glass model. *Physical Review B*, 68(21):214403, 2003. [70](#)
- [3] K. Do Ba, P. Indyk, E. Price, and D. P. Woodruff. Lower bounds for sparse recovery. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '10, pages 1190–1197, 2010. [129](#)
- [4] D. Baron, S. Sarvotham, and R. Baraniuk. Bayesian Compressive Sensing Via Belief Propagation. *IEEE Trans. on Signal Proc.*, 58:269–280, 2019. [127](#), [131](#)
- [5] M. Bayati, M. A. Erdogdu, and A. Montanari. Estimating Lasso risk and noise level. In *Advances in Neural Information Processing Systems*, pages 944–952, 2013. [40](#)
- [6] M. Bayati, M. Lelarge, and A. Montanari. Universality in message passing algorithms. arXiv:1207.7321, 2012. [69](#), [70](#), [122](#), [182](#)
- [7] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. on Inform. Theory*, 57:764–785, 2011. [24](#), [122](#), [136](#), [148](#), [154](#), [156](#), [182](#), [186](#)
- [8] M. Bayati and A. Montanari. The LASSO risk for gaussian matrices. *IEEE Trans. on Inform. Theory*, 58:1997–2017, 2012. [44](#), [61](#), [65](#), [69](#), [70](#), [72](#), [85](#), [86](#), [120](#), [147](#), [192](#), [193](#)

- [9] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Img. Sci.*, 2(1):183–202, 2009. 15
- [10] R. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961. 4
- [11] A. Belloni and V. Chernozhukov. Least squares after model selection in high-dimensional sparse models. *Bernoulli*, 19(2):521–547, 2013. 40
- [12] R. Berinde, A. C. Gilbert, P. Indyk, H. Karloff, and M. J. Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *47th Annual Allerton Conference*, pages 798 – 805, Monticello, IL, September 2008. 128
- [13] P. J. Bickel and E. Levina. Regularized estimation of large covariance matrices. *The Annals of Statistics*, pages 199–227, 2008. 71
- [14] P. J. Bickel, Y. Ritov, and A. B. Tsybakov. Simultaneous analysis of Lasso and Dantzig selector. *The Annals of Statistics*, 37:1705–1732, 2009. 9, 34, 60, 99
- [15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2011. 15
- [16] P. Bühlmann. Statistical significance in high-dimensional linear models. *Bernoulli*, 19(4):1212–1242, 2013. xiii, xiv, xvi, 17, 58, 59, 81, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 97
- [17] P. Bühlmann, M. Kalisch, and L. Meier. High-dimensional statistics with a view toward applications in biology. *Annual Review of Statistics and Its Application*, 1(1):255–278, 2014. 82, 84
- [18] P. Bühlmann and J. Mandozzi. High-dimensional variable screening and bias in subsequent inference, with an empirical comparison. *Computational Statistics*, 29:407–430, 2014. 56
- [19] P. Bühlmann and S. van de Geer. *Statistics for high-dimensional data*. Springer-Verlag, 2011. 28, 33, 38, 98

- [20] T. Cai, W. Liu, and X. Luo. A constrained  $\ell_1$  minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association*, 106(494):594–607, 2011. 71
- [21] A. T. Campo, A. Guillen i Fabregas, and E. Biglieri. Large-system analysis of multiuser detection with an unknown number of users: A high-snr approach. *Information Theory, IEEE Transactions on*, 57(6):3416–3428, 2011. 70
- [22] E. Candès and M. Davenport. How well can we estimate a sparse vector? arXiv:1104.5246v3, 2011. 128
- [23] E. Candès, J. K. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. on Inform. Theory*, 52:489 – 509, 2006. 20, 177, 178, 185
- [24] E. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59:1207–1223, 2006. 20
- [25] E. Candès and T. Tao. The Dantzig selector: statistical estimation when  $p$  is much larger than  $n$ . *The Annals of Statistics*, 35:2313–2351, 2007. 9, 60
- [26] E. J. Candès and Y. Plan. Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Transactions on Information Theory*, 57(4):2342–2359, 2011. 8
- [27] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 9(6):717–772, 2009. 8
- [28] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. on Inform. Theory*, 51:4203–4215, 2005. 20
- [29] O. Chapelle, B. Schölkopf, A. Zien, et al. *Semi-supervised learning*. MIT Press, Cambridge, 2006. 71
- [30] S. S. Chen and D. L. Donoho. Examples of basis pursuit. In *Proceedings of Wavelet Applications in Signal and Image Processing III*, San Diego, CA, 1995. 6, 20

- [31] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998. 20
- [32] L. H. Dicker. Residual variance and the signal-to-noise ratio in high-dimensional linear models. *arXiv preprint arXiv:1209.0012*, 2012. 40
- [33] D. Donoho, I. Johnstone, and A. Montanari. Accurate Prediction of Phase Transitions in Compressed Sensing via a Connection to Minimax Denoising. *arXiv:1111.1041*, 2011. 122, 130, 131
- [34] D. Donoho and J. Tanner. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1906):4273–4293, 2009. 23, 24
- [35] D. L. Donoho. Aide-memoire. high-dimensional data analysis: The curses and blessings of dimensionality, 2000. 5
- [36] D. L. Donoho. Compressed sensing. *IEEE Trans. on Inform. Theory*, 52:489–509, April 2006. 20, 128
- [37] D. L. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete Comput. Geometry*, 35:617–652, 2006. 20, 61, 128
- [38] D. L. Donoho and M. Elad. Maximal sparsity representation via  $l_1$  minimization. *Proceedings of the National Academy of Sciences*, 100:2197–2202, March 2003. 20
- [39] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845–2862, 2001. 32
- [40] D. L. Donoho, A. Javanmard, and A. Montanari. Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1231–1235, 2012. 25
- [41] D. L. Donoho, A. Javanmard, and A. Montanari. Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing. *IEEE Transactions on Information Theory*, 59(11):7434–7464, 2013. 25, 176, 178

- [42] D. L. Donoho, A. Maleki, and A. Montanari. Message Passing Algorithms for Compressed Sensing. *Proceedings of the National Academy of Sciences*, 106:18914–18919, 2009. [22](#), [24](#), [85](#), [86](#), [115](#), [121](#), [122](#), [130](#), [131](#), [132](#), [136](#), [138](#), [147](#), [153](#), [184](#)
- [43] D. L. Donoho, A. Maleki, and A. Montanari. Message Passing Algorithms for Compressed Sensing: I. Motivation and Construction. In *Proceedings of IEEE Inform. Theory Workshop*, Cairo, 2010. [22](#), [122](#), [131](#), [132](#)
- [44] D. L. Donoho, A. Maleki, and A. Montanari. The Noise Sensitivity Phase Transition in Compressed Sensing. *IEEE Trans. on Inform. Theory*, 57:6920–6941, 2011. [xiv](#), [20](#), [64](#), [86](#), [130](#), [131](#), [193](#)
- [45] D. L. Donoho and J. Tanner. Neighborliness of randomly-projected simplices in high dimensions. *Proceedings of the National Academy of Sciences*, 102(27):9452–9457, 2005. [61](#), [128](#)
- [46] D. L. Donoho and J. Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *Journal of American Mathematical Society*, 22:1–53, 2009. [61](#)
- [47] D. L. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete & Computational Geometry*, 43(3):522–541, 2010. [130](#)
- [48] D. L. Donoho and J. Tanner. Precise undersampling theorems. *Proceedings of the IEEE*, 98:913–924, 2010. [61](#)
- [49] J. Fan, S. Guo, and N. Hao. Variance estimation using refitted cross-validation in ultrahigh dimensional regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(1):37–65, 2012. [40](#)
- [50] J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001. [40](#)
- [51] J. Fan and J. Lv. Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5):849–911, 2008. [40](#)

- [52] J. Fan, R. Samworth, and Y. Wu. Ultrahigh dimensional feature selection: beyond the linear model. *The Journal of Machine Learning Research*, 10:2013–2038, 2009. 40
- [53] A. J. Felstrom and K. S. Zigangirov. Time-varying periodic convolutional codes with low-density parity-check matrix. *IEEE Trans. on Inform. Theory*, 45:2181–2190, 1999. 131, 144
- [54] S. Foucart. Sparse recovery algorithms: Sufficient conditions in terms of restricted isometry constants. In M. Neamtu and L. Schumaker, editors, *Approximation Theory XIII: San Antonio 2010*, volume 13 of *Springer Proceedings in Mathematics*, pages 65–77. Springer New York, 2012. 20
- [55] A. Frank and A. Asuncion. UCI machine learning repository (communities and crime data set). <http://archive.ics.uci.edu/ml>, 2010. University of California, Irvine, School of Information and Computer Sciences. 92, 95
- [56] J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22, 2010. 82
- [57] R. G. Gallager. *Low-Density Parity-Check Codes*. MIT Press, Cambridge, Massachusetts, 1963. Available online at <http://web./gallager/www/pages/ldpc.pdf>. 146
- [58] E. Greenshtein and Y. Ritov. Persistence in high-dimensional predictor selection and the virtue of over-parametrization. *Bernoulli*, 10:971–988, 2004. 9
- [59] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inf. Theory*, 57(3):1548–1566, 2011. 8
- [60] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in mathematical physics*, 233(1):1–12, 2003. 70
- [61] D. Guo, D. Baron, and S. Shamai. A Single-letter Characterization of Optimal Noisy Compressed Sensing. In *47th Annual Allerton Conference*, Monticello, IL, September 2009. 70
- [62] D. Guo, S. Shamai, and S. Verdú. Mutual information and minimum mean-square error in gaussian channels. *IEEE Trans. Inform. Theory*, 51:1261–1282, 2005. 165

- [63] D. Guo and S. Verdu. Randomly Spread CDMA: Asymptotics via Statistical Physics. *IEEE Trans. on Inform. Theory*, 51:1982–2010, 2005. 70, 194
- [64] S. H. Hassani, N. Macris, and R. Urbanke. Coupled graphical models and their thresholds. In *Proceedings of IEEE Inform. Theory Workshop*, Dublin, 2010. 132, 134, 144, 147
- [65] A. E. Hoerl and R. W. Kennard. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970. 6
- [66] P. J. Huber and E. Ronchetti. *Robust Statistics (second edition)*. J. Wiley and Sons, 2009. 65, 139
- [67] I. Ibragimov and R. Z. Khasminskii. *Statistical estimation: asymptotic theory*. Springer, 1981. 4
- [68] P. Indyk, E. Price, and D. P. Woodruff. On the Power of Adaptivity in Sparse Recovery. In *IEEE Symposium on the Foundations of Computer Science, FOCS*, October 2011. 132
- [69] L. Jacob, G. Obozinski, and J.-P. Vert. Group lasso with overlap and graph lasso. In *Proceedings of the 26th Annual International Conference on Machine Learning, ICML '09*, pages 433–440, New York, NY, USA, 2009. ACM. 7, 8
- [70] A. Javanmard and A. Montanari. Subsampling at information theoretically optimal rates. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2431–2435, 2012. 25, 144, 147
- [71] A. Javanmard and A. Montanari. Confidence intervals and hypothesis testing for high-dimensional statistical models. In *Advances in Neural Information Processing Systems*, pages 1187–1195. Curran Associates, Inc., 2013. 17
- [72] A. Javanmard and A. Montanari. Hypothesis testing in high-dimensional regression under the gaussian random design model: Asymptotic theory. *arXiv preprint arXiv:1301.4240*, (To appear in *IEEE Transaction on Information Theory*), 2013. 18, 37, 38

- [73] A. Javanmard and A. Montanari. Model selection for high-dimensional regression under the generalized irrepresentability condition. In *Advances in Neural Information Processing Systems*, pages 3012–3020. Curran Associates, Inc., 2013. 10
- [74] A. Javanmard and A. Montanari. State Evolution for General Approximate Message Passing Algorithms, with Applications to Spatial Coupling. *Journal of Information and Inference*, 2(2):115–144, 2013. 25, 122, 134, 148, 154, 156, 182
- [75] A. Javanmard and A. Montanari. Confidence Intervals and Hypothesis Testing for High-Dimensional Regression. *arXiv preprint arXiv:1306.3171*, (To appear in *Journal of Machine Learning Research*), 2014. 17
- [76] Y. Kabashima, M. Vehkaperä, and S. Chatterjee. Typical  $l_1$ -recovery limit of sparse vectors represented by concatenations of random orthogonal matrices. *J. Stat. Mech.*, page P12003, 2012. 70
- [77] Y. Kabashima, T. Wadayama, and T. Tanaka. A typical reconstruction limit for compressed sensing based on  $L_p$ -norm minimization. *J. Stat. Mech.*, page L09003, 2009. 70
- [78] U. S. Kamilov, V. K. Goyal, and S. Rangan. Message-Passing Estimation from Quantized Samples. *arXiv:1105.6368*, 2011. 131
- [79] B. S. Kashin and V. N. Temlyakov. A remark on compressed sensing. *Mathematical Notes*, 82:748–755, 2007. 128
- [80] Y. Kim, J. Kim, and Y. Kim. Blockwise Sparse Regression. *Stat. Sinica*, 16:375–390, 2006. 7
- [81] K. Knight and W. Fu. Asymptotics for lasso-type estimators. *The Annals of Statistics*, pages 1356–1378, 2000. 9
- [82] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborova. Statistical physics-based reconstruction in compressed sensing. *arXiv:1109.4424*, 2011. 21, 22, 127, 130, 132, 144, 147, 166, 176, 178, 186, 187
- [83] S. Kudekar, C. Measson, T. Richardson, and R. Urbanke. Threshold Saturation on BMS Channels via Spatial Coupling. In *Proceedings of the International Symposium on Turbo Codes and Iterative Information Processing*, Brest, 2010. 132, 134, 144

- [84] S. Kudekar and H. D. Pfister. The effect of spatial coupling on compressive sensing. In *48th Annual Allerton Conference*, pages 347–353, 2010. 21, 132
- [85] S. Kudekar, T. Richardson, and R. Urbanke. Threshold Saturation via Spatial Coupling: Why Convolutional LDPC Ensembles Perform So Well over the BEC. *IEEE Trans. on Inform. Theory*, 57:803–834, 2011. 21, 131, 145
- [86] S. Kudekar, T. Richardson, and R. Urbanke. Spatially coupled ensembles universally achieve capacity under belief propagation. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pages 453–457. IEEE, 2012. 132, 134, 144
- [87] L. Le Cam. On the asymptotic theory of estimation and testing hypotheses. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 129–156. University of California Press Berkeley, CA, 1956. 44
- [88] E. Lehmann and G. Casella. *Theory of point estimation*. Springer, 2 edition, 1998. 45
- [89] E. L. Lehmann and J. P. Romano. *Testing statistical hypotheses*. Springer, 2005. 48, 112
- [90] M. Lentmaier and G. P. Fettweis. On the thresholds of generalized LDPC convolutional codes based on protographs. In *IEEE Intl. Symp. on Inform. Theory*, Austin, Texas, August 2010. 131
- [91] A. Maleki, L. Anitori, A. Yang, and R. Baraniuk. Asymptotic Analysis of Complex LASSO via Complex Approximate Message Passing (CAMP). arXiv:1108.0477, 2011. 187
- [92] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993. 32
- [93] E. Matusiak and Y. C. Eldar. Sub-nyquist sampling of short pulses. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3944–3947, 2011. 179
- [94] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke. The Generalized Area Theorem and Some of its Consequences. *IEEE Trans. on Inform. Theory*, 55(11):4793–4821, 2009. 146

- [95] N. Meinshausen and P. Bühlmann. High-dimensional graphs and variable selection with the Lasso. *The Annals of Statistics*, 34:1436–1462, 2006. 9, 10, 61, 71
- [96] N. Meinshausen, L. Meier, and P. Bühlmann. p-values for high-dimensional regression. *Journal of the American Statistical Association*, 104(488):1671–1681, 2009. 17, 57, 81, 84
- [97] M. Mézard and A. Montanari. *Information, Physics and Computation*. Oxford, 2009. 194, 199
- [98] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, 1987. 194, 199
- [99] A. Montanari. Graphical Models Concepts in Compressed Sensing. In Y.C. Eldar and G. Kutyniok, editors, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012. 121, 122, 138
- [100] A. Montanari and D. Tse. Analysis of belief propagation for non-linear problems: the example of CDMA (or: how to prove Tanaka’s formula). In *Proceedings of IEEE Inform. Theory Workshop*, Punta de l’Este, Uruguay, 2006. 70
- [101] S. Negahban and M. J. Wainwright. Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Annals of Statistics*, 39(2):1069–1097, 2011. 8
- [102] S. N. Negahban and M. J. Wainwright. Simultaneous support recovery in high dimensions: Benefits and perils of block  $\ell_1/\ell_\infty$ -regularization. *IEEE Trans. Inf. Theor.*, 57(6):3841–3863, 2011. 7
- [103] Y. E. Nesterov. A method for solving the convex programming problem with convergence rate  $O(1/k^2)$ . *Dokl. Akad. Nauk SSSR*, 269:543?547 (in Russian), 1983. 15
- [104] Y. E. Nesterov. Gradient methods for minimizing composite objective function. <http://www.ecore.be/DPs/dp-1191313936.pdf>, 2007. CORE report. 15
- [105] G. Obozinski, M. J. Wainwright, and M. I. Jordan. Union support recovery in high-dimensional multivariate regression. *Annals of Statistics*, 39(1):1–47, 2011. 7

- [106] S. Oymak, C. Thrampoulidis, and B. Hassibi. The squared-error of generalized lasso: A precise analysis. *arXiv:1311.0830*, 2013. 69, 70
- [107] D. Panchenko. *The Sherrington-Kirkpatrick model*. Springer, 2013. 70
- [108] S. Rangan. Generalized Approximate Message Passing for Estimation with Random Linear Mixing. In *IEEE Intl. Symp. on Inform. Theory*, St. Perersbourg, August 2011. 131, 136
- [109] S. Rangan, A. K. Fletcher, and V. K. Goyal. Asymptotic Analysis of MAP Estimation via the Replica Method and Applications to Compressed Sensing. In *NIPS*, Vancouver, 2009. 70
- [110] G. Raskutti, M. J. Wainwright, and B. Yu. Minimax rates of estimation for high-dimensional linear regression over  $\ell_q$ -balls. In *47th Annual Allerton Conf.*, Monticello, IL, September 2009. 9, 128
- [111] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010. 8
- [112] S. Reid, R. Tibshirani, and J. Friedman. A study of error variance estimation in Lasso regression. *arXiv preprint arXiv:1311.5274*, 2013. 40
- [113] Z. Ren, T. Sun, C.-H Zhang, and H. Zhou. Asymptotic Normality and Optimalities in Estimation of Large Gaussian Graphical Model. *arXiv:1309.6024*, 2013. 61, 71
- [114] A. Rényi. On the dimension and entropy of probability distributions. *Acta Mathematica Hungarica*, 10:193–215, 1959. 123, 124
- [115] T. J. Richardson and R. Urbanke. *Modern Coding Theory*. Cambridge University Press, Cambridge, 2008. 135, 178
- [116] M. Rudelson and Z. Shuheng. Reconstruction from anisotropic random measurements. *IEEE Transactions on Information Theory*, 59(6):3434–3447, 2013. 34, 99
- [117] P. Schniter. Turbo Reconstruction of Structured Sparse Signals. In *Proceedings of the Conference on Information Sciences and Systems*, Princeton, 2010. 131
- [118] P. Schniter. A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels. *arXiv:1101.4724*, 2011. 131

- [119] P. J. Schreier and L. L. Scharf. *Statistical signal processing of complex-valued data : the theory of improper and noncircular signals*. Cambridge University Press, Cambridge, 2010. 181
- [120] S. Som, L. C. Potter, and P. Schniter. Compressive Imaging using Approximate Message Passing and a Markov-Tree Prior. In *Proc. Asilomar Conf. on Signals, Systems, and Computers*, November 2010. 131
- [121] A. Sridharan, M. Lentmaier, D. J. Costello Jr, and K. S. Zigangirov. Convergence analysis of a class of LDPC convolutional codes for the erasure channel. In *43rd Annual Allerton Conference*, Monticello, IL, September 2004. 131, 144
- [122] N. Städler, P. Bühlmann, and S. van de Geer.  $\ell_1$ -penalization for mixture regression models (with discussion). *Test*, 19(2):209–256, 2010. 40
- [123] M. Stojnic, F. Parvaresh, and B. Hassibi. On the reconstruction of block-sparse signals with an optimal number of measurements. *IEEE Transactions on Signal Processing*, 57(8):3075–3085, 2009. 7
- [124] T. Sun and C.-H. Zhang. Scaled sparse linear regression. *Biometrika*, 99(4):879–898, 2012. 40, 41, 65, 189
- [125] K. Takeda and Y. Kabashima. Statistical mechanical analysis of compressed sensing utilizing correlated compression matrix. In *IEEE Intl. Symp. on Inform. Theory*, june 2010. 70, 194
- [126] M. Talagrand. *Mean Field Models for Spin Glasses: Volume I*. Springer-Verlag, Berlin, 2010. 70
- [127] T. Tanaka. A Statistical-Mechanics Approach to Large-System Analysis of CDMA Multiuser Detectors. *IEEE Trans. on Inform. Theory*, 48:2888–2910, 2002. 70, 194
- [128] T. Tanaka and M. Okada. Approximate belief propagation, density evolution, and statistical neurodynamics for cdma multiuser detection. *Information Theory, IEEE Transactions on*, 51(2):700–706, 2005. 70
- [129] R. Tibshirani. Regression shrinkage and selection with the Lasso. *J. Royal. Statist. Soc B*, 58:267–288, 1996. 6

- [130] A. N. Tikhonov. On the stability of inverse problems. *Doklady Akademii nauk SSSR*, 39(5):195–198, 1943. 5
- [131] A. Tulino, G. Caire, S. Shamai, and S. Verdú. Support recovery with sparsely sampled free random matrices. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 2328–2332. IEEE, 2011. 70
- [132] B. A. Turlach, W. N. Venables, and S. J. Wright. Simultaneous variable selection. *Technometrics*, 47(3):349–363, 2005. 7
- [133] S. van de Geer and P. Bühlmann. On the conditions used to prove oracle results for the Lasso. *Electron. J. Statist.*, 3:1360–1392, 2009. 9, 10, 33, 34, 86
- [134] S. van de Geer, P. Bühlmann, Y. Ritov, and R. Dezeure. On asymptotically optimal confidence regions and tests for high-dimensional models. *arXiv preprint arXiv:1303.0518*, (To appear in *Annals of Statistics*), 2014. 33, 37, 38, 39, 45, 56, 72, 73
- [135] A. W. Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000. 15, 44
- [136] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Y.C. Eldar and G. Kutyniok, editors, *Compressed sensing: Theory and Applications*, pages 210–268. Cambridge University Press, 2012. 100, 101, 102
- [137] J. Vila and P. Schniter. Expectation-maximization bernoulli-gaussian approximate message passing. In *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, 2011. 130
- [138] C. Villani. *Optimal transport: old and new*, volume 338. Springer, 2008. 120
- [139] M. J. Wainwright. Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting. *IEEE Trans. on Inform. Theory*, 55:5728–5741, 2009. 128
- [140] M. J. Wainwright. Sharp thresholds for high-dimensional and noisy sparsity recovery using  $\ell_1$ -constrained quadratic programming. *IEEE Trans. on Inform. Theory*, 55:2183–2202, 2009. 9, 10

- [141] L. Wasserman. *All of statistics: a concise course in statistical inference*. Springer Verlag, 2004. 27
- [142] L. Wasserman and K. Roeder. High dimensional variable selection. *The Annals of statistics*, 37(5A):2178, 2009. 17, 56
- [143] Y. Wu and S. Verdú. Rényi Information Dimension: Fundamental Limits of Almost Lossless Analog Compression. *IEEE Trans. on Inform. Theory*, 56:3721–3748, 2010. 21, 22, 123, 124, 128, 146, 182
- [144] Y. Wu and S. Verdú. MMSE dimension. *IEEE Trans. on Inform. Theory*, 57(8):4857–4879, 2011. 123, 124, 125, 165
- [145] Y. Wu and S. Verdú. Optimal phase transitions in compressed sensing. *Information Theory, IEEE Transactions on*, 58(10):6241, 2012. 70
- [146] Z. Yang, C. Zhang, and L. Xie. On phase transition of compressed sensing in the complex domain. *IEEE Signal Processing Letters*, 19:47–50, 2012. 187
- [147] M. Yuan and Y. Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society B*, 68:49–67, 2010. 7
- [148] C-H. Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894–942, 2010. 40
- [149] C.-H. Zhang and S. S. Zhang. Confidence intervals for low-dimensional parameters in high-dimensional linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76:217–242, 2014. xiii, xiv, xvi, 17, 38, 39, 58, 59, 72, 73, 85, 86, 87, 88, 89, 90, 92, 93, 94
- [150] P. Zhao and B. Yu. On model selection consistency of Lasso. *The Journal of Machine Learning Research*, 7:2541–2563, 2006. 9, 10